

Základy forenzních databází

Tereza Uhlíková

verze 2.0

Kdo jsem

Tereza Uhlíková

Ústav analytické chemie

skupina teoretické spektroskopie

místnost A277

<https://web.vscht.cz/~uhlikovt/>

tereza.uhlikova@vscht.cz

1. lekce

- Co je to databáze
- Proč, kdo, kdy, jak ...
- Něco málo z historie
- CAP problém

Dobrodružství kriminalistiky - televizní seriál
Dějiny psané Římem - Vojtěch Zamarovský

2. lekce

- Základy informatiky
- Software
 - informace
 - záznam informace
 - číselné soustavy
 - písmenné kódy
- Hardware
 - Alan Turing, John von Neumann a počítač
 - historie vývoje počítače & super počítač
 - procesor a datová uložení
- Architektura databází

Alan Turing - Enigma

Contact film z roku 1997; seti@home

3. lekce

- Algoritmus
 - vlastnosti
 - zápis
 - struktura
- Datové typy

Arthur C. Clarke - Devět miliard božích jmen

4. lekce

- Výroková logika
- Databáze
 - DB a SŘBD
 - databázové modely
- Relační model
 - návrh tabulky

Aghata Christie - Hercules Poirot

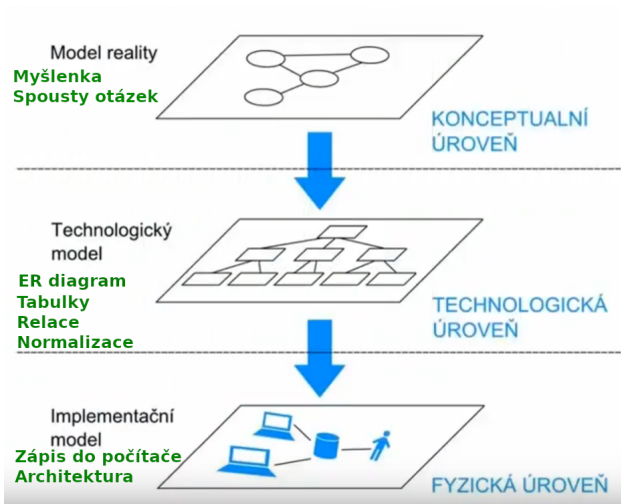
Co bude dnes

- ERA model
- Klíče, integrita a kardinalita
- Normalizace databáze

ERA MODEL

ERA MODEL

Návrh relační databáze



ERA model

ERA model

- **Entity** (objekt z reálného světa) → tabulka
- **Relationship** (vztah mezi entitami) → vztahy mezi tabulkami
- **Attribute** (vlastnosti entit) → sloupeček tabulky

- grafická reprezentace struktury databáze

Modeluje „oblast zájmu“ (např. podnik) jako kolekci entit a vztahů (relationships) mezi nimi.

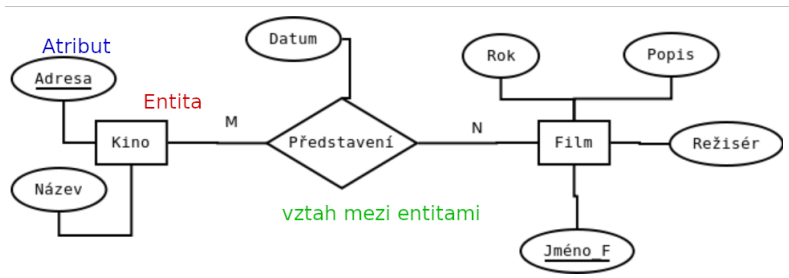
Dělá se před vytvářením databáze!! Slouží k rozmyšlení, jak bude databáze vypadat.

ERA příklad

Vytvořte ERA model pro databázi kin a promítaných filmů.

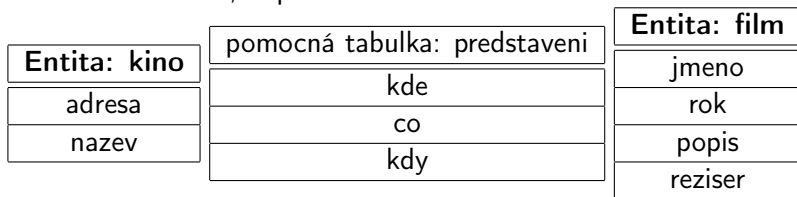
ERA model

Klasický formát



ERA model

tabulková varianta; implementační



KLÍČE A INTEGRITA

Integrita

Integrita

- celistvost, soudržnost, neporušenost
- územní integrita – územní celistvost státu v mezinárodním právu
- morální integrita – čestnost, poctivost, bezúhonnost v mravním smyslu
- vědecká integrita – celistvost a jednoznačnost principů vědeckého výzkumu
- matematická integrita – zásady dokazování a systém, který na základě něj byl vybudován
- datová integrita – v informatice kvalifikuje jistotu nebo záruku, že daná data byla přijata/přečtena bez chyb
- **databázová integrita** – kvalifikuje správné a nenarušené vztahy mezi jednotlivými záznamy v databázi

Primární klíč

Primární klíč

Entitní integrita – jedinečná identifikace každé entity/ každého záznamu v entitě - pomocí primárního klíče (PK)

Primární klíč

Entitní integrita – jedinečná identifikace každé entity/ každého záznamu v entitě - pomocí primárního klíče (PK)

Primární klíč

- zajistí jednoznačnost záznamů v tabulce
- hodnota primárního klíče se v tabulce nesmí opakovat nebo mít hodnotu Null
- teoreticky může být libovolného typu
- může být složený z více sloupců – unikátnost je pak brána ze všech najednou, v jednotlivých sloupcích se tedy mohou opakovat stejné hodnoty
- AUTO INCREMENT - zajistí automatické číslování

Referenční klíč

Referenční klíč

Referenční integrita – udržuje správné vztahy v relačně propojených databázových tabulkách - pomocí cizího klíče (FK)

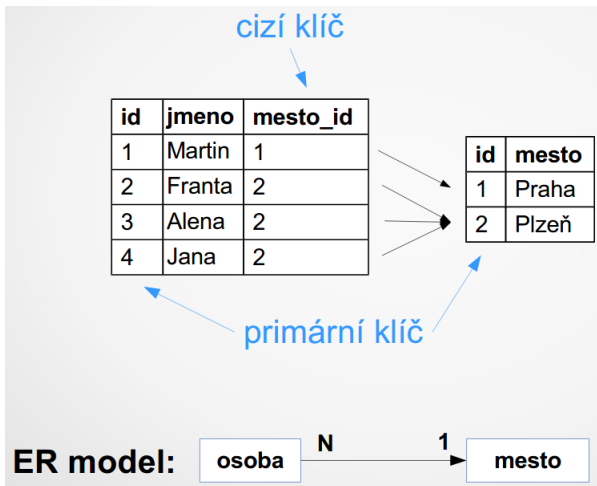
Cizí klíč

- odkazuje se na primární klíč v jiné tabulce
- odkazuje se na právě 1 tabulku
- dvojice primární a cizí klíč tvoří vazbu (relaci) mezi tabulkami
- cizí klíč musí být stejného typu a velikosti jako primární:

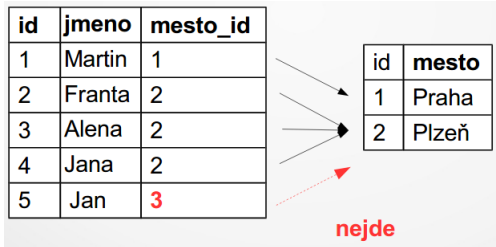
primární klíč AUTO INCREMENT \Rightarrow cizí klíč dlouhé celé číslo

primární klíč složený \Rightarrow struktura cizího klíče stejná

Tvorba relace



Vztahy klíčů



Do sloupce s definovaným cizím klíčem nelze vložit hodnotu, která není vložena v nadřazeném sloupci primární tabulky.

Zánik relace

Zánik relace

Restrict Je zakázáno smazat nadřízený (primární) záznam, pokud existují záznamy jemu podřízené (cizí).

Zánik relace

- Restrict** Je zakázáno smazat nadřizený (primární) záznam, pokud existují záznamy jemu podřizené (cizí).
- Cascade** Proublání informace o smazání nadřizeného záznamu a smazání i záznamů jemu podřizených.

Zánik relace

- Restrict** Je zakázáno smazat nadřazený (primární) záznam, pokud existují záznamy jemu podřízené (cizí).
- Cascade** Proublání informace o smazání nadřazeného záznamu a smazání i záznamů jemu podřízených.
- Set_null** Při mazání nadřazeného záznamu je jemu podřízeným záznamům odebráno propojení na tento samotný záznam je smazán.

Kardinalita vztahu

Kardinalita vztahu

maximální počet vztahů instancí, kterých se může entita účastnit
mocnost vztahu mezi entitami

Kolik položek z tabulky A se váže na kolik položek z tabulky B?

Kardinalita vztahu

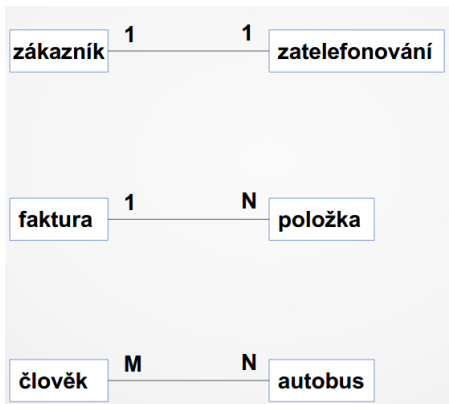
maximální počet vztahů instancí, kterých se může entita účastnit
mocnost vztahu mezi entitami

Kolik položek z tabulky A se váže na kolik položek z tabulky B?

3 typy

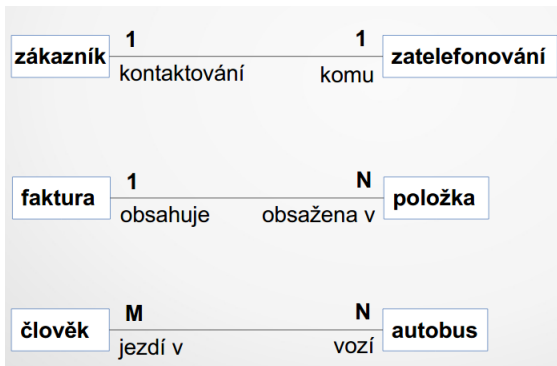
- 1:1 - na 1 položku z tabulky A připadá maximálně 1 položka z tabulky B, ta patří jen té jediné z A, žena-muž (ČR)
- 1:N, N:1 - na 1 položku z tabulky A připadá libovolné množství položek z tabulky B, na 1 položku z tabulky B připadá maximálně 1 položka z tabulky A
- M:N - bez omezení množství vazeb v obou směrech, žena-muž (ČR) v čase

Kardinalita vztahu



Kardinalita vztahu

pomůže lepšímu porozumění ER(A) modelu



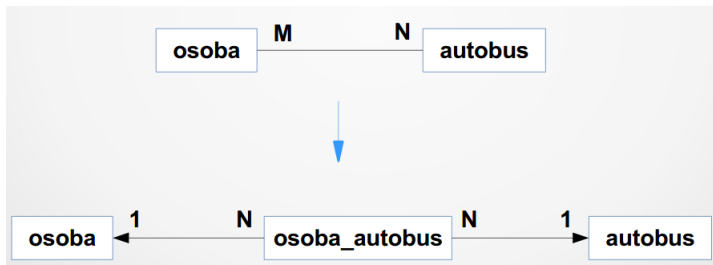
Kardinalita 1:N

- přiřazuje jednomu záznamu více záznamů z jiné tabulky
- např. vztah třída - žák
 - jeden žák navštěvuje jednu třídu
 - jedna třída se skládá z více žáků

Kardinalita M:N

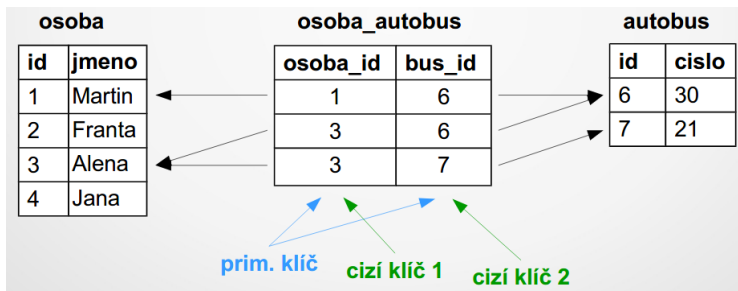
- umožňuje každému záznamu z jedné tabulky přiřadit libovolný počet záznamů z druhé tabulky
- např. vztah student - předmět
 - jeden student studuje více předmětů \wedge jeden předmět navštěvuje více studentů
 - nejsložitější případ vazby

vazbu M:N nelze vytvořit přímo mezi tabulkami, je potřeba rozložit pomocí nové tabulky na dvě vazby 1:N



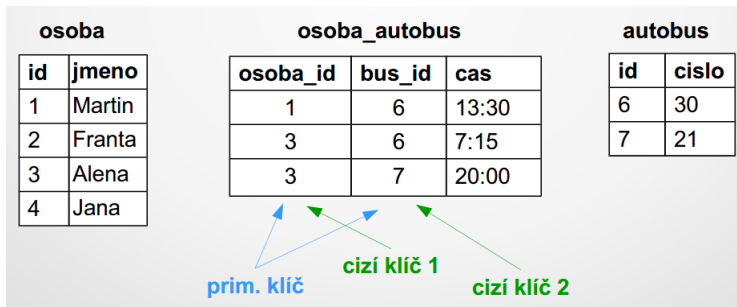
Kardinalita M:N

rozkladová tabulka může / nemusí neobsahovat další informace



Kardinalita M:N

- spojovací tabulka může obsahovat další informace
- např. v kolik hodin a kdo jakým autobusem jezdí
 - není nutné kvůli tomu vytvářet další tabulku



Příklad

Doplňte vazby a jejich kardinality do ER modelu – Základní zadání:
zaměstnanci používají služební auta.

Příklad

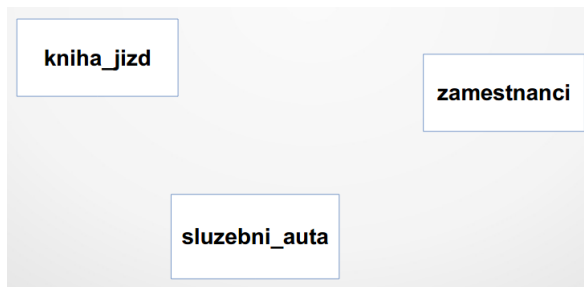
Doplňte vazby a jejich kardinality do ER modelu – Základní zadání:
zaměstnanci používají služební auta.

Rozšířené: Každý zaměstnanec může použít několik služebních aut a každé auto může používat i třeba během dne více zaměstnanců (ale ne zároveň).
Každá jízda s časem a délkou se musí zapsat do knihy jízd.

Příklad

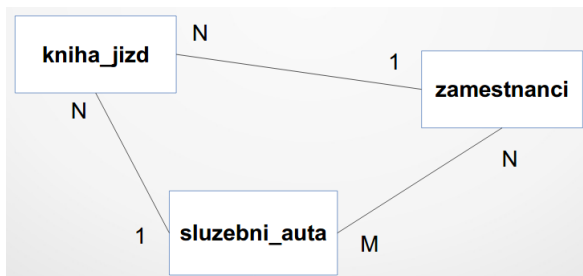
Doplňte vazby a jejich kardinality do ER modelu – Základní zadání:
zaměstnanci používají služební auta.

Rozšířené: Každý zaměstnanec může použít několik služebních aut a každé auto může používat i třeba během dne více zaměstnanců (ale ne zároveň).
Každá jízda s časem a délkou se musí zapsat do knihy jízd.



Kardinalita vztahu

výsledek



příklady

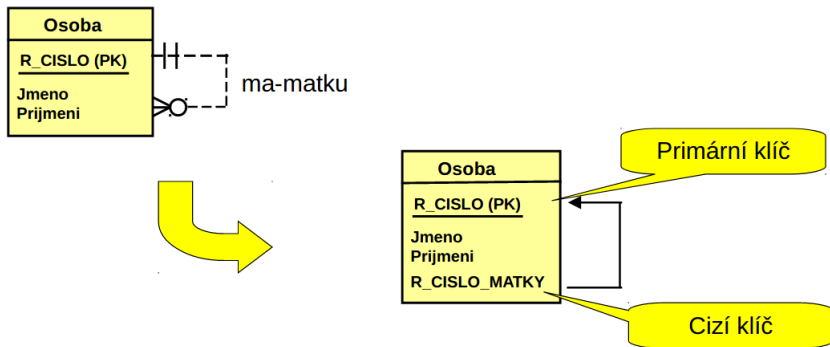
- Číšník a Stůl – 1 stůl může být obsluhován více číšníky – pamatujeme si i čas obsluhy
- Student a Semestrální práce – seznam témat a jedno si vybere
- Jídlo a Ingredience - před a po rozkladu
- Student a Termín zkoušky

Rekurze

Rekurze

vztah mezi osobou a matkou dané osoby

Matka osoby je rovněž osobou, nemůže být tudíž v jiné tabulce!



NORMALIZACE

NORMALIZACE

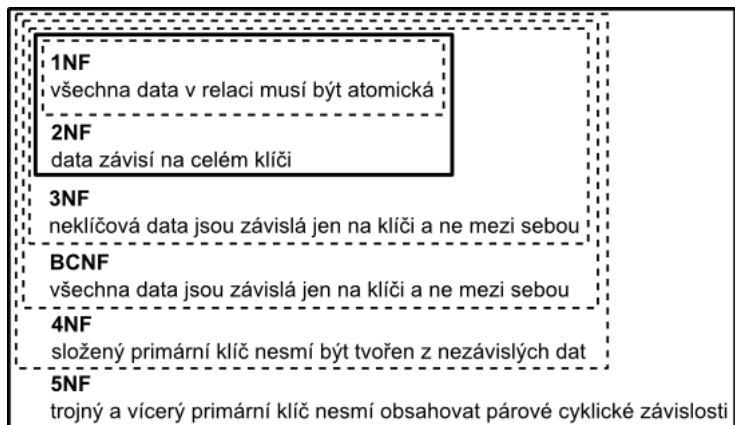
Normalizace

Slouží k jasnému vyřešení problémů s nadbytečností.

odstranění redundance - opakovaných výskytů stejného údaje v databázi

- eliminovat duplicitní (opakující se) sloupce v jednotlivých tabulkách;
- vytvořit pro každou skupinu dat s jasně vymezeným významem samostatnou tabulku a každý řádek opatřit jedinečným primárním klíčem;
- obsahem jednotlivých sloupců tabulky by měla být jednoduchá, dále nedělitelná informace
- podmnožinu dat se shodnou hodnotou pro určitý sloupec tabulky převést do samostatné tabulky a spojit s původní tabulkou cizím klíčem;
- odstranit z tabulky sloupce, které jsou přímo závislé na jiné skupině sloupců tabulky než pouze na primárním klíči

Pravidla normalizace



Tabulka půjček

Záznamy o výpůjčkách						
Číslo záznamu	Jméno zákazníka	Bydliště	Vypůjčil si			Vrátil
			DVD	Název filmu	Dne	
1	Lukáš Pavel	Studentská 25, Ostrava Přívoz, 708 31	21	Piráti z Karibiku I.	5.1.2012	9.1.2012
2	Lukáš Pavel	Studentská 25, Ostrava Přívoz, 708 31	5	Honza málem králem	5.1.2012	8.1.2012
3	Pavel Kvapil	Karvinská 4, Karviná-Ráj, 735 64	21	Piráti z Karibiku I.	10.1.2012	11.1.2012
4	Pavel Kvapil	Karvinská 4, Karviná-Ráj, 735 64	22	Piráti z Karibiku II.	10.1.2012	11.1.2012
5	Pavel Kvapil	Karvinská 4, Karviná-Ráj, 735 64	23	Piráti z Karibiku III.	10.1.2012	
6	Levá Pavla	Hornosušská 24, Horní Suchá, 735 35	51	Popelka	10.1.2012	10.1.2012
7	Pavel Lukáš	Studentská 25, Ostrava-Přívoz, 708 31	51	Popelka	15.2.2012	19.2.2012
8	Levá Pavla	Studentská 1, Ostrava-Přívoz, 708 31	5	Honza málem králem	15.2.2012	16.2.2012
9	Karel Novák	Vnitřní 1151, Těrlicko, 735 65	22	Piráti z Karibiku II.	16.2.2012	16.2.2012
...						

Úprava databázové tabulky dle normálních forem

Záznamy o výpůjčkách						
Číslo záznamu	Jméno zákazníka	Bydliště	Vypůjčil si			Vrátil
			DVD	Název filmu	Dne	
1	Lukáš Pavel	Studentská 25, Ostrava Přívoz, 708 31	21	Piráti z Karibiku I.	5.1.2012	9.1.2012
2	Lukáš Pavel	Studentská 25, Ostrava Přívoz, 708 31	5	Honza málem králem	5.1.2012	8.1.2012
3	Pavel Kvapil	Karvinská 4, Karviná-Ráj, 735 64	21	Piráti z Karibiku I.	10.1.2012	11.1.2012
4	Pavel Kvapil	Karvinská 4, Karviná-Ráj, 735 64	22	Piráti z Karibiku II.	10.1.2012	11.1.2012
5	Pavel Kvapil	Karvinská 4, Karviná-Ráj, 735 64	23	Piráti z Karibiku III.	10.1.2012	
6	Levá Pavla	Hornosušská 24, Horní Suchá, 735 35	51	Popelka	10.1.2012	10.1.2012
7	Pavel Lukáš	Studentská 25, Ostrava-Přívoz, 708 31	51	Popelka	15.2.2012	19.2.2012
8	Levá Pavla	Studentská 1, Ostrava-Přívoz, 708 31	5	Honza málem králem	15.2.2012	16.2.2012
9	Karel Novák	Vnitřní 1151, Těrlicko, 735 65	22	Piráti z Karibiku II.	16.2.2012	16.2.2012
...						

Co je špatně

- nejednoznačnost** údajů – vyhledávací anomálie – není jednoznačně určeno, co je jméno a co příjmení zákazníka (viz záznamy 1, 2 a 7), název města v záznamech 1, 2 a 7 je odlišný i přesto, že jde o stejné město;
- redundance** záznamů – když si zákazník vypůjčí více filmů najednou, pak se pro každý záznam opakují pole bydliště a datum vypůjčení, totéž při dalších výpůjčkách stejného zákazníka;
- nekonzistence** záznamů – pokud zákazník změní bydliště (viz záznamy č. 6 a 8), pak je nutné provést změnu ve všech záznamech.

1 NF

Relace je v první normální formě, pokud každý její atribut obsahuje jen atomické hodnoty.

Řešením je rozklad polí Jméno zákazníka a Bydliště na několik dalších polí

Záznamy o výpůjčkách										
Číslo záznamu	Jméno zákazníka	Příjmení zákazníka	Ulice	Číslo	Město	PSČ	Vypůjčil si			Vrátil
							DVD	Název filmu	Dne	
1	Lukáš	Pavel	Studentská	25	Ostrava Přívoz	708 31	21	Piráti z Karibiku I.	5.1.2012	9.1.2012
2	Lukáš	Pavel	Studentská	25	Ostrava Přívoz	708 31	5	Honza málem králem	5.1.2012	8.1.2012
3	Pavel	Kvapil	Karvinská	4	Karviná-Ráj	735 64	21	Piráti z Karibiku I.	10.1.2012	11.1.2012
4	Pavel	Kvapil	Karvinská	4	Karviná-Ráj	735 64	22	Piráti z Karibiku II.	10.1.2012	11.1.2012
5	Pavel	Kvapil	Karvinská	4	Karviná-Ráj	735 64	23	Piráti z Karibiku III.	10.1.2012	
6	Pavla	Levá	Hornosušská	24	Horní Suchá	735 35	51	Popelka	10.1.2012	10.1.2012
7	Lukáš	Pavel	Studentská	25	Ostrava-Přívoz	708 31	51	Popelka	15.2.2012	19.2.2012
8	Pavla	Levá	Studentská	1	Ostrava-Přívoz	708 31	5	Honza málem králem	15.2.2012	16.2.2012
9	Karel	Novák	Vnitřní	1151	Těrlicko	735 65	22	Piráti z Karibiku II.	16.2.2012	16.2.2012
...										

2 NF

Každý neklíčový atribut je plně závislý na primárním klíči

Rozložení tabulky pro data s jasně vymezeným významem/ jednoznačnou entitou.

Zákazník						
Číslo záznamu	Jméno zákazníka	Příjmení zákazníka	Ulice	Číslo	Město	PSČ
1	Lukáš	Pavel	Studentská	25	Ostrava-Přivoz	708 31
2	Pavel	Kvapil	Karvinská	4	Karviná-Ráj	735 64
3	Pavla	Levá	Studentská	1	Ostrava-Přivoz	708 31
4	Karel	Novák	Vnitřní	1151	Těrlícko	735 65
5	...					

Výpůjčka		
ID výpůjčky	Zákazník	Dne
1	1	5.1.2012
2	2	10.1.2012
3	3	10.1.2012
4	1	15.2.2012
5	3	15.2.2012
6	4	16.2.2012
...		

Položka				
Číslo položek	DVD	Název filmu	Výpůjčka	Vráceno
1	21	Piráti z Karibiku I.	1	9.1.2012
2	5	Honza málem králem	1	8.1.2012
3	21	Piráti z Karibiku I.	2	11.1.2012
4	22	Piráti z Karibiku II.	2	11.1.2012
5	23	Piráti z Karibiku III.	2	
6	51	Popelka	3	10.1.2012
7	51	Popelka	4	19.2.2012
8	5	Honza málem králem	5	16.2.2012
9	22	Piráti z Karibiku II.	6	16.2.2012
10	...			

3 NF

Žádný z jejich atributů nevykazuje tranzitivní závislost, tzn., že všechny neklíčové atributy jsou navzájem nezávislé

Vytvoření nových tabulek *Město* a *DVD*

Zákazník					
Číslo záznamu	Jméno zákazníka	Příjmení zákazníka	Ulice	Číslo	PSČ
1	Lukáš	Pavel	Studentská	25	708 31
2	Pavel	Kvapil	Karvinská	4	735 64
3	Pavla	Levá	Studentská	1	708 31
4	Karel	Novák	Vnitřní	1151	735 65
5	...				

Město	
PSČ	Město
708 31	Ostrava Přívoz
735 64	Karviná-Ráj
735 35	Horní Suchá
735 65	Těrlicko
...	

Vypůjčka		
ID vypůjčky	Zákazník	Dne
1	1	5.1.2012
2	2	10.1.2012
3	3	10.1.2012
4	1	15.2.2012
5	3	15.2.2012
6	4	16.2.2012
...		

DVD		
Číslo DVD	Název filmu	Kategorie
...
5	Honza málem králem	Pohádka
...
21	Piráti z Karibiku I.	Komedie
22	Piráti z Karibiku II.	Komedie
23	Pirátki z karibiku III.	Komedie
...
51	Popelka	Pohádka

Položka			
Číslo položek	DVD	Vypůjčka	Vráceno
1	21	1	9.1.2012
2	5	1	8.1.2012
3	21	2	11.1.2012
4	22	2	11.1.2012
5	23	2	
6	51	3	10.1.2012
7	51	4	19.2.2012
8	5	5	16.2.2012

Pravidla normalizace

I. pravidlo: Jedinečnost polí. Každé pole v tabulce by mělo představovat jedinečný typ informace.

II. pravidlo: Primární klíče. Každá tabulka musí mít jednoznačný identifikátor neboli primární klíče, který je vytvořen z jednoho nebo více polí v této tabulce.

III. pravidlo: Funkcionální závislost. Pro každou jedinečnou hodnotu primárního klíče se musí hodnoty v datových sloupcích týkat předmětu tabulky a musí tento předmět úplně popisovat.

IV. pravidlo: Nezávislost polí. Pokud se provede změna dat v libovolném poli (ne v primárním klíči), pak nebudou ovlivněna data v jakémkoliv jiném poli.

Skripta, stránky a materiály

<https://www.itnetwork.cz/mysql/restricke-v-databazi-sql-on-delete-a-on-update>

<http://books.fs.vsb.cz/dbacc20/Welcome.htm>

http://projekty.fs.vsb.cz/463/edubase/VY_01_044/

[Datab%C3%A1zov%C3%A9%20syst%C3%A9my/02%20Text%20pro%20e-learning/](#)

[Datab%C3%A1zov%C3%A9%20](#)

[syst%C3%A9my%2005.%20Normalizace%20dat.pdf](#)