

Technická univerzita v Liberci

FAKULTA PŘÍRODOVĚDNĚ-HUMANITNÍ A PEDAGOGICKÁ

Katedra: Geografie
Studijní program: Učitelství pro 2. stupeň základní školy
Studijní obor (kombinace) Informatika - Zeměpis

**ANALÝZA DAT MATRIKY TUL METODAMI GIS
TECHNICAL UNIVERSITY REGISTER DATA
ANALYSIS USING GIS METHODS**

Diplomová práce: 09-FP-KGE-09

Autor:
Zbyněk Prokop

Podpis:

Adresa:
Sulova 1250
156 00, Praha 5 - Zbraslav

Vedoucí práce: Mgr. Jiří Šmída, Ph.D.

Konzultant: Mgr. Martin Slavík, Ph.D., Mgr. Leoš Křeček,

Počet

stran	grafů	obrázků	tabulek	pramenů	příloh
90	0	24	11	26	10

V Liberci dne: 25. 5. 2009

Original zadání DP

Prohlášení

Byl(a) jsem seznámen(a) s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé diplomové práce pro vnitřní potřebu TUL.

Užiji-li diplomovou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Diplomovou práci jsem vypracoval(a) samostatně s použitím uvedené literatury a na základě konzultací s vedoucím diplomové práce a konzultantem.

Datum

Podpis

Poděkování

Děkuji vedoucímu práce Mgr. Jiřímu Šmídovi, Ph.D. za pomoc při zpracování této práce, za jeho cenné rady, pomoc a podporu. Rovněž děkuji všem, kteří mě během celého studia podporovali.

Analýza dat matriky TUL metodami GSI

Abstrakt:

Předmětem této diplomové práce je analyzovat data uložená v matrice TUL metodami GIS. Na datech FP TUL ze let 2004-2008 je ověřována prostorovost informací v matrice. Dílčím cílem je nalézt postup pro opakovanou aplikaci analýz na další soubory dat a její další využití. Výsledky jsou vizualizovány formou tématických map.

Klíčová slova: matrika, databáze, vizualizace, analýza, ArcGIS, mapa, relace, vektor, rastr

Technical University Register Data Analysis Using GIS Data

Abstract:

The aim of the Diploma thesis is to analyse the data which are saved in the register of the Technical university in Liberec with methods of GIS. Spatiality is verified by the data of the faculty of education of the TUL from 2004 to 2008. The result is to find the procedure for the repeated application of analyses, which are effective on the other files of data and its other benefits. The results are visualised in a form of the thematic map.

Key words: register, database, visualisation, analysis, ArcGIS, map, relation, vector, raster

Análisis de los datos de la matrícula de la Universidad Técnica en Liberec por los métodos GIS

Abstracto:

El objeto de este trabajo es analizar los datos que están en el registro de la Universidad Técnica en Liberec por los métodos GIS. En los datos de la Facultad Pedagógica es verificada la angularidad de las informaciones en el registro. El propósito parcial es encontrar el progreso para la aplicación repetida de los análisis por otros conjuntos de los datos, y su uso. Los resultados son visualizados por la forma de los mapas temáticas.

Palabras de cerradura: registro civil, databaze, vizualizace, análisis, ArcGIS, carta, relace, camello, rastr

Obsah

1	ÚVOD	8
2	CÍLE	9
3	METODY	10
3.1	STUDIUM ODBORNÉ LITERATURY	10
3.2	VSTUP DAT.....	10
3.3	ANALÝZA DAT	11
3.4	VÝSTUPY A VIZUALIZACE DAT.....	11
4	REŠERŠE	12
4.1	DATABÁZOVÉ SYSTÉMY.....	12
4.2	GIS A KARTOGRAFIE	14
4.3	VIZUALIZACE DATABÁZE POMOCÍ GIS.....	15
5	AKADEMICKÝ MARKETING	18
6	GEOGRAFICKÉ INFORMAČNÍ SYSTÉMY	21
6.1	POHLEDY NA GIS	22
6.2	SLOŽKY GIS.....	23
7	DATABÁZOVÉ SYSTÉMY	25
7.1	RELAČNÍ MODEL DAT.....	26
7.2	DOTAZOVACÍ JAZYK SQL	28
8	DATA	30
8.1	PROSTOROVÁ DATA	30
8.1.1	<i>Vektorová data</i>	31
8.1.2	<i>Rastrová data</i>	31
8.2	NEPROSTOROVÁ DATA.....	32
9	METODY ANALÝZY	33
9.1	PROSTOROVÁ ANALÝZA.....	33
9.2	SÍŤOVÁ ANALÝZA	34
9.3	GEOSTATISTICKÁ ANALÝZA.....	34
9.4	3D ANALÝZA.....	34
9.5	INDEX POTENCIÁLU OBCE	35
10	DATA Z MATRIKY TUL	36
10.1	POPIS PRVOTNÍCH TABULEK	37
10.1.1	<i>Tabulka UCHAZECI</i>	37
10.1.2	<i>Tabulka VYSLEDKY_PR</i>	37

10.1.3	Tabulka STUDIJNI_PROGRAMY_FP	38
10.1.4	Tabulka OBORY_STUDIA	38
10.1.5	Tabulka KOMBINACE_ST_OBORU	39
10.1.6	Tabulka ROZHODNUTI	39
10.1.7	Tabulka ZAPIS	39
11	DOPLŇKOVÁ DATA	40
11.1	DATA Z ČSÚ	40
11.2	DOPRAVNÍ SÍŤ	41
11.3	TOPOGRAFICKÝ PODKLAD	42
11.3.1	Geoportál CENIA	42
11.3.2	ArcČR500	42
12	ANALÝZA DAT MATRIKY	44
13	PŘEDPŘÍPRAVA DAT	45
13.1	ÚPRAVA DAT	45
13.1.1	Tabulka UCHAZECI	45
13.1.2	Tabulka STUDIJNI_PROGRAMY_FP	47
13.1.3	Tabulka OBORY_STUDIA	48
13.1.4	Tabulka KOMBINACE_ST_OBORU	48
13.1.5	Tabulka VYSLEDKY_PR	48
13.2	TVORBA DATABÁZE	50
13.3	VÝBĚR DAT	50
14	VSTUP DAT DO GIS	53
15	ANALÝZA	55
15.1	ČASOVÁ A VZDÁLENOSTNÍ DOSTUPNOST	55
15.2	SERVISNÍ OBLAST	59
15.3	KRIGING	62
15.4	3D MODEL ÚZEMÍ	64
15.5	INDEX POTENCIÁLU OBCE	66
16	VÝSTUPY A VYHODNOCENÍ VÝSLEDKŮ ANALÝZY	67
17	DISKUZE	71
18	ZÁVĚR	72
19	UŽITÉ ZKRATKY	73
20	SEZNAM OBRÁZKŮ	74
21	ZDROJE DAT	75

1 Úvod

Spojitost mezi slovy *marketing* a *vysoká škola* byla z valné většiny brána jako nabízený studijní obor. V současné době se ovšem nabízí i další pohled. Aby mohla univerzita prosperovat, otevírat každoročně desítky programů a oborů pro stovky až tisíce studentů a také konkurovat svým rivalům, dalším univerzitám, musí znát své schopnosti, a možnosti. Každý podnik, který má zájem uspět na jakémkoliv trhu, musí umět pojmenovat své přednosti, odhalit slabé stránky a pracovat na jejich odstranění. Zároveň přináší nové metody, postupy a produkty k uspokojení stále větší nebo naopak kvalitnější skupině zákazníků. Vysoká škola je takovému podniku velice podobná, ačkoliv neprodukuje žádný hmotný artikl. Jejimi produkty jsou kvalitní studijní programy a obory, které reflektují moderní pokrok a vývoj společnosti. Odměnou jí je základna zákazníků, v případě školy studentů, kteří využívají jejich produktů. Aby takový subjekt mohl co nejlépe nabízet své produkty, je třeba vhodně analyzovat poptávku a potenciál trhu.

Tématem mé diplomové práce je tedy analýza dat matriky Technické univerzity v Liberci metodami GIS. Jedná se o marketingovou studii, v odborné literatuře označena jako „*academic marketing*“ dle Ayada (2007). Pro potřebu této práce byla z matriky vybrána data pro týkající se Fakulty přírodovědně-humanitní a pedagogické v letech 2004 až 2008.

2 Cíle

Cílem předložené diplomové práce je:

1. Popsat strukturu databáze uchazečů o studium a studentů TUL vedené matrikou TUL
2. Navrhnout možná využití položek databáze pro prostorové analýzy v prostředí GIS
3. Vytvořit soubor grafických prezentací (map, grafů) a jejich interpretací

Za dílčí cíle jsem si stanovil:

1. Ověřit zda jsou data z matriky TUL prostorová
2. Vytvořit metodický postup pro opakované použití
3. Navrhnout další využití analýzy

3 Metody

Mezi vstupem prvotních a výstupem finálních dat je třeba vytvořit řadu podpůrného materiálu. Pro jednotlivé kroky jsou při této práci použity rozdílné programy a metody.

3.1 Studium odborné literatury

Na počátku práce je třeba shromáždit co největší množství informací a zdrojů, které se věnují stejným nebo ekvivalentním tématům jako tato diplomová práce. Výsledkem hledání je rešerší práce (kapitola 4), pro jejíž vypracování využito předplacených databází Web of Knowledge a Web of Science. Zároveň se některé zdroje podařilo vyhledat pomocí nástroje Scholar společnosti Google, která vyhledává na internetu vědecké práce.

3.2 Vstup dat

Vstupní data jsou prvotně zpracována v programu Microsoft Excel 2003. Tento program je využit jako konvertor vstupního formátu CSV (souborový formát pro výměnu tabulkových dat) do formátu XLS. Jedná se o důležitý krok, jelikož zvolený databázový software nedokáže pracovat s původním formátem CSV.

Poté následuje tvorba databáze pomocí Microsoft Access 2003. Program je využíván pro tvorbu relační databáze a následné selekci dat do požadovaného tvaru. Vstupním formátem je formát XLS a data jsou po úpravě konvertována do formátu DBF (database file). Pro třídění a výběr dat je využit standardizovaný dotazovací jazyk SQL.

Data, vyexportovaná z předchozího programu, budou v programu ArcGIS 9.3 (podkapitola 3.3) nástrojem JOIN prostorově lokalizována a připravena na analýzu. Majitelem licencí na všechny programy je FP TUL (ArcGIS 9.3. je zakoupen v licenci ArcEditor).

3.3 Analýza dat

Analytická část práce proběhne v programu společnosti ESRI ArcGIS 9.3. Databáze zde dostává svou prostorovou složku. Vhodnými metodami jsou data analyzována a následně utvářena do výstupů ve formě tématických map. Pro potřeby analýzy jsou využity extenze Spatial Analyst, Network Analyst, 3D Analyst a Geostatistical Analyst. Extenze je nadstavbový modul, který základnímu programu propůjčí další specifické nástroje pro analýzu. Vstupními formáty souborů jsou DBF, XLS a také SHP (shapefile). Formát SHP je určen pro prostorová vektorová data GIS.

Poslední metodou, která není součástí ArcGIS, je gravity index, tedy model přitažlivosti, který určuje míru potenciálu pro nábor nových studentů.

Ke spracování dat byly využity dva počítače:

- AMD Turion 64 x2 1,6 GHz, RAM 1 GB, OS Windows Vista Home Premium
- Intel Pentium4 3 GHz, RAM 1 GB, OS Windows XP Professional

Pro všechny úkony byly tyto prostředky dostačující. Výkon počítače je v relaci s časem, za který daná funkce proběhne. Proto pro některé náročnější úkony vnitřních analýz (hlavně program ArcGIS 9.3) je příhodné využít výkonnější PC.

3.4 Výstupy a vizualizace dat

Výstupní data jsou zpracována v programu ArcGIS 9.3, ve kterém vzniknou všechny tématické mapy. Metodami tématické kartografie, kterými budou výstupní data vizualizována, jsou kartogramy a kartodiagramy.

Zdroji použitých dat bylo:

- Matrika TUL
- Databáze obcí ČR Českého statistického úřadu
- ArcČR 500

4 Rešerše

Rešeršní práce je rozdělena do tří kategorií. Jelikož se tato diplomová práce zabývá analýzou dat uložených v rozsáhlé databázi bude první část věnována databázím. Druhá část se bude zabývat zpracováním dat a jejich vizualizací, k čemuž poslouží program ArcGIS. Ve třetí kategorii jsou citovány články, které se bezprostředně věnují problematice vizualizace dat z databází metodami GIS.

4.1 Databázové systémy

RIORDAN, R.M.: *Vytváříme relační databázové aplikace*. 1. vyd. Praha: Computer Press, 2000. 280 s. ISBN 80-7226-360-9.

V publikaci se autorka nejprve věnuje základním databázovým pojmům. Poté postupuje chronologicky od struktury, vztahů, relací a datové integrity přes relační algebru k návrhu samostatné relační databáze. Ve třetí části knihy se věnuje návrhu uživatelského rozhraní, kde se v poslední kapitole zabývá návrhy o pomoci uživatelů databází. V publikaci se nachází mnoho praktických ukázek v Microsoft Accessu a SQL Serveru.

BERKA, P.: *Dobývání znalostí z databází*. 1. vyd. Praha: Academia, 2003. ISBN 80-200-1062-9.

Jedná se o přehledovou publikaci, která postihuje hlavní rysy získávání informací z databází a strojového učení. Kniha je rozdělena na čtyři části. První dvě jsou spíše úvodní, aby byl srozumitelný výklad systému pro dobývání informací, kterému se velice podrobně věnuje část třetí. V obecné rovině je zde popisován postup, který by se měl využít nezávisle na úloze či aplikaci.

SHEPHERD, J.C.: *Database Management: Theory and Application*. 1. vyd. Homewood, Illinois: Richard D. Irwin, Inc., 1990. 781 s. ISBN 0-256-07829-7.

Stejně jako každá publikace o informatickém tématu se autor v první kapitole věnuje úvodu do databází. Poté se zabývá hlavně daty, které srovnává s informacemi, řeší atributy a datové modely. Ve třetí kapitole je vysvětleno, jak implementovat relace do datové struktury. V druhé polovině se autor zabývá hlavně

designem databáze a také databázové koncepci a normalizaci při rozlišení logické a fyzické složky. Poslední částí knihy je věnována velice pozorně relačnímu modelu dat, což je završeno úvodem do užití databází v jazyce SQL (Structured Query Language). Následně je práce rozebírána v hned v několika jazycích včetně Cobolu, 4GLs či Codasyl.

GILFILLAN, I. *Introduction to Relational Databases* [online]. 24.6.2002, 24.6.2002 [cit. 11.12.2008].

<<http://www.databasejournal.com/sqletc/article.php/1469521/Introduction-to-Relational-Databases.htm>>.

Tento článek se zabývá základy práce s relačními databázemi. Dokonce je zde možnost naučit se práci v různých systémech. Nechybí definice základní terminologie. Stránky a materiály tohoto typu budou využity hlavně při analýze matriky a pochopení vztahů uvnitř databáze.

***Teorie relačních databází: Relační model dat* [online]. 12.1.2006 [cit. 11.12.2008]. <<http://www.manualy.net/article.php?articleID=9>>.**

Jedná se o stručný manuál, ve kterém je popsán relační model dat, základní databázové pojmy jakým je doména, atribut, relace a také pravidla, který musí splňovat relační databáze. Stejně jako předchozí, poskytne tento článek sumarizaci platných pravidel pro tvorbu zpracovávané databáze, díky nimž budu analyzovat strukturu matriky.

***Relační vs. objektově-relační vs. objektové databáze* [online]. 1997 [cit. 11.12.2008]. <<http://www.fi.muni.cz/~xbatko/oracle/compare.html>>.**

Tato webová stránka srovnává různé typy přístupů k tvorbě a k práci s daty. V každém datovém systému zároveň pojmenovává základní pojmy každého systému, kdy lze vyvodit klady a zápory použití jednotlivých databázových systémů. Při analýze matriky bude možnost v inovaci ukládání dat a zvoleného systému dat.

4.2 GIS a kartografie

TUČEK, J.: *Geografické informační systémy – Principy a praxe*. 1. vydání. Praha: Computer Press, 1998. 424 s. ISBN 80-7226-091-X.

Výše zmíněná publikace se věnuje problematice práce se systémy GIS od jejího počátku. Autor vše rozdělil do čtyř částí. V první oddílu se věnuje úvodu do problematiky, ve které vysvětluje, co vlastně jsou prostorové informační systémy. V dalších třech kapitolách rozebírá, co je to GIS a jaké má součásti. Jako jedna z podkapitol je i historie GIS. Prostor je věnován i rozboru hardwarových a softwarových prostředků GIS a práci s prostorovými údaji. V druhé části se přesuneme k základům geoinformatiky, kdy autor vysvětlí práci při modelování geografických objektů. V další kapitole je poučeno o reprezentaci prostorových objektů a druhou závěrečnou část knihy završí rozsáhlá kapitola o databázových systémech, které jsou pro GIS nepostradatelné. V poslední řadě se věnuje tvorbě databáze, včetně jejího plnění daty, restrukturalizací dat, poté jejich analýzou a syntézou a vše je završenou kapitolou o vizualizaci dat a vytváření výstupů. Závěrečná část knihy se zabývá hodnocením kvality geografických údajů spolu s implementací a využíváním systémů včetně rozboru, dle autora, nejrozšířenějších softwarových systémů.

MACHALOVÁ, J.: *Prostorově orientované systémy pro podporu manažerského rozhodování*. 1. vyd. Praha: C. H. Beck, 2007. 141 s. ISBN 978-80-7179-463-9.

Autorka se v publikaci věnuje hlavně aplikaci GIS v různých sférách, kdy je každé sféře aplikace věnována jedna kapitola. V úvodu se zabývá geografickými informacemi, jejich validitou a prolnutím GIS do IS (Informační systém) různých organizací. Závěrem se publice zaobírá problematikou modelování terénu, či virtuální realitou, nebo také ekonomickými aspekty zavádění GIS.

ŠÍMA, J.: *Geoinformační terminologie pro geodety a kartografy*. 1. vyd. Zdíby: VÚGTK, 2003. 87 s. ISBN 80-85881-20-9.

Tato publikace je spíše informativní, jelikož se jedná o souhrnný přehled základních pojmů a termínů z oblasti geoinformatiky, který obsahuje nad 200 termínů s jejich anglickými, německými a ruskými ekvivalenty. Tento materiál bude

důležitý při práci s cizojazyčnou literaturou, konkrétně při překladu odborných termínů.

LANGR, J. T-MAPY spol. s r.o. Hradec Králové – Budování geodatabáze [online]. 1999 [cit. 11.12.2008].

<http://www.tmapy.cz/public/tmapy/cz/_aktualne/_clanky/budovani_geodatabaze.html>.

Tento příspěvek se zabývá budováním geodatabáze. Je zde navržena vhodná struktura databáze s ohledem na obsahové a formální požadavky. Dále zde jsou rozebrány produkty společnosti ESRI (Environmental Systems Research Institute, Inc.), datové modely a nástroje pro digitální kartografii.

CRONER, C. M.; SPERLING, J.; BROOME, F. R.: Geographic information systems (GIS): New perspective in understanding human health and environmental relationships. In: *Statistics in medicine*. Ročník 15. US, 1996, s. 15 – 30.

Příspěvek posunuje vnímání GIS jako nástroje pro práci s prostorovými daty a usnadnění hledání vztahů mezi nimi. Autoři navazují na odkaz Johna Snowa, který objevil klasickou asociaci mezi cholerou a vodní pumpou ve stanici Broad Street. Východiskem k práci budou hlavně poznatky o využití GIS při hledání vztahů mezi jevy, jejich klasifikaci a generalizaci.

4.3 Vizualizace databáze pomocí GIS

ZAJÍCOVÁ, Z.: Návrh zpracování dat SCIO v prostředí GIS. In *GISáček 2006*. Ročník 2006. VSB Ostrava, 2006.

GLACOVÁ, M; HORÁK, J.: Analýza situace trhu práce v měřítku okresního města. In Sbor. ref. conference GIS Seč 2002, Seč, 12.-14.6.2002, 2002, s. 12.

V obou článcích se nachází návody a možnosti vizualizace dat v GIS, což je pro jednu z fází mé práce stěžejní. Zvláště v druhém článku najdeme rady o výběru vhodných dat a jejich zpracování. Samotné vizualizaci je věnován opravdu velký prostor, kde se zde objevují i grafické příklady. V úplném závěru jsou příklady

rozebrány podle socioekonomického hlediska, což by mi mohlo pomoci při samotném hodnocení mých vlastních výstupů.

TANG, H.; MCDONALD, S. *Intergrating GIS and spatial data mining technique for target marketing of university courses*. Symposium on Geospatial Theory, Processing and Applications. Ottawa, 2002.

Článek se zabývá marketingovou využitelností GIS při analýze struktury studentů na Charles Sturt University v Albury a Austrálii. Zkoumají, kde studenti bydlí a odkud dojíždí. V článku je odkaz na mnoho prací různých autorů, které jsou úzce spojené s tématem. Studie krok po kroku rozebírá, ale nikoliv detailně, jak využít jednotlivé prvky analýzy. Zpočátku hledají spojení GIS se získáváním prostorových dat. Rozsáhlá část práce je věnována zkoumání prostorových asociací, kde jsou popsány tři různé metody přístupu k datům. V závěru charakterizují metody předpovídání potenciálních „*market areas*“.

AYAD, Y. M. *Challenges in student recruitment for educational institutions: Materials and methods: Case study of Clarion University of PA: Analysis of 2004-2005 school year data*. ESRI User conference 2007.

Autor se v této studii zabývá metodami akademického marketingu. Postup analýzy je rozebrán krok po kroku, dále jsou zapracovány a srovnány práce a myšlenky jiných autorů. Jako hlavní činitel, který strukturuje zkoumané území je tzv. model přitažlivosti, který zohledňuje počty studentů, vzdálenost dojížděky či průměrný příjem rodiny. Tato studie se snaží nalézt obecný postup doplněný praktickými příklady, které jsou prezentovány formou tématických map.

CHRISTIE, R.; FERRIS, M. *Spatial analysis for enrollment planing in higher education*. ESRI Education User Conference Proceedings 2004, San Diego, CA.

Studie se zabývá spojením GIS a prostorových dat při analýze statistických údajů. Stejně jako předešlé studie se snaží popsat postup pro vytvoření marketingového průzkumu možných zdrojů nových studentů. Autoři popisují, jakou roli hraje dojíždění pro studenta vyššího vzdělávání, klasifikují vzdálenosti a hledají asociace. Je zde uvedeno, z jakých dat čerpají informace, které transformují do

podoby prostorových dat. V příloze je přiloženo 6 tématických map, které alespoň částečně prezentují výsledky zjištěné pro California State University.

MARBLE, D. F.; MORA, V. J.; HERRIES, J. P. *Applying GIS technology to the freshman admissions process at a large university*. ESRI User Conference. 22 – 25. 5. 1995. Palm Springs, CA, USA. Dostupné z:

<<http://www.isprs.org/commission4/proceedings02/pdfpapers/553.pdf>>.

Úvod práce je věnován historickému kontextu přijímacího řízení na Ohio State University. Autoři studie se zpočátku věnují stanovením úkolů pro univerzitu. V druhém kroku je důležitá změna z pasivní na aktivní pozice, což znamená aktivně oslovovat dané lokality a adepty o studium. Je zde zohledněno pohlaví, věk, dojíždka a rasa. Zároveň se zohledňuje změna ve výchově. Nabírání nových studentů je velký problém a technologie (jako GIS) jsou uváděny jako léky náboru. Druhá polovina studie se věnuje zapojení technologie GIS do přijímacího řízení, strategie a taktiky užití GIS a završuje ji pár grafických příkladů na relevantních datech.

5 Akademický marketing

Pro začátek je nutné definovat pojem matrika. *Matrikou se rozumí úřední kniha určená k evidenci osob a skutečností spojených s osobním stavem občanů (osobní údaje)*¹. K datům uloženým v matrice lze stanovit premisu: za předpokladu propojení s mapou jsou prostorově lokalizovatelná a s využitím vhodných softwarových nástrojů lze na ně aplikovat analýzu.

Školní matrika je databáze, v níž se uchovávají informace o všech uchazečích o studium, ať jsou zapsáni do studia nebo ne. Každá škola vede svou matriku, které jsou napojeny na centrální matriku vedenou MŠMT. Pro práci s matrikou existuje prostředník, který vizualizuje editační funkce databáze. V případě TUL jím je IS STAG. Jedná se o informační systém studijní agentury, vytvořený Západočeskou univerzitou v Plzni.

Propojení matriky a IS STAG je jasné. Matrika je databáze, tedy soustava tabulek, propojených relacemi. IS STAG je klient, který zobrazuje informace obsažené v matrice. Poskytuje nástroje pro vkládání, vizualizaci, editaci a mazání záznamů.

V zásadě je možné uchazeče o studium na škole pojmenovat jako zákazníka. Škola pak může být označena za právnickou osobu nabízející službu. Tím se dostává škola a uchazeč do obchodního marketingového vztahu. Podnik (škola) tedy nabízí produkt a chce, aby o něj měl zájem zákazník (uchazeč). Obchodní vztahy jsou analyzovány s cílem dosáhnout ideálního stavu (vyššího zisku, nižší náklady, atp.). Takové výzkumy jsou nazývají marketingové analýzy.

Akademický marketing je definován jako překročení hranice mezi univerzitou a marketingem, mezi činnostmi, napříč disciplínami, kulturou, a ve vztahu ke společnosti. Poskytuje účelná měření, hypotézy a reflexi pro vedení školy a studenty (Tang, 2002). Marble (1995) jako první použil geografické a demografické údaje k analýze **market areas**, tedy míst s vysokým potenciálem pro nábor studentů. Tématu akademického marketingu se věnují v USA Ayad (2007), Christie (2004),

¹ MALÁTKOVÁ, J.: *Ottova encyklopedie A-Ž*. 1. vyd. Praha: Ottovo nakladatelství, s.r.o. 1144 s. ISBN 80-7360-014-5

Croner (1996), v Austrálii (Tang, 2002), případně i další v Číně, Singapuru, atd. V Londýně společnost Marketability, Ltd. Nabízí kurs akademické marketingu.

Metodami akademické marketingu jsou hlavně statistické výpočty a marketingové metody (audit, analýza trhu) a SWOT analýza. S rozvojem GIS přichází nový rozměr vizualizace výsledků analýz. Jako vizualizační nástroj je využit právě GIS, který přidává možnosti prostorových síťových analýz a tvorbu výstupů formou tématických map.

Role GIS v akademickém marketingu se stává více důležitou. GIS usnadňuje přesnou lokalizaci výsledků analýz a dat. Nejprve jsou hledány asociace v prostoru (Croner, 1996) a poté identifikovány **market areas** (Tang, 2002). Všechny zahraniční studie, které na téma akademického marketingu vznikly od roku 2002, využívají metod GIS.

Je běžnou praxí tvořit marketingové studie územních jednotek, podniků či služeb. Marketingové studie škol jsou tedy prozatím zahraniční záležitostí, což je souzeno dle uveřejněných výsledků. Pro územní České republiky taková analýza vytvořena nebyla, či nebyla uveřejněna. Akademický marketing je omezen na statistické údaje či marketingové výpočty, v žádném zdroji není pracováno s GIS jako funkčním nástrojem pro analýzu. Na Vysokém učení technickém v Brně, České zemědělské univerzitě v Praze či Západočeské univerzitě v Plzni již probíhají přednášky o principech tvorby strategií rozvoje univerzit, což je možné označit za počáteční fázi aplikace akademického marketingu.

Počátkem práce na studii akademického marketingu je export dat z matricy, kdy vznikne několik tabulek, které nemají žádný fyzický vztah. Pro úspěch jakékoliv práce s těmito daty je třeba vytvořit databázi, tedy propojenou množinu všech tabulek. Spojujícím prvkem jsou relace. Vznik relací doprovází řada standardů a integritních omezení, které popisují Shepherd (1990) a Písek (2007). Shepherd (2007) se věnuje zároveň problematice výběru dat z obsáhlých databází. Pro selekci dat využíván jazyk SQL, který informace, obsažené v matrice, vybírá dle vypsání kritérií a zároveň provádí sumarizaci, je-li požadována.

Croner (1996) tvrdí, že je GIS vhodný analytický nástroj pro zpracování prostorových dat a hledání asociací a vztahů. Proto je v dalším kroku je využit softwarový produkt ArcGIS společnosti ESRI. Ještě než jsou data dále zpracovávána, je třeba vybudovat geodatabázi. Langr (1999) stanoví formální i

obsahové požadavky na strukturu databáze, která bude propojena s prostorovou datovou vrstvou, čímž se data z matricy lokalizují a získají svou prostorovou složku.

Ve třetí fázi práce je nutné určit metody prostorové analýzy. Tuček (1998) odkrývá metody, které jsou součástí GIS. Také dle Zajícové (2006) a Glocové (2002) jsou vhodnými metodami síťové analýzy, které jsou běžně využívanými marketingovými nástroji. Stejně metody využívají Christie (2004) a Ayad (2007), kteří zkoumají zdroje studentů pro svou univerzitu v závislosti na dostupnosti do sídla univerzity. V programu ArcGIS tyto funkce zajišťuje extenze Network Analyst (vektorová data) a Spatial Analyst (rastrová data).

Další metodou je podle Tanga (2002) je předpověď **market areas**, kde je vysoký potenciál v počtu možných uchazečů o studium. Zde jsou zjištěná data interpolována. Tato interpolace bude zajištěna metodou kriging, jenž obsahuje extenze programu ArcGIS Geostatistical Analyst.

Tuček (1998) navrhuje jako jednu z možností výstupu dat vytvořit 3D model území. Nemusí se tedy nutně jednat o výškové body, 3D model může být brán jako další vhodný způsob zobrazení určitého jevu (počet uchazečů z obce, atp.).

Ayad (2007), který vytvářel marketingovou studii Clarion University v USA, přichází s další metodou pro analýzu dat. Jedná se o model přitažlivosti (**gravity index**). Autor zde dává do přímého vztahu dvě proměnné (počet studentů z obce a vzdálenost obce, ze které jsou, od univerzity v km), které upravuje koeficient. Průměrný měsíční příjem rodiny je zde klasifikován do čtyř skupin, kdy má každá skupina určenou hodnotu, jež je dosazována jako koeficient do rovnice. Jak sám uvádí, je možné tyto hodnoty zaměnit s jinými, ovšem při dodržení struktury vzorce. Lze tedy vložit do přímé závislosti počet uchazečů o studium z obce a počet obyvatel (věkové skupiny 15-19 let) obce, upravené koeficientem vzdálenosti do sídla univerzity. Vzdálenost je tedy rovněž klasifikována do skupin stejně jako autorův případ.

6 Geografické informační systémy

Pojem Geografické informační systémy (dále jen GIS) je standardním označením počítačových systémů, které se orientují na zpracování geografických dat prezentovaných především mapovou formou. Zároveň označuje i novou a velice progresivní, vědeckou disciplínu (Tuček (1998)). Jednoznačná a všeobecně přijatelná definice pojmu GIS však neexistuje.

Příklad: Vybrané definice GIS dle Tučka (1998)

„Databázový systém, ve kterém je většina údajů prostorově definována a na jejíž zpracování lze použít procedury dotazů na prostorové entity v databázi.“

SMITH, 1987

„Systém pro podporu rozhodování, který umožňuje integraci prostorově definovaných údajů v prostředí řešení problémů.“

COWEN, 1988

Dle Rapanta (2002) by mohla nejvhodnější definice GIS znít následovně:

„GIS je funkční celek vytvořený integrací technických a programových prostředků, dat, pracovních postupů, obsluhy, uživatelů a organizačního kontextu, zaměřený na sběr, ukládání, správu, analýzu, syntézu a presentaci prostorových dat pro potřeby popisu, analýzy, modelování a simulace okolního světa s cílem získat nové informace potřebné pro racionální správu a využívání tohoto světa.“

Existují rovněž i různé pohledy na chápání tohoto pojmu. Jeden pohled nám GIS vykládá jako technologii, aplikační nástroj a vědecký obor (Tuček (1998)), druhý naopak jako software, konkrétní aplikaci a informační technologii (Rapant (2002)). Je patrné, že i kdyby zde bylo uvedeno více možností, všechny by se lišily v drobnostech. Každá varianta má společné to, že GIS je prvním místem střetu řady moderních společenskovědeckých trendů – matematiky, geografie, kartografie, geodézie a informatiky.

6.1 Pohledy na GIS

Všechna výše uvedená fakta se shodují ve třech základních úhlech pohledu na GIS. Označují se jako kartografický, databázový a analytický. Každý specifický pohled může být zpracován soliterním softwarem, avšak právě GIS všechny tyto prvky integruje do jediné složky.

Kartografický pohled

Kartografický pohled dominuje u většiny uživatelů, kteří hledají v GIS kartografický aspekt či se snaží kvalitně prezentovat výsledky zpracované integrovanými analýzami. Všeobecně je tento pohled znám jako práce s digitálními mapami.

Databázový pohled

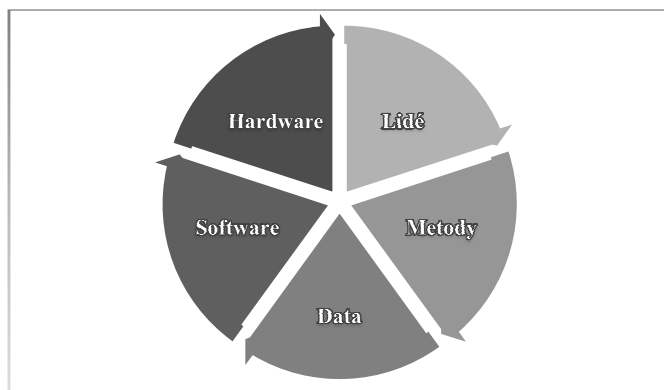
Při práci s větším množstvím dat je třeba využít správně navrhnuté a zorganizované databáze. Hlavními potřebami takového systému je shromažďování, identifikace, třídění, selekce a prezentace dat. GIS přidává nadstavbovou složku, kterou je vazba údajů na zemský povrch, dle jedinečného klíče. Toto využití nalezneme při účelových analýzách a jako integrovanou vlastnost informačních systémů měst, řízení průmyslových objektů, vyšších územních jednotek, apod.

Analytický pohled

Analytický pohled přináší s sebou prostorové analýzy, syntézy poznaků a modelování území. V tomto pohledu jsou GIS nejvíce geografické, jelikož dovolují hledat prostorové asociace mezi probíhajícími jevy a krajinou.

6.2 Složky GIS

Každý systém GIS se musí skládat z daných komponent, aby byl správně a účelně použitelný. Jednotlivé prvky jsou navzájem propojeny a ovlivňují se (Obr. 1).



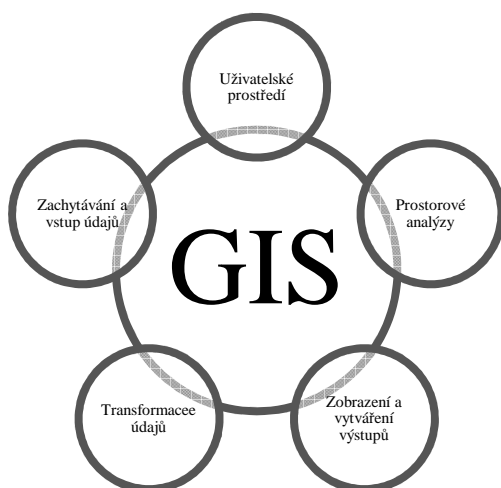
Obr. 1: Složky GIS

Hardware

V současnosti je GIS natolik variabilní, že je k dispozici na nejrůznějších počítačových platformách. Může být použit jak na osobních počítačích, tak i na víceuživatelských systémech. V mnoha případech pokročilejších analýz vyžaduje připojení specifických vstupních a výstupních zařízení, jako jsou geodetické přístroje, digitizéry a plotry. Je ovšem možné jej připojit na domácím počítači k běžné tiskárně a vytvářet tak mapové výstupy, méně náročné na tisk.

Software

Vlastní program pro práci s geodaty je postaven modulárně. Základním pilířem systému je jádro programu, které obsahuje standardní funkce. K jádru je poté možné připojit velké množství modulů a programových nadstaveb (Obr. 2) pro konkrétní specifické úlohy (práce ve 3D, prostorové, síťové a statistické analýzy, kartografická vizualizace, atp.). Mezi konkrétní softwarové produkty patří GRASS GIS, Janitor, Quantum GIS (všechny tři jsou distribuovány v licenci freeware) a ArcGIS od společnosti ESRI (Environmental Systems Research Institute, Inc. – distribuován pouze v placené verzi).



Obr. 2: Hlavní skupiny softwarových modulů GIS (zpracováno dle Tučka, 1998)

Data

Data jsou nejdůležitější část GIS, která tvoří až 80% finančních nákladů na provoz (dle Tučka (1998) 70%, dle Rapanta (2002) 90%). Jedná se o prostředky na získání a obnovu dat. Datům je věnována vlastní kapitola, viz. kapitola 5.

Metody

Jedná se o postupy a zapojení systému do funkčních schémat a modelů. Metodika práce s GIS je z hlediska praxe velice komplikovaná a náročná, proto distributoři (ESRI, Arcdata) nabízí specializovaná, samozřejmě placená, školení uživatelů – geoinformatiků.

Lidé

Lidé, kteří s těmito systémy pracují se nazývají geoinformatiči. Dají se rozdělit do kategorií, dle jejich úrovně práce s GIS. Ať se jedná o programátory, analytiku či koncové uživatele, všichni musí ovládat potřebné metody ke zvládnutí zbylých elementů GIS.

7 Databázové systémy

Databáze je systematicky uspořádaná množina dat (informací). Jedná se o komplikovanou centrální strukturu dat, která sjednocuje izolované soubory. Správa databáze je realizována prostřednictvím specializovaného programového vybavení, které se nazývá systém řízení báze dat (dále jen SŘDB). Při propojení databáze s SŘDB vzniká specifický produkt – databázový systém (Písek, 2007).

Databáze by měla být navržena tak, aby minimalizovala redundantní data, tj. data, která se objevují v databázi vícekrát (Tuček, 1998). Důležitou, ne-li kritickou, vlastností báze dat je vnitřní integrita.

Integrita databáze je stav, v němž jsou všechna data v plném rozsahu správná a využitelná v aplikačních programech. Pokud je tato vlastnost narušena, jedná se zpravidla o sémantickou chybu v záznamech (jednotlivou položku databáze). Další příčiny mohou být chyby technického a programového vybavení, případně chyby v aplikačním prostředí či datech.

Součástí SŘDB proto musí být kontrola vstupujících dat. Stejně tak jsou implementovány nástroje pro ukládání výstupních informací, aby je bylo možné znovu využít bez chyb. Jedná-li se o velké databáze, je nutné zajistit zálohování také specifickými hardwarovými prostředky, tj. externími disky na plnou zálohu či žurnálovou záložní paměť, která uchovává vždy údaje před změnou a po změně.

Tab. 1: Tabulka databáze s narušenou integritou (Písek, 2007)

Objednávka	Jméno	Adresa	Výrobek
1	Petr Sedlák	Luční 8, Praha	Tenisová raketa
2	Jana Suchá	Patočkova 6, Liberec	Floorbalová hůl
3	Jiří Dlouhý	Pod Bezem 49, Újezd	Squashová raketa
4	Petr Sedlák	Luční 8, Praha	Tenisový míček

V tabulce 1 je zřetelný problém se záznamy, které popisují objednávku pod čísly 1 a 4. Jedná se o typický problém, protože v tuto chvíli vidí SŘDB oba záznamy jako dvě odlišné objekty. Pokud by se ovšem jednalo o jednu a tutéž osobu (což by mohla, ale nemusí potvrdit adresa, která je u obou stejná), tak by systém nevypsal objednávku 4 při výpisu objednávek Petra Sedláka a naopak. U malých databází je možnost ruční opravy chyb, ale i zde je to problém. Takto neřešitelná komplikace nastává u databází řádově od desítek po tisíce záznamů. Univerzálním řešením by

v tomto případě mohlo být přidání jedinečného číselného identifikátoru ke každému jménu, čímž by se značně snížila hrozba podobné sémantické chyby.

7.1 Relaçní model dat

Pro potřeby této diplomové práce se budeme zabývat pouze databázemi založenými na relačním modelu dat.

Relaçní model dat je specifický druh databázové technologie. Je založený na matematické teorii množin a predikátové logice. Určuje si vlastní intergritní omezení (obsahuje soubor šesti normálních norem, které zjednodušují a optimalizují návrh databázové tabulky tak, aby obsahovala minimální počet redundantních dat) a způsob práce s informacemi (Vebloud, 2006). Základem relační databáze je tabulka s údaji, která se skládá ze sloupců a řádků. Je pravidlem, že relační databáze má 2 a více tabulek.

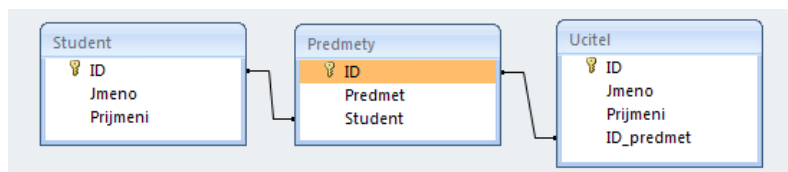
Sloupcem rozumíme svislou množinu prvků. Každý sloupec musí mít jedinečný (nezaměnitelný) název a určený datový typ, který bude obsahovat.

Řádek (záznam) je vodorovná množina prvků (atributů). Každý řádek by měl mít jedinečný identifikátor, který jednoznačně určí právě jeden záznam. Tuto vlastnost obhospodařují tzv. *klíče*.

Klíčem se rozumí unikátní atribut, případně i jejich kombinace. První úlohou klíče je reprezentovat asociace mezi záznamy. Druhým úkolem klíče je výše uvedená jednoznačná identifikace. Takový klíč se nazývá primární klíč (PK).

Dle Tučka (1998) má atribut následující dvě vlastnosti:

- 1) Je jednoznačný – neexistuje v tabulce jiný záznam, který by měl stejný PK.
- 2) Je minimální – při odebírání atributů není možné porušit pravidlo č.1.



Obr. 3: Ukázka relační databáze

V tabulce *Predmety* je soupis předmětů, na které student dochází (označenou ID studenta). V tabulce *Student* je uvedeno ID každého studenta, dále jeho jméno a příjmení. Tabulka *Ucitel* je téměř ekvivalentní s tabulkou *Student* s rozdílem uvedeného vyučovaného předmětu (ID předmětu). Černá linie naznačuje relaci (propojení) mezi jednotlivými tabulkami.

Dle obrázku 3 je patrné, že v každé tabulce je primárním klíčem záznam ID a je označen žlutým klíčem. Jedná se o zobrazení v programu Microsoft Access 2003.

V tabulkách může zároveň existovat tzv. *cizí klíč* (též nevlastní klíč). Tento specifický klíč pomáhá určovat záznamy, které spolu navzájem nesouvisí. Dovoluje propojit tabulky tak, že pokud by byl určitý záznam v cizí tabulce smazán, v rámci zachování integrity, se záznam zdrojové tabulky, obsahující jako atribut smazaný záznam, smaže také.

Typy relací

Dle Píska (2007) rozlišujeme následující 4 druhy:

- 1) **1:1** – Jeden záznam v tabulce A odpovídá jednomu záznamu tabulky B. Tento typ relace se nevyskytuje často, jelikož je možné ekvivalentní řešení v podobě jedné tabulky. Příklad užití: *Jeden zaměstnanec má jednu identifikační kartu. Při propuštění zaměstnance se smaže jeho záznam v databázi, avšak záznam o identifikační kartě zůstává a může být přiřazen jinému zaměstnanci.*
- 2) **1:N** – Jeden záznam v tabulce A odpovídá více záznamům tabulky B. Je to nejčastěji se vyskytující relační spojení v databázích. Příklad užití: *Jeden studovaný obor může studovat více studentů, naopak jeden student může studovat jen jeden obor.*
- 3) **M:N** – Několik záznamů v tabulce A odpovídá více záznamům tabulky B. Toto propojení lze vytvořit kombinací dvěma relacemi: 1:N a 1:M – vzniká třetí tabulka, tzv. vazební. Příklad užití: *Jeden výrobek může mít více vlastností, stejně tak může mít jednu vlastnost více výrobků.*
- 4) Žádný vztah mezi tabulkami

7.2 Dotazovací jazyk SQL

SQL (Structured Query Language) je standardní dotazovací jazyk v relačních a objektově relačních databázích (Písek, 2007).

Tento nástroj vznikl v roce 1974 a byl několikrát standardizován. V roce 1986 byl přijat standard s označením SQL89, v roce 1992 přišla na trh opravená verze SQL92 a v roce 1999 vychází verze SQL3 (SQL99). Poslední standardizace proběhla v roce 2003, kdy vzniká tzv. SQL 2003. Důležitým mezníkem je právě SQL3, kdy byla přidána sada ADT, obsahující procedury a funkce pro práci s prostorovými daty.

SQL se skládá z několika částí:

- DDL (Data Definition Language) – Dovoluje uživatelům vytvářet databázové objekty a upravovat jejich strukturu.
- DQL (Data Query Language) – Řídí načítání dat z databáze. Dovoluje kombinovat atributy propojených tabulek a vytvářet tak dotazy.
- DML (Data Manipulation Language) – Umožňuje uživatelům přidávat, odebírat a měnit data v databáze.
- DCL (Data Control Language) – Povoluje správcům možnost omezit a řídit přístup k datům v databázi a využívat systémová oprávnění SRDB.

SQL obsahuje řadu příkazů pro práci s daty. Jedním z nich je příkaz SELECT, který je pro tuto analýzu stěžejní. Je základním prvkem pro dotazování v relačních databázích. Jeho použití dovoluje zobrazovat požadovaná data. Výsledkem dotazu je strukturovaná množina dat, nejčastěji reprezentována formou tabulky.

Příkaz SELECT obsahuje kromě klíčového slova SELECT i další klíčová slova, obecně nazývaná jako klauzule. Některé klauzule jsou pro zápis přímo povinné, jiné variabilní. Klauzule v příkazu SELECT jsou rozděleny do pěti skupin:

- SELECT

Jedná se o povinnou klauzuli, která umožňuje specifikovat sloupce výsledné množiny dat. Sloupce mohou být vybírány z kterékoliv propojené tabulky, definované příkazem FROM. Je možné vkládat sloupce vzniklé množinovými operacemi nad více tabulkami (disjunkce, konjunkce, atd.).

- FROM

Druhá povinná klauzule, která specifikuje tabulku, z níž se vybírají sloupce.

- WHERE

První volitelná klauzule, která umožňuje filtrovat řádky zdroje. Technika filtrování se zadává pomocí predikátu, nabývající hodnoty *true*, *false* nebo *unknown*.

- GROUP BY

Další volitelná klauzule pro tvoření samostatných podmnožin v rámci sloupce, kdy sjednocuje všechny parametry odpovídající zadanému klíči. Pokud je v příkazu SELECT využita agregační funkce, stává se tato klauzule povinnou.

- HAVING

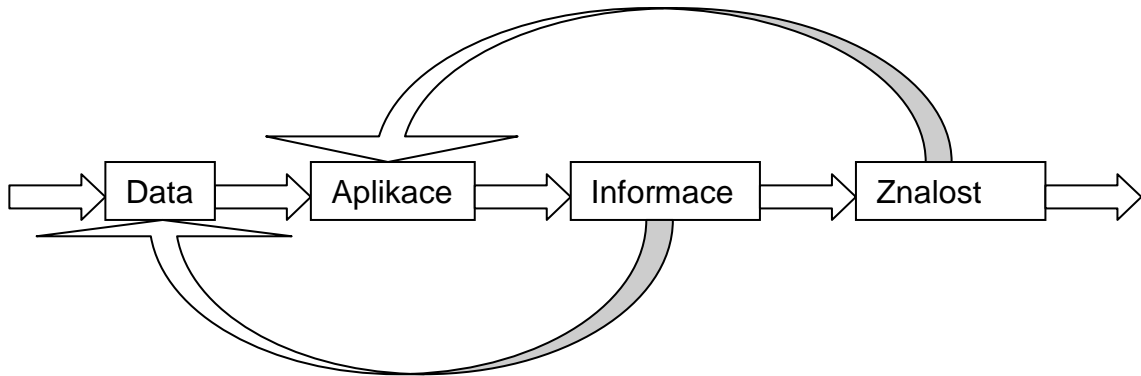
Poslední volitelná klauzule, která přímo závisí na klauzuli GROUP BY. Svou funkcí je ekvivalentní ke klauzuli WHERE, avšak slouží k filtrování dat rozdělených do podmnožin klauzulí GROUP BY.

Pro snadnější představu následuje stavba příkazu SELECT se všemi klauzulemi.

SELECT položka **FROM** tabulka **WHERE** atribut = xxx **GROUP BY** atribut **HAVING** atribut1 = yyy

8 Data

Data je výraz pro údaje, které popisují nějaký jev nebo vlastnost pozorovaného jevu. Aby bylo možné správně rozhodnout jaká data máme k dispozici, je potřeba vycházet z následující obecná posloupnosti (obr. 4):



Obr. 4: Datová posloupnost (vytvořeno dle Tučka ,1998)

Vstupní množinou této posloupnosti jsou data, která jsou zpracována pomocí aplikace, a na základě toho vznikají informace. Tyto informace mohou být přetransformovány jako data, která aplikace znovu zpracuje jako data zdrojová. Pokud informace analyzujeme, vzniká znalost. Výslednou znalost je možné využít jako výstupní hodnotu posloupnosti či naopak z ní modifikovat aplikaci, kterou transformujeme data.

Všechna data mají i svou popisnou část, tzv. *metadata*. *Metadata* informují o obsahu, v případě prostorových dat (geodat) o referenčním prostorovém systému (geografickém souřadnicovém systému), kvalitě, atp.

Je mnoho způsobů dělení dat na různé skupiny. Jednou z nich může být rozdělení na data prostorová (*spatial data*) a data neprostorová (*non-spatial data*).

8.1 Prostorová data

Prostorová data jsou jakákoliv data, která obsahují geografickou lokalizaci. Každý záznam v tabulce se zpravidla odkazuje na právě jedno místo, často určeno formou souřadnic a topologie (bod, linie, polygon). Nástroj, který zajišťuje vazbu dat na konkrétní místo v prostoru se nazývá georeference. Georeferencovaný prvek je

v ideálním případě lokalizován přímo souřadnicemi v mapě. Většinou je to zajištěno nepřímo, tj. adresou, číslem parcely, identifikačním číslem územní jednotky, apod. Synonymem pro prostorová data je hojně užívaný termín geodata.

Prostorová data dělíme dle způsobu reprezentace objektů a jevů reálného světa na vektorová a rastrová data.

8.1.1 Vektorová data

Vektorová data vyjadřují geometrické vlastnosti jevů na zemském povrchu pomocí lineárních charakteristik. Tato data členíme na tři základní prvky (Tuček, 1998):

- Bod – přesně určený prvek souřadnicemi x , y (z)
- Linie – soubor bodů (minimálně dva body – počáteční a koncový), které jsou určeny souřadnicemi x, y (z), kdy spojnice mezi dvěma body je buď nejkratší vzdáleností mezi nimi, nebo přesně definovanou křivkou
- Polygon – uzavřený soubor bodů, které jsou určeny souřadnicemi x, y (z)

Každý prvek nese informaci o své poloze, tvaru a dalších charakteristikách (atributy). Všechny tyto atributy jsou uloženy v tabulce, která je jedinečným identifikátorem provázána s daným objektem.

Výhodou vektorových dat je možnost pracovat s jednotlivými objekty jako s celky. Další výhodou je menší náročnost na paměť. Vzhledem ke GIS mají vektorová data tu vlastnost, že se při změně měřítka nezmění čitelnost znaku.

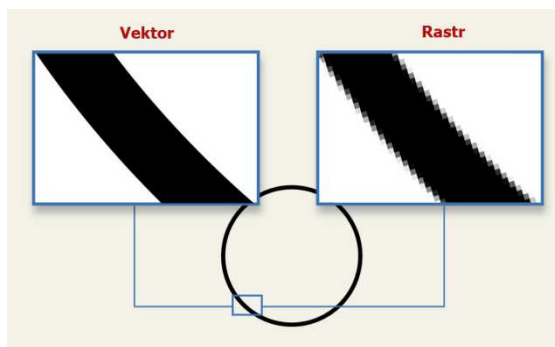
8.1.2 Rastrová data

Rastrová data se vztahují k celku. Nelze tedy snadno zjistit, z jakých obrazců se výsledek skládá. Soubor obsahuje informace o velikosti obrazu, způsobu komprese a kódování barvy. Samotný obraz je uložen jako matice, kde je každý prvek roven jednomu bodu. Do rastrových formátů se ukládají data, jež se nebudou dále editovat v systému, který je vytvořil.

Jak již bylo poznamenáno, rastrová data se skládají z buněk (pixelů). Každý pixel má svou polohu určenou souřadnicemi x a y . Buňky mohou být čtvercové,

trojúhelníkové či šestiúhelníkové. Způsob zobrazení funguje tak, že každý jeden pixel má určenou hodnotu, která je představována při zobrazení rastru přiřazenou barvou nebo odstínem barvy. Při zobrazení z větší perspektivy se rastrový obraz jeví jako vektor, ovšem při přiblížení se zhoršuje kvalita rozlišení a objevují se nerovnosti v podobě jednotlivých pixelů (obr. 4).

Výhodou rastrových dat snadná lokalizace polohy pixelu s jasnou informací o obsahu pixelu.



Obr. 5: Rozdíl mezi rastrem a vektorem (dostupné z <http://www.stargen.cz/images/rozdil-vektor-rastr-full.jpg>)

8.2 Neprostorová data

Neprostorová data jsou všechna data, která nejsou lokalizována geografickou polohou místa na zemském povrchu. Jedná se také o tzv. data popisná, která představují vlastnosti nebo jevy, jenž chceme k prostorovým datům připojit.

Neprostorová data lze s prostorovými daty spojit po splnění jedné podmínky, kterou je existence jedinečného identifikátoru objektu. Tím je v podstatě vytvořena relace mezi tabulkou atributů a lokalizovaným prvkem.

9 Metody analýzy

V GIS je vyžadováno rozšíření dotazů na databázi pro možnost tyto údaje analyzovat ve vztahu k jejich lokalizaci. Základní analytickou částí GIS je Analyst². Jedná se o prostředí definující a využívající topologii vztahů mezi objekty ve vektorové reprezentaci a manipulaci s jejich atributovými popisy.

Při tvorbě všech analýz bylo třeba maximálně využít potenciálu dat a nástrojů GIS. Softwarový produkt ArcGIS je velmi bohatý na analyzové prostředky a obsáhnout znalosti potřebné k využití všech by bylo na několik prací, z tohoto důvodu byly vybrány čtyři, které nejlépe poslouží k dosažení stanoveného cíle. Jedná se o nadstavbové prvky Spatial Analyst, Network Analyst, 3D Analyst a Geostatistical Analyst.

Využitými metodami pro analýzu dat, a které budou rozebrány v další kapitole, jsou:

- Krigování
- Prostorová časová a vzdálenostní dostupnost (cost weighted)
- Síťové rozložení servisních oblastí
- 3D modelace
- Model přitažlivosti

9.1 Prostorová analýza

Prostorová analýza je tvořena v nadstavbě Spatial Analyst systému ArcGIS. Spatial Analyst umožňuje vytvářet plnohodnotná rastrová data, provádět nad nimi dotazy a analýzy a využívat i nových možností zobrazování těchto dat. Zahrnuje funkce jako je vytváření a správa rastrových dat, konverze vektorových témat na grid, analýzy povrchu, tvorby zón vzdáleností od prvků, určování blízkosti k prvkům, odvozování povrchu z hustoty prvků a analýzy dostupnosti místa, modelování terénu (sklon, orientace, vytváření vrstevnic, stínování svahů). Lze provádět lokální a zonální analýzy, překlasifikování rastrů a mnoho dalších analýz.

² (Tuček, 1998) překlad slova *analyst* = analytik, laborant

9.2 Sít'ová analýza

Sít'ová analýza je tvořena v nadstavbě Network Analyst systému ArcGIS. Network Analyst umožňuje řešit řadu problémů založených na geografických sítích jako jsou uliční a silniční sítě, říční síť, inženýrské sítě, produktovody a podobně. Úlohy zahrnují např. nalezení nejefektivnějšího průjezdu městem, vyhledání nejbližšího bodu od události, nebo definování servisních či prodejních oblastí na základě cestovního času. Umožňuje tedy především nacházet optimální cesty mezi skupinou bodů, nacházet nejbližší zařízení od události a provádět analýzy cestovního času.

9.3 Geostatistická analýza

Geostatistická analýza je tvořena v nadstavbě Geostatistical Analyst systému ArcGIS. Geostatistical Analyst je určen pro tvorbu spojitého povrchu z hodnot naměřených v rozptýlených bodech. Spolehlivě odhaduje hodnoty povrchu použitím interpolace kriging. Průzkumné nástroje pro analýzu prostorových dat umožňují pochopit podstatu dat: jejich rozložení, globální a lokální odchylky, globální trendy, úroveň prostorové autokorelace, odchylky mezi datovými sadami a další.

9.4 3D analýza

3D analýza je tvořena v nadstavbě 3D Analyst systému ArcGIS. 3D Analyst přináší do ArcGIS třetí rozměr. Pomocí jeho nástrojů je možné vytvářet, zobrazovat a analyzovat trojrozměrná data. Lze vytvářet a editovat modely terénu (GRID), generovat vrstevnice, vypočítávat svažitosť či viditelnost terénu, vytvářet řezy atd. Pomocí nástrojů 3D Analystu se lze na geografická data dívat z různých perspektivních pohledů. Možné je také rotovat scénami a prolétávat se nad nimi. Kromě 3D dat můžeme do těchto scén přidávat i ostatní vektorová a rastrová data.

9.5 Index potenciálu obce

Identifikovat potenciální zdroj uchazečů, resp. oblast našeho zájmu, je možné, pokud se podaří nalézt výchozí charakteristiky studentů a jejich regionů, společných asociací a přístupu. Data vztažená ke studentské dostupnosti jsou velice obsáhlá a obsahují veliké množství různých proměnných. Pro najetí správných propojení z dat, je nutné využít prostorových technik získávání informací. Nejprve je třeba charakterizovat zapsané studenty ve studiu pro uplynulý rok (příp. i za předešlé roky). K rozboru je možné využít statistická data z matriky a SLDB (např. počet studentů, studované obory, bydliště a SŠ studentů). Jako druhý krok se vybíjí charakterizovat administrativní jednotky, v rámci kterých je analýza prováděna (např. průměrný měsíční výdělek rodiny, celkový počet studentů z regionu, vzdálenost od školy, atp.). Tento krok následuje aplikace modelu, které zjistí velikost potenciálu oblasti. Pak už následuje interpretace výsledků a jejich zavedení do praxe.

Model míry přitažlivosti je velice uznávanou a užívanou metodou marketingových analýz, Ayad (2007) je uvádí jako základní nástroj akademického marketingu. Tento obecný model určuje pravděpodobnost zájmu zákazníka, který je umístěn uprostřed dvou zájmových míst, rozhodnout se právě pro jedno konkrétní za jasných podmínek, nikoliv pro druhé na základě náhody. Najít právě Index přitažlivosti je určen následující obecnou rovnicí:

$$I_i = X_i \times \frac{Y_i}{Z_i}$$

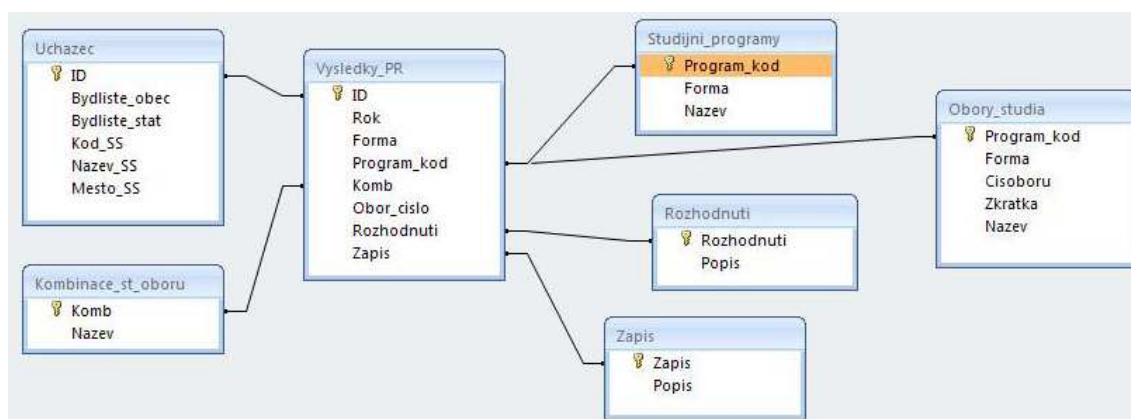
Tento index je počítán vždy pro každý region (i). V základě se jedná o vztah dvou proměnných Y a Z, mezi kterými je hledán vztah. Zpravidla je tento vztah doplněn další proměnnou, která reflektuje důležitost dalších faktorů ovlivňujících váhu jednotlivých oblastí.

10 Data z matriky TUL

Důležitým úkolem pro vytvoření prostorové analýzy dat z matriky FP TUL bylo získat potřebná data. Požadavkem bylo získat co možná nejkomplexnější informace o studentech (resp. uchazečích o studium) tak, aby nebyl porušen zákon č. 101/2000 Sb. o ochraně osobních údajů. K porušení zákona může dojít, pokud by data obsahovala osobní údaje uchazečů (jméno, příjmení, rodné číslo, datum narození, přesné bydliště).

Následně byl vytvořen balíček dat s údaji z let 2004-2008 (obr. 6), který byl plně k dispozici pro další práci. Podrobnějšímu rozboru dat v jednotlivých tabulkách a propojení relační databáze bude věnován prostor v dalších částech této práce. Obr. 6 nezobrazuje kompletní strukturu databáze matriky (chybí údaje, kterými by mohl být porušen zákon č. 101/2000 Sb.).

Matrika je naplňována daty přes IS STAG. Jde o informační systém studijní agentury vytvořený Západočeskou univerzitou, který využívá i TUL. Systém obsahuje řadu uživatelských funkcí, kterou jsou děleny dle funkce uživatele. Je možné vést evidenci o studentech, zadávat a editovat předměty, programy, obory, tisknout různé sestavy, atd.



Obr. 6: Struktura prvotních dat z matriky FP TUL

Po prvotním rozboru těchto dat bylo třeba učinit řadu opatření a úprav pro zajištění správnosti a přesnosti vstupních i výstupních dat k analýze. Jednotlivý úpravám bude věnován dostatek prostoru dále.

Pro následující podkapitoly je třeba ujednotit a rozlišit následující pojmy:

- **student** – posluchač zapsaný v některém studijním programu formou prezenční i kombinovanou;
- **uchazeč** – osoba účastnící se přijímacího řízení do některého ze studijních programů ve formě prezenční i kombinované

10.1 Popis prvotních tabulek

Tato podkapitola se bude věnovat datům, získaným z matricy FP TUL. Budou rozebrány jednotlivé tabulky včetně užitých datových typů a relačních propojení. Je důležité pečlivě prozkoumat všechny podrobnosti a potenciál informací, které jsou k dispozici. Tabulky byly z matriční databáze vyexportovány ve formátu CSV.

Tabulky obsahují dva typy dat, jedná se o number (číslo) a text (řetězec znaků).

10.1.1 Tabulka UHAZECI

Tabulka UHAZECI, která obsahuje základní informace o jednotlivých žadatelích ucházejících se o studium. Každý uchazeč je uveden právě jednou. V souboru jsou použity následující atributy (za názvem atributu je v závorce uveden použitý datový typ):

- **ID** (number) – jedinečný identifikátor uchazeče
- **BYDLISTE_OBEC** (text) – název obce trvalého bydliště uchazeče
- **BYDLISTE_STAT** (text) – název státu trvalého bydliště uchazeče
- **KOD_SS** (number) – kód ukončené SŠ, pokud byl uveden (jedná se o IZO, jedinečný identifikátor dle školského rejstříku MŠMT)
- **NAZEV_SS** (text) – název ukončené SŠ
- **MESTO_SS** (text) – sídelní obec ukončené SŠ

10.1.2 Tabulka VYSLEDKY_PR

Tabulka VYSLEDKY_PR, obsahuje výsledky přijímacího řízení uchazeče za jednotlivé akademické roky, ve kterých se hlásil na FP TUL. Zároveň je každý

uchazeč uveden pokaždé, kdy podal přihlášku ke studiu. V souboru jsou uvedeny následující atributy (za názvem atributu je v závorce uveden použitý datový typ):

- **ID** (number) – jedinečný identifikátor uchazeče
- **ROK** (number) – rok přijímacího řízení
- **FORMA** (text) – forma studia
- **PROGRAM_KOD** (text) – kód studijního programu
- **KOMB** (text) – označení kombinace studijních oborů
- **OBOR_CISLO** (text) – číslo studijního oboru (uveden pokud není kombinace)
- **ROZHODNUTI** (number) – kód rozhodnutí o přijetí uchazeče
- **ZAPIS** (number) – kód zápisu do studia

10.1.3 Tabulka **STUDIJNI_PROGRAMY_FP**

Tabulka **STUDIJNI_PROGRAMY_FP** obsahuje přehled studijních programů FP TUL. V souboru jsou uvedeny následující atributy (za názvem atributu je v závorce uveden použitý datový typ):

- **PROGRAM_KOD** (text) – kód studijního programu
- **FORMA** (text) – forma studia
- **NAZEV** (text) – název studijního programu

10.1.4 Tabulka **OBORY_STUDIA**

Tabulka **OBORY_STUDIA** zachycuje přehled studijních oborů v jednotlivých oborech na FP. V souboru jsou uvedeny následující atributy (za názvem atributu je v závorce uveden použitý datový typ):

- **PROGRAM_KOD** (text) – kód studijního programu
- **FORMA** (text) – forma studia
- **CISOBORU** (text) – kód studijního oboru
- **ZKRATKA** (text) – zkratka studijního oboru
- **NAZEV** (text) – název studijního oboru

10.1.5 Tabulka KOMBINACE_ST_OBORU

Tabulka KOMBINACE_ST_OBORU obsahuje přehled kombinací studijních oborů na FP. V souboru jsou uvedeny následující atributy (za názvem atributu je v závorce uveden použitý datový typ):

- **KOMB** (text) – označení kombinace studijních oborů
- **NAZEV** (text) – název kombinace studijních oborů

10.1.6 Tabulka ROZHODNUTI

Tabulka s názvem ROZHODNUTI obsahuje přehled kódů rozhodnutí o přijetí uchazeče. V souboru jsou uvedeny následující atributy (za názvem atributu je v závorce uveden použitý datový typ):

- **ROZHODNUTI** (number) – kód rozhodnutí o přijetí uchazeče
- **POPIS** (text) – popis rozhodnutí o přijetí uchazeče

10.1.7 Tabulka ZAPIS

Tabulka ZAPIS obsahuje přehled kódů zápisu uchazeče do studia. V souboru jsou uvedeny následující atributy (za názvem atributu je v závorce uveden použitý datový typ):

- **ZAPIS** (number) – kód zápisu uchazeče do studia
- **POPIS** (text) – popis kódu zápisu uchazeče do studia

11 Doplnková data

V této chvíli jsou data získaná z matricy kompletní, ale ke komplexní analýze ještě není databáze dostačující. Je nutné připojit další data, která doplní informace o jednotlivých územních jednotkách, případně pro vznik dalších atributů vytvoří vstupní data. Pakliže je cílem této práce vytvořit prostorovou analýzu geodatabáze, je tedy třeba zapojit mezi zpracovávaná data i demografické údaje územních jednotek. Jako další vstupní data v analýze dostupnosti bude využita dopravní síť (silnice a železnice). Při jednotlivých analýzách a tvorbě výstupů bude třeba vhodného topografického podkladu.

11.1 Data z ČSÚ

Při tvorbě analýzy dat z matricy je velice důležité zohlednit demografické údaje, které přísluší k danému územnímu celku. Tyto údaje potom dávají práci nový rozměr. Všechny informace, které se do této doby podařilo vybrat z databáze, jsou uvedeny v absolutních hodnotách. Právě díky zapojení takových dat, jako jsou data demografická, lze vytvářet výstupy v číslech relativních, tj. vyjadřující vlastnost nějakého prvku v závislosti na druhotném jevu.

Bylo tedy potřebné databázi z matricy obohatit o údaje z Českého statistického úřadu (dále jen ČSÚ). Data z ČSÚ byla získána bezplatně po domluvě s Mgr. Karlou Švehlovou, zaměstnankyní ČSÚ v Liberci.

Po identifikaci potřebných dat bylo možné připojit k tabulce UCHAZECI_UPR dvě nové tabulky, POČTY_OBYVATEL a OBYV_2008. Obě tabulky byly relací propojeny stejným klíčem, tj. atributem obsahující ICZÚJ. Obsah jednotlivých tabulek je následující:

OBCE

- **Kod obce** (text) – jedinečný identifikátor ICZÚJ
- **Název obce** (text) – název obce
- **2002** (number) – počet obyvatel v obci v roce 2002
- **2003** (number) – počet obyvatel v obci v roce 2003
- **2004** (number) – počet obyvatel v obci v roce 2004

- **2005** (number) – počet obyvatel v obci v roce 2005
- **2006** (number) – počet obyvatel v obci v roce 2006
- **2007** (number) – počet obyvatel v obci v roce 2007
- **Soucasny kod obce** (text) – současný kód ICZÚJ pro případ vzniku či zániku územní jednotky, primární klíč tabulky
- **Soucasny nazev obce** (text) – současný název obce pro případ vzniku či zániku územní jednotky
- **Existuje** (text) – označení číslicí 1 nebo 0 pro existenci obce
- **ORP_kod** (text) – identifikátor nadříděné obce s rozšířenou působností (dále jen ORP)
- **LAU1** (text) – kód okresu dle klasifikace místní samosprávy
- **NUTS3** – kód kraje dle klasifikace NUTS

Druhá tabulka, OBYV_2008, obsahuje pouze pět atributů, a sice:

- **LAU1** (text) – kód okresu dle klasifikace místní samosprávy
- **NUTS3** (text) – kód kraje dle klasifikace NUTS
- **NAZEOBCE** (text) – platný název obce
- **CISOBCE** (text) – jedinečný identifikátor ICZÚJ a zároveň primární klíč tabulky
- **STAV_31_12** (number) – počet obyvatel v obci k 31.12. 2008

Díky těmto tabulkám se celková velikost databáze rozrostla na dvanáct tabulek z původních sedmi (Obr. 7). Je tedy k dispozici dostatečné množství dat pro vyhotovení všech zamýšlených analýz.

11.2 Dopravní síť

Pro vytvoření dostupnostní analýzy do Liberce ze všech obcí v České republice je potřeba najít vhodnou dopravní síť. Tato analýza bude zpracována na podkladových datových sadách z balíku ArcČR 500, konkrétně silniční síti. V shapefile SILNICE je obsažen atribut TRIDA_SIL, který uchovává informaci o třídě silnice:

- D - Dálnice
- R – Rychlostní silnice
- 1 – Silnice I. třídy
- 2 – Silnice II. třídy
- O – Silnice ostatních kategorií

11.3 Topografický podklad

Pro tvorbu analýz a výstupů bude třeba využít dvou rozdílných topografických dat. Jako hlavní podkladový materiál bude sloužit veřejně přístupný mapový server společnosti CENIA. K přípravě síťových a prostorových analýz bude využít datový balík ArcČR 500.

11.3.1 Geoportál CENIA

Portál veřejné správy je velice moderním řešením řady aspektů. Jedná se o tzv. „úřad na internetu“. Je zde možné vyřizovat a zjišťovat řadu informací. Zároveň je standardně nabízena i řada služeb jakou je aplikace Mapové služby Portálu veřejné správy. Mapové služby lze připojit do silného klienta (ArcGIS) standardem WMS (Web Map Service). Data jsou přenášena jako rastrový obraz.

Služby využití pro potřebu diplomové práce:

- cenia_arccr_admin – administrativní členění České republiky
- cenia_dmu25 – digitální model území

11.3.2 ArcČR500

Jedná se o digitální vektorovou databázi pro území České republiky, zpracovanou v měřítku 1 : 500 000. Autorem datové sady je společnost Arcdata a ČÚZK (Český ústav zeměměřičský a katastrální). Obsahuje tři hlavní tematické skupiny map:

- Základní geografické prvky – silniční síť, lesní plochy, železniční síť, vodní toky, vrstevnice, atd.
- Administrativní členění – administrativní hranice územních jednotek
- Rozšiřující tématické informace – zeměpisná síť, letiště, hraniční přechody

Všechny vektorové vrstvy jsou zpracovány ve formátu SHP (shapefile) nebo coverage a rastrové vrstvy jsou uloženy do formátů GRID nebo TIFF. Vložené atributy jsou připojeny jako soubor INFO nebo DBF. ArcČR500 je zpracována v souřadném systému S-JTSK.

Licence je zakoupena Fakultou přírodovědně-humanitní a pedagogickou Technické univerzity v Liberci.

12 Analýza dat matriky

Cílem je popsat možný postup analýzy dat školní matriky tak, aby jej bylo možné použít opakovaně na další soubory dat. Vlastní analýze musí předcházet úprava dat do formátu vhodného pro zpracování v GIS. Výsledky analýz lze vizualizovat. Z toho vycházejí 4 fáze:

1. Předpříprava dat

První fáze se zabývá prvotním průzkumem tabulek, nalezením a opravením chyb, vytvořením relací a využitím jazyka SQL pro výběr sumarizovaných množin dat.

2. Vstup dat do GIS

Druhá fáze řeší konverzi vstupních souborů v prostředí ArcGIS do formátů DBF a SHP.

3. Analýza

Třetí fáze se zabývá analytickými metodami: krigováním, síťovou analýzou, analýzu servisních oblastí, 3D model a index přitažlivosti.

4. Výstup

Čtvrtá fáze řeší vizualizaci dat a jednotlivé výstupy na základě metod tématické kartografie.

13 Předpříprava dat

Na počátku první fáze jsou k dispozici jednotlivé tabulky bez propojení relacemi. Data z matrice TUL obsahují prostorovou informaci (obec střední školy uchazeče, obec bydliště uchazeče). V první fázi analýzy je třeba zpracovat a upravit tabulky a záznamy z matrice. Předpříprava dat se zabývá průzkumem tabulek, opravou chyb a tvorbou relací. Před ukončením první fáze jsou pomocí jazyka SQL z databáze vybrány potřebné sumarizované údaje v podobě několika tabulek.

13.1 Úprava dat

Ačkoliv jsou data velice obsáhlá, pro aplikaci analýzy bylo nutné provést řadu úprav některých atributů. Zároveň se některé tabulky ukázaly jako nedostatečné, po obsahové stránce, kdy díky těmto neúplnostem docházelo k chybám v analýzách, které později zkreslovaly výsledky. V následujících podkapitolách je rozebráno, co konkrétně bylo doplněno do obsahu jednotlivých tabulek a příslušné úpravy dat.

Tab. 2: Přehled chyb v datech

Chyba v datech	Tabulka
Typografické chyby	UCHAZECI
Nejednoznačný identifikátor obce	UCHAZECI
Nejednoznačný identifikátor střední školy	UCHAZECI
Nejednoznačný identifikátor studijního programu	STUDIJNI_PROGRAMY_FP
Nejednoznačný identifikátor studijního oboru	OBORY_STUDIA

13.1.1 Tabulka UCHAZECI

Tabulka UCHAZECI obsahuje podrobnější informace o jednotlivých uchazečích o studium. Jedná se o data z let 2004-2008. Každý uchazeč je vždy uveden pouze jednou. Tabulka obsahuje 11 072 záznamů.

Z toho počtu nemá 39 uchazečů v databázi uvedenou adresu, tudíž jsou tyto záznamy nepoužitelné pro další analýzu. Při lokalizaci bydlišť uchazečů je třeba

pracovat s daty vztahujícími se k České republice, jelikož v databázi je 34 uchazečů ze jiných států. V tabulce chybí jednoznačný identifikátor obce. Díky absenci tohoto prvku došlo ke zkreslení dat. V České republice je 275 obcí, které sdílí svůj název minimálně s jednou další obcí. V naší databázi se jedná o chybné zařazení 1225 uchazečů, což je 11 % ze všech přihlášených.

Ideálním řešením by tedy bylo připojení jedinečného identifikátorů obcí. Jelikož existuje všeobecně užívaný model identifikace územních jednotek, podaří se tímto způsobem upravit data. Na základě těchto úprav již bylo možné všechny obce přesně lokalizovat.

Do tabulky tedy přibyly následující atributy:

- **OKRES** (number) – jedná se o číselné označení NUTS4 (LAU1), kam spadá příslušná obec
- **OBEC_KOD** (text) – jednoznačný identifikátor obce nazývaná ICZÚJ, což je všeobecně užívané šestimístné *Identifikační číslo základní územní jednotky*
- **CAST_OBCE_KOD** (text) – identifikace části obce, kde má uchazeč trvalé bydliště

Pro potřebu dalšího zpracování dat bylo třeba určit, jaký typ střední školy jednotliví uchazeči navštěvovali. Předpokládaným vhodným klíčem, který obsahuje informaci o typu školy, by mělo být IZO, identifikační číslo ve školském rejstříku MŠMT. Bohužel je tento identifikátor přidělován popořadě bez klasifikace školy, čímž se stává bezcenným pro analýzu, nicméně je využit alespoň jako primární klíč tabulky. Jedinou možností na vyřešení tohoto problému je získání údajů o kódu studovaného oboru na SŠ. V rámci této práce byly jednotlivé střední školy označeny typem instituce (tab. 2). Tím vznikly dvě další tabulky připojené k databázi. První přiřazuje typ školy jednotlivým středním školám (tabulka SS_TYP) a druhá klasifikuje jednotlivé typy škol (tabulka TYP_SKOLY). Získáváme tak novou tabulku nazvanou UCHAZECI_UPR.

Tab. 3: Klasifikace typů středních škol

Typ školy	Označení
Gymnázium	1
Střední průmyslová škola	2
Obchodní akademie	3
Střední odborná škola	4
Konzervatoř	5
Střední odborné učiliště	6
Zahraniční škola	7
Integrovaná střední škola	8
Střední škola MV, MO ³	9

13.1.2 Tabulka STUDIJNI_PROGRAMY_FP

Tabulka STUDIJNI_PROGRAMY_FP obsahuje 28 studijních programů, které FP v letech 2004-2008 nabízela. Jedná se o programy jak prezenční, tak i kombinované. Každý studijní program je identifikován kódem XSSYY, kde X znamená typ studijního programu, SS je označení skupiny oborů, které dominují v daném programu, a YY znamená číslo studijního programu v rámci skupiny. Toto označení se nazývá K KOV. Pozice X může nabývat následujících hodnot:

- **B** – bakalářský studijní program
- **M** – magisterský studijní program
- **N** – magisterský studijní program navazující na bakalářský studijní program
- **P** – doktorský studijní program

Studijní program není rozlišen, zda jde o kombinovanou či prezenční formu. Proto bylo třeba vytvořit jedinečný identifikátor, označný jako KOD_PROG, v tomto případě textový řetězec složený z kódu K KOV studijního programu a zkratky formy programu (K – kombinovaná forma, P – prezenční forma). Pro názorný příklad lze uvést: prezenční forma studijního programu učitelství pro základní školy bude označen *M7503P* (příloha 1).

³ MV – Ministerstvo vnitra, MO – ministerstvo obrany

13.1.3 Tabulka OBORY_STUDIA

V tomto souboru dat jsou uloženy informace o oboru studia. Každý obor má jako jeden z atributů kód studijního programu do kterého spadá. Tato tabulka bohužel není příliš využitelná pro analýzu, jelikož se zde nenachází žádný jedinečný identifikátor. Je zde spousta oborů, které mají stejné číslo identifikační, ačkoliv se věnují jiným předmětům. Např. všechny prohloubené studijní programy mají totožné číslo oboru 7503T047 (příloha 2).

13.1.4 Tabulka KOMBINACE_ST_OBORU

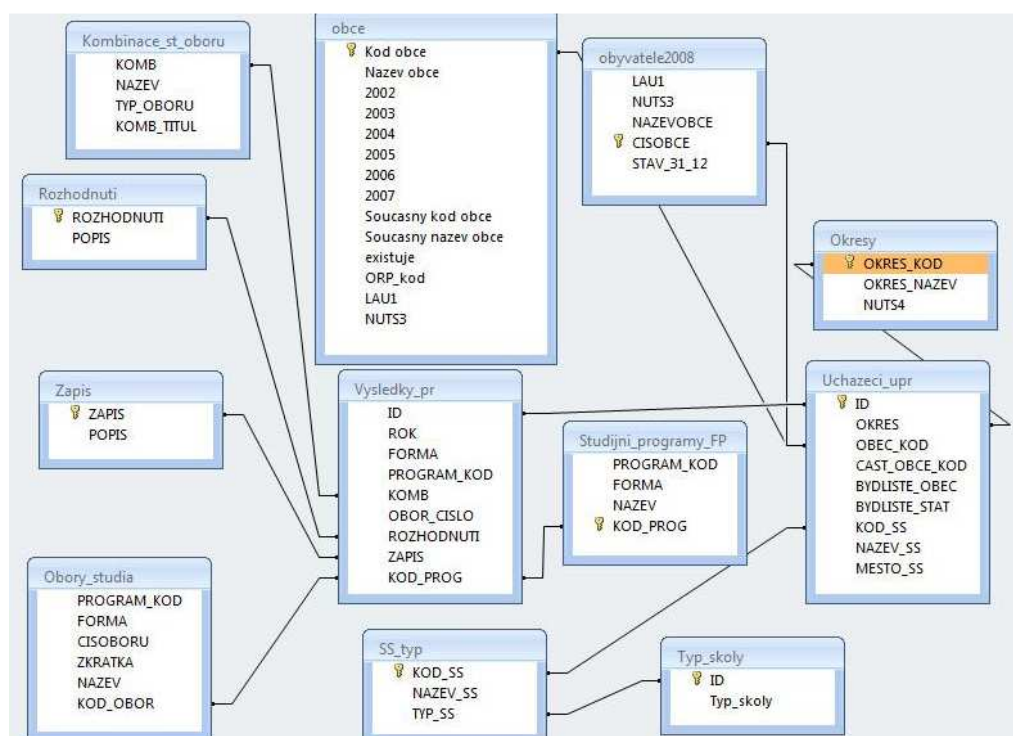
Tato tabulka KOMBINACE_ST_OBORU obsahuje označení a pojmenování všech 147 kombinací studijních oborů, které nabízí FP TUL. Atribut KOMB je vyjádřen ve tvaru XX-YY, kdy XX a YY jsou zkratky jednotlivých studijních programů. Není zde ale ošetřeno, zda se jedná o program magisterský, bakalářský či doktorský. Atribut NAZEV obsahuje název kombinace studijních oborů s tím, že bakalářský studijní program je doplněn za názvem zkratkou *bak.*. Bylo tedy třeba ručně u všech oborů nastavit atribut TYP_OBORU, kde je uvedeno buď MGR pro magisterský, BC pro bakalářský nebo PHD pro doktorský program. Jedinečný identifikátor KOMB_TITUL tak vznikne spojením KOMB a TYP_OBORU do jednoho textového řetězce, např. kombinace CHE-ZE znamená kombinaci studijních oborů Chemie – zeměpis, v magisterském programu a výsledný kód kombinace bude následně CHE-ZE_MGR.

13.1.5 Tabulka VYSLEDKY_PR

Tabulka VYSLEDKY_PR má v rámci celé databáze důležité postavení. Jak je patrné dle obr. 6, tvoří střed celé soustavy tabulek. Většina tabulek je přímo napojena na tuto tabulku a skládá se tak převážně z primárních klíčů relačně připojených dat. Bylo však nutné přidat nové atributy, které vznikly v jednotlivých dílčích tabulkách jako jedinečné klíče. Propojení jednotlivých částí je proto zajištěno přes následující primární klíče:

Tab. 4: Klíče pro relační spojení tabulek

Klíč (tabulka)	Klíč (tabulka)
ID (VYSLEDKY_PR)	ID (UCHAZECI_UPR)
KOMB (VYSLEDKY_PR)	KOMB (KOMBINACE_ST_OBORU)
ROZHODNUTI (VYSLEDKY_PR)	ROZHODNUTI (ROZHODNUTI)
ZAPIS (VYSLEDKY_PR)	ZAPIS (ZAPIS)
KOD_PROG (VYSLEDKY_PR)	KOD_PROG (STUDIJNI_PROGRAMY_FP)
OBOR_CISLO (VYSLEDKY_PR)	CISOBORU (OBORY_STUDIA)



Obr. 7: Struktura kompletní databáze

Toto funkční schéma s upravnými atributy již poskytuje plnou propojenost celé databáze a je možné vytvářet výběry dat podle kritérií, anglicky nazváno *query*.

13.2 Tvorba databáze

Pro propojení jednotlivých tabulek, které jsou zpravidla k dispozici ve formátu XLS, je výhodné využít některého databázového prostředku jako je Microsoft Access. Nejen že obsahuje řadu pomocných funkcí a průvodců, které usnadní práci i méně pokročilým uživatelům, ale také vyhledavací mechanismus (Query builder), vybírající data pomocí nastavených podmínek. Tento prostředek interpretuje příkaz SELECT na databázi, což snižuje nároky na znalost struktury tohoto příkazu.

13.3 Výběr dat

Před vstupem do další fáze analýzy byla provedena selekce následujících množin dat.

Platí, že všechny množiny jsou vytvářeny pro jednotlivé obce v letech 2004-2008.

- Počet uchazečů
- Počty uchazečů s údajem o zápisu do studia
- Počet uchazečů s údajem o rozhodnutí přijímacího řízení
- Počet uchazečů o studijní programy
- Počet uchazečů o studijní program a kombinaci studijních oborů
- Počet uchazečů o studijní program, studijní obor a formu studia
- Počty obyvatel ve věkové skupině 15-19 let a počty uchaze

Dále jsou data vybrána pro jednotlivé okresy v letech 2004-2008.

- Počty uchazečů na studijní programy
- Počty uchazečů s údajem o absolvované střední škole

Pro výběr dat k jednotlivým analýzám je využito příkazů SELECT v následující ukázkové struktuře.

- **Uchazeči, jenž jsou zapsáni do studia, dle obcí v roce 2004**

```

SELECT Uchazeci_upr.BYDLISTE_STAT, Count(Vysledky_pr.ID) AS CountOfID,
Uchazeci_upr.OBEC_KOD, Zapis.ZAPIS, Vysledky_pr.ROK
FROM Uchazeci_upr INNER JOIN (Rozhodnuti INNER JOIN (Zapis INNER JOIN
Vysledky_pr ON Zapis.ZAPIS = Vysledky_pr.ZAPIS) ON Rozhodnuti.ROZHODNUTI
= Vysledky_pr.ROZHODNUTI) ON Uchazeci_upr.ID = Vysledky_pr.ID
GROUP BY Uchazeci_upr.BYDLISTE_STAT, Uchazeci_upr.OBEC_KOD,
Zapis.ZAPIS, Vysledky_pr.ROK, Uchazeci_upr.BYDLISTE_OBEC
HAVING (((Uchazeci_upr.BYDLISTE_STAT)="česká republika") AND
((Zapis.ZAPIS)=1) AND ((Vysledky_pr.ROK)=2004));

```

- **Studijních programy a kombinace, na které se hlásí uchazeči z dané obce v jednotlivých letech**

```

SELECT Uchazeci_upr.BYDLISTE_STAT, Vysledky_pr.ROK,
Uchazeci_upr.OBEC_KOD, Uchazeci_upr.BYDLISTE_OBEC,
Studijni_programy_FP.NAZEV, Kombinace_st_oboru.NAZEV, Vysledky_pr.FORMA
FROM Uchazeci_upr INNER JOIN (Studijni_programy_FP INNER JOIN
(Kombinace_st_oboru INNER JOIN Vysledky_pr ON Kombinace_st_oboru.KOMB =
Vysledky_pr.KOMB) ON Studijni_programy_FP.KOD_PROG =
Vysledky_pr.KOD_PROG) ON Uchazeci_upr.ID = Vysledky_pr.ID
WHERE (((Uchazeci_upr.BYDLISTE_STAT)="česká republika"))
ORDER BY Vysledky_pr.ROK, Uchazeci_upr.OBEC_KOD;

```

- **Obce s informací o typu SŠ, které studovali jednotliví uchazeči**

```

SELECT Typ_skoly.Typ_skoly, Count(Typ_skoly.Typ_skoly) AS CountOfTyp_skoly,
Uchazeci_upr.OBEC_KOD, Uchazeci_upr.BYDLISTE_OBEC
FROM Typ_skoly INNER JOIN (SS_typ INNER JOIN (Uchazeci_upr INNER JOIN
Vysledky_pr ON Uchazeci_upr.ID = Vysledky_pr.ID) ON SS_typ.KOD_SS =
Uchazeci_upr.KOD_SS) ON Typ_skoly.ID = SS_typ.TYP_SS
GROUP BY Typ_skoly.Typ_skoly, Uchazeci_upr.OBEC_KOD,
Uchazeci_upr.BYDLISTE_OBEC
ORDER BY Uchazeci_upr.BYDLISTE_OBEC;

```

- **Počtech obyvatel v jednotlivých obcích v jednotlivých letech**

```
SELECT obce.[Kod obce], obce.[2004], obce.[2005], obce.[2006], obce.[2007],  
obyvatele2008.STAV_31_12  
FROM obce INNER JOIN obyvatele2008 ON obce.[Kod obce] =  
obyvatele2008.CISOBCE  
GROUP BY obce.[Kod obce], obce.[2004], obce.[2005], obce.[2006], obce.[2007],  
obyvatele2008.STAV_31_12, obce.[Nazev obce];
```

14 Vstup dat do GIS

Vybraná data byla vždy exportována do formátu XLS, poté otevřena v programu ArcGIS a převedena do databázového formátu DBF. Export z Microsoft Access přímo do formátu DBF nebyl možný z důvodu nekompatibility verzí DBF. Důvodem pro export XLS do DBF je kompatibilita formátu DBF ve velkém množství programů na rozdíl od XLS.

Provedením funkce JOIN jsou tato data připojena k vrstvě OBCE.SHP, čímž dojde ke spojení prostorových dat (obce.shp) s atributovými (matrika). Spojenou tabulku je možné trvale uložit exportem do nového SHP.

Všechny záznamy v obce.shp, pro které není záznam v matrice, jsou vyplněny hodnotou <null> (Obr. 8). Tato hodnota označuje prázdnou hodnotu. Pro analýzu je ovšem třeba namísto hodnoty <null> dosadit nulu, jelikož <null> není počítáno mezi analyzované hodnoty. Tento problém řeší výše popsany export do nového SHP, kdy všechny hodnoty <null> jsou nahrazeny nulou (Obr. 9).

FID	Shape	ICZUJ	NAZEV	NUTSS	OID	BYDLISTE S	CountOID	OBEC KOD	ZAPIS	ROK
1	Polygon	545708	Lobendava	CZ0421545708	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
2	Polygon	562661	Lipová	CZ0421562661	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
3	Polygon	562958	Sluknov	CZ0421562958	90	Česká republika	3	562958	1	2004
4	Polygon	562441	Doňm Poustevna	CZ0421562441	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
5	Polygon	562912	Veslý Šenov	CZ0421562912	83	Česká republika	1	562912	1	2004
6	Polygon	562947	Vátemov	CZ0421562947	94	Česká republika	1	562947	1	2004
7	Polygon	562581	Jiřkov	CZ0421562581	85	Česká republika	1	562581	1	2004
8	Polygon	562751	Mikulášovice	CZ0421562751	87	Česká republika	1	562751	1	2004
9	Polygon	562823	Staré Křečany	CZ0421562823	89	Česká republika	1	562823	1	2004
10	Polygon	549996	Černousy	CZ0513549996	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
11	Polygon	562777	Rumburk	CZ0421562777	88	Česká republika	6	562777	1	2004
12	Polygon	564036	Habartice	CZ0513564036	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
13	Polygon	564494	Všitkov	CZ0513564494	114	Česká republika	2	564494	1	2004
14	Polygon	563935	Butovka	CZ0513563935	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
15	Polygon	564311	Pertolbice	CZ0513564311	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
16	Polygon	562611	Krásná Lpa	CZ0421562611	86	Česká republika	3	562611	1	2004
17	Polygon	564079	Horní Ransnice	CZ0513564079	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
18	Polygon	530417	Douboice	CZ0421530417	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
19	Polygon	562862	Varnsdorf	CZ0421562862	91	Česká republika	5	562862	1	2004
20	Polygon	564133	Jindřichovice pod Smrkem	CZ0513564133	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
21	Polygon	563994	Doňm Ransnice	CZ0513563994	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
22	Polygon	562513	Hřensko	CZ0421562513	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
23	Polygon	562556	Jetřichovice	CZ0421562556	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
24	Polygon	564168	Krásný Les	CZ0513564168	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
25	Polygon	530433	Kunratice	CZ0513530433	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
26	Polygon	564028	Fřýdlant	CZ0513564028	107	Česká republika	6	564028	1	2004
27	Polygon	562505	Horní Podluží	CZ0421562505	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
28	Polygon	562793	Rybníště	CZ0421562793	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
29	Polygon	544680	Janov	CZ0421544680	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
30	Polygon	562432	Doňm Podluží	CZ0421562432	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
31	Polygon	564265	Nové Město pod Smrkem	CZ0513564265	111	Česká republika	1	564265	1	2004
32	Polygon	544701	Labská Stráž	CZ0421544701	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
33	Polygon	562335	Děčín	CZ0421562335	82	Česká republika	7	562335	1	2004
34	Polygon	568900	Růžová	CZ0421568900	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
35	Polygon	562530	Chřebá	CZ0421562530	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
36	Polygon	562343	Arnolice	CZ0421562343	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
37	Polygon	562572	Jiřetin pod Jedlovou	CZ0421562572	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>
38	Polygon	564371	Raspennava	CZ0513564371	113	Česká republika	3	564371	1	2004

Obr. 8: Atributová tabulka SHP po připojení XLS nebo DBF

FID	Shape *	ICZUJ	NAZEV	MUTS5	OID	RYDLISTE \$	CountOfID	OBEC KOD	ZAPIS	ROK
0	Polygon	545708	Lobendava	CZ0421545708	0		0		0	0
1	Polygon	562861	Lipová	CZ0421562861	0		0		0	0
2	Polygon	562858	Šluknov	CZ0421562858	90	Česká republika	3	562858	1	2004
3	Polygon	562441	Dolní Poustevna	CZ0421562441	0		0		0	0
4	Polygon	562912	Velký Šenov	CZ0421562912	93	Česká republika	1	562912	1	2004
5	Polygon	562947	Vilémov	CZ0421562947	94	Česká republika	1	562947	1	2004
6	Polygon	562581	Jířkov	CZ0421562581	85	Česká republika	1	562581	1	2004
7	Polygon	562751	Mikulášovice	CZ0421562751	87	Česká republika	1	562751	1	2004
8	Polygon	562823	Staré Křečany	CZ0421562823	89	Česká republika	1	562823	1	2004
9	Polygon	545996	Cernousy	CZ0513545996	0		0		0	0
10	Polygon	562777	Rumburk	CZ0421562777	88	Česká republika	6	562777	1	2004
11	Polygon	564036	Habartice	CZ0513564036	0		0		0	0
12	Polygon	564494	Višňová	CZ0513564494	114	Česká republika	2	564494	1	2004
13	Polygon	563935	Bulovka	CZ0513563935	0		0		0	0
14	Polygon	564311	Pertolice	CZ0513564311	0		0		0	0
15	Polygon	562611	Krásná Lípa	CZ0421562611	86	Česká republika	3	562611	1	2004
16	Polygon	564079	Horní Řasnice	CZ0513564079	0		0		0	0
17	Polygon	530417	Doubovice	CZ0421530417	0		0		0	0
18	Polygon	562882	Varnsdorf	CZ0421562882	91	Česká republika	5	562882	1	2004
19	Polygon	564133	Jindřichovice pod Smrkem	CZ0513564133	0		0		0	0
20	Polygon	563994	Dolní Řasnice	CZ0513563994	0		0		0	0
21	Polygon	562513	Hřensko	CZ0421562513	0		0		0	0
22	Polygon	562556	Jetřichovice	CZ0421562556	0		0		0	0
23	Polygon	564168	Krásný Les	CZ0513564168	0		0		0	0
24	Polygon	530433	Kunratice	CZ0513530433	0		0		0	0
25	Polygon	564028	Frydlant	CZ0513564028	107	Česká republika	6	564028	1	2004
26	Polygon	562505	Horní Podluží	CZ0421562505	0		0		0	0
27	Polygon	562793	Rybníště	CZ0421562793	0		0		0	0
28	Polygon	544680	Janov	CZ0421544680	0		0		0	0
29	Polygon	562432	Dolní Podluží	CZ0421562432	0		0		0	0
30	Polygon	564265	Nové Město pod Smrkem	CZ0513564265	111	Česká republika	1	564265	1	2004
31	Polygon	544701	Labská Stráň	CZ0421544701	0		0		0	0
32	Polygon	562335	Děčín	CZ0421562335	82	Česká republika	7	562335	1	2004
33	Polygon	566900	Růžová	CZ0421566900	0		0		0	0
34	Polygon	562530	Chřibská	CZ0421562530	0		0		0	0
35	Polygon	562343	Arnoltice	CZ0421562343	0		0		0	0
36	Polygon	562572	Jířetín pod Jedlovou	CZ0421562572	0		0		0	0
37	Polygon	564371	Raspenava	CZ0513564371	113	Česká republika	3	564371	1	2004

Obr. 9: Nově vyexportovaný SHP

Důležitým aspektem tvorby relační databáze je tvorba relací a identifikace primárních klíčů. Zvláště pokud je propojováno více tabulek z různých zdrojů, je třeba dbát na totožný datový typ propojených klíčů. Dva prvky s různými datovými prvky nelze funkčně spojit relací. S tím souvisí i předpoklad dalšího využití, jelikož některé analýzy nelze provádět s určitými typy dat. Jelikož byla všechna exportovaná data dále zpracovávána v programu ArcGIS, dalo se případné přetypování atributů řešit interními funkcemi přímo zde (v atributové tabulce lze přidat pole, do kterého se uloží přetypovaná hodnota pomocí funkce Field calculator).

15 Analýza

Ve třetí fázi dochází k samotné analýze dat z matriky. V této kapitole jsou popsány dílčí kroky každé z pěti aplikovaných analýz.

15.1 Časová a vzdálenostní dostupnost

Pomocí extenze Spatial Analyst (Wilson, 2001) je možné vytvořit analýzu dostupnosti vzdálenostní nebo časové do jednoho bodu, v tomto případě do Liberce. Jelikož data vyžadují postupné úpravy, budou dále popsány jednotlivé kroky analýzy (Obr. 10). Analýza bude provedena na liniové vrstvě silniční sítě. Ekvivalentně se provádí s údajem o dojezdovém čase a případně lze celou analýzu zopakovat na železniční síti.

Úprava vstupních dat

Nejprve je třeba upravit vstupní vrstvu, tedy *cestni_sit.shp*. Vrstva obsahuje analyzovaná data a do atributové tabulky jsou přidány dvě pole, které budou obsahovat hodnoty pro pixely nově vytvářeného rasteru. Pole pojmenované *bunka* vyjadřuje vzdálenost jednoho bodu a bude tedy nabývat hodnoty 100 (1 bod = 100 m) Druhé pole, nazvané *rychlost*, určuje jakou rychlostí lze projet jedním bodem. Tento údaj samozřejmě závisí na typu komunikace. Doplněná hodnota bude ve formátu $vv/rr*1000$, kde *vv* je délka jednoho bodu a *rr* je průměrná rychlost komunikace, která bodem prochází (Tab. 4).

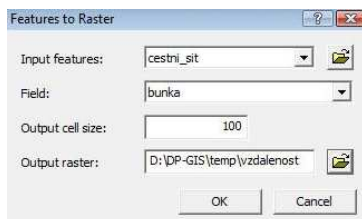
Tab. 5: Průměrné rychlosti na komunikacích⁴

Typ komunikace	Průměrná rychlost
Dálnice	90 km/h
Rychlostní silnice	70 km/h
Silnice I. třídy	60 km/h
Silnice II. třídy	45 km/h
Ostatní komunikace	40 km/h

⁴ PEŇÁZ, T.: *Analýzy dopravní dostupnosti v prostředí ArcGIS*. Arcvue, 2006, no.1, s. 9-11. ISSN 1211-2135

Rasterizace

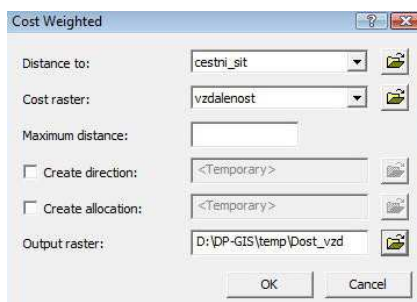
Dalším krokem je předvod vrstvy do rastru pomocí nástroje *Feature to Raster*. Zde je třeba nastavit jako *Field: bunka* a *Output cell size: 100* (obr. 10).



Obr. 10: Nastavení nástroje Features to Raster

Ohodnocení rastru

Následuje funkce *Cost Weighted (Cost Distance)*, kde je nastaveno místo, kam je zjišťována vzdálenost (Liberec), a *Cost raster (vzdalenost)* (Obr. 11).



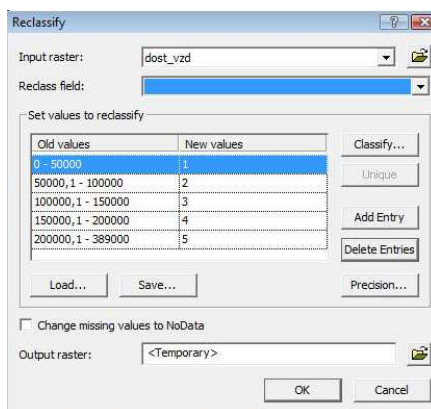
Obr. 11: Nastavení nástroje Cost Weighted

Rekalkulace vzdáleností

Vznikl rastr, který má ohodnocenu každou buňku vzdáleností do Liberce. Aby vzdálenost byla v metrech, je třeba funkcí *Raster calculator* hodnotu vzdálenosti vydělit 100. Výstupem bude ten samý rastr přepočítaný na nové hodnoty.

Reklasifikace hodnot

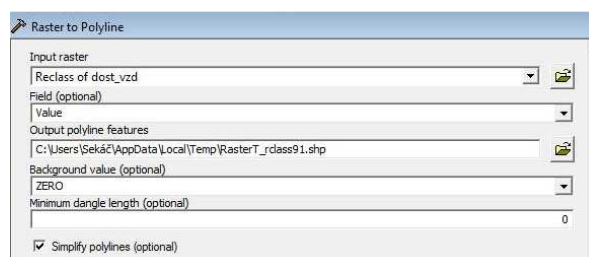
Dalším krokem je užití *Reclassify*, kterým jsou reklasifikována data uložena v rastru *Calculation* (Obr. 12). Nové rozptyly vzdáleností v metrech budou: 0-50 000,1 (1), 50 000,1-100 000 (2), 100 000,1-150 000 (3), 150 000,1-200 000 (4) a 200 000,1-389000 (5).



Obr. 12: Nastavení nástroje Reclassify

Vektorizace

Následně je nutné data převést zpět do vektorové formy (obr. 13), a to nástrojem *Raster to Polyline* (umístěn v ArcToolbox).



Obr. 13: Nastavení nástroje Raster to Polyline

Rozřezání komunikací

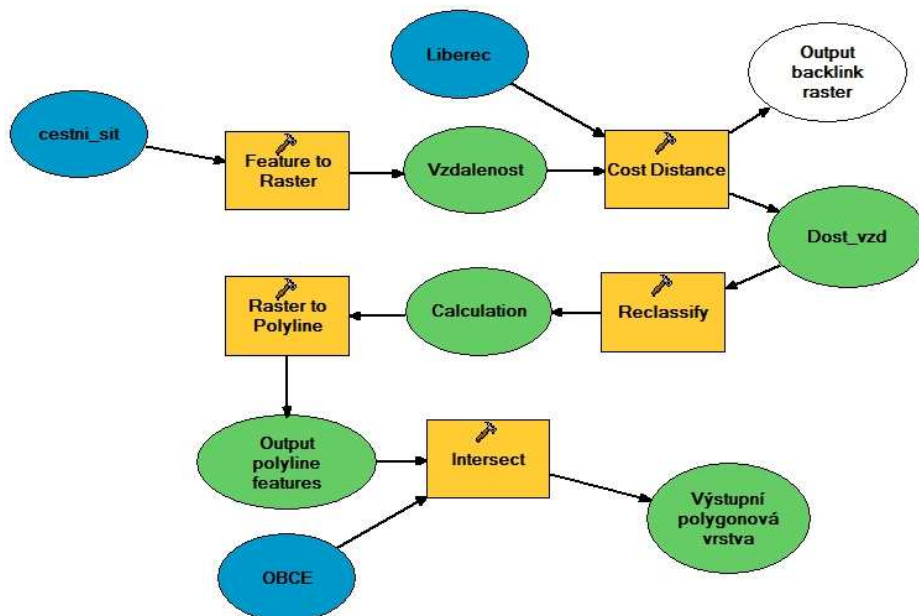
Na výslednou vektorovou vrstvu silniční sítě spolu s polygonovou vrstvou obcí je použit prostředek *Intersect* (umístěn v ArcToolbox), který silniční síť rozdělí na linie ohraničené jednotlivými obcemi. Každá obec tak získá atribut s informací, jak daleko leží od Liberce (hodnota 1-5). Ukázka: území obce Kolín (obr. 14) obsahuje komunikace s označením 2 (zelená) a 3 (žlutá).

GRID CODE	FROM NODE	TO NODE	FID OBCE	AREA	PERIMETER	OBCE	OBCE_ID	ICZUJ	NAZEV
3	9715	9722	1895	35288072	40392,945	1896	1896	533165	Kolín
3	9722	9723	1895	35288072	40392,945	1896	1896	533165	Kolín
3	9715	9723	1895	35288072	40392,945	1896	1896	533165	Kolín
3	9723	9729	1895	35288072	40392,945	1896	1896	533165	Kolín
3	9723	9729	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9780	9675	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9799	8737	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9799	9600	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9800	9801	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9801	9780	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9801	9802	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9780	9802	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9780	9803	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9799	9812	1895	35288072	40392,945	1896	1896	533165	Kolín
2	9812	9800	1895	35288072	40392,945	1896	1896	533165	Kolín

Obr. 14: Ohodnocené komunikace v obci Kolín a ukázka atributové tabulky

Sumarizace

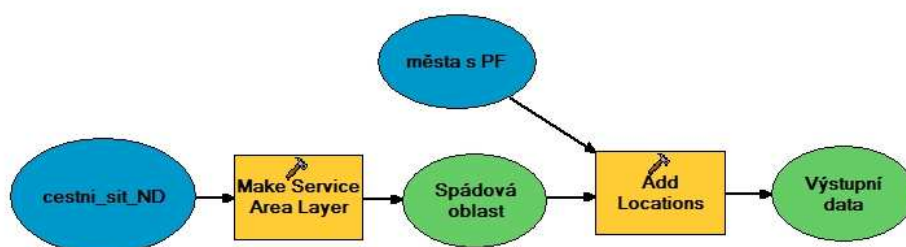
Takto rozsekanou vrstvu je třeba sumarizovat (v atributové tabulce příkaz *Summarize*), aby každá obec měla pouze jeden atribut, obsahující vzdálenost s hodnotu, a to vždy nejnižší možný.



Obr.15: Modelové zobrazení prostorové analýzy

15.2 Servisní oblast

Extenze Network Analyst (Mandloi, 2008) bude využit pro síťovou analýzu, ve které budou vytvořeny spádové oblasti jednotlivých pedagogických fakult v České republice. Konkrétní aplikací bude hledání *servisních oblastí*, na základě vzdálenosti od sídla fakulty po silniční síti. Je tedy důležité připravit dva druhy dat. První bodová vrstva bude obsahovat sídla pedagogických fakult, kterých je v České republice 9 (Liberec, Ústí nad Labem, Praha, Plzeň, České Budějovice, Hradec Králové, Brno, Olomouc, Ostarava). Druhou připravenou vrstvou je silniční síť, na které bude analýza provedena⁵ (Obr. 16).



Obr. 16: Modelové zobrazení síťové analýzy

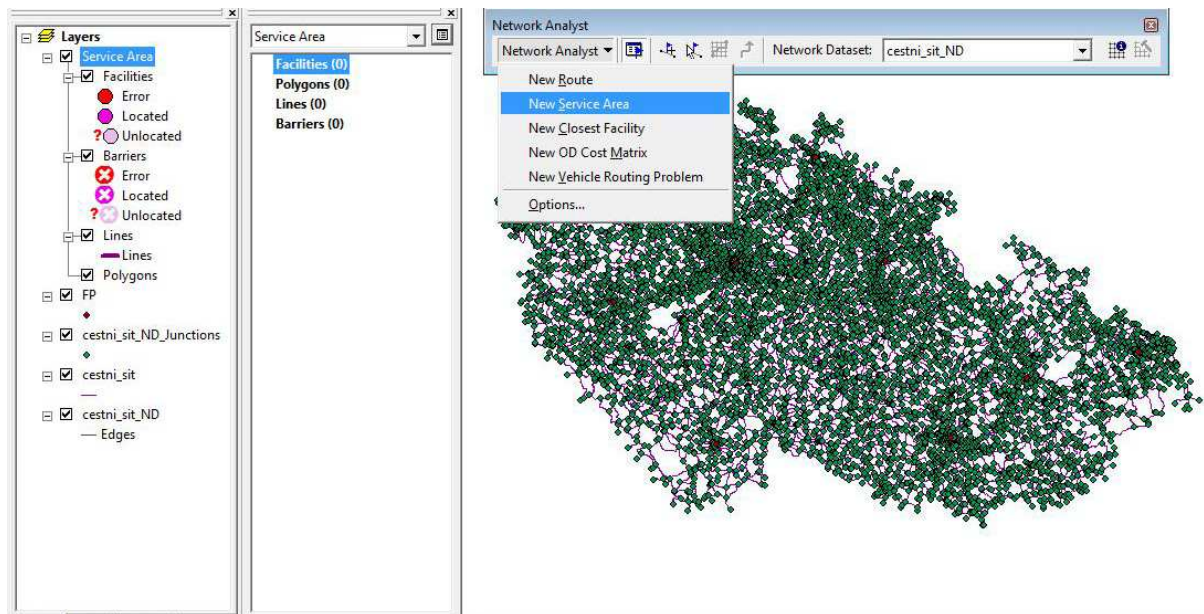
Network dataset

Prvním důležitým krokem je úprava dat silniční sítě. V ArcCatalogu (součást balíčku ArcGIS) je třeba vytvořit na silniční síti **network dataset**. Jedná se o specifický formát souboru, díky kterému lze aplikovat tuto síťovou analýzu. Network dataset je odlišný od shapefilu tím, že se do liniové struktury dat implementuje síťový model. Transformovaná data se budou nazývat *cestni_sit_ND*

Servisní oblasti

Poté již v ArcMapu není problém tuto vrstvu analyzovat. Je využit nástroj *New service area*, který vytvoří kostru pro doplnění dat a poté ji vyhodnotí. K těmto úkonům je potřeba aktivovat *Network Analyst Window*, v němž probíhají další kroky analýzy (obr. 17).

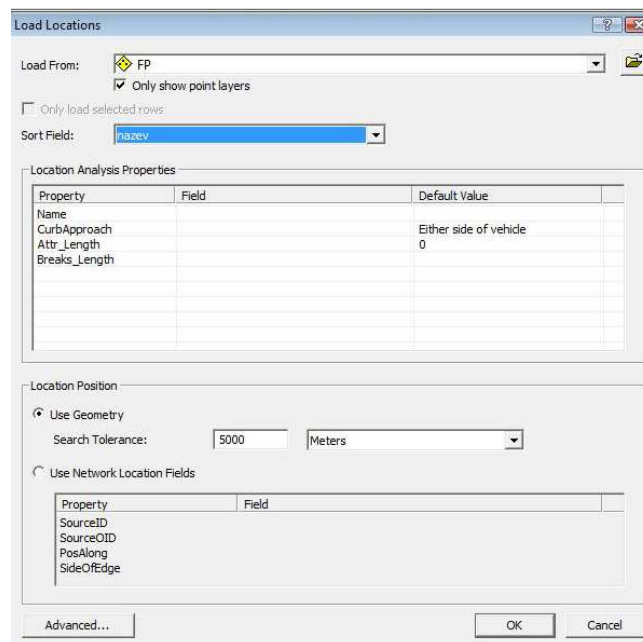
⁵ Vrstva ArcČR 500 Silnice.shp



Obr. 17: Spuštění vyhodnocení servisních oblastí

Lokalizace servisních středisek

Po vytvoření kostry je možné vkládat *Facilities* (servisní střediska, kolem nichž je vytvořena spádová oblast), *Barriers* (bariéry bránící průjezdu) a *Lines* (doplňující linie). Zde je tedy důležité vložit bodovou vrstvu fakult (FP.shp) do *Facilities*. Přes kontextové menu zvolíme u *Facilities* možnost *Load Locations* (obr. 18).



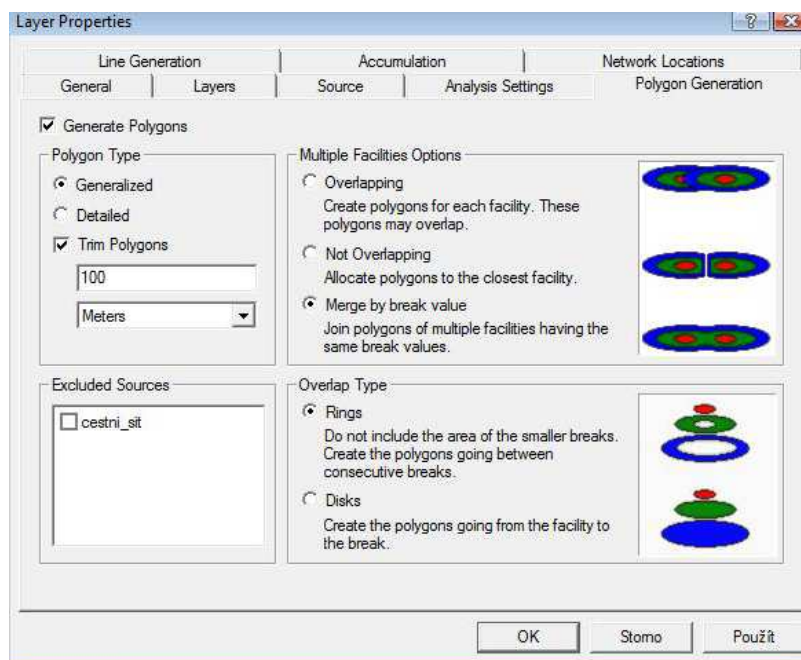
Obr. 18: Přiřazení sídel fakult jako servisních středisek

Řešení analýzy

Poté je nutné provést dodatečné úpravy vrstvy přes Layer properties (stále v Network Analyst Window) a jako poslední krok spustit řešení analýzy *Solve* v nástrojové liště Network Analyst.

Vizualizace servisních oblastí

Posledním krokem síťové analýzy je zobrazení servisních oblastí jednotlivých pedagogických fakult v ČR. Nastavení probíhá přes vlastnosti vrstvy Service Area (obr. 19). V záložce Polygon Generation je nastavováno zobrazení servisní oblasti. V záložce Analysis Settings je nutné nastavit Impedanci (v metrech) a hranice intervalů vzdáleností (25 000, 50 000, 75 000, 100 000).



Obr. 19: Nastavení zobrazení servisní oblasti

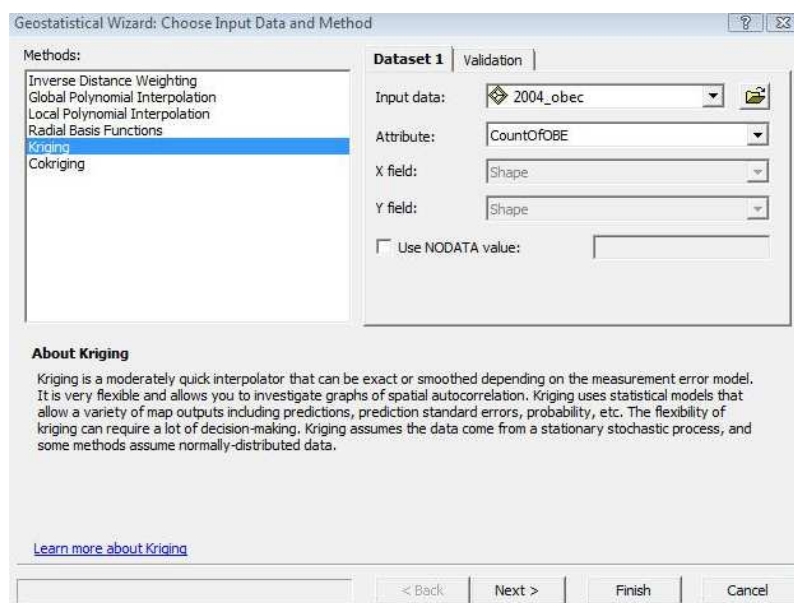
15.3 Kriging

Nástroj Geostatistical Analyst (Johnston, 2001) bude využit k analýze nazývané Kriging (krigování). Jedná se o interpolační metodu, která využívá geostacionární metody lokálního odhadu. Lokální odhad znamená výpočet pravděpodobné hodnoty proměnné buď v bodě, kde nebylo provedeno měření (či nemáme data), nebo v relativně malé ploše. Výchozí podmínky krigování jsou dány buď semivariogramem, nebo kovariací. Semivariogram je míra variace, která je určena vztahem mezi sousedními daty. Kovariance je míra vzájemné vazby mezi dvěma prvky.

Po spuštění Geostatistical Analystu je třeba spustit Geostatistical Wizard, který provádí celým procesem analýzy. Tento průvodce připraví celou analýzu ve čtyřech krocích.

Volba metody

Nejprve je třeba zvolit funkci Kriging. Dále je zde určena vstupní vrstva, což bude SHP s údajem o počtu uchazečů z obcí v jednotlivých letech a analyzovaný atribut (obr. 20).



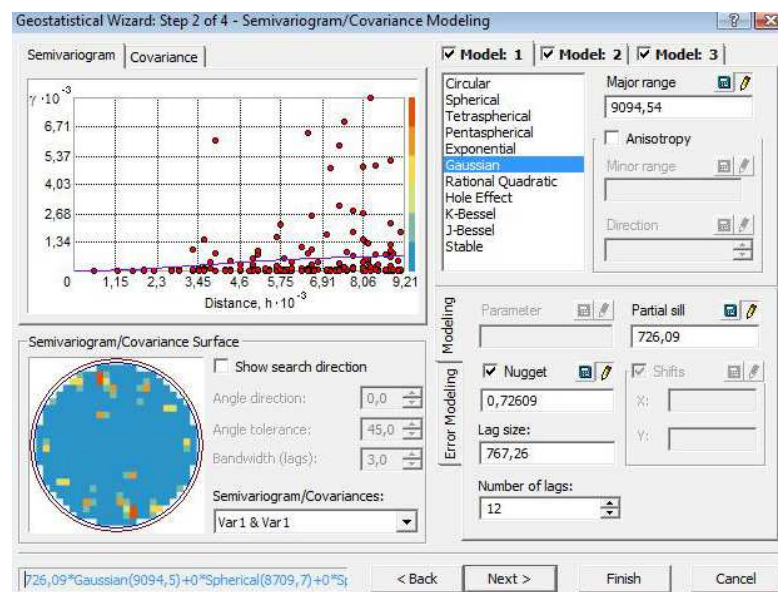
Obr. 20: Geostatistical wizard - volba metody

Geostatistická metoda

Metoda Kriging má mnoho podob, dle počtu vstupní dat a složitosti výpočtu, pro tuto analýzu bude využita *Predition map* (Předpovědní mapa) z nabídky Ordinary Kriging.

Styl modelování

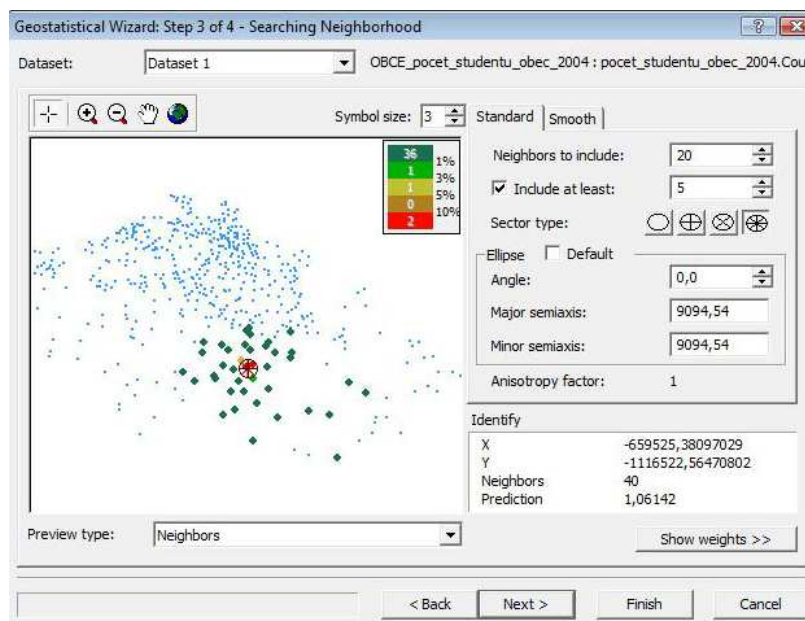
Následuje výběr mezi semivariogramem a kovariací. Zvolena bude semivariace. Je možné nastavit model funkce (semivariogram nebo kovariance), čímž dodáme výsledku určitý průběh. Lze kombinovat až tři různé modely (obr. 21).



Obr. 21: Geostatistical wizard - styl modelování

Hledání sousedů

O krok dále dominuje dialogovému oknu pole s červeným křížem, který vyznačuje místo, jenž nemá stanovenou hodnotu a bude vyhodnoceno. Vyznačené body jsou sousední místa s naměřenou hodnotou. Změnou parametru *Neighbours to include* lze nastavit, z kolika okolních bodů bude výsledná hodnota počítána. Počet zahrnutých sousedů do výpočtu bude 20 (obr. 22).



Obr. 22: Geostatistical wizard - hledání sousedů

Přehled užitých metod

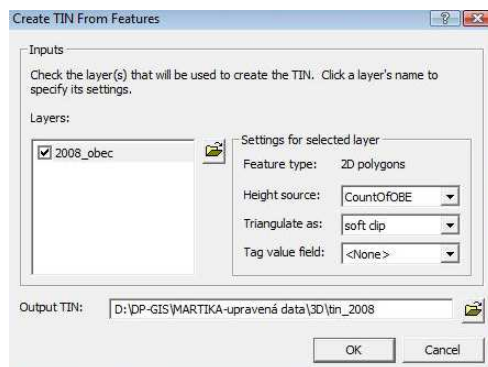
Hned v dalším okně je z grafu patrné, kde se pohybují předpovězené hodnoty (modrá linie). Poté již vznikne model analýzy, který barevně odlišuje různě ohodnocené oblasti zkoumaného území.

15.4 3D model území

Nadstavba 3D Analyst (MacDonald, 2002) dovoluje vytvořit prostorové zobrazení analyzovaných dat. Pro potřebu této práce postačí vizualizace informace, v jakém poměru jsou uchazeči a počet obyvatel (ve věku 15-19 let) z jednotlivých obcí. Jedná se spíše o informativní pohled. 3D model lze v prostředí ArcGIS vytvořit z rastrového formátu TIN, což je nepravidelná trojúhelníková síť.

Tvorba vrstvy TIN

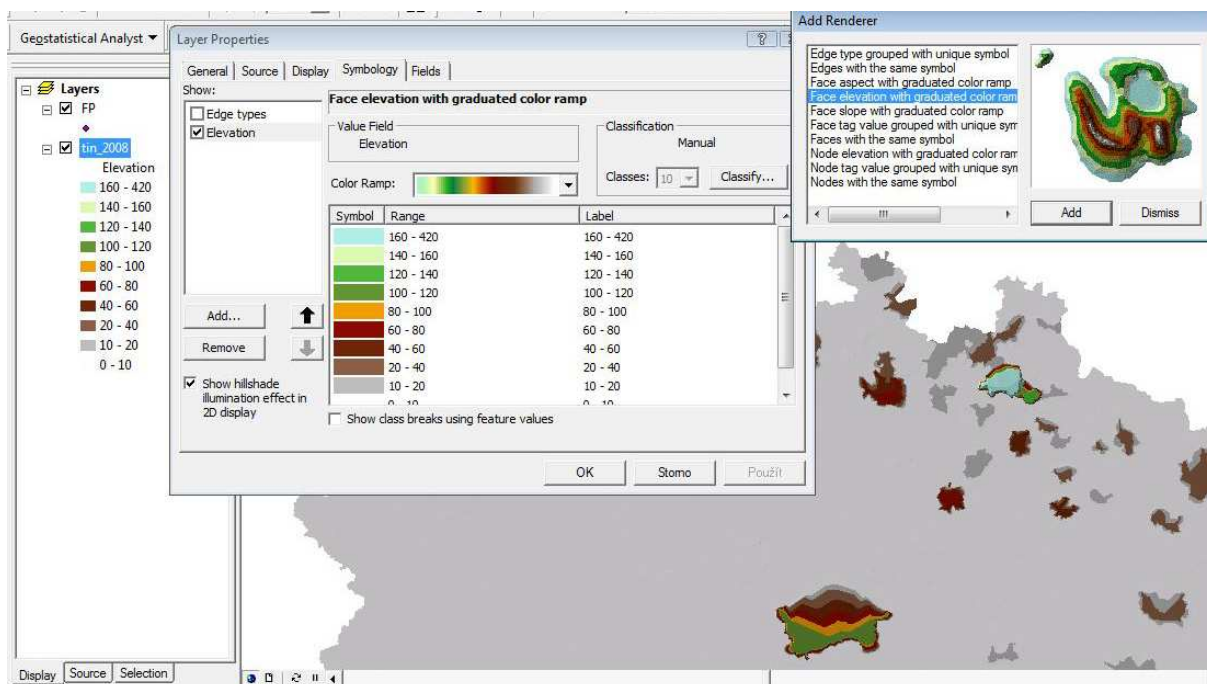
Vstupními daty pro 3D analýzu bude SHP s počty uchazečů o studium z obcí za rok 2008. V programu ArcMap je vybrán nástroj *Create TIN from Features* (obr. 23) s nastavením hodnoty, která bude určovat výšku (počet uchazečů o studium).



Obr. 23: Nastavení pro tvorbu TIN

Nastavení vizualizace

Po vytvoření vrstvy TIN je možné v Properties nastavit pod záložkou Symbology barvy jednotlivým výškovým stupňům (obr. 24).



Obr. 24: Nastavní vizualizace TIN s ukázkou 3D modelu počtu obyvatel

15.5 Index potenciálu obce

Dle Ayada (2007) index potenciálu obce (gravity index) vyjadřuje předpoklad zdrojového území pro nábor nových studentů.

Příprava dat

Do SHP obcí ČR jsou přes nástroj JOIN přidána data počtu uchazečů o studium, počtu obyvatel a vzdáleností, přičemž všechny údaje jsou dány pro každý rok a každou obec. Zároveň je vytvořeno nové pole, které je pojmenováno GI (gravity index) a datovým typem je Double.

Kalkulace gravity indexu

Tyto údaje jsou pak složeny do patřičné závislosti dle obecného vzorce (Ayad, 2007), čímž vznikne výpočet:

$$I_i = \frac{\text{počet uchazečů}_i}{\text{počet obyvatel}_i} \times \text{koeficient vzdálenosti}_i ,$$

kde počet obyvatel_i je dán počtem obyvatel ve věku od 15 do 19 let a koeficient vzdálenosti je určen dle tabulky 6. Hodnota *i* vyjadřuje rok.

Tab. 6: Klasifikace vzdáleností obcí dle Ayada (2007)

ID	Vzdálenost v km	Koeficient
1	0 – 50	10000
2	51 – 100	1000
3	101 – 150	100
4	151 – 200	10
5	201 - 389	1

16 Výstupy a vyhodnocení výsledků analýzy

Za uplynulých 5 let bylo podáno 14 338 přihlášek ke studiu v některém studijním programu na FP TUL. Celkem se hlásilo 11 072 uchazečů, což znamená, že na jeden uchazeč podal 1,23 přihlášky ke studiu. Z toho vyplývá, že každý čtvrtý uchazeč podal více jak jednu přihlášku. Přitom je průměrně přijato ke studiu 50% uchazečů, z čehož je zapsáno 65% (32% z přihlášených k přijímacímu řízení) dle tab. 7.

Kromě uchazečů z ČR se na FP TUL hlásí několik studentů ze zahraničí. Jedná se velice malé počty, nicméně je velice zajímavé odkud tito studenti jsou a kolik jich je (rok 2004 – 13 studentů, rok 2005 – 10 studentů, rok 2006 – 13 studentů, rok 2007 – 9 studentů, rok 2008 – 11 studentů). Uchazeči o studium jsou z Albánie, Etiopie, Kolumbie, Kypru, Polska, Ruska, Slovenska, USA, Německa, Ukrajiny a Vietnamu.

Tab. 7: Srovnání úspěšnosti uchazečů v letech 2004-2008

Rok	Přihlášení	Přijetí		Zapsaní		
		Počet	% z přihlášených	Počet	% z přihlášených	% z přijatých
2004	2037	780	38,3	594	29,2	76,2
2005	3703	1701	45,9	1008	27,2	59,3
2006	2823	1292	45,8	862	30,5	66,7
2007	3177	1765	55,6	1156	36,4	65,5
2008	2598	1516	58,4	991	38,1	65,4
Celkem	14338	7054	49,2	4611	32,2	65,4

Dle tabulky 8 je zjevný vývoj struktury studijních programů FP TUL. Ke standardnímu učitelskému studiu byla přidána řada odborných nečitelských programů. V roce 2004 je uchazečům nabízeno 5 programů, přičemž v roce 2008 jich je 14. Znatelný posun začíná v roce 2006 a 2007. V těchto letech začíná hluboký propad zájmu o Učitelství pro základní školy, které mezi lety 2005 a 2007 přišlo o 87% uchazečů. Tento propad částečně zmírňuje studijní program Specializace v pedagogice, na který se hlásí téměř 1000 žadatelů. V příloze 9 je vizualizováno rozmístění zastoupení jednotlivých studijních programů dle obce bydliště jednotlivých žadatelů o studium.

V období 2004-2006, tedy v době čtyřletých magisterských programů, patří mezi nejžádanější obory Český jazyk, Anglický jazyk a Občanská výchova. Tyto aprobační předmety ještě doplňuje Dějepis, Tělesný výchova a Zeměpis. Po přechodu na strukturované studium, tedy v letech 2007-2008, rovněž vzniká řada nových oborů a tak se mezi nejžádanější studijní obory na FP TUL, kromě osvědčeného Českého jazyka a Anglického jazyka, dostávají Humanitní studia.

Tab. 8: Přehled studijního programu a počtu uchazečů v jednotlivých letech

Studijní program	2004	2005	2006	2007	2008
Aplikované vědy v inženýrství		2	1	2	
Ekonomika a management	188	197	221	247	220
Filologie				63	122
Filozofie				169	131
Fyzika				1	4
Geografie				44	95
Historická studia			72	69	65
Matematika					8
Pedagogika	32	25	23	49	52
Sociální práce	648	726	521	562	412
Specializace v pedagogice				952	705
Speciální pedagogika	280	561	348	383	313
Tělesná výchova a sport				217	166
Učitelství pro základní školy	889	2192	1367	279	147
Vychovatelství			270	140	158

Z tabulky 9 je zřejmé jak velký podíl uchazečů o kombinované studium se hlásí v jednotlivých letech. V roce 2004 se téměř polovina uchazečů hlásila na prezenční studium a druhá polovina na kombinované studium. Ve srovnání s rokem 2008 je počet uchazečů o kombinovanou formu skoro třetinový oproti přihlášeným na prezenční studium.

Tab. 9: Počet uchazečů v prezenčním a kombinovaném studiu v jednotlivých letech

Počet studentů	2004	2005	2006	2007	2008
Prezenční studium	1109	2416	1954	2125	1825
Kombinované studium	928	1287	869	1052	773

Tabulka 10 doplňuje tabulku 9 údaji o studijních programech v kombinované formě studia. Oproti jasné převaze Sociální práce v roce 2004 mají v roce 2008 oba

obory vyrovnávající tendenci. V letech 2007 a 2008 je navíc doplňuje obor učitelství pro základní školy.

Tab. 10: Počet uchazečů o studijní programy v kombinované formě

Studijní program	2004	2005	2006	2007	2008
Sociální práce	648	726	521	562	412
Speciální pedagogika	280	561	348	383	313
Učitelství pro základní školy				107	48

Uchazeči z gymnázií tvoří nejpočetnější skupinu ze všech přihlášených k přijímacímu řízení na FP TUL. Další objemnou skupinou jsou absolventi Středních odborných škol, Obchodních akademií a Středních průmyslových škola (tab. 11). Rozmístění typů středních škol dle počtu uchazečů z jednotlivých okresů je vykresleno v příloze 10.

Tab. 11: Srovnání počtu uchazečů dle typů SŠ

Typ školy	2004	2005	2006	2007	2008
Gymnázium	716	1736	1190	1344	1078
Integrovaná střední škola	30	64	46	61	33
Konzervatoř	2	5	6	8	2
Obchodní akademie	193	391	365	386	358
Střední odborná škola	429	875	783	917	601
Střední odborné učiliště	118	175	118	102	63
Střední průmyslová škola	311	377	265	290	275
Škola MV a MO	29	36	10	14	6
Zahraniční školy	4	12	18	20	5

Stav procentuálního zastoupení věkové skupiny 15-19 v celkové skladbě obyvatelstva zůstává za celých 5 let téměř neměnný. Nejvyšší počet obyvatel této věkové skupiny (5001 a více) má pouze 10 krajských měst. Města a obce s 501-5000 obyvateli ve věku 15-19 let jsou nejvíce soustředěny v Podkrušnohoří, kraji Vysočina a na území mezi Olomoucí a Ostravou (příloha 3).

Interpolovaná data z matriky ilustrují vývoj rozsahu působení FP TUL. Na počátku sledovaného období byla působnost téměř po celé ploše Čech, čím dále od Liberce, tím více soustředěna do větších měst. V roce 2008 ovšem dochází k tomu, že hlavní základna uchazečů o studium je tvořena Libereckým, Středočeským, Královohradeckým, Pardubickým a Ústeckým krajem a Prahou (příloha 4). Tuto mapu doplňuje příloha 8, která zobrazuje rozmístění uchazečů o studium v % z celkového počtu.

Při výpočtu indexu potenciálu obcí došlo k potvrzení toho, že nejvyšší potenciál mají obce blízko k sídlu FP TUL (příloha 5). Čím dále od Liberce obce leží, tím nižší potenciál mají.

Zajímavý výsledek vytvořila síťová analýza, kdy se určovali servis areas, tedy území, kde nejvíce působí některá pedagogická fakulta. FP TUL působí téměř výhradně na území Libereckého kraje, což doplňuje druhá mapa, pouze Libereckého kraje, v příloze 6.

Analýza dostupnosti do Liberce po silniční síti ukazuje území, která jsou dostupná v určité vzdálenosti a čase. Ačkoliv je analýza zkrácena absencí zpřesnění silniční sítě o intravilán obce, tedy území obce, kde je průměrná rychlost nižší.

17 Diskuze

Základní otázkou je, zda lze pracovat s daty školní matriky pracovat efektivně. Prozatím byla data matriky využívána jako podklad pro statistické výpočty, které jsou součástí každé výroční zprávy, či zpráv reflektující činnost. Využití, které navrhuje tato diplomová práce, je efektivnější, protože z databáze využije prostorové složky záznamů (adresa bydliště, adresa absolvovaného nižšího stupně školy), čímž přináší nové možnosti prostorové vizualizace výsledků.

Pokud má být matriční databáze efektivně využita k prostorovým analýzám, je potřeba překonat některá úskalí práce s geodatabází. Důležitým požadavkem je svědomité vedení a údržba dat v matrice. Konkrétně se jedná o vedení záznamů, které jsou prostorově lokalizovatelné, tedy přesnou adresu bydliště (nižšího stupně školy) uchazeče o studium (tento klíč musí být vždy jedinečný). Je nutné využít jedinečných klíčů pro identifikaci jednotlivých tabulek a propojení funkčními relacemi. S tím souvisí využití GIS. Tedy využití nástroje, který je vhodný pro práci s prostorovými daty. Nejen že má implementovány nástroje pro práci s databázemi, rovněž obsahuje metody pro analýzu těchto dat a jejich vizualizaci. Je tedy vhodným prostředkem pro práci s daty matriky.

Předloženou analýzu akademického marketingu je možné aplikovat na kterýkoliv stupeň školy. Všechny školy jsou ekonomické subjekty, poskytující produkty, a mají zájem o žáky a studenty, jejich zákazníky. Každá škola musí vést svou matriku, čímž je dána možnost jí využít pro prostorovou vizualizaci, která může pomoci přípravě marketingu školy.

Hlavním doporučením, které vyplývá ze získaných map, je udržet propagaci FP TUL na regionální úrovni. Platí zde přímá úměrnost, kdy působení blíže k sídlu fakulty je širší a zasahuje i do malých obcí, avšak při větší vzdálenosti se působení omezuje na větší města. Vše závisí i na časové dostupnosti obcí. Avšak nejvíce přitažlivé obce s potenciálními uchazeči jsou v blízkém okolí sídla FP TUL. Lze konstatovat, že stěžejními regiony, ve kterých je přínosné široce působit na uchazeče o studium, jsou Liberecký, Ústecký a Královehradecký kraj, také sever Středočeského kraje a Praha.

18 Závěr

Tato diplomová práce ukazuje možnosti práce s daty, uloženými v matrice FP TUL. Data byla vždy využívána jako statistický soubor, nikoliv však jako nástroj pomoci při náboru nových studentů. Po vyřešení všech problémů se spojením s prostorovými daty lze tvořit plnohodnotné analýzy za pomoci nástroje GIS. Na základě tohoto tvrzení je splněn dílčí cíl, potvrzující, že jsou data v matrice skutečně využitelná pro prostorové analýzy a obsahují tedy prostorové údaje.

V matrice je uložena velká spousta informací, je jen potřeba pojmenovat, jaký výběr bude pro práci uchopitelný a vypovídající. Tím byl splněn první cíl práce, popsat strukturu databáze uchazečů o studium a studentů TUL vedené matrikou TUL. V této práci jsou uvedeny mnohé poznatky, které vznikly během jednotlivých fází výzkumu, a které napomohou k dalšímu využití celé práce při ekvivalentních analýzách v budoucnu, což splňuje dílčí cíl o nalezení metodického postupu pro opakované využití. Rovněž

Zvolené softwarové prostředky jsou plně vyhovující pro tvorbu všech aspektů aplikovaných analýz, avšak při tvorbě složitějších, či detailnějších, výpočtů je potřeba zvážit užití hardwarové prostředky.

V rámci práce vznikl soubor tématických map, které pomáhají reprezentovat výsledky jednotlivých analýz, čímž byl splněn cíl práce vytvořit soubor grafických prezentací a jejich interpretací. Zároveň vyjadřují postavení FP TUL a rozsah jejího vlivu. Lze tedy považovat tuto diplomovou práci za jiný úhel pohledu na marketing fakulty. Může přinést nové náměty a postupy pro nábor nových studentů ke studiu a zviditelnění trendů, jež se odráží v zájmu o jednotlivé studijní programy a obory. Navrhuje možná využití položek databáze pro prostorové analýzy v prostředí GIS. V závěrečné diskuzi je nastíněno i další využití analýzy i na základních a středních školách.

19 Užité zkratky

ArcČR 500	Digitální vektorová geografická databáze pro území ČR v měřítku 1:500 000
ČSÚ	Český statistický úřad
DB	Databáze
DBS	Databázový systém
DP	Diplomová práce
ESRI	Environmental Systems Research Institute, Inc.
FP	Fakulta přírodovědně-humanitní a pedagogická
GIS	Geografické informační systémy
IMS	Internetová mapová služba
KGE	Katedra geografie
LAU	Místní samosprávné jednotky – klasifikace obcí a okresů
MO	Ministerstvo obrany
MŠMT	Ministerstvo školství, mládeže a tělovýchovy
MV	Ministerstvo vnitra
NUTS	Normalizovaná klasifikace územních celků ČR
ORP	Obec s rozšířenou působností
SLDB	Sčítání lidu, domů a bytů
SŘDB	System řízení báze dat
SŠ	Střední škola
TUL	Technická univerzita v Liberci
VŠ	Vysoká škola
WMS	Web map service

20 Seznam obrázků

Obr. 1: Složky GIS	23
Obr. 2: Hlavní skupiny softwarových modulů GIS (zpracováno dle Tučka, 1998)	24
Obr. 3: Ukázka relační databáze	26
Obr. 4: Datová posloupnost (vytvořeno dle Tučka ,1998).....	30
Obr. 5: Rozdíl mezi rastrem a vektorem (dostupné z http://www.stargen.cz/images/rozdil-vektor-rastr-full.jpg)	32
Obr. 6: Struktura prvotních dat z matricy FP TUL.....	36
Obr. 7: Struktura kompletní databáze	49
Obr. 8: Atributová tabulka SHP po připojení XLS nebo DBF	53
Obr. 9: Nově vyexportovaný SHP	54
Obr. 10: Nastavení nástroje Features to Raster	56
Obr. 11: Nastavení nástroje Cost Weighted	56
Obr. 12: Nastavení nástroje Reclassify.....	57
Obr. 13: Nastavení nástroje Raster to Polyline.....	57
Obr. 14: Ohodnocené komunikace v obcích Kolín a ukázka atributové tabulky	58
Obr.15: Modelové zobrazení prostorové analýzy	58
Obr. 16: Modelové zobrazení síťové analýzy	59
Obr. 17: Spuštění vyhodnocení servisních oblastí.....	60
Obr. 18: Přiřazení sídel fakult jako servisních středisek	60
Obr. 19: Nastavení zobrazení servisní oblasti	61
Obr. 20: Geostatistical wizard - volba metody	62
Obr. 21: Geostatistical wizard - styl modelování	63
Obr. 22: Geostatistical wizard - hledání sousedů	64
Obr. 23: Nastavení pro tvorbu TIN.....	65
Obr. 24: Nastavní vizualizace TIN s ukázkou 3D modelu počtu obyvatel.....	65

21 Zdroje dat

- AYAD, Y. M. *Challenges in student recruitment for educational institutions: Materials and methods: Case study of Clarion University of PA: Analysis of 2004-2005 school year data*. ESRI User conference 2007.
- BATKO, M.: *Relační vs. objektově-relační vs. objektové databáze* [online]. 1997 [cit. 11.12.2008]. <<http://www.fi.muni.cz/~xbatko/oracle/compare.html>>.
- BERKA, P.: *Dobývání znalostí z databází*. 1. vyd. Praha: Academia, 2003. ISBN 80-200-1062-9.
- CHRISTIE, R.; FERRIS, M. *Spatial analysis for enrollment planing in higher education*. ESRI Education User Conference Proceedings 2004, San Diego, CA.
- CRONER, C. M.; SPERLING, J.; BROOME, F. R.: Geographic information systems (GIS): New perspective in understanding human health and enviromental relationships. In: *Statistics in medicine*. Ročník 15. US, 1996, s. 15 – 30.
- GILFILLAN, I.: *Introduction to Relational Databases* [online]. 24.6.2002, 24.6.2002 [cit. 11.12.2008]. <<http://www.databasejournal.com/sql/etc/article.php/1469521/Introduction-to-Relational-Databases.htm>>.
- GLACOVÁ, M; HORÁK, J.: Analýza situace trhu práce v měřítku okresního města. In Sbor. ref. conference GIS Seč 2002, Seč, 12.-14.6.2002, 2002, s. 12.
- JOHNSTON, K. a kol.: *Geostatistical Analyst Tutorial* [online]. 2001 [cit. 5.1.2009]. <http://webhelp.esri.com/arcgisdesktop/9.2/pdf/Geostatistical_Analyst_Tutorial.pdf>.
- LANGR, J. *T-MAPY spol. s r.o. Hradec Králové – Budování geodatabáze* [online]. 1999 [cit. 11.12.2008]. <http://www.tmapy.cz/public/tmapy/cz/_aktualne/_clanky/budovani_geodatabaze.html>.
- MACDONALD, L. a kol.: *3D Analyst Tutorial* [online]. 2002 [cit. 5.1.2009]. <http://webhelp.esri.com/arcgisdesktop/9.2/pdf/3D_Analyst_Tutorial.pdf>.
- MACHALOVÁ, J.: *Prostorově orientované systémy pro podporu manažerského rozhodování*. 1. vyd. Praha: C. H. Beck, 2007. 141 s. ISBN 978-80-7179-463-9.

- MANDLOI, D. a kol.: *Network Analyst Tutorial* [online]. 2008 [cit. 5.1.2009]. <http://webhelp.esri.com/arcgisdesktop/9.2/pdf/Network_Analyst_Tutorial.pdf>.
- MARBLE, D. F.; MORA, V. J.; HERRIES, J. P. *Applying GIS technology to the freshman admissions process at a large university*. ESRI User Conference. 22 – 25. 5. 1995. Palm Springs, CA, USA. Dostupné z: <<http://www.isprs.org/commission4/proceedings02/pdfpapers/553.pdf>>.
- PÍSEK, S. *Mircosoft Access 2007 – podrobný průvodce*. 1. vydání. Praha: GRADA, 2007. 200 s. ISBN 978-80-247-1967-2.
- RAPANT, P. *Úvod do geografických informačních systémů*. Skripta PGS. 1. vydání. Ostrava: VŠB, 2002.
- RIORDAN, R.M.: *Vytváříme relační databázové aplikace*. 1. vyd. Praha: Computer Press, 2000. 280 s. ISBN 80-7226-360-9.
- SHEPHERD, J.C.: *Database Management: Theory and Application*. 1. vyd. Homewood, Illinois: Richard D. Irwin, Inc., 1990. 781 s. ISBN 0-256-07829-7.
- ŠÍMA, J.: *Geoinformační terminologie pro geodety a kartografy*. 1. vyd. Zdíby: VÚGTK, 2003. 87 s. ISBN 80-85881-20-9.
- TANG, H.; MCDONALD, S. *Intergrating GIS and spatial data mining technique for target marketing of university courses*. Symposium on Geospatial Theory, Processing and Applications. Ottawa, 2002.
- TUČEK, J.: *Geografické informační systémy – Principy a praxe*. 1. vydání. Praha: Computer Press, 1998. 424 s. ISBN 80-7226-091-X.
- VEBLOUD.: *Teorie relačních databází: Relační model dat* [online]. 12.1.2006 [cit. 11.12.2008]. <<http://www.manualy.net/article.php?articleID=9>>.
- WILISON, J. : *Spatial Analyst Tutorial* [online]. 2001 [cit. 5.1.2009]. <http://webhelp.esri.com/arcgisdesktop/9.2/pdf/Spatial_Analyst_Tutorial.pdf>.
- ZAJÍCOVÁ, Z.: *Návrh zpracování dat SCIO v prostředí GIS*. In *GISáček 2006*. Ročník 2006. VSB Ostrava, 2006.

Zdroje dat:

Databáze počtu obyvatel v obcích pro roky 2004-2008 [databáze]. Liberec: Český statistický úřad, 2008 [cit. 1.1.2009]. <<http://www.czso.cz>>. Demografická ročenka ČR.

Databáze uchazečů o studium na FP TUL v letech 2004-2008 [databáze]. Liberec:
TUL, 2003 [cit. 1.1.2009]. <<http://stag.tul.cz>>. Matrika FP TUL.
ArcČR 500 [databáze]. Praha: Arcdata a ČÚZK.

Přílohy k diplomové práci Z. Prokopa
Analýza dat matriky TUL metodami GIS

Seznam příloh

Příloha 1: Tabulka studijních programů FP TUL

Příloha 2: Přehled oborů studia FP TUL

Příloha 3: Vývoj rozmístění obyvatelstva v ČR v letech 2004-2008

Příloha 4: Počet uchazečů o studium v letech 2005-2008

Příloha 5: Index přitažlivosti obcí ČR pro rok 2009

Příloha 6: Spádová oblast pedagogických fakult v roce 2009 dle silniční sítě

Příloha 7: Dostupnost do Liberce po silniční síti v roce 2009

Příloha 8: Rozmístění uchazečů o studium v letech 2004-2008

Příloha 9: Podíl typů středních škol v krajích ČR v letech 2004 a 2008

Příloha 10: CD

Příloha 1: Tabulka studijních programů FP TUL

PROGRAM_KOD	FORMA	NAZEV	KOD_PROG
B1101	P	Matematika	B1101P
B1103	P	Aplikovaná matematika	B1103P
B1301	P	Geografie	B1301P
B1701	P	Fyzika	B1701P
B6101	P	Filozofie	B6101P
B6208	P	Ekonomika a management	B6208P
B7106	P	Historická studia	B7106P
B7310	P	Filologie	B7310P
B7401	P	Tělesná výchova a sport	B7401P
B7501	P	Pedagogika	B7501P
B7501	K	Pedagogika	B7501K
B7502	K	Sociální péče	B7502K
B7502	P	Sociální péče	B7502P
B7505	P	Vychovatelství	B7505P
B7505	K	Vychovatelství	B7505K
B7506	K	Speciální pedagogika	B7506K
B7507	P	Specializace v pedagogice	B7507P
B7508	K	Sociální práce	B7508K
B7531	K	Předškolní a mimoškolní pedagogika	B7531K
DVPP	D	Další vzdělávání pedagogických pracovníků	DVPPD
M7503	K	Učitelství pro základní školy	M7503K
M7503	P	Učitelství pro základní školy	M7503P
M7504	P	Učitelství pro střední školy	M7504P
N1701	P	Fyzika	N1701P
P1103	P	Aplikovaná matematika	P1103P
P1103	K	Aplikovaná matematika	P1103K
P3901	K	Aplikované vědy v inženýrství	P3901K
P3901	P	Aplikované vědy v inženýrství	P3901P

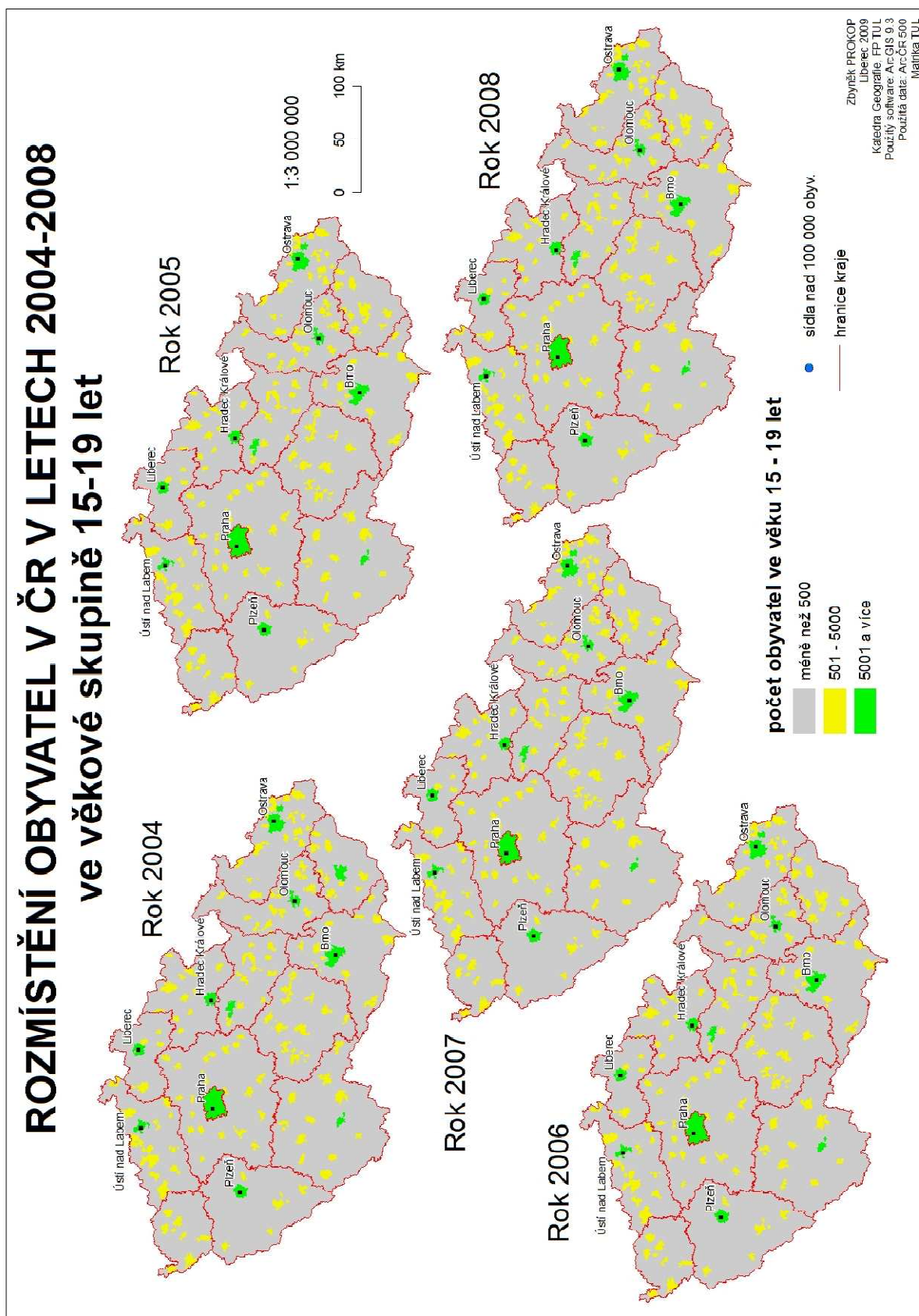
Příloha 2: Přehled oborů studia FP TUL

PROGRAM_K OD	FORM A	CISOBORU	ZKRATKA	NAZEV	KOD_OBO R
B1101	P	1101R016	MA	Matematika	B1101P
B1103	P	1103R022	PM	Pojistná matematika	B1103P
B1301	P	1301R022	AG	Aplikovaná geografie	B1301P
B1701	P	1702R001	AF	Aplikovaná fyzika	B1701P
B6101	P	6101R026	FHV	Filozofie humanitních věd	B6101P
B6208	P	6208R048	MG	Management sportovní	B6208P
B7106	P	7105R062	KM	Kulturněhistorická a muzeologická studia	B7106P
B7310	P	7310R033	FIL	Český jazyk a literatura	B7310P
B7401	P	1802R023	IF	Informatika se zaměřením na vzdělávání	B7401P
B7401	P	6107R023	HU	Humanitní studia se zaměřením na vzdělávání	B7401P
B7401	P	7105R056	DJ	Historie se zaměřením na vzdělávání	B7401P
B7401	P	7401R014	TV	Tělesná výchova se zaměřením na vzdělávání	B7401P
B7401	P	7504R006	FY	Fyzika se zaměřením na vzdělávání	B7401P
B7401	P	7504R009	CH	Chemie se zaměřením na vzdělávání	B7401P
B7401	P	7504R015	MA	Matematika se zaměřením na vzdělávání	B7401P
B7401	P	7504R180	ZE	Geografie se zaměřením na vzdělávání	B7401P
B7401	P	7504R269	CJ	Český jazyk se zaměřením na vzdělávání	B7401P
B7401	P	7504R300	SJ	Španělský jazyk se zaměřením na vzdělávání	B7401P
B7401	P	7507R	SP	Společný základ se zaměřením na vzdělávání	B7401P
B7401	P	7507R036	AJ	Anglický jazyk se zaměřením na vzdělávání	B7401P
B7401	P	7507R041	NJ	Německý jazyk se zaměřením na vzdělávání	B7401P
B7501	K	7505R004	PVC-K	Pedagogika volného času	B7501K
B7501	P	7501R030	PPZ	Pedagogicko-psychologická způsobilost	B7501P
B7501	P	7504R016	DPS	Souběžné doplňkové pedagogické studium	B7501P
B7502	K	7502R010	PG	Sociální pedagog	B7502K
B7502	K	7502R022	PR	Sociální pracovník	B7502K
B7502	P	7502R010	PG	Socialní pedagog	B7502P
B7502	P	7502R022	PR	Socialní pracovník	B7502P
B7505	K	7505R004	PVC-K	Pedagogika volného času	B7505K
B7505	P	7505R004	PVC	Pedagogika volného času	B7505P
B7506	K	7506R012	SPD	Speciální pedagogika předškolního věku	B7506K
B7506	K	7506R029	SPV	Speciální pedagogika pro vychovatele	B7506K
B7507	P	1802R023	INF	Informatika se zaměřením na vzdělávání	B7507P
B7507	P	6107R023	HU	Humanitní studia se zaměřením na vzdělávání	B7507P
B7507	P	7105R056	DJ	Historie se zaměřením na vzdělávání	B7507P
B7507	P	7504R006	FY	Fyzika se zaměřením na vzdělávání	B7507P
B7507	P	7504R009	CH	Chemie se zaměřením na vzdělávání	B7507P
B7507	P	7504R015	MA	Matematika se zaměřením na vzdělávání	B7507P
B7507	P	7504R180	ZE	Geografie se zaměřením na vzdělávání	B7507P
B7507	P	7504R269	CJ	Český jazyk se zaměřením na vzdělávání	B7507P
B7507	P	7504R300	SJ	Španělský jazyk se zaměřením na vzdělávání	B7507P
B7507	P	7507R	PS	Společný základ se zaměřením na vzdělávání	B7507P

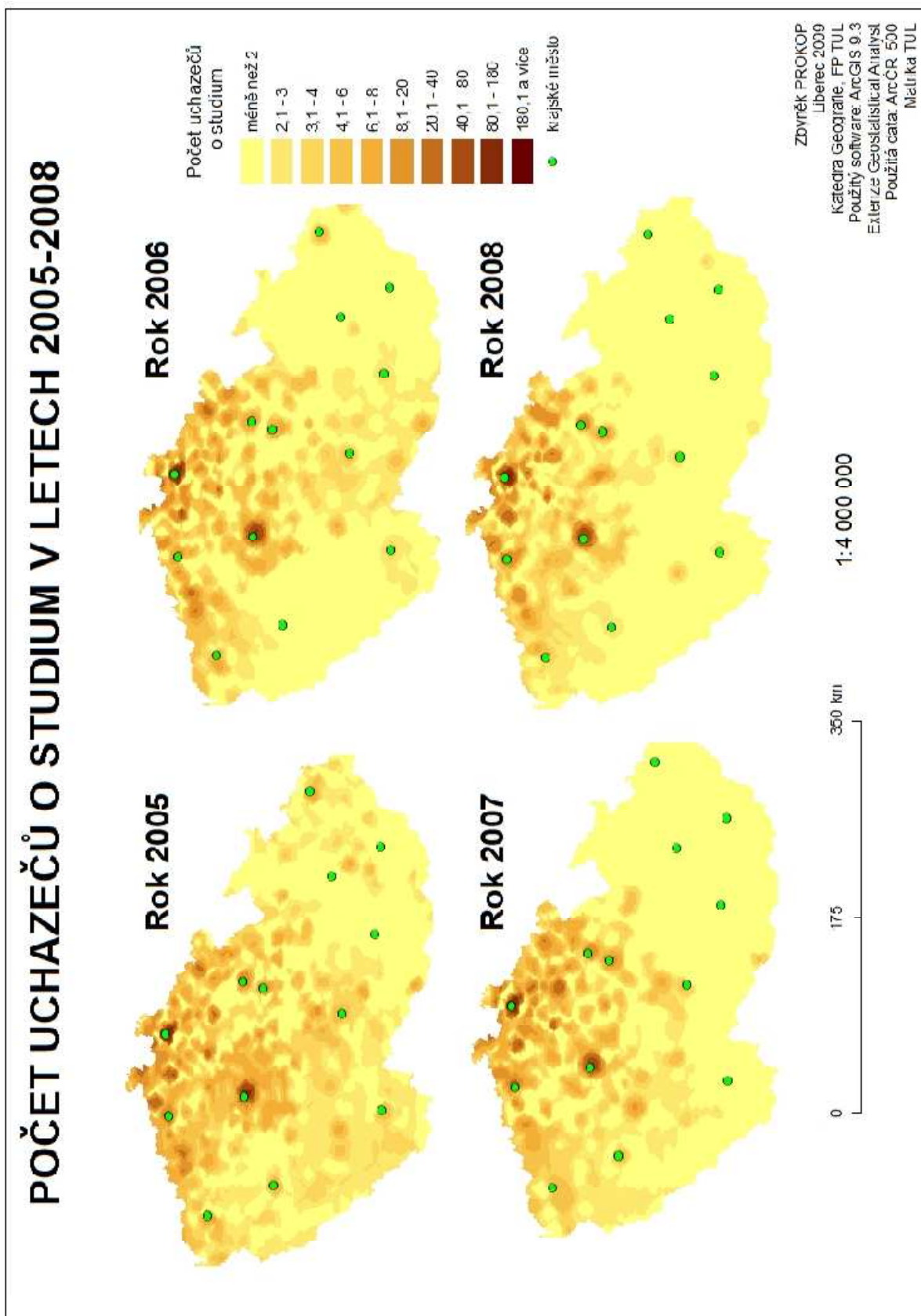
B7507	P	7507R036	AJ	Anglický jazyk se zaměřením na vzdělávání	B7507P
B7507	P	7507R041	NJ	Německý jazyk se zaměřením na vzdělávání	B7507P
B7508	K	7502R022	SOP	Sociální pracovník	B7508K
B7508	K	7502R023	PEP	Penitenciární péče	B7508K
B7531	K	7506R010	MS	Speciální pedagogika pro učitele MŠ a výchovné pracovníky	B7531K
M7503	K	7503T	PS	Profesní studium pro 2. stupeň základní školy	M7503K
M7503	K	7503T047	SPUCH	Prohloubený st.program - SPUCH	M7503K
M7503	K	7503T047	NS	Učitelství pro 1.stupeň základní školy	M7503K
M7503	K	7503T047	NS-AJ	Prohloubený studijní program - Angličtina	M7503K
M7503	K	7503T047	NS-ČJ	Prohloubený studijní program - Český jazyk	M7503K
M7503	K	7503T047	NS-MA	Prohloubený studijní program - Matematika	M7503K
M7503	K	7503T047	NS-HV	Prohloubený studijní program - Hudební výchova	M7503K
M7503	K	7503T047	NS-VV	Prohloubený studijní program - Výtvarná výchova	M7503K
M7503	K	7503T047	NS-TV	Prohloubený studijní program - Tělesná výchova	M7503K
M7503	K	7503T047	NS-SPUCH	Prohloubený studijní program - Specifické poruchy učení a chování	M7503K
M7503	K	7503T047	NS-NV	Prohloubený studijní program - Náboženská výchova	M7503K
M7503	K	7503T047	1ST	Učitelství pro 1. stupeň základní školy - kombinované studium	M7503K
M7503	K	7503T047	1ST-ČJ	Prohloubený st.program - Český jazyk	M7503K
M7503	K	7503T047	1ST-VV	Prohloubený st.program - Výtvarná výchova	M7503K
M7503	K	7503T047	1ST-HV	Prohloubený st.program - Hudební výchova	M7503K
M7503	K	7503T047	1ST-MA	Prohloubený st.program - Matematika	M7503K
M7503	K	7503T047	1ST-NJ	Prohloubený st.program - Němčina	M7503K
M7503	K	7503T047	1ST-AJ	Prohloubený st.pogram - Angličtina	M7503K
M7503	P	7503T	PS	Profesní studium pro 2. stupeň základní školy	M7503P
M7503	P	7503T009	AJ	Učitelství anglického jazyka pro 2. stupeň základní školy	M7503P
M7503	P	7503T021	CJ	Učitelství českého jazyka pro 2. stupeň základní školy	M7503P
M7503	P	7503T023	DJ	Učitelství dějepisu pro 2. stupeň základní školy	M7503P
M7503	P	7503T026	FJ	Učitelství francouzského jazyka pro 2. stupeň základní školy	M7503P
M7503	P	7503T028	FY	Učitelství fyziky pro 2. stupeň základní školy	M7503P
M7503	P	7503T030	ZE	Učitelství geografie pro 2. stupeň základní školy	M7503P
M7503	P	7503T036	CH	Učitelství chemie pro 2. stupeň základní školy	M7503P
M7503	P	7503T038	INF	Učitelství informatiky pro 2. stupeň základní školy	M7503P
M7503	P	7503T039	MA	Učitelství matematiky pro 2. stupeň základní školy	M7503P
M7503	P	7503T043	NJ	Učitelství německého jazyka pro 2. stupeň základní školy	M7503P
M7503	P	7503T045	OV	Učitelství občanské výchovy pro 2. stupeň základní školy	M7503P
M7503	P	7503T047	1ST-HV	Prohloubený studijní program - Hudební výchova	M7503P
M7503	P	7503T047	NS-TV	Prohloubený studijní program - Tělesná výchova	M7503P
M7503	P	7503T047	1ST-TV	Prohloubený studijní program-tělesná výchova	M7503P
M7503	P	7503T047	1STK	Učitelství pro 1. stupeň základní školy	M7503P
M7503	P	7503T047	1ST-NV	Prohloubený studijní program - náboženská výchova	M7503P
M7503	P	7503T047	NS-SPUCH	Prohloubený studijní program - Specifické poruchy učení a chování	M7503P
M7503	P	7503T047	1ST	Učitelství pro 1. stupeň základní školy	M7503P

M7503	P	7503T047	NS	Učitelství pro 1. stupeň základní školy	M7503P
M7503	P	7503T047	NS-VV	Prohloubený studijní program - Výtvarná výchova	M7503P
M7503	P	7503T047	1ST-VV	Prohloubený studijní program - Výtvarná výchova	M7503P
M7503	P	7503T047	1ST-NJ	Prohloubený studijní program - Němčina	M7503P
M7503	P	7503T047	1ST-MA	Prohloubený studijní program - Matematika	M7503P
M7503	P	7503T047	SPUCH	Prohloubený studijní program - Specifické poruchy učení a chování	M7503P
M7503	P	7503T047	NS-ČJ	Prohloubený studijní program - Český jazyk	M7503P
M7503	P	7503T047	NS-AJ	Prohloubený studijní program - Angličtina	M7503P
M7503	P	7503T047	NS-MA	Prohloubený studijní program - Matematika	M7503P
M7503	P	7503T047	NS-HV	Prohloubený studijní program - Hudební výchova	M7503P
M7503	P	7503T047	NS-NV	Prohloubený studijní program - Náboženská výchova	M7503P
M7503	P	7503T047	1ST-AJ	Prohloubený studijní program - Angličtina	M7503P
M7503	P	7503T047	1ST-ČJ	Prohloubený studijní program - Český jazyk	M7503P
M7503	P	7503T100	TV	Učitelství tělesné výchovy pro 2. stupeň základní školy	M7503P
M7504	P	7503T009	Aj	Učitelství anglického jazyka pro 2. stupeň základní školy	M7504P
M7504	P	7503T021	Cj	Učitelství českého jazyka pro 2. stupeň základní školy	M7504P
M7504	P	7503T023	Dj	Učitelství dějepisu pro 2. stupeň základní školy	M7504P
M7504	P	7503T028	F	Učitelství fyziky pro 2. stupeň základní školy	M7504P
M7504	P	7503T030	Ze	Učitelství geografie pro 2. stupeň základní školy	M7504P
M7504	P	7503T036	Ch	Učitelství chemie pro 2. stupeň základní školy	M7504P
M7504	P	7503T038	I	Učitelství informatiky pro 2. stupeň	M7504P
M7504	P	7503T039	M	Učitelství matematiky pro 2. stupeň základní školy	M7504P
M7504	P	7503T043	Nj	Učitelství německého jazyka pro 2. stupeň základní školy	M7504P
M7504	P	7503T045	Ov	Učitelství občanské výchovy pro 2. stupeň základní školy	M7504P
M7504	P	7503T098	Tv	Učitelství tělesné výchovy pro 2. stupeň základní školy	M7504P
M7504	P	7504T	Ps	Profesní studium pro střední školy	M7504P
M7504	P	7504T055	Fy	Učitelství fyziky pro střední školy	M7504P
M7504	P	7504T077	Inf	Učitelství informatiky pro střední školy	M7504P
M7504	P	7504T089	Ma	Učitelství matematiky pro střední školy	M7504P
N1701	P	1702T001	AF	Aplikovaná fyzika	N1701P
P1103	K	1103V035	AM_K	Matematické modely a jejich aplikace	P1103K
P1103	P	1103V035	AM	Matematické modely a jejich aplikace	P1103P
P3901	K	3901V012	FI_K	Fyzikální inženýrství	P3901K
P3901	P	3901V012	FI	Fyzikální inženýrství	P3901P
P3901	P	3901V012	FI_K	Fyzikální inženýrství	P3901P

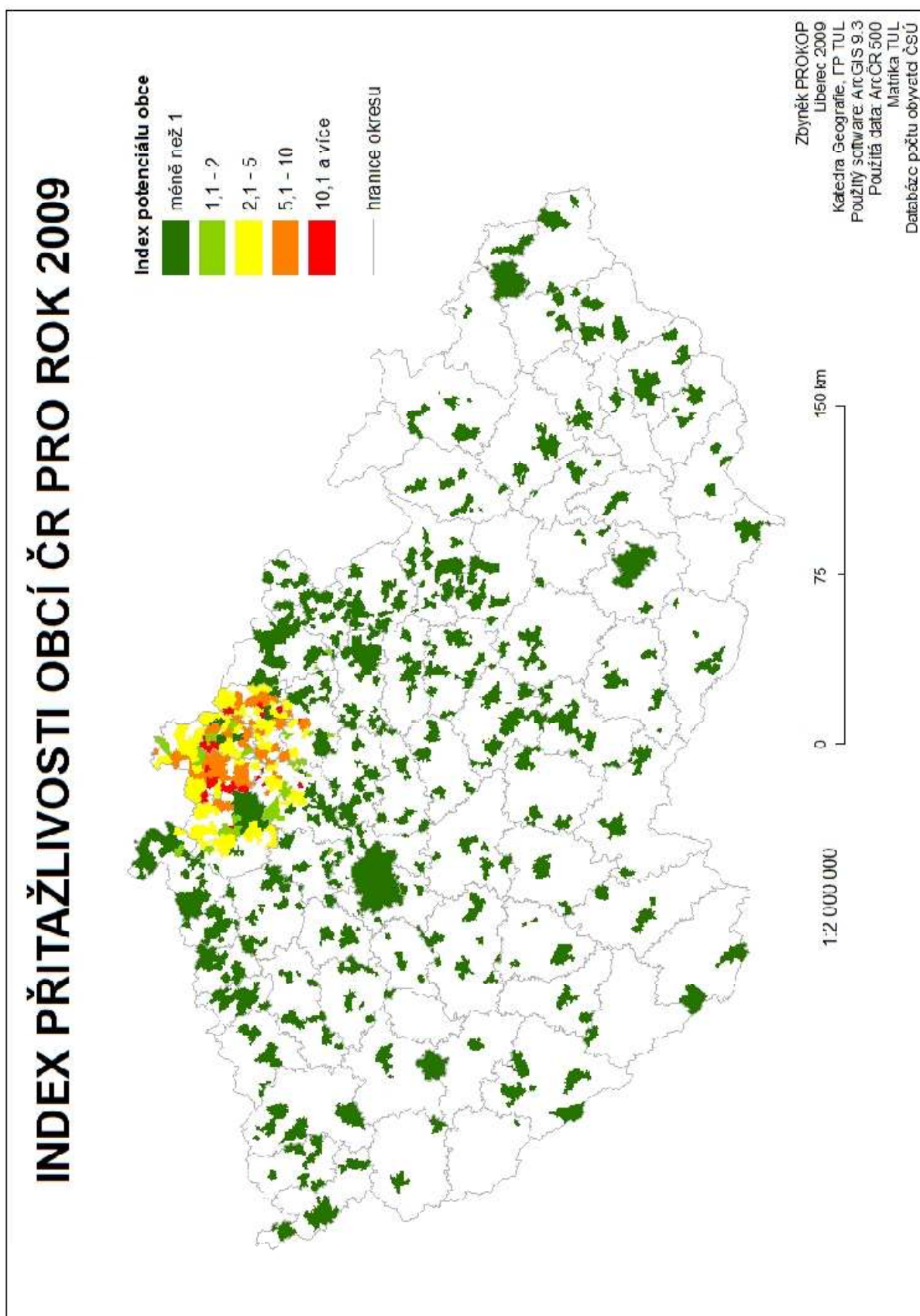
Příloha 3: Vývoj rozmístění obyvatelstva v ČR v letech 2004-2008



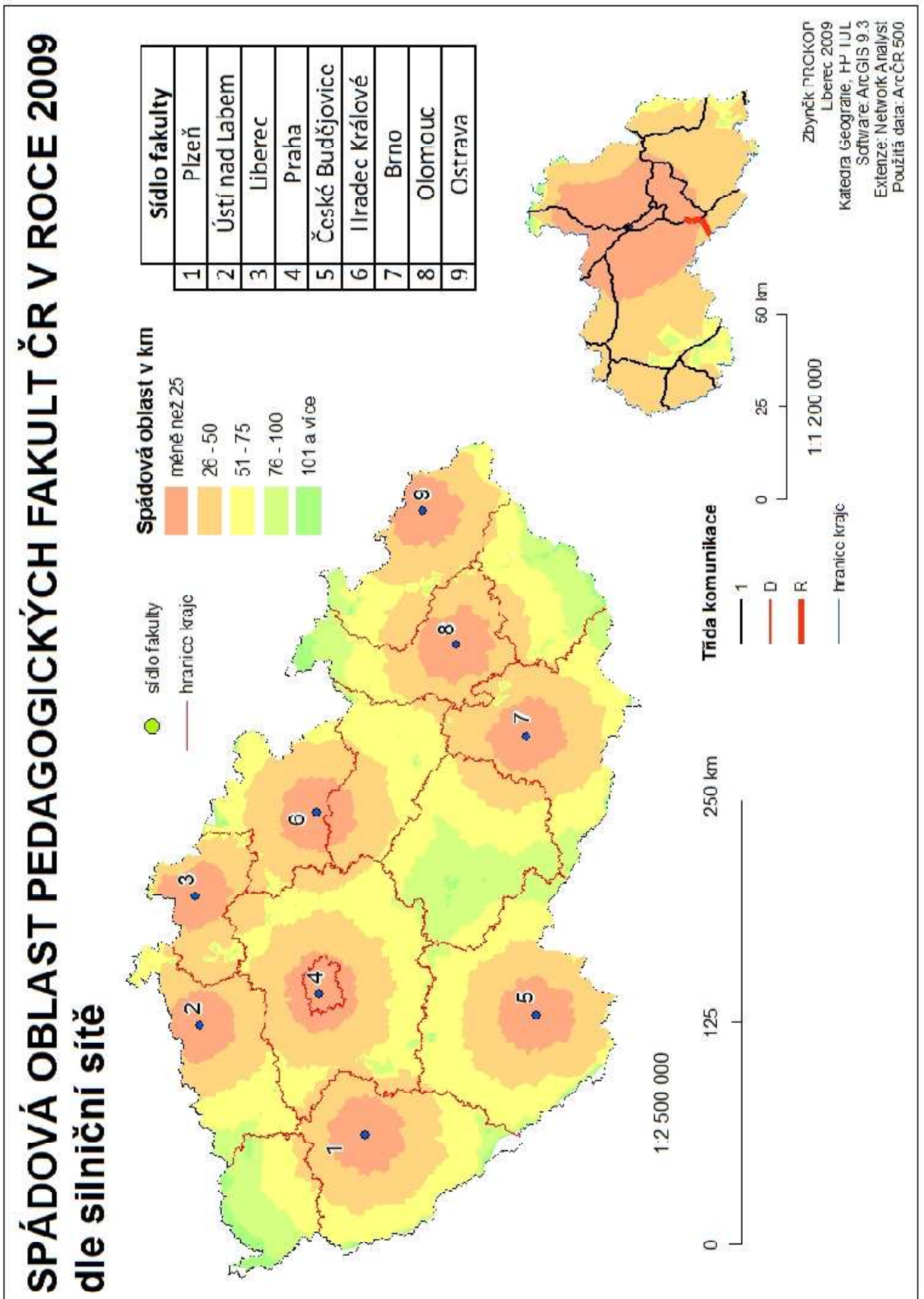
Příloha 4: Počet uchazečů o studium v letech 2005-2008



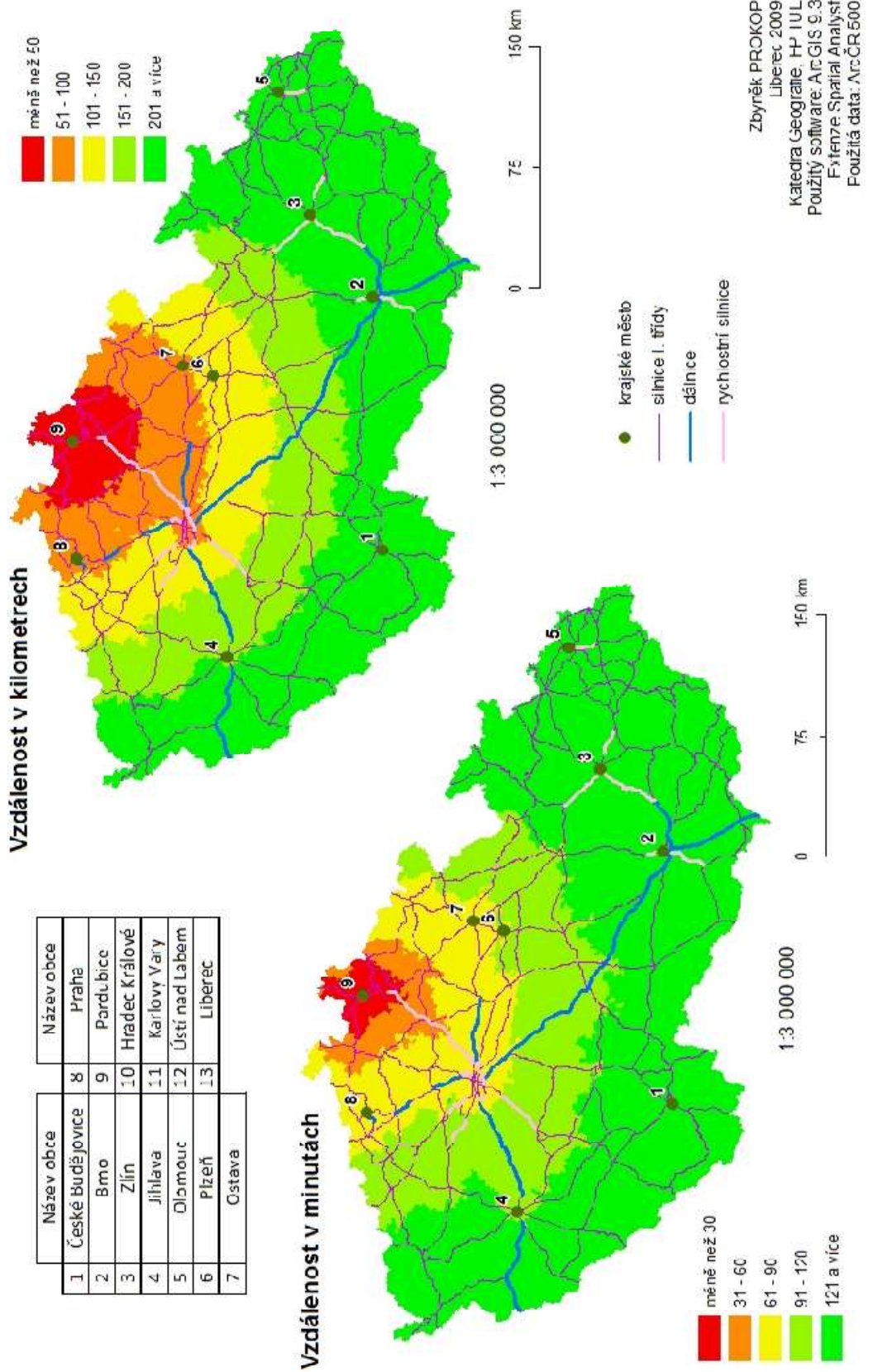
Příloha 5: Index přitažlivosti obcí ČR pro rok 2009



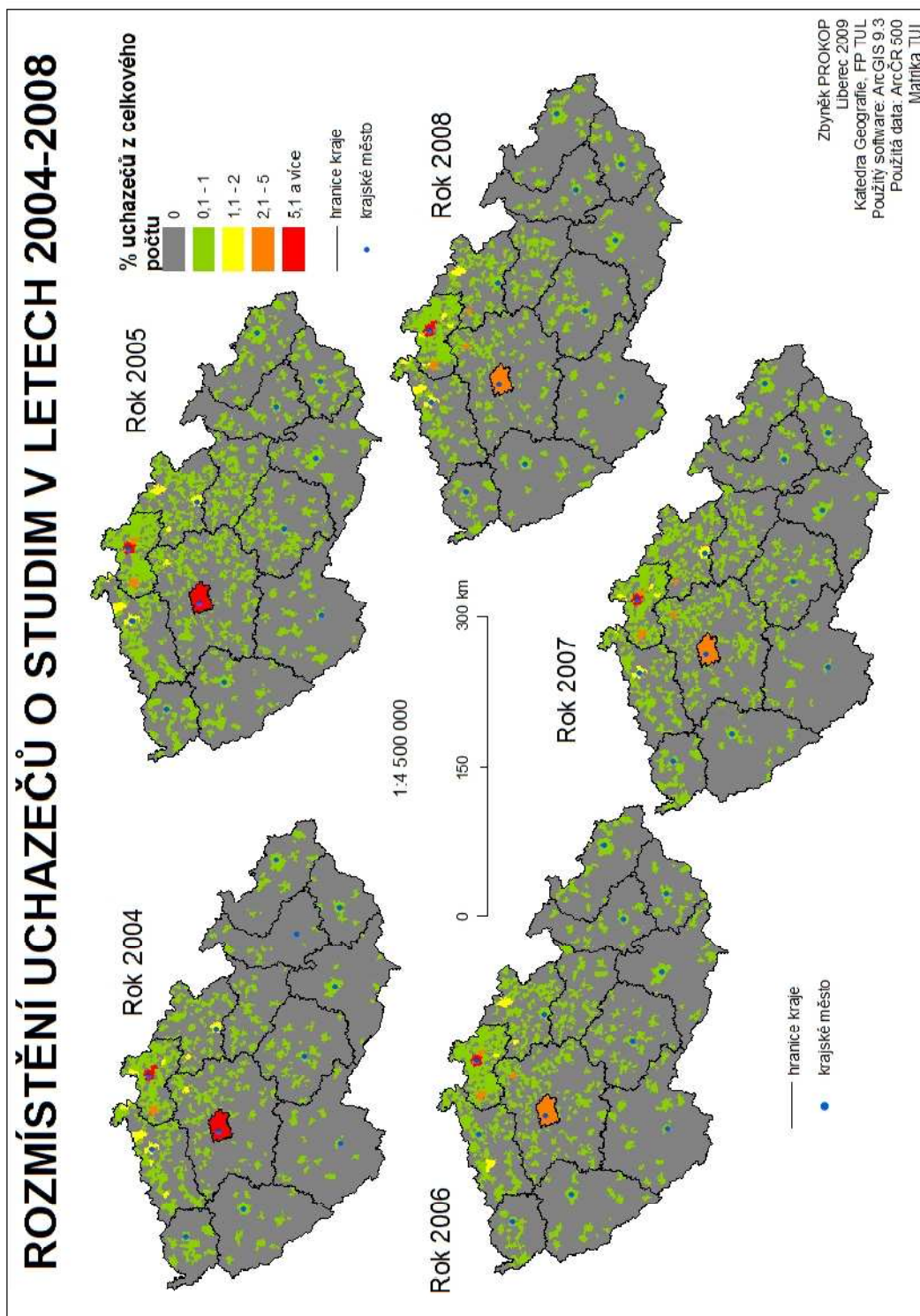
Příloha 6: Spádová oblast pedagogických fakult v roce 2009 dle silniční sítě



DOSTUPNOST DO LIBERCE PO SILNIČNÍ SÍTI V ROCE 2009

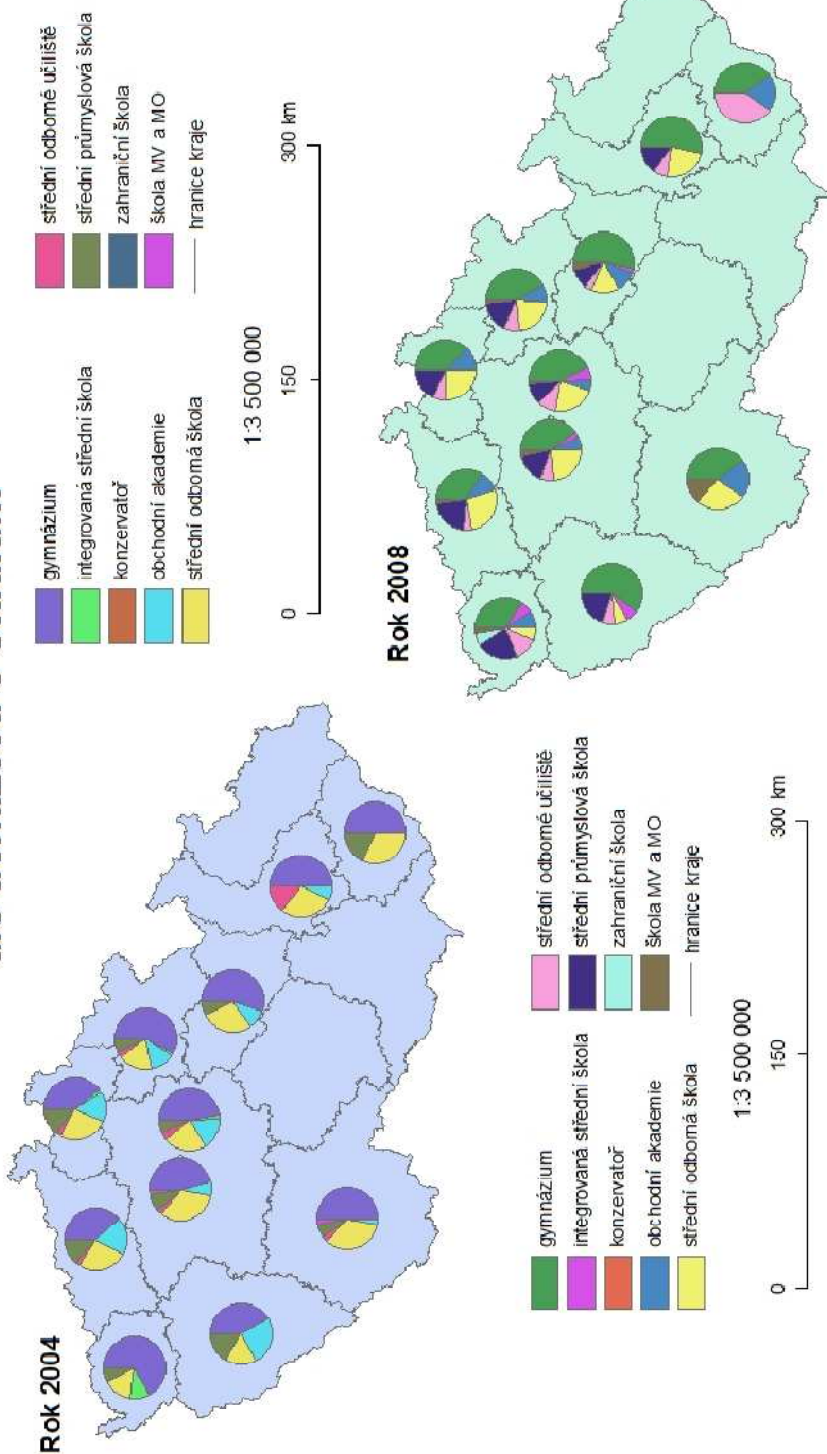


Příloha 8: Rozmístění uchazečů o studium v letech 2004-2008



Příloha 9: Podíl typů středních škol v krajích ČR v letech 2004 a 2008

PODÍL TYPŮ STŘEDNÍCH ŠKOL V KRAJÍCH ČR V ROCE 2004 A 2008 dle uchazečů o studium



Kraje, ve kterých není zobrazen žádný kartodiagram, nemají daným roce žádného uchazeče o studium na FP TUL.

Zbýněk PROKOP
Liberec 2009
Katedra Geografie, FP TUL
Software: ArcGIS 9.3
Zdrojová data: AICR 500
Matrka TUL