

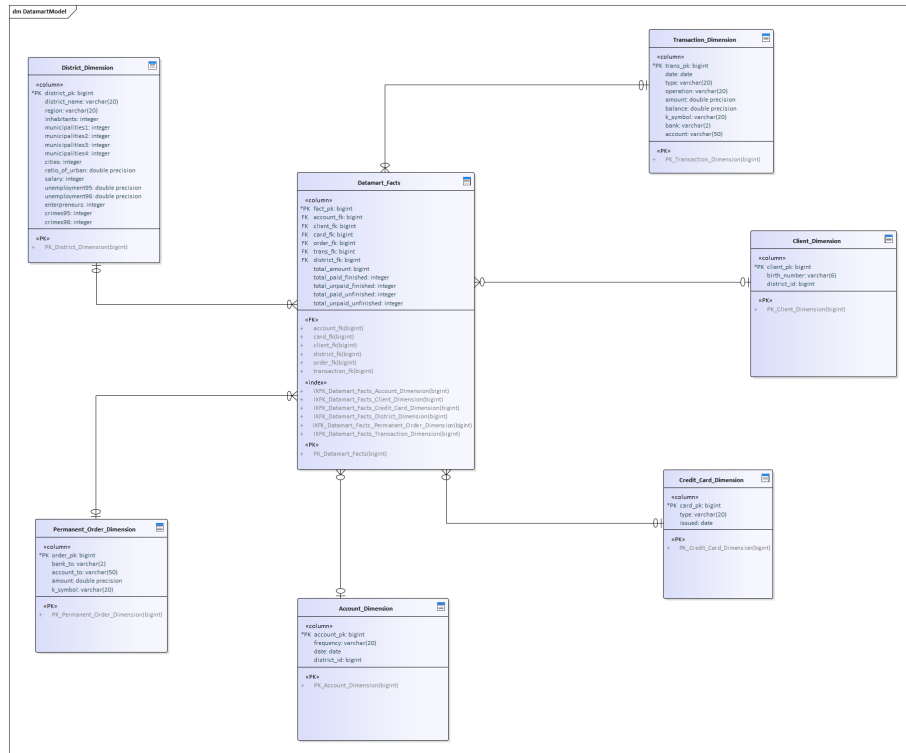
EDW Financial datamart - Business popis

Tomáš Chvosta, Petr Hanzl

Květen 2020

Na následujícím obrázku je zobrazen model datového tržiště použitý pro naši úlohu:

Obrázek 1: Datové tržiště - Financial



Můžeme si všimnout, že schéma datového tržiště je typu star a obsahuje jednu faktovou tabulku a šest dimenzionálních tabulek. Jednotlivé dimenze představují:

- **Account_Dimension** - dimenze týkající se účtu klienta
- **Client_Dimension** - dimenze týkající se klienta

- `Transaction_Dimension` - dimenze týkající se transakcí klienta
- `Credit_Card_Dimension` - dimenze týkající se kreditních karet
- `Permanent_Order_Dimension` - dimenze týkající se trvalých příkazů
- `District_Dimension` - dimenze týkající se okresů

Atributy jednotlivých dimenzí jsou popsány v 1. části semestrální práce v dokumentu s popisem dat. Názvy atributů jsou shodné se stage modelem. Jediný rozdíl je v názvech primárních klíčů, ty jsou v tomto modelu označeny pomocí „pk“.

Dále se ve schématu nachází faktová tabulka `Datamart_Facts`, která obsahuje následující atributy:

- `fact_pk` - uměle vytvořený primární klíč tabulky
- `account_fk` - cizí klíč do dimenze účtu
- `client_fk` - cizí klíč do dimenze klienta
- `card_fk` - cizí klíč do dimenze kreditní karty
- `order_fk` - cizí klíč do dimenze trvalého příkazu
- `trans_fk` - cizí klíč do dimenze transakce
- `district_fk` - cizí klíč do dimenze okresu
- `total_amount` - metrika představující celkové množství půjčených peněz
- `total_paid_finished` - metrika představující celkový počet splacených půjček s ukončenou smlouvou
- `total_unpaid_finished` - metrika představující celkový počet nesplacených půjček s ukončenou smlouvou
- `total_paid_unfinished` - metrika představující celkový počet půjček s právě probíhající smlouvou, kdy klient dodržuje splátkový kalendář
- `total_unpaid_unfinished` - metrika představující celkový počet půjček s právě probíhající smlouvou, kdy je klient v dluhu

V datovém tržišti jsou data napočítána pro lepší analýzu klientů a jejich půjček. Pomocí této analýzy můžeme jednodušeji kontrolovat klienty, kteří pro finanční instituci představují nějaké riziko, například mají příliš sjednaných půjček, ale žádnou z nich nedokáží splácet. Samozřejmě máme přehled i o klientech, kteří jsou naprosto bezproblémoví. Pomocí jednotlivých dimenzí můžeme rozšířit data obsažená ve faktových tabulkách o další popisné informace. Zároveň pomocí atributů, které jsou v dimenzích, můžeme filtrovat požadovaná data z faktových tabulek.