

ANALÝZA VYTÍŽENOSTI ŽELEZNIČNÍ INFRASTRUKTURY

Cíle

Cílem práce je provést časovou analýzu vytiženosti železniční infrastruktury. Na základě poskytnutých dat od Správy železnic bude provedeno vyhodnocení dat na železniční infrastrukturu v České republice za období roku 2016 až 2019. V první části budou data předzpracována do vhodné podoby pro další zpracování v této diplomové práci i mimo ni. V další části bude provedena analýza časové složky dat, tak aby bylo možné popsat časovou variabilitu vybraných úseků v rámci ČR. Poté bude provedeno několik kategorizací bodových vstupních dat, na základě rozdílných metod a atributů, které budou popsány. Kategorizace a analýza časových řad budou doprovázeny vhodnými mapovými výstupy a grafy.

Postup zpracování

teoretická část

- legislativa týkající se železnice v ČR
- seznámení s železniční sítí v ČR a jejím správcem
- seznámení se s odbornými a akademickými pracemi zabývajícími se podobnými tématy
- seznámení se s problematikou shlukování a časových řad

praktická část

- předzpracování dat – základní seznámení s daty a zpracování do podoby vhodné pro další analýzy
- tvorba základních vizualizací – tvorba vizuálních pomůcek pro pochopení problematiky, kterou data obsahují
- kategorizace pomocí shlukování – tvorba několika kategorizací železničních dat na základě různé parametrizace
- analýza časových řad – popis časové složky dat pomocí rozkladu časové řady na komponenty
- tvorba grafů a mapových výstupů v prostředí GIS – zobrazení výsledků pomocí grafů a vhodných kartografických metod
- tvorba posteru a webové stránky

Použitá data

Klíčovou částí této diplomové práce jsou data o vytižení železniční infrastruktury poskytnutá státní organizací SŽ ve formátu XLSX. Data se používají jako primární zdroj při výpočtu ceny za využití dráhy jednotlivými dopravci. Data jsou reprezentována měsíčními úhrny informací o vytižení za jednotlivé druhy kolejové dopravy. Jsou rozdělena na dvě části s odlišnou strukturou. První částí jsou data za jednotlivé úseky mezi monitorovacími body (2016-2019), druhým typem dat jsou úhrny za jednotlivé monitorovací body (2016 a prvních 10 měsíců 2017). Atributy se dají rozdělit do čtyřech částí. V první části je označení období, tedy měsíce a roku z kterého datový záznam pochází. V druhé části jsou popisné informace úseku, jde o informace o poloze, kódovým označení a názvu počátečního a koncového bodu. Ve třetí části jsou obsaženy informace o kategorii dopravy daného záznamu. V poslední části jsou samotná číselná data o využití železniční infrastruktury. Některé atributy jsou absolutní za celý měsíc, některé průměrné.

Kategorizace

Pro zařazení všech monitorovacích bodů do kategorií byla použita metoda shlukování a zařazení na základě četnosti nákladní a osobní dopravy. Data byla použita z roku 2016, protože jde o jediná bodová data za celý rok, která byla poskytnuta. Bohužel nelze z úsekových dat získat data bodová, jinak by bylo možné kategorizovat na základě



nejnovějších dat z roku 2019. Do kategorizace byly vybrány různé kombinace atributů a následně byla zvolena optimální řešení a ta interpretována. Byly použity jak roční agregace rozdílných atributů, tak měsíční agregace stejných atributů. Pro shlukování byl použit software Orange. Konkrétně šlo o hierarchické shlukování Wardovou metodou při použití euklidovské metriky. Výsledný počet shluků byl nastaven na pět. Tabulka s přidaným sloupcem, který označoval jednotlivé shluky, byla zobrazena v ArcGIS Pro a na jejím základě byl vytvořen mapový výstup. Výsledkem této části jsou mapové výstupy všech pěti způsobů kategorizace, jejich interpretace a vzniklé mapové sady.

Analýza časových řad

Vzhledem k tomu, že získaná úseková data o vytižení železniční infrastruktury, byla získána v měsíčních úhrnech za 4 roky, od roku 2016 do roku 2019, bylo vhodné je dále analyzovat z hlediska času. Pro tento úkol je vhodná zejména analýza časových řad, kde bylo možné popsat chování jednotlivých ukazatelů napříč časem, sledovat opakující se vzorce v datech, jejich trend a další prvky. Kvůli velkému množství poskytnutých dat nebylo možné provést analýzu každého monitorovacího bodu nebo úseku na železniční síti v ČR. Konkrétně se jednalo o 3800 úseků za každý rok. Bylo tedy nutné na základě znalosti dat získané předchozím zkoumáním a vlastní znalostí zmenšit zkoumané množství úseků. Bylo to provedeno vydefinováním význačných kategorií, které mají zajímavý roční průběh, nebo jsou zajímavé z hlediska nějakého fenoménu v železniční přepravě.

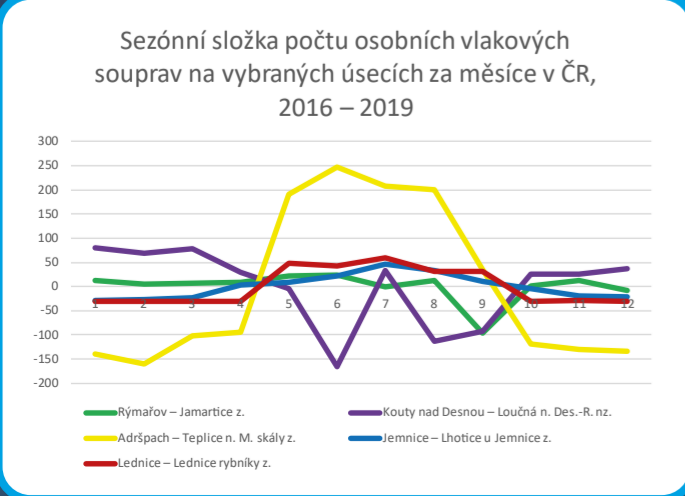
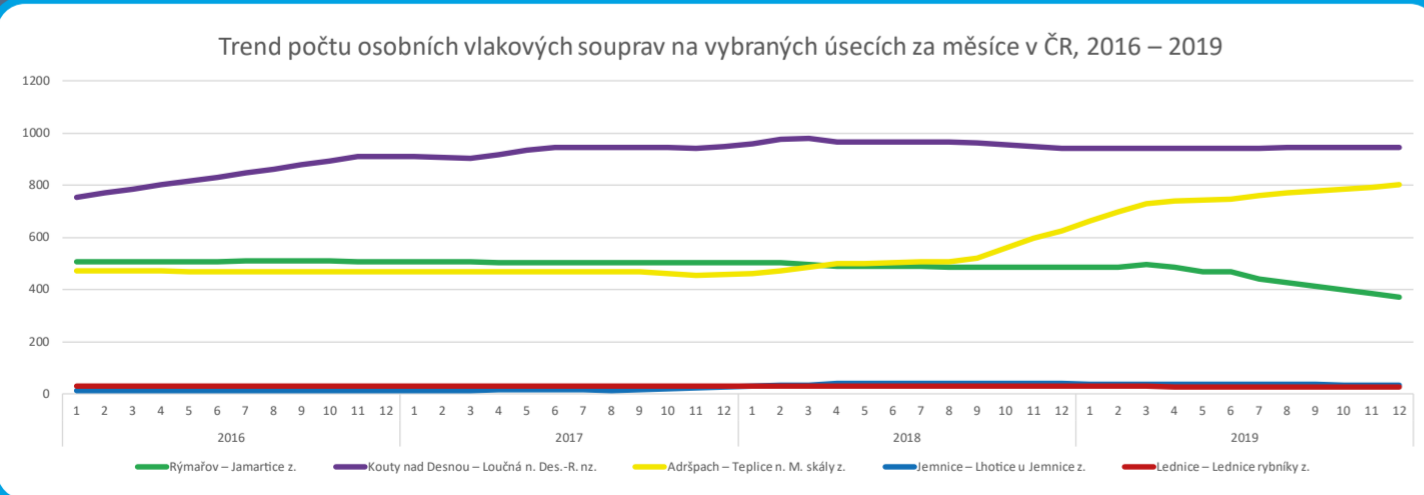
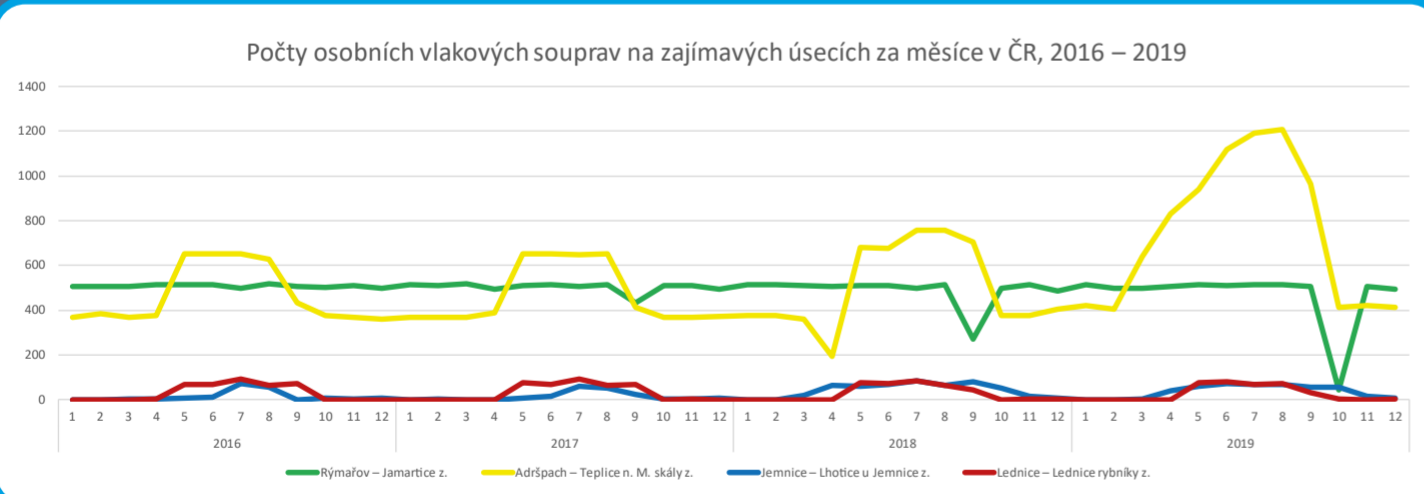
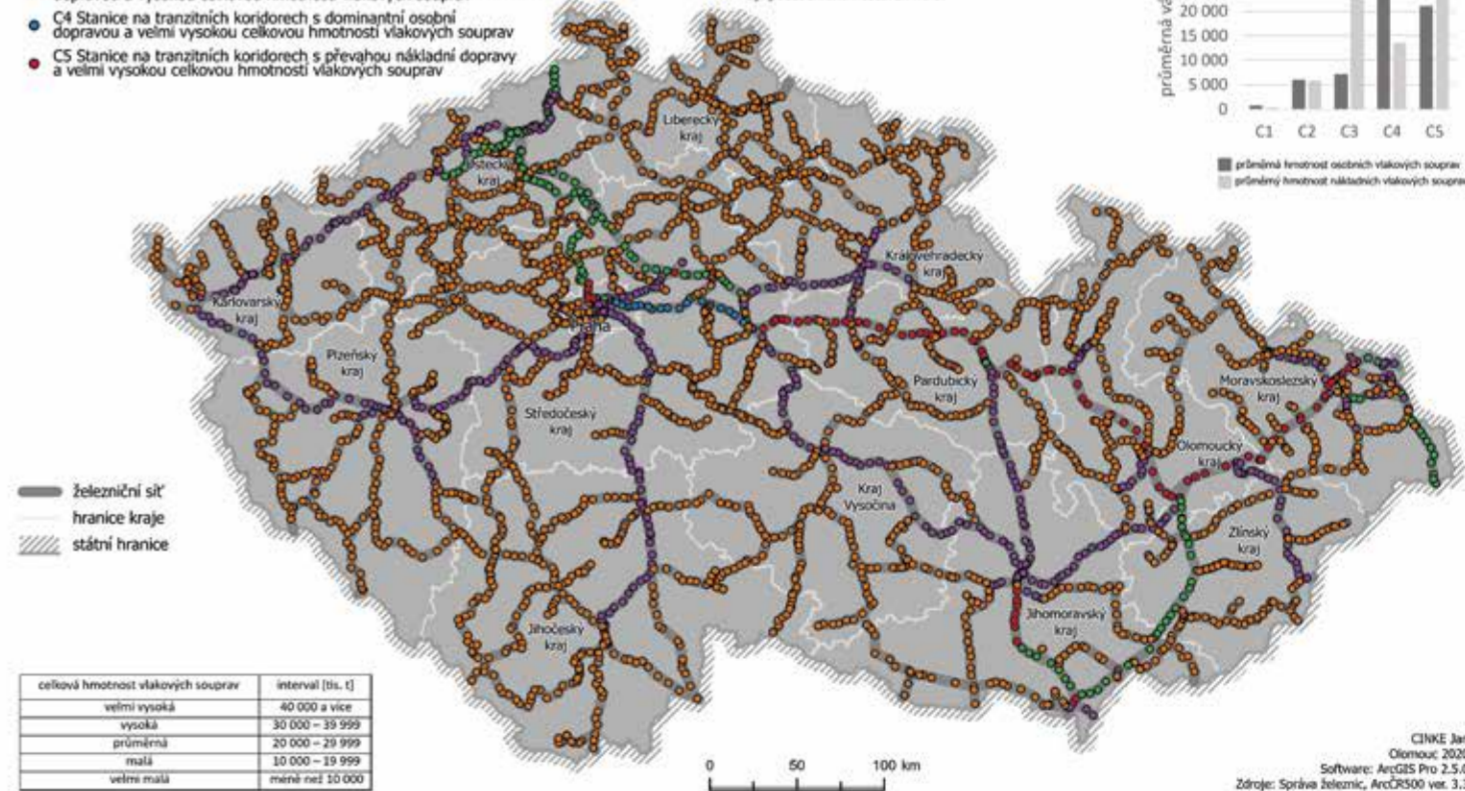
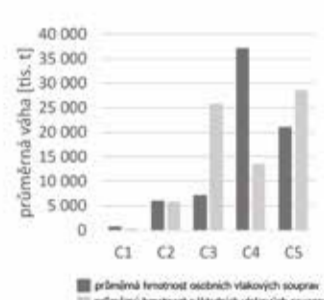
První vybranou kategorií byly **tranzitní koridory**. Jde o nejvytiženější trasy napříč ČR s regionální, národní i mezinárodní dopravou. Vybrány byly takové úseky, které leží mezi významnými železničními uzly na již zmíněných tranzitních koridorech. Další a velmi zajímavou kategorií jsou **hraniční přechody**. Zde je použití úseků velmi vhodné. Protože přes hranice vede vždy v datech pouze jedna linka. Zároveň je úsek vždy vybrán až po státní hranici, takže nejsou uvažovány vlakové soupravy, které končí na území ČR, ale pouze ty vlakové soupravy, které hranice reálně přejíždějí. V této kategorii byly vybrány úseky u všech hraničních přechodů, které leží na tranzitních koridorech. Poslední kategorií byla společná pro úseky z hlediska osobní a nákladní dopravy a byla pojmenována **zajímavé úseky**. Vybrány byly takové úseky, ve kterých je zvýšený nákladní provoz nebo zajímavý osobní provoz. Původcem jsou zde především tratě s regionálním osobním provozem a jinak využívané tratě pro nákladní dopravu.

KATEGORIZACE ŽELEZNIČNÍCH STANIC, DOPRAVNĚ ZAJÍMAVÝCH A TARIFNÍCH MÍST Z ČÍSELNÍKU SR70 V ČR V ROCE 2016

B2. Hmotnost osobních a nákladních vlakových souprav za rok

- C1 Stanice na regionálních a nadregionálních tratích s převahou osobní dopravy a velmi malou celkovou hmotností vlakových souprav
- C2 Stanice na osobních a nákladních koridorech a významných tratích s nepatrnou převahou osobní dopravy a malou celkovou hmotností vlakových souprav
- C3 Stanice na osobních a nákladních koridorech s dominantní nákladní dopravou a vysokou celkovou hmotností vlakových souprav
- C4 Stanice na tranzitních koridorech s dominantní osobní dopravou a velmi vysokou celkovou hmotností vlakových souprav
- C5 Stanice na tranzitních koridorech s převahou nákladní dopravy a velmi vysokou celkovou hmotností vlakových souprav

Tato mapa vznikla jako výstup diplomové práce Analýza vytiženosti železniční infrastruktury. Kategorie byly získány na základě shlukování atributů hmotnosti osobních a nákladních vlakových souprav. Bylo použito hierarchické shlukování, Euklidovská metrika a Wardova metoda, kde byl počet shluků nastaven na 5.



Pro analýzu časových řad byla vybrána analýza formou aditivní dekompozice. Pro realizaci analýzy byl vybrán software Orange. Konkrétně šlo o doplněk Timeseries s nástroji pro zpracování časových řad. Výsledkem této části jsou interpretace jednotlivých kategorií pro osobní a nákladní dopravu realizované na základě průběhu měsíčních hodnot, trendu a sezónní složky. Například jde o porovnání hraničních přechodů z hlediska nákladní dopravy. Pro porovnání průběhu hodnot v rámci celého časového období, byly použity grafy s vnesenými hodnotami počtu osobní a nákladní dopravy vyrovnanými na ekvivalentní délku měsíce 30 dní, aby nebyly počty zkreslené.

Závěr

Tato práce je pilotním příkladem zpracování těchto unikátních dat o vytižení železniční infrastruktury v takovémto formátu, která byla takto uvolněna zcela poprvé. Výsledné analýzy nabízejí zajímavý pohled na situaci železniční dopravy na sítích SŽ v ČR. Kompletní pohled na variabilitu monitorovacích bodů dávají kategorizace, u kterých je v rámci práce nabídnuto hned několik variant. První částí je kategorizace na základě procentuálního vyjádření poměru typu dopravy z hlediska počtu nebo hmotnosti vlakových souprav. Další částí jsou kategorizace na základě shlukování počtu nebo hmotnosti měsíčních nebo ročních úhrnů dat. Všechny kategorizace jsou podpořeny mapovými výstupy a náležitě okomentovány. Bližší pohled na časovou variabilitu vybraných kategorií stanic ze čtyř po sobě jdoucích let, dává analýza časových řad. Zde je situace chování v rámci tohoto časového období popsána s přesností na jednotlivé měsíce v rámci čtyř let, s důrazem kladeným zejména na hlavní tranzitní koridory, hraniční přechody a dopravně zajímavé monitorovací body.