**03 Datamart + Dashboard**

Rozhodli jsme se data využít k zodpovězení těchto otázek:

1. Jací klienti jsou nejrizikovější, pokud budete převážně brát data z demografických údajů?
2. Jací klienti jsou nejrizikovější, pokud budete převážně brát data o transakcích?

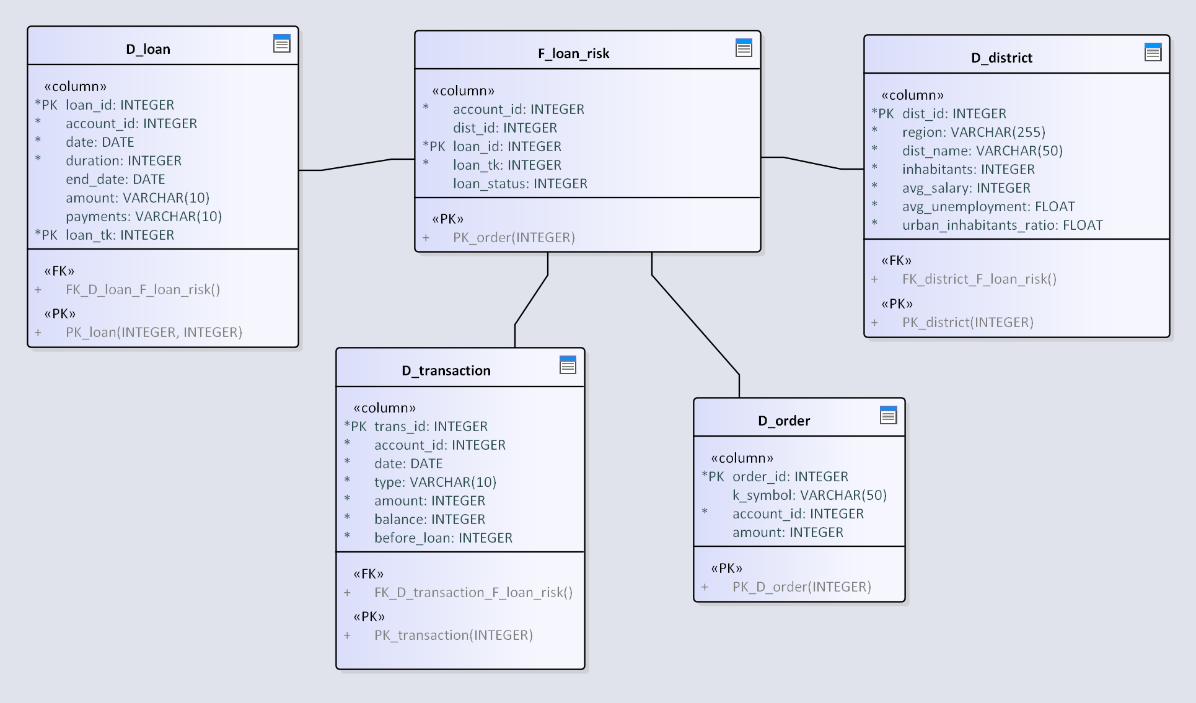
Datamart

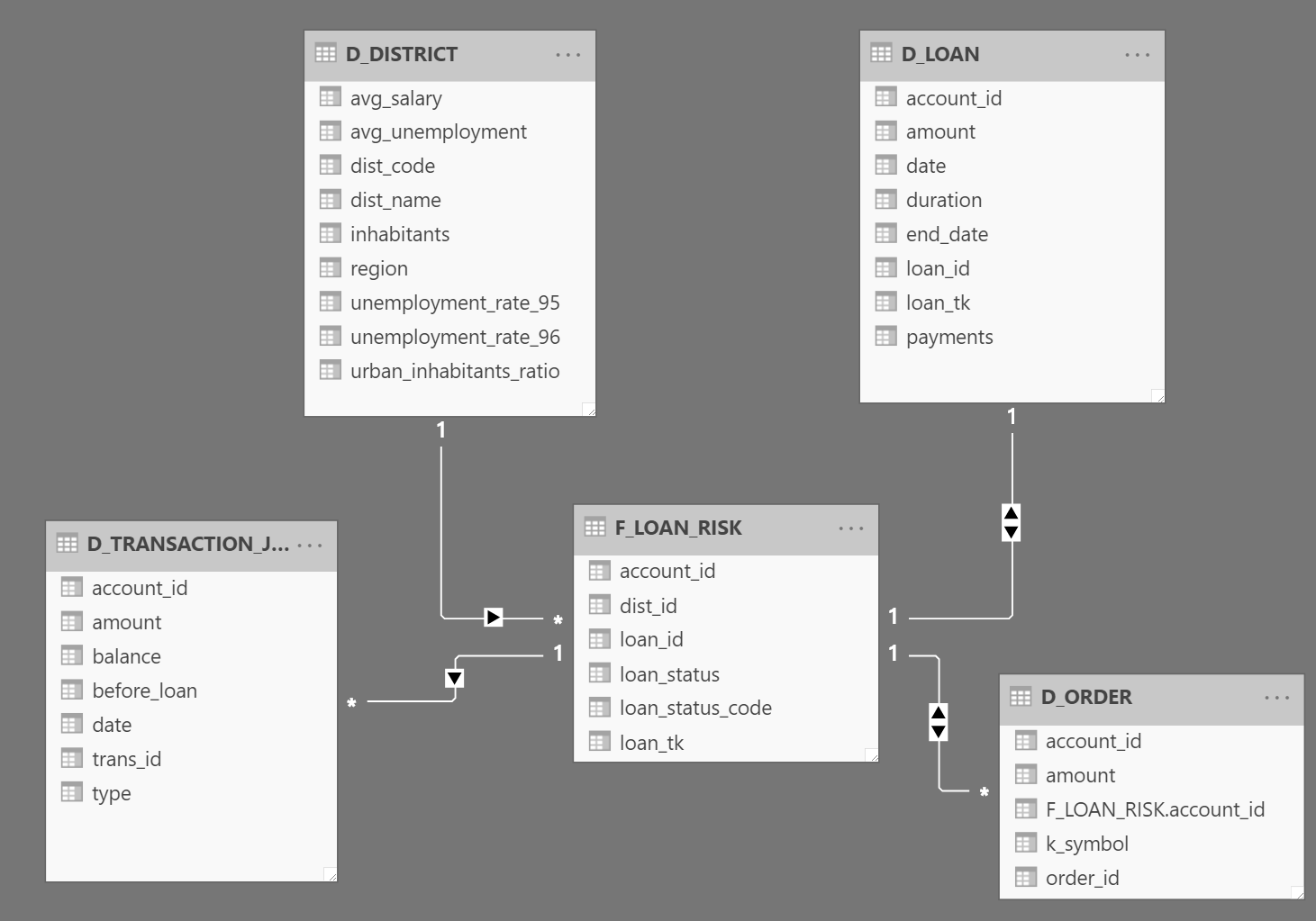
K zodpovězení otázek jsme potřebovali v datamartu udržet informaci o status každé půjčky a příslušné informace z tabulky democraphic\_data, transaction a případně i permanent\_order. K propojení tabulek jsme si ještě přidali identifikátor account\_id.

Dopočítali jsme také některé přídavné sloupce:

* avg\_unemplyment: průměrná nezaměstnanost v letech 95,96
* end\_date: datum ukončení půjčky
* before\_loan: indikátor zdali půjčka právě probíhá (ten sloupec by se mohl jmenovat i líp)

**Schéma:**





Dashboard

Demografická data



Data jsme pro lepší přehlednost zkoumali zejména podle regionů. Můžeme vidět, že největší podíl neúspěšných půjček, ať už ukončených nebo právě probíhajících je v regionech severní Moravy, západních a jižních Čech. Nezdá se, že by průměrná nezaměstnanost v region nebo zastoupení městských obyvatel mělo souvislost s úspěšností splácení půjčky. Naopak v grafu “Průměrný plat a výše splátek v regionech” je vidět mírný “spike” v průměrné výši měsíční splátky (tmavě modrá) právě v regionech s největšim poměrem neúspěšných půjček. Mezi těmito regiony je ale i Praha, proto jsme si do grafu vynesli i křivku průměrného platu v regionu (světle modrá). Zde už je vidět, že průměrná mzda v regionech severní Moravy, jižních a západních Čech neroste úměrně k výši splátky tak, jako v Praze.

Transakční data



Jak z tabulky tak z grafu “Průměrná částka a balance na účtu” si můžeme všimnout, že průměrná částka transakce se mezi různě úspěšnými dlužníky nemění, nicméně průměrná balance na účtu je pro úspěšné dlužníky výrazně vyšší. Zároveň platí, že úspěšní dlužníci mají v průměru nižší měsíční splátku, zatímco neúspěšní mají v průměru měsíční splátku vyšší než 5 tisíc.

Pokud se podíváme do levého dolního rohu na graf “Poměr typů transakce”, vidíme, že úspěšní dlužníci mají vyšší poměr INCOMING transakcí a také percentuálně méně výběrů. Zdálo by se tedy, že jednodušše mají průměrně vyšší příjmy. Pokud se podíváme hned vedle na graf “Průměrná částka transakce”, všimneme si, že průměrná výše příchozí částky není o mnoho vyšší pro neúspěšné dlužníky.

Když se podíváme na graf vpravo dole „Průměrná bilance účtu“, vidíme, že bilance neúspěšných dlužníků v době kdy nemají půjčku je sice nižší než bilance úspěšných dlužníků, ale pouze o cca 5 tisíc. To se ale změní po dobu trvání půjčky, kde už je v průměru rozdíl větší než 10 tisíc. To potvrzuje take plošný graf “Průměrná balance vs. splátka”, který ukazuje, že neúspěšní dlužníci mají v průměru neúměrně vyšší splátky ku průměrné bilanci účtu vůči úspěšným splátcům.

Toto potvrzuje také tabulka Key influencers, která ukazuje, že status FAIL u splátky můžeme očekávat v případě, že měsíční splátka je vyšší než 6930 a potom také pokud výrazně klesne průměrná bilance na účtu po transakci nebo pokud výrazně stoupne průměr částky transkace (odchozí transakce [OUTGOING, WTHDRAWAL] jsou v drtivé většině).

Pro informativnější výsledky v této tabulce doporučuji přepínat mezi typy transakcí a mezi časovým omezením transakcí ve filtru before\_loan.