4IT436 – Zadání příkladu pro řešení na cvičeních

(celkové řešení tohoto příkladu se odevzdává do konce výuky v semestru + jeho vybrané části jsou kontrolovány a hodnoceny v rámci kontrolních dnů v průběhu semestru)

**Cíl:**

* Analýza dat firmy MATASTAV
* Návrh datového skladu pro firmu MATASTAV

Obsah

[1. Popis firmy Matastav 2](#_Toc33699587)

[2. Zadání cvičného příkladu 6](#_Toc33699588)

[2.1 Společná část 6](#_Toc33699590)

[2.1.1 Zadání společné části 6](#_Toc33699591)

[2.2 Individuální část 7](#_Toc33699592)

[3. Vstupní data a další nutné podmínky pro řešení příkladu 10](#_Toc33699593)

[4. Kontrolní bod 1 11](#_Toc33699594)

[4.1. Analýza datového skladu a řešení BI 11](#_Toc33699595)

[4.2. Tabulky DWH 11](#_Toc33699596)

# Popis firmy Matastav

Firma Matastav se zabývá pronájmem stavební mechanizace a stavebních dělníků (dále jen zdroje). Svoji činností pokrývá celou ČR. Centrála firmy sídlí v Praze a dále má celkem 5 dalších poboček po celé ČR.

* Centrála a každá z poboček může pronajímat všechny zdroje, které firma Matastav vlastní.
* Na centrále i na jednotlivých pobočkách pracují zaměstnanci s různým základním platem, který je zvyšován o ohodnocení spojené s výkonem určité funkce.
* Zdroje jsou rozděleny do kategorií (Speciální stroj, Nákladní auto, Pracovní četa – zedníci, Pracovní četa – kopáči, Pracovní četa – malíři, Pracovní četa – nerozlišeno)
* Na každé z poboček a na centrále existuje jeden nebo více obchodníků, kteří se starají o svoje zákazníky.
* Zákazník vždy patří pod patronaci jednoho obchodníka (může sídlit i v jiném kraji než obchodníkova pobočka)
* Zákazníci jsou rozdělení do kategorií (Normal, Malozákázník, Velkozákazník, VIP)
* O každém pronájmu zdroje se sepíše "**Kontrakt**", ve kterém je uvedeno který zdroj, od kdy, do kdy a za jakou cenu je danému zákazníkovi pronajímán. Cena pronájmu se počítá z doporučené ceny uvedené u každého zdroje (tato cena je případně zvyšována či snižována podle toho do jaké kategorie daný zákazník spadá). Ve chvíli uzavření se kontrakt zařadí do plánu příjmů a výdajů (s kontraktem jsou také spojeny určité náklady) firmy. Po uskutečnění kontraktu se na něj vystaví faktura (tabulka t\_faktury\_vydane), která má datum splatnosti a také se u ní sleduje datum skutečného zaplacení.
* Údaje o pronájmech se sledují v tabulce Kontrakty, **jejich plánovaná a skutečná finanční realizace** v tabulce t\_prijmy\_plan a t\_prijmy\_skutecnost (**POZOR!! - údaje cena a náklady** v tabulce Kontrakty jsou pouze **plánované hodnoty** ve chvíli uzavírání smlouvy o kontraktu, nikoliv skutečně realizované – ty je třeba hledat ve fakturách a účetních knihách) .
* Firma vede jednoduché účetnictví (daňovou evidenci) tzn., účtují se příjmy a výdaje a z jejich rozdílu se určuje hrubý zisk
* Všechny realizované účetní operace spojené s činností firmy a jejích poboček jsou účtovány do peněžního deníku (jako tzv. **Položky**). Kromě toho se ještě vedou knihy závazků a pohledávek, kde jsou uvedeny plánované, ale doposud neuhrazené platby – budoucí příjmy a výdaje.
* Každá položka skutečného plnění v peněžním deníku má u sebe referenci na položku plánovaného plnění v knize závazků a pohledávek, ke které je připojená (např. mzda vyplacená se odkazuje na mzdu plánovanou k vyplacení)
* Kromě toho firma vede ještě deník vystavených faktur.

**Peněžní deník** se fyzicky skládá ze záznamů v tabulkách:

t\_vydaje\_osobni\_skutecnost – osobní náklady refundované zaměstnancům – cestovné, stravné, refundace ošacení apod.

t\_vydaje\_rezie\_skutecnost – režie poboček (plyn, elektřina, úklid, kancelářské potřeby apod.)

t\_vydaje\_najem\_skutecnost – nájem poboček (tam kde pobočka nesídlí v prostorách vlastněných firmou)

t\_vydaje\_mzdy\_skutecnost – mzdy vyplácené zaměstnancům

t\_vydaje\_zdroje\_skutecnost – výdaje spojené se zdroji (údržba, platby za pronájem třetí straně apod.)

t\_prijmy\_skutecnost – příjmy inkasované za jednotlivé kontrakty od zákazníků

tyto tabulky mají následující datovou strukturu

* idpol – interní číslo dokladu
* datum – datum realizace transakce
* castka – částka transakce
* druh – druh transakce
* protistrana – kdo položku platí či komu je placena
* ucet – způsob realizace – platba z/na účet, hotovost na pokladně apod.
* pozn – bližší popis transakce

**Knihy závazků a pohledávek** se fyzicky skládají ze záznamů v tabulkách

t\_vydaje\_osobni\_plan

t\_vydaje\_rezie\_ plan

t\_vydaje\_najem\_ plan

t\_vydaje\_mzdy\_ plan

t\_vydaje\_zdroje\_ plan

t\_prijmy\_ plan

jejich obsah je stejný jako u tabulek peněžního deníku.

Ve chvíli skutečné realizace pohledávky či závazku se její záznam/řádek označí jako „realizovaný“ a převede se do peněžního deníku. Oproti „standardním“ účetním postupům se ale záznamy z knihy závazků a pohledávek nemažou.

**Deník vystavených faktur** se fyzicky skládá z tabulky

t\_faktury vydane

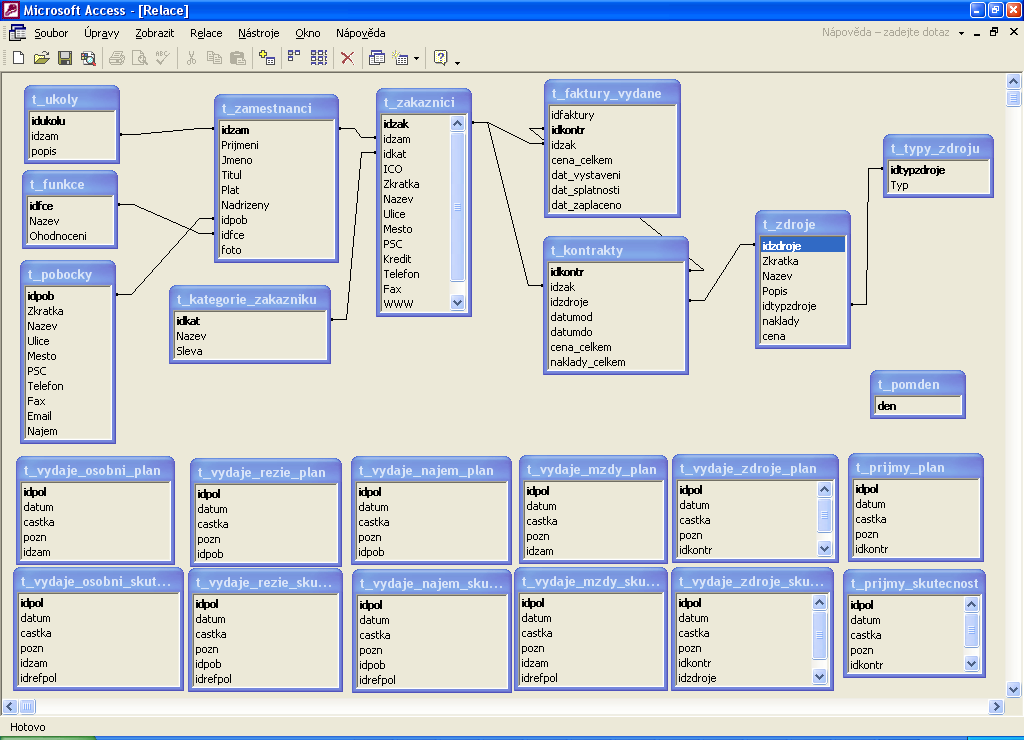
kde jsou také uvedena data vystavení, splatnosti a skutečného zaplacení

Dále jsou v jednotlivých tabulkách peněžního deníku či závazků a pohledávek vyplněny následující informace (vždy jen ty, které se vztahují k danému typu operací).

* + - idzam (id zaměstnance jehož se účetní položka týká)
    - idpob (id pobočky, které se účetní položka týká)
    - idzdroje (id zdroje kterého se účetní položka týká)
    - idkontr (id kontraktu kterého se účetní položka týká)

tedy např. k záznamům o pobočkách se nepřipojuje id zaměstnance apod.

* Podrobné údaje o kontraktech (kontrahovaný zdroj, zákazník atd.) jsou uvedeny v tabulce "Kontrakty" (viz výše), ostatní účetní údaje o jejich účetní realizaci jsou uvedeny pouze z tabulek " t\_faktury\_vydane ", " t\_faktury\_vydane ", "t\_prijmy\_plan", "t\_prijmy\_skutecnost", "t\_vydaje\_zdroje\_plan", "t\_vydaje\_zdroje\_skutecnost". **Pozor – velmi často dochází k rozdílu mezi kontrahovanou cenou v tabulce kontrakty a skutečně fakturovanou cenou.**
* Veškerá existující data, která Matastav udržuje, naleznete v souboru MS Access, který si můžete stáhnout ze stránek tohoto předmětu. Výjimkou jsou informace o uzavřených smlouvách, které jsou k dispozici ve formě přiložených textových souborů. **Tabulka t\_kontrakty je tedy prázdná** a je nutno ji z těchto souborů naplnit.
* schéma existujících dat firmy Matastav:



# Zadání cvičného příkladu



## Společná část

*Tato část zadání je společná pro všechny studenty. Cílem je zpracovat dimenzionální analýzu (viz kapitola 3), datové pumpy, OLAP kostky a výstupy v programu MS Excel a Tableau tak, aby firma Matastav mohla provádět analýzy dle následujícího zadání.*

Obsah požadovaných výstupů je vedením firmy specifikován následovně *(v závorkách je uvedena požadovaná forma výstupu).*

### Zadání společné části

**Vedení firmy Matastav chce mít přehled o následujících faktech:**

*(Předveďte v nástroji* ***Excel****)*

* v jaké výši (finančně) má uzavřeny smlouvy na kontrakty
* jaká je plánovaná marže firmy na těchto kontraktech
* jaká je průměrná doba tvání kontraktů
* jak jsou výše uvedené objemy kontraktů rozloženy
  + v čase - v pohledu rok, měsíc a v pohledu rok, kvartál, měsíc (*výstup formou čárového grafu)*
  + mezi zákazníky a jejich kategorie (*výstup formou kontingenční tabulky*)
  + mezi zdroje a jejich typy (*výstup formou kontingenční tabulky)*
  + mezi jednotlivé kraje a okresy ČR (zařaďte vždy podle sídla zákazníka)
* jaký je vývoj všech příjmů a výdajů společnosti v rozložení

*(formu a rozsah výstupu zde zvolte sami – předveďte v nástroji* ***Tableau*** *– výsledný výstup ale budete muset interpretovat.)*

* + podle jednotlivých druhů příjmů nebo výdajů:
    - operace týkající se zdrojů (příjmy a výdaje)
    - operace týkající se nájmu (výdaje)
    - operace týkající se režie poboček (výdaje)
    - operace týkající se mezd (výdaje)
    - operace týkající se osobních výdajů (výdaje)
  + v časovém rozlišení rok, kvartál, měsíc
  + v rozlišení plán/skutečnost – **zde spočítejte také rozdíl oproti plánu**
  + podle jednotlivých poboček (tam kde to má smysl)

## Individuální část

*Další část* *zadání je individuální pro každého studenta. V systému INSIS se každý student přihlašujete na jedno z následujících zadání. To potom vypracuje a odevzdává společně se zadáním pro všechny studenty (viz výše Společná část). Je na zvážení studenty zda výstupy předveďte v nástroji* ***Tableau*** *(preferováno) nebo v Excel.*

1. Zpracujte přehled vyplácených mezd na jednotlivých pobočkách (s možností rozpadu na jednotlivé zaměstnance poboček) a porovnejte ho s plánem v jednotlivých měsících. Je tam nějaký problém? Pokud ano, tak kdy a u koho? Zpracujte graficky přehled vývoje skutečné (absolutní) výše mezd v jednotlivých letech, zobrazte i meziroční rozdíly v % (tento meziroční rozdíl musí být možné zobrazit jak na úrovni pobočky, funkce nebo pro každého zaměstnance zvlášť).
2. Porovnejte *průměrné* mzdy u jednotlivých poboček a ukažte jejich vývoj v čase, v rámci jednotlivých funkcí, v rámci pracovníků pracujících v těchto funkcích. Kromě toho také zpracuje ukazatele % meziročního rozdílu v těchto mzdách.
3. Zpracujte přehled režijních nákladů a plateb nájemného jednotlivých poboček a jejich porovnání s plánem – režijní náklady a nájemné poboček sdružujte do celků na základě jejich sídla v určitém okrese či kraji. Je u některé pobočky problém? Pokud ano, tak u které? Zobrazte meziroční rozdíly v %, spočítejte také klouzavé 5 měsíční průměry.
4. Zobrazte přehled vyplácených částek (mzda + úhrada osobních výdajů) pracovníkům v jednotlivých funkcích v rámci jednotlivých poboček. Je zde nějaký rozdíl mezi plánovanou a skutečně vyplácenou částkou? Které funkce a v jakém období se to týká nejvíce? Zpracujte ukazatele, který zobrazí tento rozdíl v absolutním i v procentním vyjádření (přes pobočku, přes funkce, přes celou firmu). Dále zpracujte ukazatel minimálního a maximálního rozdílu mezi skutečnou a plánovanou mzdou v % (umožňující zjistit jaký byl např. maximální procentní přeplatek v daném měsíci v určité pobočce či funkci)
5. Graficky zobrazte skutečné vytížení jednotlivých skupin zdrojů a konkrétních zdrojů v průběhu sledovaného období. Které zdroje jsou vytěžovány nejméně? Vytížení počítejte jako počet dní, kdy byl v měřeném období zdroj aktivován na kontraktu. Pozor na kontrakty přecházející mezi měsíci či roky (např. kontrakt od 20.4. do 15.5.) – vytížení z takového kontraktu je třeba rozpočítat do obou období.
6. Firma Matastav z určitých důvodů občas zpožďuje platby nákladů některých zdrojů oproti skutečnosti.. Pro který zdroj je situace nejhorší? Jaké poplatky by Matastav musel platit, pokud by za každý den prodlení platil penále ve výši 100 % diskontní úrokové sazby České národní banky platné v den uzavření kontraktu? Penále u faktur, které nebyly dosud zaplaceny se počítá k aktuálnímu datu (ke dni generování výstupu). Graficky znázorněte, jak je velká tato doba zpoždění a potenciální penále pro jednotlivé zdroje a pro skupiny zdrojů. (Pozn: Tabulku pro sazby ČNB můžete vytvořit ručně – musí v ní být ale skutečné sazby, nikoliv vymyšlené)
7. Spočítejte a zobrazte, jak velké finanční částky by firma Matastav získala, pokud by zákazníkům účtovala poplatky z prodlení. Firma Matastav počítá penále za každý den prodlení počínaje dnem následujícím po dni splatnosti až do dne platby včetně, a to ode dne původního dne splatnosti. Penále je stanoveno za každý den prodlení ve výši 0,1 % z dlužné částky (do ní se nezahrnuje nabíhající penále). Penále u faktur, které nebyly dosud zaplaceny se počítá k aktuálnímu datu (ke dni generování výstupu). Zpracujte výstup tak, aby umožňoval pohled z časového hlediska přes jednotlivé skupiny zákazníků a dále i přes konkrétní zákazníky. Vytvořte výstup tak, že bude vždy (přes jakékoliv vybrané dimenze) možno porovnat částku plánovanou, skutečně zaplacenou a zaplacenou v případě započtení poplatků z prodlení.
8. Pro každou pobočku (dále v podrobnějším dělení na její obchodníky) spočítejte, kolik jejích kontraktů bylo v daném roce nebo měsíci v prodlení a kolik z těchto kontraktů nebylo doposud zaplaceno. Spočítejte také ukazatel průměrné výše nezaplaceného či zpožděného kontraktu za sledované období (přes všechny dimenze). Zobrazte meziroční rozdíly v %, spočítejte také klouzavé 5 měsíční průměry. Zpracujte přehled graficky.
9. Rozdělte kontrakty do skupin 0-1milKč, 1-2 mil Kč, 2-3 mil. Kč, 3-10 mil. Kč, 10+ mil Kč. Zpracujte přehled, který ukáže, kolik kontraktů daného typu realizovala určitá kategorie zákazníků (s umožněním rozpadu na jednotlivé zákazníky) a jaké *skutečné* příjmy a jaký skutečný hrubý zisk byl z těchto kontraktů realizován (celkově a v přepočtu na jeden kontrakt). Pro skutečný hrubý zisk zpracujte ukazatel kumulativního součtu v rámci každého roku (tedy např. v březnu bude ukazovat kumulativní zisk za leden, únor a březen) a klouzavý 3 měsíční průměr.
10. Pro každou skupinu nákladových typů účetních operací zpracujte přehled ukazující, celkovou sumu zpožděných plateb tohoto typu nákladu oproti plánu a průměrnou částku na jeden zpožděný kontrakt. To samé zpracujte v rozlišení přes jednotlivé pobočky. Zpracujte také ukazatel kumulativní ukazatel těchto rozdílů v rámci každého roku (tedy např. v březnu bude ukazovat kumulativní zisk za leden, únor a březen) a klouzavý 3 měsíční průměr. Zobrazte graficky.
11. Jaká výše *skutečného* hrubého zisku připadala na jeden uskutečněný kontrakt v rámci jednotlivých skupin zákazníků, a v rámci geografického rozlišení na kraje a okresy v jednotlivých letech? Pro skutečný hrubý zisk zpracujte také ukazatel kumulativního součtu v rámci každého roku (tedy např. v březnu bude ukazovat kumulativní zisk za leden, únor a březen) a klouzavý 3 měsíční průměr. Zobrazte graficky.
12. Zpracujte přehled, který ukáže, jaký celkový *skutečný* hrubý zisk realizovala určitá kategorie zákazníků (s umožněním rozpadu na jednotlivé zákazníky) porovnejte tento zisk se ziskem plánovaným. Pro oba hrubé zisky (skutečný a plánovaný) zpracujte ukazatel kumulativního součtu v rámci každého roku (tedy např. v březnu bude ukazovat kumulativní zisk za leden, únor a březen) a klouzavý 3 měsíční průměr.
13. Zpracujte seznam obchodníků seřazený podle uzavřených (počet) kontraktů v jednotlivých letech a v daných regionech (kraj, okres). U každého obchodníka dále spočítejte, kolik těchto kontraktů bylo v daném roce v prodlení s platbou. Spočítejte i kumulativní počty v rámci daného roku a klouzavé 4měsíční průměrné počty.
14. Zpracujte přehled o počtech nasazení jednotlivých zdrojů a jejich typů v  okresech a krajích (podle sídla zákazníka). Zobrazte podle jednotlivých let a u každého zdroje uveďte i součet *skutečně realizovaného* hrubého zisku (podle času a podle geografického rozlišení). Pro skutečně realizovaný hrubý zisk zpracujte ukazatel kumulativního součtu v rámci každého roku (tedy např. v březnu bude ukazovat kumulativní zisk za leden, únor a březen) a klouzavý 3 měsíční průměr.
15. Firma Matastav z určitých důvodů občas zpožďuje platby mezd. Graficky znázorněte, jak je velká tato doba zpoždění pro jednotlivé funkce a pro jednotlivé pobočky. Dále zobrazte, kolik to dělá procent celkového objemu skutečně vyplacených mezd. Zpracujte dále ukazatel klouzavého 6 měsíčního průměru doby zpoždění.
16. Vytvořte si dva vlastní indexy na jejichž základě spočítáte efektivnost práce jednotlivých poboček, zobrazte vývoj těchto indexů v jednotlivých pobočkách (v čase). Ta samá pravidla aplikujte na firmu Matastav jako celek – jaké budou výsledky?
17. Zobrazte graf závislosti výše skutečně vyplacených částek (mzdy + osobní náklady) v závislosti na výši osobního ohodnocení funkce, ve které zaměstnanec pracuje – tento výstup vytvořte pro jednotlivé funkce, pobočka a konkrétní zaměstnance (pouze pro roky 2003 a 2004). Která funkce je na tom nejlépe a která nejhůře? (Pozn. jde o typ výstupu Advanced Analysis tools -> Perspective z ProClarity).
18. Pro každou skupinu výdajových účetních operací zpracujte přehled ukazující, o kolik dní byla zpožděna platba tohoto typu nákladu oproti plánu (v dělení podle jednotlivých poboček, zdrojů a pracovníků). Pozor – některé platby nebyly zaplaceny vůbec – i s tím se musíte nějak vyrovnat a do přehledu to nějakým způsobem zařadit. To samé zpracujte v časovém rozlišení.
19. Jaké *skutečné* denní náklady připadaly na jeden uskutečněný kontrakt v rámci jednotlivých skupin zákazníků, a v rámci geografického rozlišení na kraje a okresy v jednotlivých letech? Zobrazte graficky.
20. Spočítejte a zobrazte, kolik zákazníků bylo v prodlení s platbou za kontrakty v jednotlivých skupinách zákazníků v jednotlivých měsících. Dále zobrazte graf závislosti velikosti dlužné částky a na výši povoleného kreditu, kterou MATASTAV přidělil jednotlivým zákazníkům – tento graf zpracujte ve dvou úrovních členění – pro jednotlivé kategorie zákazníků a pro jednotlivé zákazníky. (Pozn. jde o typ výstupu Advanced Analysis tools -> Perspective z ProClarity).
21. Jaká výše *skutečného* hrubého zisku připadala na konkrétní zdroje v rámci jednotlivých skupin zákazníků, a v rámci geografického rozlišení na kraje a okresy v jednotlivých letech? Pro skutečný hrubý zisk zpracujte také ukazatel kumulativního součtu v rámci každého roku (tedy např. v březnu bude ukazovat kumulativní zisk za leden, únor a březen) a klouzavý 3 měsíční průměr. Zobrazte graficky.
22. Rozdělte kontrakty do skupin 0-1milKč, 1-2 mil Kč, 2-3 mil. Kč, 3-10 mil. Kč, 10+ mil Kč. Zpracujte přehled, který ukáže, kolik kontraktů daného typu realizovala určitá pobočka (s umožněním rozpadu na jednotlivé zaměstnance) a jaké *skutečné* příjmy a jaký skutečný hrubý zisk celkově a v přepočtu na jeden kontrakt realizovala. Pro skutečný hrubý zisk zpracujte ukazatel kumulativního součtu v rámci každého roku (tedy např. v březnu bude ukazovat kumulativní zisk za leden, únor a březen) a klouzavý 3 měsíční průměr.
23. Pro každou skupinu výdajových účetních operací zpracujte přehled ukazující, o kolik dní byla *průměrně* zpožděna platba tohoto typu nákladu oproti plánu (v dělení podle jednotlivých poboček, zdrojů a pracovníků). Pozor – některé platby nebyly zaplaceny vůbec – i s tím se musíte nějak vyrovnat a do přehledu to nějakým způsobem zařadit. To samé zpracujte v časovém rozlišení.
24. Pro každý zdroj spočítejte jeho *skutečný* průměrný hrubý denní zisk (vypočte se ze skutečných příjmů a nákladů)

# Vstupní data a další nutné podmínky pro řešení příkladu

* V rámci řešení zpracujete datovou pumpu (pumpy), které ze zdrojových dat v MS Access, MS Excel a textových souborů naplní výše uvedené tabulky datového skladu.
* Strukturu tabulek zpracujte tak, že bude pracovat s tabulkami, propojenými prostřednictvím **integritních omezení** a s nastavenými primárními klíči (***database diagram*** *v MS SQL Serveru – Management Studio*).
* Pozor – **pumpu musí být možné spustit i vícekrát za sebou** bez toho, aby došlo k duplikaci dat (musí obsahovat logiku, která tomu zamezí
* Do tabulek **datového skladu** importujte **pouze položky (atributy), které jsou tam potřeba** (ostatní položky nechte ve zdrojových souborech). V datovém skladu nesmí být v tabulkách uloženy atributy, které nejsou vzhledem k zadání potřeba!
* pro vygenerování seznamu krajů použijte soubor .xls soubor získaný od České Pošty (vazba PSČ na okresy) a .xls soubor získaný od ČSÚ (vazba Okresy – kraje). Tyto soubory naimportujete také prostřednictvím pumpy do datového skladu**. Žádné zdrojové soubory, které jsou součástí zadání, se nesmí jakkoliv manuálně upravovat**. Všechny úpravy dat musí být provedeny automaticky pomocí ETL datových pump.

# Kontrolní bod 1

## Analýza datového skladu a řešení BI

*Součástí řešení je i multidimenzionální analýza a zpracování struktury datového skladu (diagram v MS SQL serveru).*

Zpracujte strukturu datového skladu, která bude obsahovat tabulky s potřebnými sloupci, ze kterých bude možné vytvořit multidimenzionální kostku. V rámci analýzy budoucího řešení vyplňte přiloženou šablonu (.xls soubor pro multidimenzionální analýzu obsahující listy Dimenze, Ukazatelé - stejný jako na 2.cvičení)

**Vyplněnou šablonu přiložte k tomuto souboru a odevzdejte v rámci prvního kontrolního bodu.**

## Tabulky DWH

Na tomto místě doplňte seznam tabulek, které bude obsahovat váš DWH. Nezapomeňte, že všechna jména tabulek musí začínat vaším xname.

Tabulky faktů:

Tabulky dimenzí:

**Zde vložte screenshot schematu vašich tabulek uložených v MS SQL serveru, který je součástí prvního kontrolního bodu.**