

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačních technologií



Bakalářská práce

Analýza cenotvorby apartmánového bydlení v České republice na vybrané platformě

Jan Kudláček

© 2024 ČZU v Praze

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Jan Kudláček

Podnikání a administrativa

Název práce

Analýza cenotvorby apartmánového bydlení v České republice na vybrané platformě

Název anglicky

Analysis of apartment housing pricing in the Czech Republic on a selected platform

Cíle práce

Hlavním cílem práce je ekonomická analýza podnikání v rámci apartmánového bydlení na zvolené platformě v České republice. V dílčích cílech bude podrobně popsáno zkoumané prostředí a následně bude vyhodnocen vztah vlastností apartmánu k ceně za zvolené časové období.

Metodika

Metodika práce se skládá ze dvou částí. V úvodní teoretické části bude analyzována zkoumaná problematika na základě odborných zdrojů. V praktické části bude na podkladě získaných dat vyhodnocena průměrná cena v každém regionu. Následně bude provedena regresní analýza (OLS) v rámci, které budou zkoumány a hodnoceny vlastnosti apartmánů a jejich vliv na cenu. Data budou získána pomocí skriptu, jenž bude automaticky vyhledávat na stránce Airbnb.com a výsledky budou použity k následnému vyhodnocení. Z provedené analýzy bude výstup pro tvorbu cen apartmánů pro jednotlivé oblasti (kraje) České republiky a tímto se zpřehlední zkoumaný segment trhu. Na podkladě výsledků analýzy budou v závěru uvedeny přínosy práce.

Doporučený rozsah práce

40

Klíčová slova

Apartmánové bydlení, airbnb, regresní analýza

Doporučené zdroje informací

Benitez-Aurioles, Beatriz, and Iis Tussyadiah. "What Airbnb does to the housing market." *Annals of Tourism Research* 90 (2020): 103108.

Lladós-Masllorens, Josep, Antoni Meseguer-Artola, and Inma Rodríguez-Ardura. "Understanding peer-to-peer, two-sided digital marketplaces: pricing lessons from Airbnb in Barcelona." *Sustainability* 12.13 (2020): 5229.

Perez-Sanchez, V. Raul, et al. "The what, where, and why of Airbnb price determinants." *Sustainability* 10.12 (2018): 4596.

Předběžný termín obhajoby

2023/24 LS – PEF

Vedoucí práce

Ing. Karel Kubata, Ph.D.

Garantující pracoviště

Katedra informačních technologií

Elektronicky schváleno dne 4. 7. 2023

doc. Ing. Jiří Vaněk, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 3. 11. 2023

doc. Ing. Tomáš Šubrt, Ph.D.

Děkan

V Praze dne 03. 12. 2023

Čestné prohlášení

Prohlašuji, že svou bakalářskou práci "Analýza cenotvorby apartmánového bydlení v České republice na vybrané platformě" jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce. Jako autor uvedené bakalářské práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 15. 3. 2024

Poděkování

Rád bych touto cestou poděkoval Ing. Karlu Kubatovi, Ph.D. za odborné vedení mé bakalářské práce, za rady a pomoc při zpracování této práce. Děkuji také firmě Rushmore s.r.o. za poskytnutí databáze dat získaných z platformy Airbnb.com.

Analýza cenotvorby apartmánového bydlení v České republice na vybrané platformě

Abstrakt

Tato bakalářská práce se zabývá analýzou faktorů apartmánového bydlení a metodikám, které jsou použity pro nalezení výsledků. Teoretická část práce se věnuje studiu odborných publikací v kontextu daného tématu a zkoumá obdobná řešení problematiky. V této části je vysvětlen princip funkcionality platformy Airbnb, její historie, možnosti a koncept cenotvorby. Teoretická část rovněž definuje a popisuje metody používané v praktické části. Ve vlastní práci probíhá analýza průměrné ceny prodaných nocí pro kraje České republiky v roce 2023. Dále je zde obsažena analýza vztahu vlastností k průměrné ceně volných nocí. Tato analýza odhaluje pozitivní vliv na průměrnou cenu při růstu počtu ložnic, vybaveností vířivkou či klimatizací a statusu „superhostitele“. Dle výsledků má na průměrnou cenu za noc záporný vliv rostoucí vzdálenost od turistické lokality či možnost bezplatného parkování.

Klíčová slova: apartmán, Airbnb, regresní analýza OLS, ekonomická analýza, vlastnosti apartmánu, průměrná cena

Analysis of apartment housing pricing in the Czech Republic on a selected platform

Abstract

This bachelor thesis is focusing on the analysis of factors of apartment housing and the methods that are used to get results. Theoretical part of this thesis is concentrated on the study of scientific papers and literature in the context of designated topic as well as researching similar solutions. This part also explains the working principle of Airbnb platform, its brief history and methods of pricing. Another topic this part is focusing on is the definition and explanation of methods used in latter part of this thesis. The practical part analyses the average price of sold nights for the regions of the Czechia in 2023. Another analysis examines relationship between chosen apartment characteristics and average prices. The results of this analysis show that amenities like hot tub or air conditioning, rising number of bedrooms or the superhost badge have a positive impact on average prices. Results also show that with a growing distance from a tourist site the average price decreases. Negative impact on the average prices is showing for apartments that have free parking options.

Keywords: apartment, Airbnb, regression analysis OLS, economic analysis, apartment attributes, average price

Obsah

1 Úvod	10
2 Cíl práce a metodika	11
2.1 Cíl práce	11
2.2 Metodika	11
3 Teoretická východiska	12
3.1 Informatika a informace	12
3.2 Systém a informační systém	12
3.3 Uživatelské rozhraní a API	13
3.4 Online platformy pro pronájem apartmánů	14
3.4.1 Ubytovací zařízení	14
3.4.2 Historie online platforem pro pronájem apartmánů	15
3.4.3 Airbnb	16
3.4.4 Rizika pro hotelový průmysl	17
3.5 Koncept cenotvorby	18
3.5.1 Obecná tvorba cen	18
3.5.2 Specifika cenotvorby apartmánového bydlení na online platformách	19
3.6 Automatizovaný sběr dat z online platforem a zpracování jeho výsledků	20
3.6.1 Sběr dat	20
3.6.2 Metody zpracování nasbíraných dat	23
3.7 Metoda nejmenších čtverců	26
3.8 Shrnutí	27
4 Vlastní práce	28
4.1 Selekce dat	28
4.2 Analýza průměrné ceny v krajích České republiky	29
4.2.1 Rozdělení apartmánů	29
4.2.2 Výběr vhodné ceny	30
4.2.3 Průměrná cena apartmánů s 1 ložnicí	31
4.2.4 Průměrná cena apartmánů se 2 ložnicemi	33
4.2.5 Průměrná cena apartmánů se 3 až 6 ložnicemi	34
4.2.6 Průměrná cena za noc pro jednotlivé typy apartmánů	35
4.2.7 Mapa České republiky a průměrná cena v krajích	35
4.3 Regresní analýza atributů a jejich vliv na cenu	36
4.3.1 Výběr proměnných	36
4.3.2 Distribuce reziduí	41
4.3.3 Výsledky analýzy	43
4.3.4 Vliv vzdálenosti na průměrnou cenu v Praze	45

5	Výsledky a diskuse	47
5.1	Analýza průměrné ceny	47
5.1.1	Výsledky pro apartmány s 1 ložnicí	47
5.1.2	Výsledky pro apartmány se 2 ložnicemi	47
5.1.3	Výsledky pro apartmány se 3 až 6 ložnicemi	48
5.1.4	Mapa krajů ČR s průměrnými cenami	48
5.2	Analýza vlivu atributů na průměrnou cenu	48
5.2.1	Výsledky analýzy se zaměřením na Českou republiku	49
5.2.2	Výsledky analýzy se zaměřením na Prahu	49
5.3	Komparativní analýza průměrné ceny s výsledky regresní analýzy	49
6	Závěr	51
7	Seznam použitých zdrojů	53
	Seznam obrázků, tabulek, grafů a zkratk	56
7.1	Seznam obrázků	56
7.2	Seznam tabulek	56
7.3	Seznam grafů	56
7.4	Seznam použitých zkratk	56
Přílohy		57
	Příloha A Grafy používané při analýze průměrných cen krajů ČR	58

1 Úvod

Bakalářská práce se zabývá problematikou cenotvorby apartmánového bydlení v České republice na platformě Airbnb s vazbou na technologické řešení analýzy cen a sběru dat. Práce se zaměřuje na základní pojmy z oblasti informatiky a informačních systémů. Dále se zabývá konceptem online platform pro pronájem apartmánů, včetně jejich historie, principu fungování a cenotvorby. Zvolenou online platformou je Airbnb, která na principu peer-to-peer spojuje majitele nemovitostí a osoby se zájmem o cestování. Platforma v současné době stále roste v popularitě mezi turisty a majiteli nemovitostí, proto je toto téma aktuální. Jsou také vysvětleny metody, kterými se došlo k výsledkům. Vlastní práce obsahuje analýzu dat získaných z Airbnb a je provedena aplikováním několika postupů. Data zahrnují informace o apartmánech, jejich vlastnostech, hostitelích a další. Práce popisuje použitý soubor dat a jeho vlastností z hlediska rozdělení a obsažených hodnot. Soubor dat je pro účely práce poskytnut firmou operující v tomto sektoru podnikání. Data byla získána pomocí softwaru, jehož princip je vysvětlen v teoretické části práce. Vzhledem k absenci přístupu k aplikačně-programovému rozhraní platformy Airbnb jsou data extrahována pomocí specifického druhu skriptu. Dále se práce soustředí na zhodnocení průměrných cen pro jednotlivé kraje České republiky za rok 2023 a také na analýzu cen v závislosti na různých faktorech, jako např. umístění apartmánu, úroveň hostitele a další vybrané vlastnosti. Zpracovávají se průměrné ceny prodaných i neprodaných nocí. Zkoumané vlastnosti jsou vybrány na základě studia požadavků od hostů a na dostupnosti těchto údajů z poskytnuté databáze. Analýzy jsou provedeny použitím statistických metod zpracovaných vlastním skriptem. Jeho nasazení a podrobný popis je obsažen v teoretické části. Tento skript je řešen v programovacím jazyce Python a umožňuje úzkou specifikaci na řešené téma. Zhodnocené ceny a vlastnosti poskytují výstup pro tvorbu cen apartmánového bydlení v sektoru podnikání krátkodobého ubytování.

2 Cíl práce a metodika

2.1 Cíl práce

Hlavním cílem práce je ekonomická analýza podnikání v rámci apartmánového bydlení na zvolené platformě v České republice. V dílčích cílech bude podrobně popsáno zkoumané prostředí a následně bude vyhodnocen vztah vlastností apartmánu k ceně za zvolené časové období.

2.2 Metodika

Metodika teoretické části využívá studium odborných publikací, článků a studií zabývajících se ekonomickým aspektem podnikání v oblasti krátkodobého ubytování. Tato část vysvětluje princip metod použitých v následující části práce. Tím poskytuje pohled na zkoumané prostředí pro praktickou část. V praktické části je provedena analýza průměrné ceny za noc v jednotlivých oblastech České republiky a kvantitativní analýza, kde je hlavním nástrojem regresní analýza metodou nejmenších čtverců. Ta posuzuje vztah mezi charakteristikami apartmánů a jejich cenou na základě dat získaných z platformy Airbnb, poskytnutých od vybrané firmy. Data jsou extrahována pomocí skriptu, který automaticky vyhledává nabídky na dané platformě se zaměřením na Českou republiku. Tento skript shromažďuje relevantní informace o cenách, lokalitě, velikosti, výbavě a dalších vlastnostech apartmánu. Informace jsou uloženy do strukturované podoby v relační databázi. Pomocí vlastního skriptu, zčásti popsaného v teoretické části, probíhají s těmito daty výběrové a transformační operace nutné pro následnou aplikaci analýz. Skript využívá několika knihoven pro vyjmenované postupy. Zpracované hodnoty jsou prezentovány využitím grafů, tabulek a obrázků. Výsledky regresní analýzy jsou interpretovány s cílem identifikace determinantů ceny a poskytnutí doporučení pro stanovení cen apartmánů.

3 Teoretická východiska

3.1 Informatika a informace

Informace je základním stavebním kamenem informatiky. Její definice, dle Norberta Wienera, říká, že informace je označení pro obsah, který se vymění s vnějším světem při jeho ovlivňování. Informace se tedy týká obsahu a výměny tohoto obsahu. [1]

Informatika je vědou o zpracování, přenášení informací v určitém systému [1]. Zabývá se tvorbou informačních systémů aplikováním inženýrských a komunikačních technologií [2]. Na jejím vzniku se podílí kybernetika a systémová věda. Věda o kybernetice zkoumá princip řízení živých systémů a strojů. Je také základem pro vznik počítačů. Systémová věda se zabývá řízením podniků a inženýrských procesů. Poskytuje nástroje pro analýzu a návrh komplexních systémů, včetně informačních systémů. Systémovou vědu je možné rozdělit na systémovou analýzu a systémové inženýrství. Analýza zkoumá problémy týkající se řízení organizace a systémové inženýrství poté navrhuje a vytváří komplexní systémy. [1]

3.2 Systém a informační systém

Systém je modelem části světa a je tvořen prvky, jež jsou vzájemně propojeny vazbami. Prvky a jejich vlastnosti neboli atributy, jsou v systému složeny a uspořádány určitým způsobem. Tento způsob vyjadřuje strukturu systému. K určitému okamžiku se v systému zkoumá stav a ten je tvořen hodnotami atributů. Chování systému představuje akce a reakce na vzniklé podněty ze systému nebo z okolí. Jako příklad systému lze uvést hotel, jehož prvky tvoří zaměstnanci, hosté, vybavení hotelu. Tyto prvky mezi sebou mají určitý vztah představující vazbu mezi prvky. Součástí tohoto systému jsou i ostatní hotely. Ty se obecně označují pojmem okolí systému. Jsou to ostatní příbuzné systémy, se kterými systém může či nemusí komunikovat. Při komunikaci s okolím se jedná o otevřený systém, pokud žádná interakce neprobíhá je to uzavřený systém. Prvky komunikující s vnějším okolím se značí pojmem hraniční prvek. [1][3]

Informační systém, zkráceně IS, je systém zaměřující se na uchování, sběr, výměnu informací s okolím. Jeho vstupy i výstupy jsou informace, prvky informačního systému tvoří lidé, informační a komunikační technologie a data. Procesy definují, jak se s informacemi zachází, konkrétně se jedná se o zpracování nebo výměnu informace lidmi neformálním způsobem, výměnu založenou na předepsaných pracovních a informačních tocích či kompletní zpracování informace a vyhodnocení pouze pomocí počítače. Správně

nastavené procesy výměny informací a jejich zpracování umožňuje informačnímu systému dosáhnout potřebných cílů. Cílem může být například zvýšení efektivity provozu, zlepšení komunikace, uchování dat, anebo jiné specifické cíle. [1][2]

Jedním příkladem informačního systému jsou platformy pro pronájem apartmánů. Jeho prvky jsou majitelé apartmánů, hosté, apartmány a jejich kalendáře spolu s cenami. Tento informační systém propojuje jednotlivé hosty s hostiteli a jejich apartmány. Kromě vyjmenovaných prvků tento systém také zkoumá preference zákazníka, měří výkon apartmánů, vyhodnocuje ukončené rezervace, konverzace a další. Veškeré tyto operace většinou probíhají na pozadí bez vědomí uživatele informačního systému. Pro práci s informačním systémem uživatel využívá speciálně navrženého rozhraní. Může se jednat o webové, aplikační, či jiné rozhraní. Účelem rozhraní je poskytnout uživateli informačního systému přehledný způsob zadávání informací, jejich analýzu a následné zobrazení výstupů.

3.3 Uživatelské rozhraní a API

Ať už se jedná o webovou aplikaci, hru či operační systém, ovládá se pomocí uživatelského rozhraní. Uživatelské rozhraní je například i paluba auta nebo ovladač k televizi, tato kapitola je však mířena pouze na uživatelské rozhraní v rámci oboru informatiky, konkrétněji na softwarové uživatelské rozhraní. Uživatelské rozhraní a API jsou vzájemně úzce provázány, jelikož API poskytuje uživatelskému rozhraní obsah k zobrazení v požadovaném formátu. [4]

Uživatelské rozhraní, anglicky User Interface (UI), je místo, na kterém dochází k výměně informací mezi člověkem a počítačem. Rozhraní umožňuje uživateli ovládat počítač, program, aplikaci a dosáhnout pomocí něj určitých cílů. Vzhled je první dojem, který uživatel získává o uživatelském rozhraní počítače, stránky nebo programu. Řadí se sem zvolené fonty textů, barvy elementů, obrázky a ikony, ovládací prvky a další prvky. [5]

Cílem uživatelského rozhraní je poskytnout uživateli jednoduchý a přehledný způsob, jak daný systém ovládat. UX, zkratka pro User Experience, je zkušenost, kterou uživatel získává při používání UI. UX může být pozitivní nebo negativní podle toho jak dobře je UI navrženo. Obsahuje také očekávání uživatelů, jejich obavy, kontext použití aplikace apod. Ačkoliv je UX něco neviditelného na pozadí, ve značném měřítku určuje úspěšnost produktu. Tato zkušenost uživatele je nadřazená uživatelskému rozhraní, jelikož dobrá zkušenost není možná bez kvalitně navrženého uživatelské rozhraní. [4]

API, v plném názvu Application Programming Interface, je sada pravidel, která definuje, jak aplikace komunikují a interagují s ostatními aplikacemi. API poskytují způsob, jak aplikace mohou získávat a zpracovávat data z jiných aplikací nebo jak mohou jiné aplikace ovládat funkce té druhé. Mohou pomoci zvýšit produktivitu, snížit náklady a zlepšit spolupráci. [6]

Nejvíce používané API se dají rozdělit do několika kategorií. Nejpopulárnější jsou REST API, které využívají klasických HTTP metod jako GET, POST, a jiné. Dalším druhem je SOAP, který data posílá ve struktuře XML a používá se spíše ve starších systémech nebo v korporátních firmách. GraphQL, také jedním z typů API, je výjimečný v tom, že uživateli dovoluje získat přesně ta data, která si vyžádal. Webové API jsou často používány pro vývoj webových, mobilních a dalších aplikací. [7]

API fungují na principu volání funkcí. Když chce aplikace použít API, vytvoří na ni požadavek. Tento požadavek obsahuje informace o tom, jakou funkci chce aplikace použít a jaké parametry chce předat funkci. API pak zpracovává požadavek a vrací odpověď. Odpověď může obsahovat data, která aplikace požadovala, nebo chybovou zprávu. [6][8]

3.4 Online platformy pro pronájem apartmánů

Tato kapitola je zaměřena na apartmány z hlediska krátkodobého ubytování a jejich prezentaci na internetu. Nejprve je vysvětleno, co je to apartmán a jaké druhy apartmánového bydlení existují. Dále jsou v kapitole prezentovány online platformy určené pro inzerci apartmánů. U těchto platforem je vysvětlen způsob, jakým fungují, princip generace příjmů a další.

3.4.1 Ubytovací zařízení

Podle specifikací od Asociace hotelů a restaurací České republiky je spektrum ubytovacích možností v naší zemi rozčleněno do několika základních skupin. Patří mezi ně hotely, motely, apartmánová ubytovací zařízení a penziony [9]. Tato bakalářská práce se soustředí na apartmánový typ ubytování, a proto je zásadní podrobně popsat kritéria, která jsou pro tuto kategorii stanovena.

Apartmánové ubytování se dělí na dva základní typy – apartmánový hotel a apartmánový komplex. Hotel i komplex nabízí ubytování ve studiích či apartmánech a oba také mohou být zařazeny do třídy dvou hvězdičkových až pěti hvězdičkových.

Apartmánový hotel se poté od komplexu liší v minimálním počtu ubytovacích jednotek, pro apartmánový hotel deset a pět pro apartmánový komplex. Další rozdíl je v nabízených službách. Zatímco apartmánový hotel poskytuje rozsah služeb odpovídající třídě hotelu, apartmánový komplex hostům poskytuje omezenou nabídku služeb. Ubytovací zařízení v hotelech i komplexech tohoto typu musí splňovat určité náležitosti pro získání kvalifikace. Všechna ubytovací zařízení musí být vybavena kuchyňským vybavením pro přípravu jídel a nápojů a také prostředky na úklid. K dalším podmínkám, které musí ubytovací zařízení splňovat, patří jednoznačné vymezení funkce obývacího prostoru, definovaného jako místo pro denní aktivitu návštěvníků bez umístění pevných lůžek. Dále se musí specifikovat, které části apartmánu se započítávají do celkové obytné plochy, a stanovit se musí i minimální počet jednotek přizpůsobených pro snadný přístup osobám s omezenou pohyblivostí. [9]

Zatímco Asociace hotelů a restaurací České republiky má jasně stanovené podmínky pro různé druhy ubytování, na platformě Airbnb je situace odlišná. Při vytvoření inzerátu ubytování si druh ubytování hostitel vybere ze seznamu a volí podle toho, jaká položka nejvíce odpovídá skutečnosti. Airbnb apartmán definuje jako typ nemovitosti, obvykle umístěný v budově s více jednotkami, kde žijí další lidé, přičemž se může jednat o byt nebo půdní vestavbu. [10]

3.4.2 Historie online platform pro pronájem apartmánů

Před příchodem internetu si lidé rezervovali dovolenou tradičními způsoby, jako byl telefon, cestovní kancelář nebo cestovní průvodce. Cestovní kanceláře nabízely širokou škálu možností ubytování, včetně hotelů, apartmánů a dalších. Potenciální hosté si mohli vybrat ubytování podle svých potřeb a rozpočtu.

V 90. letech 20. století se objevil internet a postupně se stal populárním způsobem rezervace dovolené. Vznikly první online cestovní agentury a lidé se naučili je více používat. Online platformy se poté pro pronájem apartmánů uchytily z několika důvodů. Nebylo nutné navštěvovat cestovní kanceláře, nový způsob přinášel větší flexibilitu realizace dovolené a také cenovou výhodnost. Tento nedávný nárůst peer-to-peer platform lidem umožnil efektivně využívat nevyužitý zdroj ve velkém měřítku a sdílet je tak s ostatními. Tím vznikla éra označovaná jako sdílená ekonomika. [11]

Online platformy pro pronájem apartmánů se začaly objevovat v počátcích 21. století. Průkopníkem v této oblasti byla společnost Airbnb, která byla založena v roce 2007. Airbnb

rychle získala popularitu a ke konci roku 2022 působila ve více než 220 zemích světa. Tato platforma nabízela dosud nerozšířený způsob krátkodobého ubytování, od ostatních se odlišovala svojí komunitou. Nejen hosté ale i hostitelé byli běžnými lidmi, narozdíl od např. platformy Booking, kde hostitelé byli podnikatelé či firmy. [12]

Booking.com™ je další významnou online platformou pro pronájem apartmánů, byla založena v roce 1996. Booking.com se původně zaměřovala na hotely, ale v posledních letech se rozšířila i na nabídku apartmánů. [13]

3.4.3 Airbnb

Airbnb, Inc. je online služba, která spojuje lidi, kteří hledají ubytování, s lidmi, kteří nabízejí své byty, chaty nebo jiné prostory k pronájmu. Působí celosvětově a k 30. 6. 2023 nabízela více než sedm milionů inzerátů. Airbnb nabízí různé druhy ubytování, od levných pokojů až po luxusní vily. Hosté si mohou vybrat ubytování podle svých potřeb a rozpočtu. Airbnb umožňuje lidem, kteří mají volný prostor, pronajmout ho dalším lidem a získat tak finanční příjem. Hostitelé mohou své ubytování inzerovat na platformě a nastavit si vlastní cenu. [12][14]

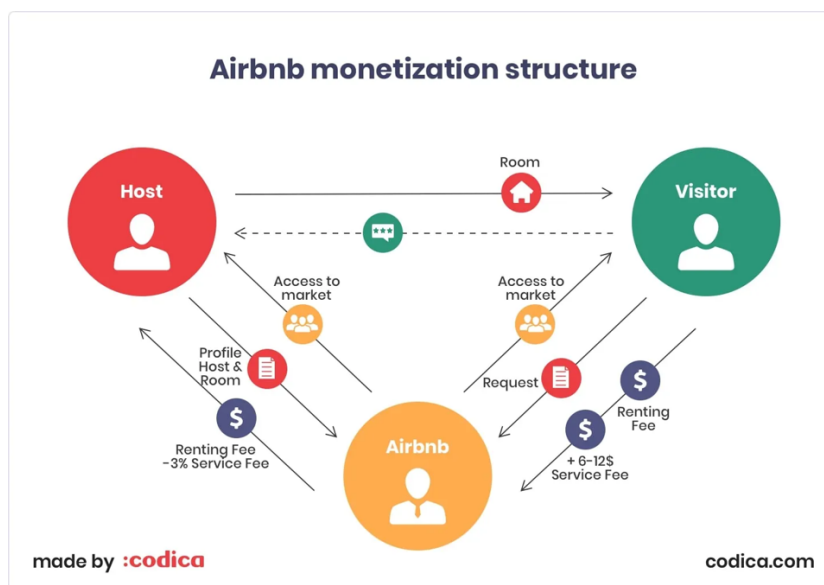
Platforma poskytuje uživatelům podrobné informace o ubytování, včetně fotografií popisu místa, dostupných vybavení a recenzí od předchozích hostů. Tímto způsobem Airbnb zvyšuje transparentnost a pomáhá hostům udělat informované rozhodnutí. [14]

Významnou charakteristikou Airbnb je také komunitní aspekt. Hostitelé a hosté komunikují přímo prostřednictvím platformy, což umožňuje oběma stranám lépe se poznat a dohodnout se na podmínkách ubytování. Tato interakce přináší osobnější zážitek v porovnání s tradičními hotelovými službami. [12]

Airbnb používá řadu technologií, včetně umělé inteligence, strojového učení, cloud computingu a big data. Umělá inteligence a strojové učení se používají k automatizaci úkolů, jako je například doporučení ubytování hostům, kontrola kvality inzerátů a řešení sporů mezi hostiteli a hosty. Cloud computing umožňuje Airbnb škálovat svůj systém tak, aby mohl obsluhovat miliony uživatelů po celém světě. Big data se používají k analýze dat o používání platformy a k identifikaci trendů, které mohou pomoci zlepšit službu. Airbnb také investuje do vývoje nových technologií, které by mohly zlepšit zážitek uživatelů z platformy [15]. Například společnost implementovala umělou inteligenci, která pomáhá důrazně dodržovat pravidla oslav Nového roku ve všech částech světa [16].

Airbnb i další podobné online platformy zpřístupňují trh hostům a hostitelům za úplatu v podobě poplatků ve dvou různých strukturách. První struktura rozděluje poplatek mezi hosta a hostitele, viz obrázek 1. Hostitel je zde reprezentovaný červeným kolečkem vlevo. Ten zveřejňuje profil pro ubytovací jednotku na Airbnb a tím získává přístup k trhu. Inzerčí ubytování podléhá poplatku za služby ve výši 3 %, který se platí společnosti Airbnb. Ústřední postavou je Airbnb, zobrazená jako oranžový kruh. Ta zprostředkovává transakce mezi hosty a hostiteli. Je tedy spojena s oběma a naznačuje tok služeb a poplatků. Třetím činitelem je host, na obrázku vpravo v zeleném kruhu. Host díky Airbnb také získává přístup k trhu tím, že může procházet nabídky a žádat o ně. Když si ubytování pronajme, zaplatí poplatek za pronájem spolu s poplatkem za služby, který se uvádí ve výši 6 až 12 % z celkových nákladů. Druhá struktura účtuje poplatky pouze hostiteli a je povinná pro tradiční pohostinské nabídky (např. hotely) a také pro hostitele připojené na software třetích stran. Typicky se poplatek pohybuje mezi 14 až 16 % z celkové částky rezervace. [17]

Obrázek 1 Zpeněžení na platformě Airbnb



Zdroj:[18]

3.4.4 Rizika pro hotelový průmysl

Fenomén sdílené ekonomiky představuje hrozbu pro tradiční hotelový průmysl, protože nabízí levnější ubytování v mnoha oblastech. Službu pronájmu apartmánu, místnosti nebo domu, může díky platformám jako Airbnb nabízet opravdu kdokoli. S provozem se nepojí žádné vysoké náklady, což je jednoznačná výhoda oproti hotelům. Hotely musí platit za udržování celkové nemovitosti, za zaměstnance a mnoho dalších služeb. Další

výhodou sdílené ekonomiky je větší rozmanitost nabídky. Hotely poskytují standardizované služby a ubytování. Platformy sdílené ekonomiky nabízejí širokou škálu možností, od levných pokojů v soukromých bytech až po luxusní domy s vlastním bazénem. Sdílená ekonomika může také změnit preference zákazníků. Hosté jsou stále více ochotni ubytovat se v soukromých bytech nebo domech, protože jim to poskytuje větší soukromí a flexibilitu. [11]

3.5 Koncept cenotvorby

Tato kapitola pojednává o obecných principech cenotvorby a také ji specifikuje pro zkoumaný sektor podnikání. Podkapitola se zabývá klíčovými aspekty, které formují cenotvorbu na Airbnb, s cílem poskytnout ucelený pohled na to, jak hostitelé určují ceny za ubytování, a jak tyto ceny ovlivňují ostatní činitele.

3.5.1 Obecná tvorba cen

Cenotvorba je proces, v němž si firma nebo podnikatel určuje částku, kterou bude zákazníkům účtovat za své produkty či služby. Při stanovování cen je zapotřebí zohlednit několik faktorů, které mají na tento proces vliv. Mezi faktory, které mají vliv na stanovení ceny patří strategie cenotvorby, míra zájmu o produkt, tržní podmínky a aktivity konkurence, stejně jako výše výrobních nákladů. Přesto ve většině průmyslových odvětví není cena stanovována subjektem samostatně. Místo toho je určována dynamikou konkurenčního trhu, která se odráží v kupní a prodejní síle. Tímto způsobem fungují všechny trhy s nabídkou identických produktů jako obilí, ropa, minerály. V ostatních případech si firma nastavuje částku podle vlastních preferencí. Většinou se jedná o prodej unikátního produktu nebo jedinečného zážitku z nákupu. Firmy využívají svého jména, které je uznávané ve společnosti, a tak mají možnost si za prodej účtovat více než obecné, přesto mnohdy identické produkty. [19]

Cíle cenotvorby se mohou napříč různými odvětvími lišit, ale obecně se firma snaží pokrýt náklady z výroby a dosáhnout zisku. Mezi další cíle může patřit získání určitého objemu trhu, naplnění kvóty prodaných produktů apod. Avšak pro správnou cenotvorbu musí mít firma pochopení pro jejich specifický trh a pro své zákazníky. [19]

3.5.2 Specifika cenotvorby apartmánového bydlení na online platformách

Vytváření cen na platformě Airbnb je proces ovlivněný mnoha faktory, od charakteristik nemovitosti po lokální tržní dynamiku a regulace. Hostitelé na Airbnb často využívají dynamické cenotvorby, aby maximalizovali své příjmy a zároveň zůstali konkurenceschopní. Tato strategie zahrnuje pružné nastavení cen na základě algoritmicke analýzy tržních dat, poptávky a sezónních trendů. Hostitelé také mohou nabízet slevy pro dlouhodobější pobyty nebo pro rezervace na poslední chvíli, aby zvýšili obsazenost. Také bylo dokázáno, že profesionální hostitelé na Airbnb používají více dynamické cenové modely na rozdíl od běžného spotřebitele (hostitele) inzerujícího např. prázdný byt pro přívůdek. Právě soutěž dvou naprosto odlišných skupin hostitelů způsobuje různé výsledky při výzkumu cenotvorby na těchto online platformách. Běžní spotřebitelé čelí úkolu nastavení správných cen navzdory chybějícím odborným znalostem a stále se zvyšujícímu počtu konkurence ve formě dalších inzerátů. [20]

Mezi faktory ovlivňující ceny na Airbnb patří lokalita, charakteristika nemovitosti či sezónnost a poptávka a také specifické atributy inzerátu, jako možnost instantní rezervace, storno podmínek, statusu hostitele a další. Předchozí studie zkoumají efekty cenotvorby na širokém rozsahu atributů, přičemž jejich výsledky se mnohdy liší. Například výzkum prováděný ve čtyřech španělských městech dokázal vyšší cenu apartmánů umístěných ve větší vzdálenosti od turistických atrakcí a také vyšší cenu pro apartmány, které jsou blíže umístěny pobřeží moře. Je to jeden z mála výzkumů, jehož výsledky dokazují vyšší cenu, čím dále je ubytování od turistických atrakcí. [21]

Jiné výzkumy ukázaly nižší cenu pro inzeráty nabízející možnost instantní rezervace či flexibilní storno podmínky. Další skupina našla pozitivní vztah mezi cenou a velikostí ubytování a také dostupnost některého vybavení, jako například bezplatné parkování. [20]

Skutečnost, že lidé nabízejí prázdné nemovitosti na Airbnb a dalších podobných platformách má také důsledky pro trh s nemovitostmi. Konkrétněji studie prováděné v Londýně ukazují větší vliv na cenu prodávaných nemovitostí nežli na cenu měsíčních nájmu. Tento efekt je způsobený, tím že se stále více majitelé domů či bytů soustředí na turistický trh, čímž nabídka nemovitostí klesá. Lokalita a přítomnost mnoha Airbnb apartmánů soustředěných v jedné oblasti zvedá hodnotu nemovitosti. Avšak nájemci bytů velmi často nemají možnost nemovitost podnajímt a tím vstoupit do turistického trhu, tudíž zažívají zvýšení cen nájmu kvůli majitelům, kteří nabídku omezují ve prospěch turistiky. [22]

3.6 Automatizovaný sběr dat z online platformem a zpracování jeho výsledků

V této kapitole je popsán proces sběru dat z online platformem a jejich následné zpracování. Sběr dat z webových stránek je možné realizovat řadou různých způsobů, avšak data nasbíraná pro tuto práci se nasbírala pomocí procesu zvaného web scraping.

3.6.1 Sběr dat

Web scraping je proces extrakce a kombinace informací z internetu v organizované podobě. Je často používán pro sběr dat pro výzkum nebo analýzu. V tomto procesu figuruje software také známý jako web robot. Má za úkol procházet webové stránky podobným způsobem jako to dělá člověk. Tedy navštívuje řadu webových stránek, přistupuje k nim přes uživatelské rozhraní a ze stránek sbírá data. Jedná se o jednu z nejstarších metod extrakce dat z webových platformem. Využívá se při nedostupnosti jiných, modernějších technik pro sběr dat, jako je například API. [23][24]

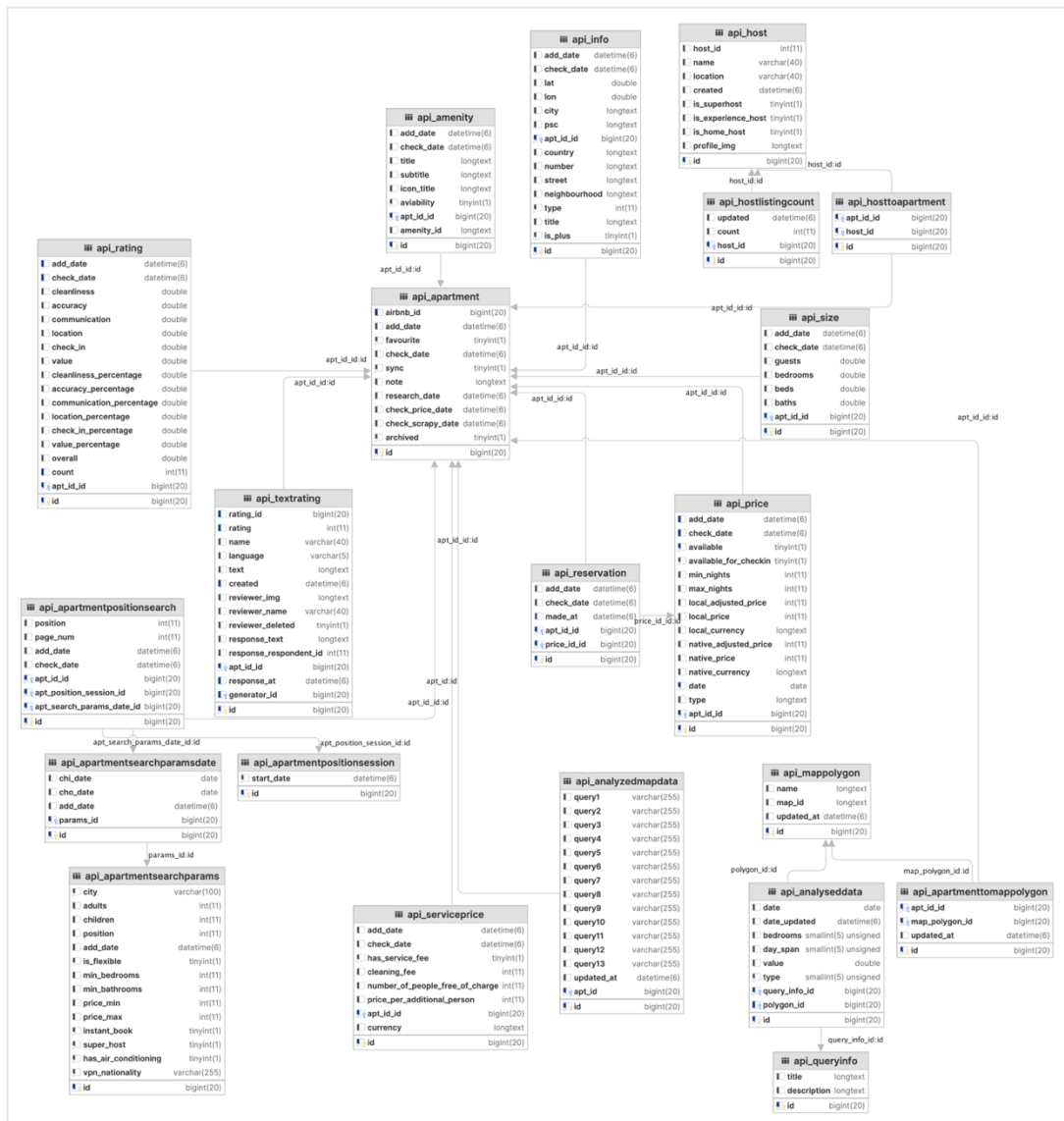
Existují tři hlavní kategorie softwaru pro web scraping: knihovny, frameworky a desktopová prostředí. Knihovny jsou nejběžnějším přístupem k vytváření webových scraperů. Poskytují základní funkčnost pro přístup na webové stránky a extrahování HTML a dat. Mezi populární knihovny patří libcurl, Apache HttpClient, jsoup a další. Frameworky jsou komplexnější než knihovny a poskytují vyšší úroveň abstrakce pro vytváření webových data-scrapers. Obvykle zahrnují komplexní funkce pro více integrované řešení definice robotů. Některé populární frameworky zahrnují Scrapy, Web-Harvest a jARVEST. Desktopová prostředí se obvykle používají pro scraping webových stránek, které vyžadují grafické uživatelské rozhraní (GUI). Poskytují uživatelsky přívětivé rozhraní pro definování a spouštění scrapů. Některá populární desktopová prostředí zahrnují Octoparse a Web Scraper. Výběr softwaru pro webový scraping závisí na konkrétních potřebách uživatele. Knihovny jsou vhodným řešením pro jednoduché úlohy scrapingu, zatímco frameworky jsou dobrou volbou pro větší nebo složitější projekty. Desktopová prostředí nejlépe slouží pro scraping webových stránek, které vyžadují GUI. [23]

API a web scraping jsou dva odlišné způsoby pro sběr informací z webových stránek. API je možné považovat za oficiální způsob, jak získávat data od poskytovatele webových stránek, zatímco web scraping je neformální způsob používaný při absenci API. V případě platformy Airbnb není API veřejně dostupné, takže web scraping bylo jedinou možností,

jak získat požadovaná data. Jelikož jsou údaje ze stránky Airbnb veřejně dostupné (stránku může navštívit kdokoli), je legální je tímto způsobem sbírat, avšak porušují se tím podmínky služby Airbnb. [25]

Pro extrakci požadovaných dat je nejprve nutné vytvořit webového robota, který prochází webové stránky. Web robot, který shromažďuje data analyzovaná v této práci, je napsán v programovacím jazyce Python. K procházení webových stránek je použit framework Scrapy, který umožňuje automatizaci interakce s webovými prohlížeči. Robot je naprogramován tak, aby opakovaně načítal webové stránky a pečlivě zachycoval a zaznamenával každý apartmán, který se objevil ve výsledcích vyhledávání. Získával podrobné informace o každé nabídce, včetně kalendáře dostupnosti, vybavení a polohy. Webový robot navíc shromažďoval komplexní údaje o různých attributech každého bytu, jeho hostitele, čímž zajistil široký soubor dat. Všechna shromážděná data jsou systematicky uspořádána a uložena ve strukturované databázi, což představuje komplexní úložiště informací pro pozdější analýzu a studium. Schéma databáze poskytnuté pro účely zpracování této práce je zobrazeno na obrázku 2. Pro analýzu průměrné ceny, zpracovanou v praktické části práce, se používají tabulky `api_apartment`, `api_price`, které obsahují apartmány a jejich ceny. Tabulky `api_size`, `api_info`, `api_mappolygon` a asociační tabulka `api_apartmenttomappolygon` obsahují další údaje o apartmánech. Vícenásobná regresní analýza prvků používá stejné tabulky spolu s ostatními, které jsou ve schématu.

Obrázek 2 Schéma relační databáze získaných dat z Airbnb



Zdroj: vlastní zpracování na podkladě dat od firmy

Web scraper nejdříve prochází výsledky vyhledávání na Airbnb. Z jednotlivých „karet“ reprezentujících inzerát extrahuje název, pozici ve vyhledávání a jeho URL adresu. Obrázek 3 zachycuje příkladnou část zdrojového kódu stránky Airbnb, kde jsou tyto atributy obsaženy. Zdrojový kód je z domovské stránky Airbnb, na které jsou obsaženy inzeráty apartmánů. Ve zdrojovém kódu jsou značky meta obsahující atributy jako URL adresu, název apartmánu či pozici ve vyhledávání. Tímto způsobem se rozšiřuje objem sledovaných apartmánů a ukládá se jejich pozice ve vyhledávání. Další scrapovací proces načítá jednotlivé apartmány z uložených URL adres a ukládá si jeho atributy. Oba tyto procesy se po určité době spouští opakovaně pro zajištění aktuálnosti dat. Podobně se získávají data o hostitelích a dalších sekcích portálu Airbnb.

Obrázek 3 Snímek zdrojového kódu HTML ze stránky Airbnb.com

```
▼ <div class=" dir dir-ltr"> == $0
  ▼ <div class="c4mnd7m atm_9s_11p5wf0 atm_dz_losqo2v dir
    dir-ltr"> grid
    ▶ <div class="c1pbo4kt atm_ar_1bp4okc atm_d2_1kqhmmj a
      tm_e2_losqo2v atm_gi_idpfg4 atm_lo_idpfg4 atm_vl_15v
      qwvr atm_k4_idpfg4 atm_wq_cs5v99 cikyho8 atm_9s_glyw
      fm atm_93_15vqwvr_hp14bx atm_9s_1txwivl_hp14bx dir
      dir-ltr" style="--transition-element_transition-dela
      y: 0ms; --transition-element_transition-duration: 20
      0ms;"> flex
    ▼ <div class="c1l1h97y atm_d2_1kqhmmj dir dir-ltr"
      style="--transition-element_transition-delay: 0ms; -
      -transition-element_transition-duration: 200ms;">
      ▼ <div itemprop="itemListElement" itemscope
        itemtype="http://schema.org/ListItem">
          <meta itemprop="name" content="3FL Apt of KINGS
            + whirlpool in heart of Prague">
          <meta itemprop="position" content="1">
          <meta itemprop="url" content="www.airbnb.com/roo
            ms/18081993?adults=10&children=0&enable_m3_priva
            te_room=true&infants=0&pets=0&check_in=2024-02-0
            4&check_out=2024-02-09&source_impression_id=p3_1
            707050424_ZDMfqQMAo75f4eX&previous_page_section
            _name=1000">
          ▶ <div> </div>
          </div>
        </div>
      </div>
    </div>
  </div>
```

Zdroj: [26]

3.6.2 Metody zpracování nasbíraných dat

Tato podkapitola pojednává o některých částech skriptu používaného pro účely této práce. Obrázek 4 nabízí pohled do řešení analýzy průměrných cen. Kód v obrázku nejdříve vytváří „dataframe“ ze specifikovaného CSV souboru. Typ proměnné „dataframe“ umožňuje kódu pracovat s dvoudimenzionálním polem prvků strukturovaným do sloupců a řádků [27]. Sloupec „date“ je podroben transformaci z původního formátu string na datetime proměnnou, ze které je poté odvozován měsíc a rok každého záznamu. Z tohoto pole prvků jsou vybrána požadovaná data, v případě tohoto kódu je možné specifikovat počty ložnic a maximální sledovanou částku. Filtrované pole je poté zpracováno pro získání

průměrné hodnoty, mediánu, minimální a maximální hodnoty. Pro každý unikátní počet ložnic v poli je následně vytvořen soubor, se kterým se v jiné části kódu dále pracuje.

Obrázek 4 Část kódu pro analýzu průměrných cen

```
merged_df = pd.read_csv(filepath_or_buffer: 'source/merged_data.csv', low_memory=False)

merged_df['date'] = pd.to_datetime(merged_df['date'])
merged_df['month'] = merged_df['date'].dt.month
merged_df['year'] = merged_df['date'].dt.year

min_bedroom_count = 3
max_bedroom_count = 6
max_price = 1070

merged_df_2023 = merged_df[(merged_df['year'] == 2023) & (merged_df['bedrooms'] >= min_bedroom_count) & (merged_df['bedrooms'] <= max_bedroom_count) & (merged_df['local_price'] <= max_price)]

unique_bedrooms = merged_df_2023['bedrooms'].unique()
unique_months = merged_df_2023['month'].unique()

analysis_results = pd.DataFrame(columns=['Bedrooms', 'Month', 'Mean_Price', 'Median_Price', 'Min_Price', 'Max_Price'])

for bedrooms in unique_bedrooms:
    for month in unique_months:
        subset_df = merged_df_2023[(merged_df_2023['bedrooms'] == bedrooms) & (merged_df_2023['month'] == month)]

        groupby_results = subset_df.groupby('map_polygon_id')['local_price'].agg(['mean', 'median', 'min', 'max']).reset_index()

        groupby_results.columns = ['Map_Polygon_ID', 'Mean_Price', 'Median_Price', 'Min_Price', 'Max_Price']

        groupby_results['Bedrooms'] = bedrooms
        groupby_results['Month'] = month

        analysis_results = pd.concat(objs: [analysis_results, groupby_results], ignore_index=True)

filename = 'output/price_analysis_by_bedrooms_and_monthBR' + str(min_bedroom_count) + '-' + str(max_bedroom_count) + '_price' + str(max_price) + '.csv'
analysis_results.to_csv(filename, index=False)
```

Zdroj: vlastní zpracování

Obrázek 5 obsahuje úryvek kódu používaného pro analýzu vlastností apartmánu a jejich vlivu na cenu. Kód nejprve načítá zkoumaný soubor dat a filtruje z něj nežádané prvky. Z filtrovaného souboru dat jsou specifikovány proměnné na horizontální a vertikální ose, které se poté předávají funkci pro regresní analýzu metodou OLS. Třída „OLS“ je součástí Python knihovny „statsmodels“. Metoda je v kódu nasazena podle dokumentace knihovny „statsmodels“. Parametry třídy jsou následující [28]:

- endog – závisle proměnná
- exog – nezávisle proměnná/proměnné
- missing – příkaz, jak zacházet s chybějícími hodnotami
- hasconst – informace, jestli se funkci předává atribut konstanty

V úryvku kódu se třídě předávají parametry „Y_filtered“ (závisle proměnná) a „X_filtered“ (nezávisle proměnné). Ty jsou poté využity v metodě „fit“ pro samotné nasazení.

Na konci úryvku kódu se vypisuje přehled výsledků analýzy.

Obrázek 5 Část kódu pro regresní analýzu

```
outliers = pd.read_csv("outliers.csv")

f_params = params[~large_residuals]
filtered_params = f_params[~f_params['id'].isin(outliers['id'])]

X_filtered = filtered_params[['Bedrooms', 'Superhost', 'Rating', 'AmenitiesCount',
    'ImageCount', 'Kitchen', 'Dryer', 'Fireplace',
    'Jacuzzi', 'SelfCHI', 'FreeParking', 'Pool',
    'AC', 'TV', 'Washer', 'WiFi']]
X_filtered.loc[:, "Rating"] = X_filtered["Rating"] * X_filtered["Rating"]
Y_filtered = filtered_params['AveragePriceOfFreeNights']

X_filtered = sm.add_constant(X_filtered)
model_filtered = sm.OLS(Y_filtered, X_filtered)
results_filtered = model_filtered.fit(cov_type='HC3')

out_filtered = results_filtered.summary()
print(out_filtered)
```

Zdroj: vlastní zpracování

Předešlý obrázek přináší pohled na provedení regresní analýzy filtrovaného souboru dat. Následující obrázek (Obrázek 6) tvoří část kódu, ve které k filtraci dochází. Soubor dat je stejně jako v ostatních částech kódu uložen do typu proměnné „dataframe“, dostupného z knihovny „pandas“. Z výsledků provedené regresní analýzy kód vyhodnocuje standardizovaná rezidua pro identifikaci velkých reziduí. Tento proces probíhá filtrací těch reziduí, jejichž hodnota je vyšší než dva. S využitím funkce „get_influence“ je určena Cookova vzdálenost, vybírají se hodnoty vyšší než jedna. Výsledné body se poté ukládají do proměnné „residuals_df“ a exportují se do CSV souboru. Kód poté pokračuje další regresní analýzou pro průzkum provedených změn.

Obrázek 6 Část kódu pro filtraci nevhodných dat

```
influence = results.get_influence()
standardized_residuals = influence.resid_studentized_internal

large_residuals = np.abs(standardized_residuals) > 2

cooks_d, pvals = influence.cooks_distance
influential_points = cooks_d > 1

residuals_df = pd.DataFrame({
    'Residuals': results.resid,
    'Standardized Residuals': standardized_residuals,
    'Cooks Distance': cooks_d
})
residuals_df['Large Residual'] = large_residuals
residuals_df['Influential'] = influential_points

residuals_df.to_csv(path_or_buf="residuals_analysis.csv", index=False)

most_influential_index = residuals_df['Cooks Distance'].idxmax()
```

Zdroj: vlastní zpracování

3.7 Metoda nejmenších čtverců

Regrese metodou nejmenších čtverců (Ordinary Least Squares regression), běžně označovaná zkratkou OLS, je široce používanou statistickou metodou pro odhad koeficientů v lineárních regresních modelech. Je používána pro popsání vztahu mezi závisle proměnnou a jednou nebo více nezávisle proměnnými. Podstata metody OLS spočívá v minimalizaci součtu čtvercových rozdílů mezi pozorovanými a předpovídanými hodnotami, který je znám jako součet čtvercových chyb. [29]

Pro účely této práce je model OLS používán pro analýzu vybraných vlastností a atributů apartmánu a jejich vliv na cenu. Lineární regrese, a tím i OLS, nachází uplatnění v různých oblastech. Například meteorologové ji používají k předpovídání prvků počasí, jako je teplota nebo srážky, na základě vnějších proměnných. V biologii může předpovídat populaci druhů na základě faktorů, jako je počet predátorů nebo dostupné přírodní zdroje. V oblasti ekonomie se kromě mnoha jiných aplikací používá k předpovídání příjmů společnosti na základě údajů o prodeji. [29]

Při provádění lineárního regresního modelu pomocí OLS musí být splněno několik předpokladů, včetně nezávislosti pozorování, homogenity rozptylu a normálního rozdělení reziduí. Tyto předpoklady jsou rozhodující pro platnost závěrů modelu. [29]

3.8 Shrnutí

Teoretická část práce zkoumá klíčové aspekty informatiky a systémů se zaměřením na informační systémy, uživatelská rozhraní, a API. Tato část zahrnuje teoretickou analýzu konceptů, jako jsou systémy, informační systémy, a jejich význam v kontextu informatiky. Důraz je kladen na význam a funkci uživatelských rozhraní a API v moderních technologických aplikacích. Teoretická část se dále zabývá tvorbou cen a specifikami online platforem pro pronájem apartmánů, zejména v kontextu sdílené ekonomiky. Metody a techniky, jako je automatizovaný sběr dat z online platforem a metoda nejmenších čtverců, jsou prozkoumány pro porozumění principu a využití v praktických aplikacích.

Práce se dále detailně zaměřuje na shrnutí získaných dat a jejich analýzu. Nejprve je popsán zkoumaný soubor dat, jeho rozdělení, objem a uložené hodnoty. Následně je provedena analýza průměrné ceny za noc v jednotlivých regionech České republiky a v různých počtech pokojů. Práce také zkoumá vybrané vlastnosti nabídek apartmánů a určuje jejich vliv na cenu za noc.

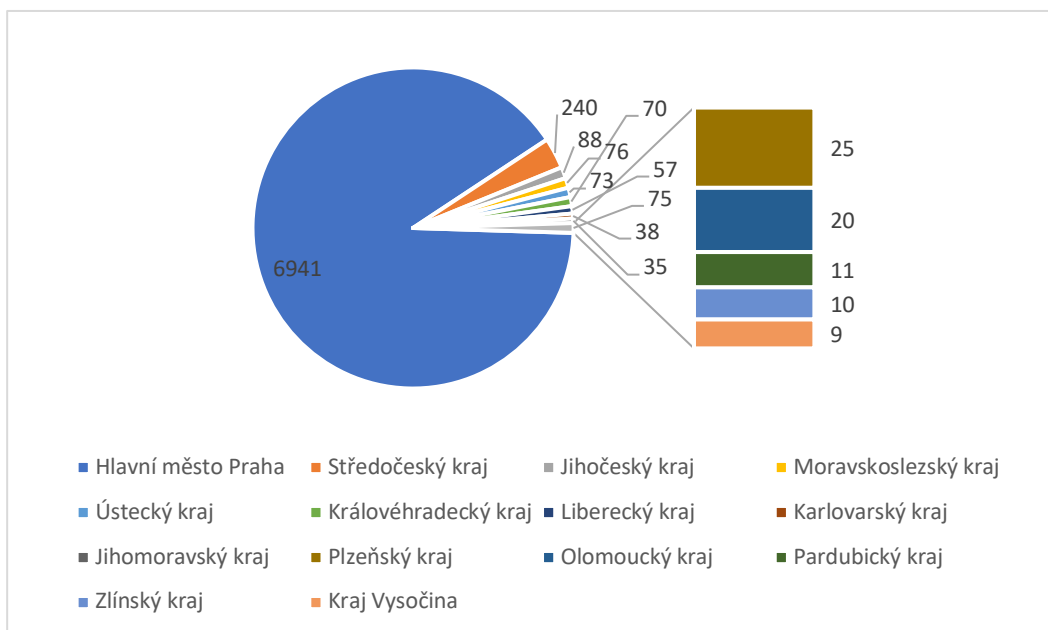
4 Vlastní práce

Ekonomická analýza je zaměřena především na podnikání v rámci apartmánového bydlení, a to z pohledu zhodnocení průměrných cen pro jednotlivé kraje České republiky a posouzení faktorů, které mohou ovlivňovat průměrnou cenu za noc. Výsledky těchto analýz jsou následně komparovány pro vybraný atribut. Postup je realizován dle zkušeností z dobré praxe v podnikání v sektoru apartmánového bydlení. Vlastnosti jsou vybrány na základě popularity u hostů, jak je popsáno na straně 36. Nejedná se o klasickou ekonomickou analýzu, nýbrž o analýzu hlavních prvků v podnikání v rámci krátkodobého ubytování v apartmánovém bydlení na zvolené platformě v České republice.

4.1 Selekcce dat

Data poskytnutá pro účely této práce jsou výsledkem sběru dat od firmy, která se zabývá podnikáním v oblasti apartmánového bydlení v Praze. Vzhledem k tomu sběr dat soustředila zejména na odpovídající oblast. Graf 1 ukazuje absolutní počet zkoumaných apartmánů napříč kraji České republiky. Kraje jsou seřazené podle počtu apartmánů. Je patrné, že naprostá většina z nich je umístěna v Praze, v procentuální hodnotě je podíl pražských apartmánů vůči ostatním 90,2 %. V absolutní hodnotě je v Praze 6 941 apartmánů, na druhém místě s nejvyšším počtem apartmánů je Středočeský kraj s 240 sledovanými apartmány. Větší počet apartmánů v ostatních krajích České republiky nebylo možné v požadované době sledovat.

Graf 1 Počet apartmánů v krajích ČR



Zdroj: vlastní zpracování

Databáze výsledků sběru dat poskytuje rozsah sledovaných aspektů v časovém období od roku 2022 až do roku 2025. Analýza průměrných cen za noc pro kraje České republiky zpracovává informace o umístění, velikosti a cenách apartmánů.

4.2 Analýza průměrné ceny v krajích České republiky

Tato kapitola se věnuje zhodnocení průměrné ceny za noc pro jednotlivé kraje České republiky pro rok 2023. Nejprve se kapitola zabývá rozdělením apartmánů, poté druhem ceny vhodného pro analýzu, následně jsou zpracovány analýzy průměrné ceny pro jednotlivé kategorie.

4.2.1 Rozdělení apartmánů

V rámci celé České republiky je největší počet sledovaných apartmánů v Praze. Tím je Praha v rámci analýzy poměrně znevýhodněna, jelikož obsahuje mnoho různých cenových tříd apartmánů. Z toho důvodu jsou v analýze průměrných nocí apartmány rozděleny nejen podle krajů České republiky ale také podle počtu ložnic. Počty apartmánů pro příslušný počet ložnic jsou znázorněny v tabulce 1. Apartmány s jednou ložnicí mají nejvyšší zastoupení a tvoří 67,1 % celého zkoumaného souboru, u ostatních kategorií s vyšším počtem ložnic počet apartmánů vždy klesá. Proto je analýza soustředěna na dvě nejpočetnější skupiny, jedno a dvou ložnicové apartmány. Ostatní skupiny

jsou následně také popsány, avšak do menší hloubky. Počty ložnic se v průběhu roku pro jednotlivé apartmány mění, a proto je pro zjednodušení a zachování kategorizace apartmánů zvolen první zaznamenaný počet ložnic.

Tabulka 1 Rozdělení apartmánů podle počtu ložnic

Počet ložnic	Počet apartmánů v kategorii
1	4 934
2	1 611
3	534
4	157
5	87
6	35

Zdroj: vlastní zpracování

4.2.2 Výběr vhodné ceny

Z dostupného souboru dat je možné analyzovat dva typy cen – cena za prodanou noc a cena za volnou noc. Pro výběr té správné je provedena komparativní analýza počtu prodaných a počtu volných nocí, viz tabulka 2. Ve většině případů chybí více záznamů pro volné noci než prodané, chybějící prodané noci téměř vždy činí 15 specifických datumů. Celkový počet chybějících prodaných nocí je menší a objevuje se pouze v Pardubickém kraji a v kraji Vysočina, které ve stejném pořadí dosahují hodnot 26 a 4. Na základě této analýzy je cena prodaných nocí vybrána jako sledovaný parametr.

Některé apartmány mají více záznamů pro jeden konkrétní den, jelikož se soubor dat aktualizoval a kontroloval změnu cen či dostupnost tohoto data. Z toho důvodu je také nutné vybrat vyhovující záznam. Jelikož se některé záznamy mění i přes to, že nejsou dostupné k rezervaci, je vybrán první záznam ceny nedostupné (prodané) noci, který je k dispozici. Tímto způsobem je zajištěna hodnota nejvíce odpovídající ceně, za kterou je noc prodaná.

Ceny ubytování jsou uvedeny v různých měnách. Kromě českých korun, které dominují, se ceny u některých apartmánů objevují v eurech, amerických dolarech a dalších měnách. Databáze zachycuje všechny tyto údaje a pro zjednodušení srovnání a dalšího zpracování je při sběru dat nutné všechny ceny převést na jednotnou měnu. Tato jednotná referenční měna je americký dolar. Převod cen do USD je realizován pomocí směnného kurzu platného vždy ke dni provedení sběru dat.

Při prvotním průzkumu a analýze je zjištěno, že soubor cen obsahuje hodnoty mimo normální rozdělení, což vedlo ke zkreslení výsledků. K určení horní hranice výše ceny

za noc je použit statistický percentilový přístup. Jeho použitím jsou stanoveny dvě hranice: 95. percentil ve výši 443 USD a 99. percentil ve výši 1 070 USD. Na základě těchto hodnot je pro apartmány s jednou a dvěma ložnicemi přijata konzervativnější hodnota horní hranice na úrovni 95. percentilu, což odpovídá ceně 443 USD za noc. Pro druhou skupinu tří až šesti ložnic se přijatá hranice rovná hodnotě 99. percentilu odpovídající 1 070 USD. Z osobní zkušenosti v podnikání v tomto sektoru se dá výskyt extrémních částek vysvětlit tím, že hostitelé v některých případech nastaví neúměrně vysokou částku pro zamezení prodeje konkrétních nocí, tento způsob slouží jako alternativa k jejich zablokování.

Tabulka 2 Počty chybějících hodnot pro typy cen

Jméno kraje	Počet chybějících volných nocí	Počet chybějících prodaných nocí
Středočeský kraj	1	0
Jihočeský kraj	14	0
Plzeňský kraj	15	0
Karlovarský kraj	15	0
Ústecký kraj	15	0
Liberecký kraj	15	0
Královéhradecký kraj	0	0
Pardubický kraj	15	26
Kraj Vysočina	15	4
Jihomoravský kraj	15	0
Olomoucký kraj	15	0
Moravskoslezský kraj	13	0
Zlínský kraj	15	0
Hlavní město Praha	0	0

Zdroj vlastní zpracování

4.2.3 Průměrná cena apartmánů s 1 ložnicí

První analyzovaná kategorie zkoumá apartmány s jednou ložnicí. Výsledky analýzy jsou vyobrazené v grafu 2, ve kterém horizontální osa reprezentuje jednotlivé měsíce roku 2023 a osa vertikální je tvořena výší průměrné finanční částky za noc v amerických dolarech.

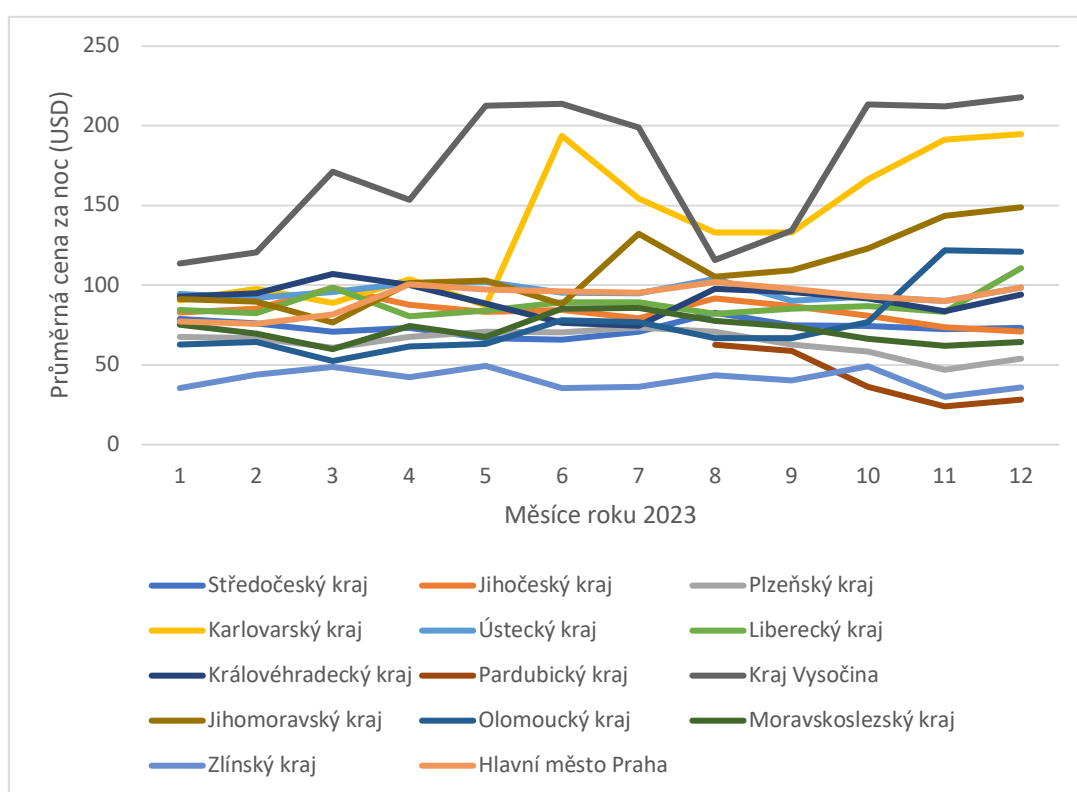
Nejvyšších hodnot dosahuje kraj Vysočina, jeho křivka je umístěna nad všemi ostatními. V polovině a na konci roku dosahuje nejvyšších průměrných hodnot, ty se pohybují okolo 200 USD za noc. Počet zkoumaných apartmánů umístěných

na Vysočině je však velmi malý, tudíž i jeden samotný apartmán s vyššími cenami může zkreslit výsledky průměrných cen.

V druhé polovině roku má, hned po Vysočině, nejvyšší ceny Karlovarský kraj, vrcholící v červnu v době konání mezinárodního filmového festivalu. V červnu je průměrná zaznamenaná cena za noc 193 USD, cena z tohoto bodu následně klesá až do září, kdy naopak znovu začíná růst a vrcholí v prosinci na úrovni 194 USD. Jihomoravský kraj vykazuje podobný trend v menší míře a začínající v červenci. Ostatní kraje se pohybují bez jasného trendu s výjimkou Prahy, kde je možné sledovat mírný náznak trendu. V prvním čtvrtletí jsou ceny nižší a postupně se zvedají s vrcholy v létě a v prosinci.

Výsledek průměrné ceny v kraji Vysočina je velmi pravděpodobně zkreslený kvůli jednomu či více apartmánům operujících ve vyšší cenové hladině. Proto je vhodné výsledky doplnit také o další hodnoty proměnných z tohoto vzorku, konkrétně o medián a maximální hodnotu za sledované období.

Graf 2 Průměrná cena za noc v krajích ČR pro apartmány s 1 ložnicí (USD)



Zdroj: vlastní zpracování

Pro lepší pochopení výsledků analýzy průměrných cen je zpracován medián, viz graf 1 v přílohách (Příloha A). Horizontální osa představuje měsíce roku 2023, vertikální osa hodnotu mediánu, ceny jsou také uvedeny v amerických dolarech.

Zatímco všechny křivky krajů se podobají křivkám v grafu průměrných cen, Vysočina se vymyká. V lednu medián v kraji Vysočina dosahuje hodnoty 113, poté stoupá na hodnotu 212 v únoru a tuto hodnotu vykazuje konstantně až do konce roku. Obecně je možné vyvodit, že v tomto regionu je ve sledovaném období několik extrémně položených hodnot, jež ovlivňují průměrnou cenu, a tak je medián stabilnějším ukazatelem zkoumaných cen. I přesto je tento kraj na vertikální ose umístěn nejvýše. To může naznačovat výskyt ubytování s vysokou cenovou hladinou a tím, že je na Vysočině nejmenší počet zkoumaných apartmánů, vyšší cenová hladina jednoho je schopna ovlivnit výsledky celého kraje.

Graf 2 v přílohách (Příloha A) poskytuje přehled maximální zaznamenané ceny apartmánů s 1 ložnicí v jednotlivých měsících roku 2023 v rámci České republiky. Vertikální osa zobrazuje maximální cenu v USD a horizontální osa měsíce roku 2023. Tento graf doplňuje analýzu průměrné ceny a zdůrazňuje výjimky. Hlavní město Praha dosahuje nejvyšších maximálních cen, a to konstantní hodnoty 443 USD za noc, která odpovídá horní hranici sledovaného cenového rozpětí. Ačkoli se průměrná cena a medián v Praze pohybují na srovnatelné úrovni s ostatními kraji, Praha se vyznačuje nabídkou ubytování za nejvyšší ceny. Zbytek České republiky sleduje výkyv cen v prosinci a lednu, kdy ceny mnohdy překonaly úroveň celého roku. V některých krajích se maximální cena vychyluje i v jiných měsících, například v Jihočeském kraji v červnu a v Jihomoravském kraji v květnu a říjnu. Tyto výkyvy však nikdy nedosahují horní hranice sledovaných cen, takže se jedná o skutečný nárůst cenové hladiny v daných měsících.

4.2.4 Průměrná cena apartmánů se 2 ložnicemi

Následující analyzovanou kategorií je ubytování obsahující dvě ložnice, jak je prezentováno na grafu 3 v přílohách (Příloha A). Horizontální osa zobrazuje jednotlivé měsíce roku 2023 a vertikální průměrnou cenu v USD. Ve srovnání s předešlou kategorií jsou ceny v této kategorii více homogenní a dochází k menším výkyvům. Je rovněž zřejmé, že tato skupina dosahuje vyšších průměrných cenových úrovní. Zatímco předchozí kategorie často zaznamenává cenové hladiny v rozmezí 50 až 100 USD, aktuálně zkoumaná skupina se umísťuje v horní polovině tohoto rozmezí a běžně tuto hranici přesahuje.

Nejnižších hodnot dosahuje Zlínský a Pardubický kraj, oba se pohybují pod hranicí 100 dolarů za noc, Zlínský kraj ke konci roku klesá až pod hodnotu 50 dolarů. Polovina oblastí zažívá nejnižší noc v listopadu, další v únoru či březnu ale také v letních měsících červnu a červenci, navzdory vyšší turistice v tomto období. Je pozoruhodné, že Liberecký

kraj, s nejnižší cenou právě v červenci, vykazuje nejvyšší hodnotu v listopadu, fakticky nejvyšší cenu celé České republiky za rok 2023 v této kategorii. Druhá nejvyšší cena je v květnu v Jihomoravském kraji s rozdílem dvou dolarů oproti listopadovému výkyvu Libereckého kraje. Na horní hranici cen se více než polovinu sledovaného období pohybuje Jižní Morava, po té většinu roku Praha a pozoruhodně také Ústecký kraj, jež se v předchozí kategorii pohybuje na nejnižší úrovni. Další rozdíl je možné sledovat u Karlovarského kraje, ve kterém po celý rok hodnota kolísá okolo 100 USD, kromě toho ani nevykazuje výkyv cen v červnu. Průměrné ceny a pozice krajů se tedy v porovnání s minulou kategorií liší, ani jeden kraj konzistentně nedosahuje v obou kategoriích nejvyšších výsledků.

4.2.5 Průměrná cena apartmánů se 3 až 6 ložnicemi

Jak již bylo zmíněno, s rostoucím počtem ložnic apartmánu klesal jejich sledovaný počet, z tohoto důvodu jsou zbývající skupiny sloučeny do jedné. Navzdory tomu je rovněž provedena analýza s dodržáním předchozího rozdělení (tedy vždy pro konkrétní počet ložnic), která ukazuje, že některé kraje se ve výsledcích analýzy vůbec nevyskytují, proto je zvolen tento, vhodnější, přístup.

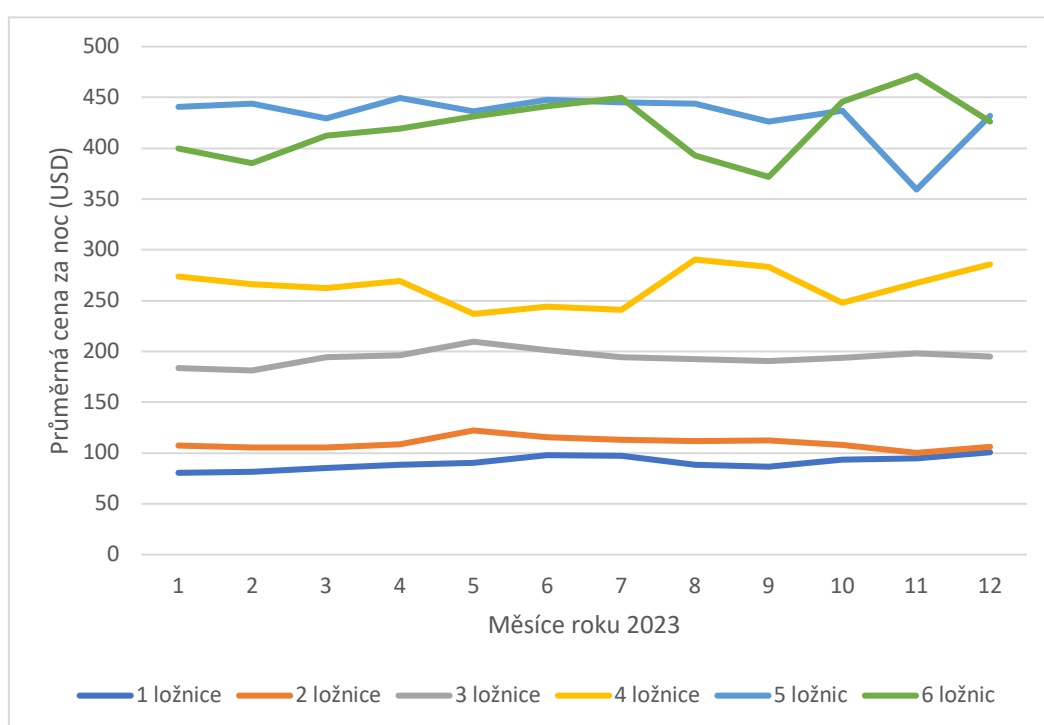
Graf 4 v přílohách (Příloha A) tedy zobrazuje průměrnou cenu za noc pro apartmány se třemi až šesti ložnicemi, na horizontální ose tohoto grafu leží měsíce roku 2023 a vertikální osa představuje cenu v USD. Podobně jako tomu je u apartmánů s jednou ložnicí, i pro tuto širší kategorii je jednou z nejnvýše položených křivek kraj Vysočina. Tento jev je opět způsoben nízkým počtem sledovaných apartmánů, který má vliv na výsledky analýzy. Na podobné hladině se pohybuje také Pardubický kraj, jehož celkový počet sledovaných apartmánů (bez ohledu na počet ložnic) se rovná 11, jak je znázorněno na grafu 1.

V krajích České republiky pro tuto skupinu není sledován jasný trend v průměrné cenové hladině, dochází zde k náhodným výkyvům v ceně. V několika případech se vyskytuje rostoucí i klesající výkyv v jednom měsíci pro různé kraje, jako například v červnu. Plzeňský a Jihomoravský kraj vykazují pokles průměrné ceny, zatímco Zlínský a Olomoucký kraj naopak růst. Většina krajů se pohybuje nad hranicí 200 USD s některými výjimkami.

4.2.6 Průměrná cena za noc pro jednotlivé typy apartmánů

Graf 3 poskytuje informace o vývoji průměrných cen podle rozdělení apartmánů na unikátní počty ložnic. Skupiny jedno a dvou ložnicových apartmánů operují s průměrným rozdílem 19,3 USD cenové hladiny, rozdíl dvou a tří ložnicových je několikrát větší a rovná se 84,6 USD. Skupina tří a čtyř ložnicových apartmánů má mezi sebou menší rozdíl, v průměru 69,7 USD. Skupiny s největším počtem ložnic, pěti a šesti ložnicové apartmány, se během roku pohybují poměrně podobně, avšak skupina s pěti ložnicemi tu vyšší převyšuje průměrně o 12 USD.

Graf 3 Průměrná cena za noc pro typy apartmánů (USD)



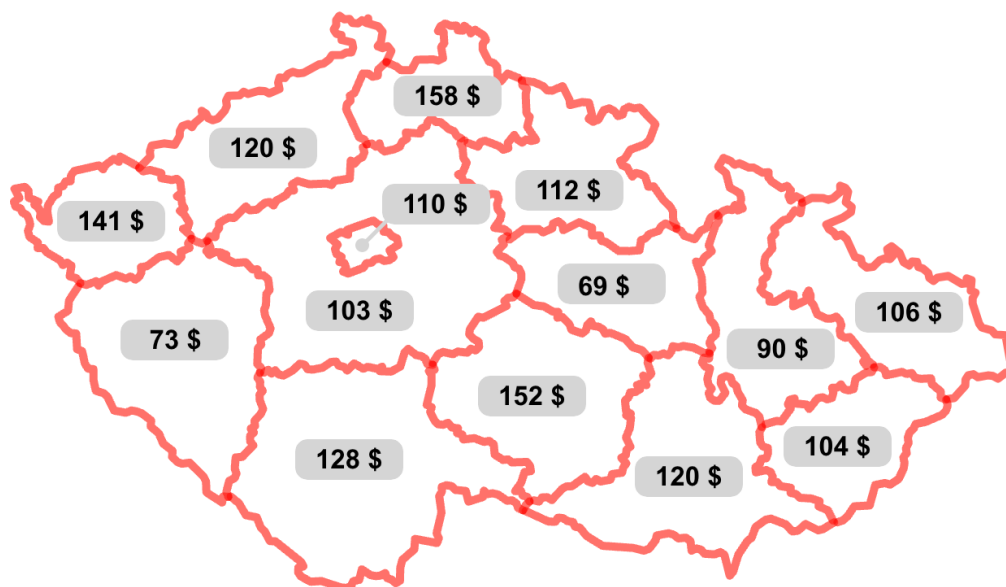
Zdroj: vlastní zpracování

4.2.7 Mapa České republiky a průměrná cena v krajích

Pro zakončení kapitoly průměrné ceny je vytvořena mapa České republiky rozdělena na kraje s hodnotami průměrných nocí pro celý rok 2023, viz obrázek 7. Jedná se o kompletní soubor apartmánů od jedné ložnice po šest ložnic. Ceny jsou, stejně jako u předešlých grafů, uváděny v USD a jsou omezeny dříve stanovenou hranicí 443 USD. Nejvyšší ceny jsou dosaženy v Libereckém kraji, poté v kraji Vysočina a Karlovarském kraji v uvedeném pořadí. Nejnižší ceny jsou zaznamenány v Pardubickém, Plzeňském a Olomouckém kraji. Opět je však nutné podotknout, že výsledky analýzy pravděpodobně

nereflektují skutečnou situaci zcela výstižně kvůli výrazným rozdílům v objemu sledovaných apartmánů.

Obrázek 7 Mapa krajů ČR a jejich průměrná cena apartmánu za noc v roce 2023 (USD)



Zdroj: vlastní zpracování, [30]

4.3 Regresní analýza atributů a jejich vliv na cenu

Tato kapitola se věnuje analýze vlastností inzerovaných apartmánů na platformě Airbnb za použití regresní analýzy metodou OLS. Je zkoumán vliv jednotlivých vlastností na cenu, výsledky odhalují např. koeficient růstu průměrné ceny při zvýšení počtu ložnic o jednotku, zatímco jsou ostatní vlastnosti konstantní. Analýza se zaměřuje na celou Českou republiku, není rozdělena na kraje. Důvodem je nedostatek zkoumaných dat v ostatních krajích.

4.3.1 Výběr proměnných

Výběr vlastností, které se v této práci zkoumají, je ovlivněn několika faktory. Předpokladem pro výběr vlastností je jejich dostupnost v databázi. Vlastnosti jako hodnocení apartmánu, počet ložnic nebo lokalita jsou běžně považované za rozhodující faktory při nastavení cen. Specifické vybavení apartmánu může zvýšit prodej a zájem hostů, proto jsou zde zkoumány méně se vyskytující položky jako například vířivka, bazén či klimatizace. Ačkoli se vyskytují málo, hosté o ně mají zájem a řadí se mezi nejhledanější vybavení na Airbnb [31].

Tabulka 3 obsahuje informace o proměnných používaných v regresní analýze, jejich rozdělení do kategorií, název, typ a definici proměnné. Závisle proměnnou je průměrná cena volných nocí s názvem „AveragePricesOfFreeNights“. Tato hodnota je přímo obsažena v databázi, ze které prováděné analýzy čerpají data a je uvedena v USD. Hodnoty této proměnné jsou vytvořeny průměrem všech záznamů o cenách neprodaných nocí pro každý apartmán odděleně. Každý apartmán má tedy proměnnou „AveragePricesOfFreeNights“, která značí průměr všech cen apartmánu za různé časové období, zpravidla od roku 2022 do roku 2025. Nejedná se o časovou řadu, nýbrž o agregované ceny. Nezávisle proměnné jsou rozděleny do dvou kategorií, na „reklamní a hostitelské funkce“ a „charakteristika ubytování“, přičemž celkový počet sledovaných proměnných je 16.

Tabulka 3 Definice proměnných používaných ve vícenásobné regresní analýze

Kategorie	Jméno proměnné	Typ proměnné	Definice
Závisle proměnná	AveragePricesOfFreeNights	Kvantitativní	Průměrná cena za noc pro apartmán v USD
Reklamní a hostitelské funkce	Superhost	Kategoriální	Status superhostitele; 0 = není superhostitel; 1 = je superhostitel
	Rating	Kvantitativní	Hodnocení apartmánu v intervalu 0-5
	ImageCount	Kvantitativní	Počet fotografií ubytovací jednotky
Charakteristika ubytování	Bedrooms	Kategoriální	Počet ložnic v ubytovací jednotce
	AmenitiesCount	Kategoriální	Počet vybavení ubytovací jednotky
	Kitchen	Kategoriální	Vybavenost kuchyní; 0 = není dostupné; 1 = je dostupné
	Dryer	Kategoriální	Vybavenost sušičkou; 0 = není dostupné; 1 = je dostupné
	Fireplace	Kategoriální	Vybavenost krbem; 0 = není dostupné; 1 = je dostupné
	Jacuzzi	Kategoriální	Vybavenost vířivkou; 0 = není dostupné; 1 = je dostupné
	SelfCHI	Kategoriální	Možnost self-checkin; 0 = není dostupné; 1 = je dostupné
	FreeParking	Kategoriální	Přítomnost bezplatného parkování; 0 = není dostupné; 1 = je dostupné
	Pool	Kategoriální	Vybavenost bazénem; 0 = není dostupné; 1 = je dostupné
	AC	Kategoriální	Vybavenost klimatizací; 0 = není dostupné; 1 = je dostupné
	TV	Kategoriální	Přítomnost televize; 0 = není dostupné; 1 = je dostupné
Washer	Kategoriální	Vybavenost pračkou; 0 = není dostupné; 1 = je dostupné	
WiFi	Kategoriální	Přítomnost Wi-Fi internetového připojení; 0 = není dostupné; 1 = je dostupné	

Zdroj: vlastní zpracování

Tabulka 4 obsahuje výsledky popisné statistiky vybraných proměnných. U kvantitativních proměnných jsou uvedeny informace o průměrné, minimální a maximální hodnotě a o standardní odchylce. U kategoriálních proměnných je předložen počet unikátních hodnot a jejich frekvence v absolutním i relativním vyjádření.

Tabulka 4 Deskriptivní statistika proměnných

Proměnná	Kvantitativní proměnné				Kategoriální proměnné				
	Průměrná hodnota	Std. error	Min	Max	Počet unikátních hodnot	Frekvence	Frekvence (%)		
AveragePriceOfFree Nights	118,632	63,072	14,395	407,931		-			
Bedrooms		-			5	1,0	2562	1,0	73,03
						2,0	691	2,0	19,70
						3,0	189	3,0	5,39
						4,0	53	4,0	1,51
						5,0	13	5,0	0,37
Superhost		-			2	1,0	1863	1,0	53,11
						0,0	1645	0,0	46,89
Rating	4,751	0,233	2,0	5,0		-			
AmenitiesCount	37,932	10,452	13,0	60,0		-			
ImageCount	20,221	10,161	1,0	52,0		-			
Kitchen		-			2	1,0	3279	1,0	93,47
						0,0	229	0,0	6,53
Dryer		-			2	0,0	2717	0,0	77,45
						1,0	791	1,0	22,55
Fireplace		-			2	0,0	3325	0,0	94,78
						1,0	183	1,0	5,22
Jacuzzi		-			2	0,0	3420	0,0	97,49
						1,0	88	1,0	2,51
SelfCHI		-			2	1,0	1927	1,0	54,93
						0,0	1581	0,0	45,07
FreeParking		-			2	0,0	2709	0,0	77,22
						1,0	799	1,0	22,78
Pool		-			2	0,0	3454	0,0	98,46
						1,0	54	1,0	1,54
AC		-			2	0,0	2749	0,0	78,36
						1,0	759	1,0	21,64
TV		-			2	1,0	2789	1,0	79,5
						0,0	719	0,0	20,5
Washer		-			2	1,0	2459	1,0	70,1
						0,0	1049	0,0	29,9
WiFi		-			2	1,0	3462	1,0	98,69
						0,0	46	0,0	1,31

Zdroj: vlastní zpracování

Soubor dat je cíleně omezen vrchním limitem u vybraných proměnných, viz tabulka 5. Pro čtyři sledované proměnné jsou vypočteny 95. a 99. percentily, které poskytují informaci o tom, kolik procent hodnot se vyskytuje v určité hranici. Pro proměnné „AveragePricesOfFreeNights“, „Amenities Count“ a „ImageCount“ je přijata konzervativnější hodnota na úrovni 95. percentilu, 99. percentil pouze u proměnné „Bedrooms“.

Tabulka 5 Maximální hodnoty proměnných

Jméno proměnné	Maximální přijatá hodnota	Percentil
AveragePricesOfFreeNights	417.3	95
Bedrooms	5	99
AmenitiesCount	60	95
ImageCount	52	95

Zdroj: vlastní zpracování

Před provedením regresní analýzy je zkoumána vzájemná korelace vlastností, viz tabulka 6. Nejvíce korelované jsou proměnné „Superhost“ a „Rating“, koeficient se rovná hodnotě 0,426. V tomto případě je korelace předpokládána, jelikož podmínka, jak se stát tzv. superhostitelem na platformě Airbnb je vysoké hodnocení (>4,9). Dalšími páry s vyšší korelací jsou „Washer“ a „Dryer“ s hodnotou 0,325, či dvojice „Kitchen“ a „Washer“ dosahující 0,302. Žádné proměnné nejsou vysoce korelované.

Tabulka 6 Korelační tabulka proměnných

	Bedrooms	Superhost	Rating	AmenitiesCount	ImageCount	Kitchen	Dryer	Fireplace	Jacuzzi	SelfCHI	FreeParking	Pool	AC	TV	Washer
Superhost	-0,06														
Rating	0	0,426													
AmenitiesCo	0,08	0,284	0,27												
ImageCount	0,226	0,113	0,136	0,236											
Kitchen	0,094	0,065	0,122	0,22	0,108										
Dryer	0,032	0,061	0,066	0,089	0,061	0,101									
Fireplace	0,181	-0,02	0,087	0,077	0,117	0,04	0,006								
Jacuzzi	0,067	-0,048	-0,001	0,058	0,057	0,016	0,034	0,17							
SelfCHI	-0,065	0,071	-0,052	0,227	-0,017	0,019	0,025	-0,088	-0,063						
FreeParking	0,066	-0,049	0,102	0,084	0,061	0,01	0,013	0,247	0,091	-0,126					
Pool	0,084	-0,054	0,035	-0,004	0,072	-0,008	0,031	0,173	0,137	-0,056	0,233				
AC	-0,023	0,085	0,079	0,076	0,028	-0,111	0,039	-0,027	0,025	0,079	-0,085	-0,012			
TV	0,032	0,068	0,093	0,127	0,096	0,033	0,059	-0,063	0,015	0,064	-0,044	-0,008	0,123		
Washer	0,132	0,044	0,064	0,154	0,142	0,302	0,325	-0,041	0	-0,03	-0,04	-0,016	-0,118	0,079	
WiFi	-0,004	0,037	-0,021	0,077	0,02	0,055	0,056	-0,171	-0,04	0,071	-0,121	-0,04	0,038	0,153	0,109

Zdroj: vlastní zpracování

4.3.2 Distribuce reziduí

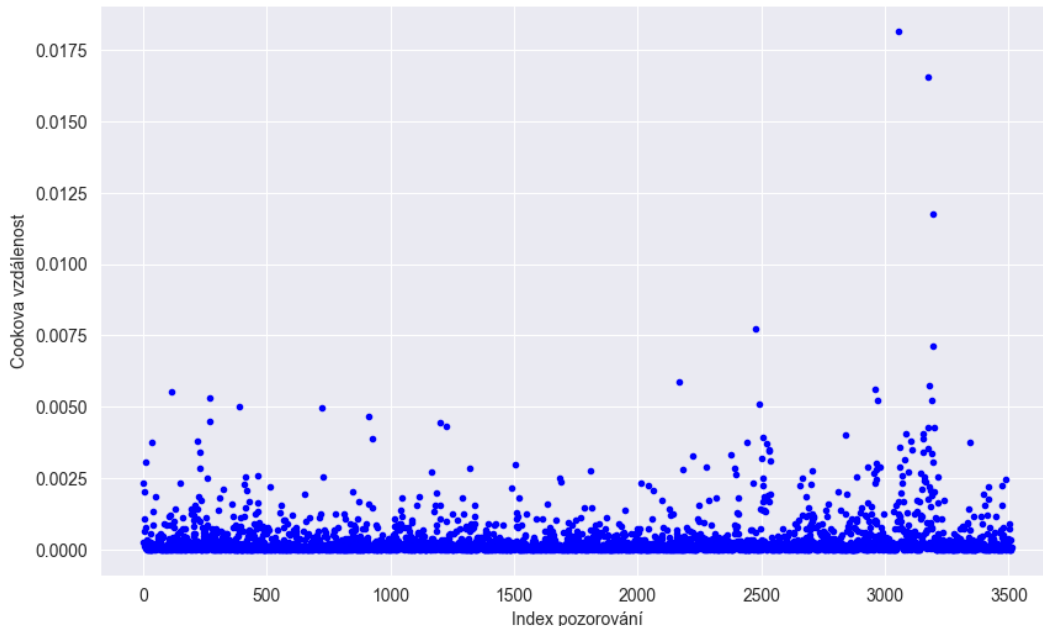
Pro použití statistické metody OLS používané pro výzkum vlastností je předpoklad normální rozdělení reziduí. V případě zkoumaného souboru dat však tento předpoklad není splněn. Proto byl proveden dodatečný průzkum dat a testy normality.

Prvním krokem k určení normálního rozdělení dat je Shapiro-Wilkův test, jehož výsledky zamítají normální distribuci (testová statistika: 0,990850031375885; p-

hodnota: $9,316291770151974e-15$). Pro výzkum distribuce reziduí probíhá identifikace velkých reziduí, odlehlých hodnot a vlivných bodů. Velká rezidua jsou identifikována porovnáním se standardními rezidui a v případě rozdílu většího než 2 se považuje za velké. Odlehlé hodnoty jsou zjištěny pomocí porovnání rezidua a dvojnásobku standardní chyby. Vlivné body jsou poté určeny pomocí Cookovy vzdálenosti, přičemž bod se vzdáleností větší než 1 je považován za vlivný. Všechny hodnoty identifikované těmito procesy jsou odstraněny ze zkoumaného souboru dat, a následně je provedena regresní analýza znovu bez těchto hodnot. I přes tyto úpravy však rezidua stále nevykazují normální rozdělení s těmito výsledky Shapiro-Wilkova testu: testová statistika: $0,9657636880874634$; p-hodnota: $8,583603983801585e-28$, proto výsledky regresní analýzy nemusí odpovídat skutečnosti.

Vlivné body jsou ze souboru dat odstraněny, zbývající body jsou znázorněny na grafu Cookovy vzdálenosti, viz graf 4. Na horizontální ose je index pozorování a vertikální osa představuje Cookovu vzdálenost. Jak je z grafu patrné, žádné z reziduí se 1 nepřibližuje, maximální vzdálenost je mírně větší než 0,0175.

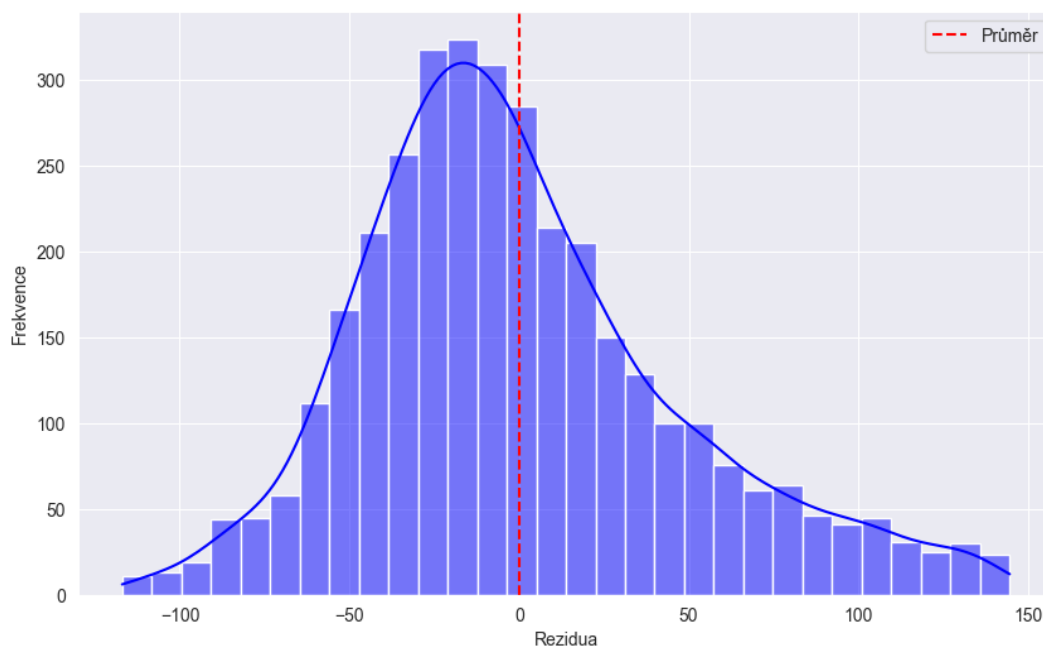
Graf 4 Cookova vzdálenost reziduí



Zdroj: vlastní zpracování

Po všech těchto uskutečněných úpravách je pro vizualizaci distribuce reziduí zpracován graf, viz graf 5. Horizontální osa grafu představuje samotná rezidua a vertikální osa zobrazuje jejich frekvenci. Rozdělení se zdá být mírně zkreslené a jsou v něm znatelné odchylky od symetrické Gaussovy křivky očekávané v normální distribuci.

Graf 5 Distribuce reziduí



Zdroj: vlastní zpracování

Úpravy zkoumaného souboru dat jej zmenšují o 216 záznamů. Hodnota R^2 se zvyšuje o 2 desetiny, hodnota F-statistiky se také mění, rovněž i koeficienty zkoumaných proměnných. Výsledný soubor dat přináší přesnější výsledky než neupravený soubor dat. Metoda OLS je provedena s kovariačním typem HC3, který zajišťuje že jsou standardní chyby robustní vůči heteroskedasticitě, výsledky tedy i přes nenormální rozdělení reziduí jsou v rámci okolností maximálně přesné.

4.3.3 Výsledky analýzy

V této části jsou popsány výsledky regresní analýzy metodou OLS, viz tabulka 7. Tabulka ukazuje přehled o koeficientu, standardní odchylce, hodnotě z, signifikanci a intervalu v němž se koeficient nachází. Koeficient dané proměnné značí změnu závisle proměnné (v tomto případě průměrné ceny za noc) při změně proměnné o jednu jednotku.

Hodnota R^2 (míra kvality modelu) se rovná 0,431, tudíž 43,1 % průměrné ceny volných nocí je vysvětleno nezávisle proměnnými v modelu. Obecně je model statisticky významný (F-statistic = 124,2; p-hodnota = 0,00) a s většinou nezávisle proměnných také statisticky významných.

Hodnota koeficientu konstanty je rovna cca -7 USD, což reprezentuje odhadovanou hodnotu průměrné ceny, kdyby všechny nezávisle proměnné byly rovny nule. V tomto případě je však odhad statisticky nevýznamný. Statisticky nevýznamné jsou také

odhadované koeficienty u proměnných „Kitchen“, „Fireplace“, „Pool“, „Washer“ či „WiFi“. Důvodem může být jejich nerovnoměrně rozdělená frekvence hodnot, jako například u „WiFi“, která se vyskytuje v 98,7 % apartmánů nebo proměnné „Pool“, ta se naopak vyskytuje pouze u 1,54 % apartmánů. Počet ložnic, vybavenost vířivkou a klimatizací mají dle odhadu pozitivní a statisticky významný vliv na průměrnou cenu.

Z kategorie reklamních a hostitelských funkcí jsou všechny proměnné statisticky významné a mají pozitivní vliv na průměrnou cenu, s nejvyšším koeficientem u proměnné „Superhost“. Proměnná „Rating“ byla pro účely modelu transformována umocněním, kvůli vysokému číslu podmínek „Condition number“ při použití netransformované hodnoty. V obou případech však byl odhadován pozitivní a statisticky významný vliv. Slabší pozitivní vliv má počet fotografií apartmánu v inzerátu.

Ze skupiny vybavení má, kromě dříve zmíněných, vysoký vliv přítomnost televize a sušičky. Celkový počet vybavení v apartmánu má pozitivní, avšak slabý vliv na cenu. Negativní a statisticky významný vliv má možnost bezplatného parkování. Zdá se, že tato proměnná výrazně snižuje průměrnou cenu. Možnost samostatného odbavení (check-in) zvyšuje průměrnou cenu a je statisticky významná.

Tabulka 7 Výsledky regresní analýzy OLS

Proměnná	Koeficient	Std. Error	z	P> z	[0,025	0,975]
(Constant)	-7,0742	12,388	-0,571	0,568	-31,354	17,206
Bedrooms	47,2374	1,636	28,869	0	44,03	50,444
Superhost	11,5394	1,822	6,332	0	7,968	15,111
Rating	1,0836	0,452	2,399	0,016	0,198	1,969
AmenitiesCount	0,337	0,088	3,85	0	0,165	0,509
ImageCount	0,1828	0,086	2,127	0,033	0,014	0,351
Kitchen	0,073	3,416	0,021	0,983	-6,621	6,767
Dryer	16,2897	2,213	7,36	0	11,952	20,628
Fireplace	6,0597	4,302	1,409	0,159	-2,372	14,491
Jacuzzi	29,8867	6,461	4,626	0	17,224	42,55
SelfCHI	8,7412	1,708	5,118	0	5,394	12,088
FreeParking	-30,3614	1,924	-15,78	0	-34,132	-26,59
Pool	-1,5035	7,8	-0,193	0,847	-16,792	13,785
AC	28,1173	2,111	13,322	0	23,981	32,254
TV	20,1148	1,852	10,863	0	16,486	23,744
Washer	-3,3082	1,941	-1,704	0,088	-7,113	0,497
WiFi	-8,4348	7,405	-1,139	0,255	-22,948	6,078

Zdroj: vlastní zpracování

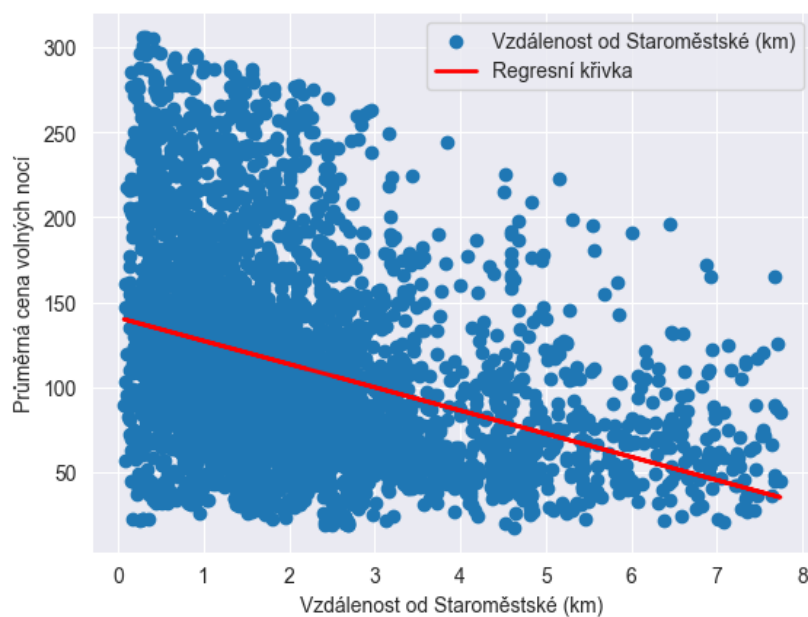
4.3.4 Vliv vzdálenosti na průměrnou cenu v Praze

Jediný kraj České republiky s dostatkem dat pro samostatnou regresní analýzu je hlavní město Praha. Vzhledem k tomu, že většina apartmánů v předchozí regresní analýze je právě z Prahy, tato podkapitola zkoumá vliv pouze jedné proměnné. Zvolenou proměnnou je vzdálenost (km) od stanice metra Staroměstská. Ta se nachází v městské části Staré Město, což je jedna z nejvyhledávanějších turistických lokalit v Praze. Tato proměnná je s některými proměnnými v předchozí analýze vysoce korelovaná, proto se zde soustředí pouze na ni.

Podobně jako u předešlé analýzy jsou proměnné omezeny maximální přijatou hodnotou pro odstranění odlehlých hodnot. Hranice vzdálenosti je stanovena 95. percentilem na hodnotu 7,774 km, průměrná cena volných nocí také 95. percentilem rovnající se 418.388 USD. Ze zkoumaného souboru dat jsou také odstraněny hodnoty s velkým reziduem.

Pro regresní analýzu je v tomto případě také nasazena metoda OLS, avšak s nerobustními chybami. Analýza má celkem 4 344 pozorování, míra kvality (R^2) tohoto modelu je 12 %. Model je statisticky signifikantní, s hodnotami F-statistic 594,7 a p-hodnotou $3.32e-123$. Koeficient pro konstantní hodnotu se rovná 140,9675 USD a je statisticky významný. Rostoucí vzdálenost apartmánu od zvolené lokace má záporný vliv na průměrnou cenu volných nocí a je statisticky významná (koeficient = -13,6532; p-hodnota = 0,00). Tudíž čím blíže je apartmán stanici Staroměstská, je cena vyšší. Graf 6 ukazuje předpovídanou regresní křivku, která je klesající a sledované hodnoty vzdálenosti s cenami. Na horizontální ose je vzdálenost od Staroměstské, na ose vertikální průměrná cena volných nocí. Data ze zkoumaného souboru jsou prezentována modrými tečkami a z grafu je patrné, že jsou rozptýlena a nesoustředěna v bezprostředním okolí regresní křivky.

Graf 6 Graf regresní analýzy



Zdroj: vlastní zpracování

5 Výsledky a diskuse

5.1 Analýza průměrné ceny

Tato bakalářská práce zkoumala průměrné ceny prodaných nocí apartmánů na platformě Airbnb pro rok 2023 v jednotlivých krajích České republiky. Apartmány byly rozděleny nejen podle krajů, ale také podle ložnic. Bylo zjištěno, že zkoumaný soubor dat je z většiny tvořen apartmány umístěnými v Praze, tímto jevem jsou výsledky zkresleny. Výsledné hodnoty jsou v této kapitole převedeny z amerických dolarů na české koruny pomocí průměrného kurzu (1 USD = 22,21 Kč) roku 2023 [32].

5.1.1 Výsledky pro apartmány s 1 ložnicí

Nejvyšších průměrných cen apartmánů s jednou ložnicí v průběhu roku 2023 dosahovaly kraje Vysočina a Karlovarský, viz graf 2. Průměrná hodnota ceny prodané noci celého roku se pro kraj Vysočina rovnala hodnotě 3 842 Kč, v Karlovarském kraji poté 3 021 Kč. Značný výkyv v průměrné ceně Karlovarský kraj zaznamenal v červnu, cena stoupla až o 122 % oproti květnu. Nejnižších hodnot v tomto počtu ložnic dosahoval Zlínský kraj, pohybující se pod hranicí 1 110 Kč.

Ostatní statistické veličiny, konkrétně medián a maximální sledované hodnoty, byly pro tuto kategorii taktéž podrobeny zkoumání. Kraj Vysočina, který vykazoval nejvyšší průměrnou cenu v průběhu roku, měl konstantní úroveň mediánu i maximální hodnoty téměř celý rok. Vzhledem k tomu, že tento kraj měl nejmenší počet sledovaných apartmánů, výsledky průměrné ceny byly ovlivněny malým počtem apartmánů s vyšší cenovou hladinou. Záměrem bylo tuto skutečnost omezit rozdělením do skupin podle počtu ložnic a také stanovením limitu průměrných cen používaných v analýze. Nicméně ani to nepomohlo, jelikož kraj Vysočina se v podobě neočekávaného výsledku průměrných cen objevil ve dvou skupinách i po limitování cen. Pozoruhodné také je porovnání úrovně křivky průměrné ceny hlavního města Prahy s křivkou maximální hodnoty. Zatímco křivka průměrné ceny se pohybuje téměř na stejné hladině jako většina ostatních krajů, křivka maximální hodnoty je po celý rok konstantní (= 9 839 Kč) a jednoznačně nejvyšší.

5.1.2 Výsledky pro apartmány se 2 ložnicemi

V porovnání s předchozí skupinou apartmánů s jednou ložnicí se cenová hladina v této skupině obecně zvýšila, kromě některých výjimek se cena nepohybuje pod hranicí 1110 Kč.

Několik krajů zároveň zaznamenává rostoucí výkyv průměrných cen v květnu, kdy Jihomoravský kraj dosáhl 2. nejvyšší úrovně pro tuto skupinu apartmánů a to 4 675 Kč, což je o 54 Kč méně než 1. nejvyšší hodnota, která se vyskytla v Libereckém kraji (4729 Kč). Hodnota se objevila v listopadu, kdy většina ostatních krajů naopak zaznamenala značný propad v průměrných cenách. V tomto měsíci se objevila jak nejvyšší, tak i nejnižší hodnota skupiny apartmánů. Nejnižší byla v Plzeňském kraji, na úrovni 533 Kč je o 88,7 % menší než nejvyšší hodnota této skupiny.

5.1.3 Výsledky pro apartmány se 3 až 6 ložnicemi

Při počtu ložnic, blízcímu se maximální hranici šesti, ubývalo hodnot a analýza neměla velkou vypovídací hodnotu. Proto se apartmány se 3 až 6 ložnicemi sloučily do jedné kategorie. S poměrně velkou rezervou nad ostatními kraji se pohybuje kraj Vysočina a Pardubický kraj, viz graf 4 v přílohách (Příloha A). Průměr celého roku pro tyto dva kraje se rovná 11 489 Kč a 12 635 Kč ve stejném pořadí. Ostatní kraje a jejich průměrná cena prodaných nocí se v průběhu roku většinou pohybovaly v intervalu od 2221 do 8884 Kč. Nejnižší hodnotu 422 Kč sledoval Olomoucký kraj, v tomto případě se však pravděpodobně jedná o chybu, jelikož tato hodnota je o 89,8 % menší než průměr tohoto kraje za celý rok.

5.1.4 Mapa krajů ČR s průměrnými cenami

Průměrná cena byla zpracována i pro celý rok 2023 bez rozlišení měsíců a bez rozdělení na jednotlivý počet ložnic. Obrázek 7 vyobrazuje mapu České republiky rozdělenou na kraje s uvedenou průměrnou cenou u každého z nich. Kraj s nejvyšší průměrnou cenou byl Liberecký (3 509 Kč), Vysočina (3 376 Kč) a Karlovarský kraj (3 132 Kč). Ostatní kraje a jejich průměrná cena za noc byly v intervalu od 1 999 po 2 843 korun. Pardubický a Plzeňský kraj projevují nejnižší průměrné ceny ve výši 1532 korun a 1 621 korun v uvedeném pořadí.

5.2 Analýza vlivu atributů na průměrnou cenu

Tato podkapitola pojednává o výsledcích analýz za použití metody nejmenších čtverců. V analýzách byl zkoumán vliv vybraných vlastností apartmánu na průměrnou cenu za volnou noc, přičemž bylo zkoumáno celkem 17 různých vlastností apartmánu. Výběr dat, filtrování, transformace a samotné analýzy byly provedeny uplatněním vlastního skriptu napsaného v programovacím jazyce Python.

5.2.1 Výsledky analýzy se zaměřením na Českou republiku

Statisticky signifikantní model OLS s mírou kvality 43,1 % a celkovým počtem pozorování 3508 poskytl výstup o koeficientech proměnných a jejich statistické významnosti. Výsledky jsou zaznamenány v tabulce 7. Většina sledovaných proměnných byla statisticky signifikantní až na výjimky, jako např. „WiFi“, „Kitchen“ či „Fireplace“.

Proměnné z první kategorie reklamních a hostitelských vlastností mají kladný vliv na průměrnou cenu a jsou statisticky signifikantní. Proměnná „Superhost“ (koef. = 11,5394; p-hodnota = 0) má však prokazatelně vyšší vliv než dvě ostatní proměnné, „Rating“ (koef. = 1,0836; p-hodnota = 0,016) a „ImageCount“ (koef. = 0,1828; p-hodnota = 0,033).

Z kategorie charakteristiky ubytování měly signifikantní a v rámci koeficientu vysoký kladný vliv proměnné „Bedrooms“, „Dryer“, „Jacuzzi“, „AC“ a „TV“. Proměnná „Bedrooms“ byla v tomto modelu zahrnuta především jako kontrolní. O něco menší statisticky významný kladný vliv má možnost self-checkinu. Vliv blížící se nule má počet položek vybavení s hodnotou koeficientu 0,337 a p-hodnotou 0. Záporný vliv na průměrnou cenu má možnost bezplatného parkování s koeficientem -30,3614 a p-hodnotou 0. Tento jev může být způsoben rozdělením dat do krajů. Většina apartmánů ze zkoumaného souboru dat je umístěna v Praze, kde většinou není možnost bezplatného parkování. Je však dostupné v ostatních krajích, kde je průměrná cena za noc obecně nižší.

5.2.2 Výsledky analýzy se zaměřením na Prahu

Další analýza zkoumala pouze apartmány umístěné v Praze a do 7,7km od turistické destinace (Staré Město). Právě vliv vzdálenosti apartmánu od této lokality byl v analýze zkoumán. Výsledky analýzy ukázaly, že s rostoucí vzdáleností od zvoleného místa klesá průměrná cena volné noci. Opačný jev sledovala studie prováděna ve Španělsku, jejíž výsledky sdělují, že s rostoucí vzdáleností od turistických atrakcí se cena zvedá [21]. V případě této práce bylo vyzkoumáno, že se cena při růstu vzdálenosti o 1 km sníží o 13,65 USD, avšak míra kvality modelu je pouze 12 %.

5.3 Komparativní analýza průměrné ceny s výsledky regresní analýzy

Tabulka 8 obsahuje výsledky analýzy průměrných cen a odhadované hodnoty průměrné ceny s využitím koeficientu z regresní analýzy. Data kategorizuje podle počtu ložnic apartmánu. Sloupec sledovaných hodnot obsahuje údaje převzaté z grafu 3, zatímco předpokládané hodnoty jsou odvozeny přičtením zaokrouhlené hodnoty koeficientu (47,2)

pro každou další ložnici. Pro skupinu apartmánů s jednou ložnicí je hodnota sloučena, aby sloupce začínaly na stejné úrovni. Je nutné zmínit, že hodnoty ve sloupci sledovaná hodnota obsahují průměrné ceny prodaných nocí, zatímco sloupec předpokládaná hodnota vychází z průměrné ceny volných nocí. Nejmenší rozdíl 9,5 USD se vyskytuje u skupiny se třemi ložnicemi, největší pak u skupiny s pěti (153,4 USD). Regresní analýza předpovídala konzistentní nárůst ceny při zvýšení počtu ložnic, avšak analýza průměrných cen vykazovala různé hodnoty změny mezi jednotlivými skupinami. Z toho důvodu se odhadované hodnoty liší od sledovaných.

Tabulka 8 Průměrné ceny typů apartmánu a jejich předpovídané hodnoty (USD)

Počet ložnic	Sledovaná hodnota	Předpokládaná hodnota
1	90,4	
2	109,7	137,6
3	194,3	184,8
4	264	232
5	432,6	279,2
6	420,5	326,4

Zdroj: vlastní zpracování

Vlastní část práce je založena na rozsáhlé práci s daty pomocí vlastního skriptu. V rámci analýzy průměrných cen skript umožnil stanovit typ analyzované průměrné ceny a pro cenu vypočítat limit pomocí percentilu. Byl využit pro sloučení několika tabulek databáze dohromady, pro následnou filtraci a výběr požadovaných dat. Z konečného souboru dat byla poté analyzována průměrná cena, medián a minimální a maximální hodnoty pro několik různých rozdělení. Výsledky bylo možné kategorizovat podle typu apartmánu, podle krajů České republiky nebo podle měsíců. Zmíněný skript byl aplikován také na regresní analýzu a řadu dalších metod s ní související, jako např.: analýza distribuce reziduí nebo identifikace odlehlých hodnot. Zpracované analýzy určují klíčové faktory cenotvorby apartmánového bydlení v rámci krátkodobého ubytování na platformě Airbnb. Kód byl na míru vyvinutý pro účely této práce, tím byla zajištěna vysoká koncentrovanost na dosažení konkrétních výsledků.

6 Závěr

V rámci práce byly zhodnoceny důležité metodické postupy a atributy apartmánů, které se používají při tvorbě cen. Analýza se soustředila na apartmány v České republice inzerované na platformě Airbnb. Vlastní práce zkoumala průměrnou cenu pro jednotlivé kraje České republiky a analyzovala vztah mezi vlastnostmi apartmánu a jejich cenami za noc. Ekonomická analýza nebyla opomenuta, ale byla provedena specificky podle potřeb zkoumaného prostředí podnikání v rámci apartmánového bydlení v České republice na platformě Airbnb.

Nejvyšších průměrných cen pro celý rok 2023 v rámci zkoumaného souboru dat dosahoval Liberecký kraj, kraj Vysočina a Karlovarský kraj, jak je uvedeno na straně 36. Při rozdělení apartmánů do skupin podle počtu ložnic byla nejpočetnější skupinou ta s apartmány s jednou ložnicí. S každou další ložnicí průměrná cena narůstá, jak je popsáno na straně 35. Ve skupině apartmánů s jednou ložnicí nejvyšších průměrných hodnot dosahoval kraj Vysočina. Obecně nejnižších hodnot dosahoval Pardubický kraj.

Tato práce také vysvětlila postup sběru dat z online platformy při nedostupnosti aplikačně-programového rozhraní. Pro sběr dat byl využit software typu web scraper. Výsledky byly uloženy do databáze, ze které byla čerpána data pro analýzy. Zásadní význam pro tuto práci měl v rámci provedených analýz speciálně navržený skript, popsáný na stranách 23, 24 a 25. Umožňoval k analýzám přistupovat inovativním způsobem, s osobním porozuměním a maximální možností přizpůsobení na daný sektor podnikání.

Z výsledků analýzy vzešly údaje o vlastnostech, jejich statistické významnosti a vztahu ke zvolené závislé proměnné – průměrné ceně neprodaných nocí. Bylo zjištěno, že kladný vliv na závisle proměnnou má počet ložnic a některé specifické vybavení, jako např.: vířivka, klimatizace, televize či sušička. Z toho vyplývá, že pokud je apartmán některou z těchto položek vybaven, jeho průměrná cena za noc roste. Stejně tak s každou další ložnicí se zvyšuje průměrná cena. Záporný vliv na průměrnou cenu za noc měla možnost bezplatného parkování. Z kategorie vlastností inzerátu apartmánu na Airbnb měl pouze status „superhostitele“ poměrně vyšší koeficient oproti 2 ostatním proměnným – počet fotografií a hodnocení. Tyto údaje plynou ze strany 45, kde jsou uvedeny výsledky analýzy. Statisticky nevýznamné se prokázaly být některé proměnné obsahující informaci o dostupnosti vybavení.

Další analýza zkoumala vliv vzdálenosti apartmánu od pražské turistické destinace (Staré Město), byl tedy omezena pouze na apartmány umístěné v hlavním městě Praze. Z výsledků uvedených na straně 46 vyplývá, že čím je vzdálenost mezi těmito dvěma body vyšší, tím je průměrná cena volných nocí menší. Lze tedy říci, že apartmány umístěné v městské části Staré Město či apartmány v blízké vzdálenosti k tomuto místu mají vyšší průměrnou cenu než ty odlehlejší.

Přínosem této práce je poskytnutí informací a metodického postupu orientace v tomto sektoru podnikání. Zájemce o podnikání získává klíčové atributy apartmánů, na které se zaměřit pro nastavení odpovídající ceny. Taková osoba by se, podle strany 45, měla zaměřit na poskytnutí vhodného počtu ložnic, vybavení jako je vířivka nebo klimatizace a také by se obecně měla zaměřit na důvěryhodné a spolehlivé poskytování služeb, které rozhodují o statusu „superhostitele“. Vhodné by bylo také doporučení kraje, kde by své podnikání provozovala, avšak analýza průměrných cen neposkytuje dostatečně ověřitelné údaje. Proto je v tomto případě, dle strany 46, možné doporučit orientovat se především na umístění v blízkosti turistické destinace vybraného města.

7 Seznam použitých zdrojů

- [1] GÁLA, Libor, Jan POUR a Zuzana ŠEDIVÁ. *Podniková informatika: počítačové aplikace v podnikové a mezipodnikové praxi*. 3., aktualizované vydání. Praha: Grada Publishing, 2015. Management v informační společnosti. ISBN 978-802-4754-574.
- [2] BRUCKNER, Tomáš. *Tvorba informačních systémů: principy, metodiky, architektury*. Praha: Grada, 2012. Management v informační společnosti. ISBN 978-80-247-4153-6.
- [3] BUCHALCEVOVÁ, Alena. *Metodiky vývoje a údržby informačních systémů: kategorizace, agilní metodiky, vzory pro návrh metodiky*. Praha: Grada, 2005. Management v informační společnosti. ISBN 80-247-1075-7.
- [4] ŠTRÁFELDA, Jan. *Co je User Interface (UI)* [online]. Dostupné z: <https://www.strafelda.cz/ui>
- [5] INTERACTION DESIGN FOUNDATION. *User Interface (UI) Design. The Interaction Design Foundation* [online]. 2016. Dostupné z: <https://www.interaction-design.org/literature/topics/ui-design>
- [6] LANE, Kin. *API 101. API Evangelist* [online]. Dostupné z: <https://apievangelist.com/info/101/>
- [7] POSTMAN, INC. *What is an API? Postman API Platform* [online]. 2024. Dostupné z: <https://www.postman.com/what-is-an-api/>
- [8] DE, Brajesh a Brajesh DE. *API Management*. In: *API Management* [online]. Berkeley, CA: Apress, 2017, s. 15-28. ISBN 978-1-4842-1306-3. Dostupné z: doi:10.1007/978-1-4842-1305-6_2
- [9] HOTELSTARS UNION. *Kritéria pro apartmánové hotely/komplexy 2021 - 2025* [online]. 2021, 23. Dostupné z: <https://hotelstars.cz/Upload/2022/11/22/apartmany-q1-2022.pdf>
- [10] AIRBNB, INC. *How to define what kind of space you're hosting. Airbnb* [online]. 2022. Dostupné z: <https://www.airbnb.com/resources/hosting-homes/a/how-to-define-what-kind-of-space-youre-hosting-361>
- [11] ROMA, Paolo, Umberto PANNIELLO a Giovanna LO NIGRO. *Sharing economy and incumbents' pricing strategy: The impact of Airbnb on the hospitality industry. International Journal of Production Economics* [online]. 2019, **214**, 17-29. ISSN 09255273. Dostupné z: doi:10.1016/j.ijpe.2019.03.023

- [12] AIRBNB, INC. About us. *Airbnb* [online]. 2023. Dostupné z: <https://news.airbnb.com/about-us/>
- [13] BOOKING.COM™. About Booking.com™. *Booking.com* [online]. Dostupné z: <https://www.booking.com/content/about.html>
- [14] AIRBNB, INC. *How Airbnb works* [online]. Dostupné z: <https://www.airbnb.com/e/howairbnbworks>
- [15] AIRBNB, INC. Sharing More About the Technology That Powers Airbnb. *Airbnb Newsroom* [online]. 2018. Dostupné z: <https://news.airbnb.com/sharing-more-about-the-technology-that-powers-airbnb/>
- [16] AIRBNB, INC. Airbnb using AI to help clamp down on New Year's Eve parties globally. *Airbnb Newsroom* [online]. Dostupné z: <https://news.airbnb.com/airbnb-using-ai-to-help-clampdown-on-new-years-eve-parties-globally/>
- [17] AIRBNB, INC. Airbnb service fees. *Airbnb* [online]. Dostupné z: <https://www.airbnb.com/help/article/1857>
- [18] CODICA.COM. How Much Does It Cost to Build a Rental Marketplace Like Airbnb. *Codica* [online]. Dostupné z: <https://www.codica.com/blog/how-to-build-website-like-airbnb-step-by-step-guide/>
- [19] WEYGANDT, Jeremy, Paul KIMMEL a Donald KIESO. *Weygandt's Managerial Accounting: Tools for Business Decision Making*. Global edition. John Wiley, 2017. ISBN 9781119419655.
- [20] LLADÓS-MASLLORENS, Josep, Antoni MESEGUER-ARTOLA a Inma RODRÍGUEZ-ARDURA. Understanding Peer-to-Peer, Two-Sided Digital Marketplaces: Pricing Lessons from Airbnb in Barcelona. *Sustainability* [online]. 2020, **12**(13). ISSN 2071-1050. Dostupné z: doi:10.3390/su12135229
- [21] PEREZ-SANCHEZ, V., Leticia SERRANO-ESTRADA, Pablo MARTI a Raul-Tomas MORA-GARCIA. The What, Where, and Why of Airbnb Price Determinants. *Sustainability* [online]. 2018, **10**(12). ISSN 2071-1050. Dostupné z: doi:10.3390/su10124596
- [22] BENÍTEZ-AURIOLES, Beatriz a Iis TUSSYADIAH. What Airbnb does to the housing market. *Annals of Tourism Research* [online]. 2021, **90**. ISSN 01607383. Dostupné z: doi:10.1016/j.annals.2020.103108
- [23] GLEZ-PEÑA, Daniel, Anália LOURENÇO, Hugo LÓPEZ-FERNÁNDEZ, Miguel REBOIRO-JATO a Florentino FDEZ-RIVEROLA. Web scraping technologies in an API world. *Briefings in Bioinformatics* [online]. 2014, 2014-09-01, **15**(5), 788-797. ISSN 1477-4054. Dostupné z: doi:10.1093/bib/bbt026

- [24] KHDER, Moaiad. Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application. *International Journal of Advances in Soft Computing and its Applications* [online]. 2021, 2021-12-30, **13**(3), 145-168. ISSN 27101274. Dostupné z: doi:10.15849/IJASCA.211128.11
- [25] MERIS, Ramon. Legalities of Airbnb API Scraping. *Airbtics | Airbnb Analytics* [online]. Dostupné z: <https://airbtics.com/airbnb-api-scraping/>
- [26] AIRBNB, INC. Learn about Guest Favorites, the most loved homes on Airbnb. *Airbnb* [online]. 2024. Dostupné z: <https://www.airbnb.com/>
- [27] PANDAS. DataFrame. *Pandas API Reference* [online]. 2024. Dostupné z: <https://pandas.pydata.org/docs/reference/frame.html>
- [28] PERKTOLD, Josef, Skipper SEABOLD, Jonathan TAYLOR a statsmodels-developers. Linear Regression. *Statsmodels* [online]. 2023. Dostupné z: <https://www.statsmodels.org/stable/regression.html#model-classes>
- [29] LUMIVERO. Ordinary Least Squares regression (OLS). *XLSTAT, Your data analysis solution* [online]. Dostupné z: <https://www.xlstat.com/en/solutions/features/ordinary-least-squares-regression-ols>
- [30] SEZNAM.CZ. *Kraje ČR* [online]. 2024. Dostupné z: <https://mapy.cz/zakladni?source=area&id=585504>
- [31] AIRBNB, INC. The amenities guests want. *Airbnb* [online]. 2020. Dostupné z: <https://www.airbnb.com/resources/hosting-homes/a/the-amenities-guests-want-25>
- [32] KURZY.CZ. USD průměrné kurzy 2023, historie kurzů měn. *Kurzy.cz* [online]. 2024. Dostupné z: <https://www.kurzy.cz/kurzy-men/historie/USD-americky-dolar/2023/>

Seznam obrázků, tabulek, grafů a zkratek

7.1 Seznam obrázků

Obrázek 1 Zpeněžení na platformě Airbnb	17
Obrázek 2 Schéma relační databáze získaných dat z Airbnb	22
Obrázek 3 Snímek zdrojového kódu HTML ze stránky Airbnb.com	23
Obrázek 4 Část kódu pro analýzu průměrných cen.....	24
Obrázek 5 Část kódu pro regresní analýzu.....	25
Obrázek 6 Část kódu pro filtraci nevhodných dat.....	26
Obrázek 7 Mapa krajů ČR a jejich průměrná cena apartmánu za noc v roce 2023 (USD).	36

7.2 Seznam tabulek

Tabulka 1 Rozdělení apartmánů podle počtu ložnic	30
Tabulka 2 Počty chybějících hodnot pro typy cen	31
Tabulka 3 Definice proměnných používaných ve vícenásobné regresní analýze	38
Tabulka 4 Deskriptivní statistika proměnných.....	40
Tabulka 5 Maximální hodnoty proměnných	41
Tabulka 6 Korelační tabulka proměnných	41
Tabulka 7 Výsledky regresní analýzy OLS.....	45
Tabulka 8 Průměrné ceny typů apartmánu a jejich předpovídané hodnoty (USD).....	50

7.3 Seznam grafů

Graf 1 Počet apartmánů v krajích ČR	29
Graf 2 Průměrná cena za noc v krajích ČR pro apartmány s 1 ložnicí (USD).....	32
Graf 3 Průměrná cena za noc pro typy apartmánů (USD)	35
Graf 4 Cookova vzdálenost reziduí	42
Graf 5 Distribuce reziduí.....	43
Graf 6 Graf regresní analýzy	46

7.4 Seznam použitých zkratek

API – Application Programming Interface

CSV – Comma Separated Values

GUI – Graphical User Interface

IS – Informační Systém

OLS – Ordinary Least Squares

UI – User Interface

URL – Uniform Resource Locator

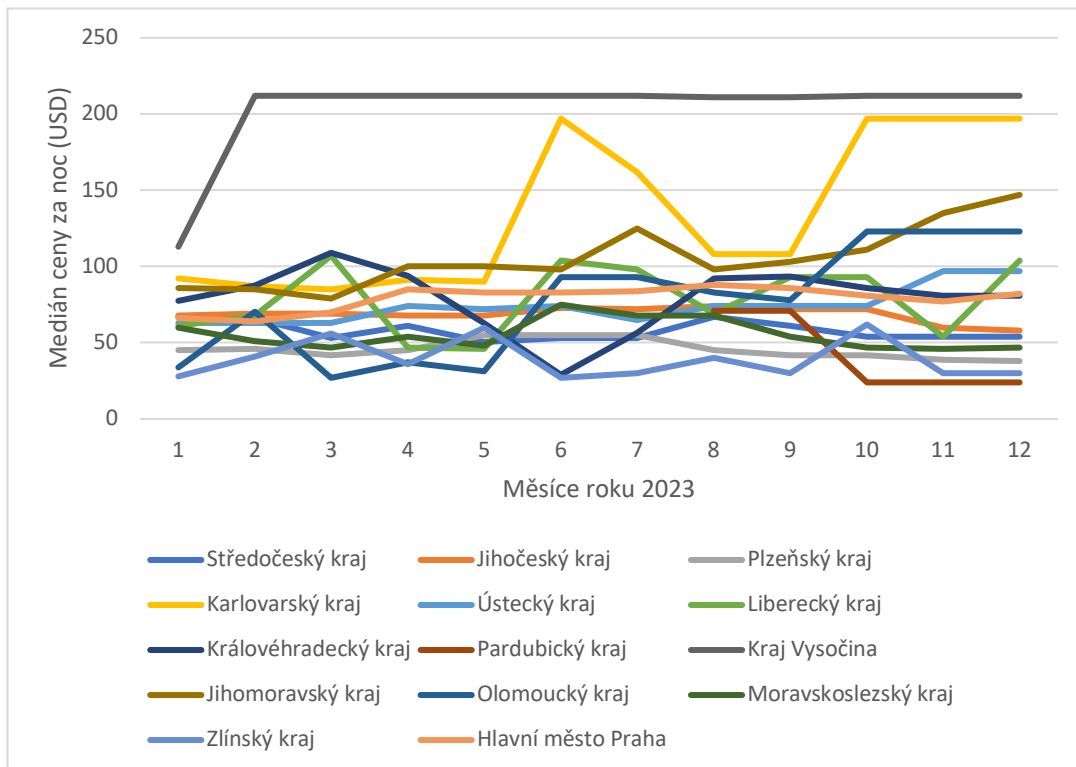
UX – User Experience

Přílohy

Příloha A Grafy používané při analýze průměrných cen krajů ČR

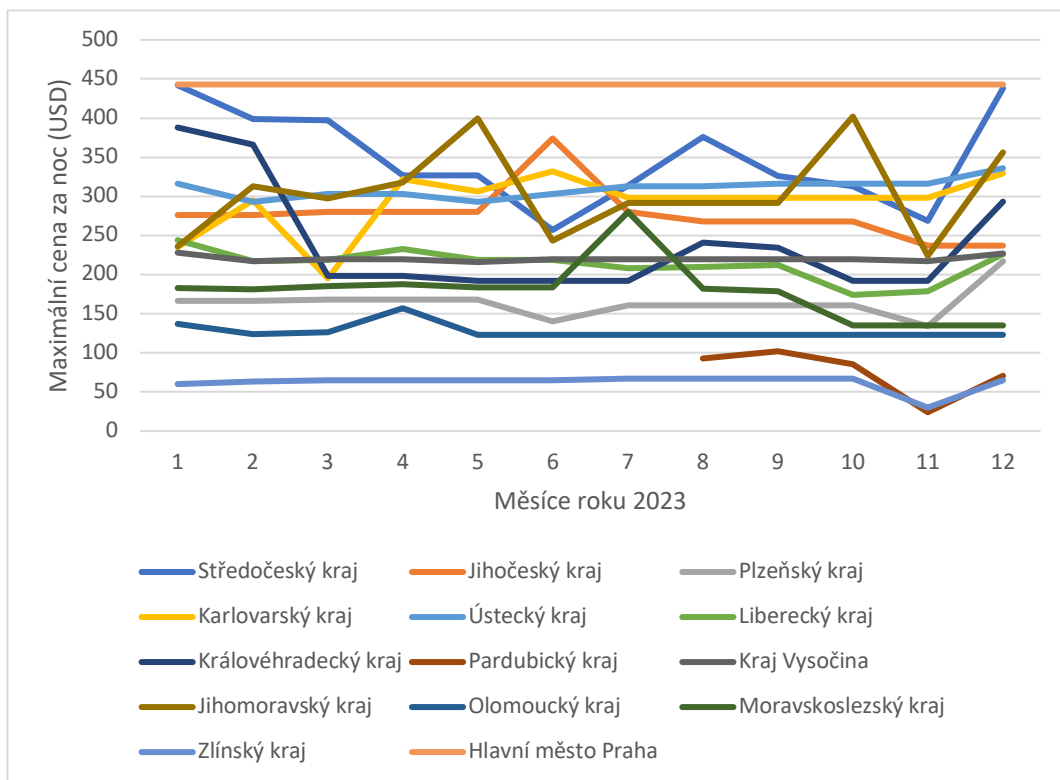
Příloha A Grafy používané při analýze průměrných cen krajů ČR

Graf 1 Medián ceny za noc v krajích České republiky pro apartmány s 1 ložnicí



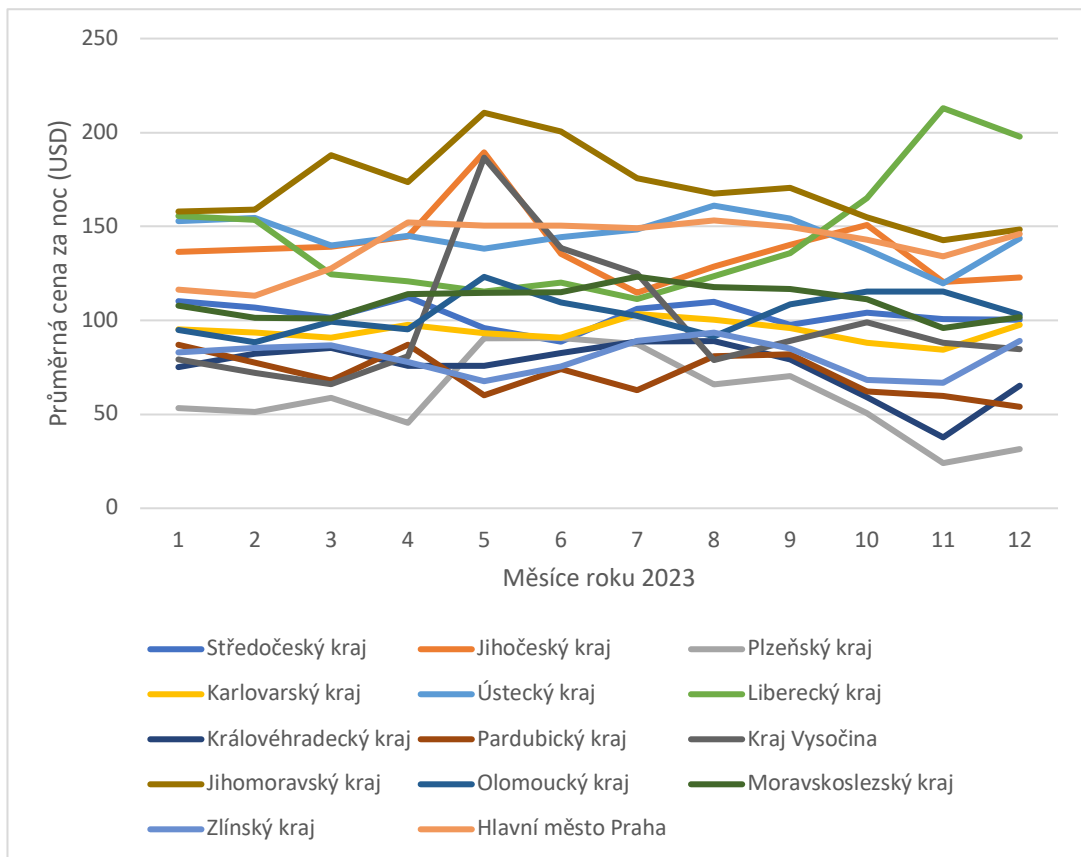
Zdroj: vlastní zpracování

Graf 2 Maximální dosažené ceny v krajích České republiky pro apartmány s 1 ložnicí



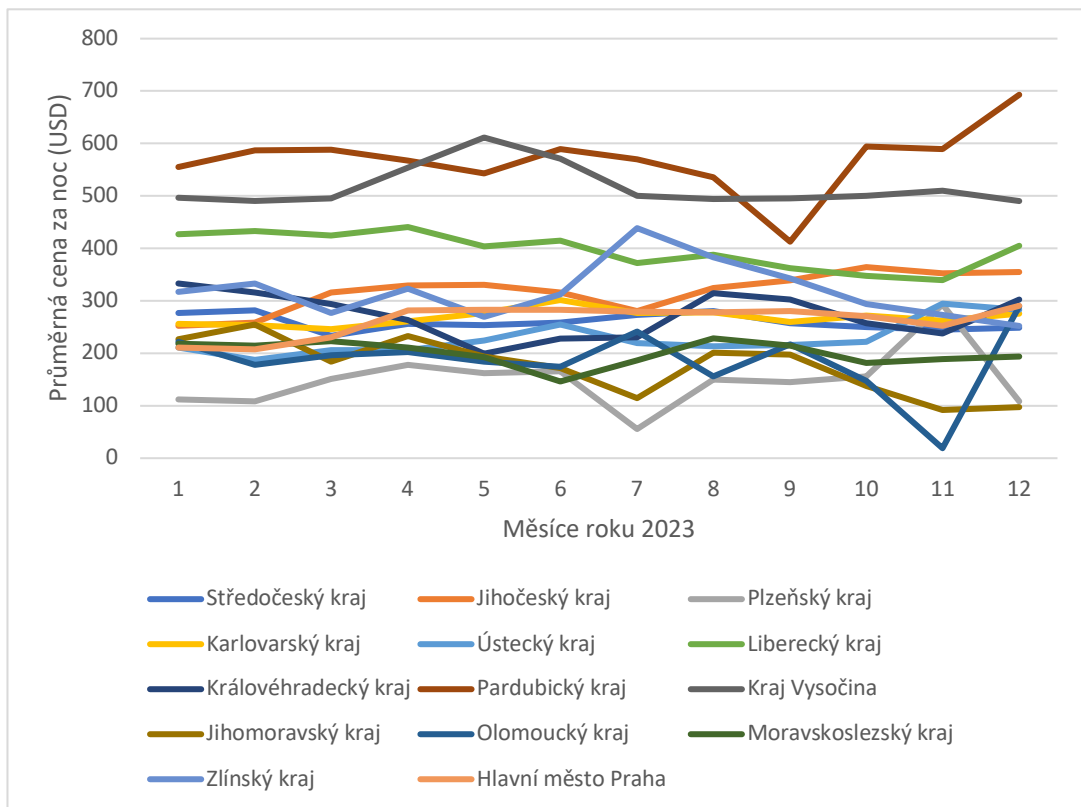
Zdroj: vlastní zpracování

Graf 3 Průměrná cena za noc v krajích České republiky pro apartmány se 2 ložnicemi



Zdroj: vlastní zpracování

Graf 4 Průměrná cena za noc v krajích České republiky pro apartmány se 3–6 ložnicemi



Zdroj: vlastní zpracování