**Financial Problem**

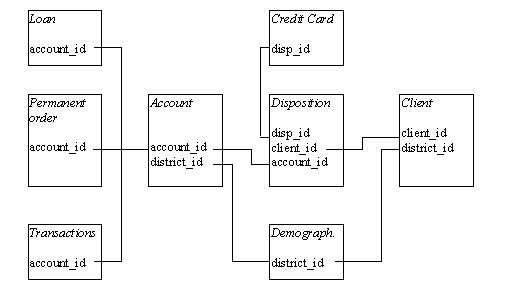
**Domain**

Once upon a time, there was a bank offering services to private persons. The services include managing of accounts, offering loans, etc.

**Task description**

The bank wants to improve their services. For instance, the bank managers have only vague idea, who is a good client (whom to offer some additional services) and who is a bad client (whom to watch carefully to minimize the bank loses). Fortunately, the bank stores data about their clients, the accounts (transactions within several months), loans already granted and credit cards issued. The bank managers hope to improve their understanding of customers and seek specific actions to improve services. A mere application of a discovery tool will not be convincing for them.

**Data description (database schema)**



The data about the clients and their accounts consist of following relations:

* relation **account** (4500 objects in the file ACCOUNT.ASC) - each record describes static characteristics of an account,
* relation **client** (5369 objects in the file CLIENT.ASC) - each record describes characteristics of a client,
* relation **disposition** (5369 objects in the file DISP.ASC) - each record relates together a client with an account i.e. this relation describes the rights of clients to operate accounts,
* relation **permanent order** (6471 objects in the file ORDER.ASC) - each record describes characteristics of a payment order,
* relation **transaction** (1056320 objects in the file TRANS.ASC) - each record describes one transaction on an account,
* relation **loan** (682 objects in the file LOAN.ASC) - each record describes a loan granted for a given account,
* relation **credit card** (892 objects in the file CARD.ASC) - each record describes a credit card issued to an account,
* relation **demographic data** (77 objects in the file DISTRICT.ASC) - each record describes demographic characteristics of a district.

Each account has both static characteristics (e.g. date of creation, address of the branch) given in relation "account" and dynamic characteristics (e.g. payments debited or credited, balances) given in relations "permanent order" and "transaction". Relation "client" describes characteristics of persons who can manipulate with the accounts. One client can have more accounts, more clients can manipulate with single account; clients and accounts are related together in relation "disposition". Relations "loan" and "credit card" describe some services which the bank offers to its clients; more credit cards can be issued to an account, at most one loan can be granted for an account. Relation "demographic data" gives some publicly available information about the districts (e.g. the unemployment rate); additional information about the clients can be deduced from this.

### Relation 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 |

### Relation client

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| client\_id | client identifier |  |
| birth number | birthday and sex | 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 |  |

### Relation disposition

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| disp\_id | record 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 |

### Relation permanent order (debits only)

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| order\_id | record 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 payment  "LEASING" stands for leasing  "UVER" stands for loan payment |

### Relation Transaction

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| trans\_id | record 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 |  |

### Relation Loan

|  |  |  |
| --- | --- | --- |
| **item** | **meaning** | **remark** |
| loan\_id | record 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 |

### Relation Credit card

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

### Relation 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 |  |

Some issues to consider

* Table Trans.asc: the most frequent mode in the file transactions isVYBER (withdrawal in cash). Among these transactions the most frequent k\_simbol is SLUZBY (payment for statement).
* The use of credit cards is poor.
* 30% of transactions are performed on the 30th or 31st of a month. This is due to interest which is credit only on the last day of a month.
* Remittances (money transfer) to other banks are only performed on the 5th, 6th, .., through the 14th of each month. On these days, the number of credits in cash is significant higher.
* The number of bank transfers (remittances) is increasing during each year. Th summit is in December.
* Check credit card promotion. Study how the bank can upgrade existing cardholders.
* Study credit cardholders behaviour in terms of money withdrawals.
* Study how loans are requested and paid back!
* Try to model a system for predicting credit card ownership. Use a score to measure client interest on owning a credit card. This will be used by marketing to guide client approach: top score get a personal visit, seconds a phone call, other a letter, etc etc.
* Study how costumers default in debt.
* High value loans tend to default more frequently.
* Contracts with less than 12 month tend to cause less problem than longer contracts.
* Check accounts and second authorize user and correlation to problems (default).
* Try to answer the following questions:
  + Which accounts are likely to repay their loans? Whom should the bank watch carefully to minimize losses from unpaid loans!?
  + Which costumers not having a credit card are likely to posses one?
  + Which costumers not using the service of payment orders are likely to use it? To whom to offer the service of payment orders?