

Česká zemědělská univerzita v Praze

Technická fakulta



Analýza vývoje cen ojetých vozidel

bakalářská práce

Vedoucí bakalářské práce: doc. Ing. Boleslav Kadleček, CSc.

Kristýna Adámková

PRAHA 2014

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Katedra vozidel a pozemní dopravy

Technická fakulta

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Adámková Kristýna

Obchod a podnikání s technikou

Název práce

Analýza vývoje cen ojetých vozidel

Anglický název

Price analysis for used cars

Cíle práce

Cílem práce je popsat vývoj a dosavadní poznatky v problematice prodeje ojetých vozidel, zejména principy, metody, metodiky, technické faktory, časové a sezónní vlivy při stanovení ceny ojetých vozidel

Metodika

- prostudovat základní literaturu, normy, internetové odkazy a další dostupné prameny z celého světa a provést literární rešerši v oblasti organizace prodeje a stanovení ceny ojetých vozidel
- vyhledat a kontaktovat významné instituce, servisní organizace a soudně znalecké společnosti a jednotlivce, zabývající se stanovením ceny vozidel a problematikou organizace prodeje ojetých vozidel
- provést vlastní analýzu a uvést nové případné teoretické předpoklady a názory

Osnova práce

1. Úvod
2. Vývoj a současný stav ohodnocování ojetých vozidel
3. Přehled metodik a výpočetních aplikací
4. Trendy v oblasti modelovacích nástrojů
5. Závěr

Prohlášení

Prohlašuji, že bakalářskou práci jsem vypracovala samostatně pod vedením doc. Ing. Boleslava Kadlečka, CSc. a na základě uvedené použité literatury.

Dne 6. dubna 2014

.....

podpis studenta

Poděkování

Ráda bych touto cestou poděkovala vedoucímu bakalářské práce panu doc. Ing. Boleslavu Kadlečkovi, CSc. za pomoc při vypracování této práce a za poskytnutí potřebných podkladů.

Abstrakt

Bakalářská práce se zabývá vývojem a dosavadními poznatky v problematice prodeje ojetých vozidel. Pozornost je věnována průběžnému vývoji tohoto trhu z pohledu několika klíčových parametrů během posledních 3 let. Následně je práce zaměřena na současné dění na trhu ojetých vozidel a současný stav ohodnocování těchto vozidel. Zároveň jsou v práci popsány principy data miningu (dolování dat z databází) a s tím související přehled možných metodik, výpočetních aplikací a modelovacích nástrojů sloužících k tomuto způsobu zpracování velkého množství vstupních dat, pomocí nichž bychom měli dosáhnout zhodnocení ojetého vozidla. Poslední kapitola je věnována celkové analýze trhu ojetých vozidel v České republice za rok 2013.

Klíčová slova

prodej ojetých vozidel, rssfocus, cebia, statistika

Price analysis for used cars

Abstract

This bachelor thesis deals with the development and existing knowledge about the issue of sale of used vehicles. It is focused on the ongoing development of this market from the perspective of several key parameters during the last three years. Subsequently, the work is focused on contemporary development on the market of used vehicles and the current state of the valuation of these vehicles. At the same time, the thesis describes principles of data mining and the list of possible related methodologies, computer applications and modeling tools used for this type of processing large amounts of input data, with the help of which we should achieve increased value of a used vehicle. The last chapter is devoted to the overall analysis of the market for used cars in the Czech Republic in 2013.

Keywords

sale of used cars, rssfocus, cebia, statistics

Obsah

1	Úvod	1
2	Vývoj a současný stav prodeje ojetých vozidel.....	2
2.1	Ochrana spotřebitele.....	3
2.1.1	Sdružení na Ochranu vlastníků automobilů (SOVA).....	3
2.2	Trendy v prodeji ojetin.....	3
3	Nynější stav ohodnocování ojetých vozidel	8
3.1	Cebia, spol. s.r.o.	8
3.2	Společnost IBS automotive	8
3.3	EurotaxGlass's	9
3.4	Znalecký standard č. I/2005	10
4	Přehled metodik a výpočetních aplikací.....	11
4.1	Data mining	11
4.2	Princip data miningu	12
4.3	Metodiky data miningu	13
4.3.1	Metodika CRISP-DM.....	13
4.3.2	Metodika 5A.....	15
4.3.3	Metodika SEMMA	16
4.4	Typy úloh pro data mining.....	16
4.4.1	Predikce a klasifikace	16
4.4.2	Deskripce.....	17
4.4.3	Hledání nuggetů	17
4.5	Základní techniky data miningu.....	18
4.5.1	Statistické techniky	18
4.5.2	Techniky umělé inteligence (nestatistické metody)	19
5	Trendy v oblasti modelovacích nástrojů.....	21
5.1	Komerční systémy.....	22
5.2	Volně dostupné systémy	23
6	Současný stav prodeje ojetých vozidel.....	25
6.1	Cebia statistika	25

6.2	SDA statistika.....	34
7	Závěr.....	35
8	Seznam odborné literatury.....	37
9	Použité zkratky	40
10	Seznam obrázků.....	41

1 Úvod

Trh ojetých automobilů je ve statistikách (vedených od r. 1998) každoročně na úrovni více než 40% z celkového počtu nově registrovaných vozidel. Týká se tedy tato problematika přibližně dvou milionů vlastníků. V souvislosti se stále se vyskytující neseriózností na trhu ojetých vozidel je potřeba v zájmu mnohých z nás tomuto dění věnovat značnou pozornost. Situace na českém trhu s ojetými vozidly je stále vážná a to hlavně z důvodu klamání spotřebitele, kterého se dopouštějí mnohé autobazary či prodejci ojetých automobilů.

Hlavním cílem bakalářské práce je podat ucelený pohled v oblasti organizace prodeje a stanovení ceny ojetých vozidel. Podání uceleného pohledu znamená představit dosavadní vývoj tohoto trhu, především z pohledu klíčových parametrů. Trh ojetých vozidel je tak rozsáhlý a zasahuje až do oblastí, které stále nejsou tolik prozkoumány, jako například malé autobazary, prodeje mezi soukromými osobami a tak nezbyvá nic jiného, než výsledky posoudit odhadem. V neposlední řadě je velmi důležitá orientace na spotřebitele tohoto trhu, tedy případné uchazeče o koupi ojetého vozu. Tedy dostatečná informovanost těchto uchazečů o možných nástrahách ze strany autobazarů či prodejců ojetých vozidel je důležitá. Další pozornost je směřována na současný stav ohodnocování ojetých vozidel, čemuž se v dnešní době věnuje spousta institucí, které mají v zájmu poskytovat zákazníkům informace a poradenství ohledně této problematiky. k odhadu ceny ojetého vozidla je potřeba velké množství vstupních dat. Ke zpracování těchto dat je nezbytné analyzovat tzv. „data mining“ (dolování dat z databází), které se při ohodnocování ojetých automobilů používá. V souvislosti s tím možné metodiky, úlohy, techniky a modelovací nástroje využívané při tomto zpracování dat. Poslední kapitola je věnována souhrnným informacím o trhu ojetých vozidel v České republice za rok 2013.

Zaměření bakalářské práce je teoretické. Práce je vypracována formou rešerše informačních zdrojů a doplněna o závěry ze studia příslušné literatury.

2 Vývoj a současný stav prodeje ojetých vozidel

V průběhu posledních let začal v celé Evropě včetně České republiky klesat zájem o nová vozidla a růst daleko větší zájem o vozidla ojetá. To vše v důsledku ekonomické recese. Zákazníci, kteří si dříve pořizovali nová vozidla, se v současné době, a to především z finančních důvodů orientují spíše na vozidla ojetá a vyhledávají kvalitní zánovní ojetá vozidla s prokazatelnou historií. Ve prospěch a v zájmu zákazníka je, aby vůbec věděl, co kupuje, je na místě provést pečlivou kontrolu těchto vozidel (CEBIA, 2013).

Základním přezkoumáním vozidla je diagnostikovat co nejvíce údajů týkajících se jejich historie. V první řadě údaje vztahující se především ke skutečně najetým kilometrům, jelikož přetáčení tachometrů, by se v dnešní době dalo označit jako jeden z nejčastějších prohřešků autobazarů, co se týče prodeje ojetých vozidel (CEBIA, 2013).

Nejčastější prohřešky autobazarů:

- přetáčení tachometrů,
- umělé omlazování vozidel,
- komisioní prodej v bazaru,
- kamuflované bouračky,
- spojování dvou aut do jednoho,
- legalizovaná kradená auta,
- dovoz ze zahraničí bez zaplacení DPH,
- auta z leasingu prodaná do bazaru (ačkoliv patří leasingové společnosti),
- auta v exekuci (PALLA, 2013).

V současné době i nadále český trh s ojetými vozidly představuje mnohá nebezpečí. Je tedy důležité, aby spotřebitelé, tedy zákazníci, kteří mají v plánu si pořizovat ojetý vůz, byli dostatečně informováni o možných nástrahách autobazarů či prodejců těchto vozů.

2.1 Ochrana spotřebitele

Ve spojitosti s prodejem ojetých vozidel mohou spotřebitelé narazit na mnohá rizika. Je v tomto odvětví neustále co zlepšovat a předcházet tak možným hrozbám ze strany nepoctivých autobazarů nebo individuálních dovozců. V důsledku těchto nástrah vznikají různé asociace či sdružení, která se tyto problémy snaží jak jen je to možné odstranit.

2.1.1 Sdružení na Ochranu vlastníků automobilů (SOVA)

Cílem tohoto sdružení, jak již bylo řečeno, je zajistit co největší informovanost spotřebitele, co se týče českého trhu s ojetými vozidly. Dále seznámení s nejčastějšími triky autobazarů či prodejců ojetých vozů, poskytovat rady všem po celé ČR. Poté aby bylo zajištěno, že se v ČR budou prodávat pouze vozidla kvalitní, technicky způsobilá a prověřené ojeté vozy, jejichž provoz nebude pro řidiče představovat životu nebezpečný hazard nebo riziko zabavení způsobené nejasným nebo v horším případě kriminálním původem (SOVA, © 2004 - 2014).

Situace v tuzemsku se daleko více zhoršila po vstupu ČR do EU (1. 5. 2004), čímž na hranicích přestaly probíhat kontroly dovážených ojetých vozidel. To způsobilo, že se tak na české silnice dostalo velké množství aut ze všech koutů EU, naprosto staré, nezpůsobilé a v některých případech až životu nebezpečné (SOVA, © 2004 - 2014).

Smyslem tohoto sdružení je tedy spolupracovat s médii, která by na jednotlivé případy upozorňovala. Taktéž jsou vyzýváni majitelé ojetých automobilů, aby se tomuto sdružení svěřovali se svými negativními zkušenostmi a mohlo tak být o těchto praktikách informováno více lidí. A donutit tímto způsobem prodejce ojetých vozidel, aby od těchto praktik ustoupili (SOVA, © 2004 - 2014).

Dalším sdružením, na které se spotřebitelé mohou bezprostředně obrátit je Sdružení obrany spotřebitelů (SOS), Věci automobilové a mnoho dalších.

2.2 Trendy v prodeji ojetin

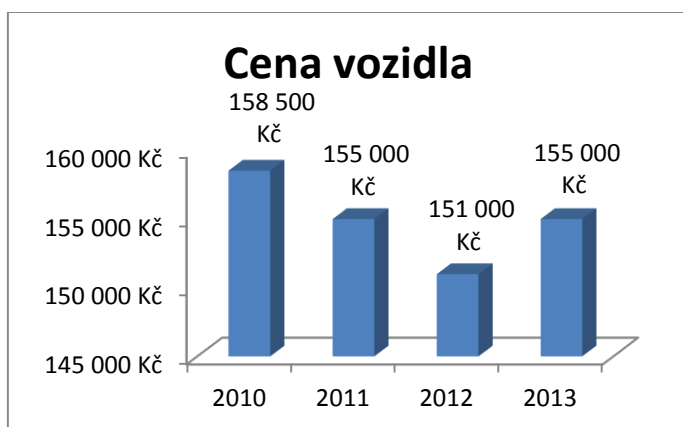
Nyní se budeme soustředit na trh s ojetými vozidly a sledovat trendy v této oblasti. Jde o oblast stále málo prozkoumanou, kde chybí řada přesných údajů a čísel. Například ani co se týče celkového objemu trhu ojetých vozidel, tedy počtu ročně prodaných ojetých vozidel, nejsou k dispozici přesná fakta. Kdežto vzhledem ke snižujícímu se zájmu o nová vozidla (a to i firem) se jedná o trh rostoucí a prosperující. Proto aktuálnější a podstatnější

se stává zkvalitňování tohoto trhu a odstranění řady podvodů, ke kterým bohužel i nadále dochází. Lze očekávat, že zákazníci (a to i firmy), kteří se původně soustředili na pořizování nových vozů a levnější varianta je pro ně ojetý vůz, se budou především orientovat na zánovní, kvalitní a prošetřené vozy s jednoznačnou historií (CEBIA, 2013c).

V následující analýze je srovnáno několik klíčových parametrů: prodejní cena, obsah motoru, stáří vozidla, stav tachometru, podíl typu motoru a karoserií automobilů. Je zde sledován vývoj a snaha dospět k možným příčinám aktuálních trendů ohledně prodeje ojetých vozidel. Zkoumali se skutečně prodané vozy s cenami, za které byly skutečně prodány. Ale co se týče dalších informací, které se udávají při prodeji, je opravdu potřeba počítat s tím, že se velmi často liší od reality. Analýza probíhala v různých autobazarech v České republice za poslední 3 roky (rok 2013 za prvního půl roku) a byla provedena na poměrně rozsáhlém vzorku a to u více než 75 tisíc ojetých osobních a užitkových vozidel (CEBIA, 2013c).

Ze sloupcového grafu na obrázku č. 1 je zřejmé, že průměrné prodejní ceny ojetých vozidel od roku 2010 klesaly. Lze říci, že ceny ojetých vozidel kopírují ceny nových automobilů, které během posledních let klesly zhruba o 5%. Ovšem pozoruhodný je v tomto kontextu opětovný nárůst průměrných prodejních cen v první polovině roku 2013. Kde se může jednat o krátkodobou výchylku nebo tedy o zmíněný zájem zákazníků o dražší zánovní a kvalitní ojetá vozidla, která si pořizují z ekonomických důvodů místo nových vozidel (CEBIA, 2013c).

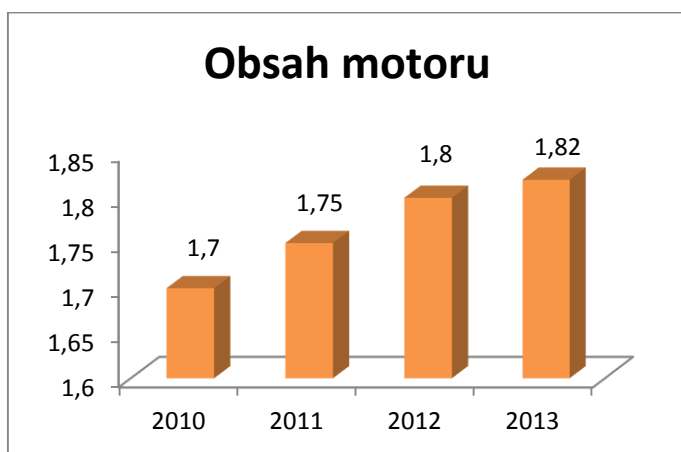
Obrázek 1 Cena vozidla



Zdroj: (CEBIA, 2013c)

Co se týče obsahu motoru (obrázek č. 2), jedná se o průměrnou kubaturu a to jak u diesellového, tak benzinového motoru. Jak vyplývá z analýzy, neustále větší zájem projevují zákazníci o motory se stále vyšším obsahem. Dokonce i přesto, že motory jsou stále výkonnější a vozidlo s vyšším obsahem motoru má dražší povinné ručení (CEBIA, 2013c).

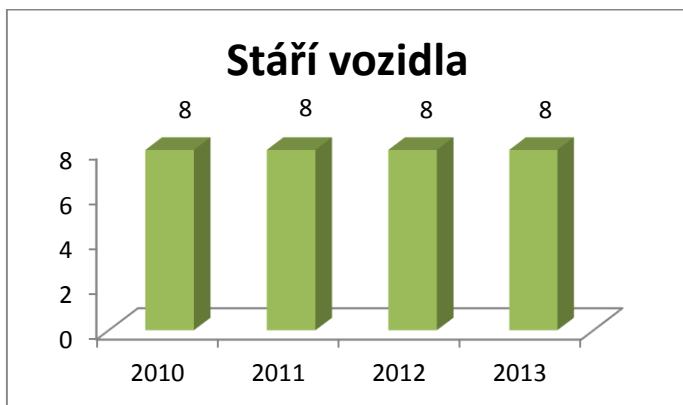
Obrázek 2 Obsah motoru



Zdroj: (CEBIA, 2013c)

Další parametr, tedy průměrné stáří ojetých vozidel bylo u pozorovaného vzorku během všech let shodné a to 8 let (obrázek č. 3). Dalo by se předpokládat, že zájem zákazníků se postupem času bude zaměřovat spíše na mladší vozidla, ale jak již můžeme vidět, data nám to nepotvrzují. Vozový park se tak v české republice nákupem 8letých aut průběžně příliš neomlazuje, jelikož průměrné stáří aut v České republice je necelých 14 let (CEBIA, 2013c).

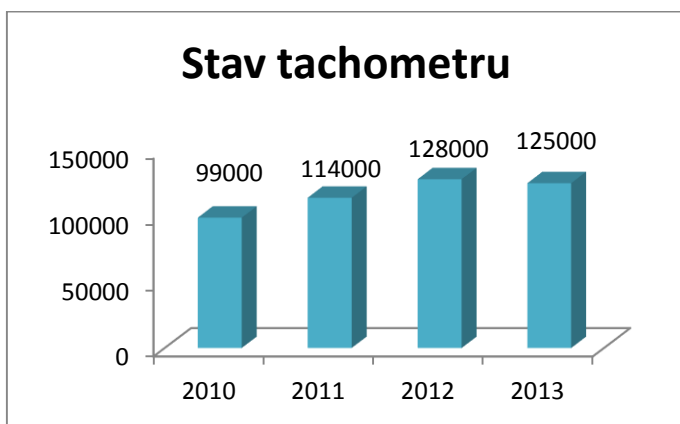
Obrázek 3 Stáří vozidla



Zdroj: (CEBIA, 2013c)

V průběhu posledních let lze zaznamenat serióznější přístup prodejců k uváděným stavům tachometru a to v důsledku existence a působení společností zabývajících se ověřováním původu a historie vozidel. Prodejce se díky těmto společnostem už tak často neopováhá auto tzv. "stočit". Výsledek analýzy si lze paradoxně interpretovat např. tak, že zákazníci více kupují vozidla s vyšším stavem kilometrů tachometru, tedy více ojetá (obrázek č. 4) (CEBIA, 2013c).

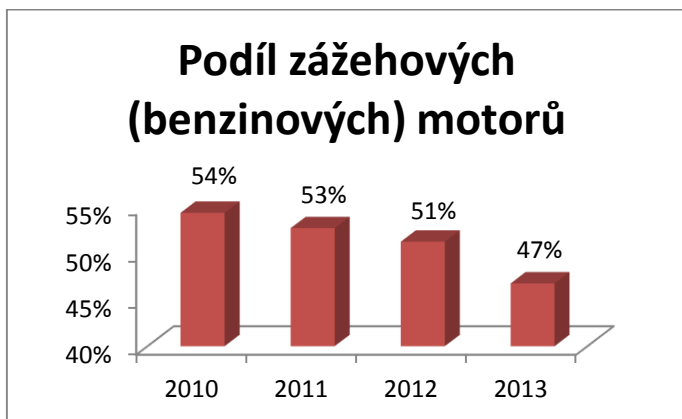
Obrázek 4 Stav tachometru



Zdroj: (CEBIA, 2013c)

Z grafu na obrázku č. 5 je patrné, že prodej benzinových motorů se od roku 2010 snížil o 7% a v současné době zákazníci kupují spíše vozidla s dieselovými motory. Zřejmě právě ekonomická situace zákazníků (i firem) je v tomto případě hlavním důvodem. Zákazníci tak díky nižší spotřebě a levnější naftě snižují své provozní výdaje (CEBIA, 2013c).

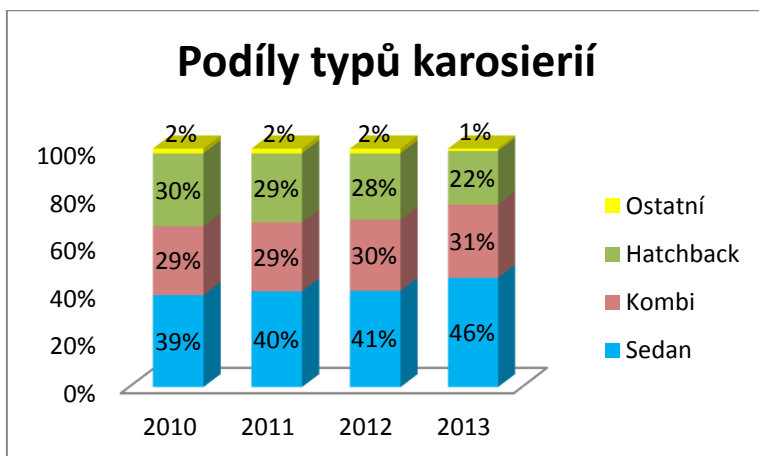
Obrázek 5 podíl zážehových (benzinových) motorů



Zdroj: (CEBIA, 2013c)

Z grafu na obrázku č. 6 je vidět rostoucí zájem o sedany a klesající zájem o hatchbacky. Vozy kombi lákají stále téměř třetinu zákazníků (CEBIA, 2013c).

Obrázek 6 Podíly typů karoserií



Zdroj: (CEBIA, 2013c)

3 Nynější stav ohodnocování ojetých vozidel

Existuje celá řada společností a to i s mezinárodní působností, která poskytuje svým klientům poradenství a informace v oblasti automobilové a dopravní problematiky. Dále je mnoho dalších způsobů či služeb, na základě kterých se lze pro danou koupi rozhodnout. Někteří dali a dávají na své subjektivní rozhodnutí, někteří důvěřují radám známých a jiní využijí například posudku soudního znalce. V dalších podkapitolách jsou uvedeny nejvýznamnější instituce zabývající se touto problematikou.

3.1 Cebia, spol. s.r.o.

Cebia, spol. s.r.o., patří mezi špičku na trhu už více jak 20 let a to v oblasti ověřování původu a historie vozidel. Zároveň tak společnost pomáhá dlouhodobě svým klientům se spolehlivým a hlavně bezpečným nákupem ojetých vozidel. Je poskytována komplexní služba Autotracer, která pracuje v režimu on-line a je připravena poskytnout okamžitou pomoc každému, kdo si chce pořídit ojeté vozidlo (CEBIA, 2012b).

Autotracer Vám podá maximum podstatných informací ohledně Vámi zvoleného vozidla, jako jsou například záznamy o ověření roku výroby, o stavu tachometru, dále informace o kontrole odcizení, o kontrole leasingu, návodu na kontrolu identifikátorů vozidla, servisní historii a mnoho dalších užitečných informací. Tento systém je vymezen zejména leasingovým společností, všem motoristům, pojišťovnám a prodejčům ojetých vozidel. Prostřednictvím osvědčení Cebia Report Vám také může pomoci s prodejem Vašeho stávajícího vozu. Kromě této služby poskytuje i řadu dalších služeb sloužících k prověřování původu vozidel. Tyto služby jsou efektivním nástrojem pro spolehlivý nákup ojetého automobilu, jelikož jsou zaměřeny na odhalování takto „znehodnocených“ vozidel (CEBIA, 2012a).

3.2 Společnost IBS automotive

IBS automotive je společnost s mezinárodní působností, která svým klientům poskytuje informace a poradenství ohledně automobilové a dopravní problematiky. Díky vlastnímu vývoji v oblasti informační techniky a specializaci na moderní komunikační technologie se z odborně orientované firmy v průběhu několika let vytvořil poskytovatel informačních služeb v sektoru automobilového průmyslu a obchodu. Součástí produktové

skupiny této společnosti je systém pro ohodnocování ojetých vozidel TAXexpert (SIKORA, 2013).

Je založen na poměrně rozsáhlé databázi vozidel, jejich technických parametrech včetně výchozích cen a seznamů standardní a doplňkové výbavy. Pro seznámení se s tímto systémem, je zde možnost stáhnutí demoverze programu na stránkách společnosti. V první řadě, je potřeba do programu nastavit vstupní informace, např. druh vozu, model, výrobce, rok výroby, výbava, poškození různých částí vozu, atd., na základě kterých TAXexpert provede propočty a navrhne dvě hodnoty: prodejní a nákupní cenu (IBS automotive, © 2013).

Jak se ukázalo během konzultace s pracovníky jednoho středně velkého autobazaru, není tento systém v praxi mezi prodejci používán. Systém slouží spíše bankovním institucím jakožto poskytovatelům úvěrů. Tito poskytovatelé se na základě toho, co program vyhodnotí rozhodnou, zda kupujícímu úvěr poskytnou a v jaké výši. Taktéž z konzultace byly zjištěny skutečnosti, proč prodejci tento systém v praxi nevyužívají. Jelikož prodejci automobilů rozhodují dle svých zkušeností a opravdu pěkné automobily ocení vyšší, než průměrnou cenou, systém TAXexpert nemá tuto možnost subjektivního ohodnocení, z čehož vyplývá, že velice pěkný vůz podhodnotí. Zákazník tak tedy nedostane na tento vůz úvěr v potřebné výši a zbývající částku tak musí u poskytovatele složit ve formě akontace.

3.3 EurotaxGlass's

EurotaxGlass's patří mezi přední evropské poskytovatele informací týkajících se automobilového obchodu, především vztahujících se k oceňování ojetých vozidel. Zajišťuje nezávislé, odborné ohodnocení vozidla, jeho opravy nebo odhad vývoje ceny vozidla do budoucna. A to díky široké databázi informací o dopravních prostředcích pokrývající osobní, terénní, dodávkové a nákladní automobily, motocykly a specifické dopravní prostředky jako jsou zemědělské stroje nebo lodě. Dále podrobná technická data, která jsou neustále aktualizována. A to celé prostřednictvím širokého portfolia online softwarů (ERE on-line, XchangeNET, Konto on-line ocenění), offline softwarů (ERE manager, EGIS, Forecast, Autowert), dále knižních publikací a individuálním přístupem (EUROTAXGLASS'S, 2008).

3.4 Znalecký standard č. I/2005

Ohodnocení ojetého vozidla je důkladně popsáno ve znaleckých standardech vyhrazených k oceňování motorových vozidel vydaných v roce 2005. V těchto standardech jsou naprosto striktně nastavené stupnice, v jakých může znalec dané vozidlo ohodnotit. Na základě jeho posouzení udává do vzorců popsaných ve standardech různou výši opotřebení. Pomocí těchto vzorců následně dopočítá srážku z původní ceny a tím ohodnotí posuzovaný automobil. Tato možnost ohodnocení není tak často využívána, neboť se za znalecký posudek vynakládají značné finanční prostředky (KREJČÍŘ a BRADÁČ, 2005).

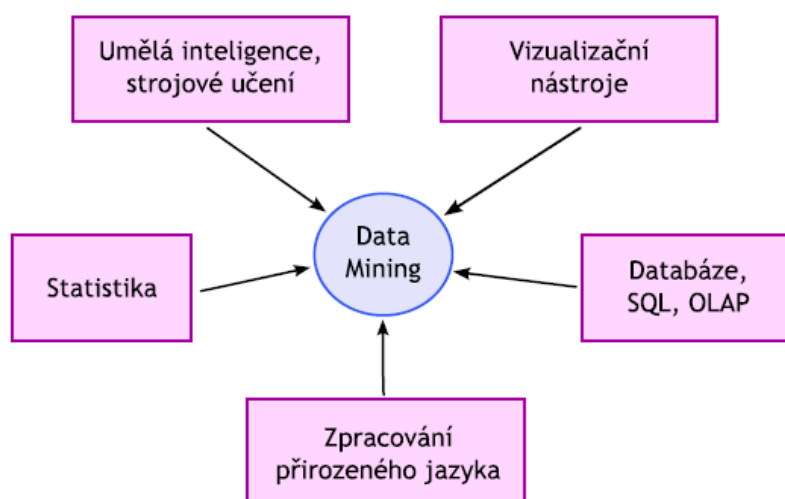
4 Přehled metodik a výpočetních aplikací

Dobývání znalostí z databází, tato oblast informatiky má své zdroje v odvětvích: databáze, statistika a umělá inteligence (strojové učení). Vývoj těchto odvětví probíhal nezávisle do chvíle, kdy velikost automaticky sbíraných dat začala uživatelům přerůstat přes hlavu. Dále když vznikla potřeba používat tato data pro podporu (strategického) rozhodování ve firmách a v době, když se metody strojového učení začaly využívat v oblasti ekonomie.

4.1 Data mining

Data mining můžeme chápat jako obor, který se začal značně vyvíjet na počátku devadesátých let minulého století. Tenkrát docházelo k pozvolnému sjednocování technik a technologií, které se v tehdejší době vyvíjely, co se týče závislosti na sobě zcela samostatně. Databázové technologie představují již několik desetiletí ověřený způsob, jak shromažďovat rozsáhlé objemy dat a tato data pomocí speciálních dotazovacích jazyků (např. SQL) filtrovat. Na druhé straně existuje již několik desetiletí řada matematicko-statistických metod, které jsou prokázaným nástrojem pro analýzu a modelování závislostí v datech. Právě propojením těchto technik a technologií vznikl na počátku devadesátých let minulého století již zmíněný data mining. Na obrázku č. 7 jsou představeny hlavní zdroje data miningu.

Obrázek 7 Zdroje data miningu

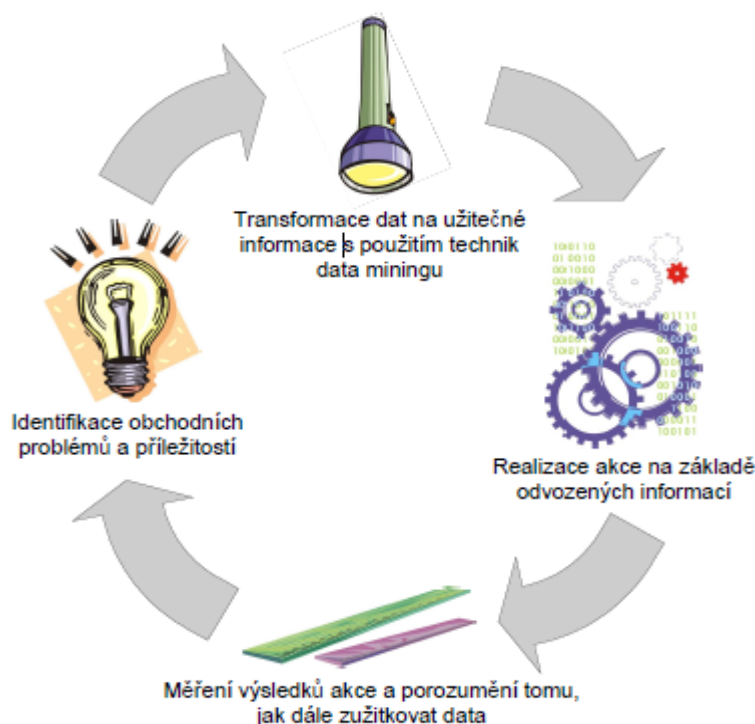


Zdroj: (POŠÍK, 2010)

4.2 Princip data miningu

Podstatným přínosem data miningu je v procesu podpory rozhodování nalezení závislostí a vztahů ukrytých uvnitř velkých množství dat. Samotná identifikace závislostí a vztahů však k získání úplného efektu použití data miningu nestačí. Pro firmy je velmi důležité, aby uměly proměnit informace získané z dat v akci a tuto akci následně v hodnotu. Berry a Linoff označují tento proces jako virtuální cyklus data miningu. (obrázek č. 8) (BERRY a LINOFF, 1997).

Obrázek 8 Jednotlivé fáze virtuálního cyklu DM



Zdroj: (BERRY a LINOFF, 1997)

Z obrázku č. 8 je zřejmé, že virtuální cyklus data miningu je sestaven ze čtyř fází:

1. identifikace obchodního problému, resp. příležitosti,
2. použití technik data miningu k transformaci dat na užitečné informace,
3. realizace obchodní akce na základě odvozených informací,
4. měření výsledků obchodní akce a snaha o identifikaci nového problému nebo příležitosti.

4.3 Metodiky data miningu

S postupem času v oblasti data miningu začaly pro lepší přehlednost a efektivnost vznikat metodiky, které uživatelům jednak přináší jednotný rámec pro řešení různých úloh z oblasti získávání znalostí z dat a dále představují konkrétní využití jednotlivých fází výše zmíněného virtuálního cyklu. V následujících třech podkapitolách jsou charakterizovány metodiky, z nichž nejrozšířenější je CRISP-DM. Jako další budou uvedeny metodiky SEMMA a 5A (BERKA, 2003).

4.3.1 Metodika CRISP-DM

Metodologie CRISP-DM (CROSS-Industry Standard Process for Data Mining) vlastněná firmami NCR Corporation, SPSS Inc., Daimler Chrysler AG a OHRA. SPSS se doložáním z dat zabývá a vytvořila softwarový nástroj IBM SPSS Modeler (Clementine). CRISP-DM realizuje kompletní průběh virtuálního cyklu. Vývoj byl zahájen jako projekt evropské komise definující model standardního postupu při vytváření data miningových projektů (CHAPMAN a kol., © 2000).

Účelem této metodiky je vytvořit návrh všeobecného postupu pro řešení projektů, který bude použitelný ve více komerčních softwarových aplikacích. S tím související návrh na řešení problémů, se kterými se během projektu můžeme setkat. Data miningové úlohy řešené touto metodikou jsou rychlejší, efektivnější a spolehlivější. Taktéž jsou na ně vynaloženy nižší náklady.

Metodologie CRISP-DM rozděluje celý proces data miningového projektu do šesti základních etap, v rámci nichž dále rozlišuje další kroky. Těmito etapami jsou:

Etapa 1 Porozumění problému

Podstatný v této etapě je jednak účel projektu a jeho jednotlivé náležitosti z hlediska obchodního. Musí se zde počítat jak s možnostmi obchodních příležitostí, tak s možnými riziky. Tyto poznatky se dále zformulují na požadavky z analytického hlediska a určí se prozatímní plán prací. V této fázi je taktéž velice podstatné nastavení kritérií pro hodnocení celého projektu (CHAPMAN a kol., © 2000).

Etapa 2 Porozumění datům

Tato etapa začíná bezprostředně získáváním dat potřebných pro analýzu, následně je potřeba se s danými daty seznámit a prozkoumat je například prostřednictvím popisné

statistiky (minima, průměry, četnosti hodnot, maxima atd.) a to vše za účelem zjištění zajímavých podmnožin, zhodnocení kvality atd. (CHAPMAN a kol., © 2000)

Etapa 3 Příprava dat

V této etapě jsou zahrnuty veškeré činnosti, které jsou nutné k vytvoření finální datové sady, která bude zpracovávána data miningovým softwarem. Tedy se kterou bude pracovat data miningový software. Činnosti, které jsou v této fázi prováděny, zahrnují selekce a čištění dat, transformace dat, odvozování dat atd. Tato fáze se obvykle provádí opakovaně, jelikož bývá nejpracnější (CHAPMAN a kol., © 2000).

Etapa 4 Modelování

Zde v první řadě dochází k výběru nejvhodnějšího modelovacího algoritmu pro následnou analýzu dat, což se nejčastěji provádí prostřednictvím data miningového softwaru. Jestliže to situace resp. problém dovoluje, lze pro vyřešení jednoho data miningového úkolu vyzkoušet více technik. Vytvořit tedy více modelů s využitím různých algoritmů, porovnání výsledků, případně jejich kombinace. Některé z metod mají speciální požadavky na formu dat, což si často může vyžádat návrat o krok zpět do etapy přípravy dat (CHAPMAN a kol., © 2000).

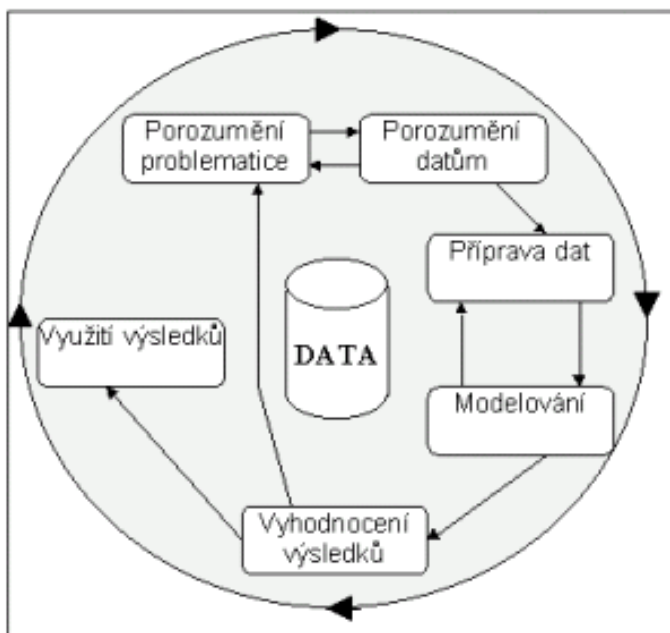
Etapa 5 Vyhodnocení výsledků

Co se týče této etapy, samozřejmě je v tuto chvíli k dispozici vytvořený model (resp. modely), který by se z pohledu analýzy měl jevit jako kvalitní. Cílem je zhodnotit jeho správnost z hlediska obchodního, zkrátka zda bylo dosaženo všech předem stanovených cílů. Pokud tedy model, který byl označen za výsledný a tak tedy vyhovující a schopný plnit obchodní potřeby, je na řadě kompletní revize data miningové úlohy, aby se prověřilo, zdali nebyl opomenut nějaký důležitý úkol (CHAPMAN a kol., © 2000).

Etapa 6 Využití výsledků

Implementace výsledků do obchodních procesů je fáze, kde se výsledky data miningové analýzy zavádějí do obchodních procesů firmy, současně se stanoví plány týkající se kontroly a údržby vytvořeného modelu (CHAPMAN a kol., © 2000).

Obrázek 9 Jednotlivé fáze metodiky CRISP-DM



Zdroj: (EASY MINER, 2013)

Na obrázku č. 9 je zachycen životní cyklus projektu data miningu podle metodiky CRISP-DM. Sled jednotlivých fází není nijak striktně určen, jelikož závisí na výstupu každé z nich a na základě těchto výstupů se rozhoduje o postupu k další fázi, nebo naopak o návratu k fázi předchozí. Uvedené šipky ve schématu znázorňují nejdůležitější a současně také nejčastější závislosti mezi fázemi. Co se týče vnější kružnice, představuje virtuální cyklus data miningu. Po uplatnění této metodiky a měření výsledků jsou k dispozici nové obchodní příležitosti, které následně směřují k provedení další data miningové úlohy.

4.3.2 Metodika 5A

Metodika využívaná firmou SPSS, tedy tvůrcem data miningového softwaru. Název metodiky je zkratkou pro jednotlivě prováděné kroky (BERKA, 2003):

- Assess (posouzení) – zhodnocení požadavku projektu.
- Access (získávání) – seskupení požadovaných dat
- Analyze (analyzování) – provedení analýz.
- Act(provedení) – přeměna znalostí na znalosti vykonávané
- Automate (automatizace) – interpretace výsledku analýzy v praxi a následné uplatnění

4.3.3 Metodika SEMMA

Metodika využívaná firmou SAS, vycházející z vlastní metodiky pro získávání znalostí z databází. Jednotlivá písmena názvu SEMMA opět vystihují jednotlivě prováděné kroky (BERKA, 2003):

- Sample (vzorek) – volba příslušných dat
- Explore (poznávání) – prozkoumání, redukce a vizualizace těchto dat
- Modify (úprava) – tvorba veličin, transformace dat, vytváření veličin
- Model (modelování) – datová analýza
- Assess (posouzení) – srovnání a srozumitelné vyjádření modelů

4.4 Typy úloh pro data mining

Data mining mnohdy řeší problémy ze spousty různých oborů, tedy mnohdy zdánlivě nesouvisející problémy. Je zapotřebí přiřadit příslušnému problému skupinu, do které spadá a to ve chvíli, než je zvolena konkrétní metoda pro modelování problému. Abychom docílili splnění data miningového problému, je možné zvolit více metod, které do skupiny spadají. Je mnohdy žádoucí, co se týče závěru mít k dispozici srovnání více výsledků různých metod. Samozřejmě je vhodné k porovnání metod znát jejich výhody a nevýhody. Ucelená podoba, tedy přesné rozdělení typů úloh neexistuje, avšak nejčastěji se v literatuře uvádějí základní typy úloh, které umožňují řešit techniky data miningu, patří sem především (BERKA, 2003):

- Klasifikace
- Predikce
- Hledání nuggetů

Na obrázku č. 10 jsou pro bližší představu v rámci jednoho konceptu zobrazeny znalosti, které jsou v jednotlivých typech úloh řešeny.

4.4.1 Predikce a klasifikace

Predikce představuje prognózu, tedy předpověď hodnot budoucích na základě znalosti hodnot dalších, ve většině případů minulých. Má obdobnou funkci jako klasifikace nebo odhad. Nalézání znalostí, které pomohou nová data roztřídit (klasifikovat) do předem

určených skupin je cílem právě těchto úloh. Princip těchto úloh spočívá ve výběru jednoho cílového atributu A, modelování vlivu ostatních atributů na atribut A. Na úkor srozumitelnosti je dána přednost většímu počtu znalostí, které pokrývají daný problém.

Mezi nejčastěji používané techniky pro predikci a klasifikaci patří např. rozhodovací stromy, neuronové sítě, regrese. Blíže je o těchto technikách pojednáno v následující podkapitole (BERKA, 2003).

4.4.2 Deskripce

Jak je patrné z názvu, deskripce charakterizuje (popisuje) data jako celek. Účelem této úlohy je získávání spíše méně, za to opravdu srozumitelných znalostí vystihující celý problém. Odhaluje nové, ještě nepoznané zákonitosti (znalosti), které se podílejí na rozvoji lidských poznatků ve zkoumané oblasti.

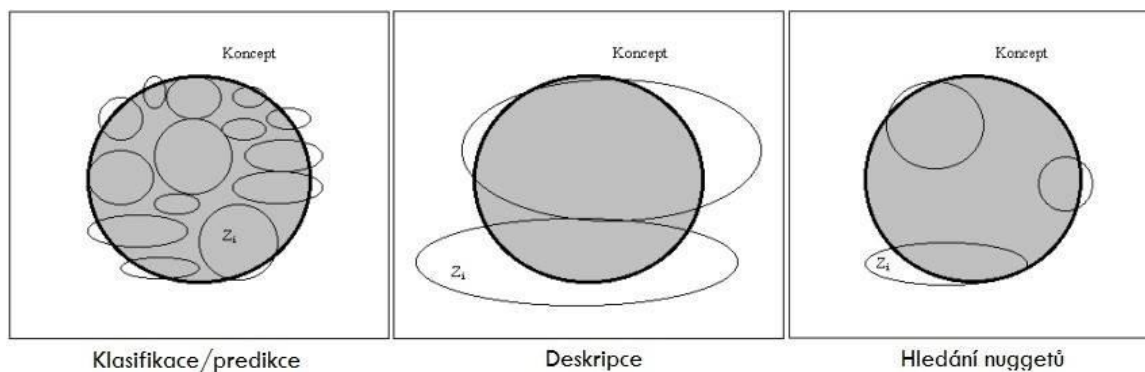
Mezi vhodné techniky pro úlohu deskripce patří např. shluková analýza, sumarizace, vizualizace (BERKA, 2003).

4.4.3 Hledání nuggetů

Hledání nuggetů je do značné míry podobné deskripci. Cílem této úlohy je taktéž získávání nových srozumitelných znalostí, ovšem s tím rozdílem, že tyto znalosti nemusí pokrývat celý problém, ale pouze jeho zajímavé části.

Techniky, které se nejčastěji využívají při hledání nuggetů, jsou např. asociační pravidla, segmentace, vzorkování, shluková analýza (BERKA, 2003).

Obrázek 10 Zobrazení typů data miningových úloh



Zdroj: (BERKA, 2003)

4.5 Základní techniky data miningu

K dosažení cílů data miningových úloh, lze využít různé techniky, které se mohou ve všech výše uvedených typech úloh prolínat. z tohoto důvodu nemůžeme každé úloze přesně přiřadit jen jednu techniku, proto je v dalších dvou podkapitolách uvedený přehled vhodných technik, pomocí kterých bychom měli dosáhnout cílů výše uvedených úloh.

4.5.1 Statistické techniky

Statistické metody jsou teoreticky dobře prozkoumané a z praxe prověřené výpočetní algoritmy sloužící pro analýzu dat. Pro dolování dat mají význam (BERKA, 2003):

- **Diskriminační analýza** – metoda, která se využívá ke klasifikaci pozorování do dvou nebo více skupin. Vybraný počet členění na skupiny závisí na očekávaném cíli analýzy a na uspořádání dat, která jsou k dispozici.
- **Regresní analýza** – statistická metoda, která představuje celou řadu známých, v minulosti naměřených hodnot a prostřednictvím této analýzy předvídáme vývoj hodnot budoucích.
- kde známe celou řadu v minulosti naměřených hodnot a pomocí regresní analýzy předpovídáme vývoj hodnot do budoucna.
- **Shluková analýza** – cílem je rozčlenit datový soubor do tzv. „shluků“ (skupin) a to na základě jejich podobnosti. Cílem shlukování je tedy současně vytvořit a nalézt tyto skupiny, které nejsou předem známé.

Mezi další techniky, které bychom mohli uplatnit, jsou: korelační analýza, faktorová analýza, analýza rozptylu či kontingenční tabulky.

4.5.2 Techniky umělé inteligence (nestatistické metody)

Neuronové sítě

Využitím neuronových sítí se pokoušíme přiblížit činnosti běžných neuronů mozku. Tedy neuronové sítě pracují analogicky s lidským mozkiem. Obdobně jako mozek jsou tvořeny spoustou vzájemně propojených neuronů. V těchto sítích je neuron považován jako buňka, která přijímá podněty od dalších neuronů, které jsou k ní napojeny na vstupu. Neuronová síť transformuje (přetváří) vstupy na výstupy prostřednictvím různých spojení mezi neurony. Tento výsledek je poté srovnán s výsledkem reálně naměřeným. Pokud účinek vstupních dat překročí stanovenou mez, neuron se aktivuje a začne sám svým výstupem působit na další neurony. z hlediska dolování dat je jednou z důležitých vlastností schopnost těchto modelů se učit z příkladů. Proces učení probíhá po krocích, ve kterých jsou upravovány váhy vazeb (spojení) mezi neurony, až do té doby, dokud celková chyba neklesne pod předem stanovenou a přijatelnou hladinu. Až poté je možné u takto připravené sítě prezentovat vstupy modelované situace a odečítat výstupy. Neuronové sítě jsou nejčastěji využívány pro vyhledávání podobností, vzorů, tvorbu prediktivních modelů. Výsledky se však často těžko interpretují (BERKA, 2003).

Rozhodovací stromy

Prediktivní model zobrazující data v podobě stromu. Popis modelu pomocí rozhodovacího stromu je řadou jednoduchých pravidel. Na vrcholu stromu stojí nadřazená všeobecná množina, která se poté dělí do množin dle určitého kritéria. Tyto množiny se dále dělí do podmnožin opět dle dalšího kritéria a to tak, že předešlé množiny jsou jim nadřazené. Jak z textu vyplývá, data v jednotlivých množinách se vyznačují stejnými vlastnostmi. Nespornou výhodou rozhodovacích stromů je jejich schopnost vysvětlit výsledky (BERKA, 2003).

Genetické algoritmy

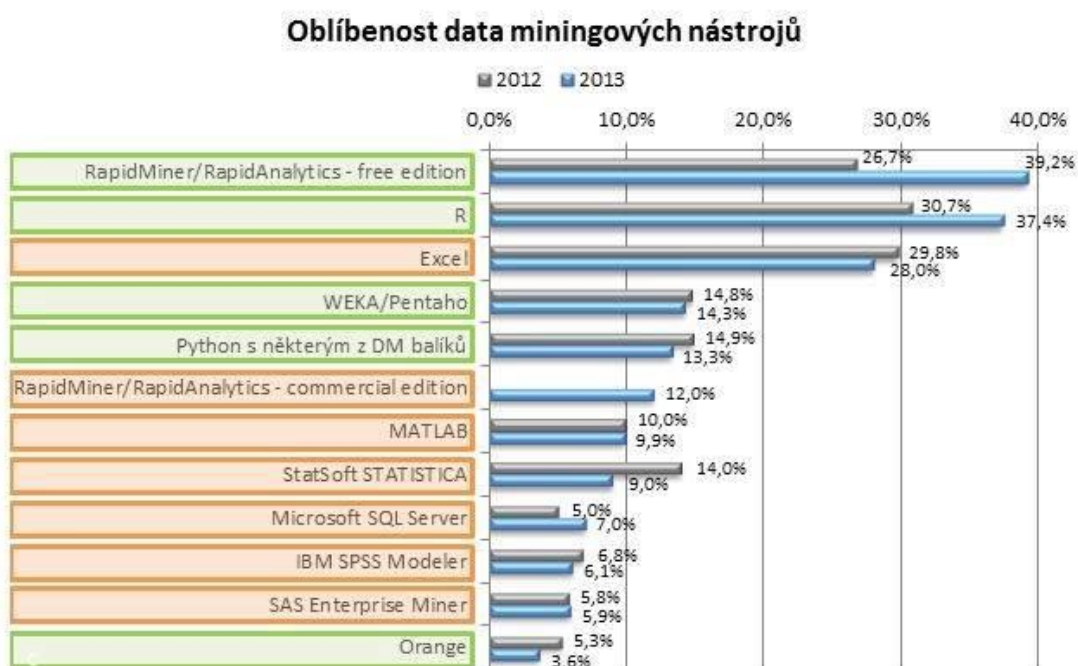
Patří mezi evoluční algoritmy, tedy výpočty (mechanismy, modely) založené na evolučním procesu v přírodě, tzn. pomocí párování, mutací a klonování, ve kterém DNA představují 0 a 1, se z původních modelů vyvíjí čím dál tím vhodnější model pro zadanou úlohu. Původní modely, které stály na začátku evoluce, mohou být řešeny různými technikami dolování dat. Tato technika vyžaduje výkonově náročné počítačové zpracování, což není v dnešní době, kdy výkonnost počítačů stále roste, žádný problém. Proto si také tato metoda získává čím dál větší popularitu (BERKA, 2003).

5 Trendy v oblasti modelovacích nástrojů

V současné době existuje celá řada neustále se rozvíjejících softwarových systémů sloužících pro řešení data miningových úloh. Uživatelé si mohou opatřit jednak volně šiřitelné systémy, které jsou zdarma k dispozici na internetu, stejně tak mohou využívat placené verze těchto systémů, které oproti neplaceným nabízejí širší nabídku v rámci modelovacích nástrojů.

Na obrázku č. 12 je zobrazena oblíbenost využívaných data miningových nástrojů. Jedná se o výzkum, ve kterém hlasovalo 1880 uživatelů, kteří odpovídali na otázku: „Jakého analytického data miningového nástroje Vámi bylo využito v průběhu tvorby reálného projektu během posledních dvanácti měsíců? (květen 2012 - 2013)?“ Bylo vybráno dvanáct produktů, které se umístily mezi prvními dvaceti. Výsledky jsou porovnávány s tentýž výzkumem prováděným o rok dříve (květen 2011 – 2012). Volně šiřitelné systémy jsou označeny zelenou barvou a systémy placené oranžovou (KD nuggets, 2013).

Obrázek 11 Využití data miningových nástrojů dle oblíbenosti



Zdroj: (KD nuggets, 2013)

Pro další snadnější orientaci v těchto systémech, je jejich rozdělení do několika skupin dle společných znaků. Právě jedním z těchto znaků je účel, se kterým jsou tyto softwary vytvářeny. Producenti vyvíjejí jednak softwary sloužící primárně pro data mining a v druhé řadě softwary s jiným primárním využitím, např. matematickým, statistickým atd. (HOLEŇA, 2006).

V praxi se zhruba ve 30-40% pro dobývání znalostí z databází používají i systémy, které nejsou vytvářeny primárně za tímto účelem, ale nabízejí data mining pouze jako doplňkové rozšíření. Což se projevuje i na obrázku č. 12, kde se na prvních příčkách objevují i programy, které nejsou čistě jen záležitostí data miningu. Jsou to tyto typy softwarů (HOLEŇA, 2006):

- Databázové systémy: Oracle, Microsoft SQL Server
- Statistické systémy: systémy SAS, StatSoft Statistica, SPSS, projekt R
- Matematické systémy: MATLAB

Ve dvou níže uvedených podkapitolách jsou představeny tři nejznámější komerční systémy a tři nejznámější volně dostupné systémy.

5.1 Komerční systémy

IBM SPSS Modeler (Clementine)

V současnosti je SPSS Modeler v oblasti data miningu jedním z nejvíce rozšířených systémů. Vytvořen byl nejprve firmou Integral Solutions a to pod názvem Clementine. Roku 1999 se firma Integral Solutions spolu se světoznámou firmou SPSS, která se zabývá vývojem statistického softwaru, sjednotily. Od této doby zastřešuje vývoj softwaru firma SPSS. V roce 2009 byla firma SPSS zakoupena proslulou firmou IBM. V důsledku této změny se upustilo od názvu Clementine a je využíváno názvu - IBM SPSS Modeler. Poskytuje celou řadu vstupů dat do programu, je zde možné využití mnoha data miningových modelovacích metod, přičemž výsledné modely si může uživatel zobrazovat prostřednictvím databází, tabulek a různých vizualizačních prostředků. Celý tento systém je založen na metodologii CRISP-DM (HOLEŇA, 2006), (KHU DHUR, 2009).

SAS Enterprise Miner

Enterprise Miner je dalším z řady data miningových systémů. Jeho vývoj probíhá od roku 1998 firmou SAS, která patří mezi špičku na trhu zabývající se statistickým softwarem. Dále je Enterprise Miner jedním z mnoha modulů, které lze v rámci rozsáhlého softwaru SAS zakoupit. s dalšími produkty této firmy je ovšem plně kompatibilní. Nabízí využití mnoha implementovaných metod s eventuelním využitím vnitřního programovacího jazyka. Prostředí pro uživatele je grafické, tedy uživatelsky příjemné. Tento systém vychází z metodiky SEMMA (HOLEŇA, 2006), (SAS, 2014).

StatSoft STATISTICA Data Miner

STATISTICA Data Miner je třetím nejrozšířenějším nástrojem zaměřeným na data mining. Je produktem firmy StatSoft, která byla založena roku 1984 v USA. Obdobně jako oba výše uvedené příklady těchto nástrojů, je i tento vyvíjen firmou, která se ve své činnosti prvotně věnuje statistice. O uživatelském prostředí se i zde dá říci, že je uživatelsky příjemné a přizpůsobivé. Metod data miningu je v tomto programu nabízena široká škála. Finální modely je možné vytvořit jako spustitelný kód v mnoha programovacích jazycích, např. Visual Basic, Java apod. V případě tohoto nástroje je výhodou možnost využití systému Web Statistica, který umožňuje v prostředí internetového prohlížeče projektovat a řídit práce vztahující se k data miningu na jiném počítači (STATSOFT, 2014).

5.2 Volně dostupné systémy

Rapid Miner 6

Rapid Miner je v současné době jedním z nejvyhledávanějších a tak tedy nejvyužívanějších volně šiřitelných softwarů pro data mining. To dokazují i výsledky výzkumu na obrázku č. 12, kde můžeme vidět, že se Rapid Miner umístil na prvním místě. Dále můžeme vidět, že ani placená verze tohoto programu nezůstává pozadu a co se týče využití, nachází se zhruba uprostřed. Navíc placená verze je rozšířena o další využití. Lze ho spustit téměř na jakémkoli operačním systému, jelikož je kompletně napsán v programovacím jazyce Java (RAPID MINER, 2013).

WEKA

Systém WEKA zastupuje kategorii nekomerčních, volně dostupných systémů. Tento program byl původně vyvíjen za vysokoškolskými účely na univerzitě Waikato, nachá-

zející se na Novém Zélandě. Hlavním účelem bylo vytvořit sadu technik data miningu pro posílení novozélandské ekonomiky, zejména zemědělství. Program poskytuje nástroje pro klasifikaci, předzpracování, shlukování, regresi a vizualizaci. V průběhu vývoje byl přepsán do jazyka Java, byl přizpůsoben všeobecnému využití a je šířen pomocí GNU General Public License (GNU GPL). Weka je složkou svobodného softwaru nazývaného Pentaho, jenž se orientuje na business inteligenci a sjednocuje několik zprvu samostatných produktů. Projektu se stále daří i přes konkurenci v podobě Rapid Mineru (PENTAHO, 2006), (WEKA 3, 2013).

Orange

Orange stejně jako nadcházející program vznikl v prostředí univerzity ležící ve slovinské Lublani a to přímo na fakultě zabývající se počítačovými a informačními vědami. Stejně tak patří mezi volně dostupné programy. Program podporuje vizuální programování či skriptovací jazyk Python. k naprogramování dalších nových funkcí lze využít jazyka C++ (GOOGLE SUMMER OF CODE, 2013).

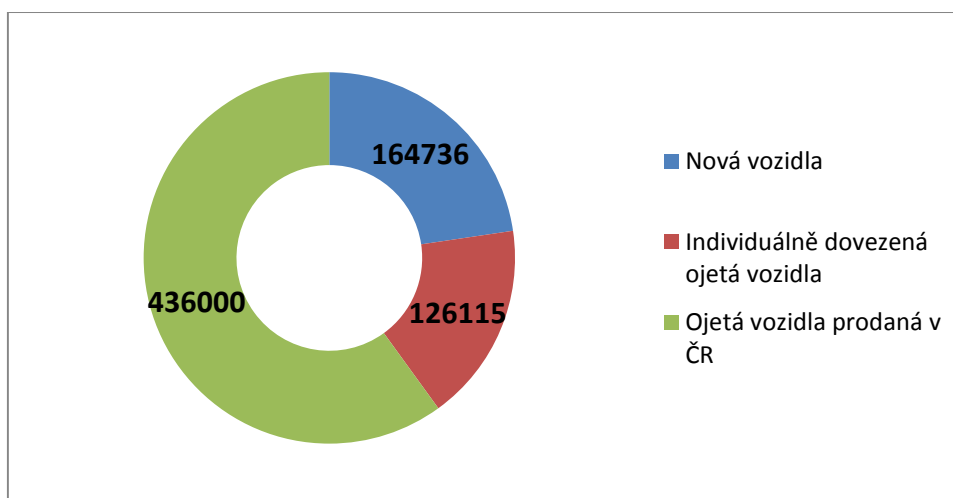
6 Současný stav prodeje ojetých vozidel

V této kapitole si zobrazíme v první řadě souhrnné statistiky o trhu ojetých vozidel za rok 2013, které byly zpracovány a jak je uvádí společnost Cebia, která se touto problematikou zabývá. V druhé řadě si pro srovnání, co se týče celkového počtu registrovaných vozidel v České republice za rok 2013, uvedeme statistiku prvních registrací vozidel poskytovanou Svazem dovozců automobilů (SDA).

6.1 Cebia statistika

Data byla zpracovávána na vzorku ojetých vozidel prodaných v České republice v loňském roce. Dále bylo čerpáno z dat SDA, informací uveřejněných autobazary a z dat dovozců vozidel. Jelikož je trh ojetých vozidel tak rozsáhlý a proniká do neprozkoumaných oblastí, např. malých autobazarů nebo prodejců, mezi soukromými osobami, nezbývá nic jiného, než výsledky zhodnotit odhadem. Na obrázku č. 12 je zobrazen celkový prodej vozidel v České republice za rok 2013 (CEBIA, 2014).

Obrázek 12 Celkový prodej vozidel v ČR za rok 2013



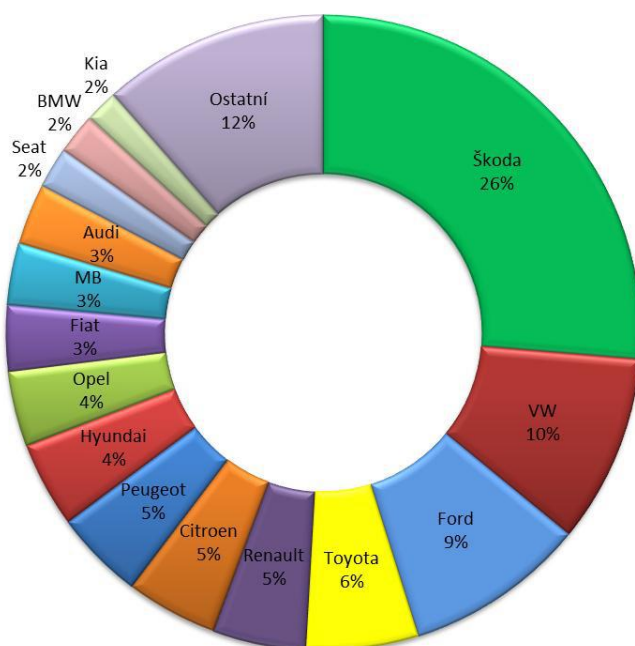
Zdroj: (CEBIA, 2014)

Tržní podíly značek ojetin

I v tomto odvětví se začal zaznamenávat nárůst prodeje nových vozidel značek Hyundai nebo Kia. Rostoucí značka Dacia se prozatím na trhu ojetých vozidel příliš neuplatnila. z hlediska tržních podílů jednotlivých značek byla v roce 2013 na prvním místě Škoda, i když v průběhu let, se její podíl postupně snižuje. Pro představu v roce 2011 její podíl dosahoval 30%, v roce 2012 činil 28%. Podíly ostatních jednotlivých značek a jejich

pořadí se nijak značně nemění. Byl zaznamenán nárůst prodeje značek Hyundai nebo Kia v oblasti ojetých vozidel, kdežto například rostoucí značka Dacia se na trhu ojetých vozidel zatím moc neprosadila (CEBIA, 2014).

Obrázek 13 Tržní podíly značek ojetin

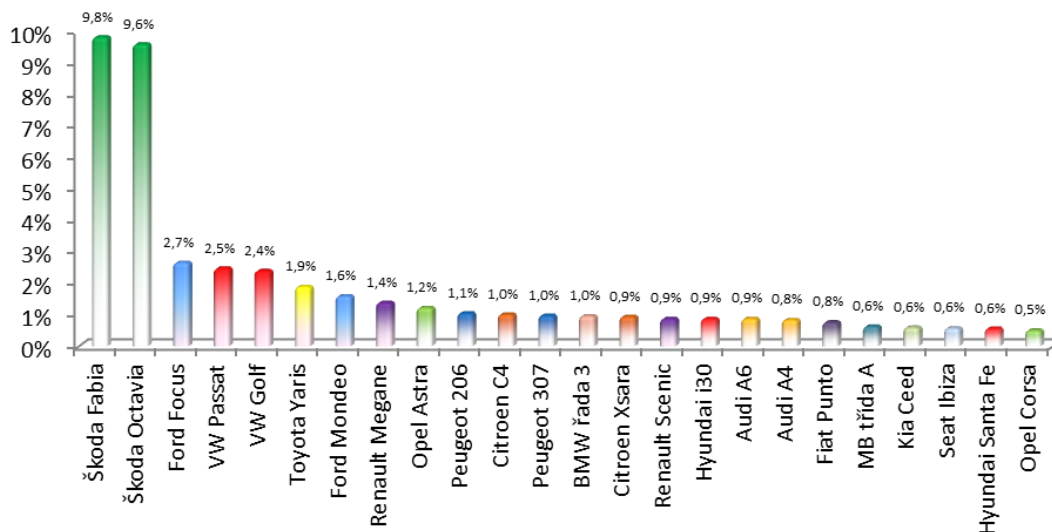


Zdroj: (CEBIA, 2014)

Nejprodávanejší modely

Jak můžeme zaznamenat na obrázku č. 15, mezi nejprodávanejší modely za rok 2013 opět suverénně patří modely Fabia a Octavia značky Škoda, s takřka stejným 10 % podílem. Avšak oproti předchozím letům vymizely nyní staré modely, jako například Felicia a naproti tomu mezi nejprodávanejší modely se dostaly Hyundai i30 nebo Kia Ceed (CEBIA, 2014).

Obrázek 14 Nejprodávanější modely

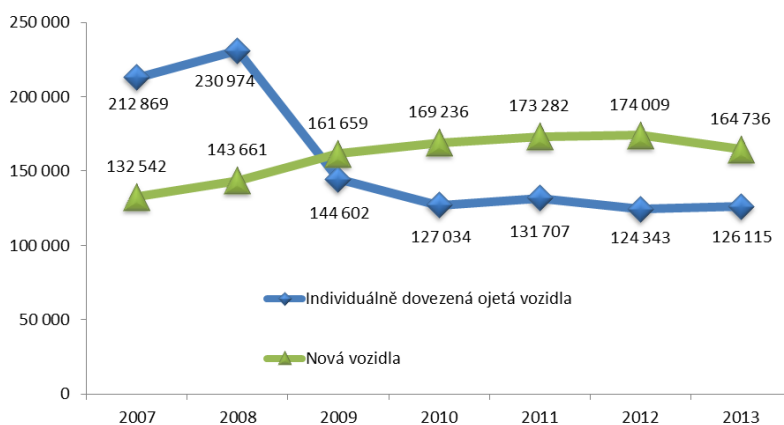


Zdroj: (CEBIA, 2014)

Vývoj osobních dovozů vozidel

V této statistice se pojednává o dovezené osobní automobily prvně registrované v České republice. z obrázku č. 16 lze usoudit, že co se týče posledních 4 let, počet dovezených ojetých vozů se ustálil a ročně se tedy dováží přibližně stejné množství těchto vozidel. Na druhou stranu, prodej nových osobních automobilů po letech růstu v roce 2013 poklesl a to o více jak 5% (CEBIA, 2014).

Obrázek 15 Vývoj individuálních dovozů vozidel



Zdroj: (CEBIA, 2014)

Stáří vozového parku v ČR

Stáří vozového parku v České republice se postupně zvyšuje. Za následek to mají nízké prodeje nových aut, dovoz starých ojetých automobilů ze západu a nadále rostoucí zájem českých zákazníků o ojeté vozy. z celkového počtu skoro 35% představovaly vozy starší 15 let, 27 % vozy starší 10 – 15 let. Od konce roku 1998 je to poprvé, kdy se průměrný věk osobních aut dostal nad hranici 14 let (CEBIA, 2014).

Obrázek 16 Stáří vozového parku



Zdroj: (CEBIA, 2014)

Trh ojetých vozidel z pohledu dalších parametrů

Nejžádanějšími barvami ojetých vozidel jsou šedá a modrá. Pro představu celosvětově nejoblíbenějšími barvami jsou u nových vozidel s 25% barva bílá, stříbrná a černá s 18%. Dle statistik SDA v roce 2012 bylo v České republice téměř 22% všech prodaných aut bílé barvy, kdežto ještě dva roky předtím to bylo přibližně o polovinu méně. Dalo by se tedy očekávat, že až tyto vozy budou v průběhu několika dalších let prodávány jako ojeté, vzroste procentuální zastoupení v prodeji bílých vozidel (CEBIA, 2014).

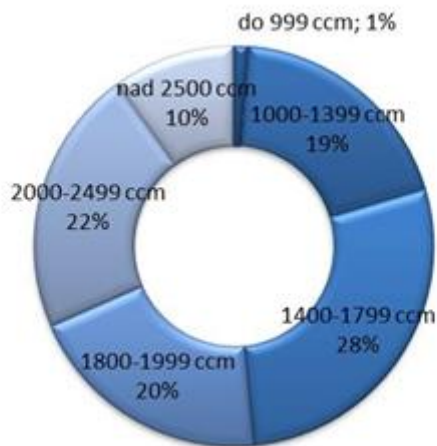
Obrázek 17 Barvy



Zdroj: (CEBIA, 2014)

Za rok 2013 se prodávala vozidla s průměrnou hodnotou kubatury 1800 ccm. Je tomu tak jak u dieselových, tak u benzinových motorů. Naprosto minimální prodej byl zaznamenán u aut s obsahem do 1 litru. z průzkumu tedy vyplývá, že zákazníci mají zájem o automobily s motory o stále vyšších výkonech i přesto, že vozidla s vyšším obsahem motoru mají dražší povinné ručení (CEBIA, 2014).

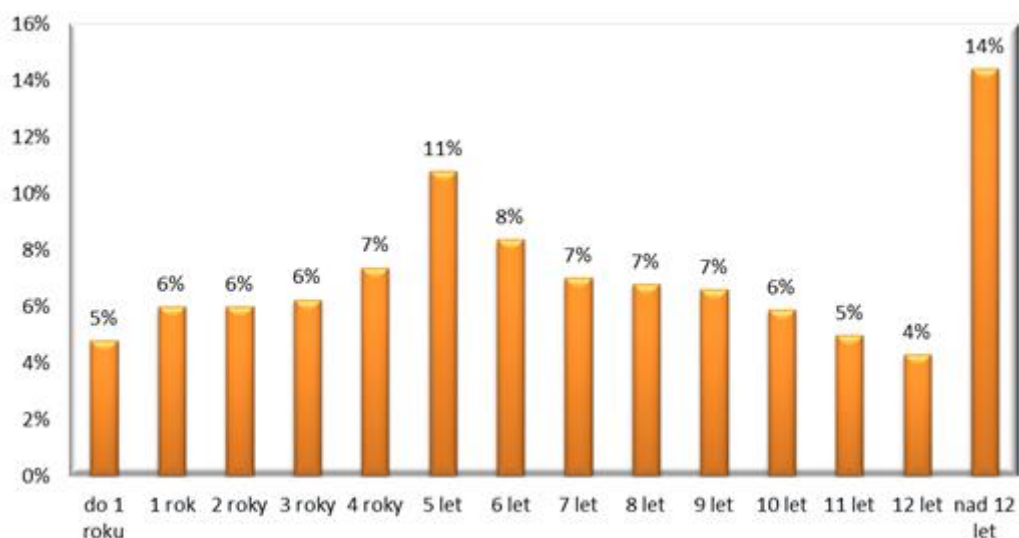
Obrázek 18 Obsah motoru



Zdroj: (CEBIA, 2014)

Nejvíce prodaných vozidel během roku 2013 bylo 5 let starých, tzn., že byla vyrobena v roce 2008. Avšak průměrné stáří vozidel bylo 7,5 roku. Mezi těmito údaji byl rozdíl zapříčiněn vcelku velkým množstvím vozů starších 12 let, což způsobilo zvýšení průměrného stáří. Podstatným parametrem, jak můžeme shlédnout na obrázku č. 20 je četnost prodeje automobilů konkrétního stáří (CEBIA, 2014).

Obrázek 19 Stáří ojetých vozidel



Zdroj: (CEBIA, 2014)

U bezmála 50 % v loňském roce prodaných automobilů se uváděný stav tachometru pohyboval mezi 100 – 200 tisíci kilometry. Průměrná hodnota uvádějící se na tachometru představovala 123 tisíc kilometrů. Pouze 12% vozidel mělo na tachometru uvedeno přes 200 tisíc, což však neodpovídá realitě. Právě vozidla, která opravdu mají přes 200 tisíc najetých kilometrů, jsou „stáčena“ na hodnotu pod 100 tisíc a můžeme říci, že graf za minulý rok by vypadal zcela jinak (CEBIA, 2014).

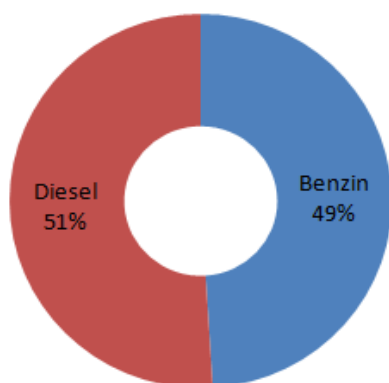
Obrázek 20 Stav na tachometru



Zdroj: (CEBIA, 2014)

Zdali se jedná o automobily se zážehovými (benzinovými) nebo vznětovými (naftovými) motory, je v tomto ohledu zájem zákazníků rozdělen vcelku rovnoměrně. Od roku 2010 se prodej benzinových motorů snížil o 7% a v roce 2013 se prodávaly spíše dieselové motory (obrázek č. 22). Pravděpodobně tomu tak bylo z důvodu ekonomické situace zákazníků či firem. Právě v důsledku této situace a zároveň nižší spotřebě a levnější naftě se tak snížili jejich provozní výdaje. s ohledem na loňské srovnání cen obou paliv a zároveň snižující se spotřebě benzinových motorů s malým objemem, lze očekávat, že opět poroste zájem o benzinové motory (CEBIA, 2014).

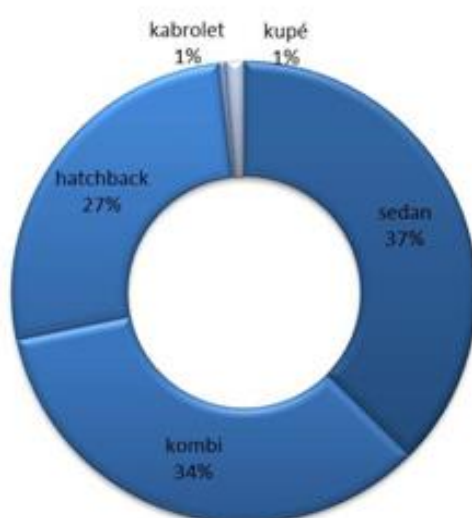
Obrázek 21 Palivo



Zdroj: (CEBIA, 2014)

Stejně jako v předchozích letech, ani v roce 2013 se nijak značně nezměnil zájem o ojetá vozidla, co se týče typu karoserie. Největší zájem zákazníci mají o sedany, o něco menší zájem o vozy kombi a nejmenší zájem je o hatchbacky. Ovšem dá se říci, že poměr mezi těmito segmenty je vcelku souměrný. Kupé a kabriolety se podílejí na celkových prodejkách pouze kolem 1% (CEBIA, 2014).

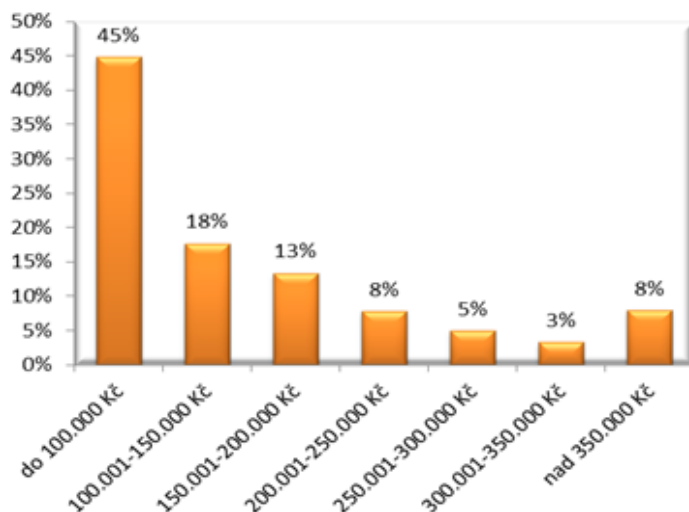
Obrázek 22 Karoserie



Zdroj: (CEBIA, 2014)

Zcela převažující je zájem zákazníků o ojetá vozidla do 100 tisíc korun. Jak je zřejmé na obrázku č. 24, prodejní cena klesá se zvyšující se prodejní cenou. V roce 2013 byla průměrná prodejní cena 155 tisíc korun (CEBIA, 2014).

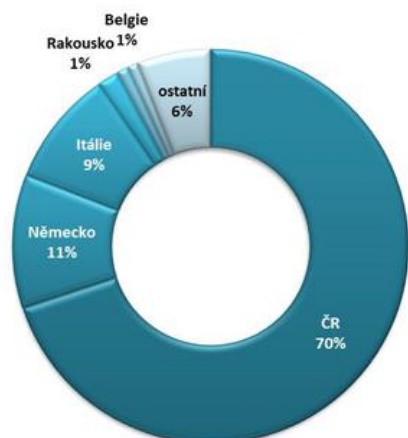
Obrázek 23 Prodejní cena



Zdroj: (CEBIA, 2014)

Zhruba tři čtvrtiny ojetých vozidel pochází z tuzemska, to znamená, že vozidla byla v roce 2013 koupena v ČR. Přibližně 60% těchto vozidel bylo v minulosti dovezeno do České republiky a zbývajících 40% bylo původně v České republice prodáno a zaregistrováno jako nová vozidla. Zcela určitě se většina českých zákazníků při koupi vozidla poohlíží v České republice. A to může být jednak z důvodu, že pocítují určité obavy z individuálního dovozu, ovšem svou roli určitě hraje pohodlí pořízení ojetého automobilu v tuzemsku. Ohledně vozidel provozovaných v České republice je pro zákazníka většinou přístupných mnohem více informací a možnost snadnějšího ověření historie a původu vozidla a tak minimalizace rizik spojených s nákupem vozidla (CEBIA, 2014).

Obrázek 24 Původ vozidel

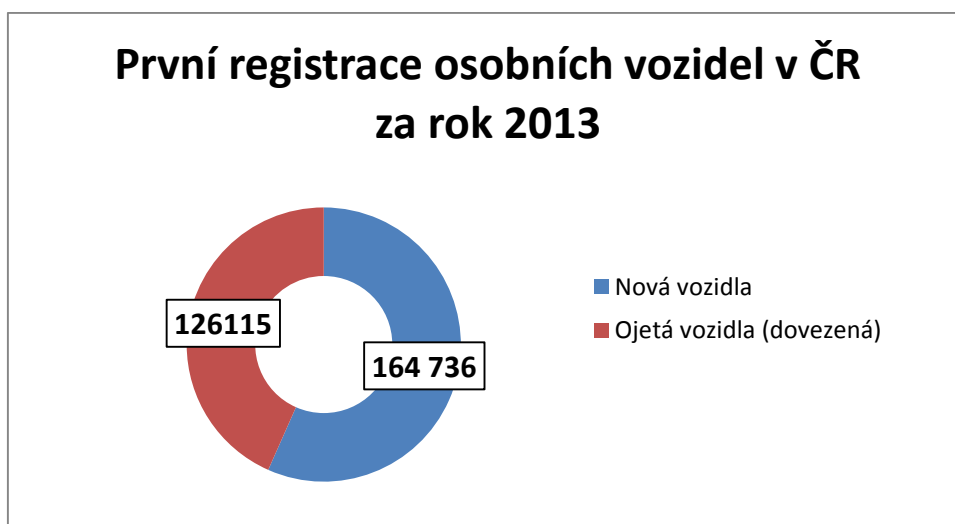


Zdroj: (CEBIA, 2014)

6.2 SDA statistika

Z obrázku č. 25 můžeme vidět celkový počet prvních registrací osobních automobilů v České republice za rok 2013. Za tento rok je registrováno celkem 290 851 osobních automobilů, z toho 164 736 nových vozidel a 126 115 ojetých vozidel (dovezených). Podíl ojetých vozidel z dovozu představuje zhruba 43 %. Oproti roku 2012 to představuje zhruba 2% roční nárůst (SDA, 2014).

Obrázek 25 Registrace osobních vozidel v ČR za rok 2013



Zdroj: (SDA, 2014)

7 Závěr

Záměrem bakalářské práce byla analýza českého trhu prodeje ojetých vozidel a možné způsoby ohodnocování těchto vozidel. Z provedené rešerše vyplývá, že k serióznímu provedení odhadu ceny ojetého vozidla je nutné zpracování velkého množství vstupních dat. Jedná se o tzv. „data mining“ (dolování dat z databází), přičemž pro jejich zpracování existuje řada metodik a modelovacích nástrojů.

Provedená analýza trhu ojetých vozidel v ČR přinesla několik dalších zajímavých zjištění. Problematikou ojetin se zabývá celá řada společností (např. CEBIA, EUROTAXGLASS'S, SDA, Auto SAP, Svaz dovozců automobilů, SDA, Sdružení ochrany vlastníků automobilů a řada dalších). Ve většině případů se shodují v hodnocení rizik a neseřízností mnohých prodejců, avšak výrazně se rozcházejí v hodnocení počtu prodeje ojetých vozidel. Souhrnné statistiky, které byly vyhodnoceny ze strany Svazu dovozců automobilů, oproti informacím společnosti Cebia vykazují značné rozdíly. SDA čerpá data z centrálního registru vozidel, který je považován dle mého názoru za důvěryhodnější zdroj. Dle SDA bylo za rok 2013 registrováno 290 851 vozidel a z toho 126 000 ojetých, kdežto Cebia za tento rok vyhodnotila celkový počet prodaných ojetých automobilů na více než 526 000.

Z hlediska analýzy trhu ojetých vozidel v ČR za rok 2013, která byla realizována společností Cebia z pohledu několika klíčových parametrů, bylo dále dosaženo těchto výsledků. Nejprodávanější značkou byla vyhodnocena Škoda a to s modely Fabia a Octavia. Počet ojetých vozidel, která byla do České republiky individuálně dovezena, dlouhodobě stagnuje. Skoro tři čtvrtiny automobilů pochází z tuzemska. Stáří vozového parku v České republice překročilo 14 let. Mezi nejoblíbenější barvy ojetých vozidel patří šedá a modrá. Nejvíce se během roku 2013 kupovala ojetá vozidla s obsahem motoru od 1400 do 2000ccm. Nejvíce se za tento rok prodávala vozidla vyrobená v roce 2008. U bezmála 50% prodaných automobilů se uváděný stav tachometru pohyboval mezi 100 – 200 tisíci kilometry. Zcela převládá zájem o ojeté vozy do 100 tisíc korun. Zákazníci projevovali zájem takřka stejnoměrně o automobily s benzinovými a dieselovými motory, stejně tak rovnoměrný byl zájem o sedany a vozy kombi, o něco méně o hatchbacky.

Trh ojetých vozidel v ČR za rok 2013 z hlediska jeho velikosti představuje tyto charakteristiky. Během roku 2013 se v České republice prodalo více než 562 000 ojetých

osobních automobilů, což představuje 8% meziroční růst. Celý objem trhu představoval bezmála 90 miliard korun při průměrné ceně vozidla 159 000 korun. Během loňského roku bylo osobně dovezeno a zaregistrováno 126 115 ojetých automobilů a zbylých 436 000 zakoupeno v tuzemsku, což je zhruba 3,5 krát více. Mezi soukromými osobami se z celkového počtu prodaných ojetých vozidel prodalo 302 000 a 260 000 bylo prodáno prodejci ojetých vozidel, především autobazary. Co se týče nových osobních automobilů, těch bylo v roce 2013 prodáno 164 736, z čehož lze říci, že ojetých vozidel se prodává přibližně 3,5 krát více než nových.

Nemohou být opomenuty stinné stránky tohoto trhu. Je stále co zlepšovat a předcházet tak mnohým nástrahám vůči zákazníkům ze strany nepoctivých autobazarů a prodejců, se kterými se na tomto trhu i nadále setkáváme. Tato rizika je třeba určitým způsobem eliminovat. Vozidla, která jsou dovážena do České republiky, by měla podléhat neúprosné kontrole původu. Měla by být náležitě odhlášena původním majitelem, k jejich registraci by mělo být povinné potvrzení o tom, že tato vozidla nejsou kradená, bouraná, že na nich nezůstává žádný závazek (dluh) jako například nesplacený leasing apod.

Na základě provedené analýzy lze konstatovat, že zákazníci, kteří si dříve pořizovali nová vozidla, se v současné době a to především z finančních důvodů orientují spíše na vozidla ojetá a mají zájem o kvalitní ojetá vozidla s prokazatelnou historií. Jedná se tak pro ně o levnější variantu. Tudíž poptávka o tato vozidla roste a s přehledem lze tento trh považovat za prosperující.

8 Seznam odborné literatury

BERKA. *Dobývání znalostí z databází*. Vyd. 1. Praha: Academia, 2003, 366 s. ISBN 80-200-1062-9.

BERRY a LINOFF. *Data mining techniques: for marketing, sales, and customer support*. New York: Wiley, c1997, x, 454 p. ISBN 04-711-7980-9.

CEBIA. *Cebia Autotracer: Informace o systému Autotracer* [online]. © 2012a [cit. 2014-01-04]. Dostupné z: <http://www.zkontrolujsiauto.cz/info-o-systemu-autotracer>.

CEBIA. *Cebia summary 2/2013: Informace, statistiky a zajímavosti z oblasti prodeje ojetých vozidel* [online]. Praha, 2013c [cit. 2014-01-05]. Dostupné z: http://www.cebia.cz/o-spolecnosti/napsali-o-nas/tiskove-zpravy.html/1_308-cebia-summary-2-2013/1.

CEBIA. *Cebia summary 3/2013: Informace, statistiky a zajímavosti z oblasti prodeje ojetých vozidel* [online]. 2013d [cit. 2014-01-04]. Dostupné z: http://www.cebia.cz/o-spolecnosti/napsali-o-nas/tiskove-zpravy.html/1_330-cebia-summary-3-2013/1.

CEBIA. *Cebia summary 4/2013: Informace, statistiky a zajímavosti z oblasti prodeje ojetých vozidel* [online]. 2014 [cit. 2014-01-11]. Dostupné z: http://www.cebia.cz/o-spolecnosti/napsali-o-nas/tiskove-zpravy.html/1_355-cebia-summary-4-2013/1.

CEBIA. *Profil společnosti* [online]. © 2012b [cit. 2014-01-11]. Dostupné z: <http://www.cebia.cz/o-spolecnosti/historie-a-profil.html>.

EASY MINER. *Výuka data miningu* [online]. 2013 [cit. 2014-01-05]. Dostupné z: <http://zt.vse.cz/easyminer/>.

EUROTAXGLASS'S. *O společnosti Eurotaxglass's* [online]. © 2008 [cit. 2014-01-05]. Dostupné z: http://www.eurotaxglass.cz/various/about_us.php.

GOOGLE SUMMER OF CODE. *Orange* [online]. 2013 [cit. 2014-02-15]. Dostupné z: <http://orange.biolab.si/trac/wiki/GSoC>.

HOLEŇA, Martin. *Statistické aspekty dobývání znalostí z dat*. 1. vyd. Karolinum, 2006, 106 s. ISBN 80-246-1186-4.

CHAPMAN a kol. *Step-by-step data mining guide* [online]. © 2000 [cit. 2014-02-08]. Dostupné z:

<ftp://ftp.software.ibm.com/software/analytics/spss/support/Modeler/Documentation/14/UseManual/CRISP-DM.pdf>.

KD NUGGETS. *What Analytics, Big Data, Data mining, Data Science software you used in the past 12 months for a real project?* [online]. 2013 [cit. 2014-02-15]. Dostupné z:

<http://www.kdnuggets.com/polls/2013/analytics-big-data-mining-data-science-software.html>.

KHUDHUR. CIO BUSINESS WORLD. *Novým vlastníkem společnosti SPSS je IBM* [online]. 2009 [cit. 2014-02-22]. Dostupné z: <http://businessworld.cz/aktuality/novym-vlastnikem-spolecnosti-spss-je-ibm-5234>.

KREJČÍŘ a BRADÁČ. *Znalecký standard číslo I/2005*. Brno: Cerm, 2005. ISBN 978-80-7204-370-5.

PALLA. SDRUŽENÍ OBRANY SPOTŘEBITELŮ. *Vyjádření pro Svět motorů k prohrěškům autobazarů* [online]. 2013 [cit. 2014-03-29]. Dostupné z:

<http://www.asociace-sos.cz/vyjadreni-pro-svet-motoru-k-prohreskum-autobazaru/>.

PENTAHO. *Pentaho Acquires Weka Project* [online]. 2006 [cit. 2014-02-15]. Dostupné z:

http://www.pentaho.com/press-room/releases/20060919_pentaho_acquires_weka.

POŠÍK. *Data mining, strojové učení, základní úlohy, učení s učitelem a bez učitele* [online]. 2010 [cit. 2014-03-08]. Dostupné z:

https://cw.felk.cvut.cz/wiki/_media/courses/y33aui/01introtoml.pdf.

RAPID MINER. *Rapid Miner* [online]. 2013 [cit. 2014-02-16]. Dostupné z:

<http://rapidminer.com/news-posts/rapidminer-v6-0/>.

SAS. *SAS Enterprise Miner* [online]. 2014 [cit. 2014-02-16]. Dostupné z:

http://www.sas.com/en_us/software/analytics/enterprise-miner.html#section=1.

SVAZ DOVOZCŮ AUTOMOBILŮ. *Výroční zprávy* [online]. 2014 [cit. 2014-03-23]. Dostupné z: <http://portal.sda-cia.cz/clanky/download/2013-Vyrocní-zprava-SDA-Annual-report.pdf>

SDRUŽENÍ NA OCHRANU VLASTNÍKŮ AUTOMOBILŮ. *O nás* [online]. © 2004 - 2014 [cit. 2014-03-16]. Dostupné z: <http://www.sdruzeni-sova.cz/clanek/proc-existuje-sova-2002>.

SIKORA. IBS AUTOMOTIVE. *Profil společnosti* [online]. © 2013 [cit. 2014-01-25]. Dostupné z: <http://www.ibsautomotive.cz/o-spolecnosti/>.

STATSOFT. *STATISTICA Data Miner* [online]. 2014 [cit. 2014-01-25]. Dostupné z: <http://www.statsoft.cz/produkty/5-dataminingove-nastroje/21-statistica-data-miner/detail/>.

WEKA 3. *Data Mining Software in Java: WEKA: The university of Waikato* [online]. 2013 [cit. 2014-01-26]. Dostupné z: <http://www.cs.waikato.ac.nz/ml/weka/>.

9 Použité zkratky

CRISP-DM – Cross Industry Standard Process for Data Mining

ČR – Česká republika

DM – Data Mining

EU – Evropská Unie

GNU GPL – GNU General Public License

IBM – International Business Machines

MATLAB – MATrix LABoratory

SAS – Statistical Analysis System

SDA – Svaz dovozců automobilů

SOS – Self Organizing Map

SPSS – Statistical Package for the Social Sciences

SQL – Structured Query Language

WEKA – Waikato Environment for Knowledge Analysis

10 Seznam obrázků

OBRÁZEK 1 CENA VOZIDLA	4
OBRÁZEK 2 OBSAH MOTORU	5
OBRÁZEK 3 STÁŘÍ VOZIDLA	6
OBRÁZEK 4 STAV TACHOMETRU.....	6
OBRÁZEK 5 PODÍL ZÁŽEHOVÝCH (BENZINOVÝCH) MOTORŮ.....	7
OBRÁZEK 6 PODÍLY TYPŮ KAROSÉRIÍ	7
OBRÁZEK 7 ZDROJE DATA MININGU.....	11
OBRÁZEK 8 JEDNOTLIVÉ FÁZE VIRTUÁLNÍHO CYKLU DM	12
OBRÁZEK 9 JEDNOTLIVÉ FÁZE METODIKY CRISP-DM	15
OBRÁZEK 10 ZOBRAZENÍ TYPŮ DATA MININGOVÝCH ÚLOH	18
OBRÁZEK 11 VYUŽITÍ DATA MININGOVÝCH NÁSTROJŮ DLE OBLÍBENOSTI	21
OBRÁZEK 12 CELKOVÝ PRODEJ VOZIDEL V ČR ZA ROK 2013.....	25
OBRÁZEK 13 TRŽNÍ PODÍLY ZNAČEK OJETIN	26
OBRÁZEK 14 NEJPRODÁVANĚJŠÍ MODELY	27
OBRÁZEK 15 VÝVOJ INDIVIDUÁLNÍCH DOVOZŮ VOZIDEL.....	27
OBRÁZEK 16 STÁŘÍ VOZOVÉHO PARKU	28
OBRÁZEK 17 BARVY	29
OBRÁZEK 18 OBSAH MOTORU	29
OBRÁZEK 19 STÁŘÍ OJETÝCH VOZIDEL	30
OBRÁZEK 20 STAV NA TACHOMETRU	31
OBRÁZEK 21 PALIVO	31
OBRÁZEK 22 KAROSERIE	32
OBRÁZEK 23 PRODEJNÍ CENA	33
OBRÁZEK 24 PŮVOD VOZIDEL.....	34
OBRÁZEK 25 REGISTRACE OSOBNÍCH VOZIDEL V ČR ZA ROK 2013	34