

Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Katedra geoinformatiky

**KLASIFIKACE A ČASOVÁ ANALÝZA OSOBNÍ
A NÁKLADNÍ VLAKOVÉ DOPRAVY
REGIONÁLNÍCH TRATÍ**

Diplomová práce

Bc. Lukáš POSPÍŠIL

Vedoucí práce doc. Ing. Zdena Dobešová, Ph.D.

Olomouc 2022

Geoinformatika a kartografie

ANOTACE

Diplomová práce popisuje klasifikaci a časovou analýzu provozu osobní a nákladní železniční dopravy na regionálních a celostátních tratích Olomouckého a Moravskoslezského kraje. Využita byla data provozu na železnici za období 2016–2021, obsahující měsíční agregace údajů o vlakové dopravě v České republice. Tato data nejsou veřejně dostupná a jejich poskytovatelem je státní organizace Správa železnic.

V teoretické části jsou uvedeny metody zpracování práce i odborné studie, ve kterých již byla tato specifická data analyzována. Dále je zde uvedena kategorizace železniční infrastruktury České republiky, údaje o dopravě z Českého statistického úřadu a úvod do problematiky časových řad a shlukování.

V praktické části byl na úvod popsán postup selekce dat pouze pro požadované území, a možnosti očištění dat od kalendářních vlivů. Vyhotoveny byly mapové vizualizace struktury dopravy v letech 2016 a 2021. Navrženy byly metody identifikace výkyvů v provozu a provedena časová analýza spočívající v tvorbě časových řad s jejich následnou dekompozicí. Klasifikace tratí byla realizována několika metodami shlukování. Jejich výsledky byly interpretovány a vizualizovány na mapových výstupech. Práce obsahuje také diagramy a tabulky se statistickými údaji.

KLÍČOVÁ SLOVA

vlaková doprava; železnice; časová řada; shlukování

Počet stran práce: 95

Počet příloh: 12 (z toho 2 volné a 3 elektronické)

ANOTATION

The thesis describes the classification and time analysis of passenger and cargo train traffic on regional and national lines of the Olomouc and Moravian-Silesian regions. Railway traffic data for the period 2016-2021, containing monthly aggregations of train traffic data in the Czech Republic, were used. This data is not publicly available and is provided by the state organisation Railway Administration.

In the theoretical part, the methods of work processing and expert studies in which these specific data have already been analysed are presented. Furthermore, the categorisation of the railway infrastructure of the Czech Republic, the traffic data from the Czech Statistical Office and an introduction to the issues of time series and clustering are presented.

In the practical part, the procedure of selecting data only for the required area and the possibilities of data adjustment from calendar effects were described. Map visualizations of the traffic structure in 2016 and 2021 were produced. Methods for identifying fluctuations in traffic were proposed and a time series analysis was performed consisting in the generation of time series with their subsequent decomposition. Classification of the lines was carried out using several clustering methods. Their results were interpreted and visualized on map outputs. The work also contains diagrams and tables with statistical data.

KEYWORDS

train transport; railways; time series; clustering

Number of pages: 95

Number of appendixes: 12

Prohlašuji, že

- diplomovou práci včetně příloh, jsem vypracoval(a) samostatně a uvedl(a) jsem všechny použité podklady a literaturu.

- jsem si vědom(a), že na moji diplomovou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevýdělečně, ke své vnitřní potřebě, diplomovou práci užívat (§ 35 odst. 3),

- souhlasím, aby jeden výtisk diplomové práce byl uložen v Knihovně UP k prezenčnímu nahlédnutí,

- souhlasím, že údaje o mé diplomové práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé diplomové práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé diplomové práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne

Bc. Lukáš POSPÍŠIL

Za pomoc a rady při řešení diplomové práce děkuji vedoucí práce doc. Ing. Zdeně Dobešové, Ph.D. a konzultantovi Mgr. Michalovi Kučerovi, kterému současně děkuji za poskytnutí dat potřebných pro zpracování této práce.

UNIVERZITA PALACKÉHO V OLMOUCI

Přírodovědecká fakulta

Akademický rok: 2020/2021

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: Bc. Lukáš POSPÍŠIL
Osobní číslo: R200618
Studijní program: N0532A330009 Geoinformatika a kartografie
Studijní obor: Geoinformatika a kartografie
Téma práce: Klasifikace a časová analýza osobní a nákladní vlakové dopravy regionálních tratí
Zadávající katedra: Katedra geoinformatiky

Zásady pro vypracování

Cílem práce je provést klasifikaci a časovou analýzu provozu na vybraných celostátních a regionálních železničních tratích Olomouckého a Moravskoslezského kraje. Student použije víceletou datovou sadu agregovaných měsíčních údajů o osobní a nákladní dopravě na regionálních tratích (bez železničních koridorů). Provede časové vyhodnocení naměřených dat, klasifikaci tratí s určením tratí s podobným časovým průběhem, včetně identifikace tratí se specificky odlišným provozem.

Celá práce (text, přílohy, výstupy, zdrojová a vytvořená data) se odevzdá v digitální podobě na paměťovém nosiči (CD, DVD, SD karta, flash disk). Text práce s vybranými přílohami bude odevzdán ve dvou svázaných výtiscích na sekretariát katedry. O diplomové práci student vytvoří webovou stránku v souladu s pravidly dostupnými na stránkách katedry. Práce bude zpracována podle zásad dle Voženílek (2002) a závazné šablony pro diplomové práce na KGI. Povinnou přílohou práce bude poster formátu A2.

Rozsah pracovní zprávy: max. 50 stran
Rozsah grafických prací: dle potřeby
Forma zpracování diplomové práce: tištěná

Seznam doporučené literatury:

CINKE J. Analýza dat vytíženosti železniční infrastruktury, Univerzita Palackého v Olomouci, Katedra geoinformatiky PŘF, diplomová práce, 2020
DOBEŠOVÁ Z., KUČERA M. Time series of workload on railway routes, In Silhavy, R. (ed.) CSOC 2019. Artificial Intelligence and Algorithms in Intelligent Systems, vol. 985, pp. 370-380, 2020, Springer, Cham. DOI: 10.1007/978-3-030-19810-7_37
HANČLOVÁ J., TVRDÝ L.: Úvod do analýzy časových řad, Ekonomická fakulta VŠB-TU Ostrava, 2003
KŘIVÝ, Ivan, 2012. Analýza časových řad. Ostrava: Ostravská univerzita v Ostravě, 2012.
PETR. P. Metody Data Miningu (část 1, část 2). Pardubice, Univerzita Pardubice. 2014
Správa železnic. Statistické ročenky
Správa železnic. Popis sítě. Mapy. Dostupné na: <https://provoz.spravazeleznic.cz/Portal/ViewArticle.aspx?oid=594598>
ŠARMANOVÁ J. Metody analýzy dat, VŠB-TU, Ostrava, 2012, ISBN 978-80-248-2565-6.
VOŽENÍLEK, V. Diplomové práce z geoinformatiky. Olomouc, Univerzita Palackého v Olomouci, 2002

Vedoucí diplomové práce: doc. Ing. Zdena Dobešová, Ph.D.
Katedra geoinformatiky

Datum zadání diplomové práce: 9. listopadu 2020
Termín odevzdání diplomové práce: 6. května 2022

L.S.

doc. RNDr. Martin Kubala, Ph.D.
děkan



prof. RNDr. Vít Voženílek, CSc.
vedoucí katedry

OBSAH

SEZNAM POUŽITÝCH ZKRATEK	10
ÚVOD	11
1 CÍLE PRÁCE.....	12
2 METODY A POSTUPY ZPRACOVÁNÍ.....	13
3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY	17
3.1 Správa železnic, státní organizace	17
3.2 Kategorizace železniční infrastruktury	17
3.3 Řešené práce	19
3.4 Průběh pandemie nemoci covid-19 v ČR.....	20
3.5 Doprava z Českého statistického úřadu.....	21
3.6 Časové řady.....	24
3.6.1 Problémy analýzy časových řad.....	24
3.6.2 Základní analýzy časových řad	25
3.7 Shluková analýza.....	26
3.7.1 Nehierarchické a hierarchické metody	27
3.7.2 Dynamické borcení časové osy.....	28
4 PŘÍPRAVA DAT	29
4.1 Výběr monitorovacích bodů SR70.....	29
4.2 Výběr železničních úseků.....	31
4.3 Tvorba agregačních tabulek	33
4.4 Očištění dat	33
5 STRUKTURA ŽELEZNIČNÍ DOPRAVY	35
5.1 Železniční doprava v roce 2016	36
5.2 Železniční doprava v roce 2021	38
5.3 Železniční doprava v březnu–květnu roku 2020 a 2021.....	40
6 ČASOVÁ ANALÝZA.....	43
6.1 Výkyvy v provozu	43
6.2 Meziroční tempo růstu	46
6.3 Dekompozice časových řad.....	47
6.4 Analýza vybraných úseků a stanic.....	48
6.4.1 Celostátní úseky	49
6.4.2 Regionální úseky.....	52
6.4.3 Průmyslové úseky	55
6.4.4 Specifické úseky.....	58
6.4.5 Stanice na státní hranici	61
7 KLASIFIKACE TRATÍ.....	64
7.1 Hierarchické shlukování	65
7.1.1 Roční agregace	67
7.1.2 Měsíční agregace.....	73
7.2 Shlukování metodou k-means.....	76
7.2.1 Roční agregace	77

7.2.2	Měsíční agregace	79
7.3	Shlukování podle časových řad	80
7.3.1	Časové řady počtu vlaků bez nulových hodnot	81
7.3.2	Časové řady počtu vlaků za období 2016–2019	82
7.3.3	Časové řady počtu vlaků v roce 2020	84
7.3.4	Časové řady počtu vlaků v roce 2021	87
8	VÝSLEDKY	90
9	DISKUZE	93
	ZÁVĚR	95
	POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE	
	PŘÍLOHY	

SEZNAM POUŽITÝCH ZKRATEK

Zkratka	Význam
CZ-NACE	Klasifikace ekonomických činností
ČSÚ	Český statistický úřad
ČD	České dráhy
DTW	Dynamic Time Warping
MDČR	Ministerstvo dopravy České republiky
MVČR	Ministerstvo vnitra České republiky
MZČR	Ministerstvo zdravotnictví České republiky
NZIP	Národní zdravotnický informační portál
PAM	Partition around medoids
SR70	Číselník železničních stanic a dopravně významných míst
SR72	Číselník železničních drah
SŽ	Správa železnic
VDB	Veřejná databáze ČSÚ

ÚVOD

Železniční doprava je v České republice významným druhem osobní i nákladní dopravy. Česká republika disponuje nejhustší železniční sítí v Evropě, jejíž celková délka v roce 2021 byla 9 358 km (Správa železnic, 2022a). Cestování vlakem je v ČR považováno za jeden z nejpobulárnějších způsobů dopravy. Právě díky rozsáhlé síti železničních tratí je železniční doprava dostupná na většině území. Nákladní doprava bývá často využívána pro přepravu zboží o velkých hmotnostech na delší vzdálenosti, Při přepravě velkých objemů komodit na velké vzdálenosti jsou náklady na železniční přepravu nižší, nežli na silniční nebo leteckou. V porovnání se silniční dopravou není však železniční doprava tak komplexní a je omezena výchozí a koncovou stanicí (Besta, 2022).

Význam železniční dopravy se v České republice v posledních letech neustále zvyšoval, což od roku 2015 dokládají rostoucí hodnoty přepravených osob v osobní dopravě, a rostoucí hmotnost přepravených věcí u nákladní dopravy (ČSÚ, 2021). Rostoucí trend v osobní dopravě potvrzuje i národní dopravce České dráhy, který je největším provozovatelem osobních vlaků na českých železnicích. Lidé také cestují na delší vzdálenosti než dříve, jelikož v dnešní době jsou ochotni za práci dojíždět dále, a také je vlak oblíbenou alternativou místo auta na delší trasy (České dráhy, 2020).

V roce 2020 však dorazila do České republiky celosvětová pandemie virového onemocnění covid-19. Proto vláda zavedla sérii opatření, která měla šíření viru zamezit. Tato opatření omezila mobilitu obyvatel, a tím poptávku po osobní dopravě, a současně snížila přepravu zboží po železnici. Po zavedení opatření klesl počet cestujících ve vlacích Českých drah o 90 %. Úbytek cestujících zaznamenali také ostatní dopravci. Malý počet cestujících způsobil dopravcům velké ztráty v tržbách. Na nákladní dopravu měla vliv ekonomická situace. Utlumení průmyslové výroby způsobilo pokles přeprav komodit (Haspra, 2021).

1 CÍLE PRÁCE

Hlavním cílem diplomové práce je provést klasifikaci a časovou analýzu provozu osobní a nákladní dopravy na vybraných celostátních a regionálních železničních tratích (bez železničních koridorů) Olomouckého a Moravskoslezského kraje. K dosažení cílů bude využita víceletá datová sada agregovaných měsíčních údajů o vlakové dopravě na tratích České republiky. Nejprve bude provedeno předzpracování dat, zahrnující selekci údajů pouze pro definované území, vytvoření prostorových vrstev pro mapové vizualizace a očištění dat od kalendářních vlivů. Poté budou pro vybrané úseky vytvořeny časové řady, na jejichž základě bude vyhodnocován provoz osobní a nákladní dopravy od roku 2016 do roku 2021. Nakonec bude provedena klasifikace metodou shlukování, spočívající v identifikaci tratí s podobným a současně odlišným provozem, a určení tratí s podobným časovým průběhem provozu.

Dalšími cíli práce je hodnocení vlivu dopadů pandemie covidu-19 na osobní a nákladní dopravu v letech 2020 a 2021. Dále zjišťování trendů v provozu na tratích pomocí dekompozice časových řad. V případě, že to bude realizovatelné, tak bude sledována i složka sezónní. Při shlukování budou využity a porovnány přístupy hierarchické metody a metody k-means. U hierarchické metody budou testovány různé metody spojování shluků a vzdálenostních metrik, z nichž bude nakonec vybrána nejvhodnější kombinace pro konkrétní analyzovaná data. Posledním cílem je využití shlukování časových řad metodou dynamického borcení časové osy pro hledání tratí s podobným časovým průběhem.

Přínosem práce budou jak nově vzniklé prostorové vrstvy železniční infrastruktury v Olomouckém a Moravskoslezském kraji, tak samotný text práce, který obsahuje podrobný popis vývoje a situace v železniční dopravě ve sledovaném území, doplněný diagramy, statistickými údaji a mapovými výstupy.

2 METODY A POSTUPY ZPRACOVÁNÍ

V úvodu práce byla provedena rešerše literatury, seznámení se s tématem práce, metodami zpracování, důležitými pojmy a povahou poskytovaných dat. Podstatným dokumentem pro zpracování této práce je diplomová práce Jana Cinkeho (2020), ve které jsou taktéž využita data provozu na železnici, avšak pro území celé České republiky. Jelikož se jedná o specifická data, která nejsou běžně poskytována, je tato práce ojedinělým zdrojem poznatků o charakteru vyhodnocovaných dat a postupů pro jejich prvotní zpracování. Další informace k tématu provozu na železnici v ČR poskytuje článek *Time Series of Workload on Railway Routes* (Dobešová, Kučera, 2019) a studie *Železniční doprava v roce 2020* (Kučera, Dobešová, 2022).

Pro problematiku časové analýzy a tvorby časových řad byla nastudována odborná literatura, a to konkrétně publikace *Úvod do analýzy časových řad* (Hančlová, Tvrдый, 2003) a *Analýza časových řad* (Křivý, 2012). Tématem klasifikace a shlukové analýzy se zabývá učební text *Metody analýzy dat* (Šarmanová, 2012). Jelikož zpracování časové i shlukové analýzy bylo realizováno v programu Orange, byla využita učebnice *ORANGE Praktický návod do cvičení předmětu Data Mining* (Dobešová, 2022).

Podpůrnými materiály tématu železniční infrastruktury jsou mapy železnice ČR, které vydává Správa železnic (SŽ) a pro veřejnost je poskytuje na svých webových stránkách (Správa železnic, 2022b). Správa železnic dále vydává vnitřní instruktážní dokumenty ve formě služebních rukověti, které jsou volně dostupné. Prvním z nich je *Služební rukověť SŽ SR70 – Číselník železničních stanic a dopravně významných míst* (Svoboda, 2021a). Číselník obsahuje železniční stanice, dopravně zajímavá a tarifní místa na železniční síti vlastněné Českou republikou. Druhým dokumentem je *Služební rukověť SŽ SR72 – Číselník železničních drah* (Svoboda, 2021b), který obsahuje seznam železničních drah a vleček.

Po rešerši byla realizována příprava dat zahrnující tvorbu agregací pomocí kontingenčních tabulek a očištění dat na stejnou délku měsíce pro potřeby tvorby časových řad. Dále byla provedena tvorba jednoduchých mapových vizualizací, časová analýza spočívající v tvorbě časových řad a jejich dekompozice na časové složky, klasifikace a shlukování tratí a stanic pro identifikaci tratí nebo stanic s podobným i odlišným provozem. Poslední fází byla tvorba finálních mapových výstupů.

Použité metody

Data provozu na železnici byla poskytována ve formě tabulkových souborů. Proto byly dále zpracovány v tabulkovém procesoru Microsoft Excel verze 2102, s využitím kontingenčních tabulek pro tvorbu agregací za jednotlivé roky a měsíce pro osobní i nákladní vlakovou dopravu. V GIS software ArcGIS Pro 2.8.3. byly z dat vytvořeny dvě prostorové vrstvy. Bodová vrstva monitorovacích bodů na území České republiky a liniová vrstva železničních úseků České republiky. Poté na základě podkladových dat proběhla selekce monitorovacích bodů a úseků, které leží v Olomouckém nebo Moravskoslezském kraji a nejsou součástí železničních tranzitních koridorů. Agregovaná data ve formě tabulek byla připojena k prostorovým vrstvám a dále analyzována a vizualizována. Tyto kroky proběhly v softwaru ArcGIS Pro 2.8.3. s využitím standardních nástrojů. Analýza časových řad proběhla v programu Orange verze 3.30.2, ve kterém byla realizována dekompozice časových řad na jednotlivé časové složky, společně s využitím programu Microsoft Excel. V programu Orange byla taktéž řešena klasifikace a shluková analýza monitorovacích bodů a železničních tratí.

Výsledky klasifikace a shlukování byly vizualizovány v ArcGIS Pro 2.8.3., ve kterém proběhla tvorba všech mapových výstupů vzniklých v rámci této práce. Výsledný poster byl upraven v grafickém softwaru Inkscape 1.0.

Použitá data

Pro zpracování diplomové práce jsou stěžejní data provozu na železnici. Tato data vlastní státní organizace Správa železnic a jsou poskytnuta pouze pro účely této diplomové práce. Primárně se využívají pro výpočet poplatku za využití železniční dopravní cesty pro železniční dopravce. Data jsou poskytována ve formě tabulek ve formátu *.xlsx*, obsahující informace o využití dráhy v měsíčních agregacích za jednotlivý rok podle kategorie dopravy. Správa železnic má k dispozici i denní agregace, což však představuje velké množství dat, které nelze v rámci této práce zpracovat. V datech se rozlišuje lokomotivní, nákladní, osobní, služební a soupravná kategorie železniční dopravy. Pro zpracování práce byla poskytnuta data za období let 2016–2021.

Data provozu na železniční infrastruktuře jsou vztažena k monitorovacím bodům SR70. Monitorovacími body se rozumí železniční stanice, dopravně zajímavá místa jako jsou například hradla či dopravní a tarifní místa. Všechny tyto body jsou vedeny v číselníku železničních stanic SR70, který vydává Správa železnic. V datech je veden výchozí a koncový dopravní bod SR70. Díky tomu je možné na data nahlížet i jako na úseky mezi monitorovacími body. Jelikož se nejedná o data směrová, jde o úhrny provozu místa v obou směrech. Každý bod SR70 má v tabulce vícenásobný výskyt, jelikož má záznam hodnot pro jednotlivé kategorie vlakové dopravy v každém měsíci daného roku. Data mají podrobnou atributovou složku (Tab. 1) obsahující popisné informace monitorovacího bodu, a především údaje o počtu vlaků, průměrné délce soupravy, hmotnosti souprav, průměrné hmotnosti soupravy, průměrném počtu náprav a průměrném počtu vozů, které daným bodem projely.

Poskytovaná data obsahují údaje z celé České republiky. V rámci této práce byly použity údaje z monitorovacích bodů pouze z vybraných celostátních a regionálních železničních tratích Olomouckého a Moravskoslezského kraje pro osobní a nákladní vlakovou dopravu. Údaje z železničních tranzitních koridorů nebyly uvažovány, jelikož jejich hodnoty jsou příliš vysoké v porovnání s regionálními úseky, a tyto rozdíly by zkreslily a negativně ovlivnily výsledky analýz.

Tab. 1 Atributy dat provozu na železnici

Název atributu	Popis atributu
Datum numericky	Datum ve formátu RRRRMMDD, ke kterému je záznam vztažen
Číslo úseku dle KKM	Interní číselné označení úseku
Výchozí dopravní bod SR70	Jedinečný šestimístný číselný identifikátor výchozího monitorovacího bodu vedeného v číselníku SR70
Název výchozího dopravního bodu	Jedinečné slovní označení výchozího monitorovacího bodu
GPS souřadnice výchozího dopravního bodu	Souřadnice polohy výchozího monitorovacího bodu ve formátu SS°MM'VV,vvv"
SR70 koncového dopravního bodu	Jedinečný šestimístný číselný identifikátor koncového monitorovacího bodu vedeného v číselníku SR70
Název koncového dopravního bodu	Jedinečné slovní označení koncového monitorovacího bodu
GPS souřadnice koncového dopravního bodu	Souřadnice polohy koncového monitorovacího bodu ve formátu SS°MM'VV,vvv"
Druh vlaku	Podrobné dělení osobních, nákladních a služebních vlakových souprav pomocí zkratk definovaných předpisem SŽ
Kód kategorie dopravy	Číselný kód kategorie vlakové dopravy: 1 – osobní, 2 – nákladní, 3 – služební, 4 – lokomotivní, 5 – soupravový
Kategorie dopravy	Slovní označení kategorie vlakové dopravy: osobní, nákladní, služební, lokomotivní, soupravový
Počet vlaků	Počet vlaků za měsíc
Průměrná délka soupravy	Průměrná délka soupravy za měsíc
Hmotnost souprav	Celková hmotnost souprav za měsíc
Průměrná hmotnost soupravy	Průměrná hmotnost soupravy za měsíc
Průměrný počet náprav	Průměrný počet náprav za měsíc
Průměrný počet vozů	Průměrný počet vozů za měsíc

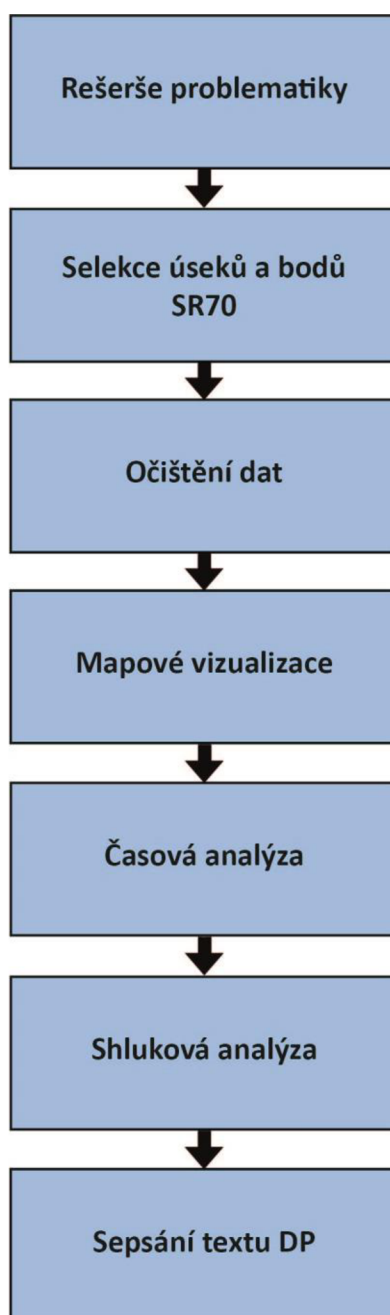
(zdroj: Cinke, 2020; autor)

Použité programy

Prvotní fáze práce spočívala v přípravě poskytnutých dat pro jejich následnou analýzu. Příprava dat probíhala v tabulkovém procesoru Microsoft Excel verze 2102 z balíku Microsoft 365 Apps, ve kterém pomocí funkce kontingenčních tabulek byly vytvářeny agregace hodnot pro osobní a nákladní vlakovou dopravu. Také zde byla data pomocí jednoduchého vzorce očištěna na stejnou délku měsíce, pro potřeby tvorby časových řad. V GIS software ArcGIS Pro 2.8.3. byly z původních dat vytvořeny nové prostorové

vrstvy, ke kterým lze připojit hodnoty z agregačních tabulek a vizualizovat je v prostoru. Časová analýza osobní a nákladní vlakové dopravy byla realizována v programech Orange 3.30.2, společně s opětovným využitím tabulkového procesoru Microsoft Excel. Software Orange byl taktéž využit pro klasifikační a shlukové analýzy měřících bodů a tratí. Shlukování podle podobnosti časových řad bylo testováno pomocí programovacího jazyka R v prostředí programu RStudio 1.4.1717. Mapové výstupy a vizualizace byly řešeny v ArcGIS Pro 2.8.3. Případné designové úpravy mapových výstupů, včetně tvorby finálního posteru, byly realizovány v grafickém softwaru Inkscape 1.0. Text diplomové práce byl sepsán v textovém procesoru Microsoft Word verze 2102 z balíku Microsoft 365 Apps.

Postup zpracování



Obr. 1 Diagram postupu zpracování diplomové práce

3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

3.1 Správa železnic, státní organizace

Správa železnic je státní organizace, která vznikla 1. ledna 2003. Jejím základním posláním je plnit funkci vlastníka a provozovatele dráhy celostátní a drah regionálních ve vlastnictví státu. Správa železnic hospodáří s majetkem, který tvoří železniční dopravní cestu, zajišťuje provozu schopnost, modernizaci a rozvoj celostátních a regionálních drah ve vlastnictví státu (Správa železnic, 2022c).

Správa železnic vydává vnitřní instruktážní dokumenty formou služebních rukovětí, které obsahují vnitřní předpisy organizace. Tyto dokumenty jsou závazné pro komunikaci zaměstnanců SŽ. *Služební rukověť SŽ SR70 – Číselník železničních stanic a dopravně významných míst*, je vnitřní instruktážní dokument pro popis a stanovení dopravních bodů. Gestorem dokumentu je Správa železnic, konkrétně Odbor řízení provozu. Dopravní body jsou vedeny v evidenci s výstupním číselníkem SR70. Číselník SR70 obsahuje železniční stanice a jiná dopravně významná místa nacházející se na železniční síti na území ČR a místa mimo tuto síť, která jsou významná pro sestavu jízdního řádu. Každému místu je přidělena jedinečná identifikace (číselný kód, název). Správa železnic nenese odpovědnost za údaje o místech mimo železniční síť provozovanou SŽ. Ostatní provozovatelé drah musí o zařazení bodu do Číselníku SR70 požádat jeho správce. Tento provozovatel poté odpovídá za aktuální údaje daného místa. Dokument je závazný pro komunikaci mezi SŽ se státními institucemi ČR a mezinárodními železničními institucemi a systémy. Tento dokument nabyl účinnosti od 12. prosince 2021 a nahradil původní služební rukověť Číselník železničních stanic, dopravně zajímavých a tarifních míst, účinný od 1. prosince 2009, který vydala tehdy ještě Správa železniční dopravní cesty (Svoboda, 2021a).

Služební rukověť SŽ SR72 – Číselník železničních drah obsahuje seznam tratí celostátní dráhy, regionálních drah, místních drah, zkušebních drah a seznam vleček, které jsou buď ve vlastnictví České republiky a právo hospodařit s nimi má Správa železnic, nebo jsou ve vlastnictví samotné Správy železnic, nebo SŽ zajišťuje jejich provoz. Číselník drah však může obsahovat i dráhy ve vlastnictví nebo provozované jinými subjekty. Číselník drah slouží pro jednoznačnou identifikaci dráhy a skládá se ze dvou částí, a to Číselníku tratí a Číselníku vleček. Účinnost SŽ SR72 nabyla 9. 8. 2021 a jejím gestorem je Správa železnic, konkrétně Odbor obchodních a smluvních vztahů (Svoboda, 2021b).

3.2 Kategorizace železniční infrastruktury

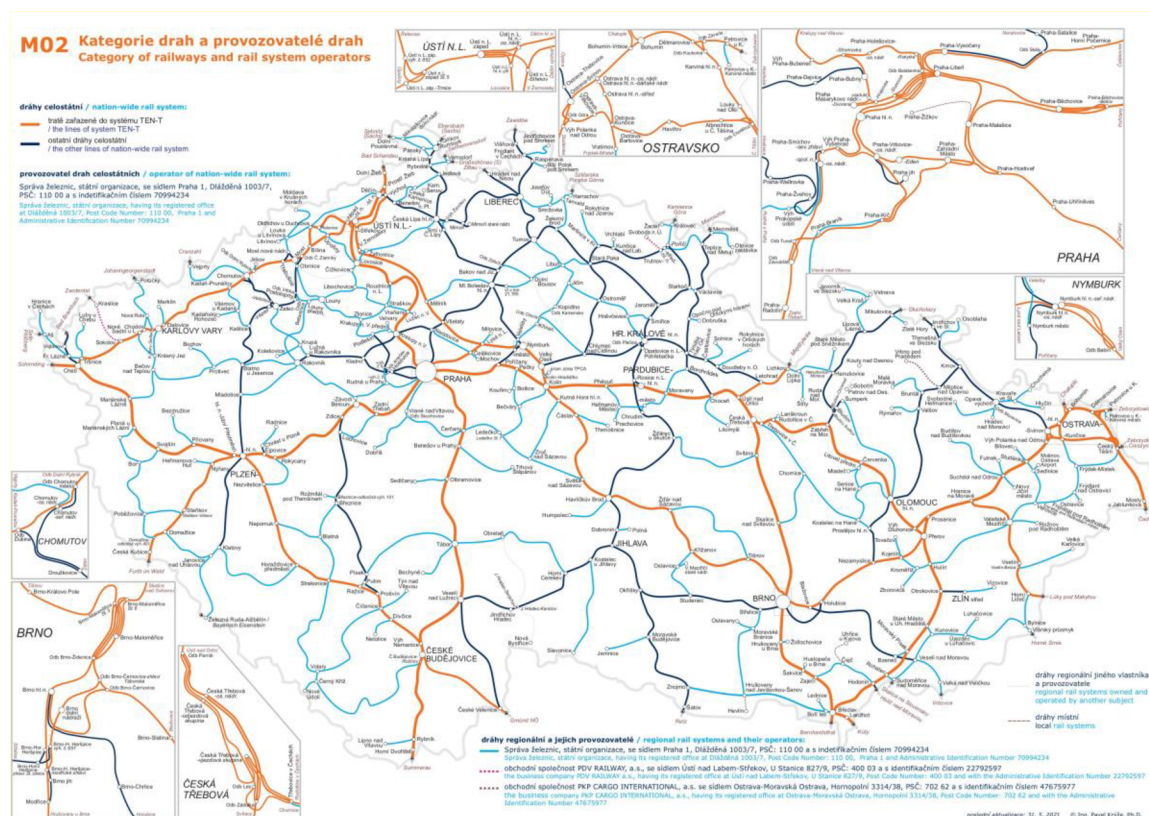
Železniční dráhy se z hlediska významu, účelu a technických podmínek, stanovených prováděcím předpisem, člení do jednotlivých kategorií. Kategoriemi železničních drah jsou podle zákona č. 266/1994 Sb., o dráhách, v platném znění:

- dráha celostátní, již je dráha, která slouží mezinárodní a celostátní veřejné železniční dopravě a je jako taková označena,
- dráha regionální, již je dráha regionálního nebo místního významu, která slouží veřejné železniční dopravě a je zaústěná do celostátní nebo jiné regionální dráhy,
- dráha místní, již je dráha místního významu oddělená od celostátní nebo regionální dráhy; dráha je oddělená, umožňuje-li přesun drážního vozidla na jinou dráhu jen s použitím zvláštního technického zařízení nebo slouží-li

výhradně provozování neveřejné osobní drážní dopravy, osobní drážní dopravy pro potřeby cestovního ruchu nebo provozované historickými vlaky,

- vlečka, již je dráha, která slouží vlastní potřebě provozovatele nebo jiného podnikatele a je zaústená do celostátní nebo regionální dráhy, nebo jiné vlečky,
- zkušební dráha, již je dráha, která slouží zejména k provádění zkušebního provozu drážních vozidel nebo zkoušek pro schválení typu nebo změny typu drážních vozidel a drážní infrastruktury,
- speciální dráha, která slouží zejména k zabezpečení dopravní obslužnosti obce.

„Ministerstvo dopravy považuje dráhu celostátní za zásadní pro dopravní obslužnost území České republiky dálkovou dopravou a rovněž pro spojení České republiky se zahraničím. Postupně na ní dochází k zavádění požadavků interoperability železničního systému tak, aby umožňovala bezpečný a nepřerušovaný provoz vlaků dosahujících stanovených úrovní výkonnosti. Dráha celostátní patří do vlastnictví České republiky. Dráhy regionální doplňují dráhu celostátní. Jsou určeny k zajištění dopravní obslužnosti regionů v oblastech, kde regionální doprava není zajišťovaná po dráze celostátní.“ (MDČR, 2022a).



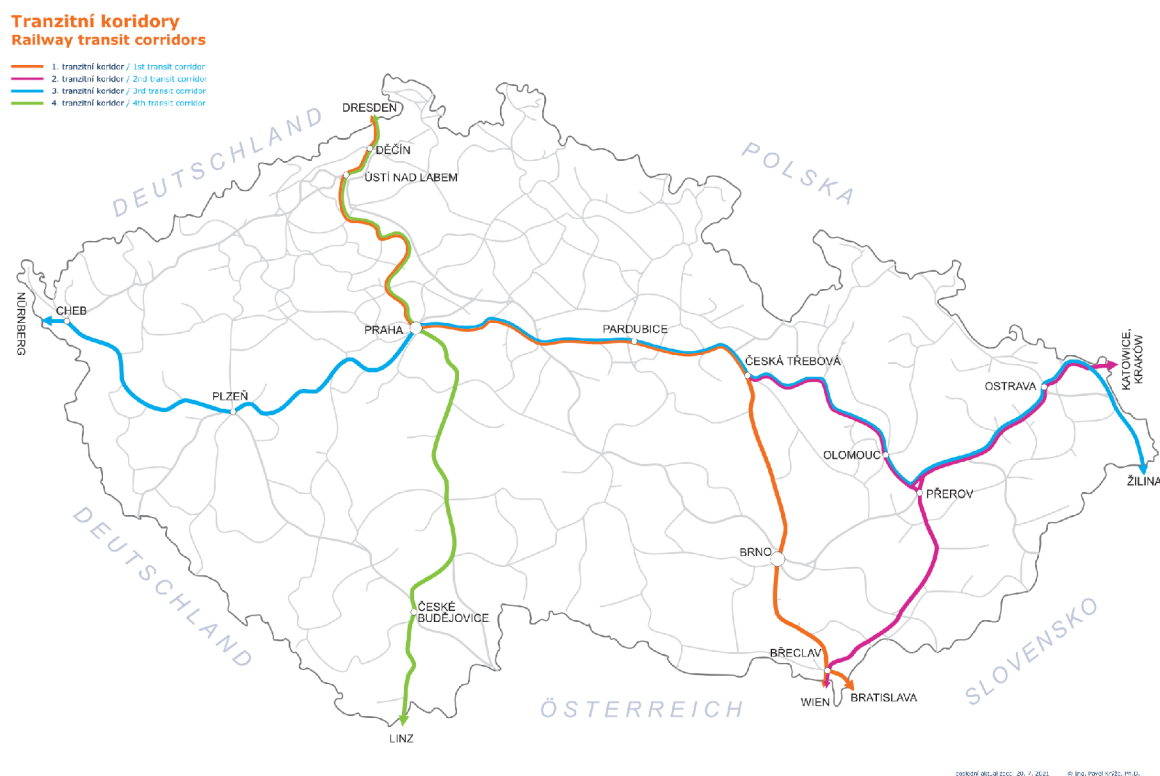
Obr. 2 Kategorie drah a jejich provozovatelé (zdroj: Správa železnic, 2021)

V Olomouckém a Moravskoslezském kraji působí tři provozovatelé drah. Společnost PKP CARGO INTERNATIONAL a.s., SART – stavby a rekonstrukce a.s. a Správa železnic, státní organizace. Společnost PKP CARGO INTERNATIONAL a.s. provozuje regionální trať 313 (označení v jízdním řádu) Milotice nad Opavou – Vrbno pod Pradědem. Jedná se o trať s délkou zhruba 20,5 km. Firma SART je provozovatelem regionální tratě 293 Šumperk – Kouty nad Desnou, včetně tříkilometrové odbočky Petrov nad Desnou – Sobotín. Souhrmně se pro tyto dva úseky tratě 293 užívá název Železnice Desná.

Vlastníkem tratě není stát, ale Svazek obcí údolí Desné. Od roku 2005 je jejím provozovatelem zmíněná firma SART – stavby a rekonstrukce a.s. (SART, 2022). Zbylé tratě na území obou krajů provozuje Správa železnic, státní organizace.

„Pro Českou republiku byly definovány čtyři tranzitní železniční koridory (Obr. 3), na nichž se provede prioritně modernizace v parametrech, které jsou obsaženy v evropských Dohodách AGC – o mezinárodních železničních magistrálách a AGTC – o nejdůležitějších trasách mezinárodní kombinované dopravy a souvisejících objektech, ke kterým ČR přistoupila a v plánech rozvoje železničních sítí vypracovaných na úrovni Evropské unie a Mezinárodní železniční unie.“ (MDČR, 2022b). Územím Olomouckého a Moravskoslezského kraje prochází dva koridory:

- 2. železniční koridor (Baltsko-jadranský koridor): Gdaňsk – Warszawa – Katowice – Petrovice u Karviné – Ostrava – Přerov – Břeclav
- 3. železniční koridor (Rýnsko-dunajský koridor): Le Havre – Paris – Frankfurt a.M. – Cheb – Plzeň – Praha – Ostrava – Žilina – Košice – Lvov



Obr. 3 Zjednodušená verze železničních tranzitních koridorů (zdroj: Správa železnic, 2019)

3.3 Řešené práce

Vyhodnocením dat provozu na železnici se již ve své diplomové práci s názvem *Analýza dat vytíženosti železniční infrastruktury* věnoval Jan Cinke (2020), student Katedry geoinformatiky Univerzity Palackého v Olomouci. Jelikož tato specifická data Správa železnic uvolnila poprvé, právě pro akademické účely na Katedře geoinformatiky v Olomouci, jedná se o ojedinělou práci na toto téma. Cílem práce bylo provést kategorizaci a časovou analýzu vytíženosti železniční infrastruktury s popsáním časové variability vybraných úseků v ČR. V rámci práce Jana Cinkeho byla vyhodnocována data pro území celé České republiky za období od roku 2016 až 2019.

V současné době se tématu provozu na železnici společně věnují doc. Ing. Zdena Dobešová, Ph.D. a Mgr. Michal Kučera z Katedry geoinformatiky Univerzity Palackého v Olomouci. V roce 2019 publikovali článek *Time Series of Workload on Railway Routes* (Dobešová, Kučera, 2019), ve kterém představují zpracování časových řad provozu železničních tratí v České republice, s popisem vybraných stanic z hlediska struktury a změn v provozu, jak v osobní, tak i nákladní vlakové dopravě. Nejnovější publikací je studie *Železniční doprava v roce 2020* (Kučera, Dobešová, 2022), která popisuje změnu v intenzitě vnitrostátní a mezinárodní vlakové dopravy v ČR v roce 2020 v porovnání s předchozími roky. Studie reaguje na světovou pandemii koronaviru a související restriktce, které omezily mobilitu obyvatel a průmyslovou výrobu, což významně ovlivnilo železniční dopravu v roce 2020.

Dopadům pandemie covidu-19 na železniční dopravě je věnován příspěvek *Dopad pandemie covid-19 na železniční sektor* (Haspra, 2021), který je součástí Vědeckotechnického sborníku Správy železnic. Příspěvek popisuje dopady protiepidemických opatření na osobní i nákladní vlakovou dopravu v České republice a Evropě.

3.4 Průběh pandemie nemoci covid-19 v ČR

„COVID-19 je označení pro infekci způsobenou koronavirem SARS-CoV-2, který se poprvé objevil koncem roku 2019 ve Wu-chanu v Číně. Během krátké doby se nové onemocnění nazvané později COVID-19 způsobené dosud nepopsaným koronavirem SARS-CoV-2 rozšířilo do celého světa.“ (NZIP, 2021). Jedná se o vysoce infekční respirační onemocnění, které se šíří především mezilidským kontaktem. Prevencí před šířením nemoci je tedy především omezení společenského kontaktu. První výskyt nemoci na území ČR byl potvrzen 1. března 2020 (MZČR, 2020).

Přehled nouzových stavů, vládních opatření a nařízení v souvislosti s pandemií koronaviru v letech 2020–2021. Vybrána jsou pouze opatření, které mají spojitost s mobilitou obyvatelstva:

- 1. nouzový stav (12. března 2020–17. května 2020)
 - 12. března 2020 - zákaz veřejných i soukromých akcí a zákaz vstupu veřejnosti do sportovních a kulturních zařízení
 - 14. března 2020 - zakázán provoz restaurací a některých obchodů
 - 16. března 2020 - omezení volného pohybu osob a uzavření státních hranic
 - 23. dubna 2020 - bylo znovu umožněno vycestovat do zahraničí
 - 11. května 2020 - otevření nákupních center a zahrádek restaurací
- 2. nouzový stav (5. října 2020–14. února 2021)
 - 5. října 2020 - zakázány hromadné akce, uzavřeny některé střední školy
 - 12. října - on-line výuka na VŠ, SŠ a ZUŠ v celé ČR; zastavení kulturních akcí, zrušeny veškeré profesionální i volnočasové aktivity sportovců
 - 23. října - omezení volného pohybu, uzavření maloobchodů a služeb
 - 28. října - zákaz vycházení ven mezi 21:00–04:59
 - 18. listopadu - do škol se vrací žáci 1. a 2. třídy, a žáci speciálních škol
 - 23. listopadu - obnovení výuky v části základních škol a skupinově na VŠ, zmírnění zákazu vycházení až od 23:00
 - 3. prosince - otevření obchodů, restaurací a služeb; konec zákazu nočního vycházení

- 18. prosince – zákaz volného pohybu osob
- 27. prosince – noční zákaz vycházení od 21:00, uzavřeny školy, obchod a služby (Vláda ČR, 2020)
- 3. nouzový stav (15. února 2021–26. února 2021)
- 4. nouzový stav (27. února 2021–11. dubna 2021)
 - 1. března – tvrdý lockdown – omezení pohybu pouze v rámci okresu
 - 12. dubna – konec lockdownu společně s omezením volného pohybu, otevřeny první stupně ZŠ
 - 10. května – otevřeny ZŠ a obchody
- 5. nouzový stav (26. listopadu 2021–25. prosince 2021) (Vláda ČR, 2021).

3.5 Doprava z Českého statistického úřadu

Český statistický úřad (ČSÚ) poskytuje statistické informace o dopravě na území ČR. Z volně dostupných informací lze sledovat vývoj různých ukazatelů o dopravě v uplynulých letech, včetně jejich předpokládaných změn v období pandemie covidu-19.

Ve Veřejné databázi ČSÚ (VDB) jsou vedeny *Indexy tržeb bez DPH v dopravě a skladování* (ČSÚ, 2022a). Indexy jsou vypočítány pro jednotlivá odvětví sekce Doprava a skladování obsažené v Klasifikaci ekonomických činností (CZ-NACE). Železniční doprava je vedena v rámci odvětví Pozemní a potrubní doprava (49) společně se silniční a potrubní dopravou. Nelze tedy vyhodnotit hodnoty pouze pro železniční dopravu, včetně jejího rozdělení na osobní a nákladní. Indexy jsou vedeny s měsíční, čtvrtletní a roční periodicitou. Uvedené indexy tržeb (Tab. 2) jsou vypočítány s bazickým indexem průměru stálých cen roku 2015 v měsíční periodě očištěné o sezónní a měsíční vlivy s měřicí jednotkou v %.

Tab. 2 Indexy tržeb bez DPH v dopravě a skladování

	2016	2017	2018	2019	2020	2021
Leden	98,3	100,0	106,3	109,1	105,4	95,4
Únor	98,2	101,8	104,9	108,3	106,4	95,2
Březen	99,0	102,8	106,4	111,2	92,4	98,0
Duben	98,7	101,2	104,8	107,9	75,8	97,6
Květen	97,6	103,9	107,1	107,7	81,7	97,2
Červen	100,6	101,0	108,1	105,4	87,4	97,8
Červenec	99,3	101,6	108,8	104,1	92,2	99,3
Srpen	101,0	103,3	108,0	103,7	93,5	98,9
Září	102,1	101,9	106,2	104,3	96,4	100,5
Říjen	101,5	102,9	106,5	103,7	96,6	100,9
Listopad	102,0	102,7	104,5	103,1	97,2	103,4
Prosinec	101,7	105,4	106,5	102,4	93,9	104,5

(zdroj: ČSÚ, 2022a)

V hodnotách indexu je viditelný propad tržeb v letech 2020 a 2021. Výrazný propad nastal v březnu a dubnu roku 2020, kdy došlo k propuknutí covidu-19 na území České republiky. Z únorové hodnoty 106,4 % se v březnu tržby propadly na 92,4 %, a v dubnu až na 75,8 %, což je absolutním minimum sledovaného období. Hodnoty indexu pod 100 % se držely až do srpna roku 2021. Od následujícího měsíce měly hodnoty stoupající trend. Hodnoty indexu tržeb v dopravě a skladování v období 2016–2021 potvrzují, že pandemie výrazně negativně ovlivnila oblast dopravy. Dá se tedy předpokládat, že ovlivnění se bude týkat i konkrétně železniční dopravy.

Další statistikou z VDB je *Přeprava věcí a osob, přepravní výkony* (ČSÚ, 2022b). Tento ukazatel již rozlišuje konkrétní druhy dopravy. Lze tedy vyhodnotit hodnoty pro osobní i nákladní železniční dopravu za celé území ČR. Pro osobní vlakovou dopravu jsou sledovány počty přepravených osob a přepravní výkon. Přepravní výkon pro osobní dopravu znamená součet vzdáleností, na které byli jednotliví cestující přepraveni. Pro nákladní dopravu je sledována hmotnost přepravených věcí a přepravní výkon. V případě nákladní dopravy přepravní výkon znamená součet jednotlivých součinů hmotnosti přepravených věcí a vzdálenosti, na kterou byli věci přepraveny. Statistické výstupy jsou poskytovány ve čtvrtletní či roční periodicitě.

Tab. 3 Přeprava věcí a osob, přepravní výkony

		Osobní železniční doprava		Nákladní železniční doprava	
		Přepravené osoby (tis. osob)	Přepravní výkon (mil. oskm)	Přepravené věci (tis. tun)	Přepravní výkon (mil. tkm)
2016	1. čtvrtletí	43 078	2 010	23 502	3 735
	2. čtvrtletí	46 174	2 241	24 636	3 900
	3. čtvrtletí	45 514	2 367	24 385	3 874
	4. čtvrtletí	44 404	2 226	25 510	4 109
2017	1. čtvrtletí	43 929	2 064	24 043	4 032
	2. čtvrtletí	46 985	2 437	23 868	3 825
	3. čtvrtletí	46 324	2 575	24 147	3 993
	4. čtvrtletí	45 786	2 421	24 458	3 993
2018	1. čtvrtletí	45 185	2 221	24 136	4 059
	2. čtvrtletí	49 107	2 667	24 152	4 063
	3. čtvrtletí	47 089	2 720	25 204	4 164
	4. čtvrtletí	48 155	2 678	25 815	4 278
2019	1. čtvrtletí	47 197	2 379	25 368	4 259
	2. čtvrtletí	50 053	2 909	25 280	4 009
	3. čtvrtletí	49 221	2 995	23 833	3 918
	4. čtvrtletí	47 740	2 787	24 046	3 884
2020	1. čtvrtletí	38 988	2 054	23 233	3 912
	2. čtvrtletí	25 934	1 137	20 869	3 527
	3. čtvrtletí	39 053	2 149	21 898	3 665
	4. čtvrtletí	25 167	1 144	24 863	4 033
2021	1. čtvrtletí	22 715	969	23 986	3 995
	2. čtvrtletí	32 512	1 488	24 622	4 144
	3. čtvrtletí	38 820	2 381	23 520	3 794

(zdroj: ČSÚ, 2022b)

U osobní vlakové dopravy došlo k prvnímu výraznému snížení počtu přepravených osob oproti předchozím obdobím v prvním čtvrtletí roku 2020. Následující čtvrtletí byl propad ještě výraznější. Ve druhém čtvrtletí roku 2020 se počet cestujících oproti stejnému období přechozího roku snížil téměř o polovinu. Nejméně přepravených osob bylo v prvním čtvrtletí roku 2021 s počtem necelých 23 milionů cestujících. V období od roku 2020 až do 3. čtvrtletí 2021 se počty cestujících nedostaly k hodnotám okolo 45 milionů, které byly běžné v předchozích letech. Opatření zavedená v souvislosti s pandemií omezila mobilitu obyvatel, kteří následně necestovali na delší trasy. To potvrzují hodnoty přepravního výkonu osobní dopravy, kdy od roku 2020 dochází k výrazným propadům. V prvním čtvrtletí 2021 je hodnota nejnižší za celé sledované období. Vysvětlením může být zákaz pohybu obyvatel mezi okresy, který byl na tři týdny vyhlášen v březnu 2021 (MVČR, 2021). U nákladní dopravy nejsou viditelné přílišné změny. K nejvýraznějšímu výkyvu v hmotnosti přepravených věcí došlo ve 2. čtvrtletí

2020, tedy na začátku pandemie v ČR. Propad oproti 2. čtvrtletí z předchozích let činí zhruba 5 milionu tun. Z hodnot přepravených věcí a přepravního výkonu lze usuzovat, že nákladní železniční doprava nebyla pandemií a opatřeními příliš ovlivněna.

3.6 Časové řady

Podle Křivého (2012) pod pojmem časová řada rozumíme data (výsledky pozorování), která jsou chronologicky uspořádána, např. seismický záznam v geofyzice, řada nejvyšších (nejnižších) denních teplot v meteorologii, vývoj koncentrace nečistot v ekologii, změny počtu jedinců nějaké populace v demografii, vývoj rozvodovosti v sociologii nebo vývoj cen v ekonomii. Máme přitom na mysli tzv. statistické (stochastické) řady, které jsou zatíženy nejistotou, nikoliv řady deterministické, jejichž chování lze jednoznačně popsat nějakým matematickým vzorcem. Jestliže vyjdeme z teorie náhodných procesů, lze říct, že časová řada představuje konkrétní realizaci odpovídajícího náhodného (stochastického) procesu. Hančlová a Tvrdý (2003) uvádí, že ze statistického hlediska může časová řada podléhat změnám v průměru či variabilitě, což je řada nestacionární, nebo být stále stejná tudíž řada stacionární. U stacionární řady nejsme na základě zjištěných statistických parametrů jako je aritmetický průměr hodnot nebo jejich rozptyl schopni rozlišit jeden úsek řady od druhého. Oproti tomu nestacionární řada vykazuje změny v chování. Například aritmetický průměr z hodnot ze začátku řady je signifikantně jiný než průměr hodnot na konci, což znamená, že taková řada vykazuje trend.

3.6.1 Problémy analýzy časových řad

Při analýze časových řad se vyskytuje množství problémů, které jsou pro tuto metodu zpracování dat specifické. Hančlová a Tvrdý (2003) ve své publikaci uvádí především problémy:

- s volbou časových bodů,
 - okamžikové
 - intervalové
- s kalendářem,
 - různá délka měsíců
 - různý počet víkendů v měsíci
 - různý počet pracovních dnů v měsíci
 - pohyblivé svátky
- s délkou časových řad,
- nesrovnalostí dat.

Při volbě časových okamžiků pozorování se doporučuje volit kompromisní řešení. Velká hustota časových bodů pozorování umožňuje dobře vystihnout charakteristické rysy časové řady, ale mohou nastat potíže při výpočtech. V každém případě je snahou volit ekvidistantní intervaly mezi sousedními pozorováními (Křivý, s. 12, 2012). Problémy s kalendářem, jako je různá délka kalendářních měsíců, různý počet pracovních dní v měsíci, pohyblivé svátky (např. Velikonoce), mohou mít překvapivé následky. Řady je však možné od těchto problémů „očistit“ jako například:

- vyrovnání různého počtu dní v měsíci,

$$y_t^{(\text{očištěná})} = y_t \frac{\bar{p}_t}{p_t}$$

kde y_t ... hodnota očišťovaného ukazatele,

p_t ... počet pracovních dní v měsíci t ,

\bar{p}_t ... průměrný počet pracovních dní v měsíci za rok či jiný základ př. 30 dní

- některé krátkodobé nepravidelnosti v kalendáři mohou být odstraněny pomocí agregace např. použití čtvrtletní agregované hodnoty namísto původních měsíčních údajů (Hančlová, Tvrđý, s. 11, 2003).

Problémy s délkou časových řad souvisí s počtem pozorování při analýze časových řad. Některé analýzy vyžadují určitou minimální délku řady (např. Boxův-Jenkinsův přístup předpokládá minimálně 50 pozorování), naopak u velice dlouhých řad je nebezpečí, že v průběhu tohoto časového období se mění charakteristiky modelu, a tudíž vnitřní struktura generující řadu se stává s rostoucí délkou obtížně modelována v případě modelů, které předpokládají stabilní chování parametrů (Hančlová, Tvrđý, s. 11, 2003). Krivý (2012) upozorňuje, že délka časové řady se definuje jako celkový počet pozorování v časové řadě, nikoliv jako časové rozpětí mezi prvním a posledním pozorováním. Problémy s nesrovnalostí jednotlivých měření souvisí s výběrovým vzorkem a zároveň reprezentativností tohoto vzorku i z hlediska časového vývoje (Hančlová, Tvrđý, s. 11., 2003).

3.6.2 Základní analýzy časových řad

Krivý (2012) ve své publikaci zmínil čtyři nejčastěji používané přístupy k analýze časových řad:

- dekompozice časové řady,
- Boxova-Jenkinsova metodologie,
- lineární kauzální (faktorové) modely,
- spektrální analýza časových řad.

Dekompozice časové řady spočívá v rozkladu časové řady na čtyři základní složky, jimiž jsou trend (Tr), sezónní složka (Sz), cyklická složka (C) a náhodná (reziduální) složka (ϵ). Z tohoto pohledu chápeme časovou řadu jako trend, na který se „nabalují“ periodické složky (sezónní a cyklická složka) a náhodná složka (nejčastěji bílý šum). Trend reprezentuje dlouhodobé změny v průměrném chování řady (dlouhodobý růst nebo pokles) a je způsoben faktory, které působí systematicky tj. ve stejném směru. Periodické změny, které se odehrávají v průběhu roku a každý rok se opakují, představují sezónní složku. Tyto změny souvisí především se střídáním ročních období. Pro studium sezónních vlivů se doporučuje pracovat s řadami měsíčních, nejvýše kvartálních měření. Cyklická složka je chápána jako nepravidelnosti okolo trendu, kdy se pravidelně střídá fáze růstu s fázemi poklesu. Délky a intenzita fází se mohou v průběhu času měnit. Příčiny, které způsobují cyklickou složku, lze jen těžko identifikovat. Nepravidelnosti způsobují také náhodnou neboli reziduální složku. Tyto

fluktuace jsou však náhodné a nemají systematický charakter. Složka zahrnuje také chyby měření či chyby ve zpracování dat. Náhodná složka může mít charakter bílého šumu, což znamená že je tvořena hodnotami nezávislých náhodných veličin s nulovou střední hodnotou a konstantním rozptylem (Křivý, 2012). Dekompoziční metoda se zabývá identifikací a modelováním především systematických složek, konkrétně trendové a sezónní složky (Hančlová, Tvrdý, 2003).

Boxova-Jenkinsova metodologie předpokládá, že časová řada je slabě stacionární. Základním prvkem modelu je zde reziduální složka. Tato metoda umožňuje modelovat časové řady, které mají zřetelný trendový či sezónní charakter. Oproti dekompozici umožňuje modelovat trendovou i sezónní složku stochasticky. Modely Boxovy-Jenkinsovy metodologie jsou proti dekompozičním modelům flexibilnější. To znamená, že se lépe adaptují na změny v průběhu časové řady (Křivý, 2012).

Lineární dynamické modely fungují tak, že se hodnoty sledované časové řady vysvětlují pomocí jiných řad, které se označují jako faktorové časové řady (Křivý, 2012). „Rozdílem od modelu Box-Jenkinse spočívá v tom, že zde kromě popisované časové řady a bílého šumu vystupují ještě další časové řady – příčinné faktory.“ (Hančlová, Tvrdý, 2003).

Spektrální analýza časových řad využívá na rozdíl od předchozích tří metod odlišný přístup, kdy se analyzovaná časová řada považuje za kombinaci sinusových a kosinusových křivek s různými amplitudami a frekvencemi. Pomocí speciálních statistických nástrojů lze získat představu o intenzitě zastoupení jednotlivých frekvencí v časové řadě (tzv. spektrum řady). Často hovoříme o tzv. fourierovské analýze (Hančlová, Tvrdý, 2003).

3.7 Shluková analýza

„Shluková analýza zkoumá, zda se množina objektů $O = \{O_1, O_2, \dots, O_m\}$ zadaných reálnými atributy $A = \{A_1, A_2, \dots, A_n\}$ přirozeně rozpadá na výrazné podmnožiny objektů si podobných, a přitom nepodobných objektům shluků ostatních. Pokud takové podmnožiny existují, nazýváme je shluky.“ Shlukování je tedy úloha, při které dochází k rozdělování objektů do skupin. Kritéria klasifikace nejsou při této úloze známa nebo jsou právě předmětem výzkumu. Současně nejsou známy ani vzory objektů reprezentujících budoucí podmnožiny. Pokud shluky existují, lze je interpretovat jako hledané klasifikační třídy (Šarmanová, 2013).

Za základní problémy vznikající při shlukové úloze Šarmanová (2013) uvádí:

- výběr atributů charakterizujících podobnost objektů,
- měření podobnosti a nepodobnosti (vzdálenosti) objektů
 - koeficienty korelace, asociace
 - metriky
- pojem shluku a jeho geometrický model,
- počet shluků rozkladu, počáteční rozklad,
- pojem vzdálenosti shluků,
- formulace řešené úlohy, důvod shlukování.

Pro hledání skupin podobných objektů v dané množině je potřeba objekty nejdříve vhodně charakterizovat takovými atributy, podle kterých se jejich podobnost rozezná. Prvním krokem je tedy volba relevantních atributů, které jsou homogenní ke zvolenému

hledisku, z jakého se mají objekty klasifikovat. Při výběru velkého počtu atributů by skupin s podobnými všemi vlastnostmi bylo příliš mnoho a byly by příliš malé, což bývá nežádoucí. „Při hledání podobných objektů vycházíme z úvahy, že dva objekty jsou si tím podobnější, čím více atributů pro ně nabývá stejných nebo blízkých hodnot.“ Podobnost lze zjednodušeně chápat jako vzdálenost mezi body v n -rozměrném prostoru, které reprezentují dané objekty. Čím větší vzdálenost, tím nepodobnější jsou si objekty. Ve většině metod shlukové analýzy se užívá pojem míra vzdálenosti nebo míra podobnosti objektů. Míra vzdálenosti neboli míra nepodobnosti je používána častěji nežli míra podobnosti. Pro vyjádření podobnosti objektů se používají koeficienty asociace nebo koeficienty korelace. Pro měření vzdálenosti objektů se používají metriky (Šarmanová, 2013).

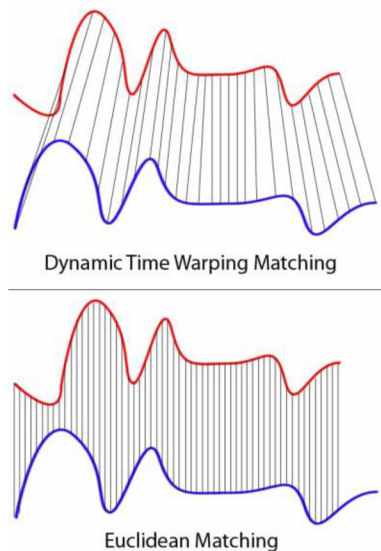
3.7.1 Nehierarchické a hierarchické metody

Podle cíle shlukování rozlišujeme nehierarchickou metodu a hierarchickou metodu shlukování. Metody nehierarchické hledají nejlepší prostý rozklad množiny objektů na shluky. Rozkládají množinu do podmnožin dle předem daného kritéria. První rozklad na podmnožiny se dále nedělí a nevytváří se hierarchická struktura. Na počátku nehierarchického shlukování je nutné zjistit optimální počet shluků, do kterých budou objekty rozřazeny. Dle počtu shluků se nehierarchické metody dále dělí na metody s pevně daným počtem shluků a na metody měnící počet shluků za běhu. K zjištění optimálního počtu shluků se využívají indexy vypočtené na základě proměnlivého faktoru K . Jedná se například o Calinski-Harabascův index, C index, nebo Goodman-Kruskal. Mezi metody s pevně definovaným počtem shluků patří například Forgýova a Janceyova metoda, MacQueenova k -středová metoda (K -means), sférické k -středové metody s opakovaným pūlením nebo fuzzy k -středové shlukování. Mezi metody s proměnlivým počtem shluků patří například MacQueenova metoda se dvěma parametry nebo metoda ISODATA (Kučera, 2008).

Principem hierarchické metody je vytváření stromové struktury jednotlivých shluků objektů. Hierarchické metody se dělí na dva druhy. Metoda aglomerativní je založena na tom, že na počátku je každý objekt samostatným shlukem. Postupně se v každém kroku spojí dva nejpodobnější shluky. První shluk je vytvořen ze dvou objektů na základě matice vzdáleností. Následně je vzdálenost dalších shluků určována zvolenou metodou pro určování vzdáleností shluků. Algoritmus končí, když jsou všechny objekty sloučeny do jednoho shluku. Divizivní metoda vychází naopak z toho, že na počátku jsou všechny objekty v jednom shluku, a v každém kroku dochází k rozdělení na další dva. Algoritmus končí v momentě, když je docíleno n jednoprvkových shluků (Petr, 2014). Hierarchické shlukování využívá několik metod, podle kterého jsou vybírány nejpodobnější shluky. Metoda nejbližšího souseda vytvoří shluk z objektů nebo shluků, které mají mezi sebou nejmenší vzdálenost v porovnání s ostatními objekty nebo shluky. Metoda nejvzdálenějšího souseda vytvoří shluk z objektů nebo shluků, které mají mezi sebou naopak největší vzdálenost v porovnání s ostatními objekty nebo shluky. Centroidní metoda, využívá euklidovské metriky pro změření vzdálenosti těžišť shluků a následného sloučení shluků, které mají nejmenší vzdálenosti mezi těžišti. Metoda průměrné vazby, počítá podobnost dvou shluků jako průměr vzdáleností mezi každými dvěma objekty patřícími do dvou různých shluků. Mediánová metoda, určuje podobnost dvou shluků ze vzdálenosti mediánů dvou shluků. Wardova metoda, využívá součet druhých mocnin odchylek každého objektu od těžiště shluku, do kterého náleží (Kučera, 2008).

3.7.2 Dynamické borcení časové osy

Dynamické borcení časové osy (Dynamic Time Warping, DTW) je algoritmus pro měření podobnosti dvou časových posloupností. Tato metoda hledá optimální zarovnání dvou časově závislých sekvencí, podle předem definovaných pravidel. Zjednodušeně se jedná o nelineární spárování sekvencí, aby si co nejvíce odpovídaly. Použití Euklidovské vzdálenosti je nevhodné při porovnání dvou sekvencí podobného tvaru, ale odlišného časového uspořádání, jelikož se jedná o lineární transformaci, která porovnává prvky na stejných indexech, tedy protější prvky obou sekvencí. Z toho důvodu dojde k ovlivnění výpočtu podobnosti i při mírném posunu sekvence, který nemusí mít v systému žádný význam (Procházka, 2013). Vlivem nestejného zarovnání na časové ose můžeme chybně označit totožné průběhy, které jsou vůči sobě jen posunuty, za nepodobné. Naopak nelineární metoda DTW umožňuje přizpůsobit tyto průběhy zavedením časově nelineární tzv. borbivé funkce (Černošus, 2016). Pro eliminaci časových rozdílů mezi sekvencemi je tedy aplikováno borcení jedné z časových os způsobem, aby byla dosažena co nejlepší možná shoda s druhou osou (Procházka, 2013).



Obr. 4 Rozdíl v porovnávání hodnot mezi DTW a Euklidovou vzdáleností (zdroj: Amidon, 2020).

Podle Senina (2008) metoda DTW funguje na principu, kdy pro vyjádření vzdálenosti dvou různých prvků sekvencí je nutné definovat metriku (*local cost measure/local distance measure*), která podobným prvkům přiřadí nízkou hodnotu, a naopak méně podobným prvkům hodnotu vysokou. Tato vzdálenost, která je vypočítána pro každou dvojici x, y obou sekvencí, je sestavena do matice vzdáleností (*cost matrix*), ve které dojde k nalezení cesty mezi sekvencemi, tzv. deformační cesty (*warping path*). Tato cesta navíc musí splňovat tři podmínky, a to hraniční podmínku, podmínku monotónnosti a podmínku velikosti kroku. Hraniční podmínka definuje, že první a poslední prvky sekvencí budou spárovány na sebe navzájem, čímž dojde k zajištění, že budou zarovnány celé sekvence a nedojde k vynechání žádného z krajních bodů. Podmínka monotónnosti zaručuje stoupající trend deformační cesty. Podmínka velikosti kroku vyjadřuje kontinuitu pohybu prvků v matici, čímž zajišťuje, aby nebyl vynechán či duplikován žádný prvek vzdálenostní matice. Vyhovujících deformačních cest může být více než jedna. Pro konečné spárování je tedy potřeba najít nejvhodnější cestu, kterou je ta, která má minimální náklady. Pro efektivnější hledání je sestavena akumulovaná matice vzdáleností (*accumulated cost matrix*), která je tvořena prvky s nejnižší možnou cenou, za kterou se lze dostat na aktuální prvek z bodu x_1, x_2 .

4 PŘÍPRAVA DAT

Prvním krokem zpracování poskytnutých dat byla jejich příprava do podoby, ve které je bude možné následně analyzovat a vytvářet z nich výstupy. Nejprve byly vyselektovány pouze údaje za monitorovací body SR70 a železniční úseky, které se nachází na regionálních a celostátních tratích v Olomouckém a Moravskoslezském kraji. Tyto vybrané prvky byly předmětem dalšího zpracování, kdy pro ně byly vytvářeny měsíční a roční agregace pomocí kontingenčních tabulek. Pro potřeby časové analýzy a tvorby časových řad byla data očištěna na stejnou délku měsíce.

4.1 Výběr monitorovacích bodů SR70

Na začátku práce bylo potřeba definovat, které monitorovací body SR70 a železniční tratě se stanou předměty zpracování. Jelikož data jsou poskytována za celou Českou republiku, bylo potřeba z nich vyselektovat údaje pouze z monitorovacích bodů ležících v Olomouckém nebo Moravskoslezském kraji. Jelikož data neobsahují atribut, podle kterého by se dalo určit ve kterém kraji či jiné administrativní jednotce se bod SR70 nachází, byly body SR70 lokalizovány pomocí atributů s GPS souřadnicemi. V první fázi tohoto procesu bylo využito dat z roku 2021, protože se jedná o nejaktuálnější data. Cílem bylo získat výpis monitorovacích bodů SR70, které jsou v datech obsaženy. Jelikož má v datech každý bod vícenásobný výskyt, z důvodů záznamu hodnot pro různé druhy vlakové dopravy a pro každý měsíc v roce, nelze sloupec s identifikátorem nebo názvem SR70 pouze zkopírovat. Problém s duplicitami monitorovacích bodů byl vyřešen využitím funkce *UNIQUE* v programu Microsoft Excel. Tato funkce byla aplikována na atribut Výchozí dopravní bod SR70, Název výchozího dopravního bodu, GPS souřadnice výchozího dopravního bodu. Do nové tabulky byly tedy vypsány pouze jedinečné monitorovací body reprezentovány číselným identifikátorem z číselníku SR70, názvem a polohou ve formě GPS souřadnic. Podle názvů monitorovacích bodů bylo zjištěno, že data obsahují údaje z bodů, které se nenachází na území České republiky, ale na území sousedních států. Body, které leží na území ČR mají číselný identifikátor začínající dvojčíslím 54. Z tabulky byly následně odstraněny záznamy s identifikátorem začínajícím dvojčíslím 51 (Polsko), 86 (Slovensko), 80 (Německo) a 81 (Rakousko). Dále byly odstraněny záznamy, které neměly uvedenou hodnotu GPS souřadnic. Aby bylo možné data nahrát do prostředí ArcGIS Pro, bylo nutné upravit formát souřadnic. Došlo k rozdělení souřadnic do dvou sloupců, aby souřadnice x a y byly uloženy v oddělených sloupcích. Poté byly všechny speciální znaky pro stupně, minuty, vteřiny a jejich desetiny nahrazeny mezerou. Desetinná čárka byla nahrazena tečkou. Takto připravená tabulka byla dále zpracována v ArcGIS Pro. Poloha monitorovacích bodů byla vizualizována pomocí nástroje *Coordinate Table to Point*. Tento nástroj umožňuje zobrazit bodová data na základě jejich souřadnic v různých formátech. Pro parametr formátu vstupních souřadnic byl nastaven parametr *Degrees, Minutes, and Seconds – Two Fields*. Ten odpovídá souřadnicím ve formátu stupňů, minut, vteřin a jejich desetiny, a současně uložení v oddělených polích. Výstupem nástroje je nově vzniklá prostorová vrstva monitorovacích bodů SR70 na území celé České republiky. Některé body byly špatně lokalizovány a nacházely se na území vedlejších států. Příčinou byla špatná konverze souřadnic, z důvodu chybějících hodnot minut v původních souřadnicích. Tyto body byly manuálně upraveny doplněním nul na pozici minut a opětovném spuštění nástroje *Coordinate Table To Point*. Z této prostorové vrstvy již bylo možné vyselektovat konkrétní monitorovací body, které splňují kritéria zadání této diplomové práce. S využitím administrativní vrstvy krajů z databáze ArcČR® 500 v 3.3,

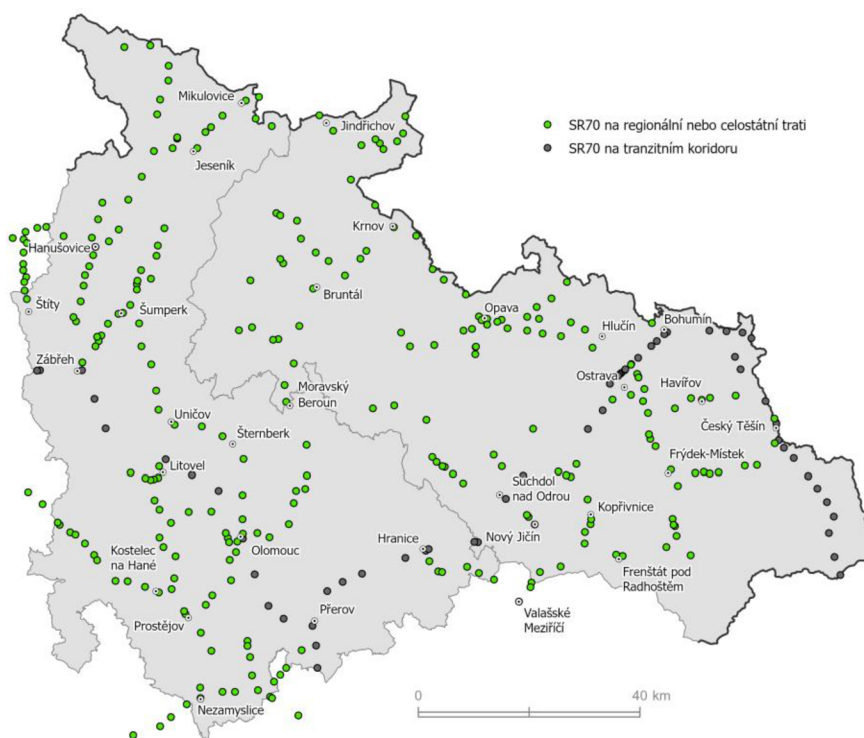
vrstvy železnice z databáze Data200 (©ČÚZK, 2022) a podkladových map Správy železnic (Správa železnic, 2022b) a Ortofota ČR (©ČÚZK, 2022) byly manuálně vybrány body, které leží na území Olomouckého nebo Moravskoslezského kraje. Hranice krajů však nebyly závazné. Do výběru byly na základě vlastního uvážení zahrnuty i některé body, které leží mimo zájmové území. Jedná se o body, které mají vazby na železniční tratě Olomouckého či Moravskoslezského kraje, a v případě jejich nezahrnutí by došlo k narušení kontinuity tratě. Příkladem je trať č. 300 Brno–Přerov, kde část úseku Přerov–Kojetín vede přes Zlínský kraj. Vynechání monitorovacích bodů z tohoto úseku by způsobilo rozdělení tratě na dvě části. Dalším příkladem je stanice Štítý, která leží na okraji Olomouckého kraje a přístupná je pouze tratěmi č. 024 Lichkov–Štítý a č. 025 Dolní Lipka–Hanušovice, které vedou přes Pardubický kraj. V případě nezahrnutí monitorovacích bodů z úseků vedoucích přes Pardubický kraj, by vznikl odlehlý bod. Z takto vybraného vzorku bodů bylo potřeba zanechat pouze ty, které leží na celostátních a regionálních tratích. Jelikož v datech nebyl dostupný klasifikátor, podle kterého by bylo možno body přiřadit k typu tratě, selekce proběhla opět manuálně, a to podle mapy železničních tranzitních koridorů a podkladových map. Výsledkem je tedy bodová vrstva 307 monitorovacích bodů SR70 z roku 2021 ležících na celostátních a regionálních tratích v Olomouckém a Moravskoslezském kraji. Zbylé body byly klasifikovány jako součást tranzitních železničních koridorů.

Situace na železnici se v průběhu času vyvíjí, a to se týká i samotných monitorovacích bodů. Některé mohou být vyřazeny z provozu nebo naopak některé nové do něj mohou být uvedeny. „Srovnání dat z různých let komplikují změny v umístění nebo označení měřících bodů, které vznikají například rekonstrukcí tratě.“ (Kučera, Dobešová, 2022). Jejich počet se tedy v průběhu let mohl změnit. Aby byla data porovnatelná, kompletní a zamezilo se, že pro daný monitorovací bod nebudou z některého roku data dostupná, bylo nutné zajistit, aby se zpracovávaný bod SR70 vyskytoval v datech z každého roku. Nejprve byl proces zpracování dat z roku 2021 aplikován i na data z předchozích let. Tím vznikly prostorové vrstvy bodů SR70 na celostátních a regionálních tratích pro každý rok a současně byl zjištěn jejich počet. Porovnáním bylo zjištěno, že počet monitorovacích bodů SR70 na regionálních a celostátních tratích (Tab. 4) se nejvíce změnil v roce 2020, kdy se oproti ostatním rokům jejich počet snížil zhruba o 20. Následně byly do nové tabulky pomocí funkce KDYŽ vypsány pouze body SR70, které se vyskytují současně v každém roku. Tuto podmínku splnilo 282 monitorovacích bodů na celostátních a regionálních tratích.

Tab. 4 Počet monitorovacích bodů SR70 v jednotlivých letech

Rok	Olomoucký a Moravskoslezský kraj	
	Počet SR70 na regionálních a celostátních tratích	Počet SR70 na tranzitních železničních koridorech
2016	302	63
2017	305	62
2018	306	64
2019	303	65
2020	285	59
2021	307	65

Výsledkem této fáze přípravy dat jsou tedy bodové vrstvy monitorovacích bodů SR70 na celostátních a regionálních tratích pro každý dílčí rok, a především společná bodová vrstva 282 monitorovacích bodů SR70 na celostátních a regionálních tratích obsažených ve všech poskytovaných letech zároveň. Pomocí těchto vrstev bylo možné vizualizovat naměřené hodnoty intenzity provozu společně s výsledky realizovaných analýz.



Obr. 5 Lokalizace monitorovacích bodů SR70 s údaji v každém roce současně

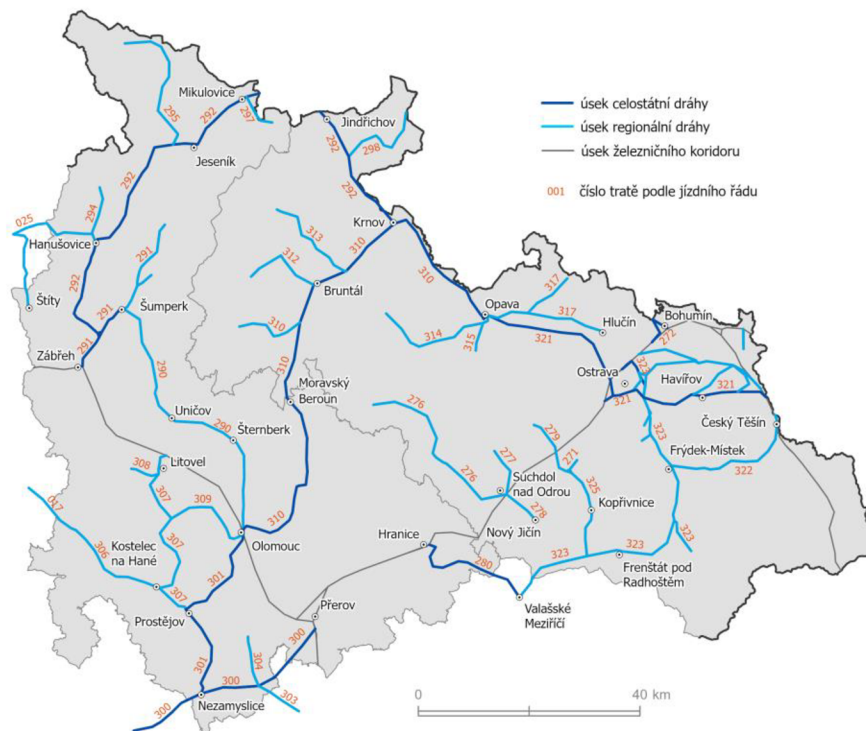
4.2 Výběr železničních úseků

Jak již bylo zmíněno, v datech jsou uvedeny výchozí a koncové dopravní body SR70. Díky tomu lze na data nahlížet i jako na úseky mezi těmito body. Z těchto úseků lze následně vytvořit jednoduchou železniční síť. Takovýto železniční úsek je v poskytovaných datech definován číselným identifikátorem společně s výchozím

dopravním bodem SR70 a koncovým dopravním bodem SR70. Železniční síť ve formě liniového modelu umožní vizualizovat údaje o provozu na železničním úseku. K výběru úsekových dat byla opět nejprve využita nejaktuálnější poskytnutá data z roku 2021. Na počátku, stejně jako při výběru monitorovacích bodů, bylo nutné odstranit duplicity a vypsát pouze jedinečné úseky uvedené v tabulce. Pro každý úsek byl zachován atribut Číslo úseku dle KKM, Výchozí dopravní bod SR70, Název výchozího dopravního bodu, GPS souřadnice výchozího dopravního bodu, SR70 koncového dopravního bodu, Název koncového dopravního bodu, GPS souřadnice koncového dopravního bodu. Stejně jako v případě výběru monitorovacích bodů, bylo nutné transformovat formát souřadnic. V tomto případě bylo potřeba zpracovat souřadnice výchozího i koncového dopravního bodu. Postup zpracování souřadnic byl totožný jako u monitorovacích bodů a je popsán v předchozí části práce. Připravená tabulka byla nahrána do ArcGIS Pro, kde byla vložena do nástroje *Coordinate Table To 2-Point Line*. Zde byly nastaveny atributy obsahující souřadnice výchozího a koncového dopravního bodu. Pomocí tohoto nástroje byla vytvořena liniová vrstva jednoduché železniční sítě celé ČR. Z této sítě byly vybrány úseky nacházející se v zájmovém území Moravskoslezského a Olomouckého kraje, které byly následně rozlišeny na úseky regionálních a celostátních tratí od úseků tvořících tranzitní koridory. Železniční síť pro rok 2021 je tvořena 462 úseky na celostátních a regionálních tratích, a 100 úseky na železničních koridorech. Pro tvorbu mapových výstupů, byla železniční síť vytvořena i z dat roku 2016. V roce 2016 síť tvořilo 459 celostátních a regionálních úseků, a 89 koridorových úseků. Ke každému úseku byl doplněn atribut čísla trati podle jízdního řádu pro rok 2021 (Správa železnic, 2021b).

Obdobně jako u monitorovacích bodů SR70, bylo snahou zajistit porovnatelnost a kompletnost dat pro analýzy. To znamená, vybrat takové úseky, pro které jsou dostupné údaje za každý rok z období 2016–2021. Hledání společných úseků probíhalo podobně jako u hledání společných monitorovacích bodů SR70. Z let 2016–2020 byly do nové tabulky vypsány unikátní úseky za celé území, a z roku 2021 pouze celostátní a regionální úseky. Pomocí funkce *KDYŽ* bylo nalezeno 434 celostátních a regionálních úseků, které jsou v každém roce. Z těchto úseků byla vytvořena železniční síť, která byla na první pohled na několika místech nespojitá. Přerušení linie tratě bylo způsobeno tím, že v datech, ze kterých byla železniční síť vytvořena, neexistoval úsek, který spojuje body, ve kterých došlo k přerušení. Daný úsek v datech nebyl, jelikož nebyl v datech pro každý sledovaný rok.

Ve výsledku tedy byla vytvořena liniová vrstva úseků regionálních a celostátních tratí v roce 2016 a 2021. Tyto vrstvy byly určeny především pro vizualizace, kdy byly analyzovány hodnoty pouze z těchto konkrétních let. Dále byl získán seznam 434 celostátních a regionálních úseků, které jsou společné pro každý rok. Tyto úseky bylo možné analyzovat napříč všemi roky, jelikož bylo zajištěno, že budou mít pro každý rok dostupná data. Z těchto úseků byla taktéž vytvořena prostorová vrstva, nicméně kvůli chybějícím úsekům nebyla vhodná k použití v mapových výstupech.



Obr. 6 Železniční síť v roce 2021

4.3 Tvorba agregačních tabulek

Zdrojová data jsou poskytována v měsíčních agregacích za konkrétní rok podle druhu dopravy. Každý monitorovací bod se v datech může vyskytovat vícekrát, jak již bylo uvedeno v části popisu dat. Pomocí kontingenční tabulky byly vytvářeny měsíční i roční agregace hodnot počtu vlaků a hmotnosti souprav pro identifikátor bodu SR70. Na tabulku byl aplikován filtr kategorie dopravy. To umožnilo vytvořit výsledné souhrny pro osobní a nákladní železniční dopravu zvlášť.

Tvorba agregačních tabulek pro úseková data byla až na jediný rozdíl totožná, jako pro monitorovací body SR70. Rozdíl spočíval v tom, že kontingenční tabulka nepočítala agregace pro identifikátor bodu SR70, ale pro číslo úseku, což je úsekový identifikátor. Podle nastavení kontingenční tabulky vznikaly měsíční a roční agregace hodnot počtu vlaků, průměrné délky soupravy, hmotnost souprav, průměrné hmotnosti soupravy, průměrného počtu náprav a průměrného počtu vozů pro úsek.

4.4 Očištění dat

Při tvorbě časových řad je nutné data ošetřit od různých kalendářních vlivů. Aby byla měsíční data porovnatelná, je potřeba je očistit na stejnou délku měsíce. Jelikož jsou zpracovávána data za období 6 let (2016–2021) nastává zde také vliv přestupných roků. V tomto případě jsou přestupné roky 2016 a 2020. Běžný nepřestupný rok má 365 dní, oproti tomu přestupný je o jeden den delší.

Časové řady byly plánovány pro ukazatele počtu a váhy osobních a nákladních vlakových souprav. Z toho důvodu stačilo očistit hodnoty pouze u těchto atributů. Přepočítání hodnot na stejnou délku měsíce bylo provedeno v MS Excel podle vzorce, který je uveden v odborné literatuře (Hančlová, Tvrđý, 2003):

$$y_t(\text{očištěná}) = y_t \frac{\bar{p}_t}{p_t}$$

kde:

- y_t hodnota očišťovaného ukazatele,
- p_t počet pracovních dní v měsíci t ,
- \bar{p}_t průměrný počet pracovních dní v měsíci za rok.

U nepřestupného roku (2017, 2018, 2019, 2021) je $\bar{p}_t = 365/12$ asi 30,41667.

U přestupného roku (2016, 2020) je $\bar{p}_t = 366/12$ tj. 30,5.

Na základě ústní konzultace vedoucí diplomové práce s Mgr. Ondřejem Vencálkem, Ph.D. z Katedry matematické analýzy a aplikací matematiky Přírodovědecké fakulty Univerzity Palackého v Olomouci, bylo z důvodu porovnatelnosti všech roků a jejich měsíčních hodnot bez ohledu na to, zda jsou přestupné nebo nepřestupné doporučeno hodnoty přepočítat konstantou průměrné délky měsíce pro 6 roků, a tedy:

$$\bar{p}_t = (366 + 365 + 365 + 365 + 366 + 365) / (6 * 12) = 2192/72 \quad \text{tj. } 30,4444$$

Tato konstanta průměrné délky měsíce byla doporučena jako vhodná pro konkrétní analyzované roky v této diplomové práci.

5 STRUKTURA ŽELEZNIČNÍ DOPRAVY

V části přípravy dat došlo k vybrání konkrétních železničních úseků a monitorovacích bodů SR70, ze kterých byly následně vytvořeny prostorové vrstvy. Tyto vrstvy bylo vhodné využít k prostorovým vizualizacím naměřených a analyzovaných hodnot. Na úvod bylo cílem vytvořit prosté vizualizace, které budou sloužit k prvotnímu seznámení s rozložením a strukturou železniční dopravy v Olomouckém a Moravskoslezském kraji. Díky tomu, byla vytvořena představa o provozu v regionu a zároveň mohly být vytipovány zajímavé železniční úseky, na které se lze zaměřit podrobněji v dalším zpracování. Nejdříve však byly vytvořeny agregační tabulky hodnot počtu či váhy vlakových souprav na úsecích, které byly následně v programu ArcGIS Pro připojeny k příslušným prostorovým vrstvám.

Celkově bylo vytvořeno 6 mapových výstupů. Jako první vznikly vizualizace celkového počtu a celkové hmotnosti osobních a nákladních vlakových souprav na úsecích pro rok 2016 a 2021, doplněny podílem jednotlivých druhů vlakové dopravy na celkovém počtu a hmotnosti. Tím byly úseky kategorizovány podle převládajícího druhu vlakové dopravy.

Dostupnost dat z let 2020 a 2021, tedy z období pandemie, bylo příležitostí pro zjištění změn a dopadu pandemie jak na osobní, tak i nákladní železniční dopravu v obou krajích. Z hlediska vládních restrikcí a samotného působení pandemie na mobilitu obyvatel byly v obou letech nejpostíženější především jarní měsíce. Předpokladem bylo, že v těchto měsících došlo v zasažených letech vůči předchozím ke snížení provozu na železniční infrastrukturu. Bylo tedy vytvořeno porovnání hodnot počtu vlakových souprav z období březen–květen 2020 a 2021 vůči průměrné hodnotě z totožného období v letech 2016–2019. Porovnání byla vytvořena pro osobní a nákladní vlakovou dopravu zvlášť, aby bylo možné postihnout vliv na jednotlivé druhy dopravy. V této kapitole jsou vloženy pouze náhledy vybraných výstupů, které byly pro dané téma nejnázornější a měly největší informační hodnotou. Všechny mapové výstupy v plné velikosti jsou přílohou této diplomové práce.

Pro mapové výstupy za rok 2016 byly využity pouze SR70 a úseky z dat roku 2016. Pro rok 2021, byly použity pouze vrstvy obsahující SR70 a úseky z roku 2021. Podíl druhu dopravy na celkovém počtu nebo hmotnosti vlakových souprav byl kvůli přehlednosti mapy vyhodnocen na základě hodnot v monitorovacích bodech SR70, ve kterých byl také zobrazen. Procentuální podíl druhu dopravy byl kategorizován do 5 kategorií po 20 %. Pokud byl v místě podíl osobní dopravy na celkovém počtu či hmotnosti 55 %, analogicky se nákladní doprava podílela 45 %.

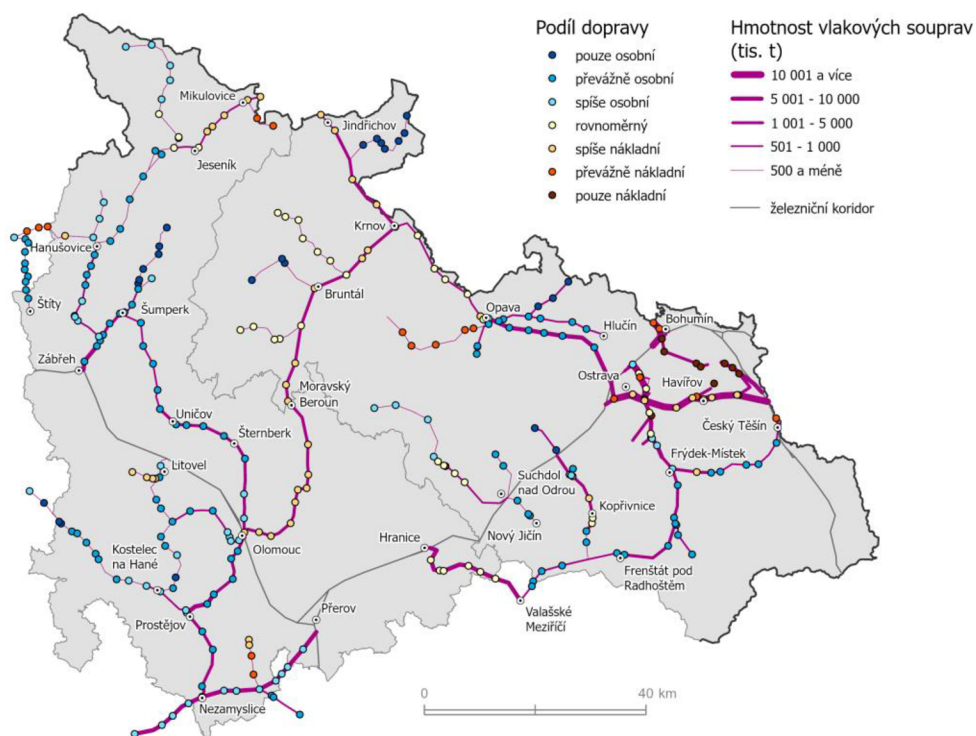
Tab. 5 Kategorizační tabulka procentuálního podílu dopravy

Podíl osobní dopravy [%]	Kategorie	Podíl nákladní dopravy [%]
100	Pouze osobní	0
79–99,9	Převážně osobní	0,1–19
60–79,9	Spíše osobní	20–39,9
40–59,9	Rovnoměrný	40–59,9
20–39,9	Spíše nákladní	60–79,9
0,1–19,9	Převážně nákladní	80–99,9
0	Pouze nákladní	100

Tab. 6 Úseky s největším počtem vlaků za rok 2016

Číslo úseku dle KKM	Název úseku	Celkový počet vlakových souprav
21320404	Ostrava-Stodolní z – Ostrava střed	45 852
21320402	Ostrava-Stodolní z – Ostrava báňské n.VOK	45 836
21320201	Ostrava uhelné n. – Ostrava báňské n.VOK	42 463
25210602	Haviřov – Šenov z	42 236
25210601	Šenov z – Ostrava-Bartovice	42 050

B. Podíl druhu dopravy na celkové hmotnosti vlakových souprav v roce 2016

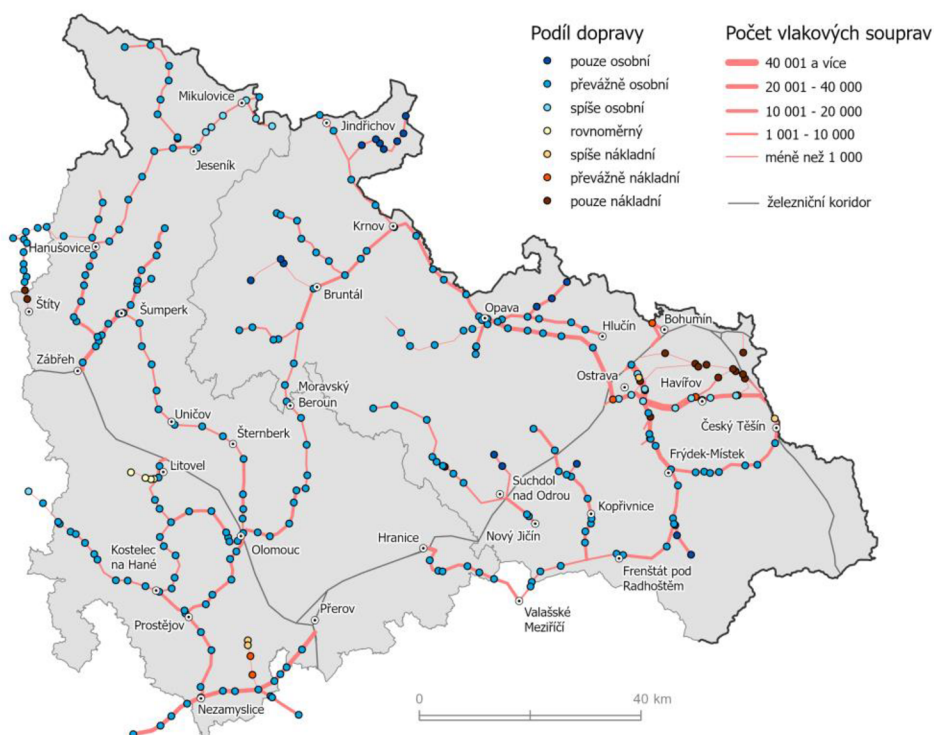


Obr. 8 Podíl druhu dopravy na celkové hmotnosti vlakových souprav v roce 2016

Železniční infrastruktura v okolí Ostravy byla v roce 2016 nejvíce frekventovaná i z hlediska celkové hmotnosti souprav. Ta na některých tratích přesáhla 10 milionu tun. Nefrekventovanějším úsekem Haviřov–Šenov z, projelo osobních a nákladních vlaků o celkové hmotnosti zhruba 25 milionu tun. Podíl na celkové hmotnosti v tomto úseku měla spíše nákladní doprava. Podle podílu dopravy na celkové hmotnosti je území heterogenní. Na Šumpersku, Olomoucku a Prostějovsku převládá podíl osobní dopravy. Vyvážený poměr hmotnosti osobních a nákladních vlaků byl například na trati mezi Hranicemi na Moravě a Valašským Meziříčím. Nákladní doprava převažovala například na trati mezi Olomoucí a Krnovem, či na zmíněném Ostravsku, kde na některých tratích nákladní vlaky tvořily celou hmotnost.

5.2 Železniční doprava v roce 2021

A. Podíl druhu dopravy na celkovém počtu vlakových souprav v roce 2021



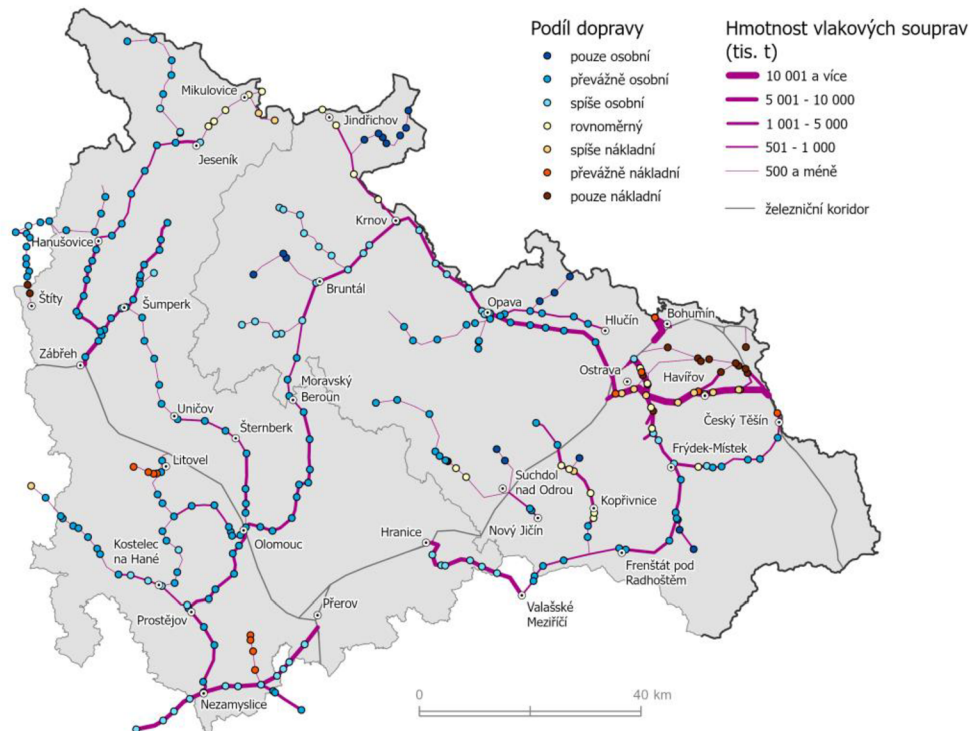
Obr. 9 Podíl druhu dopravy na celkovém počtu vlakových souprav v roce 2021

Je zřejmé, že v roce 2021 byl stejně jako v roce 2016, největší provoz na Ostravsku. Na vybraných úsecích zde projelo více než 40 tisíc osobních a nákladních vlakových souprav. Nejvyšší provoz za rok 2021 byl v úseku Ostrava-Stodolní z.–Ostrava střed, kterým projelo 52 656 vlakových souprav. Další tratě s vysokým provozem byly na Šumpersku, Opavsku a Přerovsku. Naopak nízký počet vlaků byl na regionálních tratích, především na těch s hlavovou (koncovou) stanicí, a v pohraničí. Několik úseků, kde projelo méně než 1 tisíc vlaků se nachází i v okolí Ostravy. Jedná se o úseky, které jsou obsluhovány pouze nákladní dopravou. Avšak na většině území měla na celkovém počtu největší podíl převážně osobní doprava. Na Ostravsku probíhala spíše osobní doprava, ale jelikož se jedná o významnou průmyslovou oblast, je v okolí několik úseků, kde dominovaly nebo jezdily pouze nákladní vlaky.

Tab. 7 Úseky s největším počtem vlaků za rok 2021

Číslo úseku dle KKM	Název úseku	Celkový počet vlakových souprav
21320404	Ostrava-Stodolní z – Ostrava střed	52 656
21320402	Ostrava-Stodolní z – Ostrava báňské n.VOK	49 682
21320602	Ostrava-Kunčičky z – Ostrava-Kunčice	49 222
21320601	Ostrava-Kunčičky z – Ostrava střed	48 953
21320201	Ostrava uhelné n. – Ostrava báňské n.VOK	43 063

B. Podíl druhu dopravy na celkové hmotnosti vlakových souprav v roce 2021

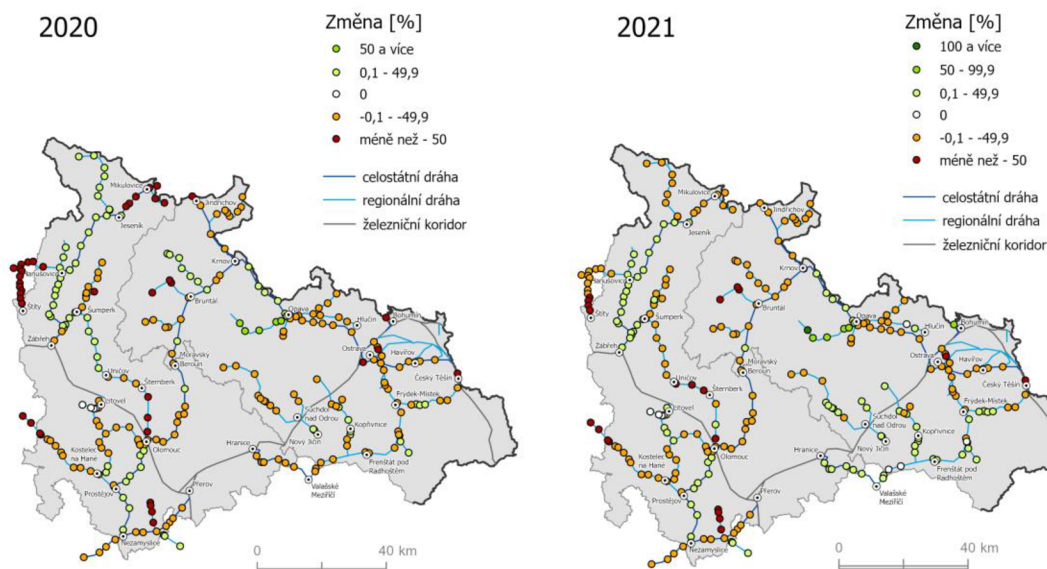


Obr. 10 Podíl druhu dopravy na celkové hmotnosti vlakových souprav v roce 2021

Celková hmotnost osobních a nákladních vlaků byla i v roce 2021 největší na Ostravsku, kde na ní většinou největší podíl měla nákladní doprava. V tomto roce byl nejvytíženější úsek Ostrava-Kunčice–Ostrava-Bartovice, s celkovou hmotností souprav necelých 22 milionu tun, což je o zhruba 3 miliony tun méně, než na nejvytíženějším úseku roku 2016. Z další části regionu patřily k nejvíce zatíženým opět úseky mezi Opavou a Ostravou, Hranicemi na Moravě a Valašským Meziříčím nebo Přerovem a Nezamyslicemi. Na první pohled oproti roku 2016 ubylo také úseků, kde byla hmotnost z větší části tvořena nákladní dopravou. Příkladem je trať vedoucí z Olomouce do Krnova. Naopak na většině území v roce 2021 převažuje podíl osobní dopravy.

5.3 Železniční doprava v březnu–květnu roku 2020 a 2021

A. Změna počtu osobních vlakových souprav v březnu–květnu 2020 a 2021 vůči průměru z let 2016–2019



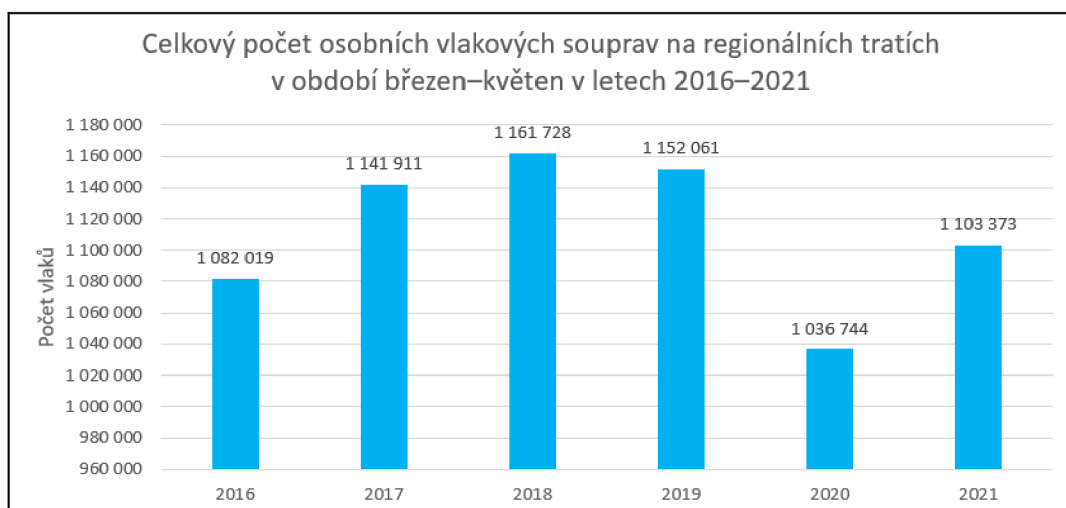
Obr. 11 Změna počtu osobních vlaků v březnu až květnu v letech 2020 a 2021

Počet vypravených osobních vlaků od března do května 2020 oproti průměru ze stejných měsíců předchozích let byl ve většině monitorovacích bodech nižší. K propadům došlo na regionálních i celostátních tratích a ve většině případů nebyly větší než 50 %. Jelikož v březnu 2020 bylo na určitou dobu vyhlášeno uzavření státních hranic kvůli pandemii koronaviru, došlo na hraničních přechodech k propadům v osobní dopravě o více než 50 %. K takto vysokým propadům došlo ještě na regionální trati do Štítů, Tovačova, z Bruntálu do Malé Morávky nebo na trati mezi Olomoucí a Šternberkem. Naopak o více než 50 % osobních vlaků projelo na regionální trati do Jakartovic na Opavsku. 1. března 2021 došlo na více než měsíc k omezení mobility obyvatel a byl až na výjimky zakázán jejich pohyb mezi okresy. To mohlo být příčinou, že i v březnu až květnu 2021 došlo ve většině monitorovacích bodech ke snížení počtu osobních vlaků vůči průměru. Oproti roku 2020 se však situace zlepšila na státních hranicích, kdy již neplatilo jejich uzavření, a propady v osobní dopravě už nebyly tak výrazné. V případě hraničního přechodu v Bohumíně naopak došlo k nárůstu osobních vlaků proti průměru o 85 %. Počty osobních vlaků narostly také na tratích na Valašsku. Nárůst o více než 100 % byl zaznamenán ve stanici Jakartovice s průměrným počtem 15 osobních souprav, kde projelo 32 osobních vlaků.

Před statistickým zhodnocením je nutné uvést, že se jedná o součet počtu vlaků ze všech úseků. Jedna konkrétní vlaková souprava může být započítána na více úsecích, kterými projela a je tedy v datech vícekrát. Nelze tedy výsledky diagramu přímo interpretovat tak, že jde o celkový počet vypravených vlaků, ale spíše o celkový počet průjezdů vlaků na celostátních a regionálních tratích.

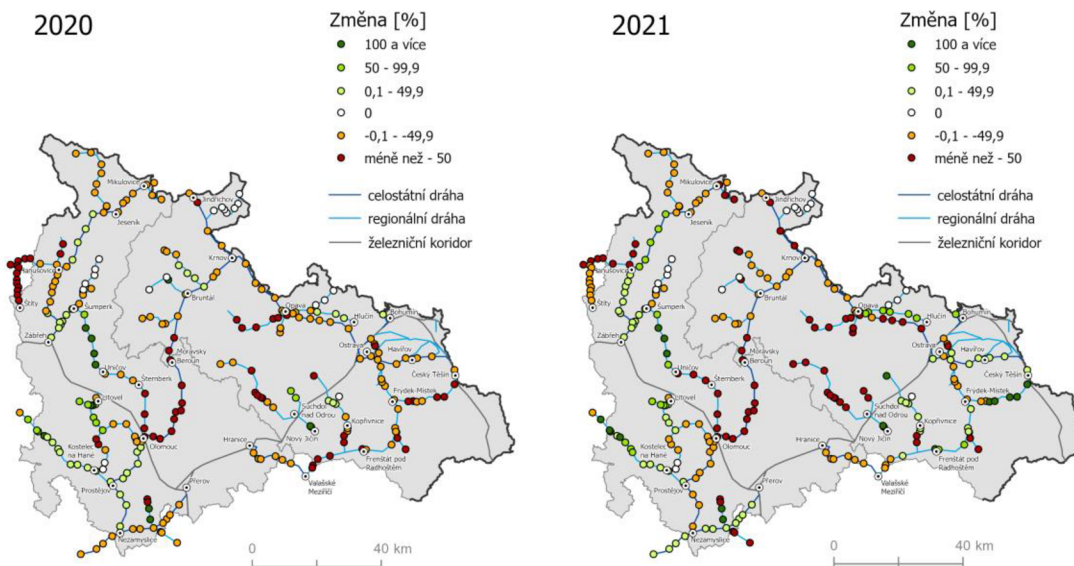
Statistické porovnání celkového počtu osobních vlakových souprav na celostátních a regionálních tratích v Olomouckém a Moravskoslezském kraji v období březen–květen let 2016–2021 vykazuje mezi roky 2016 až 2018 rostoucí trend.

V tomto období tedy provoz na železnici z hlediska osobní dopravy rostl. Po menším propadu v roce 2019 přišel výrazný propad v roce 2020, kdy se počet osobních vlakových souprav, oproti vrcholu za sledované období z roku 2018, snížil o zhruba 125 000. Jedním z očekávaných důvodů mohou být protiepidemická opatření, která omezila mobilitu obyvatel a tím i poptávku po osobní vlakové dopravě. V následujícím roce se již počet opět zvedá, avšak stále nedosahuje hodnot z předchozích let.



Obr. 12 Celkový počet osobních vlaků v březnu–květnu na všech regionálních tratích

B. Změna počtu nákladních vlakových souprav v březnu–květnu 2020 a 2021 vůči průměru z let 2016–2019

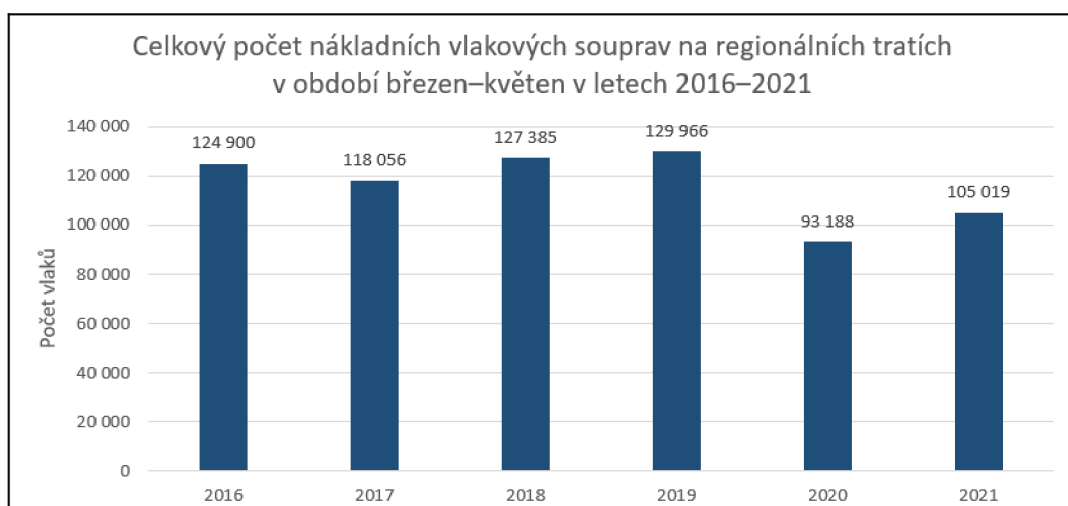


Obr. 13 Změna počtu nákladních vlaků v březnu až květnu v letech 2020 a 2021

U nákladní dopravy došlo v březnu–květnu 2020, stejně jako v případě osobní dopravy, k mnohačetným propadům počtu vypravených souprav. O méně než 50 % nákladních vlaků projelo například na úsecích v okolí Olomouce směrem na Šternberk nebo Krnov. Menší počet nákladních vlaků, než průměrně jezdil i na Ostravsku. V případě hraničních přechodů došlo v roce 2020 k poklesu i počtu nákladních vlaků. Výjimkou je

státní hranice v Bohumíně, kde došlo naopak k mírnému nárůstu. Počet nákladních vlaků se výrazně zvýšil na trati mezi Šumperkem a Uničovem, kde ve stanici Nová Hradečná projelo 103 souprav oproti průměrným 5. Větší počet nákladních vlaků je také patrný na tratích na Prostějovsku. V roce 2021 se situace na několika místech zlepšila, a to především na Ostravsku nebo na trati z Valašského Meziříčí do Frýdku-Místku. Výrazný nárůst zaznamenala například trať mezi Frýdkem-Místkem a Českým Těšínem, Opavou a Hlučínem, Přerovem a Nezamyslicemi nebo trať z Jeseníku směrem na Zábřeh na Moravě.

U celkového počtu nákladních vlakových souprav na regionálních tratích Olomouckého a Moravskoslezského kraje v období březen–květen let 2016–2021 byl rostoucí trend od roku 2017 do roku 2019. Menší propad, o necelých 7 tisíc souprav, nastal mezi lety 2016 a 2017. Rostoucí trend byl přerušen v roce 2020, kdy celkový počet nákladních vlakových souprav zaznamenal své minimum ve sledovaném období. Oproti předchozímu roku 2019, kdy byl počet ve sledovaném období nejvyšší, se v roce 2020 propadl o necelých 37 tisíc souprav na hodnotu 93 188. Obdobně jako u osobních souprav následující rok celkový počet vzrostl, nicméně nedosahuje hodnot, kterých nabýval před roky zasaženými pandemií covid-19.



Obr. 14 Celkový počet nákladních vlaků v březnu–květnu na všech regionálních tratích

6 ČASOVÁ ANALÝZA

Časová analýza osobní a nákladní vlakové dopravy na vybraných regionálních a celostátních železničních tratích Olomouckého a Moravskoslezského kraje byla realizována pomocí metody časových řad. Pro analýzu byly sledovány ukazatele počtu a váhy vlakových souprav. Cílem bylo odhalit úseky a stanice, které vykazují specifický provoz. Vytvořeny byly časové řady pro osobní i nákladní vlakovou dopravu. Ve fázi přípravy dat již došlo k ošetření problému různé délky měsíců. U vybraných úseků či stanic došlo k dekompozici časových řad, pro odhalení jednotlivých časových složek.

6.1 Výkyvy v provozu

Na počátku analyzování dat z hlediska jejich časové složky, bylo vhodné v datech vyhledat nulové hodnoty u vybraných ukazatelů. Nalezením nulových hodnot měsíčního počtu nebo váhy vlakových souprav v určitém měsíci je možné odhalit přerušení provozu na dané trati v důsledku výluky či jiné mimořádnosti. Pokud se nulové hodnoty v konkrétních měsících každý rok opakují, zřejmě se bude jednat o trať se specifickým provozem, kdy v daných měsících na trati neprobíhá osobní nebo nákladní vlaková doprava. Typicky se jedná o tratě se sezonním provozem, kdy zde probíhá osobní vlaková doprava pouze v letních měsících.

První možností hledání nulových hodnot v datech bylo využití nástrojů MS Excel. Data tedy byla analyzována ve formě tabulek, kdy první dva sloupce obsahovaly ID a název úseku či monitorovacího bodu, a první řádek obsahoval časovou složku. Hodnoty jsou očištěné od kalendářních vlivů. Pomocí podmíněného formátování s nastavením pravidel pro zvýraznění buněk, se formátovaly buňky, které jsou rovny hodnotě 0. Buňky s nulovými hodnotami byly označeny odstíny červené barvy. Tímto způsobem byly analyzovány počty osobních i nákladních vlakových souprav na úsecích i v monitorovacích bodech. Díky časové složce dat, lze snadno určit kdy k nulovému počtu došlo a zdali se jedná o ojedinělý či pravidelně se opakující výkyv v provozu. Jednoduše je tedy možné vytipovat úseky, na kterých probíhá sezonní provoz. Při aplikaci metody na data počtu osobních vlakových souprav na úsecích byl u některých úseků objeven společný vzor opakování nulových hodnot ve stejné časové období (Obr. 15). Na těchto úsecích ve stejných intervalech neprojela žádná osobní souprava. Z této informace je možné usoudit, že se jedná o úseky tvořící totožnou železniční trať.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	
1	Číslo úseku dle KKM	název_úsek	01.01.2016	01.02.2016	01.03.2016	01.04.2016	01.05.2016	01.06.2016	01.07.2016	01.08.2016	01.09.2016	01.10.2016	01.11.2016	01.12.2016	01.01.2017	01.02.2017	01.03.2017	01.04.2017	01.05.2017	01.06.2017	01.07.2017	01.08.2017	
143	21012401	Kojetín_Měrovice nad Hanou z	1763	1783	1792	1339	1930	1800	1707	1571	1864	1258	1823	1749	1808	1717	1709	1878	1776	1800	1717	1756	
144	21012402	Měrovice nad Hanou z_Němčice nad Hanou	1763	1783	1792	1339	1931	1800	1707	1570	1864	1258	1823	1749	1808	1717	1707	1878	1776	1800	1717	1756	
145	21012601	Chropyně_hr.VUSC 0710/0720 03	1753	1768	1796	1329	1921	1795	1695	1727	1866	1299	1828	1747	1798	1788	1762	1869	1765	1792	1768	1748	
146	21012603	hr.VUSC 0710/0720 03_Kojetín	1753	1768	1796	1329	1921	1795	1695	1727	1866	1299	1828	1747	1798	1787	1762	1869	1765	1792	1768	1748	
147	21012801	hr.VUSC 0710/0720 01_Velký	1754	1774	1793	1331	1921	1795	1700	1734	1866	1299	1828	1747	1796	1710	1712	1869	1763	1792	1768	1750	
148	21012803	Chropyně_hr.VUSC 0710/0720 02	1754	1774	1793	1331	1921	1795	1700	1734	1866	1299	1828	1747	1796	1710	1712	1869	1763	1792	1768	1750	
149	21013001	Věky_v_Hrozev předsínácká	1754	1774	1793	1331	1921	1795	1700	1734	1866	1299	1828	1747	1796	1710	1712	1869	1763	1792	1768	1750	
150	21110201	Kojetín_Km 2,800	0	0	0	0	24	0	0	24	0	24	0	8	0	0	0	0	0	0	0	26	
151	21110202	Uhřetice obec z_Km 2,800	0	0	0	0	24	0	0	24	0	24	0	8	0	0	0	0	0	0	0	0	26
152	21110204	Lobedice nž_Uhřetice obec z	0	0	0	0	24	0	0	24	0	24	0	8	0	0	0	0	0	0	0	0	26
153	21110401	Lobedice nž_vl_v km 8,231	0	0	0	0	24	0	0	24	0	24	0	8	0	0	0	0	0	0	0	0	26
154	21110402	Oplicany z_n_Km 8,231	0	0	0	0	24	0	0	24	0	24	0	8	0	0	0	0	0	0	0	0	26
155	21110403	Oplicany z_n_Km 8,231	0	0	0	0	24	0	0	24	0	24	0	8	0	0	0	0	0	0	0	0	26
156	21110404	Tevačov nž_Oplicany z	0	0	0	0	24	0	0	24	0	24	0	8	0	0	0	0	0	0	0	0	26
157	21120201	Kojetín_hr.VUSC 0710/0720 05	1016	1037	1017	1010	1040	985	1010	1033	1032	1025	1032	989	1035	1032	1041	1002	983	1037	1010	924	
158	21210202	Bozmorov z_Postupky z	1016	1037	1017	1010	1040	985	1010	1033	1032	1025	1032	989	1035	1032	1041	1002	983	1037	1010	924	
159	21210203	Kroměříž_Postupky z	1053	1067	1052	1042	1070	1014	1040	1051	1063	1053	1063	1063	1036	1069	1072	1050	1022	1010	1056	956	
160	21210205	hr.VUSC 0710/0720 05_Bozmorov z	1016	1037	1017	1010	1040	985	1010	1033	1032	1025	1032	989	1035	1032	1041	1002	983	1037	1010	924	
161	21310201	hr.VUSC 0720/0800 01_Hostavice	722	716	728	712	740	742	674	685	735	495	732	709	738	739	743	712	719	742	680	698	
162	21310202	AHr Křivová_Valašské Meziříčí	722	716	728	712	740	742	674	685	735	495	732	709	738	739	743	712	719	742	680	698	
163	21310205	hr.VUSC 0720/0800 01_AHr Křivová	722	716	728	712	740	742	674	685	735	495	732	709	738	739	743	712	719	742	680	698	
164	21310401	Hostavice_Mohlov hl.zraf.nz	722	741	738	727	740	742	373	31	735	682	732	716	738	739	743	712	730	743	680	698	
165	21310601	Mohlov hl.zraf.nz_Velovice	722	741	738	717	740	742	373	31	735	682	732	716	738	739	743	712	730	743	680	698	
166	21310802	Frenštát p.Rad.m.z_Velovice	1212	1232	1222	1216	1228	1222	664	82	1228	1144	1118	1194	1226	1231	1229	1210	1221	1233	1175	1187	
167	21310804	Frenštát p.Rad.m.z_Velovice	722	741	738	717	740	742	373	31	736	682	666	710	738	739	743	712	730	743	680	698	
168	21310101	Frenštát p.Rad.m.z_Kvondice pod Cindřej	1380	1401	1390	1389	1397	1311	1316	1351	1397	1398	1396	1352	1395	1400	1393	1374	1388	1400	1340	1256	
169	21310205	Frydlant n.Ostřavici_AHr Čeladná z	1380	1424	1401	1211	1417	1333	1356	1394	1418	1332	1418	1381	1415	1422	1316	1382	1407	1425	1358	1380	
170	21311202	AHr Čeladná z_Kvondice pod Cindřej	1378	1401	1382	1189	1397	1311	1316	1351	1396	1311	1396	1360	1395	1400	1292	1374	1387	1400	1340	1350	
171	21311401	vl_v km 102,205_Frydlant n.Ostřavici	1795	1799	1766	1778	1767	1696	1617	1656	1772	1623	1762	1702	1769	1778	1651	1721	1726	1808	1650	1661	
172	21311601	Balka_Přímno	1656	1701	1675	1667	1679	1613	1539	1572	1650	1543	1689	1620	1686	1695	1567	1642	1644	1705	1540	1549	
173	21311602	Balka_Přímno_Místek	1657	1703	1676	1667	1671	1682	1540	1572	1689	1643	1689	1623	1686	1695	1673	1642	1672	1705	1540	1564	
174	21320101	Ostřava levné_n_Ostřava uhebné n.	1	2	1	3	2	1	1	2	3	2	3	4	0	2	0	0	7	0	1	4	
175	21320102	Ostřava levné_n_Ostřava uhebné n.	0	0	0	0	2	1	2	1	1	1	87	191	75	67	0	94	121	67	0	6	
176	21320103	Ostřava uhebné n_Ova hl.n. Sc101,102	1179	1733	1802	1338	1185	1235	1195	1256	1214	1230	1445	1298	1284	1294	1428	1318	1351	1300	1272	1297	
177	21320201	Ostřava uhebné n_Ostřava balčíkové n.VDK	2807	2810	2818	2822	2840	2915	2710	2826	2881	2798	2896	2821	2969	2996	3021	2815	2956	3004	2834	2882	
178	21320402	Ostřava Stodolná z_Ostřava balčíkové n.VDK	2807	2813	2818	2821	2859	2915	2710	2826	2880	2797	2866	2822	2969	2996	3021	2815	2956	3003	2834	2882	
179	21320404	Ostřava Stodolná z_Ostřava střed	2806	2811	2817	2821	2860	2915	2710	2826	2880	2797	2866	2822	2969	2996	3021	2817	2956	3003	2834	2882	

Obr. 15 Nulové hodnoty počtu osobních vlakových souprav na úsecích

Výluka na trati nemusí nutně znamenat nulové hodnoty. Na trati může být zavedena pouze částečná výluka, která platí pouze pro konkrétní vlaky, v určitý časový interval nebo pouze pracovní dny. V tomto případě bude v daném měsíci výrazně nižší hodnota, avšak nebude nulová. Proto bylo nasnadě data analyzovat i metodou identifikace odlehlých hodnot, která odhalí i výrazně nižší, a zároveň i výrazně vyšší hodnoty. Identifikováním odlehlých hodnot měsíčního počtu nebo váhy vlakových souprav by mělo opět dojít k odhalení výkyvů a mimořádností v provozu na dané trati. Výrazně nízká či nulová hodnota počtu vlakových souprav v určitém měsíci může signalizovat, že v tomto období byla na daném úseku tratě zavedena výluka či došlo k jinému omezení provozu. Naopak odlehlá vysoká hodnota je znamením, že v daném úseku došlo z nějakého důvodu k prudkému nárůstu provozu. Tyto informace opět poslouží k nalezení tratí se specifickým provozem. K identifikování odlehlých hodnot je vhodné použít automatizované postupy. Řešením bylo vytvoření skriptu v programovacím jazyce R nebo sestavení modelu v programu Orange.

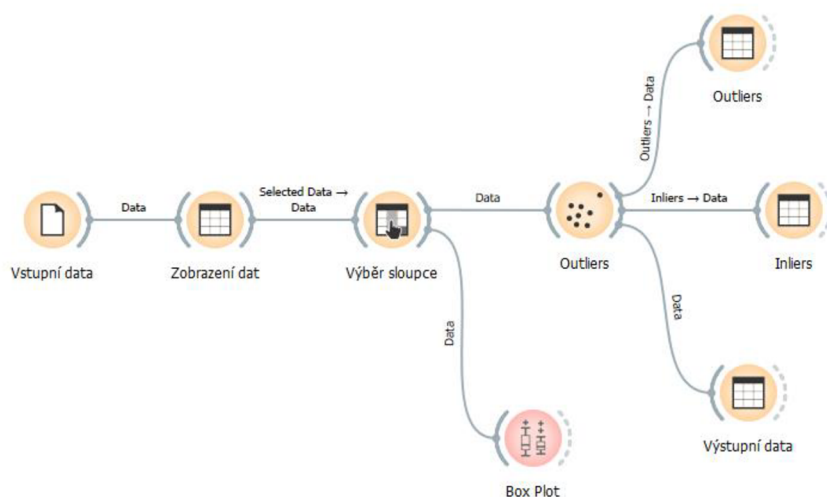
V prostředí programu R Studio byl napsán krátký skript, jehož funkcí je vyhledat odlehlé hodnoty ve vstupním datasetu. Vstupem je soubor ve formátu *.xlsx*, kterým je tabulka obsahující hodnoty počtu osobních nebo nákladních vlakových souprav na úsecích či v monitorovacích bodech. Pro správné zpracování je potřeba, aby tabulka byla transponovaná. První sloupec tabulky obsahuje datum, ostatní sloupce obsahují hodnoty pro jednotlivé úseky nebo body SR70. První řádek obsahuje názvy těchto úseků či monitorovacích bodů. Po nahrání tabulky, byl první sloupec obsahující datum použit pro názvy řádků. Poté byly vytvořeny dva prázdné seznamy (Obr. 16). Do jednoho z nich byly pomocí *for* cyklu ukládány hodnoty, které byly metodou *boxplot* identifikovány jako odlehlé. Do druhého seznamu byly indexy jejich pozic v tabulce. Pro spuštění takto připraveného skriptu byly doinstalovány dvě rozšíření. Jedná se o balíček *readxl*, který umožnil nahrání souboru *.xlsx*, a balíček *tidyverse*, pomocí jehož funkce byly přejmenovány řádky tabulky. Výsledkem skriptu je seznam odlehlých hodnot pro každý úsek nebo měřicí bod. Nevýhodou tohoto řešení je chybějící spojitost mezi časovým údajem a nalezenou hodnotou. Nelze tedy určit kdy k dané odchylce došlo, což je pro potřeby této analýzy podstatná informace.

```
9 outvalue<-list()
10 outindex<-list()
11 for(i in 1:ncol(df)){
12   outvalue[[i]]<-boxplot(df[,i])$out
13   outindex[[i]]<-which(df[,i] == outvalue[[i]])
14 }
```

Obr. 16 Část programového kódu pro detekci odlehlých hodnot

Druhým řešením bylo sestavení modelu v programu Orange. Na začátek modelu byl vložen uzel pro načtení dat. Data jsou načtena ve formátu tabulky, kde první sloupec obsahuje datum a ostatní sloupce hodnoty jednotlivých sledovaných entit. V dialogovém okně tohoto uzlu byl pro první sloupec s datumem nastaven datový typ *datetime*. Detekci odlehlých hodnot v souboru dat umožňuje uzel *Outliers*. Ten má čtyři metody detekce: *One Class SVM*, *Covariance Estimator*, *Local Outlier Factor*, *Isolation Forest*. Otestovány byly všechny dostupné metody, z nichž nejspokojivější výsledky měla metoda *Local Outlier Factor* s využitím Euklidovské metriky a 10 sousedy. Jak již bylo zmíněno, datová tabulka obsahuje několik sloupců, kde každý sloupec reprezentuje

jednotlivý úsek, v opačném případě měřicí bod. Aby nástroj *Outliers* nepracoval v domněnku, že se jedná o multidimenzionální data, bylo třeba hledat odlehlé hodnoty v každém sloupci zvlášť. Před uzlem *Outliers* byl tedy vložen uzel k výběru sloupce, a tím došlo k selekci konkrétního prvku, který má být analyzován. K tomuto uzlu byl dodatečně ještě připojen uzel *Box Plot*, pro vizualizaci krabicového grafu, podle něhož je možné odlehlé hodnoty manuálně interpretovat. Výstupem uzlu *Outliers* byly tři nové datové tabulky. Tabulka *Outliers* obsahovala hodnoty, které byly identifikovány jako odlehlé. Tabulka *Inliers* naopak obsahovala hodnoty, které nebyly označeny za odlehlé. Třetím výstupem byla tabulka obsahující všechny analyzované hodnoty, doplněné novým atributem *Outlier*, který jednoduše textem *Yes* nebo *No* rozlišuje, zdali je příslušná hodnota odlehlá nebo ne. Pomocí takto sestaveného modelu je možné najít na úseku nebo monitorovacím bodu odlehlé hodnoty v počtu vlakových souprav, včetně uvedení časového údaje, kdy byla daná hodnota naměřena. Nevýhodou tohoto řešení však je, že je potřeba tento proces opakovat pro každý úsek zvlášť a nelze jej spustit pro všechny zároveň.

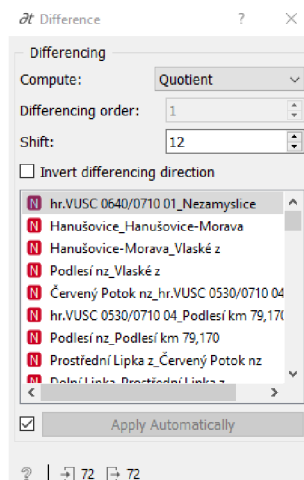


Obr. 17 Model pro detekci odlehlých hodnot v programu Orange

Cílem identifikace odlehlých hodnot bylo odhalení specifických tratí a výkyvů v provozu na železničních úsecích či v monitorovacích bodech. Pro tyto účely bylo dostatečné sledovat ukazatel počtu osobních nebo nákladních vlakových souprav. Z takto identifikovaných úseků či monitorovacích bodů byly některé reprezentativní vybrány pro následnou podrobnější časovou analýzu. Pro potřeby některých analýz, jako třeba shlukování podle podobnosti časových řad, bylo vhodné zajistit aby data neobsahovala prvky s nulovými hodnotami. Díky provedené identifikaci nulových hodnot bylo možné jednoduše tyto problémové prvky odstranit. Pro účely hledání nulových hodnot v datech provozu na železnici v Olomouckém a Moravskoslezském kraji se nejlépe osvědčilo řešení v MS Excel, které bylo taktéž využito. Jedná se o nejrychlejší a zároveň nejjednodušší způsob odhalení nulových hodnot, které signalizují výkyvy v provozu, v datech obsahujících 282 monitorovacích bodů SR70 nebo 421 železničních úseků. Metoda sestavením workflow v programu Orange, není vhodná pro hledání v takto objemných datech, ale lze ji využít při analýze menšího počtu prvků. Naopak využití skriptu v jazyce R vybízí k využití na objemná data s velkým počtem prvků, nicméně je zde zmíněný problém s chybějícím časovým údajem, který je v případě této analýzy podstatný.

6.2 Meziroční tempo růstu

Jelikož jsou data poskytnuta za období let 2016–2021, bylo možné sledovat, jak se situace na železnici vyvíjela v průběhu těchto let. Tato analýza byla opět zaměřena na měsíční hodnoty počtu vlakových souprav na jednotlivých úsecích nebo v monitorovacích bodech. Porovnávány byly hodnoty stejných měsíců dvou po sobě jdoucích roků, a tedy počítáno meziroční tempo růstu. Pro automatizovaný výpočet meziročního tempa růstu byl využit uzel *Difference* v programu Orange. Tento uzel umožňuje kromě počítání rozdílů dvou sousedních hodnot také výpočet poměru dvou hodnot v různém čase, který je označován jako tempo růstu. Pro výpočet tempa růstu byl v dialogovém okně uzlu *Difference* nastavena hodnota parametru *Compute* na možnost *Quotient*. Možnost *Percentage change* umožňuje spočítání meziroční procentuální změny. Jelikož bylo počítáno meziroční tempo růstu, byl parametr *Shift* nastaven na hodnotu 12, podle počtu měsíců v roce. Vstupními daty uzlu byla tabulka, obsahující hodnoty jednotlivých úseku či bodů ve sloupcích. Pro sloupec obsahující datum, byl nastaven datový typ *datetime*. Následně stačilo v dialogovém okně uzlu *Difference* vybrat konkrétní prvek, pro který bude vypočítáno meziroční tempo růstu.



Obr. 18 Nastavení parametrů uzlu Difference

Výsledná křivka tempa růstu byla vizualizována pomocí uzlu *Line Chart*. Výhodou tohoto uzlu je možnost zobrazení původní křivky hodnot počtu vlakových souprav současně s křivkou tempa růstu. Pokud meziročně dochází k růstu počtu souprav, nabývá tempo růstu hodnot vyšších než 1. Naopak pokud je počet nižší, dochází k poklesu tempa růstu a hodnota je nižší než 1. V případě že nedošlo k růstu ani poklesu je hodnota tempa rovna 1. Výpočet byl aplikován na železniční úseky a monitorovací body, které byly analyzovány v kapitole 6.4. Výsledky výpočtu pomohly k charakteristice provozu v daných úsecích či stanicích.

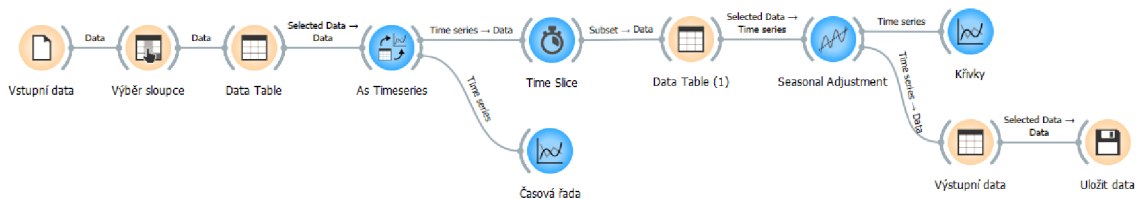


Obr. 19 Model pro výpočet tempa růstu

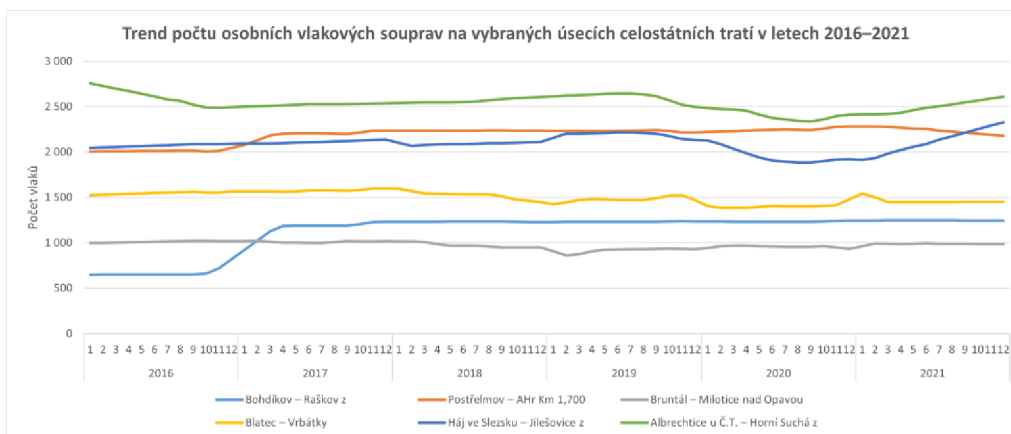
6.3 Dekompozice časových řad

Časové řady lze dále podrobněji analyzovat pomocí jejich rozkladu. Cílem dekompozice bylo získat především trendovou a sezonní složku časových řad. Na základě trendu vývoje v počtu vlakových souprav během sledovaného období lze usoudit, zdali například na celostátních tratích roste nebo klesá počet vlakových souprav a tím tedy zhodnotit jejich provoz. Sezonní složka odhalí vzory v chování, které se každý rok opakují. V případě průmyslových úseků byl předpokládán pokles osobních i nákladních vlaků v letních měsících, kdy dochází k celozávodním dovoleným. Pro sledování sezonní složky je žádoucí, aby časová řada neměla odlehlá pozorování. To znamená, aby nebyla zatížena výkyvy, které by sezonní křivku zkreslily. Proto bylo třeba dát pozor na výluky na trati a propady kvůli pandemii v roce 2020 a 2021. Z toho důvodu mělo smysl provádět rozklad na sezonní složku pouze u časových řad železničních úseků a stanic, které nebyly zasaženy výlukami a současně pouze za období od roku 2016–2019. Díky těmto opatřením bylo zajištěno, že výsledná sezonní křivka nebyla ovlivněná a lze ji považovat za relevantní. Dekompozice časových řad byla prováděna pouze pro úseky a stanice, které byly analyzovány v kapitole 6.4.

Dekompozice časových řad byla realizována v programu Orange. Pro proces dekompozice byl sestaven model obsahující několik uzlů. Na počátku jsou vloženy uzly určené k nahrání dat ve formě tabulky a výběru sloupce, který obsahuje hodnoty úseku, který chceme analyzovat. Při nahrání dat byl pro sloupec obsahující datum nastaven datový typ *datetime*. Pro časovou analýzu je důležitý uzel *As Timeseries*, který data interpretuje jako prvky časové řady. V jeho dialogovém okně byl pro parametr *Sequence* vybrán atribut obsahující datum, podle kterého bude sestrojena časová osa. Navazující uzel *Časová řada* umožňuje grafické zobrazení křivky za celé sledované období. Jak již bylo zmíněno, pro dekompozici časových řad na sezonní složku bylo vhodné, aby křivka nebyla zatížena výkyvy. Po přezkoumání křivky v tomto uzlu bylo snadné určit, zdali výkyvy obsahuje a v jakém období k nim došlo. Aby hodnoty z tohoto období nebyly při analýze uvažovány, byl využit uzel *Time Slice*. Ten umožnil nastavit časový rozsah, ze kterého budou data analyzována. Bylo tedy možné nastavit, aby byla časová řada a její rozklad proveden pouze pro hodnoty od roku 2016–2019, a to bez úpravy vstupních dat. Časový rozsah byl nastavován pouze pro potřeby rozkladu na sezonní složku, a to individuálně pro každý analyzovaný úsek, podle průběhu křivky a výskytu výkyvů. Při vyšetřování trendu bylo ponecháno celé sledované období 2016–2021. Samotnou dekompozici časové řady řeší uzel *Seasonal Adjustment*. V možnostech uzlu byla nastavena sezonní perioda na hodnotu 12, jelikož se jedná o roční data. Dále zde byl nastaven druh rozkladu. Aditivní rozklad se využívá, pokud je variabilita hodnot časové řady přibližně konstantní v čase. Naopak multiplikatívni rozklad se používá u časových řad, kde její variabilita roste v čase, nebo se v čase mění (Hančlová, Tvrdý, 2003). Při rozkladu na sezonní složku byl využíván multiplikatívni rozklad, protože jeho výsledky obsahovaly nižší náhodnou složku nežli výsledky aditivního rozkladu. Na závěr modelu byl vložen uzel na vizualizaci výsledných křivek časových složek společně s uzlem k uložení výsledných hodnot jednotlivých složek ve formě tabulky. Výsledky dekompozice byly ukládány do tabulek a v programu MS Excel byla manuálně přizpůsobována vizuální stránka křivek časových složek, jelikož možnosti úprav v programu Orange jsou omezené. Všechny výstupy vytvořené dekompozicí jsou součástí elektronických příloh (Příloha 9), která je součástí CD přiloženého k této diplomové práci.



Obr. 20 Model využitý pro dekompozici časových řad



Obr. 21 Příklad výsledného grafu po dekompozici časových řad

6.4 Analýza vybraných úseků a stanic

Sledovaných úseků železničních tratí a monitorovacích bodů SR70 je pro efektivní a plnohodnotnou časovou analýzu příliš velké množství. Proto byly vybrány, i na základě předchozí identifikace odlehklých hodnot, reprezentativní úseky a stanice, které byly rozděleny do 5 kategorií. Kategorie byly definovány dle vlastního uvážení tak, aby reprezentovaly síť železniční infrastruktury v Olomouckém a Moravskoslezském kraji. Kategorie *Celostátní úseky* obsahuje úseky tratí, které jsou dle Ministerstva dopravy (MD) kategorizovány jako tratě celostátní. Naopak kategorie *Regionální úseky* obsahuje úseky tratí, které jsou podle MDČR kategorizovány jako tratě regionální. Úseky do těchto dvou kategorií byly vybrány dle vlastního uvážení tak, aby byly rozloženy na celém sledovaném území. Dílčím cílem časové analýzy bylo zaměřit se a zjistit, zdali došlo k změnám v železniční dopravě ve spojení s průmyslovým odvětvím, především v období vlivu pandemie. Proto byla vytvořena kategorie *Průmyslové úseky*, ve které se nachází úseky, které navazují přímo do průmyslového areálu či se nachází v jeho blízkosti. Do kategorie *Specifické úseky* byly vybrány úseky, na kterých se očekával specifický provoz. Poslední kategorií jsou *Stanice na státní hranici*, ve které jsou zařazeny stanice na hraničních přechodech celostátních a regionálních tratích.

Vybrané úseky a stanice byly analyzovány metodou časových řad. Sledovány byly údaje počtu a váhy osobních a nákladních vlakových souprav. V každé kategorii byly nejprve vytvořeny časové řady dohromady ze všech úseků či stanic v rámci jednoho diagramu. Poté byl každý úsek či stanice analyzována zvlášť. Pozornost byla věnována období pandemie covidu-19. U osobní vlakové dopravy existuje domněnka, že jezdil menší počet osobních vlaků. Nicméně osobní doprava je nasmlouvaná k zajištění dopravní obslužnosti krajů, tudíž počet vypravených vlaků by neměl výrazně kolísat. Avšak je možné že kvůli malému počtu cestujících mohou jezdit kratší soupravy. Proto byl ve vybraných případech sledován i ukazatel průměrné délky souprav. Všechny vytvořené časové řady a diagramy jsou součástí elektronických příloh (Příloha 9).

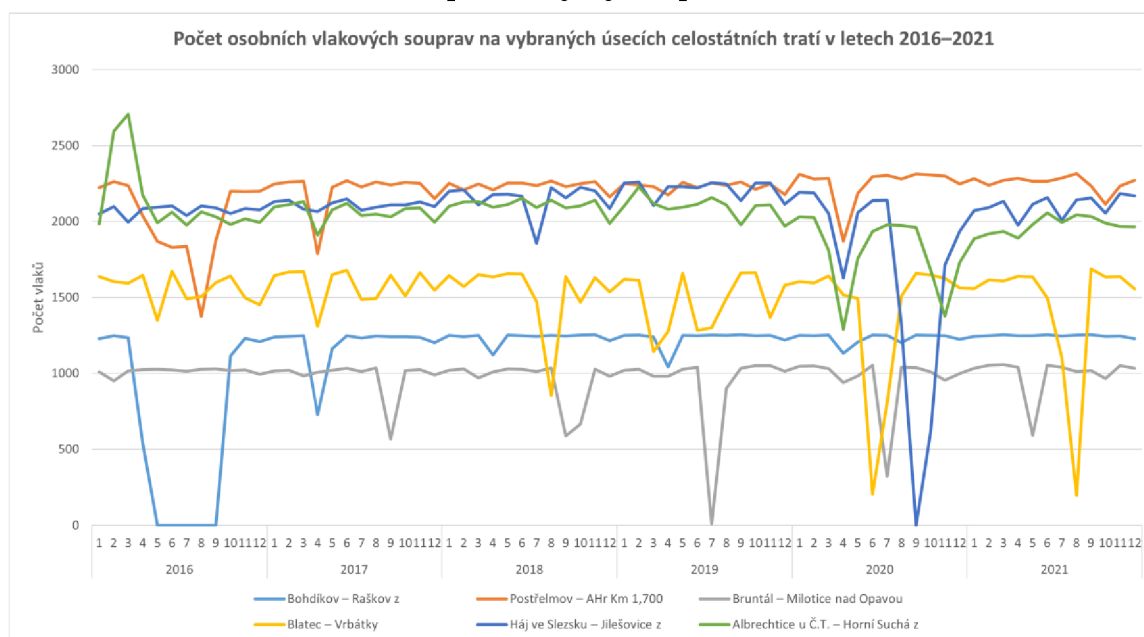
6.4.1 Celostátní úseky

Vybráno bylo celkem 6 úseků tvořících tratě, které jsou kategorizovány jako dráhy celostátní. Jedná se o tyto úseky:

- 25210406 Albrechtice u Č.T. – Horní Suchá z,
- 22011201 Blatec – Vrbátky,
- 13621902 Bohdík – Raškov z,
- 21912001 Bruntál – Milotice nad Opavou,
- 22510602 Háj ve Slezsku – Jilešovice z,
- 19310298 Postřelmov – AHr Km 1,700.

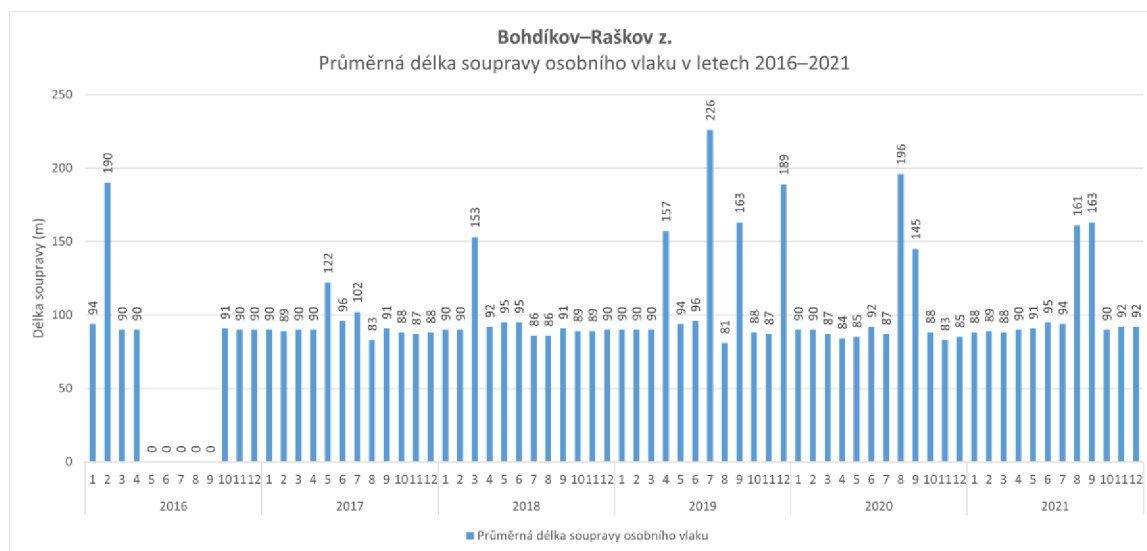
Osobní doprava

Na vybraných celostátních úsecích se počet osobních souprav, s výjimkou zřejmých výkyvů, běžně pohyboval zhruba mezi 1 000 až 2 500. Nejvíce osobních vlaků v jednom měsíci projelo úsekem Albrechtice u Č.T.–Horní Suchá z. v březnu 2016, kdy zde počet narostl až na 2 708 osobních souprav. Na všech úsecích došlo v průběhu let k několika významným propadům v osobní dopravě. V úseku Bohdík–Raškov z. došlo v roce 2016 na čtyři měsíce k úplnému zastavení osobní dopravy. Od května do září zde neprojela žádná souprava. Důvodem byla výluka na trati 292 v úseku mezi Šumperkem a Jeseníkem, kde se tento úsek vyskytuje (IDSOK, 2016a). Výluka na trati 321 v úseku mezi Opavou a Ostravou (Hrušková, 2020) v roce 2020 také výrazně ovlivnila železniční dopravu od srpna do října mezi Hájem ve Slezsku a Jilešovicemi. V září tudy neprojel žádný osobní vlak. Výrazné propady v počtu osobních vlaků postihly v červnu 2020 a srpnu 2021 úsek Blatec–Vrbátky. V obou případech to byl meziroční pokles o zhruba 85 %. V dubnu roku 2020 poklesl počet osobních vlaků na všech vybraných úsecích současně. Největší pokles byl na úsecích Albrechtice u Č.T. – Horní Suchá z, Háj ve Slezsku–Jilešovice z, Postřelmov – AHr Km 1,700, kdy se na každém úseku oproti předchozímu měsíci počet snížil zhruba o 500 souprav. V úseku Albrechtice to znamenalo meziroční pokles o 30 %. Tento hromadný propad zřejmě souvisí s dopady vládních nařízení v souvislosti se prvním výskytem pandemie v ČR.



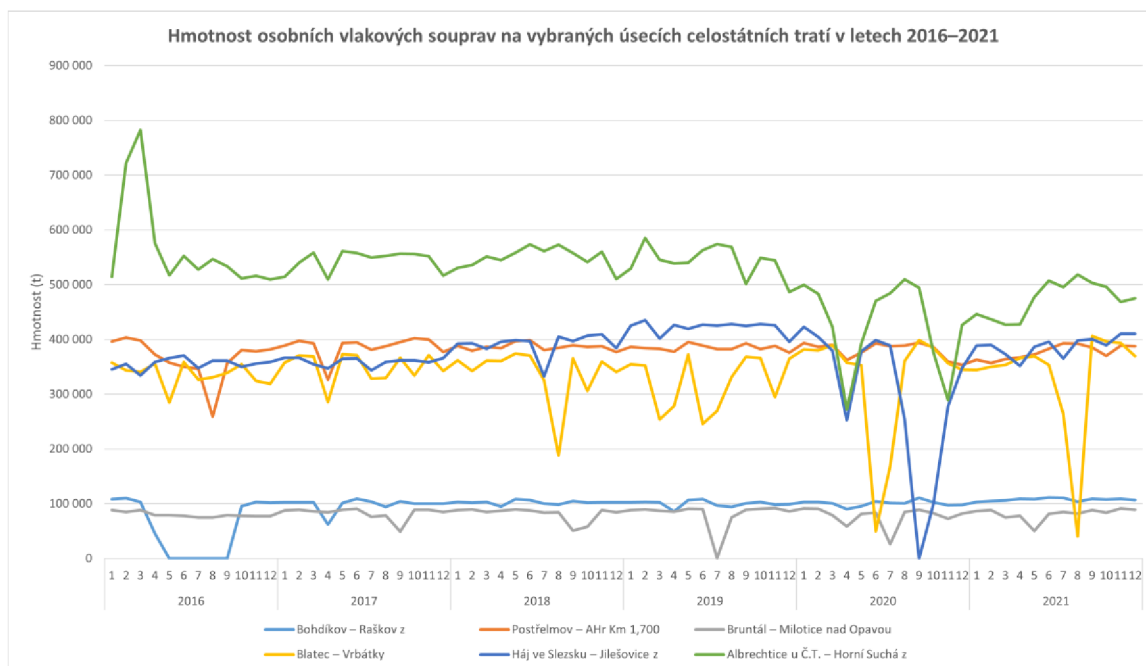
Obr. 22 Počet osobních vlaků na úsecích celostátních tratí

V případě úseku Bohdík–Raškov z., nebyl propad tak výrazný. I přes tento pokles vůči předchozímu měsíci, byl v dubnu zaznamenán meziroční růst počtu souprav. Jak již bylo zmíněno v úvodu kapitoly, více ovlivněna mohla být délka souprav než jejich počet, což se v tomto případě nepotvrdilo. Průměrná délka souprav v dubnu 2020 i následujícím měsíci se nijak nevyvíkala z hodnot předchozích měsíců, nebo stejných měsíců v předešlých letech. Mimo to byl v posledních letech objeven trend, kdy zde v srpnu a září jezdí průměrně skoro dvojnásobně delší soupravy nežli v předchozích měsících. Důvodem může být větší počet cestujících kvůli začínajícímu školnímu roku.



Obr. 23 Průměrná délka osobní soupravy v úseku Bohdík–Raškov

Zajímavé je porovnání úseků Postřelmov–AHR Km 1,700 a Albrechtice u Č.T.–Horní Suchá z. z hlediska počtu a váhy osobních souprav. Ačkoliv na úseku z Postřelmova jezdilo skoro po celých 6 let každý měsíc více osobních vlaků než na úseku v Albrechticích, tak naopak jejich hmotnost byla větší právě v úseku z Albrechtic do Horní Suché. Tento fakt indikuje, že na úseku mezi Albrechticemi a Horní Suchou jezdily delší soupravy, nežli na úseku Postřelmov–AHR Km 1,700.



Obr. 24 Hmotnost osobních vlaků na úsecích celostátních tratí

Od roku 2021 trend osobních vlaků výrazně rostl v úseku Háj ve Slezsku – Jilešovice z, a Albrechtice u Č.T. – Horní Suchá z. Dříve, od konce roku 2019, zde byl však trend klesající. Stejný průběh od konce roku 2019 byl i v úseku Blatec–Vrbátky. Od roku 2021 však počet neroste ani neklesá. Naopak v úseku Postřelmov – AHR Km 1,700 byl trend v roce 2021 klesající. Rozklad na sezónní složku byl proveden pouze pro úsek Bohdíkovo–Raškov a Háj ve Slezsku – Jilešovice. Tyto dva úseky neměly trend počtu osobních vlaků. V úseku Bohdíkovo–Raškov byly nižší hodnoty osobních vlaků v dubnu. U druhého úseku byly nízké v červenci.

Nákladní doprava

Nejvyšší provoz nákladní dopravy byl na trati Albrechtice u Č.T.–Horní Suchá z, která počtem nákladních vlaků výrazně převyšovala ostatní tratě. Nejvíce, zhruba 1 100, jich zde projelo stejně jako u osobní dopravy v březnu 2016. Velký počet nákladních vlaků v tomto úseku nebyl překvapivý, jelikož leží v průmyslové oblasti nedaleko Havířova. U ostatních úseků se většinou počet nákladních vlaků pohybuje do 200 souprav za měsíc. Výluky na trati, které omezily osobní dopravu se týkaly i nákladní dopravy. V létě 2016 na úseku Bohdíkovo–Raškov z. nejezdily ani nákladní vlaky, stejně tak v září 2020 v úseku Háj ve Slezsku–Jilešovice z. Výjimka byla v úseku Bruntál–Milotice nad Opavou, kde v červenci 2019 nejezdily osobní vlaky, ale nákladní doprava nebyla zastavena a projelo zde okolo 30 nákladních souprav.

Trend počtu nákladních vlaků byl v každém úseku rozdílný. Nejvíce proměnlivý byl v úseku Albrechtice u Č.T. – Horní Suchá. Do poloviny roku 2017 byl klesající, následně prudce rostl až do konce roku 2018, od kterého začal klesat až do jara 2020, kdy začal opět růst. Rostoucí trend nákladních vlaků byl také v úseku u Postřelmova. U zbylých úseků byl trend klesající. Rozklad na sezónní složku u nákladních vlaků nebyl realizován, jelikož byla doprava nepravidelná a obsahovala mnoho náhodných výkyvů.

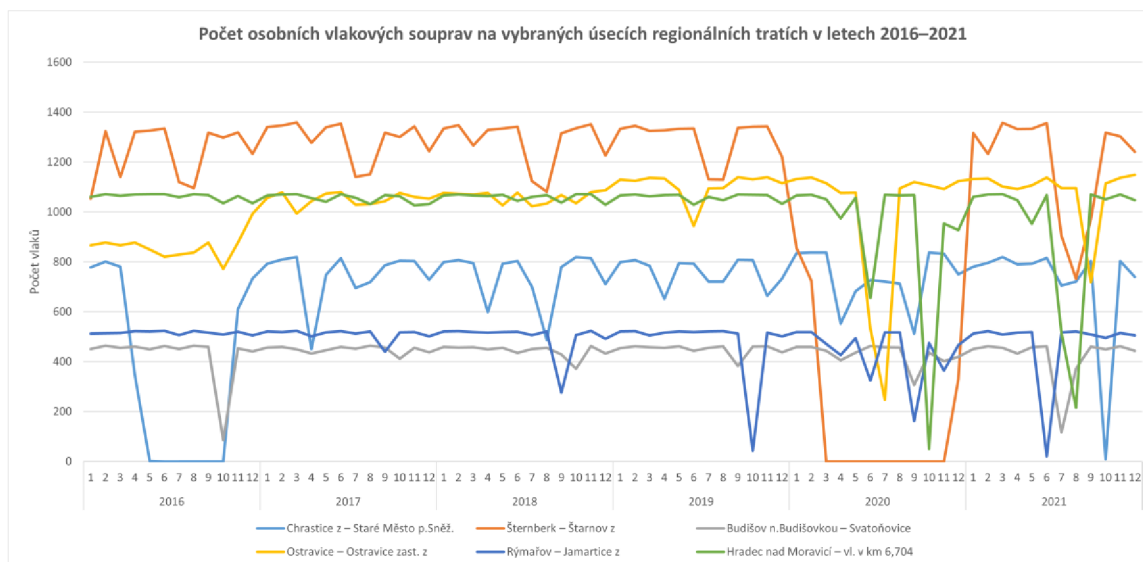
6.4.2 Regionální úseky

Vybráno bylo celkem 6 úseků tvořících tratě, které jsou kategorizovány jako dráhy regionální. Jedná se o tyto úseky:

- 13510204 Chrastice z – Staré Město p.Sněž.,
- 13610498 Šternberk – Štarnov z,
- 19611201 Budišov n.Budišovkou – Svatoňovice,
- 21610203 Ostravice – Ostravice zast. z,
- 22210606 Rýmařov – Jamartice z,
- 22610204 Hradec nad Moravicí – vl. v km 6,704.

Osobní doprava

Na vybraných úsecích regionálních tratí nebyl počet osobních vlaků za jeden měsíc větší než 1 400 souprav. Vyjma zřejmých výkyvů se obecně ze všech úseků počet osobních vlaků v jednom měsíci pohyboval v rozsahu 400 až 1 400 souprav. Nejméně vytěžovaný úsek z této šestice byl Budišov n.Budišovkou–Svatoňovice. Pokud nedošlo k výraznému výkyvu, projelo zde měsíčně zhruba 400 až 500 osobních vlaků. Nejvíce osobních vlaků jezdilo v úseku Šternberk–Štarnov z, který ostatní úseky výrazně převyšoval i celkovou hmotností osobních souprav. Hmotnost zde dosahovala až 140 tisíc tun za měsíc. Na druhém nejvytěžovanějším úseku Ostravice–Ostravice zast. z., se váha pohybovala okolo 60 tisíc tun měsíčně, což je výrazný rozdíl. V úseku Šternberk–Štarnov došlo v roce 2020 k dlouhodobému propadu osobní dopavy, z důvodu výluky na trati 290 v úseku mezi Olomoucí a právě Šternberkem (IDSOK, 2020). Od začátku roku se počet vlaků snižoval až do března, od kterého už byla osobní doprava úplně zastavena. Počet vlaků začal růst až v prosinci na konci roku. Následující měsíc již byl počet zpět na běžných hodnotách. Úplná výluka a náhradní autobusová doprava byla zavedena na 5 měsíců v roce 2016 i v úseku Chrastice z – Staré Město p.Sněž. (IDSOK, 2016b). Na tomto úseku se počet průjezdů osobních vlaků propadl na nulu i v říjnu 2021. V případě úseku Ostravice–Ostravice zast. z. bylo na přelomu let 2016 a 2017 vysoké tempo růstu. Od roku 2017 začalo jezdit více než tisíc souprav měsíčně. Předchozí rok se počet pohyboval okolo 870 souprav měsíčně. Na první pohled bylo zřejmé, že poslední dva roky byly úseky regionálních tratí zatíženy velkým množstvím výkyvů. Na všech úsecích došlo k častým propadům počtu osobních vlaků. Kromě plánovaných výluk byly zřejmě opět na vině dopady protiepidemických opatření na mobilitu obyvatel a tím i poptávku po osobní dopravě. K propadům však docházelo i v letních měsících, ve kterých restrikce nebyly tak přísné. Je možné, že na těchto úsecích došlo k přechodu na letní jízdní řád.

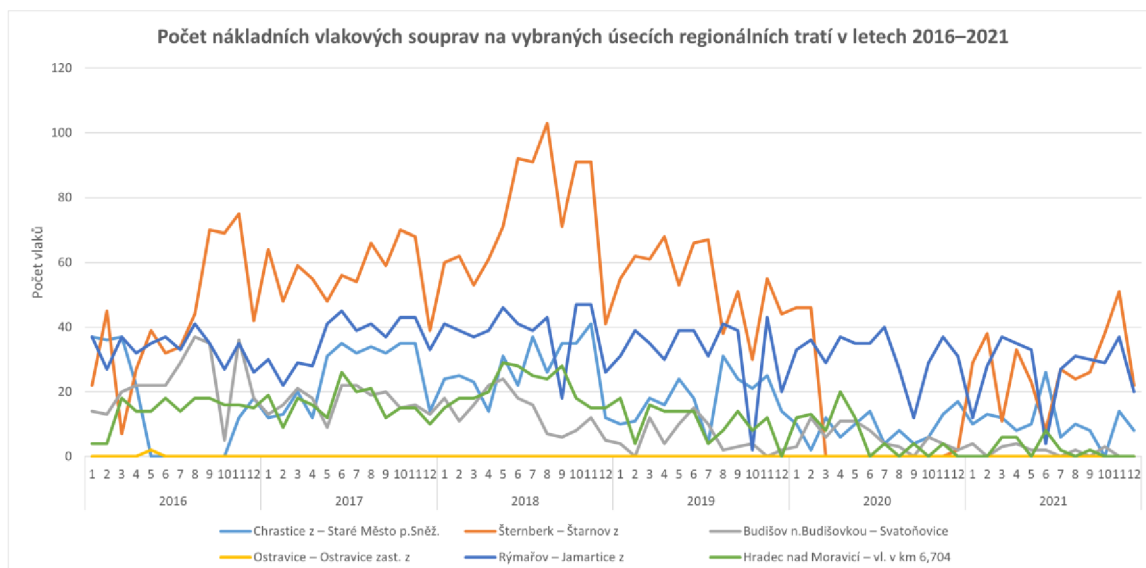


Obr. 25 Počet osobních vlaků na úsecích regionálních tratí

Trend počtu osobních vlaků se měnil především v posledních dvou letech. Rostoucí trend byl mezi roky 2016 a 2017 v úseku Chrastice – Staré Město p.Sněž, Ostravice – Ostravice zast. a Šternberk – Štarnov. Zatímco od roku 2017 do roku 2020 byl trend na všech úsecích víceméně neměnný, v roce 2020 začal klesat. Klesající trend v osobní dopravě přetrvával v roce 2021 v úsecích Ostravice – Ostravice zast. a Chrastice – Staré Město p.Sněž. U zbylých úseků byl v roce 2021 zaznamenán rostoucí trend počtu osobních vlaků. Úsek Šternberk – Štarnov byl však v roce 2020 ovlivněn dlouhodobou výlukou. V úseku Budišov n.Budišovkou – Svatoňovice pravidelně klesal počet osobních souprav v říjnu a prosinci. Naopak nárůsty počtu vykazuje srpen a září. Úsek v Hradci nad Moravicí v průběhu roku nevykazoval výrazné odchylky. Nižší hodnoty sezonní složky byly v červnu a prosinci.

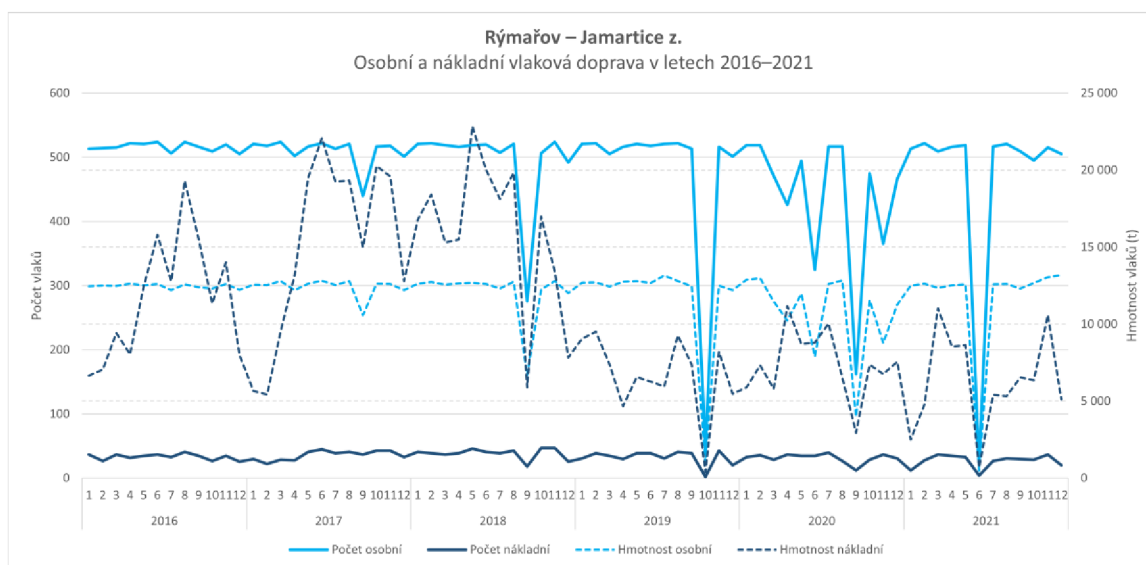
Nákladní doprava

Nákladní dopravou nebyly tyto úseky výrazně zatíženy. S výjimkou úseku Šternberk–Štarnov z. neprojelo za měsíc na žádném z úseků více než 50 nákladních vlaků. V úseku Ostravice–Ostravice zast. z., nákladní vlaky nejezdily vůbec. Pouze v květnu 2016 zde projely dvě soupravy. Nejvíce nákladních vlaků tedy jezdilo v úseku Šternberk–Štarnov z. Měsíční maximum nastalo v srpnu 2018, kdy zde projelo 103 souprav. Celkovou hmotností byl však úsek vytíženější v září 2016. I přesto, že zde projelo o 33 souprav méně, byla jejich celková hmotnost vyšší zhruba o 3 tisíce tun. Od maxima v srpnu 2018, byl trend v počtu vlaků klesající. Obecně však na všech úsecích počet nákladních vlaků v průběhu let značně kolísal. V posledních dvou letech se jejich počet snížil a nedosahoval hodnot jako v měsících předchozích let.



Obr. 26 Počet nákladních vlaků na úsecích regionálních tratí

Vývoj nákladní dopravy byl zajímavý v úseku Rýmařov–Jamartice z. Počet vlaků byl v průběhu let více méně stejný, a to okolo 30 souprav měsíčně. Nicméně jejich celková hmotnost se výrazně změnila. V letech 2016–2018 byla váha nákladních vlaků vyšší než v letech 2019–2021. Nejtěžší soupravy jezdily v letech 2017 a 2018. Toto zatížení vlaků je možné vysvětlit těžbou dřeva v okolí Rýmařova kvůli kůrovcové kalamitě, která v těchto letech sužovala podhůří Jeseníků a Bruntálsko.



Obr. 27 Osobní a nákladní doprava v úseku Rýmařov–Jamartice

Počet nákladních vlaků na úsecích regionálních tratí má klesající trend. Mírně rostoucí trend byl mezi rokem 2017 a 2018 právě na zmíněném úseku mezi Rýmařovem a Jamarticemi. Do konce roku trend rostl i v úseku Šternberk – Štarnov. Poté je trend výrazně ovlivněn výlukou v roce 2020. Sezónní složka u nákladní dopravy byla zjišťována pouze v úseku Rýmařov – Jamartice. Vyšší hodnoty vůči trendu byly od května do srpna, a poté v říjnu a listopadu. Naopak nižší od prosince až do dubna.

6.4.3 Průmyslové úseky

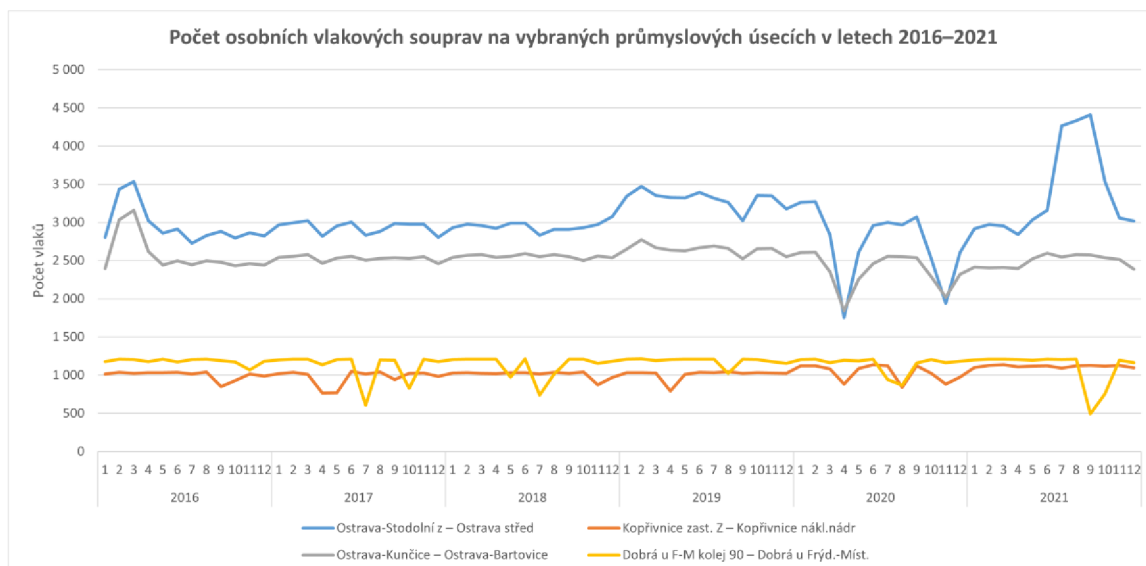
Vybrány byly celkem 4 úseky, které se nachází v blízkosti významných průmyslových objektů či oblastí. Jedná se o tyto úseky:

- 25310403 Dobrá u F-M kolej 90 – Dobrá u Frýdku-Místku
- 21720402 Kopřivnice zast. z – Kopřivnice nákl. nádr
- 21320404 Ostrava-Stodolní z – Ostrava střed
- 25210801 Ostrava-Kunčice – Ostrava-Bartovice

Jak již bylo zmíněno v úvodu kapitoly, dílčím cílem časové analýzy bylo zaměření se na úseky v blízkosti průmyslových areálů. Snahou bylo zjistit k jaké situaci došlo na železnici v kombinaci s průmyslovou výrobou, u které se v době pandemie předpokládá její omezení. Mimo dopadů pandemie lze z časových řad vypožorovat celozávodní dovolené, kdy dochází především v letních měsících k propadu počtu vlakových souprav. Železniční doprava je často využívána k přepravě výrobků automobilek, ale může být využita i k přepravě komponent (ČD Cargo, 2022). Úsek Dobrá u F-M kolej 90–Dobrá u Frýdku-Místku se nachází v blízkosti automobilového závodu Hyundai Nošovice, na který přímo navazuje vlečka z tohoto podniku. Dalším automobilovým závodem v regionu je kopřivnická Tatra, proto byl analyzován úsek Kopřivnice zast. z–Kopřivnice nákl. nádr. Úsek Ostrava-Stodolní z–Ostrava střed se nachází v centru Ostravy a navazují na něj lokální tratě vedoucí do těžebních oblastí. Poslední úsek Ostrava-Kunčice–Ostrava-Bartovice vede okolo průmyslového areálu Liberty Ostrava, který je zaměřen na hutnictví a strojírenství.

Osobní doprava

Analyzování osobní dopravy v těchto úsecích nebylo stěžejní. Avšak na úsecích Ostrava-Stodolní z–Ostrava střed a Ostrava-Kunčice–Ostrava-Bartovice nejezdil zanedbatelný počet osobních vlaků. Jelikož je Ostrava významný železniční uzel, tak zde byl největší provoz osobních vlaků ze všech úseků napříč kategoriemi. Každý měsíc zde projelo více než 2 500 osobních vlaků. V únoru a březnu 2016 došlo k nárůstu na obou úsecích. K výraznému nárůstu došlo v roce 2021 na úseku Ostrava-Stodolní z–Ostrava střed. Od července do září zde měsíčně jezdilo více než 4 tisíce osobních vlaků. Maximum nastalo v září s celkovým počtem 4 408 osobních souprav. K poklesům na obou úsecích došlo v roce 2020, kdy se v březnu počet vlaků začal snižovat až do května. Jedná se o období začátku pandemie koronaviru v ČR. V úseku Ostrava-Stodolní z–Ostrava střed v únoru projelo 3 270 vlaků, zatímco v dubnu pouze 1 753, což je propad o 46 %. Druhý významný pokles nastal v říjnu a listopadu, kdy byla znovu zavedena striktní protiepidemická opatření. Ve stejný čas na jaře i na podzim mírně poklesla osobní doprava i v úseku Kopřivnice zast. Z – Kopřivnice nákl. nádr. V případě úseku Dobrá u F-M kolej 90–Dobrá u Frýdku-Místku k žádným změnám u osobní dopravy v období protiepidemických nařízení nedošlo.

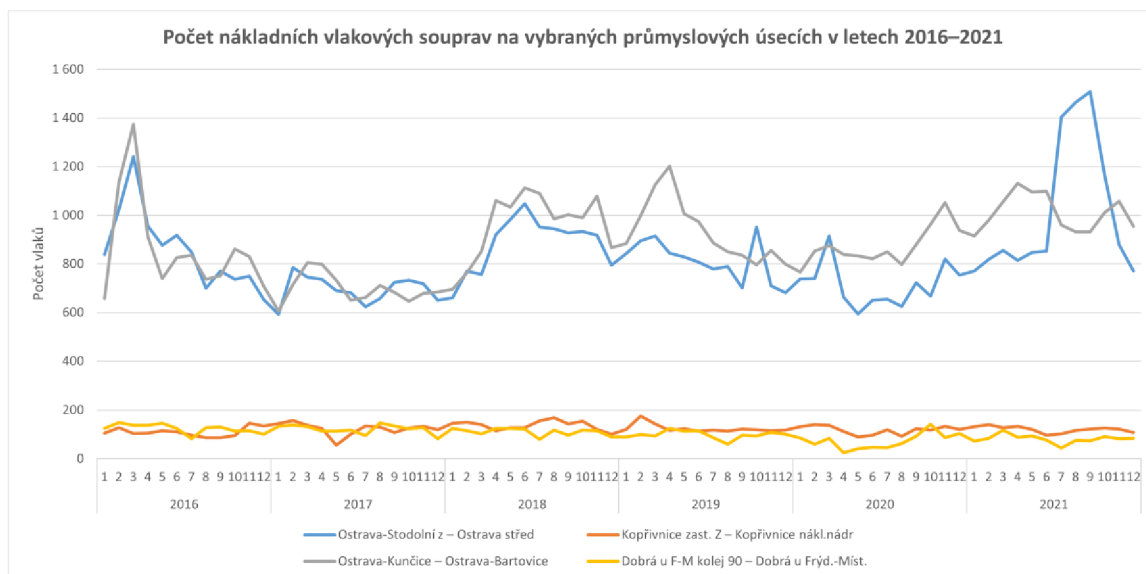


Obr. 28 Počet osobních vlaků na průmyslových úsecích

Poté, co na přelomu let 2019 a 2020 trend počtu osobních vlaků v úsecích Ostrava-Stodolní z–Ostrava střed a Ostrava-Kunčice–Ostrava-Bartovice klesal, nabral v druhé půlce roku 2020 rostoucí trend. Do konce roku 2019 byl na obou úsecích také rostoucí. V případě úseků Dobrá u F-M kolej 90 – Dobrá u Frýdku-Místku a Kopřivnice zast. z – Kopřivnice nákl. nádr. se počet vlaků příliš neměnil a oba úseky měly stacionární trend. Díky tomu byly vhodné pro dekompozici na sezónní složku. V úseku Dobrá u F-M kolej 90 – Dobrá u Frýdku-Místku jezdil menší počet vlaků pravidelně v červenci. U druhého úseku to bylo v dubnu.

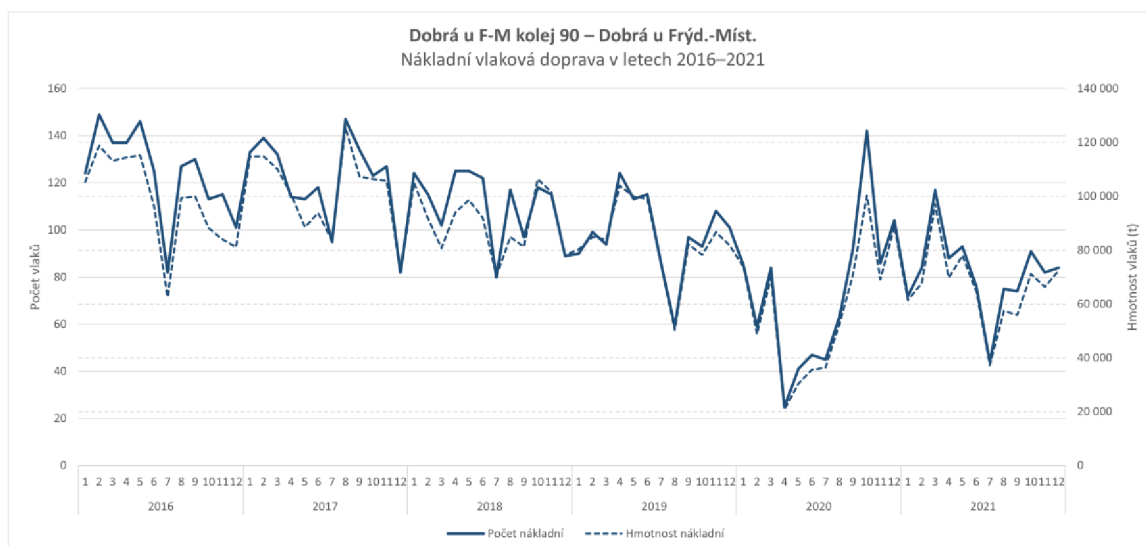
Nákladní doprava

Ostravské úseky byly také významně zatíženy nákladní dopravou, když se počet nákladních vlaků, které na nich projely, pohyboval zhruba od 600 až do 1 500 souprav za měsíc. Stejně jako u osobní dopravy, na obou úsecích okolo března 2016 počet nákladních vlaků strmě narostl. Od dubna se však počet začal snižovat. V úseku Ostrava-Kunčice–Ostrava-Bartovice došlo na začátku roku 2018 opět k narůstání počtu vlaků. Od té doby zde jezdil pravidelně větší počet nákladních vlaků než v předchozím roce. K prudkému nárůstu nákladních vlaků došlo v létě 2021 v úseku Ostrava-Stodolní z–Ostrava střed. Od května 2020, kdy se počet propadl až na 594 souprav za měsíc, což byl na tomto úseku nejmenší počet za celé sledované období, začal počet opět pomalu narůstat až do června 2021. Následně v červenci počet skokově narostl z červnových 852 na 1 403 nákladních souprav. Meziroční tempo růstu bylo na hodnotě 2,15. Vrchol nastal v září, kdy úsekem projelo 1 509 nákladních vlaků, což byl meziroční nárůst o 109 %. Na základě zvyšujícího se počtu nákladních vlaků je možné předpokládat, že v oblasti došlo ke zvýšení průmyslové výroby.



Obr. 29 Počet nákladních vlaků na průmyslových úsecích

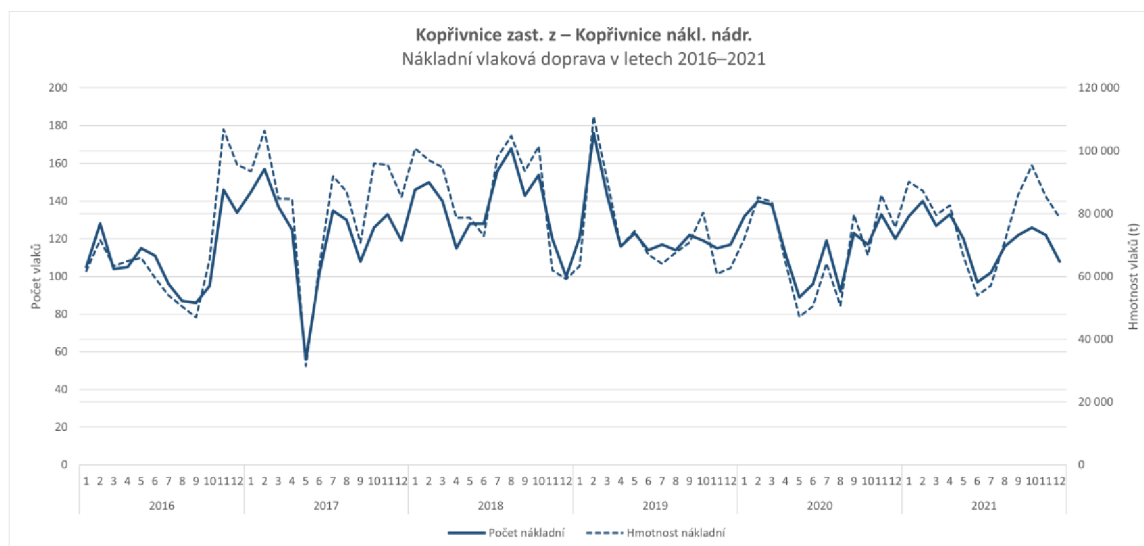
V úseku Dobrá u F-M kolej 90 – Dobrá u Frýdku-Místku sousedícího s automobilkou Hyundai byl počet nákladních vlaků v průběhu let kolísavý. V dlouhodobém horizontu však klesal. Největší počet byl zaznamenán v únoru 2016, kdy v úseku projelo 149 nákladních vlaků. Největší propad, na celkově 25 souprav, však nastal v dubnu 2020, kdy byla kvůli pandemii koronaviru na několik týdnů přerušena výroba (iDnes.cz, 2020)). Od srpna se počet nákladních vlaků prudce zvyšoval, až na říjnové maximum 142 souprav. Po omezení se dalo očekávat zvýšení produkce, aby byly dohnány ztráty. Křivka hmotnosti vlaků kopírovala křivku jejich počtu.



Obr. 30 Nákladní doprava na úseku v blízkosti automobilky Hyundai

Na úseku Kopřivnice zast. z – Kopřivnice nákl. nádr. nebyl na první pohled viditelný žádný vzor v chování z hlediska počtu nákladních vlaků. Nejméně jich projelo v květnu 2017, a to pouze 56, což proti předchozímu měsíci byl úbytek o 69 souprav. K dalšímu propadu došlo na konci roku 2018. V únoru 2019 však úsekem projelo 176 nákladních vlaků, což bylo nejvíce za celých 6 let. V letech ovlivněných pandemií covid-19 došlo

k poklesům nákladní dopravy mezi dubnem a květnem 2020 a poté květnem a červnem 2021. První pokles lze spojovat s následky pandemie a vládních opatření. Druhý pokles nastal v období, ve kterém neplatila žádná přísná opatření, která by měla ovlivnit výrobu.



Obr. 31 Nákladní doprava v blízkosti automobilky Tatra

Počet nákladních vlaků měl od polovinu roku 2020 až do konce roku 2021 rostoucí trend v úsecích Ostrava-Kunčice–Ostrava-Bartovice a Ostrava-Stodolní z–Ostrava střed. Předtím počet rostl také v roce 2018. U zbylých úseků trend od roku 2018 mírně klesal. Dekompozice časové řady Dobrá u F-M kolej 90–Dobrá u Frýdku-Místku odhalila každoroční červencové propady počtu nákladních vlaků. V červenci jezdí současně i menší počet osobních vlaků. V tomto termínu se tedy dá předpokládat celozávodní dovolená v automobilce Hyundai. Následující měsíc již byly hodnoty vyšší nežli trend. V případě úseku Kopřivnice zast. z – Kopřivnice nákl. nádr. byly počty pravidelně nižší v dubnu, květnu a září. Naopak více vlaků jezdilo v únoru. U obou úseku jezdilo méně vlaků také v prosinci.

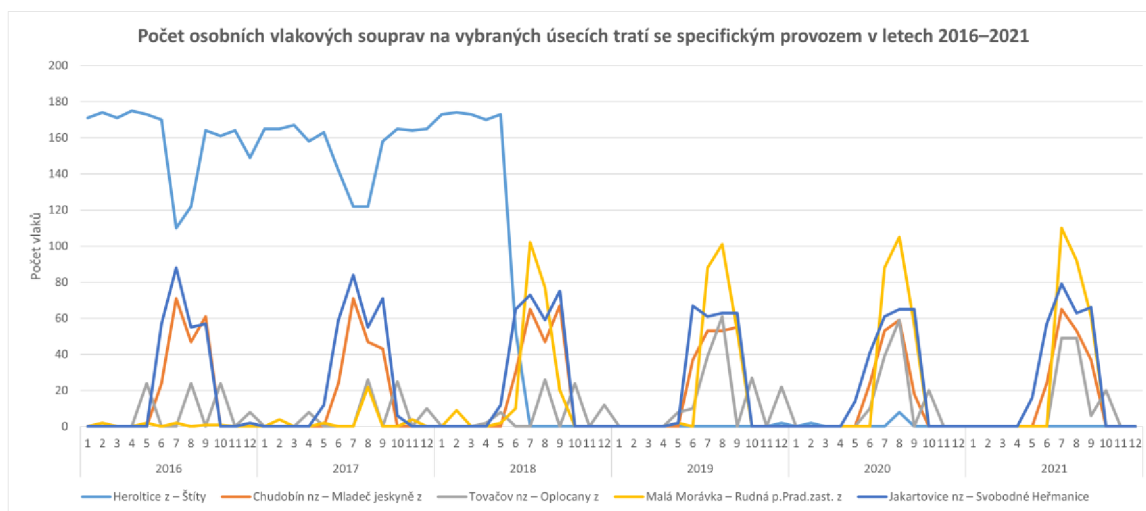
6.4.4 Specifické úseky

Vybráno bylo celkem 5 úseků, u kterých byl očekáván specifický provoz. Jedná se o tyto úseky:

- 13411001 Heroltice z – Štítý
- 19530401 Chudobín nz – Mladeč jeskyně z
- 21110403 Tovačov nz – Oplocany z
- 22310804 Malá Morávka – Rudná p.Prad.zast. z
- 22711001 Jakartovice nz – Svobodné Heřmanice

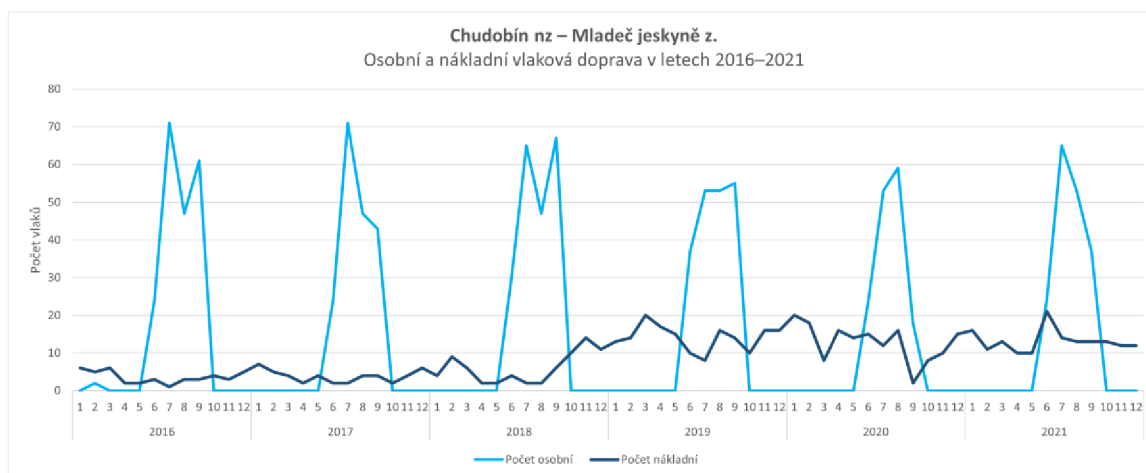
K vybrání specifických úseků byly využity výsledky při analýze výkyvů v provozu a oficiální dokument Správy železnic Seznam málo využívaných tratí (SŽ, 2022). Seznam obsahuje tratě, které jsou klasifikovány jako málo využívané. Jsou zde vedeny tratě společně s časovým rozmezím, kdy na nich neprobíhá provoz. Z tohoto seznamu lze tedy identifikovat tratě se sezonním provozem.

Provoz na vybraných úsecích je opravdu minimální. Na většině úseků jezdí osobní vlaky pouze v letních měsících. Jejich počet je v desítkách za měsíc. Výjimkou byl úsek Heroltice z–Štítý, kde jezdilo více než 100 osobních souprav měsíčně. Nicméně od července 2018 byla v tomto úseku osobní doprava oficiálně ukončena. Od té doby byl úsek obsluhován pouze nákladními vlaky, kdy jejich počet nepřesáhl 20 za měsíc. Nákladní doprava je celkově na těchto úsecích zanedbatelná. V úseku Jakartovice nz–Svobodné Heřmanice jezdily osobní vlaky od května do konce září. Podle jejich počtu, lze odhadovat, že jde pouze o víkendový provoz.



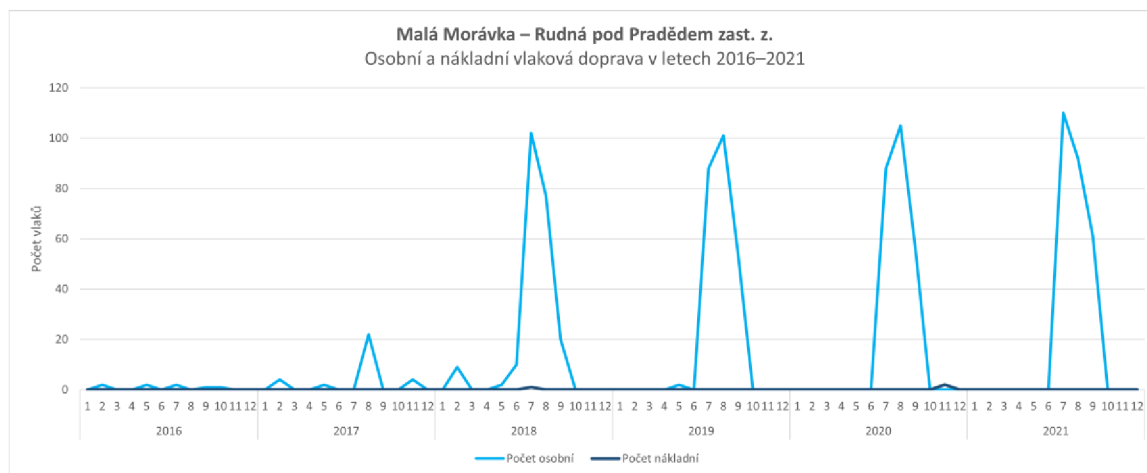
Obr. 32 Počet osobních vlaků na úsecích tratí se specifickým provozem

Úsek Chudobín nz–Mladeč jeskyně z. je součástí trati 308 z Litvle právě do obce Mladeč, která je známá díky místním jeskyním. Jedná se tedy o turisticky zajímavé místo především v letních měsících. Na trati proto fungoval sezonní provoz pouze od května do října. Podle nízkého počtu souprav i zde probíhal pouze víkendový provoz. V září 2020 byla vypočítána nízká hodnota meziročního tempa růstu. V úseku projelo pouze 18 souprav, zatímco předchozí rok jich projelo 55. Na dané trati zřejmě došlo v září 2020 k nějakým omezením provozu.



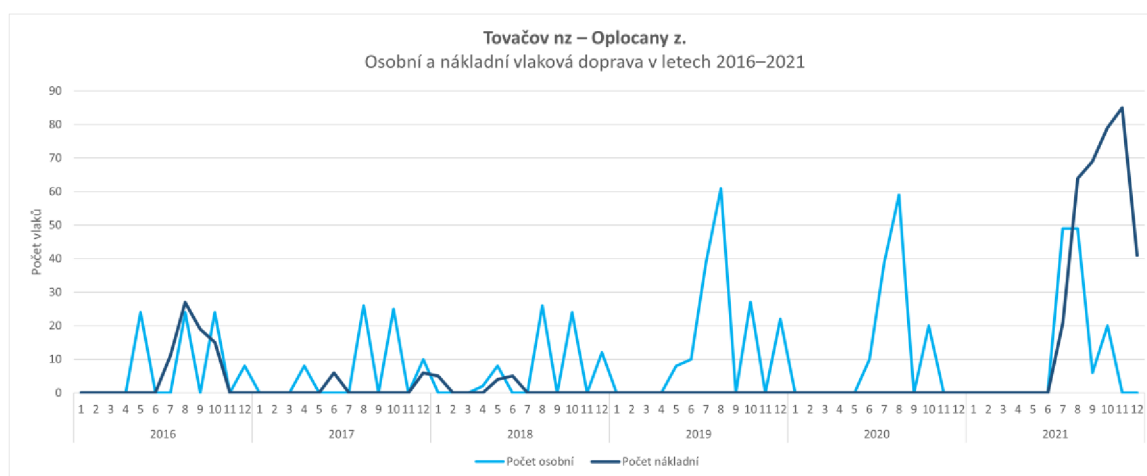
Obr. 33 Osobní a nákladní doprava v úseku Chudobín – Mladeč jeskyně

Průběh křivky počtu osobních vlaků v úseku Malá Morávka – Rudná p.Prad.zast. z naznačuje, že provoz byl v roce 2018 obnoven na trati 312 z Bruntálu do Malé Morávky. V přechozích letech zde pravidelná osobní doprava neprobíhala a nákladní vlaky zde nejezdily téměř vůbec. Od roku 2018 byl spuštěn sezónní provoz od června do září. Nízký měsíční počet souprav naznačuje, že vlaky jezdily pouze o víkend.



Obr. 34 Osobní a nákladní doprava v úseku Malá Morávka – Rudná pod Pradědem

Specifická situace byla v úseku Tovačov nz–Oplocany z. na trati 304 Kojetín–Tovačov. Od roku 2016 do roku 2019 zde jezdily osobní vlaky pouze od dubna do prosince. Nicméně nejezdily každý měsíc. Například v září osobní vlaky nejezdily až do roku 2021, a v červnu a červenci začaly jezdit v roce 2019. Od roku 2019 jezdilo nejvíce vlaků v červenci a srpnu, čímž se výrazně zvýšil provoz na trati. Největší počet byl v srpnu 2019, kdy zde projelo 61 souprav. Nízký počet vlaků znamená, že nejezdily pravidelně každý den. Nákladní doprava v tomto úseku probíhala sporadicky. Od července 2018 až do června 2021 tudy neprojela žádná nákladní souprava. Změna přišla v červenci 2021, kdy nákladní vlaky začaly jezdit a jejich počet se v listopadu vyšplhal až na 85 souprav.



Obr. 35 Osobní a nákladní doprava v úseku Tovačov – Oplocany

Počet osobních vlaků na úsecích se specifickým provozem měl rostoucí trend. Výjimkou byl úsek Heroltice – Štítý, kde byl provoz osobní dopravy od roku 2019 zastaven. U

nákladní dopravy byly trend proměnlivý. V posledních letech počet nákladních vlaků rostl v úseku Chudobín nz – Mladeč jeskyně z. V úseku Heroltice – Štítý byla sice ukončena osobní doprava, ale nákladní stále v malé míře probíhala. Nicméně počet nákladních vlaků od roku 2018 klesal. V případě těchto úseků, kde probíhá sezónní provoz, nemělo význam provádět rozklad časové řady. Sezónní složka je zřejmá již na základě křivky časové řady.

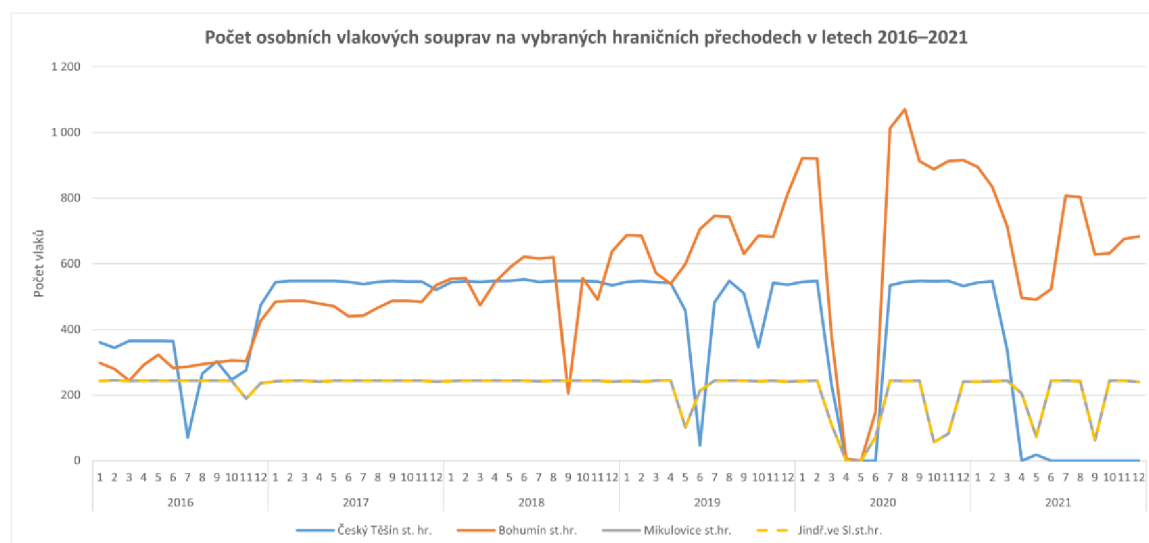
6.4.5 Stanice na státní hranici

Vybrány byly celkem 4 měřící body SR70, které se nachází na státních hranicích České republiky. Jedná se o ty body SR70:

- 54-300145-00 Český Těšín st. hr.
- 54-300244-00 Bohumín st. hr.
- 54-300426-00 Mikulovice st. hr.
- 54-301549-00 Jindřichov ve Slezsku st. hr.

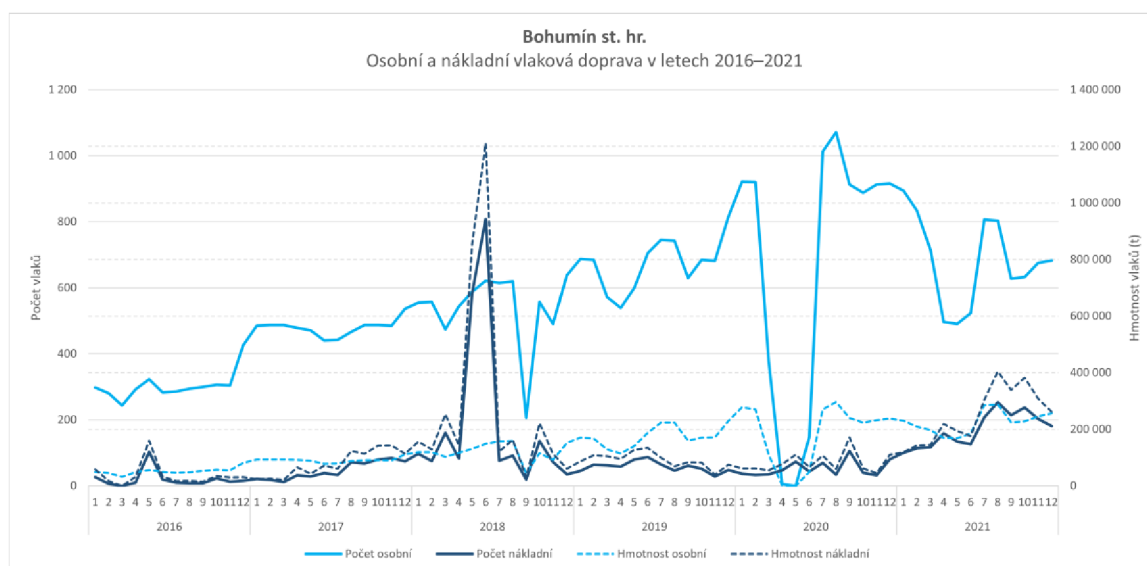
Na území Olomouckého a Moravskoslezského kraje se nachází čtyři hraniční přechody, které nejsou na železničním koridoru. Ve všech případech vede železniční trať do Polska. Osobní doprava byla největší ve stanici na státní hranici v Bohumíně. Naopak nejmenší provoz byl ve stanicích Mikulovice st. hr. a Jindřichov ve Slezsku st. hr. Tyto stanice jsou specifické tím, že leží na stejné trati. Jedná se o trať 292 Šumperk–Krnov, kde v úseku právě mezi Mikulovicemi a Jindřichovem ve Slezsku probíhá peážní doprava. To znamená, že doprava probíhá přes území jiného státu. V tomto případě přes území Polska a město Hlucholazy. Že se jedná o stanice na stejné trati potvrzují totožné průběhy počtů osobních vlaků. Pravidelně zde jezdilo 244 osobních vlaků za měsíc.

K nejvýraznější změně v osobní dopravě na státních hranicích došlo v roce 2020, kdy bylo 16. března vyhlášeno uzavření státních hranic České republiky kvůli šíření koronaviru. Počet osobních vlaků se na všech místech okamžitě propadl. V dubnu a květnu těmito hraničními přechody neprojel žádný osobní vlak. Od června byl již provoz osobních vlaků obnoven na všech stanicích, kromě Českého Těšína, kde první osobní vlaky projely až v červenci. Uzavření hranic však nemělo vliv na nákladní dopravu, která nebyla omezena. Méně vlaků jezdilo i na jaře 2021, kdy byl tvrdý lockdown.



Obr. 36 Počet osobních vlaků ve stanicích na hraničních přechodech

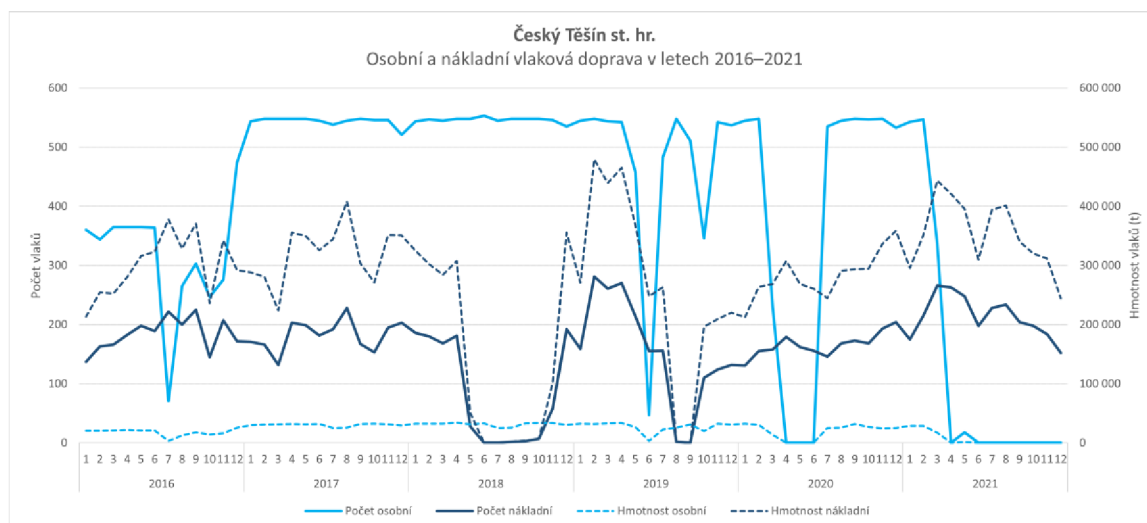
Nejvíce osobních vlaků projíždělo přes státní hranici v Bohumíně. Až do srpna 2020 tu byl rostoucí trend počtu osobních vlaků. Jak bylo uvedeno, na jaře 2020 došlo k velkému výpadku v osobní dopravě. Z únorových 920 osobních souprav se počet propadl až na nulu. Na trati byla zachována pouze nákladní doprava. Po červnovém obnovení osobní dopravy začalo jezdit ještě více vlaků nežli před lockdownem. Vůči únoru, což byl poslední měsíc neovlivněný protiepidemickými opatřeními, projelo v červenci o 93 souprav více. Nejvíce 1 071 osobních souprav projelo v srpnu. Nárůstu zřejmě pomohlo, že se jednalo o prázdninové měsíce a po období, kdy byla omezena mobilita, byla větší poptávka po cestování do ciziny. Od září se počet začal opět propadat až do května 2021, kdy počet klesl na 491 souprav za měsíc. Březen a duben 2021 byly ovlivněny tvrdým lockdownem, kdy nebylo možné cestovat mezi okresy. Proto byly omezeny i cesty mimo ČR. Nákladních vlaků zde nejezdilo více než 200 za měsíc. Rapidní nárůst nákladních vlaků nastal v květnu a červnu 2018, kdy se právě v červnu počet vyšplhal na 808 souprav. To je 873% nárůst vůči dubnu, kdy byl jejich počet pouhých 83. Tyto dva měsíce přes státní hranici přejelo více nákladních vlaků než osobních. Od listopadu 2020 provoz nákladní dopravy rostl.



Obr. 37 Osobní a nákladní doprava ve stanici Bohumín st. hr.

V Českém Těšíně podle počtu souprav převažovala osobní doprava nad nákladní. Avšak celkovou hmotností souprav byla stanice výrazně vytíženější nákladní dopravou. V měsících, kdy nedošlo k výkyvům, stanicí projíždělo okolo 550 osobních souprav měsíčně. Tolik vlaků měsíčně zde začalo jezdit od začátku roku 2017. U osobní dopravy došlo v průběhu let k několika propadům. Nejprve v červenci 2016, červnu a říjnu 2019 a poté v dubnu, květnu a červnu 2020, kdy se osobní doprava úplně zastavila, kvůli zmíněnému uzavření státních hranic. K přerušení osobní dopravy došlo ještě v dubnu 2021. Jak bylo zjištěno, v úseku Český Těšín–Cieszyn byla osobní doprava od 20. března 2021 až do odvolání přerušena na základě rozhodnutí polských úřadů a Moravskoslezského kraje (České dráhy, 2021). V porovnání s počtem osobních vlaků, byl ve stanici i velký počet nákladních vlaků. Maximální počet byl 289 v únoru 2019. U nákladní dopravy došlo ke dvěma výrazným výpadkům. Nejprve od května do listopadu 2018, kdy zde v období od června do října projelo celkem 12 nákladních vlaků. Následující rok nejezdily nákladní vlaky v srpnu a září. Protiepidemická opatření,

jako uzavření státních hranic, nijak výrazně nákladní dopravu v tomto místě neovlivnila. Naopak počet nákladních vlaků měl od roku 2020 stoupavou tendenci.



Obr. 38 Osobní a nákladní doprava ve stanici Český Těšín st. hr.

Nejvíce počet vlaků v osobní dopravě rostl ve stanici Bohumín st. hr. Počet osobních vlaků měl rostoucí trend od začátku roku 2016 až do propadu v dubnu 2020, kdy došlo k uzavření státních hranic. Poté počet opět rostl až do začátku roku 2021, od kdy začal trend prudce klesat. U stanice Český Těšín st. hr. byl rostoucí trend do poloviny roku 2017, poté nastala stagnace až do konce roku 2018, od kterého trend počtu osobních vlaků klesal. Ve významnějších stanicích na státních hranicích tedy trend osobní dopravy v roce 2021 klesal. Ve stanicích Mikulovice a Jindřichov byl podle očekávání trend totožný. Počet vlaků začal klesat v roce 2020. V roce 2021 však provoz na těchto dvou hraničních přechodech rostl. Jelikož byly řady stacionární, bylo vhodné provést jejich dekompozici. Jelikož však mají stejný průběh, stačilo rozložit pouze jednu časovou řadu a výsledky interpretovat pro obě stanice. Rozložena byla časová řada počtu osobních vlaků ve stanici Mikulovice. K odchylce docházelo v květnu a listopadu, kdy byly počty vlaků pravidelně menší. Z hlediska nákladní dopravy ve všech stanicích v posledních dvou letech provoz narůstal, na rozdíl od osobní dopravy. U řad počtu nákladních vlaků nebyla provedena dekompozice žádné stanice, jelikož řady byly nestacionární a obsahovaly velké množství výkyvů.

7 KLASIFIKACE TRATÍ

Cílem klasifikace bylo nalezení podobných a zároveň specificky odlišných železničních úseku nebo stanic podle provozu. Za první část klasifikace se dají považovat výstupy v kapitole 5 *Struktura železniční dopravy*. Zde byly úseky klasifikovány podle celkového počtu vlakových souprav a podle jejich celkové hmotnosti v letech 2016 a 2021. Monitorovací body SR70 byly klasifikovány do kategorií podle procentuálního podílu druhu dopravy na celkovém počtu vlakových souprav nebo celkové hmotnosti vlakových souprav v letech 2016 a 2021. Druhá část klasifikace regionálních a celostátních tratí Olomouckého a Moravskoslezského kraje byla realizována metodou shlukování. Při shlukové analýze železničních úseků byla využita především metoda hierarchického shlukování. Otestována byla taky nehierarchická varianta, a to k-středová metoda (k-means). Železniční úseky byly shlukovány na základě ročních a měsíčních agregací počtu a hmotnosti vlakových souprav v roce 2021. Tento rok byl vybrán, jelikož obsahuje nejaktuálnější data. Shlukováno bylo 462 železničních úseku z roku 2021.

Pro lepší interpretaci výsledných shluků, byly vytvořeny kategorizační tabulky podle počtu a hmotnosti vlakových souprav. Tyto tabulky umožnily klasifikovat shluky do 6 nadefinovaných kategorií. V případě shlukování ročních agregací byly shluky kategorizovány na základě tabulek ročních agregací počtu a hmotnosti osobních a nákladních vlakových souprav. Naopak v případě shlukování měsíčních agregací byla kategorizace shluků provedena podle kategorizační tabulky měsíčních úhrnu počtu osobních a nákladních vlakových souprav. Každý shluk byl do příslušné kategorie přiřazen na základě hodnoty mediánu sledovaného atributu. Celková charakteristika shluků byla doplněna poznatky získanými z dendrogramu v programu Orange, a také z mapových výstupů, které poskytly prostorové informace.

Tab. 7 Kategorizační tabulka podle počtu vlakových souprav za rok

Počet osobních vlaků	Kategorie	Počet nákladních vlaků
20 001 a více	Velmi vysoký	10 001 a více
15 001–20 000	Vysoký	5 001–10 000
10 001–15 000	Střední	2 001–5 000
5 001–10 000	Nízký	1 001–2 000
1 001–5 000	Velmi nízký	101– 1 000
méně než 1 000	Minimální	méně než 100

Tab. 8 Kategorizační tabulka podle hmotnosti vlakových souprav za rok

Hmotnost osobních vlaků (tis.t)	Kategorie	Hmotnost nákladních vlaků (tis. t)
5 001 a více	Velmi vysoká	15 001 a více
2 501–5 000	Vysoká	10 001–15 000
1 001–2 500	Střední	5 001–10 000
501–1 000	Nízká	1 001–5 000
101–500	Velmi nízká	101–1 000
méně než 100	Minimální	100

Tab. 9 Kategorizační tabulka podle počtu vlakových souprav za měsíc

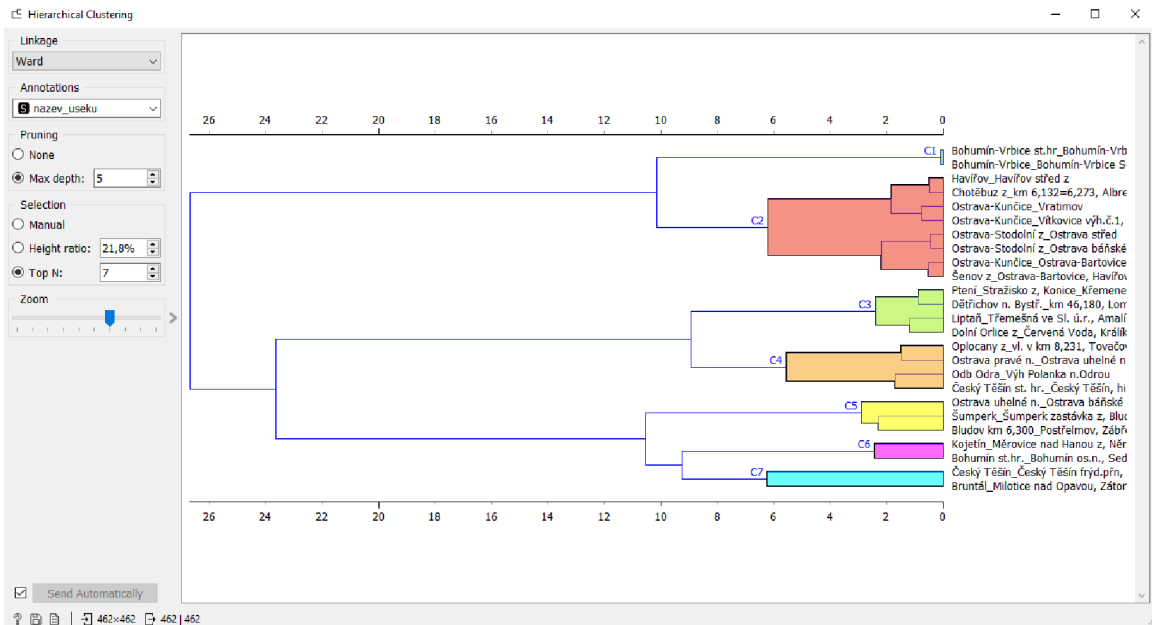
Počet osobních vlaků	Kategorie	Počet nákladních vlaků
2 501 a více	Velmi vysoký	801 a více
2 001–2 500	Vysoký	401–800
1 001–2 000	Střední	201–400
501–1 000	Nízký	101–200
101–500	Velmi nízký	51–100
méně než 100	Minimální	méně než 50

V poslední fázi klasifikace tratí bylo provedeno shlukování železničních úseků podle průběhu časových řad. Cílem shlukování podle časových řad bylo nalézt úseky, které mají podobný a zároveň rozdílný časový průběh počtu osobních nebo nákladních vlakových souprav. Testováno bylo použití metody dynamického borcení časové osy na data provozu na železnici.

7.1 Hierarchické shlukování

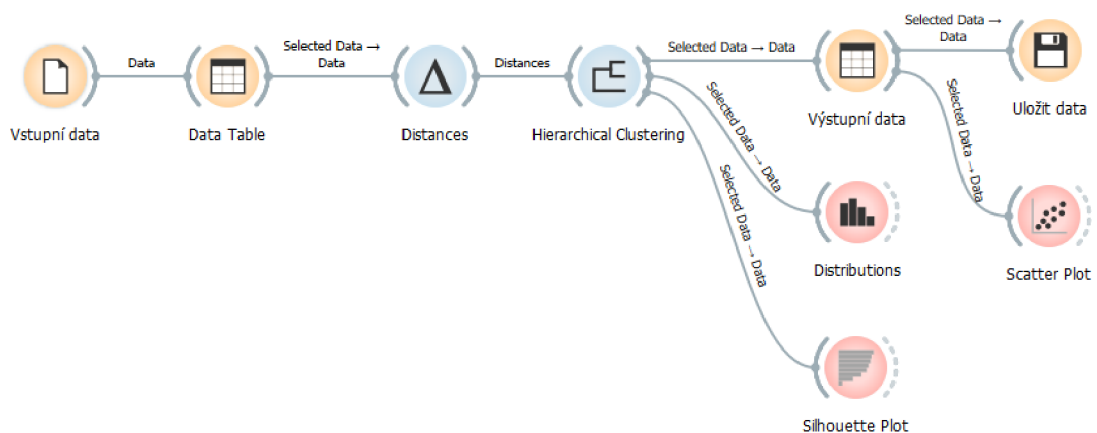
Pro shlukování železničních úseků byla nejprve použita hierarchická metoda, která byla považována za výchozí metodou shlukové analýzy. Úseky byly nejprve shlukovány na základě ročních agregací počtu a váhy osobních a nákladních vlakových souprav v roce 2021. Poté proběhlo shlukování podle hodnot měsíčních agregací z roku 2021. Testovány byly různé metody hierarchického shlukování a různé vzdálenostní metriky. Nakonec z nich byla vybrána nejvhodnější kombinace pro dané ukazatele analyzovaných dat.

Hierarchické shlukování bylo realizováno v programu Orange. Byl sestaven jednoduchý model, kde hlavní úlohu hrály uzly *Distances* a *Hierarchical Clustering*. Na počátku modelu byl uzel pro nahrání vstupních dat. Vstupní data byla ve formě tabulky, kde řádek znamenal záznam a sloupce obsahovaly hodnoty počtu nebo váhy osobních a nákladních vlaků. V uzlu *Distances* byla vybírána konkrétní vzdálenostní metrika, podle které se měřila vzdálenost objektů. Z nabídky byla testována především Euklidovská, Manhattanská, Cosinova a Mahalonobisova vzdálenost. V případě využití Euklidovské a Manhattanské vzdálenosti byla data vždy normalizována. Normalizace dat v uzlu *Distances* se vždy provádí po sloupcích. V uzlu *Hierarchical Clustering* se nastavují jednotlivé parametry hierarchického shlukování jako metoda shlukování, maximální rozvětvení shlukovacího stromu a počet výsledných shluků. V nabídce metod shlukování jsou *Single linkage* – metoda nejbližšího souseda, *Average linkage* – metoda průměrné vazby, *Weighted linkage*, *Complete linkage* – metoda nejvzdálenějšího souseda, *Ward linkage* – Wardova metoda. Otestovány byly všechny metody. V tomto uzlu se také zobrazuje výsledný dendrogram.



Obr. 39 Výsledný dendrogram hierarchického shlukování

Na uzel *Hierarchical Clustering* byl navázán uzel *Distributions*, který ukazuje rozložení objektů do výsledných clusterů, a dále uzel *Silhouette Plot*, na jehož základě byla vyhodnocována kvalita výsledných shluků. Hodnota siluety u jednotlivých objektů naznačuje, jak moc objekt zapadá do přiřazeného shluku. Hodnota blízko 1 naznačuje, že objekt je blízko středu shluku. Naopak hodnota blízko 0 naznačuje, že objekt je na hranici mezi 2 shluky (Orange Visual Programming, 2015). Čím více objektů s nízkou hodnotou siluety, tím méně byly výsledné shluky kvalitní, a bylo nutné změnit jejich počet. Výsledky shlukování byly zobrazeny v uzlu Výstupní data, kde bylo v novém sloupci *Cluster* pro každý objekt číslo přiřazeného shluku. V uzlu *Scatter plot* byly shlukované objekty zobrazeny na bodovém grafu, kdy byly barevně odlišeny podle příslušného shluku. Nakonec byly výsledky uloženy ve formě nové tabulky, a dále zpracovány v GIS softwaru, kde byly vytvořeny mapové výstupy. V textu práce jsou pouze náhledy. Mapové výstupy v plné velikosti jsou součástí příloh práce.



Obr. 40 Workflow hierarchického shlukování v programu Orange

7.1.1 Roční agregace

Železniční úseky byly hierarchicky shlukovány podle roční agregace hodnot počtu a váhy osobních a nákladních vlakových souprav z roku 2021. Byly zvoleny tři různé kombinace atributů, podle kterých byly úseky shlukovány. Nejprve podle počtu osobních a nákladních vlaků, poté podle hmotnosti osobních a nákladních vlaků, a na závěr byla vyzkoušena varianta s využitím všech 4 atributů zároveň, a tedy podle počtu a hmotnosti osobních a nákladních vlaků. Hodnoty byly před procesem shlukování normalizovány. Výsledné shluky byly kategorizovány podle kategorizačních tabulek ročních hodnot (Tab. 7 a Tab. 8).

A. Shlukování podle počtu osobních a počtu nákladních vlakových souprav za rok

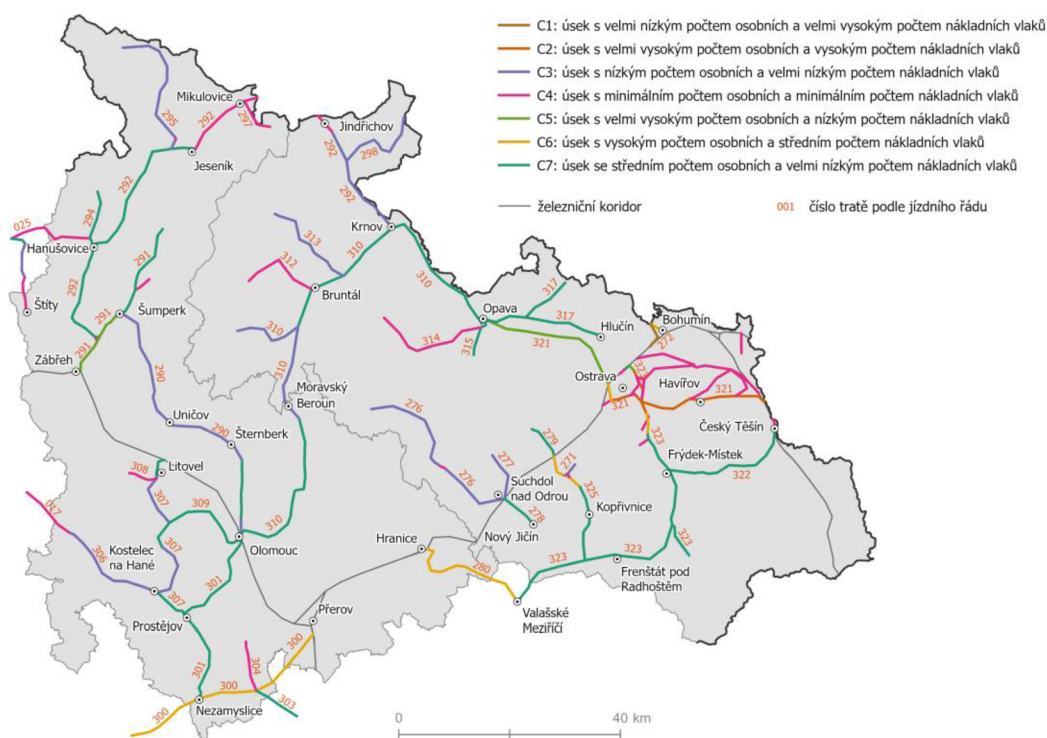
- Wardova metoda
- Mahalanobisova vzdálenost

U klasifikace úseků podle počtu osobních a nákladních vlaků byly z testovaných kombinací shlukovacích metod a vzdálenostních metrik nejlepší výsledky získány při využití Wardovy metody v kombinaci s Euklidovou a Mahalanobisovou vzdáleností do 7 cílových shluků. Volba Euklidovy vzdálenosti vykazovala na rozdíl od Mahalanobisovy vzdálenosti rovnoměrnější distribuci úseků do jednotlivých shluků. Nicméně podle skóre siluety byla podobnost úseků vyšší při využití Mahalanobisovy metriky. Jeden shluk obsahoval pouze 2 úseky. Na bodovém grafu bylo zřejmé, že shluk obsahující byl vytvořen správně, jelikož hodnoty těchto úseků byly odlehlé od zbytku datového souboru. Jednalo se o úseky Bohumín-Vrbice st. hr. – Bohumín-Vrbice S618a a Bohumín-Vrbice – Bohumín-Vrbice S618a, které se vyznačovaly nízkým počtem osobních vlaků, a naopak vysokým počtem nákladních vlaků. Pro finální zpracování byla tedy vybrána Wardova metoda s Mahalanobisovou vzdáleností, jelikož měla vyšší kvalitu výsledných shluků nežli při použití Euklidovy vzdálenosti.

Tab. 10 Hodnoty mediánu počtu osobních a nákladních vlaků

Shluk	Medián počtu osobních vlaků	Medián počtu nákladních vlaků
C1	1 534	13 698
C2	23 644	8 239
C3	6 155	169
C4	126	72
C5	25 262	1 118
C6	16 861	2 977
C7	13 501	282

- C1:** Úseky s velmi nízkým počtem osobních a velmi vysokým počtem nákladních vlakových souprav za rok
- C2:** Úseky s velmi vysokým počtem osobních a vysokým počtem nákladních vlakových souprav za rok
- C3:** Úseky s nízkým počtem osobních a velmi nízkým počtem nákladních vlakových souprav za rok
- C4:** Úseky s minimálním počtem osobních a minimálním počtem nákladních vlakových souprav za rok
- C5:** Úseky s velmi vysokým počtem osobních a nízkým počtem nákladních vlakových souprav za rok
- C6:** Úseky s vysokým počtem osobních a středním počtem nákladních vlakových souprav za rok
- C7:** Úseky se středním počtem osobních a velmi nízkým počtem nákladních vlakových souprav za rok



Obr. 41 Výsledky hierarchického shlukování podle počtu vlaků za rok

Svým provozem jsou specificky odlišné úseky C1. Jedná se o zmíněnou dvojici úseků Bohumín-Vrbice st. hr. – Bohumín-Vrbice S618a a Bohumín-Vrbice – Bohumín-Vrbice S618a, kde podle počtu souprav nákladní doprava výrazně převyšovala osobní. Největší provoz osobních i nákladních vlaků byl na úsecích C2. Ty jsou na tratích č. 323 a 321 na Ostravsku. Počtem osobních vlaků jsou úsekům C2 nejpodobnější úseky C5 a C6. Velmi vysoký počet osobních vlaků, ale nízký nákladních mají úseky C5. Tyto železniční úseky se nachází na celostátních tratích č. 321 v úseku Opava–Ostrava a trati č. 291 v úseku Zábřeh–Šumperk. Úseky C6, na kterých je vůči úsekům C5 větší počet nákladních vlaků, tvoří taktéž celostátní tratě, konkrétně č. 300 a 280. Nejmenší provoz

mají úseky C3 a C4. Jedná se o úseky na regionálních tratích. Výsledný shluk C4 není příliš kvalitní. Z hlediska osobní dopravy je homogenní, jelikož klasifikoval úseky, kde probíhá sezónní provoz (tratě č. 304, 308, 312, 314) nebo zde osobní doprava neprobíhala. Tedy počet osobních vlaků za rok byl minimální. Avšak z pohledu nákladní dopravy obsahuje úseky s počtem nákladních vlaků v jednotkách a současně úseky z Ostravska, kde jejich počet byl vyšší než tisíc souprav za rok. Úseky C7 se středním počtem osobních vlaků a velmi nízkým počtem nákladních jsou na celostátních i regionálních tratích.

B. Shlukování podle hmotnosti osobních a hmotnosti nákladních vlakových souprav za rok

- Wardova metoda
- Mahalanobisova vzdálenost

Při shlukování úseků podle hmotnosti osobních a nákladních vlaků, byly nejlepší výsledky dosaženy při volbě Wardovy metody s využitím Mahalanobisovy vzdálenosti. Úseky byly shlukovány do 7 clusterů. Tento počet vykazoval nejlepší kvalitu výsledných shluků.

Tab. 11 Hodnoty mediánu hmotnosti osobních a nákladních vlaků

Shluk	Medián hmotnosti osobních vlaků (tis.t)	Medián hmotnosti nákladních vlaků (tis.t)
C1	519	18 833
C2	5 676	10 799
C3	2 569	67
C4	4 426	2 335
C5	0	2 208
C6	118	19
C7	722	78

C1: Úseky s nízkou hmotností osobních a velmi vysokou hmotností nákladních vlakových souprav za rok

C2: Úseky s velmi vysokou hmotností osobních a vysokou hmotností nákladních vlakových souprav za rok

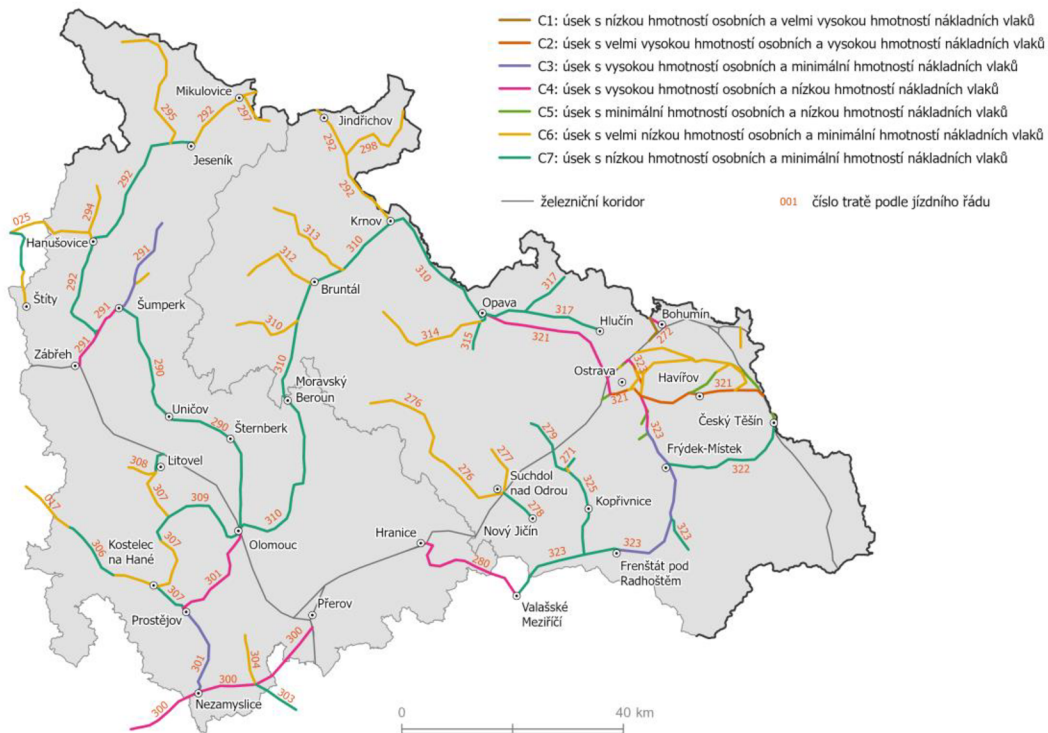
C3: Úseky s vysokou hmotností osobních a minimální hmotností nákladních vlakových souprav za rok

C4: Úseky s vysokou hmotností osobních a nízkou hmotností nákladních vlakových souprav za rok

C5: Úseky s minimální hmotností osobních a nízkou hmotností nákladních vlakových souprav za rok

C6: Úseky s velmi nízkou hmotností osobních a minimální hmotností nákladních vlakových souprav za rok

C7: Úseky s nízkou hmotností osobních a minimální hmotností nákladních vlakových souprav za rok



Obr. 42 Výsledky hierarchického shlukování podle hmotnosti vlaků za rok

Stejně jako u varianty A, vznikl shluk C1 obsahující pouze dva úseky v Bohumíně, které jsou odlišné a odlehlé i z hlediska hmotnosti souprav. Hmotnost osobních i nákladních souprav byla největší v úsecích C2. Ty jsou na celostátních tratích č. 323 a 321. Ostatní úseky jsou již zatíženy buď nízkou nebo minimální hmotností nákladních souprav. Vysoká hmotnost osobních vlaků je na úsecích C3 a C4, které jsou na celostátních tratích č. 280 z Hranic do Valašského Meziříčí, 291 z Šumperka do Zábřeha, 300 z Přerova, 301 mezi Olomoucí a Prostějovem, 321 z Opavy do Ostravy. Hmotností osobních vlaků jsou minimálně zatíženy úseky u Havířova. Důvodem je neprobíhající osobní doprava v těchto místech. Úseky C6 jsou na koncových regionálních tratích, a jsou zatíženy nízkou hmotností osobních a minimální hmotností nákladních vlaků. Nízkou hmotnost osobních vlaků mají úseky C7, které jsou na regionálních i celostátních tratích. Provozem s nízkou hmotností osobních souprav jsou si podobné tratě č. 290, 292, 310, 317, 322, 323. Výsledky shlukování varianty A a B jsou si velmi podobné. Důvodem je velká korelace mezi počtem a hmotností souprav. Čím více vlaků úsekem projede, tím větší celková hmotnost souprav.

C. Shlukování podle počtu osobních, počtu nákladních, hmotnosti osobních a hmotnosti nákladních vlakových souprav za rok

- Wardova metoda
- Euklidova vzdálenost

Vyzkoušeno bylo také shlukování úseků podle počtu a hmotnosti osobních a nákladních vlakových souprav současně. Výsledky této varianty byly při interpretaci pojmenovány jako celkový provoz, jelikož kombinují všechny sledované atributy jak osobní, tak nákladní dopravy. Data bylo nutné normalizovat. Při této variantě byly nejlepší výsledky shlukování získány opět při využití Wardovy metody, tentokrát v kombinaci s Euklidovou vzdáleností. Při volbě celkových 7 shluků byly podle hodnot silhouette výsledné shluky nejkvalitnější.

Tab. 12 Hodnoty mediánu počtu a hmotnosti osobních a nákladních vlaků

Shluk	Medián			
	Počtu osobních vlaků	Počtu nákladních vlaků	Hmotnosti osobních vlaků (tis. t)	Hmotnosti nákladních vlaků (tis. t)
C1	1 534	13 698	519	18 833
C2	37 529	11 613	8 472	12 625
C3	23 619	7 597	5 674	9 851
C4	16 722	441	2 967	118
C5	19 677	2 959	4 595	2 481
C6	774	95	27	32
C7	11 478	282	703	79

C1: Úseky s velmi nízkým počtem a nízkou hmotností osobních vlakových souprav, a velmi vysokým počtem a velmi vysokou hmotností nákladních vlakových souprav za rok

C2: Úseky s velmi vysokým počtem a velmi vysokou hmotností osobních vlakových souprav, a velmi vysokým počtem a vysokou hmotností nákladních vlakových souprav za rok

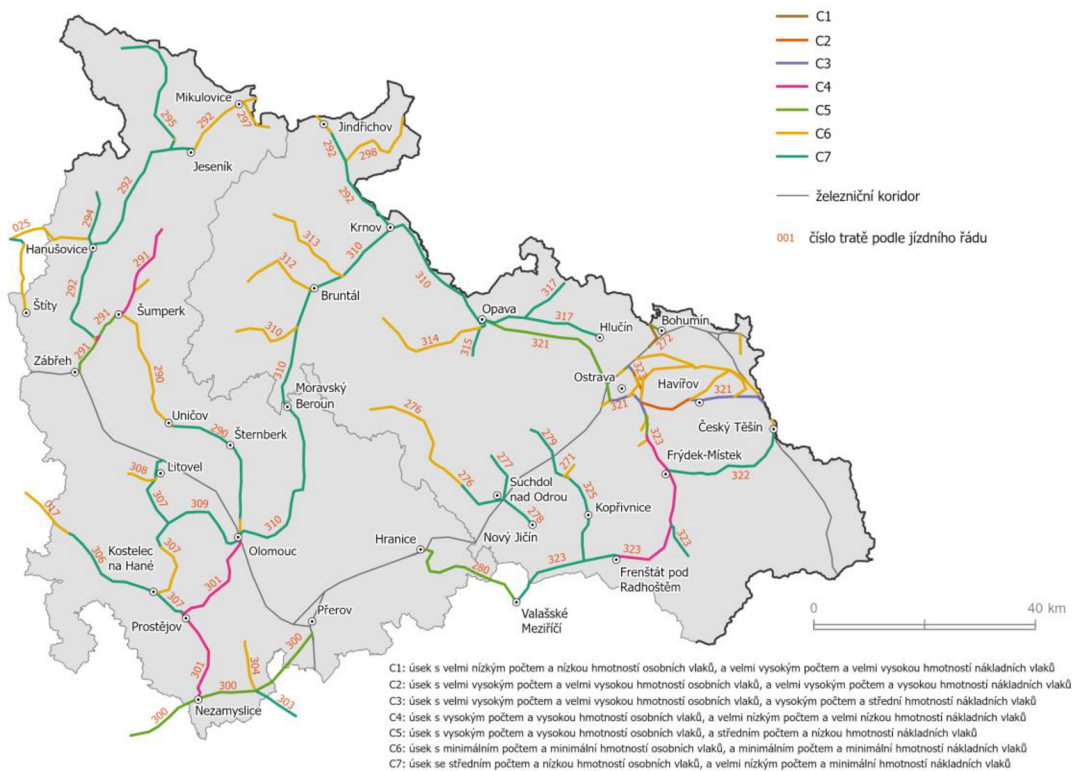
C3: Úseky s velmi vysokým počtem a velmi vysokou hmotností osobních vlakových souprav, a vysokým počtem a střední hmotností nákladních vlakových souprav za rok

C4: Úseky s vysokým počtem a vysokou hmotností osobních vlakových souprav, a velmi nízkým počtem a velmi nízkou hmotností nákladních vlakových souprav za rok

C5: Úseky s vysokým počtem a vysokou hmotností osobních vlakových souprav, a středním počtem a nízkou hmotností nákladních vlakových souprav za rok

C6: Úseky s minimálním počtem a minimální hmotností osobních vlakových souprav, a minimálním počtem a minimální hmotností nákladních vlakových souprav za rok

C7: Úseky se středním počtem a nízkou hmotností osobních vlakových souprav, a velmi nízkým počtem a minimální hmotností nákladních vlakových souprav za rok



Obr. 42 Výsledky hierarchického shlukování podle počtu a hmotnosti vlaků za rok

Celkovým provozem byly opět odlišné úseky dva úseky v Bohumíně. Ty mají nízký počet vlaků v osobní dopravě, ale velmi vysoký provoz podle počtu a hmotnosti nákladních vlaků. Celkově největší provoz je v úsecích C2. Jim nejpodobnější jsou úseky C3. Ty mají také velmi vysoký provoz osobní dopravy, rozdílem jsou nízké hodnoty nákladní dopravy. Všechny tyto úseky jsou na tratích č. 323 a 321 u Ostravy a Havířova. Úseky C4 mají stejně jako úseky C5 vysoký provoz osobní dopravy. Mírně se liší nákladní dopravou, kdy na úsecích C5 jezdí větší počet souprav. Takový provoz je především na celostátních tratích ze Šumperka do Zábřehu, z Olomouce do Prostějova, z Přerova do Nezamyslic, z Hranic do Valašského Meziříčí a z Opavy do Ostravy. Největší provoz osobní dopravy je na regionální trati č. 291 v úseku Šumperk – Kouty nad Desnou a trať č. 323 v úseku Frenštát pod Radhoštěm–Ostrava. Minimálně provoz byl na úsecích C6. Tyto úseky jsou především na regionálních tratích, v okolí Ostravy, ale kromě stanice Bohumín i na tratích s hraničním přechodem. Střední provoz osobních vlaků mají celostátní i regionální tratě na Jeseník, Olomouc, Šternberk, Krnov, Nový Jičín, Kopřivnici.

7.1.2 Měsíční agregace

Železniční úseky byly hierarchicky shlukovány i podle měsíčních agregací hodnot počtu osobních a nákladních vlakových souprav. Jednalo se o měsíční agregace po dobu roku 2021. Pro shlukování podle měsíčních agregací byl sledován pouze ukazatel počtu vlakových souprav, a byly definovány dvě varianty zpracování podle druhu vlakové dopravy. V první variantě byly úseky shlukovány podle počtu osobních vlakových souprav za měsíc, ve druhé variantě podle počtu nákladních vlakových souprav za měsíc. Jelikož byly úseky shlukovány na základě jednoho atributu se stejnou jednotkou, jehož hodnoty byly rozdělené do 12 měsíců, nebylo třeba hodnoty před zpracováním normalizovat. Výsledné shluky byly kategorizovány podle kategorizační tabulky měsíčních hodnot (Tab. 9).

A. Shlukování podle počtu osobních vlakových souprav za měsíc

- Wardova metoda
- Euklidova vzdálenost

Měsíční hodnoty počtu osobních vlakových souprav byly shlukovány Wardovou metodou s využitím Euklidovy vzdálenosti. Podle grafu siluety si úseky v jednotlivých shlucích byly nejpodobnější při volbě shlukování do 5 clusterů. Výsledné shluky vyšly optimálně, jelikož každý z nich reprezentoval jednu z definovaných kategorií počtu vlakových souprav, vyjma kategorie velmi nízkého počtu.

Tab. 13 Hodnoty mediánu počtu osobních vlaků

Shluk	Medián počtu osobních vlaků
C1	56
C2	635
C3	1 236
C4	2 861
C5	2 152

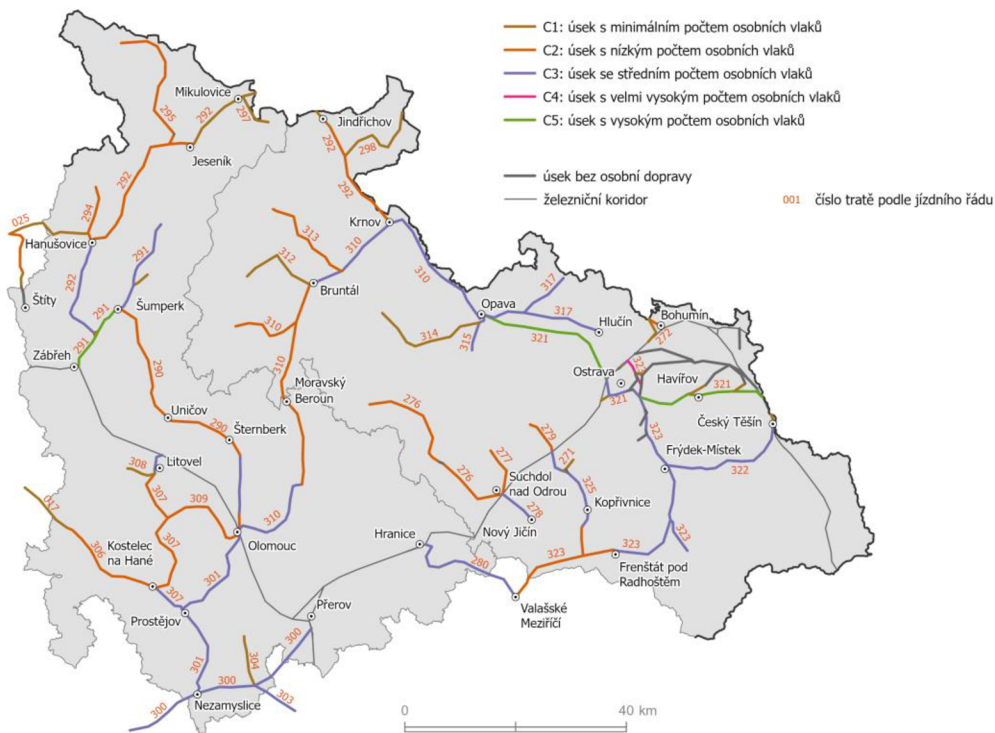
C1: Úseky s minimálním počtem osobních vlakových souprav za měsíc

C2: Úseky s nízkým počtem osobních vlakových souprav za měsíc

C3: Úseky se středním počtem osobních vlakových souprav za měsíc

C4: Úseky s vysokým počtem osobních vlakových souprav za měsíc

C5: Úseky s vysokým počtem osobních vlakových souprav za měsíc



Obr. 43 Výsledky hierarchického shlukování podle počtu osobních vlaků za měsíc

Minimální počet osobních vlaků za měsíc projede na regionálních tratích č. 017, 025 do Štítů, 304 do Tovačova, 308 do Mladče, 312 do Malé Morávky, 314 do Jakartovic, 271 na letiště do Mošnova a 298. Většinou jde o tratě se sezonním provozem. Minimální počet je i na úsecích u hraničních přechodů. Nízký počet vlaků za měsíc je především na regionálních tratích. Mezi celostátní tratě s nízkým počtem patří trať č. 292 a 310. Střední počet vlaků je na celostátních i regionálních tratích. Překvapivé je střední provoz u tratí č. 280 z Hranic do Valašského Meziříčí, 300 z Přerova do Nezamyslic a 301 z Olomouce do Prostějova, jelikož na základě ročních agregací jsou klasifikovány jako tratě s vysokým provozem. Vysoký počet je v úsecích Šumperk–Zábřeh a na trati č. 321. Velmi vysokým počtem osobních vlaků za měsíc disponuje trať č. 323 v úsecích mezi stanicemi Ostrava-střed, Ostrava-Stodolní, Ostrava-Kunčičky, Ostrava-Kunčice.

B. Shlukování podle počtu nákladních vlakových souprav za měsíc

- Wardova metoda
- Euklidova vzdálenost

Stejně jako u shlukování podle počtu osobních vlakových souprav bylo i v případě nákladních nejkvalitnějších výsledků získáno pomocí hierarchického shlukování Wardovou metodou s Euklidovou vzdálenostní metrikou. Výsledné shluky byly opět kategorizovány podle jejich mediánů. Zastoupení kategorií počtu nákladních vlaků bylo totožné jako u osobních vlaků. Zastoupeny byly všechny kategorie kromě kategorie velmi nízkého počtu.

Tab. 14 Hodnoty mediánu počtu nákladních vlaků

Shluk	Medián počtu nákladních vlaků
C1	978
C2	700
C3	14
C4	113
C5	264

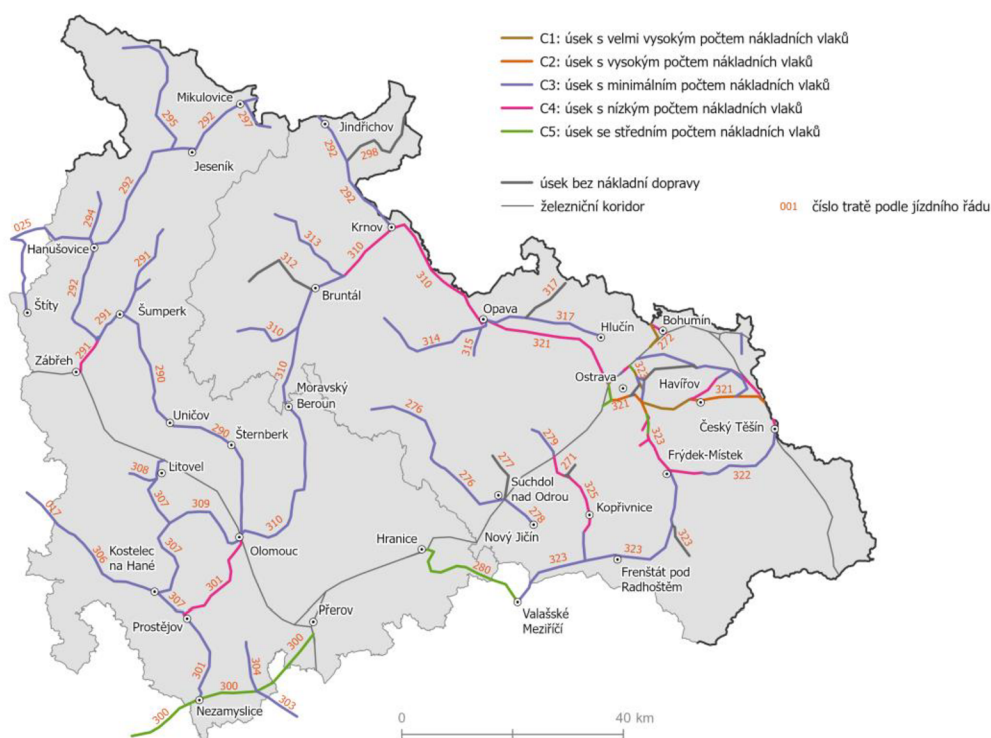
C1: Úseky s velmi vysokým počtem nákladních vlakových souprav za měsíc

C2: Úseky s vysokým počtem nákladních vlakových souprav za měsíc

C3: Úseky s minimálním počtem nákladních vlakových souprav za měsíc

C4: Úseky s nízkým počtem nákladních vlakových souprav za měsíc

C5: Úseky se středním počtem nákladních vlakových souprav za měsíc



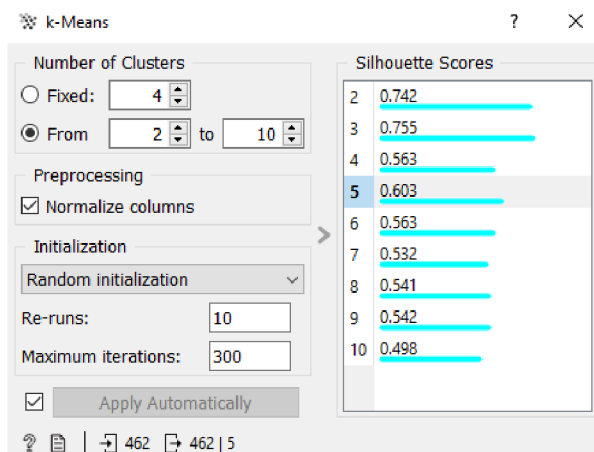
Obr. 44 Výsledky hierarchického shlukování podle počtu nákladních vlaků za měsíc

Na většině tratí, bez ohledu na to, jestli jde o regionální či celostátní, za měsíc projede minimální počet nákladních vlaků. Nízký počet nákladních vlaků měsíčně jezdí na tratích mezi Olomoucí a Prostějovem, u Krnova, Opavy, Kopřivnice, Ostravy i Frýdku-Místku. Střední počet jezdí na trati č. 300 z Přerova a na trati 280 mezi Hranicemi a Valašským Meziříčím. Úseky s nejvyšším počtem nákladních vlaků jsou podle očekávání na Ostravsku, jelikož se jedná o významnou průmyslovou oblast. Největší provoz je na tratích z Bohumína do Polska, a z Ostravy do Havířova.

7.2 Shlukování metodou k-means

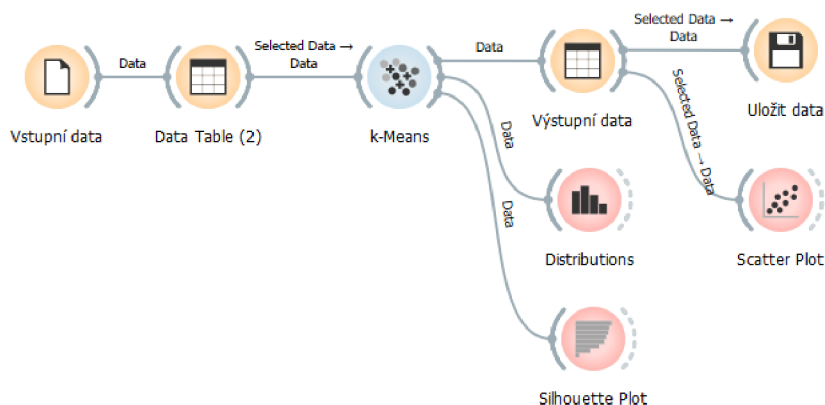
Ačkoliv byla pro shlukovou analýzu v rámci této práce výchozí hierarchická metoda shlukování, byla na stejná data otestována i nehierarchická metoda, konkrétně metoda k-means (k-středů). Cílem bylo porovnat výsledky jednotlivých metod, zdali jsou téměř shodné nebo výrazně odlišné, případně z těchto metod vybrat tu vhodnější pro zpracování dat provozu na železnici. Na rozdíl od hierarchického shlukování se u této metody předpokládá předem stanovený cílový počet shluků. Aby byly výsledky porovnatelné, byl vždy zvolen stejný počet shluků, jako u hierarchického shlukování. Pro shlukování ročních agregací se jednalo o 7 cílových shluků, u měsíčních agregací o 5 cílových shluků.

Shlukování metodou k-středů bylo realizováno opět v programu Orange. Byl sestaven podobný model jako pro hierarchické shlukování. Na začátku modelu je uzel pro načtení dat. Samotný proces shlukování provádí uzel *k-Means*. V jeho dialogovém okně je možné nastavit fixní počet shluků nebo interval, ve kterém algoritmus najde jejich nejvhodnější počet. Nejvhodnější počet shluků je označen nejvyšší hodnotou silhouette score v pravém sloupci. Dále je zde možné nastavit normalizaci dat, před procesem shlukování. V parametru *Initialization* se volí metoda určení počátečních středů shluků. Při zpracování byla zvolena metoda *Random initialization*, která středy rozmístí náhodně a poté jsou při každé iteraci aktualizovány.



Obr. 45 Dialogové okno nastavení uzlu k-Means

Na uzel *k-Means* byly napojeny uzel *Distributions*, pro informaci o rozložení úseků v jednotlivých shlučích, a také uzel *Silhouette Plot*, pro kontrolu podobnosti úseků ve shlučích. Pomocí uzlu *Scatter Plot* byly objekty vykresleny do bodového grafu. Po provedení shlukování vznikly v datech dva nové sloupce. V sloupci *Cluster* byl pro každý objekt uveden název příslušného shluku, a ve sloupci *Silhouette* skóre siluety, udávající podobnost objektu vůči shluku, do kterého byl zařazen. Výsledky byly ukládány ve formě tabulek a dále zpracovávány v GIS software, kde byly sestavovány mapové výstupy, které jsou součástí příloh této práce.



Obr. 46 Workflow shlukování k-means metodou

7.2.1 Roční agregace

Železniční úseky byly metodou k-means shlukovány podle roční agregace hodnot počtu a váhy osobních a nákladních vlakových souprav z roku 2021. Shlukovány byly podle stejných kombinace atributů, jako u hierarchického shlukování. Hodnoty byly předzpracovány pomocí jejich normalizace. Výsledné shluky byly stejně jako výsledky hierarchického shlukování kategorizovány podle kategorizačních tabulek ročních hodnot (Tab. 7 a Tab. 8).

A. Shlukování podle počtu osobních a počtu nákladních vlakových souprav za rok

Nejvyšší skóre siluety vyšlo při klasifikaci do 7 cílových shluků, což odpovídá i výslednému počtu shluků při hierarchickém shlukování. Výsledné shluky byly obdobné jako u shlukování hierarchickou metodou. I u shlukování metodou k-means byly do samostatného shluku zařazeny specifické úseky v Bohumíně. Největší provoz osobní i nákladní dopravou byl především na celostátních tratích. Nejméně naopak úseky na regionálních tratích, které jsou méně významné. Při shlukování počtu osobních a počtu nákladních vlakových souprav za rok metodou k-means bylo dosaženo téměř stejných výsledků, jako při hierarchickém shlukování.

B. Shlukování podle hmotnosti osobních a hmotnosti nákladních vlakových souprav za rok

Nejlepší skóre siluety shlukování vyšlo při 5 shlucích, ale aby bylo možné porovnat výsledek s výstupy hierarchického shlukování, bylo zvoleno shlukování do 7 clusterů. Opět jako v případě varianty A, byly výsledné shluky podobné těm, které vznikly hierarchickou metodou. Váhou osobních i nákladních vlaků byly nejvíce vytížené úseky na stejných celostátních tratích. Nejméně zase stejné úseky regionálních a vybraných celostátních tratích. V případě shlukování podle hmotnosti souprav tedy nebyly významné rozdíly mezi výsledky testovaných metod.

C. Shlukování podle počtu osobních, počtu nákladních, hmotnosti osobních a hmotnosti nákladních vlakových souprav za rok

Při shlukování této kombinace atributů byl nejvhodnější počet 7 cílových shluků. Stejný počet vyšel byl zvolen i u hierarchické metody. U této varianty již byly shluky tvořeny jinými kategoriemi nežli u hierarchické metody. Kromě hmotnosti nákladních vlakových souprav, výsledné shluky zastoupily všechny nadefinované kategorie pro ostatní

atributy. Dalo by se tedy tvrdit, že tato metoda shlukování byla citlivější a podala podrobnější výsledky.

Tab. 15 Hodnoty mediánu počtu a hmotnosti osobních a nákladních vlaků

Shluk	Medián			
	Počtu osobních vlaků	Počtu nákladních vlaků	Hmotnosti osobních vlaků (tis. t)	Hmotnosti nákladních vlaků (tis. t)
C1	14 345	359	1 207	97
C2	23 648	8 239	5 677	10 796
C3	0	2 032	0	2 208
C4	19 568	1 628	4 332	1 645
C5	266	13	18	7
C6	7 673	168	367	53
C7	1 534	13 698	519	18 833

C1: Úseky se středním počtem a střední hmotností osobních, a velmi nízkým počet a minimální hmotností nákladních vlakových souprav za rok

C2: Úseky s velmi vysokým počtem a velmi vysokou hmotností osobních, a vysokým počtem a vysokou hmotností nákladních vlakových souprav za rok

C3: Úseky s minimálním počtem a minimální hmotnost osobních, střední počet a nízká hmotnost nákladních

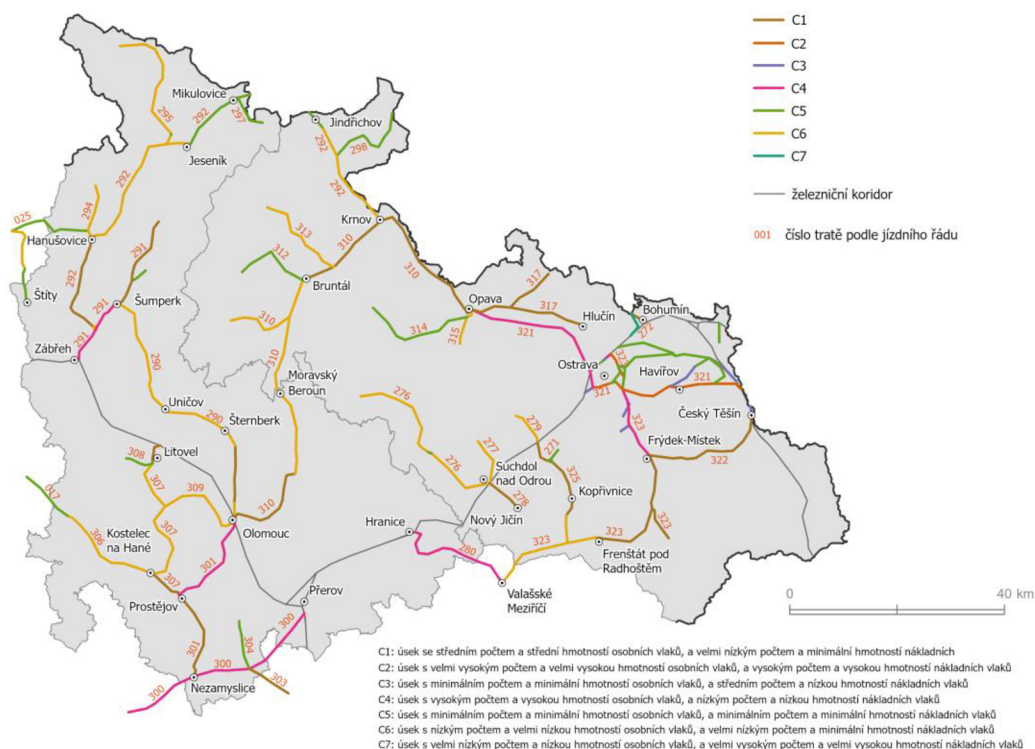
C4: Úseky s vysokým počtem a vysoká hmotností osobních, a nízkým počtem a nízkou hmotností nákladních vlakových souprav za rok

C5: Úseky s minimálním počtem a minimální hmotností osobních, a minimálním počtem a minimální hmotností nákladních vlakových souprav za rok

C6: Úseky s nízkým počtem a velmi nízkou hmotností osobních, a velmi nízkým počtem a minimální hmotností nákladních vlakových souprav za rok

C7: Úseky s velmi nízkým počtem a nízkou hmotností osobních, a velmi vysokým počtem a velmi vysokou hmotností nákladních vlakových souprav za rok

Všeobecně jsou výsledky podobné jako výsledky hierarchické metody. Největší provoz byl na stejných celostátních tratích. Nejmenší provoz byl opět na koncových regionálních tratích. Rozdíly byly v přiřazení některých úseků do shluku se středním počtem osobních vlaků, jako například trať č. 291 v úseku Šumperk–Kouty nad Desnou nebo trať č. 323 v úseku Frenštát pod Radhoštěm – Frýdek-Místek, které byly u hierarchické metody přiřazeny mezi úseky s vysokým počtem a vysokou hmotností osobních vlaků. Stejně jako u hierarchické metody, byly odhaleny odlišné úseky v Bohumíně. Ty mají nízký provoz osobní dopravy, ale velmi vysoké počty a váhu nákladních vlaků.



Obr. 47 Výsledky k-means shlukování podle počtu a hmotnosti vlaků za rok

7.2.2 Měsíční agregace

Železniční úseky byly metodou k-means shlukovány podle měsíčních agregací hodnot počtu osobních a nákladních vlakových souprav z roku 2021. Byly provedeny dvě varianty zpracování, stejně jako u hierarchické metody. V první variantě byly úseky shlukovány podle počtu osobních vlakových souprav za měsíc, ve druhé variantě podle počtu nákladních vlakových souprav za měsíc. Výsledné shluky byly taktéž kategorizovány podle kategorizační tabulky měsíčních hodnot (Tab. 9).

A. Shlukování podle počtu osobních vlakových souprav za měsíc

Doporučený počet cílových shluků byl 6, jelikož měl nejvyšší skóre siluety. Aby byly výsledky porovnatelné s předchozí hierarchickou metodou, tak byla zvoleno shlukování do 5 clusterů. Výsledky nebyly pro interpretaci shluků příznivé, neboť vznikly dva velice podobné shluky, které spadaly do kategorie se středním počtem osobních vlaků za měsíc. Nicméně úseky patřící do těchto shluků byly jako úseky se středním počtem osobních vlaků označeny i při použití hierarchické metody. Rozdílem je absence shluku úseků s velmi vysokým počtem osobních vlaků. U hierarchické metody bylo takto klasifikováno zhruba 5 úseků. Ty byly u metody k-means klasifikovány jako úseky s vysokým počtem osobních vlaků. Až na tyto rozdíly, byly výstupy obou metod stejné.

B. Shlukování podle počtu nákladních vlakových souprav za měsíc

I u shlukování podle počtu nákladních vlakových souprav bylo nejvhodnější shlukování do 5 clusterů. Kvůli zachování porovnatelnosti, byl zvolen počet 5. Na většině úseků projede nízký počet nákladních vlaků za měsíc. Nejvíce jich měsíčně jezdí v Bohumíně a na trati č. 321 v Ostravě a Havířově. Výsledky metody k-means jsou srovnatelné s výsledky hierarchické metody.

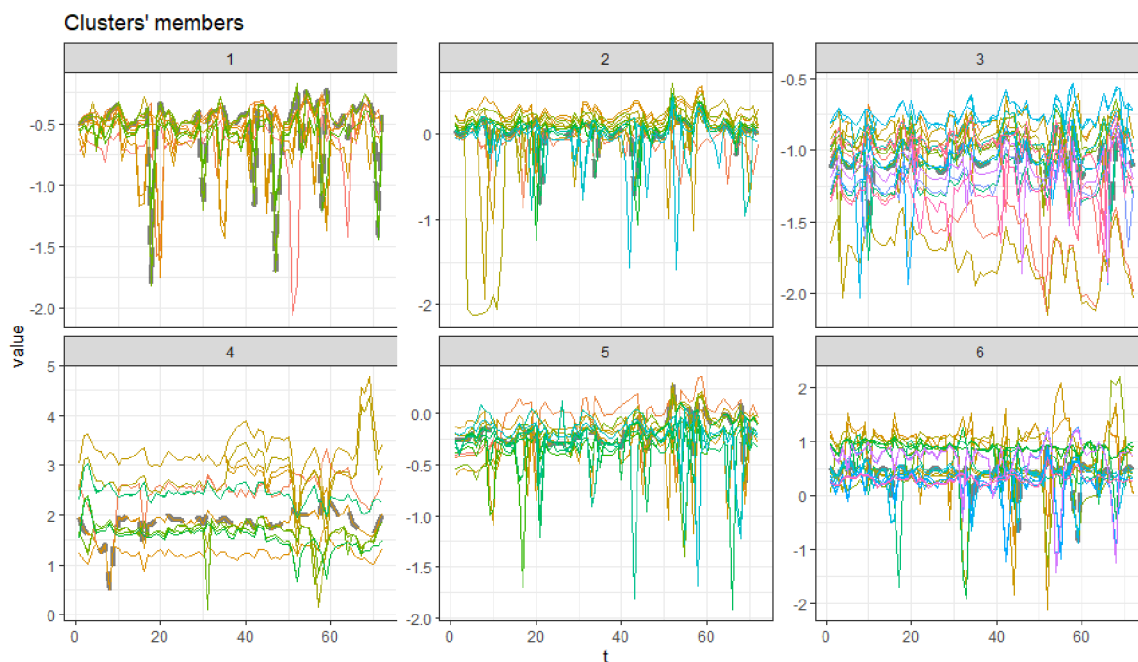
7.3 Shlukování podle časových řad

Cílem shlukování podle časových řad bylo nalézt úseky, které mají podobný a zároveň rozdílný časový průběh počtu osobních nebo nákladních vlakových souprav. Podobnost byla hledána na základě průběhu křivek očištěných měsíčních agregací počtu osobních nebo nákladních vlakových souprav na úsecích. Primárním zdrojem dat této metody shlukování bylo všech 434 regionálních a celostátních úseků, které mají údaje od roku 2016 až do roku 2021. Z primárního zdroje však bylo nadefinováno několik variant vstupních dat, které byly následně shlukovány. V rámci shlukování časových řad metodou nejbližších středů bylo záměrem otestovat metodu dynamického borcení časové osy (DTW) na data provozu na železnici, případně využít Euklidovu vzdálenost.

Shlukování podle časových řad bylo realizováno ve statistickém programu R Studio, pro který byl napsán programový kód (Příloha 7). Základem tohoto skriptu je balíček *dtwclust*, obsahující funkce pro shlukování časových řad spolu s optimalizacemi pro vzdálenost DTW. Hlavní funkce *tsclust* provádějící shlukování nabízí možnost využití i jiných metrik, včetně různých metod shlukování. Pro zpracování byla využita nehierarchická metoda nejbližších středů. Dalšími použitými balíčky byly *BBmisc* s funkcí pro normalizaci vstupních dat, *tidyverse* s funkcemi pro úpravu struktury vstupních data a balíčky *openxlsx* a *readxl* pro práci s daty ve formátu *.xlsx*.

Data byla před shlukováním normalizována odchylkou od středu. Shluky jsou tedy výsledkem podobnosti průběhu časových řad a neudávají absolutní hodnoty počtu vlaků. Pro shlukování metodou nejbližších středů byl ve funkci *tsclust* nastaven parametr *type* jako *partitional*. Pro hledání vhodného počtu cílových shluků byl pro parametr *k* zvolen rozsah 2-9 shluků, aby se snížila výpočetní náročnost a bylo možné výsledky interpretovat. Výsledný počet shluků byl vybírán na základě vyhodnocení statistiky průměrných vzdáleností mezi objekty uvnitř shluku. Poté bylo shlukování provedeno s pevně daným počtem, který vykazoval nejmenší hodnoty vzdáleností. Jako metoda určení centroidu shluků, byla v parametru *centroid* nastavena možnost *pam*, jelikož je v literatuře jednou z nejspojovanějších metod s využitím metriky DTW (Sobolewska, 2019). Metoda PAM je založena na principu, že centroidy shluků jsou vždy jednou z časových řad, která je v datech (rdrr.io, 2022). Aby shlukování proběhlo s využitím metriky DTW, byl parametr *distance* = "dtw". V případě využití Euklidovy metriky byl parametr *distance* = "euclidean". Výsledné shluky byly vizualizovány pomocí jednoduchých liniových grafů. Pro interpretaci výsledků shlukování sloužil medoid jednotlivých shluků. Medoidem je vždy jedna z časových řad, jejíž vzdálenost od ostatních objektů shluku je minimální. Tato časová řada tedy nejlépe charakterizuje příslušný shluk.

Provedeno bylo nejprve shlukování podle metody DTW a následně byly výsledky porovnány s výsledky při použití Euklidovy metriky. U obou způsobů byla pro určení centroidu použita možnost *pam*. V případě, že byly výsledky srovnatelné, byly popsány výstupy shlukování s metodou DTW.



Obr. 48 Ukázka výsledných shluků časových řad

Na mapových vizualizacích bylo viditelné, že některé úseky nebyly klasifikovány. Shlukovány byly časové řady úseků, které měly hodnoty za všechny poskytnuté roky. Pro vizualizace však byla použita železniční síť z roku 2021. Některé úseky z roku 2021 předchozí podmínku nesplňovaly, a proto nebyly součástí shlukování a nemohly být klasifikovány.

7.3.1 Časové řady počtu vlaků bez nulových hodnot

U první varianty shlukování bylo snahou hledat podobné úseky podle provozu, který nebyl zatížen výlukami v celém sledovaném období od roku 2016 do roku 2021. Pro tuto úlohu bylo nutné z datového souboru odstranit úseky, na kterých byl v některém z měsíců počet vlaků nulový. Z toho, že byly hledány pouze úseky s nulovými hodnotami, vyplývá, že bylo uvažováno pouze o úplných výlukách v trvání minimálně jednoho měsíce. Postup identifikace těchto úseků byl popsán v kapitole 6.1. Nevýhodou tohoto postupu je, že zároveň dojde k odstranění úseků, u kterých nulový počet vlaků neznamená anomálii v provozu, ale jedná se o pravidelné chování, jako například u tratí se sezónním provozem. Podmínku, aby úsek neobsahoval nulovou hodnotu počtu osobních vlaků za měsíc splnilo z celkových 434 pouze 255 úseků. U nákladních vlaků těchto úseků bylo 217.

Osobní doprava

U shlukování měsíčních agregací počtu osobních vlaků v období 2016–2021 bylo nejmenších hodnot vzdáleností docíleno při volbě 6 cílových clusterů. I přesto však byly průměrné vzdálenosti příliš vysoké a nešlo tedy očekávat vysokou kvalitu výsledných shluků. Z výsledných medoidů jednotlivých shluků byly vzory v chování dopravy těžko interpretovatelné, jelikož křivky obsahovaly mnoho náhodných výkyvů a pouze jedna z nich vykazovala nějaký trend počtu osobních vlaků. Jelikož výsledky shlukování nebyly validní a interpretace chování dopravy na úsecích byla velice omezená, nebyly nakonec výsledky uvedeny.

Nákladní doprava

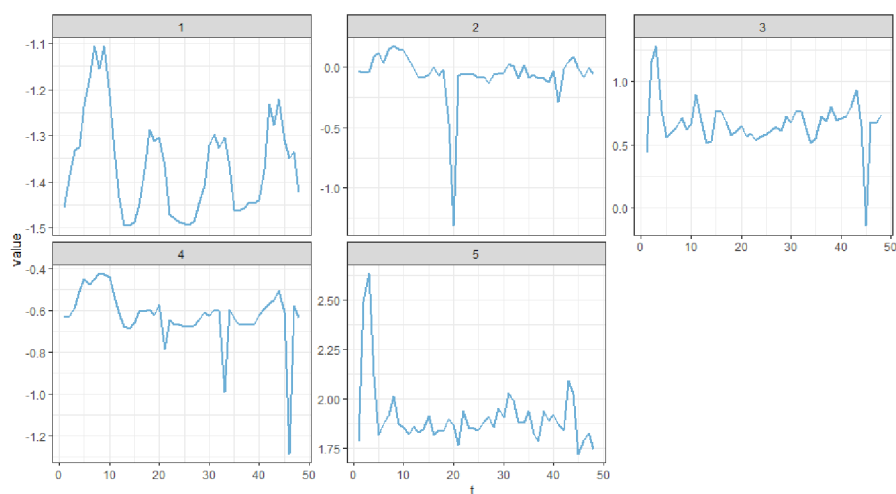
Stejně jako u osobní dopravy, byl počet nákladních vlaků shlukován do 6 cílových shluků. Přesto, že tento počet měl statisticky nejlepší výsledky byly průměrné vzdálenosti uvnitř shluku hodně vysoké. U křivek medoidu byl obdobný problém jako u osobní dopravy. Křivky byly zatížený častými výkyvy a v takto dlouhém časovém horizontu byly vzory v chování těžko interpretovatelné. Výsledky shlukování tedy nebyly kvalitní a podobnost úseků byla velmi nízká. Proto bylo od interpretace výsledků této varianty analýzy upuštěno.

7.3.2 Časové řady počtu vlaků za období 2016–2019

Cílem shlukování časových řad za období 2016–2019, bylo nalezení úseků, které měly podobný průběh provozu, než v ČR zasáhla pandemie covidu-19, který vlakovou dopravu ovlivnil. V tomto případě byly zachovány všechny úseky a nebyly odstraněny úseky s nulovými hodnotami. Při této variantě vstupních dat byly výsledky obou přístupů srovnatelné. Popsány jsou výstupy shlukování s metrikou DTW.

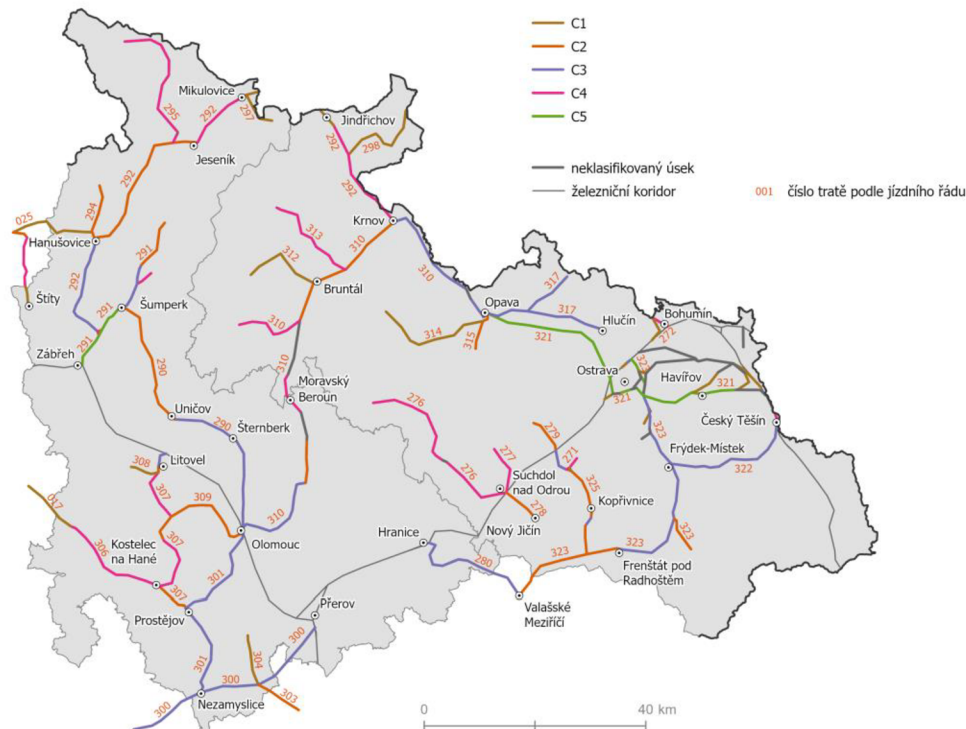
Osobní doprava

Nejlepší statistické hodnoty vyšly při volbě 5 cílových shluků. Na základě vyšší průměrné vzdálenosti byl méně kvalitní shluk C3. Ve výsledných shlucích byly dobře patrné některé specifické vzory v provozu.



Obr. 49 Medoidy shluků průběhu počtu osobních vlaků v letech 2016–2019

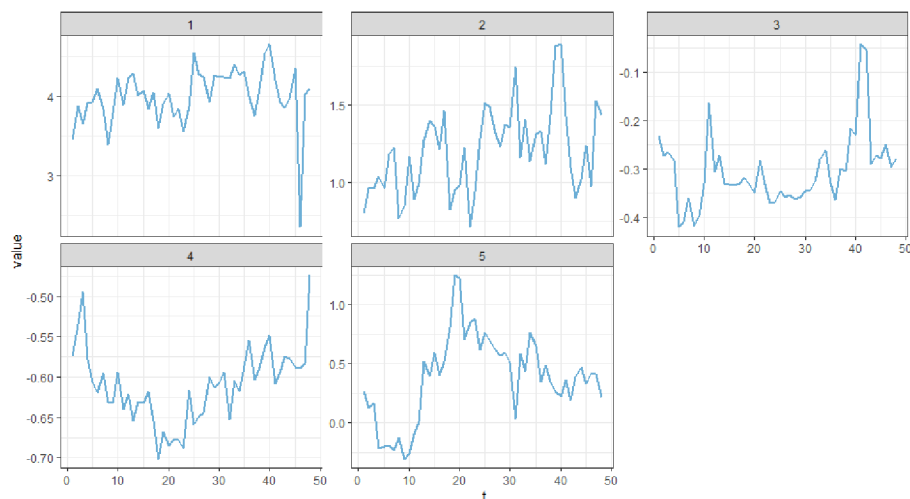
Úseky C1 se vyznačovaly pravidelně se opakujícími nárůsty počtu osobních vlaků. Průběh naznačuje, že se jedná o úseky, kde probíhal sezónní provoz. Podle mapové vizualizace bylo zjevné, že úseky C1 se skutečně nachází na tratích, kde byl podle dřívějších analýz zjištěn sezónní provoz. Jedná se například o trať č. 308 do Mladče nebo č. 312 do Malé Morávky. Na úsecích C2 došlo k poklesům počtu osobních vlaků na podzim roku 2017, jinak se jejich počet příliš nemění. U úseků C3 došlo ke zvýšení počtu vlaků na začátku roku 2016. Naopak propady počtu nastaly na podzim roku 2019. U úseků shluku C4 byl od roku 2017 klesající trend počtu vlaků. V průběhu let docházelo na podzim k pravidelným poklesům. Na počátku roku 2016 došlo na úsecích C5 k vysokému nárůstu, poté se počet opět vrátil na výchozí hodnotu. Takový průběh provozu byl na trati č. 321.



Obr. 50 Výsledky shlukování podle průběhu počtu osobních vlaků v letech 2016–2019

Nákladní doprava

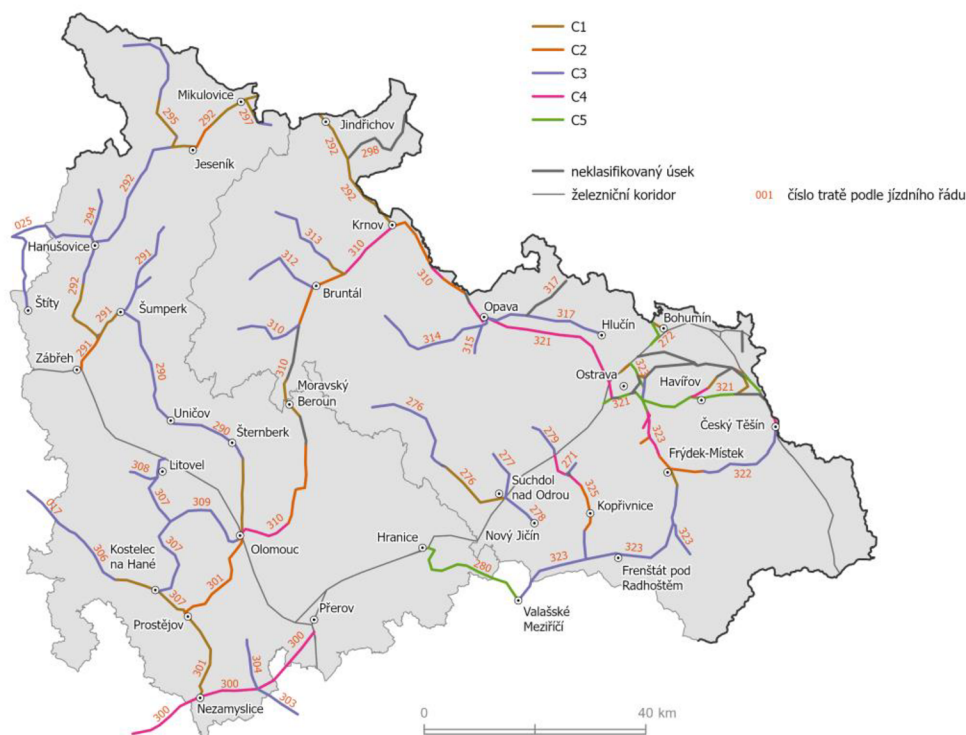
U nákladní dopravy byly časové řady klasifikovány taktéž do 5 clusterů. Uvnitř shluku C2 byla příliš velká průměrná vzdálenost mezi objekty, tudíž byla podobnost přiřazených úseků brána s rezervou.



Obr. 51 Medoidy shluků průběhu počtu nákladních vlaků v letech 2016–2019

Úseky C1 měly rostoucí trend počtu nákladních vlaků. V říjnu 2019 u těchto úseků došlo k propadu nákladní dopravy. Na úsecích C2 počet nákladních vlaků v průběhu let značně kolísá. U úseků C3 se propadla nákladní doprava v roce 2016. Naopak růst

zaznamenala v roce 2019. Na úsecích C4 byl do konce roku 2017 klesající trend počtu nákladních vlaků. Od roku 2018 však počet nákladních vlaků na těchto úsecích rostl. Na přelomu let 2016 a 2017 na úsecích C5 výrazně narostl počet nákladních vlaků. Od září 2017 však počet začal klesat. Úseky C5 se vyskytují v oblasti Ostravy a na trati č. 280 mezi Hranicemi a Valašským Meziříčím.



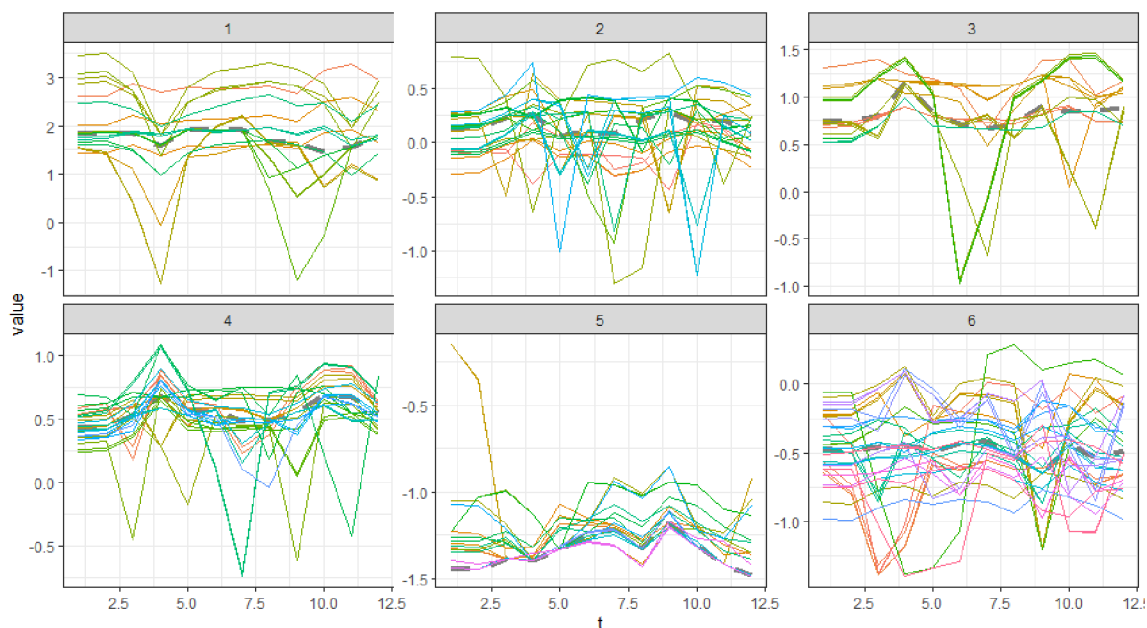
Obr. 52 Výsledky shlukování podle průběhu počtu nákladních vlaků v letech 2016–2019

7.3.3 Časové řady počtu vlaků v roce 2020

Shlukováním časových řad měsíčních agregací počtu vlaků pouze v roce 2020 byla sledována situace a změny v dopravě během prvního roku pandemie covidu-19 v ČR. U této varianty shlukování bylo očekáváno, že dojde především k identifikaci úseků, u kterých došlo v souvislosti s pandemií k výkyvům v dopravě. Při shlukování měsíčních hodnot za období jednoho roku byly lepší výsledky u osobní i nákladní dopravy získávány při využití Euklidovy metriky. Průměrné vzdálenosti uvnitř shluku byly nižší a křivky časových os ve shlucích si byly podobnější. Níže jsou popsány výstupy shlukování časových řad s využitím Euklidovy metriky.

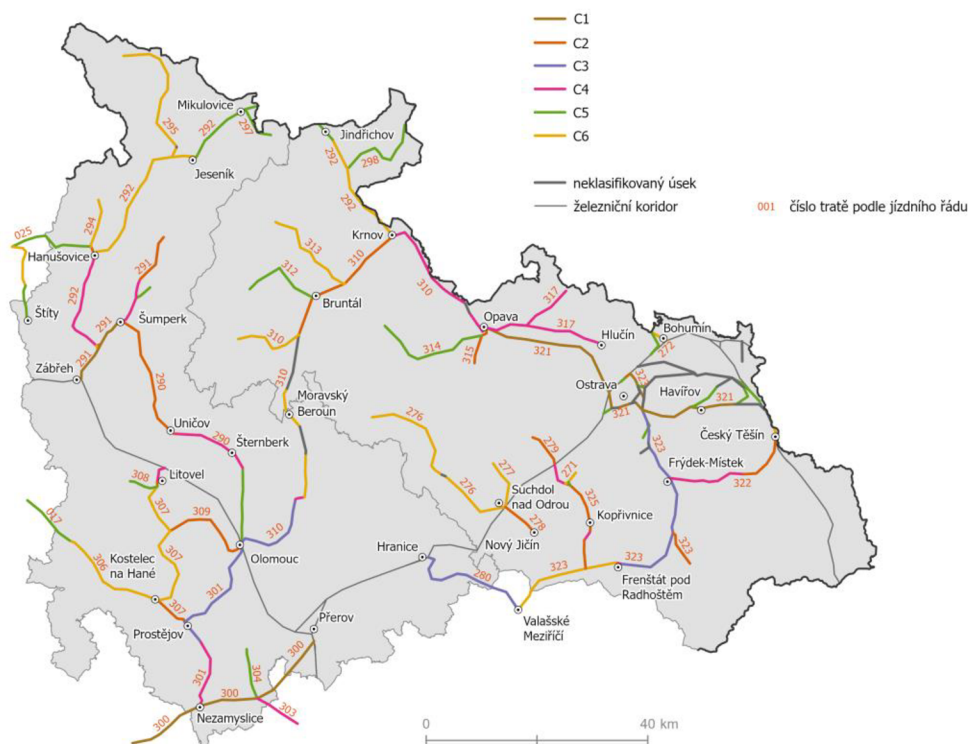
Osobní doprava

U osobní dopravy se v roce 2020 předpokládaly znatelné změny u počtu vlakových souprav, kvůli dopadům pandemie nemoci covid-19 na mobilitu obyvatel. Úseky, obsahující hodnoty počtu vlaků v každém měsíci roku 2020 byly klasifikovány do 6 shluků.



Obr. 53 Shluky podle průběhu počtu osobních vlaků v roce 2020

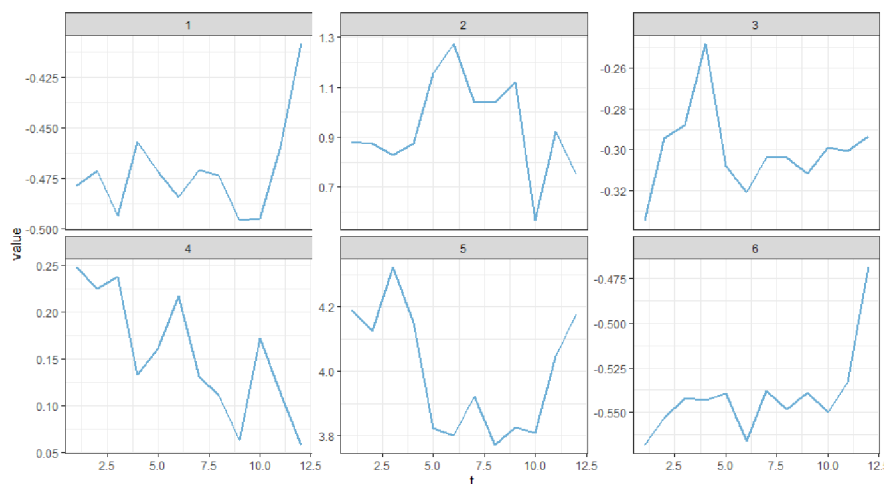
Na úsecích C1 byl zřejmý výrazný propad v období kolem měsíce dubna. To lze považovat za následek protiepidemických opatření. Po nárůstu v letních měsících nastal na podzim opět propad. Podle mapové vizualizace byly takto klasifikovány významné celostátní tratě, které byly v předozích analýzách shledány jako tratě s největším provozem v Olomouckém a Moravskoslezském kraji. Průběhy u úseků C2 byly více heterogenní. Podle meodidu byl počet vlaků nižší od května do konce července, poté následoval růst. V úsecích C3 je na jaře viditelný nárůst, který byl v létě vystřídán propadem. Úseky C4 zaznamenaly v roce 2020 dvě období růstu počtu vlaků. Na rozdíl od úseků C1, tyto úseky v dubnu zaznamenaly nárůst osobní dopravy. Druhé období růstu přišlo od konce prázdnin. Rostoucí trend počtu osobních vlaků byl do září na úsecích C5. Poté se počet snižoval až do konce roku. Jednalo se o úseky na regionálních tratích, které byly dříve označeny jako tratě se specifickým provozem. Podle grafu bylo zřejmé, že úseky C6 jsou z hlediska průběhu provozu celkem různorodé.



Obr. 54 Výsledky shlukování podle průběhu počtu osobních vlaků v roce 2020

Nákladní doprava

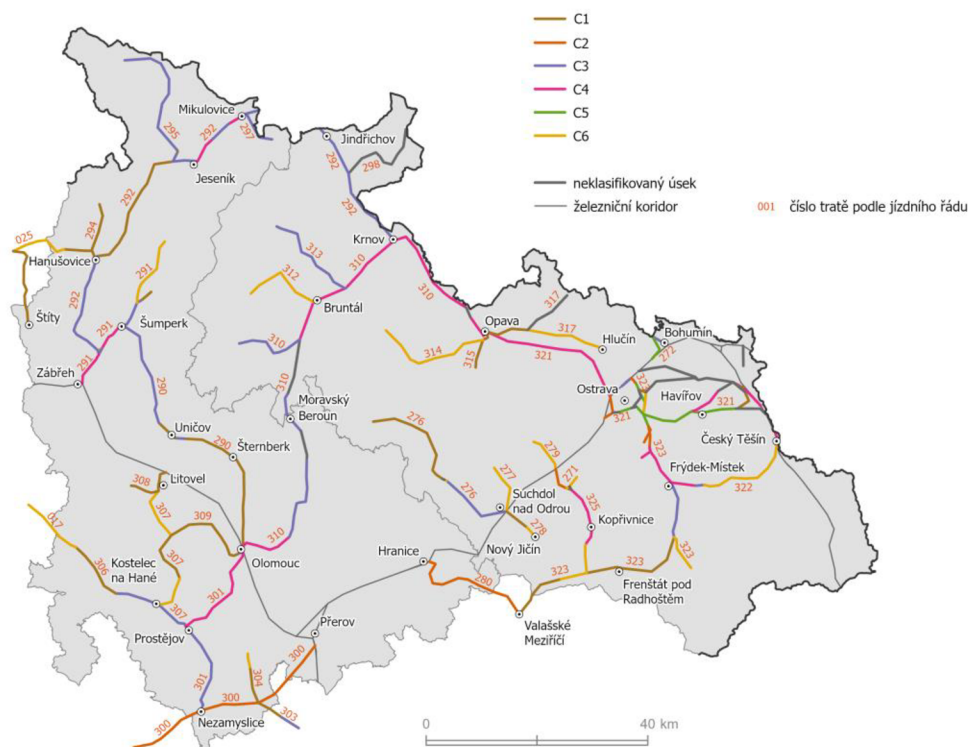
U časových řad nákladní dopravy v roce 2021 mohly být některé úseky podobné kvůli propadu počtu vlaků v březnu a dubnu, kdy byla na mnoha místech omezena průmyslová výroba. Poklesy v letních měsících mohou spojovat úseky, které jsou ovlivněny poklesem výroby a těžby například z důvodu celozávodních dovolených.



Obr. 55 Medoidy shluků průběhu počtu nákladních vlaků v roce 2020

Na úsecích C1 stoupal od listopadu provoz v nákladní dopravě. Shluk C2 klasifikoval tratě s červnovým nárůstem vlaků. Jejich počet od té doby každý měsíc klesal. Jednalo se o trať Prerov–Nezamyslice nebo Hranice–Valašské Meziříčí. Výraznější propad nastal

v listopadu. Na úsecích C3 jezdil vyšší počet nákladních vlaků na jaře než ve zbytku roku. Klesající trend počtu nákladních vlaků měly od začátku roku 2020 úseky C4. Na tratích C5 došlo od března k poklesu v provozu v nákladní dopravě. Počet vlaků začal růst až ke konci roku. Takovýto provoz byl na tratích v Bohumíně a na trati č. 321 v úseku u Ostravy a Havířova. Rostoucí trend počtu nákladních vlaků mají tratě C6. Největší růst zaznamenaly v listopadu a prosinci.



Obr. 56 Výsledky shlukování podle průběhu počtu nákladních vlaků v roce 2020

7.3.4 Časové řady počtu vlaků v roce 2021

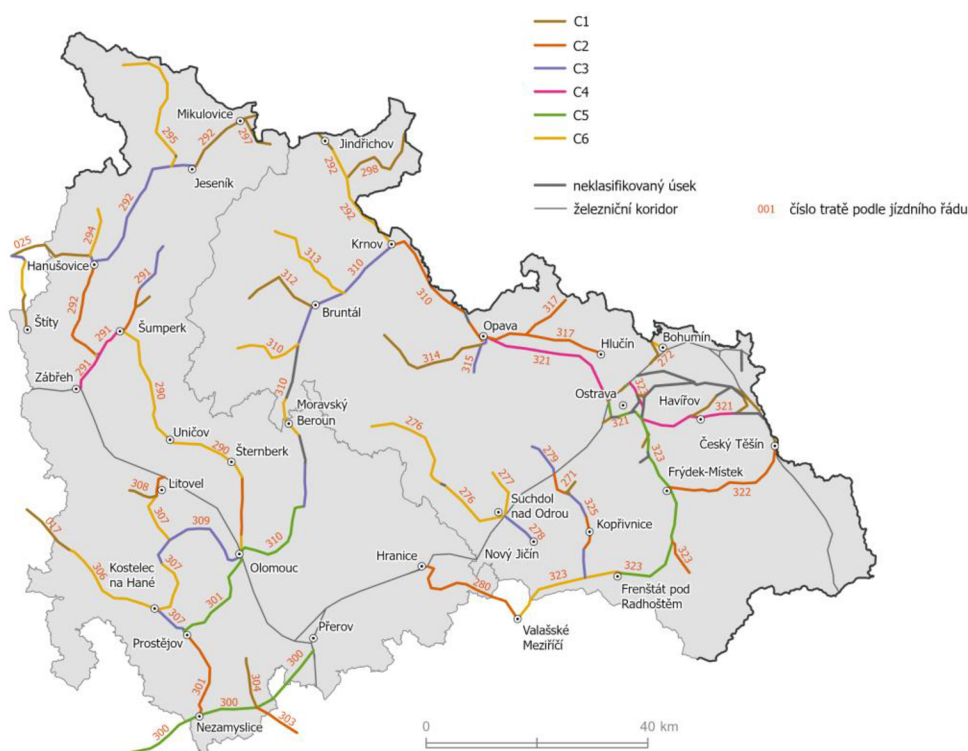
Při shlukování časových řad měsíčních agregací počtu vlaků pouze v roce 2021 byla sledována situace během druhého roku, kdy na území ČR probíhala pandemie covid-19. V roce 2021 se již neočekávaly tak tvrdé následky protiepidemických opatření na vlakovou dopravu. Stejně jako časových řad za rok 2020, byly relevantnější výsledky získány s využitím Euklidovy metriky.

Osobní doprava

U časových řad osobní dopravy se mohly projevit propady v počtu vlaků na jaře, kdy byl zaveden tvrdý lockdown a byla omezena mobilita obyvatel mezi okresy. Naopak nárůsty mohly přijít v letních měsících, kdy došlo zvolnění restrikcí.

Úseky C1 měly zvýšený provoz osobní dopravy v letních měsících, nežli ve zbytku roku. Tento provoz byl typický na tratích s minimální osobní dopravou. Vyšší počet osobních vlaků jezdil na jaře na tratích C2. Pokles vlaků v červnu a červenci zaznamenaly tratě C3. V zimě byly na těchto tratích provoz slabý. Mírný dubnový nárůst, byl výkyv v jinak klesajícím trendu počtu osobních vlaků na tratích C4. Jejich počet se začal od října zvyšovat. V dubnu počet vlaků poklesl také na tratích C5. Počet

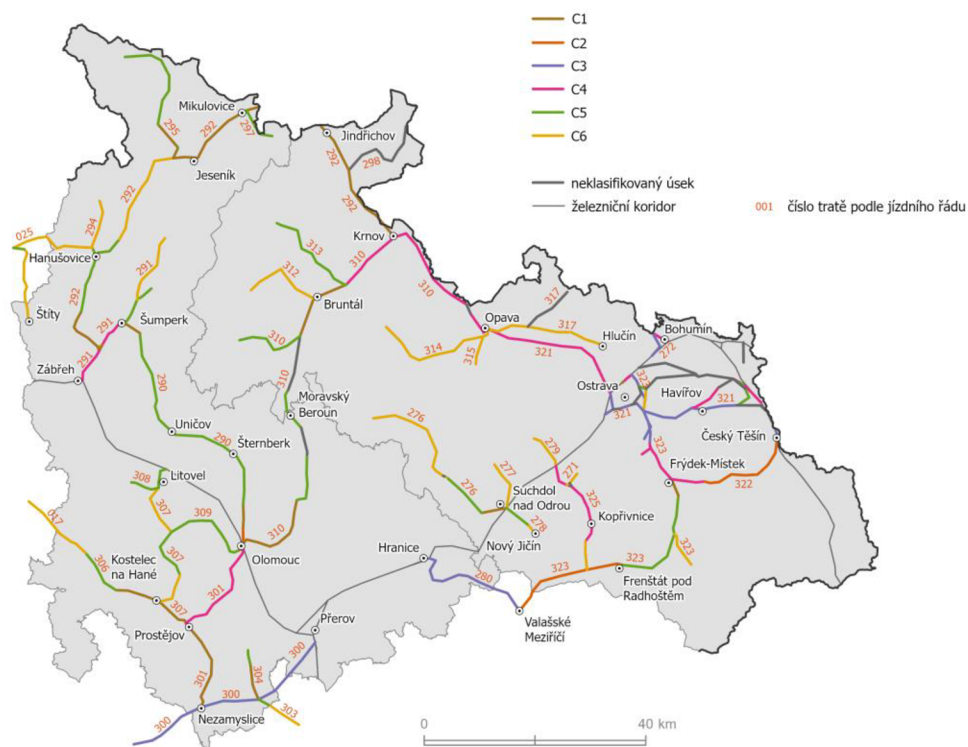
vlaků se od začátku roku až do srpna zvyšoval na tratích C6. Poté začala mít osobní doprava klesající tendenci. Takový průběh byl na vybraných regionálních tratích.



Obr. 57 Výsledky shlukování podle průběhu počtu osobních vlaků v roce 2021

Nákladní doprava

Na úsecích C1 byl zaznamenán zvýšený počet průjezdů nákladních vlaků v létě o prázdninách. Trati č. 322 z Frýdku-Místku do Českého Těšína a 323 z Valašského Meziříčí do Frenštátu pod Radhoštěm měli v březnu 2021 skokový nárůst počtu nákladních souprav. Trati C3 zaznamenaly propady v nákladní dopravě v červnu a září. Na tratích C4 byl v roce 2021 rostoucí provoz nákladní dopravy, který byl poznamenán propady mezi dubnem a květnem. Do tratí C4 spadá trať č. 291 Šumperk–Zábřeh, 301 Olomouc–Zábřeh a 310 z Ostravy do Krnova. Jde o jedny z celostátních tratí s nejvyšším provozem v regionu. V provozu úseků C5 byl odhalen nižší počet nákladních vlaků na jaře. Pro úseky C6 byl signifikantní zvýšený počet vlaků mezi červencem a zářím. Následující měsíce jezdil menší počet nákladních vlaků.



Obr. 58 Výsledky shlukování podle průběhu počtu nákladních vlaků v roce 2021

8 VÝSLEDKY

Hlavním cílem práce bylo provést klasifikaci a časovou analýzu provozu na vybraných celostátních a regionálních železničních tratích Olomouckého a Moravskoslezského kraje. Práce byla navržena a zpracována tak, že bylo docíleno několika dílčích výsledků namísto jednoho komplexního. Výsledky lze rozdělit do čtyř částí podle toho, v jakém pořadí byly zpracovány.

Příprava dat

I ve fázi přípravy dat došlo k získání několika výsledků a výstupů. Z poskytnutých dat, která byla ve formě tabulek, byly vytvořeny prostorové vrstvy, které sloužily k vizualizaci hodnot intenzity provozu a výsledků provedených analýz. Jednalo se o bodové a liniové vrstvy železniční infrastruktury Olomouckého a Moravskoslezského kraje. Bodové vrstvy obsahují lokalizované monitorovací body SR70 na celostátních a regionálních tratích za každý dílčí rok období 2016–2021, a body SR70, které mají záznamy ve všech letech současně. Dvě liniové vrstvy obsahují železniční úseky regionálních a celostátních tratí v roce 2016 a 2021. Každému úseku byl doplněn atribut čísla tratě podle jízdních řádů. Tím byla vytvořena železniční síť, složená z dílčích traťových úseků. Postup zpracování a úpravy primárních dat, byl podrobně rozepsán a může sloužit jako návod. Pomocí kontingenčních tabulek byly z primárních dat vytvořeny agregace měsíčních a ročních hodnot počtu vlaků a hmotnosti souprav v úsecích a monitorovacích bodech SR70. Nakonec byly hodnoty měsíčních agregací očištěny od kalendářních vlivů podle vzorce, který byl navržen pro použití na víceletá data.

Struktura železniční dopravy

V této části práce byly vytvořeny mapové vizualizace, které sloužily k prvotnímu seznámení s rozložením a strukturou železniční dopravy v Olomouckém a Moravskoslezském kraji. Celkem bylo vytvořeno 10 mapových výstupů, které porovnávají hodnoty provozu z let 2016, 2020 a 2021. Podle celkového počtu a hmotnosti vlakových souprav za rok 2016 a 2021 byl největší provoz zjištěn na tratích v okolí Ostravy. Na vybraných úsecích tratě č. 321 projelo více než 40 tisíc vlakových souprav za rok. Celkový počet tvořily spíše osobní vlaky, avšak hmotností souprav převyšovala doprava nákladní. Největší provoz byl v obou letech na trati Ostrava-Stodolní – Ostrava-Střed. Další tratě s vysokým provozem byly Šumperk–Zábřeh, Přerov–Nezamyslice, Olomouc–Prostějov, Opava–Ostrava, Hranice–Valašské Meziříčí. Naopak nejmenší provoz byl na regionálních tratích. Na drtivé většině všech tratí převažovala osobní doprava. Odhaleny byly také tratě, na kterých probíhala pouze osobní nebo nákladní doprava. Porovnáním hodnot počtu vlaků v období od března do května roku 2020 a 2021 vůči průměru ze stejného období předchozích let, bylo zjištěno, na kterých tratích došlo v období pandemie covidu-19 k poklesu nebo nárůstu počtu vypravených vlaků. Největší poklesy osobních vlaků o více než 50 % zaznamenaly v roce 2020 úseky na hraničních přechodech. Důvodem bylo na určitou dobu uzavření státních hranic pro osobní dopravu. Počty vlaků klesly především na regionálních tratích. Poklesy však zaznamenal i provoz na celostátních tratích. Na některých úsecích došlo naopak k vyššímu počtu průjezdů vlaků. Propady byly zaznamenány i u nákladní dopravy. V roce 2021 se situace v osobní dopravě zlepšila především na hraničních přechodech. Celkově byl počet vypravených vlaků v březnu–květnu 2020 a 2021 nižší než v předchozích letech.

Časová analýza

Navrženy byly metody pro odhalení výkyvů a výluk v provozu. První metoda spočívala v identifikaci nulových hodnot vybraného ukazatele v datovém souboru s využitím nástrojů MS Excel. Tato metoda je jednoduchá a vhodná pro hledání úplných výluk v provozu, kdy po dobu alespoň jednoho měsíce neprojde žádný vlak. Současně může posloužit při hledání úseků se sezónním provozem. Pro odhalení výkyvů a částečných výluk v provozu byly navrženy dva automatizované způsoby identifikace odlehklých hodnot. Oba způsoby však v této práci nenašly velké uplatnění. Řešení v programu Orange bylo vyhodnoceno jako nevhodné pro objemná data tohoto formátu. Výsledky R skriptu neobsahovaly důležitou časovou složku pro potřeby časové analýzy. Dále byly popsány možnosti analyzování časových řad pomocí sledování meziročního tempa růstu a dekompozice časových řad.

Hlavními výstupy časové analýzy byly samotné časové řady. Vyhodnocením časových řad měsíčních agregací byl popsán provoz na vybraných úsecích od roku 2016 do roku 2021. Interpretován byl vývoj osobní i nákladní dopravy podle počtu a hmotnosti souprav. Pomocí časových řad byly odhaleny výkyvy v provozu a také především vliv pandemie na osobní a nákladní dopravu v letech 2020 a 2021. Na úsecích celostátních tratí byl obecně velký počet osobních vlakových souprav a provoz byl zatížen menšími výkyvy. Situace na jaře 2020 byla individuální. U některých úseků došlo k poklesu až o 500 nákladních vlaků za měsíc, na jiných nebyl provoz ovlivněn. Na úsecích celostátních tratí probíhá pravidelná nákladní doprava s velkým významem. Pandemie nejvíce postihla provoz na regionálních tratích, kdy v posledních dvou letech na těchto tratích docházelo k častým a výrazným propadům v osobní dopravě. Provoz nákladní dopravy je na těchto tratích relativně nízký. Zajímavý vývoj nákladní dopravy byl odhalen na trati z Rýmařova. V letech 2017 a 2018 se zde výrazně zvýšila hmotnost nákladních souprav i přesto, že jejich počet byl zhruba stejný jako předchozí a následující roky. Nárůst hmotnosti byl připisován zvýšené těžbě dřeva kvůli kůrovcové kalamitě. Na úsecích v blízkosti průmyslových areálů byla nákladní doprava ovlivněna především na jaře 2020. Nejvíce byl zasažen úsek v blízkosti automobilky Hyundai Nošovice. Tyto zjištění potvrdily tezi o utlumení průmyslové výroby na počátku šíření covidu-19 v ČR. V roce 2021 již nebyly ovlivněny tak výrazně. Minimální dopady pandemie byly na tratě se specifickým provozem. Naopak výrazně byla ovlivněna doprava na hraničních přechodech. Na základě rozhodnutí o uzavření státních hranic neprojel hraničními přechody v dubnu a květnu 2020 žádný osobní vlak. Stanicemi projížděly pouze nákladní soupravy. Celkově byl provoz vyšší na přechodech Bohumín a Český Těšín, a to díky mezinárodní dopravě. Na přechodech Mikulovice a Jindřichov probíhá pouze vnitrostátní doprava přes území Polska. Od roku 2020 ve stanicích Bohumín a Český Těšín trend počtu osobních vlaků klesal. Naopak počet nákladních vlaků roste, což může indikovat zvýšený význam mezinárodní nákladní dopravy.

Klasifikace tratí

Za první výsledky klasifikace tratí je možné považovat výstupy v kapitole *Struktura železniční dopravy*. Úseky a monitorovací body byly manuálně kategorizovány podle intenzity provozu. Cílem klasifikace bylo nalezení tratí s podobným a zároveň specificky odlišným provozem. Hlavními výsledky této části byly klasifikované železniční úseky podle měsíčních a ročních hodnot počtu a hmotnosti osobních a nákladních vlaků. Vyzkoušeny byly dva přístupy zpracování využitím hierarchického shlukování a shlukováním metodou k-means. Pro obě metody byly sestaveny modely v programu Orange. Experimentováno bylo s různými kombinacemi metod, metrik a cílových počtů

shluků. U hierarchického shlukování se nejvíce osvědčila Wardova metoda s využitím Mahalanobisovy nebo Euklidovy vzdálenosti. Pro interpretaci výsledných shluků byly sestaveny kategorizační tabulky. Výsledky byly prezentovány na několika mapových náhledech. Specificky odlišný provoz byl odhalen na úsecích Bohumín-Vrbice st. hr. – Bohumín-Vrbice S618a a Bohumín-Vrbice – Bohumín-Vrbice S618. Tyto dva úseky měly tendenci se shlukovat do samostatného shluku. Provoz zde byl specifický velmi nízkým podílem osobní dopravy, ale velmi vysokým počtem a hmotností nákladních vlaků. Od ostatních se také odlišoval provoz na trati č. 321 v úseku Ostrava–Havířov. Úseky podobné nízkým provozem byly především na regionálních tratích. Na tratích se sezonním provozem nebo pouze nákladní dopravou byl minimální provoz. Rozdíly ve výsledcích hierarchického a k-means shlukování byly zanedbatelné. Kvůli větším možnostem a variantám shlukování, byl za vhodnější přístup vyhodnocena hierarchická metoda.

Nakonec byly tratě klasifikovány podle podobnosti průběhu časové osy. Vyzkoušeno bylo shlukování s využitím metody dynamického borcení časové osy. Bylo navrženo několik variant vstupních dat, vždy se však jednalo o měsíční agregace počtu osobních nebo nákladních vlaků. Funkce shlukování byla sepsána do programového kódu v jazyce R. Výsledky shlukování časových os s dlouhodobým časovým rozmezím nebyly relevantní kvůli vysokým průměrným vzdálenostem mezi objekty shluku a složité interpretaci. Klasifikovány byly tratě s podobným průběhem provozu v letech 2016–2019, tedy v době před pandemií covidu, v roce 2020 a poté 2021. Pro shlukování hodnot za jeden rok bylo jako lepší řešení vyhodnoceno použití Euklidovy metriky. Výsledky byly uspokojivé a lze v nich najít podobu s výsledky předchozích analýz.

9 DISKUZE

Diplomová práce byla založena na analýze dat provozu na železnici, které obsahují podrobné údaje o intenzitě železniční dopravy. Vlastníkem dat je Správa železnic, která je využívá pro výpočet poplatku za využití železniční dopravní cesty pro železniční dopravce. Tato data jsou za normálních okolností veřejně nedostupná. Na Katedru geoinformatiky v Olomouci byla poskytnuta pro akademické účely. Jedná se tedy o specifická tajná data, jejichž další šíření je zakázáno. Proto nejsou vstupní a výstupní data součástí příloh práce. Zpracováním těchto dat se na katedře dříve zabýval Jan Cinke (2020). Ten ve své práci vyhodnocoval data provozu za období 2016–2019 na celém území České republiky. Výhodou a rozdílem této diplomové práce je zaměření na menší územní celek, což umožňuje detailnější analýzu tratí nežli při zpracování v celostátním měřítku. Současně jsou k dispozici data za rok 2020 a 2021, kdy vypukla a probíhala na území ČR pandemie virové choroby covid-19, tudíž lze vyhodnotit její dopady na osobní a nákladní vlakovou dopravu v Olomouckém a Moravskoslezském kraji. Data byla poskytnuta v měsíčních agregacích. Správa železnic má k dispozici i denní agregace, ale ty by však kvůli velkému objemu dat nešlo v rámci této práce zpracovat. Data v původní podobě nešlo okamžitě použít a analyzovat. Musela být nejprve pomocí kontingenčních tabulek sumarizována a filtrována podle druhu dopravy. Jelikož atributová složka neobsahovala žádný atribut, podle kterého by se daly objekty lokalizovat, musely být úseky a monitorovací body SR70 v Olomouckém a Moravskoslezském kraji vybrány manuálně. Současně musely být podle polohy odděleny úseky a SR70 z koridorů. Nakonec byly hodnoty počtu a hmotnosti osobních a nákladních vlaků očištěny od kalendářních vlivů. Toto předzpracování dat bylo časově náročné. Postup úpravy dat do požadované formy je podrobně popsán krok po kroku, a může sloužit dalším řešitelům. Přínosem práce, je možnost se díky podrobným datům podívat na konkrétní úseky železniční sítě a vyhodnotit jejich provoz. Data o dopravě z Českého statistického úřadu jsou agregovaná a nepodávají tak detailní vhled do železniční dopravy.

Přesto, že je práce zaměřena na železniční dopravu pouze ve dvou krajích ČR, aby byly výsledky detailnější nežli v měřítku celé České republiky, nebylo možné vytvořit časové řady pro všechny úseky. Bylo jich stále více než 400. Proto bylo navrženo řešení, kdy bylo definováno 5 kategorií tak, aby reprezentovaly strukturu a rozložení železniční dopravy v Olomouckém a Moravskoslezském kraji. Do každé kategorie bylo přiřazeno několik vybraných úseků z celého regionu, které danou kategorii charakterizují.

Na časových řadách bylo viditelné, že na většině úseků docházelo v průběhu let k častým výkyvům, především propadům. Ty mohly být způsobeny například výlukami v dopravě či jinými mimořádnostmi. Tyto výluky však nelze všechny dohledat a ověřit. Propady v letech 2020 a 2021 jsou nejčastěji přisuzovány dopadům pandemie covidu-19. Ve většině případů však není ani možné přesně určit, proč ke změně došlo, a jedná se pouze o domněnky. Například odůvodnění, že zvýšení váhy nákladních souprav ve stanici Rýmařov souvisí s těžbou kůrovcového dřeva není podloženo reálnými daty, ale jedná se pouze o hypotézu. U některých případů lze určit, proč došlo k poklesu v provozu. Například nákladní doprava v úseku v blízkosti automobilky v Nošovicích, která je dominantním uživatelem, je závislá na podnikové výrobě. V době, kdy došlo k pozastavení výroby, došlo okamžitě k poklesu počtu nákladních vlaků. Cílem práce však není sledovat a zjišťovat příčiny změn v detailu.

Pro klasifikaci tratí byly vybrány standardní metody shlukování. Jejich výsledky byly víceméně totožné, a těžko zhodnotit, která metoda byla lepší. Realizace k-means metody bylo jednoduché, naopak hierarchická metoda umožňuje nastavení více parametrů a bylo možné více experimentovat jak už se vzdálenostní metrikou nebo počtem shluků. Shluky byly interpretovány na základě hodnoty mediánu. Ta na rozdíl od průměru není citlivá na odlehlé hodnoty, které hodnotu výrazně průměru ovlivňují. Interpretací podle průměru by mohlo dojít k mylné kategorizaci výsledného shluku. Jeho průměrná hodnota by mohla spadat do kategorie s vysokým počtem vlaků, avšak reálné hodnoty úseků by byly nízké.

Metodu DTW je možné použít pro shlukové analýzy časových řad v jazyce R a Python. V této práci byl využit programový kód v jazyce R. Naprogramováno bylo shlukování metodou nejbližších středů, ale DTW je možné využít i při hierarchickém shlukování. Časová náročnost výpočtu při klasifikaci do 5 až 6 shluků nebyla delší než minutu. Cílem shlukování časových řad bylo hledání úseků s podobným průběhem provozu. Při shlukování časových řad měsíčních hodnot po dobu 6 let byly výsledky neuspokojivé. Statisticky byly průměrné vzdálenosti mezi objekty shluku příliš vysoké, tudíž podoba časových řad malá, a to jak u použití DTW, tak při volbě Euklidovy metriky. U časových řad železniční dopravy, které jsou zatíženy častými a náhodnými výkyvy, se v takto dlouhém časovém horizontu hledá podobnost těžko. Proto bylo hledání úseků s podobným průběhem provozu v takto dlouhém období nemožné. Řešením by bylo zvýšení počtu cílových shluků, ale při velkém počtu shluků by nastal problém v jejich obtížné interpretaci. Obecně čím více cílových shluků, tím větší podobnost objektů uvnitř shluku. Při shlukování časových řad kratších délek již byly výsledky uspokojivé u metody DTW i Euklidovy vzdálenosti. Při shlukování hodnot za jeden rok se více osvědčilo využití Euklidovy vzdálenosti. Otázkou je, jestli bylo použití metody DTW potřeba, jelikož řady vůči sobě nebyly posunuty v čase. Pro data provozu na železnici je dostačující využití Euklidovy vzdálenosti. Hlavní překážkou shlukování nebyla ani tak délka časových řad, jako jejich vysoká náhodná složka. Nevýhodou metody je také interpretace výsledků. Ta je založena na manuálním vyhodnocení výsledných shluků a je tedy subjektivní, a při takto velkém objemu analyzovaných časových řad složitá a časově náročná. Největší podobnosti tratí by bylo dosaženo při použití velkého počtu cílových shluků. Při shlukování menšího počtu časových řad kratšího časového období by však tato metoda mohla být efektivní.

ZÁVĚR

Hlavním cílem diplomové práce bylo na základě dat provozu na železnici provést klasifikaci a časovou analýzu provozu na regionálních a celostátních tratích Olomouckého a Moravskoslezského kraje. Práce byla rozdělena do několika částí, kdy nejdůležitější poznatky přinesla časová analýza a klasifikace tratí. Výsledky byly často rozlišovány podle druhu dopravy a roku. Všechny dosažené výsledky v průběhu zpracování jsou uvedeny v textu práce.

V teoretické části byla popsána struktura poskytovaných dat, použité metody zpracování a vybraný software pro realizaci diplomové práce. Charakterizována je železniční infrastruktura v celorepublikovém měřítku, včetně platné kategorizace tratí podle Ministerstva dopravy. Uvedeny jsou také studie a odborné publikace, které se věnují zpracování dat provozu na železnici nebo tématu železniční dopravy. Nastíněn byl úvod do problematiky časových řad a shlukové analýzy.

Pro potřeby navržených analýz bylo nutné provést předzpracování dat. Z primárních dat byly vybrány pouze údaje o železnici v Olomouckém a Moravskoslezském kraji, a popsán postup tvorby prostorových dat, který zahrnuje kategorizaci celostátních a regionálních tratí. Dílčím úsekům tratí byl doplněn atribut čísla tratě podle jízdního řádu. Takto byla vytvořena prostorová data železniční sítě v regionu. Hodnoty vybraných atributů byly očištěny od kalendářních vlivů, aby nebyly ovlivněny křivky časových řad. Na vytvořených vrstvách železničních úseků a monitorovacích bodů SR70 byly vizualizovány naměřené hodnoty provozu a výsledky analýz. Výstupy rozšířily představu o struktuře a rozložení železniční dopravy v regionu.

Před zpracováním časové analýzy byly navrženy metody, kterými lze odhalit výkyvy v provozu nebo identifikovat tratě se sezónním provozem. Kvůli velkému počtu úseků bylo definováno 5 kategorií, do kterých byly přiřazeny reprezentativní úseky. Z časových řad byl vyhodnocen provoz osobní i nákladní železniční dopravy v letech 2016 až 2021. Na celostátních tratích došlo k hromadným propadům osobní i nákladní dopravy v roce 2020 kvůli pandemii covidu-19. Podle počtu nákladních vlaků však na těchto tratích od roku 2020 roste význam nákladní dopravy. Na regionálních tratích v průběhu let docházelo k častým výkyvům. Nejhorší situace nastala v době pandemie covidu-19, kdy v letech 2020 a 2021 byly obrovské propady počtu osobních vlaků. Počet nákladních vlaků na těchto tratích se v dlouhodobém horizontu snižuje.

Posledním bodem práce bylo klasifikování tratí podle podobného provozu. Výchozí metodou klasifikace bylo hierarchické shlukování. Výsledky byly porovnány s výsledky metody k-means. Tím byly klasifikovány úseky podle podobného měsíčního nebo ročního provozu. Úseky byly klasifikovány také podle podobného časového průběhu provozu pomocí shlukování podle časových řad s využitím metody dynamického borcení časové osy. Shlukování bylo realizováno pomocí naprogramovaného skriptu v jazyce R.

Hlavní cíle práce byly splněny. Provoz osobní a nákladní dopravy v regionu byl časově analyzován. Odhaleny byly výkyvy v provozu a také dopady pandemie na železniční dopravu. Celostátní a regionální tratě byly klasifikovány podle provozu osobní a nákladní dopravy. Identifikovány byly tratě s podobným a specificky odlišným provozem mezi lety 2016 a 2021.

POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

BESTA, Petr. Porovnání jednotlivých druhů dopravy. *Techportal* [online]. Praha, ©2022, 5 [cit. 2022-08-13]. Dostupné z: https://www.techportal.cz/download/enoviny/enlog/porovnaní_jednotlivých_druhu_dopravy.pdf

CINKE, Jan. *Analýza dat vytiženosti železniční infrastruktury*. Olomouc, 2020. Diplomová práce. Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra geoinformatiky

ČERNOHOUS, Lukáš. *Metoda dynamického borcení času* [online]. Brno, 2016 [cit. 2022-08-13]. Dostupné z: <https://dspace.vutbr.cz/xmlui/handle/11012/60494?locale-attribute=pl>. Bakalářská práce. VUT v Brně, Fakulta elektrotechniky a komunikačních technologií.

DOBEŠOVÁ, Z. *ORANGE, Praktický návod do cvičení předmětu Data Mining: učební text* [online]. Olomouc: Univerzita Palackého v Olomouci, 2022 [cit. 2022-08-13]. ISBN 978-80-244-6086-4. DOI: 10.5507/prf.22.2440864 Dostupné z: <http://dobesova.upol.cz/Orange/OrangeManual.pdf>

DOBEŠOVÁ Z., KUČERA M. Time series of workload on railway routes, In Silhavy, R. (ed.) CSOC 2019. Artificial Intelligence and Algorithms in Intelligent Systems, vol. 985, pp. 370-380, 2020, Springer, Cham. DOI: 10.1007/978-3-030-19810-7_37

HANČLOVÁ, Jana a Lubor TVRDÝ. *Úvod do analýzy časových řad* [online]. Ostrava, 2003. VŠB-TU Ostrava, Ekonomická fakulta. Dostupné z: https://www.fd.cvut.cz/departament/k611/PEDAGOG/VSM/7_AnalyzaCasRad.pdf

HASPRA, Daniel. Dopad pandemie covid-19 na železniční sektor. *Vědeckotechnický sborník Správy železnic* [online]. Praha: Správa železnic, 2021, **2021**(4), 2–5 [cit. 2022-08-13]. Dostupné z: <https://www.spravazeleznic.cz/documents/50004227/135082601/Dopad+pandemie+covid-19+na+zeleznicni+sektor.pdf/a017fff6-ac99-489e-a8eb-59fb74e133e4>

HRUŠKOVÁ, Jitka. Železniční dopravu omezí výluka na trati mezi Ostravou a Opavou. *Moravskoslezský deník* [online]. Ostrava: Vltava Labe Media, 2020 [cit. 2022-08-13]. Dostupné z: https://moravskoslezsky.denik.cz/zpravy_region/zeleznicni-dopravu-omezi-vyluka-na-trati-mezi-ostrovou-a-opavou-20200817.html

KRÝŽE, Pavel. Mapa uvedená v knižním jízdním řádu. In: *Správa železnic: Portál provozování dráhy* [online]. Praha: Správa železnic, 2021b [cit. 2022-08-13]. Dostupné z: <https://provoz.spravazeleznic.cz/Portal/ViewArticle.aspx?oid=594598>

KŘIVÝ, Ivan, 2012. Analýza časových řad [online]. Ostrava: Ostravská univerzita v Ostravě [cit. 2022-08-13]. Dostupné z: <http://www1.osu.cz/~bujok/files/ancas.pdf>

KUČERA, Michal a Zdena DOBEŠOVÁ. Železniční doprava v roce 2020. In: *Symposium GIS Ostrava 2022: Smart City – vize a realita* [online]. 2022, - [cit. 2022-08-13]. ISBN 978-80-248-4607-1. ISSN 1213-239X. Dostupné z: doi.org/10.31490/9788024846071-161

KUČERA, Jiří. *Metody kategorizace dat* [online]. Brno, 2008 [cit. 2022-08-13]. Dostupné z: <https://is.muni.cz/th/w8lgz/>. Bakalářská práce. Masarykova univerzita, Fakulta informatiky.

PETR, Pavel. *Metody Data Miningu* [online]. Pardubice: Univerzita Pardubice, 2014 [cit. 2022-08-13]. ISBN 978-80-7395-872-5

PROCHÁZKA, Jan. *Využití časových řad v analýze dat z eye tracking systému* [online]. Olomouc, 2013 [cit. 2022-08-13]. Dostupné z: <https://dspace.vutbr.cz/xmlui/handle/11012/60494?locale-attribute=pl>. Diplomová práce. Univerzita Palackého v Olomouci, Přírodovědecká fakulta.

SENIN, Pavel. *Dynamic Time Warping Algorithm Review* [online]. Honolulu: University of Hawaii at Manoa, 2008 [cit. 2022-08-14]. Dostupné z: https://www.researchgate.net/publication/228785661_Dynamic_Time_Warping_Algorithm_Review

SOBOLEWSKA, Ewa. Dynamic Time Warping (DTW) as a mean to cluster time series. *RPubs* [online]. 2019 [cit. 2022-08-13]. Dostupné z: <https://rpubs.com/esobolewska/dtw-time-series>

SVOBODA, Jiří. SŽ SR70: Číselník železničních stanic a dopravně významných míst. In: *Portál provozování dráhy: Popis sítě: Číselníky a změny názvů: Číselník SR 70* [online]. Praha: Správa železnic, 2021a [cit. 2022-08-13]. Dostupné z: <https://provoz.spravazeznic.cz/portal/ViewArticle.aspx?oid=34462>

SVOBODA, Jiří. SŽ SR72: Číselník železničních drah. In: *Portál provozování dráhy: Popis sítě: Číselníky a změny názvů: Číselník SR 72* [online]. Praha: Správa železnic, 2021b [cit. 2022-08-13]. Dostupné z: <https://provoz.spravazeznic.cz/portal/ViewArticle.aspx?oid=130812>

ŠARMANOVÁ, Jana. *Metody analýzy dat: učební text* [online]. Ostrava: VŠB-TU Ostrava, 2012 [cit. 2022-08-13]. ISBN 978-80-248-2565-6. Dostupné z: <http://www.person.vsb.cz/archivcd/FEI/MAD/MAD.pdf>

Automotive. *ČD Cargo: Co nabízíme: Přeprava komodit* [online]. ČD Cargo, 2022 [cit. 2022-08-13]. Dostupné z: https://www.cdcargo.cz/cs_CZ/automotive?inheritRedirect=true

Cestující ve vlacích ČD dál přibývají, loni jich bylo přes 182 milionů. *České Dráhy: Tiskové centrum: Tiskové zprávy* [online]. Praha: České dráhy, 30. 1. 2020 [cit. 2022-08-13]. Dostupné z: <http://www.ceskedrahy.cz/tiskove-centrum/tiskove-zpravy/-30954/>

COVID-19: úvod, inkubační doba, původce a sezónnost onemocnění. *NZIP* [online]. Praha: Národní zdravotnický informační portál, 2021 [cit. 2022-08-13]. Dostupné z: <https://www.nzip.cz/clanek/447-covid-19-zakladni-informace>

ČD postupně obnovují provoz většiny omezených regionálních linek. *České dráhy: Tiskové centrum: Tiskové zprávy* [online]. Praha: České dráhy, 14. 6. 2021 [cit. 2022-08-13]. Dostupné z: <http://www.ceskedrahy.cz/tiskove-centrum/tiskove-zpravy/-31627/>

Indexy tržeb bez DPH v dopravě a skladování – očištěno. *ČSÚ: Veřejná databáze: Statistiky: Doprava* [online]. Praha: Český statistický úřad, 2022a [cit. 2022-08-13]. Dostupné z: https://vdb.czso.cz/vdbvo2/faces/cs/index.jsf?page=vystup-objekt&pvo=DOP01-AOC&z=T&f=TABULKA&skupId=1873&katalog=31028&pvo=DOP01-AOC&str=v639&c=v3~4__RP2021MP12&v=v638__STOCISTENI__7604__O

Informace o kategorizaci železniční sítě. *MDČR: Dokumenty: Drážní doprava: Železniční infrastruktura* [online]. Praha: Ministerstvo dopravy ČR, ©2022a [cit. 2022-08-13]. Dostupné z: <https://www.mdcr.cz/Dokumenty/Drazni-doprava/Zeleznicni-infrastruktura/Informace-o-kategorizaci-zeleznicni-site>

Informujeme o výlukách na železnici. *IDSOK: Aktuální výluky a změny: Změny ve veřejné dopravě* [online]. Olomouc: Koordinátor Integrovaného dopravního systému Olomouckého kraje, 2016b [cit. 2022-08-13]. Dostupné z: <https://www.idsok.cz/aktuality/20160930-informujeme-o-vylukach-na-zeleznici/>

Mapy pro širokou veřejnost. *Správa železnic: Portál provozování dráhy* [online]. Praha: Správa železnic, ©2022b [cit. 2022-08-13]. Dostupné z: <https://provoz.spravazeleznic.cz/Portal/ViewArticle.aspx?oid=594598>

Na tratích se takřka na půl roku zastavuje železniční doprava. *IDSOK: Aktuální výluky a změny: Změny ve veřejné dopravě* [online]. Olomouc: Koordinátor Integrovaného dopravního systému Olomouckého kraje, 2016a [cit. 2022-08-13]. Dostupné z: <https://www.idsok.cz/vyluky/na-tratich-se-takrka-na-pul-roku-zastavuje-zeleznicni-doprava/>

Přeprava věcí a osob, přepravní výkony. *ČSÚ: Veřejná databáze: Statistiky: Doprava* [online]. Praha: Český statistický úřad, 2022b [cit. 2022-08-13]. Dostupné z: https://vdb.czso.cz/vdbvo2/faces/cs/index.jsf?page=vystup-objekt&z=T&f=TABULKA&skupId=1613&katalog=31028&pvo=DOP05-C&c=v3~8__RP2020&str=v8#w=

Silhouette Plot. *Orange Visual Programming* [online]. Orange Data Mining, 2015 [cit. 2022-08-13]. Dostupné z: <https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/widgets/visualize/silhouetteplot.html?highlight=silhouette>

Statistická ročenka České republiky - 2021: DOPRAVA, INFORMAČNÍ A KOMUNIKAČNÍ ČINNOSTI: Výkony železniční dopravy. *ČSÚ: Vydáváme: Katalog produktů* [online].

Praha: Český statistický úřad, 2021 [cit. 2022-08-13]. Dostupné z: <https://www.czso.cz/csu/czso/20-doprava-informacni-a-komunikacni-cinnosti-seufqx6g2p>

Tranzitní železniční koridory. *MDČR: Dokumenty: Drážní doprava: Železniční infrastruktura* [online]. Praha: Ministerstvo dopravy ČR, ©2022b [cit. 2022-08-13]. Dostupné z: <https://www.mdcr.cz/Dokumenty/Drazni-doprava/Zeleznicni-infrastruktura/Tranzitni-zeleznicni-koridory>

Tsclust: Time series clustering. *Rdrr.io: R Package Documentation* [online]. 2022 [cit. 2022-08-13]. Dostupné z: <https://rdrr.io/cran/dtwclust/man/tsclust.html>

V České republice jsou první tři potvrzené případy nákazy koronavirem. *MZČR* [online]. Praha: Ministerstvo zdravotnictví ČR, 2020 [cit. 2022-08-13]. Dostupné z: <https://koronavirus.mzcr.cz/v-ceske-republice-jsou-prvni-tri-potvrzene-pripady-nakazy-koronavirem/>

V Nošovicích ráno vypnuli stroje. Koronavirus zastavil výrobu hyundaiů. *IDnes* [online]. Praha: Mafra, 21. 3. 2020 [cit. 2022-08-13]. Dostupné z: https://www.idnes.cz/auto/zpravodajstvi/hyundai-nosovice-vyroba-tovarna-fabrika-koronavirus.A200321_124707_automoto_fdv

Vláda na tři týdny zakázala pohyb mezi okresy. Omezení má zabránit šíření agresivních mutací koronaviru. *Ministerstvo vnitra ČR* [online]. Praha, 2021 [cit. 2022-08-13]. Dostupné z: <https://www.mvcr.cz/clanek/vlada-na-tri-tydny-zakazala-pohyb-mezi-okresy-omezeni-ma-zabranit-sireni-agresivnich-mutaci-koronaviru.aspx>

Vládní usnesení související s bojem proti epidemii – rok 2020. *Vláda ČR* [online]. Praha: Vláda ČR, 30. 12. 2020 [cit. 2022-08-13]. Dostupné z: <https://www.vlada.cz/cz/epidemie-koronaviru/dulezite-informace/vladni-usneseni-souvisejici-s-bojem-proti-epidemii-koronaviru---rok-2020-186999/>

Vládní usnesení související s bojem proti epidemii – rok 2021. *Vláda ČR* [online]. Praha: Vláda ČR, 29. 12. 2021 [cit. 2022-08-13]. Dostupné z: <https://www.vlada.cz/cz/epidemie-koronaviru/dulezite-informace/vladni-usneseni-souvisejici-s-bojem-proti-epidemii---rok-2021-193536/>

Začíná další dlouhá výluka mezi Olomoucí a Šternberkem. *IDSOK: Aktuální výluky a změny: Změny ve veřejné dopravě* [online]. Olomouc: Koordinátor Integrovaného dopravního systému Olomouckého kraje, 2020 [cit. 2022-08-13]. Dostupné z: <https://www.idsok.cz/vyluky/zacina-dalsi-dlouha-vyluka-mezi-olomouci-a-sternberkem/>

Základní charakteristika železniční sítě. *Správa železnic* [online]. Praha: Správa železnic, ©2022a [cit. 2022-08-13]. Dostupné z: <https://www.spravazeznic.cz/o-nas/vse-o-sprave-zeleznic/zeleznice-cr/zeleznicni-sit-v-cr>

Základní údaje. *Správa železnic: O nás: Vše o Správě železnic* [online]. Praha: Správa železnic, ©2022c [cit. 2022-08-13]. Dostupné z: <https://www.spravazeleznic.cz/o-nas/vse-o-sprave-zeleznic/zakladni-udaje>

Železnice Desná. *SART* [online]. Šumperk: SART – stavby a rekonstrukce [cit. 2022-08-13]. Dostupné z: <http://www.sart.cz/zeleznice-desna/>

Ilustrace

Obr. 2:

KRÝŽE, Pavel. Kategorie drah a provozovatelé drah. In: *Správa železnic: Portál provozování dráhy* [online]. Praha: Správa železnic, 2021a [cit. 2022-08-13]. Dostupné z: <https://provoz.spravazeleznic.cz/Portal/ViewArticle.aspx?oid=594598>

Obr. 3:

KRÝŽE, Pavel. Tranzitní koridory. In: *Správa železnic: Portál provozování dráhy* [online]. Praha: Správa železnic, 2019 [cit. 2022-08-13]. Dostupné z: <https://provoz.spravazeleznic.cz/Portal/ViewArticle.aspx?oid=594598>

Obr. 4:

AMIDON, Alexandra. How to Apply K-means Clustering to Time Series Data. In: *Towards Data Science* [online]. 2020 [cit. 2022-08-13]. Dostupné z: <https://towardsdatascience.com/how-to-apply-k-means-clustering-to-time-series-data-28d04a8f7da3>

PŘÍLOHY

SEZNAM PŘÍLOH

Vázané přílohy:

Příloha 1: Podíl druhu dopravy na celkovém počtu vlaků v roce 2016

Příloha 2: Podíl druhu dopravy na celkové hmotnosti vlaků v roce 2016

Příloha 3: Podíl druhu dopravy na celkovém počtu vlaků v roce 2021

Příloha 4: Podíl druhu dopravy na celkové hmotnosti vlaků v roce 2021

Příloha 5: Změna počtu osobních vlaků v březnu-květnu 2020 a 2021

Příloha 6: Změna počtu nákladních vlaků v březnu-květnu 2020 a 2021

Příloha 7: Programový kód shlukování podle časových řad v jazyce R

Elektronické přílohy: (součást CD)

Příloha 8: Struktura železniční dopravy

Příloha 9: Časová analýza vybraných úseků a stanic

Příloha 10: Klasifikace tratí

Volné přílohy:

Příloha 11: Poster

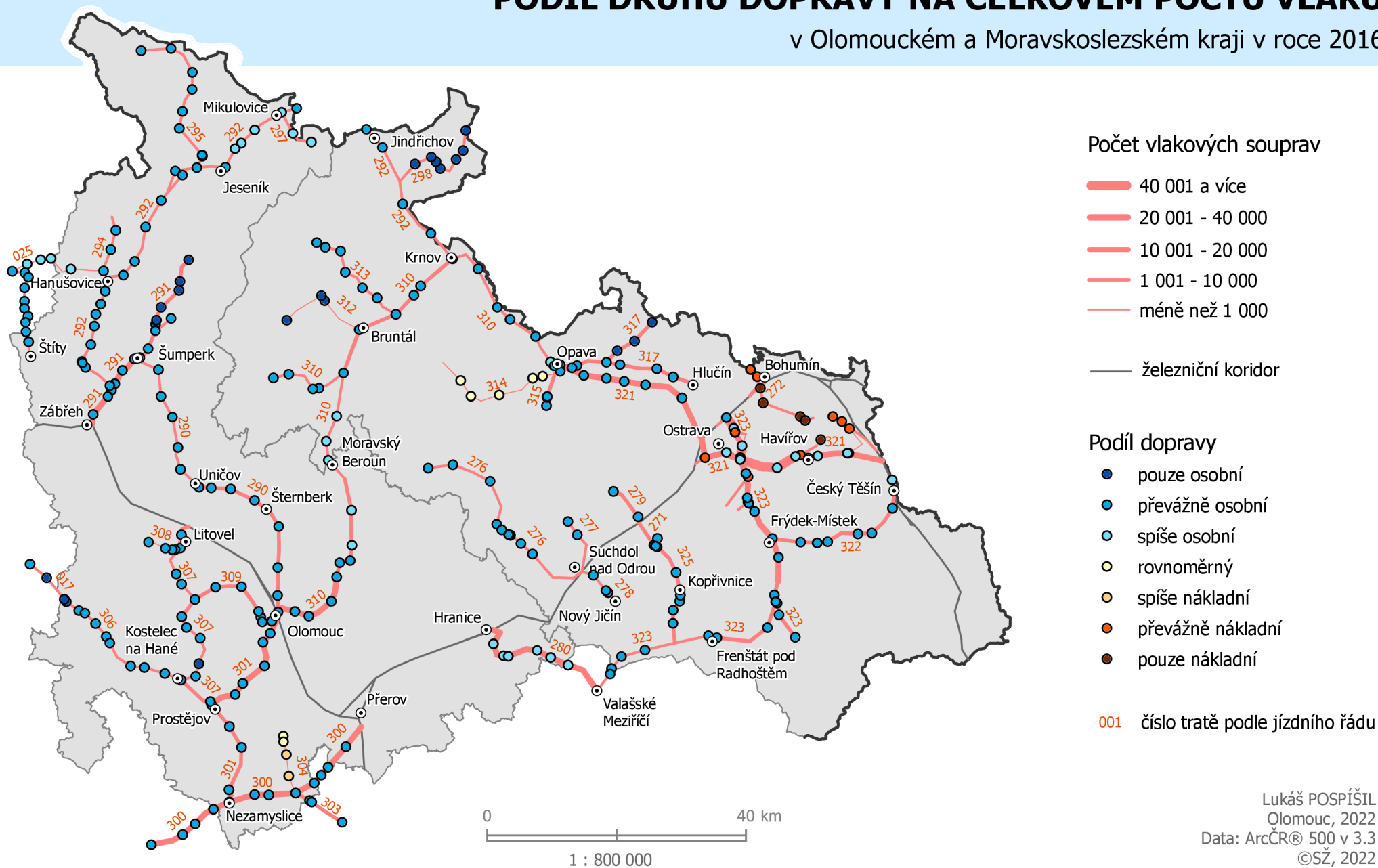
Příloha 12: CD

Popis struktury CD

- Text_prace.pdf
- Poster.pdf
- Data
 - vystupni_data.gdb
- Web
- Prilohy
 - Priloha_8.pdf
 - Priloha_9.pdf
 - Priloha_10.pdf
- Skripty_workflow
 - dekompozice.ows
 - hierarchicke_shlukovani.ows
 - kmeans_shlukovani.ows
 - time_series_clustering.R

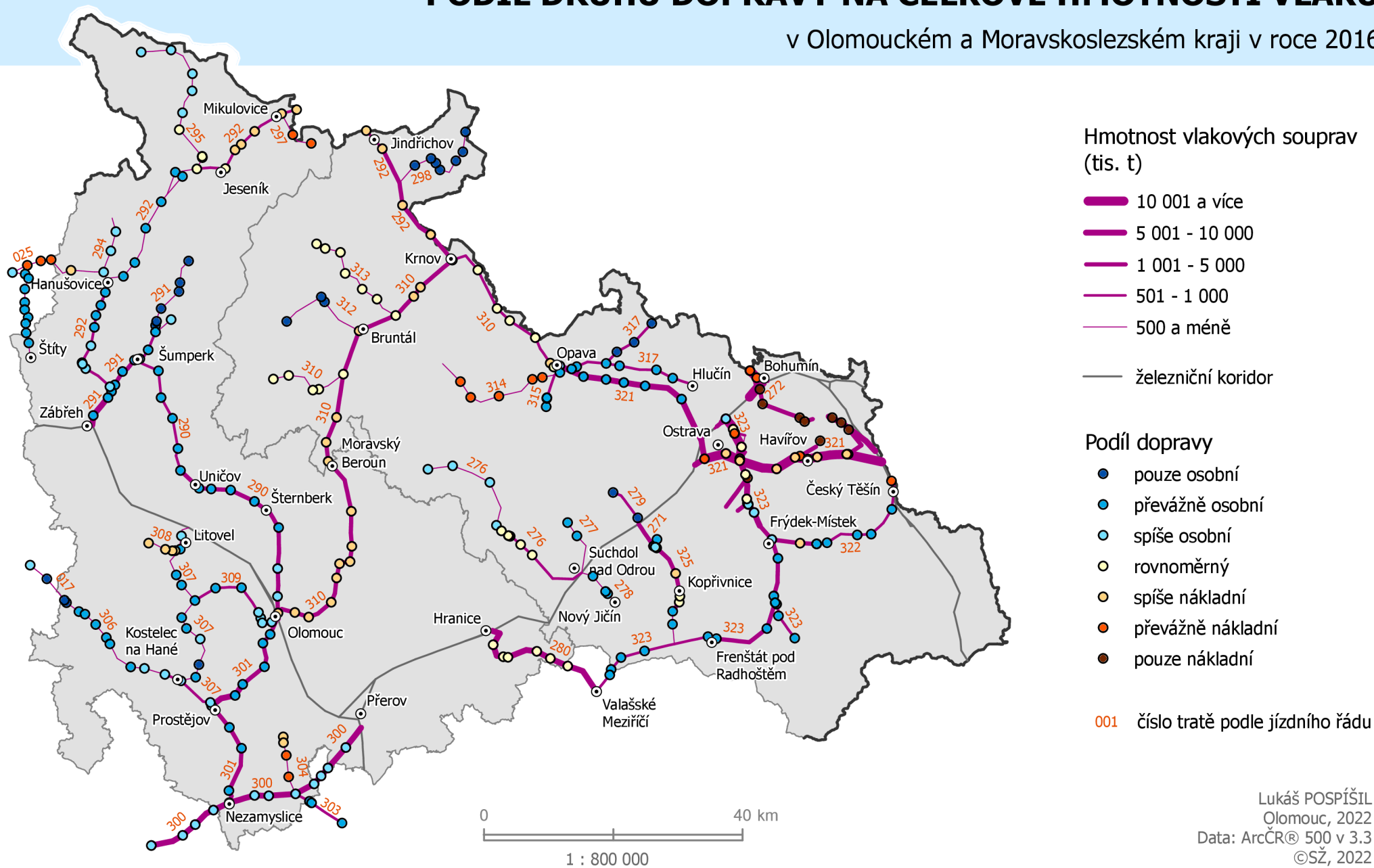
PODÍL DRUHU DOPRAVY NA CELKOVÉM POČTU VLAKŮ

v Olomouckém a Moravskoslezském kraji v roce 2016



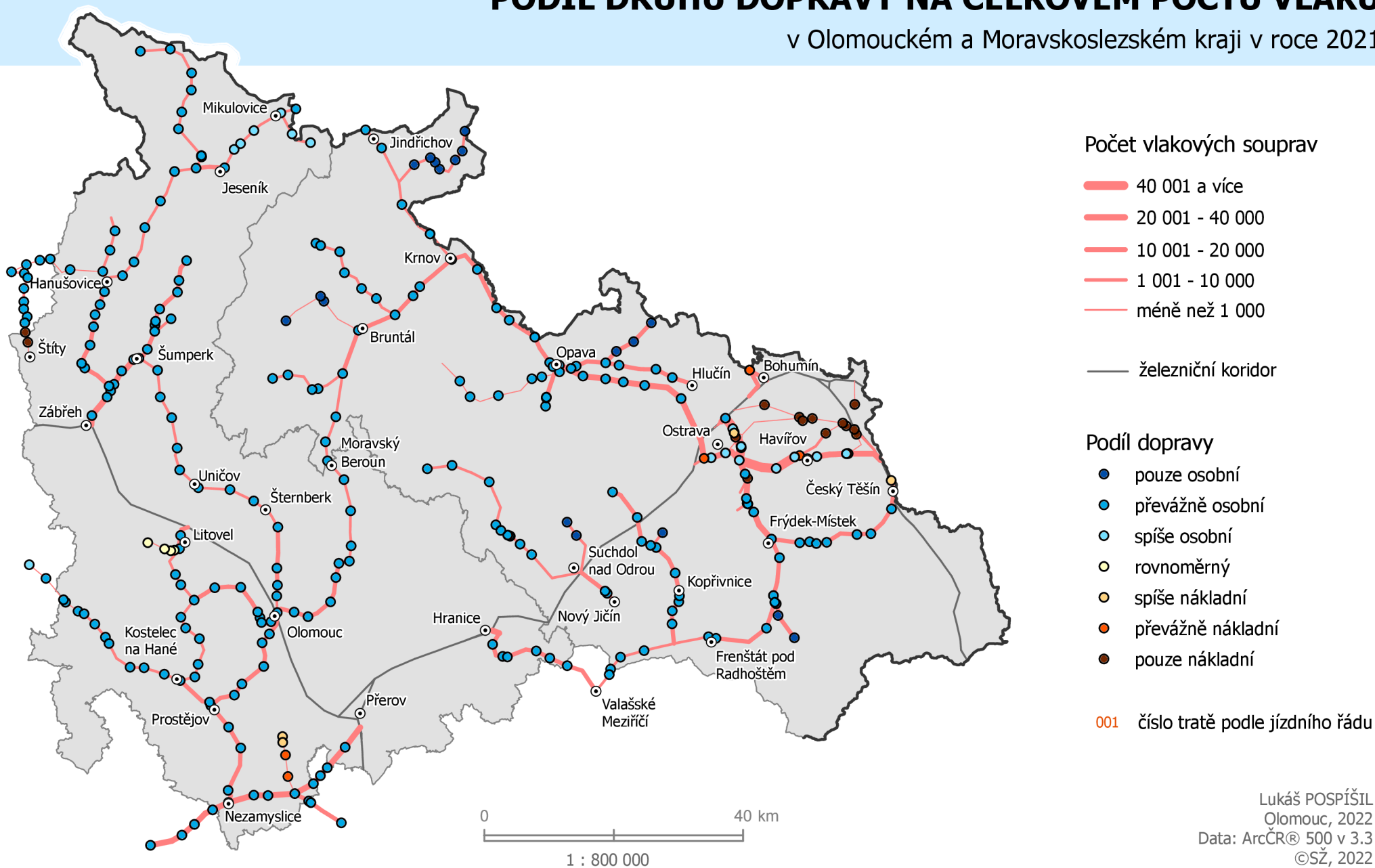
PODÍL DRUHU DOPRAVY NA CELKOVÉ HMOTNOSTI VLAKŮ

v Olomouckém a Moravskoslezském kraji v roce 2016



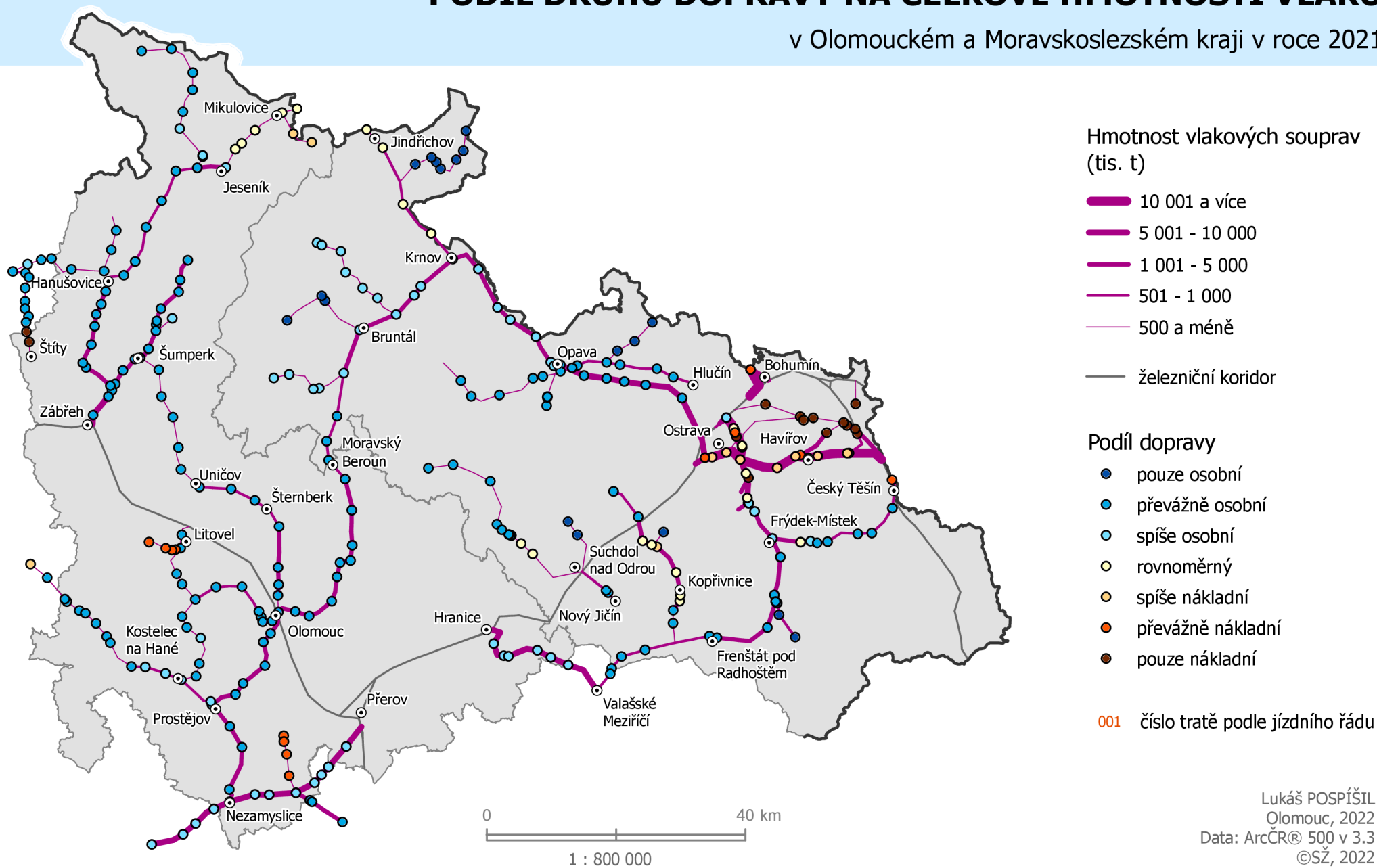
PODÍL DRUHU DOPRAVY NA CELKOVÉM POČTU VLAKŮ

v Olomouckém a Moravskoslezském kraji v roce 2021



PODÍL DRUHU DOPRAVY NA CELKOVÉ HMOTNOSTI VLAKŮ

v Olomouckém a Moravskoslezském kraji v roce 2021



ZMĚNA POČTU OSOBNÍCH VLAKŮ

v březnu–květnu roku 2020 a 2021 v Olomouckém a Moravskoslezském kraji

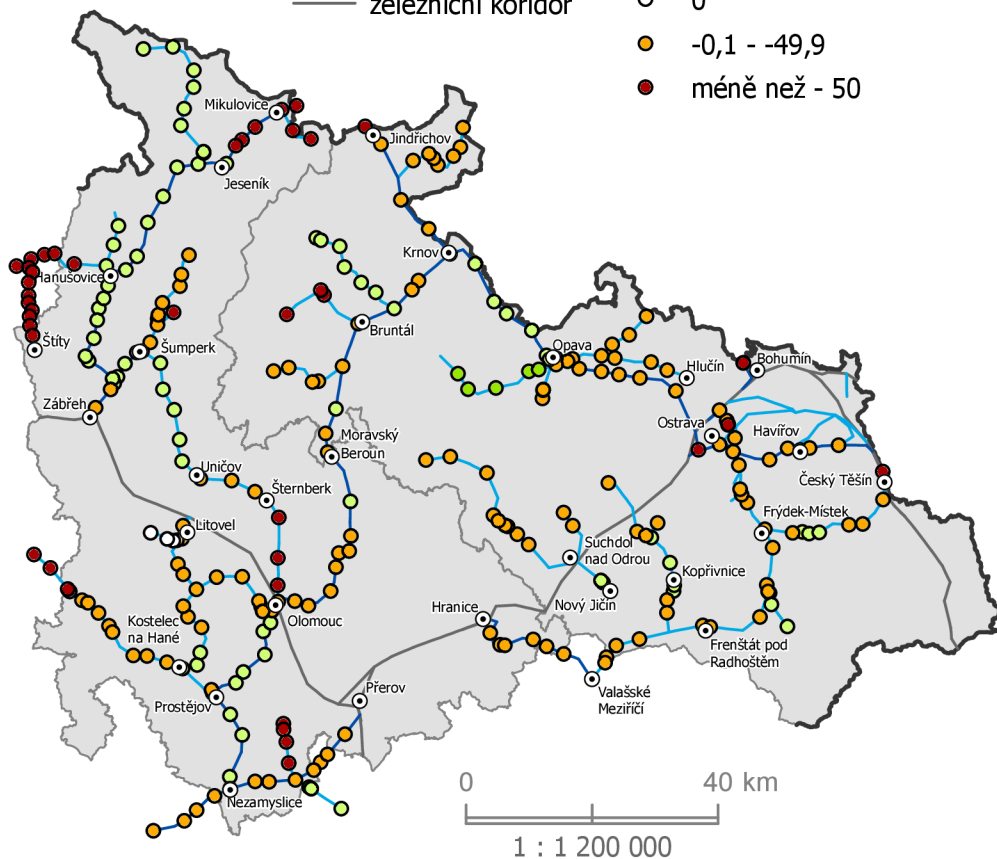
Změna je počítána vůči průměrnému počtu osobních vlaků ze stejných měsíců předchozích let 2016–2019.

2020

- celostátní dráha
- regionální dráha
- železniční koridor

Změna [%]

- 50 a více
- 0,1 - 49,9
- 0
- -0,1 - -49,9
- méně než - 50

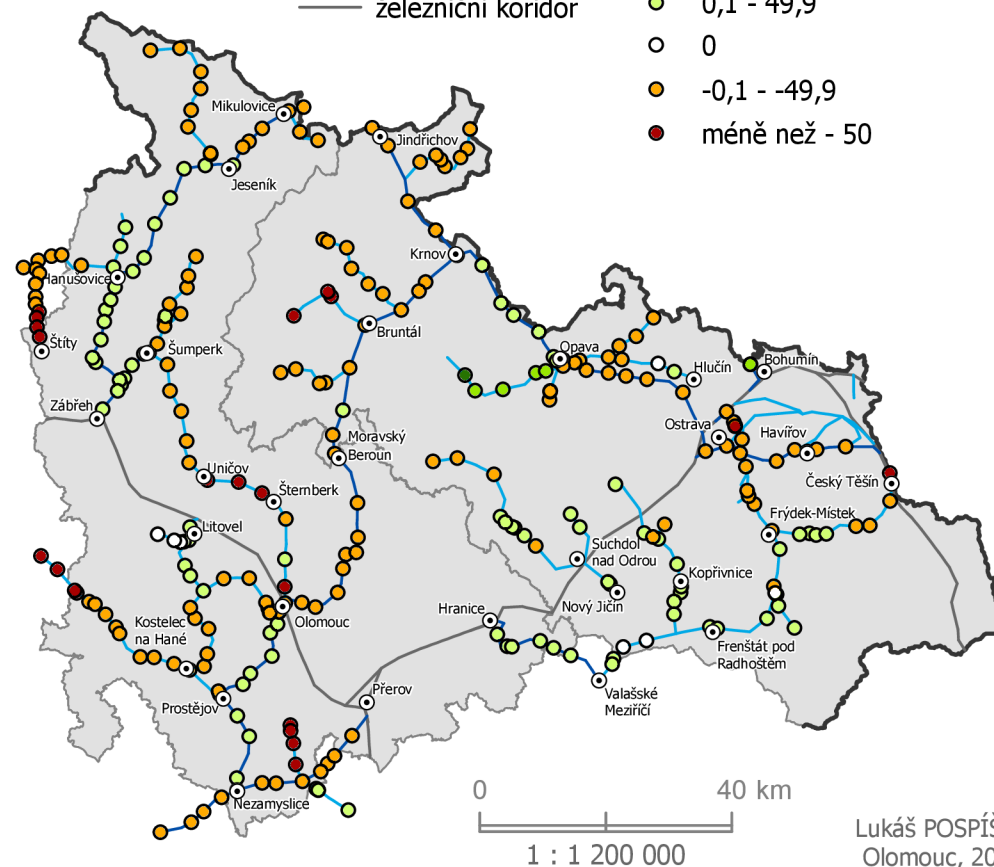


2021

- celostátní dráha
- regionální dráha
- železniční koridor

Změna [%]

- 100 a více
- 50 - 99,9
- 0,1 - 49,9
- 0
- -0,1 - -49,9
- méně než - 50

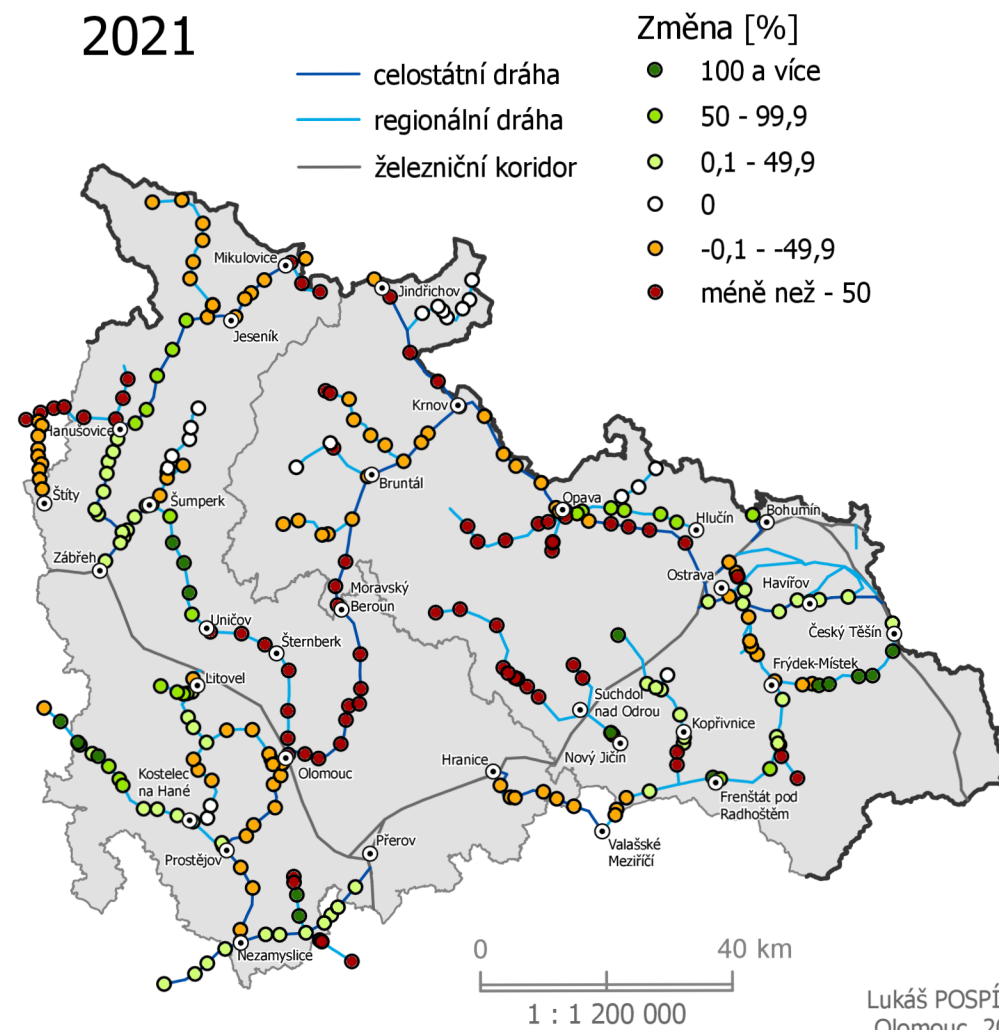
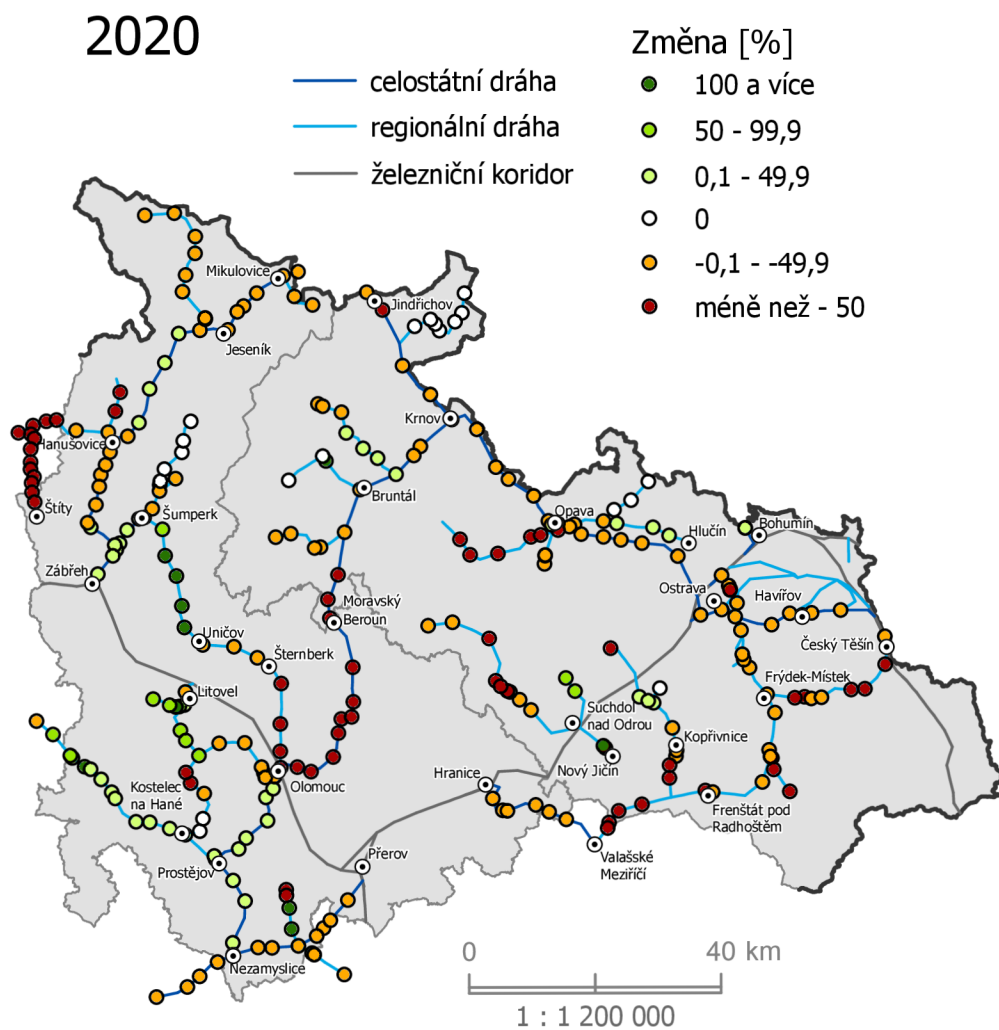


Lukáš POSPÍŠIL
Olomouc, 2022
Data: ArcCR® 500 v 3.3
©SŽ, 2022

ZMĚNA POČTU NÁKLADNÍCH VLAKŮ

v březnu–květnu roku 2020 a 2021 v Olomouckém a Moravskoslezském kraji

Změna je počítána vůči průměrnému počtu nákladních vlaků ze stejných měsíců předchozích let 2016–2019.



Príloha 7: Programový kód shlukování časových řad v jazyce R

```
#Nacteni knihoven
library(dtwclust)
library(BBmisc)
library(readxl)
library(openxlsx)
library(tidyverse)

#Nastaveni pracovniho adresare a nacteni vstupnich dat
setwd("workspace")

data <- read_excel("input.xlsx")

#Nastaveni prvnioho sloupce jako nazev radku
data <- data %>% remove_rownames %>% column_to_rownames(var="cislo_useku")

#Normalizace dat
data.norm <- BBmisc::normalize(data, method="standardize")

### Shlukovani casovych rad s vyuzitim metriky DTW ###
#Vypocet statistiky pro ruzny pocet cilovych shluku.
clust.dtw <- tsclust(data.norm, type="partitional", k=2L:9L,
distance="dtw", centroid="pam")

#Volba poctu cilovych shluku podle statistiky.
list(clust.dtw)

#Shlukování metodou nejbližsich stredu s pevnym poctem shluku.
clust.dtw <- tsclust(data.norm, type="partitional", k=6L,
distance="euclidean", centroid="pam")

#Liniove grafy vseh shluku
plot(clust.dtw, type = "sc")

#Liniove grafy medoidu vseh shluku
plot(clust.dtw, type = "centroids", lty = 1, lwd = 1)

#Vytvoreni seznamu shlukovanych objektu s cislem prirazeneho clusteru
t(cbind(data[,0], cluster = clust.dtw@cluster))

seznam.dtw <- data.frame(cbind(data[,0], cluster = clust.dtw@cluster))

#Export seznamu do formátu .xlsx
write.xlsx(seznam.dtw,"output.xlsx", rowNames=TRUE)

### Shlukovani casovych rad s vyuzitim Euklidovy metriky ###
#Vypocet statistiky pro ruzny pocet cilovych shluku.
clust.euc <- tsclust(data.norm, type="partitional", k=2L:9L,
distance="euclidean", centroid="pam")
```

```
#Volba poctu cilovych shluku podle statistiky.  
list(clust.euc)  
  
#Shlukování metodou nejbližsich stredů s pevným počtem shluku  
clust.euc <- tsclust(data.norm, type="partitional", k=6L,  
distance="euclidean", centroid="pam")  
  
#Liniove grafy vsech shluku  
plot(clust.euc, type = "sc")  
  
#Liniove grafy medoidu vsech shluku  
plot(clust.euc, type = "centroids", lty = 1, lwd = 1, col = "#6baed6")  
  
#Vytvoreni seznamu shlukovanych objektu s cislem prirazeneho clusteru  
t(cbind(data[,0], cluster = clust.euc@cluster))  
seznam.euc <- data.frame(cbind(data[,0], cluster = clust.euc@cluster))  
  
#Export seznamu do formátu .xlsx  
write.xlsx(seznam.euc, "output.xlsx", rowNames=TRUE)
```