**Banking Project Brief**

A retail bank’s services include managing of accounts, offering loans and credit cards. The objective of this project is to find opportunities to improve the bank’s services by finding interesting groups of customers (e.g. to differentiate between who may be good and bad customers) or interesting analysis of customer behaviour (e.g. actual or likely spending patterns).

The sort of info we are looking for is anything that identifies customers who may be valuable and profitable versus customers who may be a bad risk.

To support this analysis the bank stores data about their customers, account transactions, the loans already granted and the credit cards issued. They also have access to various pieces of demographic data about populations in the places where their customers live. By analysing this data we want you to provide any useful insights – **backed up with real data as evidence!** – into customer behaviour.

Download and import the data from this link:

<https://github.com/dportas1/shared/blob/master/bank_project.zip?raw=true>

The data about the customers and their accounts consist of following tables:

* **account** (4500 rows in the file ACCOUNT.ASC) - each row describes static characteristics of an account
* **client** (5369 rows in the file CLIENT.ASC) - each row describes a customer
* **disposition** (5369 rows in the file DISP.ASC) - each row relates a customer to account(s) which they hold
* **permanent order** (6471 rows in the file ORDER.ASC) - each row describes characteristics of a payment order
* **transaction** (1056320 rows in the file TRANS.ASC) - each row describes one transaction on an account
* **loan** (682 objects in the file LOAN.ASC) - each row describes a loan granted for an account
* **credit card** (892 objects in the file CARD.ASC) - each row describes a credit card issued to an account
* **demographic data** (77 objects in the file DISTRICT.ASC) - each row describes demographic characteristics of a district

Each account has both static characteristics (e.g. date of creation, address of the branch) given in the "account" table and dynamic characteristics (e.g. payments debited or credited, balances) given in the tables "permanent order" and "transaction". The "client" table describes characteristics of the customers using the accounts. One customer can have multiple accounts, several customers can hold a joint account; customers and accounts are related by the "disposition" table.

The "loan" and "credit card" tables show the loans made and cards issued to customers. Several credit cards can be issued per account. At most, one loan can be granted for an account.

The table "demographic data" gives some publicly available information about the districts where customers live (e.g. the unemployment rate). You can use this to derive any useful additional analysis of customers.

**Table: account**

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| account\_id | identification of the account |  |
| district\_id | location of the branch |  |
| date | date of creating of the account | in the form YYMMDD |
| frequency | frequency of issuance of statements | "POPLATEK MESICNE" stands for monthly issuance  "POPLATEK TYDNE" stands for weekly issuance  "POPLATEK PO OBRATU" stands for issuance after transaction |

**Table: client**

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| client\_id | row identifier |  |
| birth number | identification of client | the number is in the form YYMMDD for men, the number is in the form YYMM+50DD for women,  where YYMMDD is the date of birth |
| district\_id | address of the client |  |

**Table: disposition**

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| disp\_id | row identifier |  |
| client\_id | identification of a client |  |
| account\_id | identification of an account |  |
| type | type of disposition (owner/user) | only owner can issue permanent orders and ask for a loan |

**Table: permanent order**

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| order\_id | row identifier |  |
| account\_id | account, the order is issued for |  |
| bank\_to | bank of the recipient | each bank has unique two-letter code |
| account\_to | account of the recipient |  |
| amount | debited amount |  |
| K\_symbol | characterization of the payment | "POJISTNE" stands for insurrance payment  "SIPO" stands for household  "LEASING" stands for leasing  "UVER" stands for loan payment |

**Table: Transaction**

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| trans\_id | row identifier |  |
| account\_id | account, the transation deals with |  |
| date | date of transaction | in the form YYMMDD |
| type | +/- transaction | "PRIJEM" stands for credit  "VYDAJ" stands for withdrawal |
| operation | mode of transaction | "VYBER KARTOU" credit card withdrawal  "VKLAD" credit in cash  "PREVOD Z UCTU" collection from another bank  "VYBER" withdrawal in cash  "PREVOD NA UCET" remittance to another bank |
| amount | amount of money |  |
| balance | balance after transaction |  |
| k\_symbol | characterization of the transaction | "POJISTNE" stands for insurrance payment  "SLUZBY" stands for payment for statement  "UROK" stands for interest credited  "SANKC. UROK" sanction interest if negative balance  "SIPO" stands for household  "DUCHOD" stands for old-age pension  "UVER" stands for loan payment |
| bank | bank of the partner | each bank has unique two-letter code |
| account | account of the partner |  |

**Table: Loan**

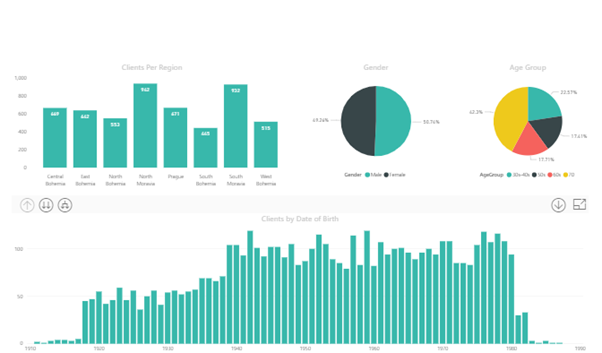
|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| loan\_id | row identifier |  |
| account\_id | identification of the account |  |
| date | date when the loan was granted | in the form YYMMDD |
| amount | amount of money |  |
| duration | duration of the loan |  |
| payments | monthly payments |  |
| status | status of paying off the loan | 'A' stands for contract finished, no problems,  'B' stands for contract finished, loan not payed,  'C' stands for running contract, OK so far,  'D' stands for running contract, client in debt |

**Table: Credit card**

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| card\_id | row identifier |  |
| disp\_id | disposition to an account |  |
| type | type of card | possible values are "junior", "classic", "gold" |
| issued | issue date | in the form YYMMDD |

**Table: Demographic data**

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| A1 = district\_id | district code |  |
| A2 | district name |  |
| A3 | region |  |
| A4 | no. of inhabitants |  |
| A5 | no. of municipalities with inhabitants < 499 |  |
| A6 | no. of municipalities with inhabitants 500-1999 |  |
| A7 | no. of municipalities with inhabitants 2000-9999 |  |
| A8 | no. of municipalities with inhabitants >10000 |  |
| A9 | no. of cities |  |
| A10 | ratio of urban inhabitants |  |
| A11 | average salary |  |
| A12 | unemploymant rate '95 |  |
| A13 | unemploymant rate '96 |  |
| A14 | no. of enterpreneurs per 1000 inhabitants |  |
| A15 | no. of commited crimes '95 |  |
| A16 | no. of commited crimes '96 |  |

****