

Data Mart Descriptive Report

EDUARDO ALFONSO RAZO, SUBHAMOY DAM, CHENXIN XIE

Introduction: The following approach has been employed while preparing the data mart. All the rows are representing a unique client id and its characteristics. Different tables have been summarized and been merged according to this methodology. Otherwise dummy variables have been added for categorical values in order to convert them into numerical. This makes the data mart ready to fit into any predictive model to derive further insights.

1. Variables in data mart

| No. | Variable Name | Values | Source table |
|-----|-----------------------|--|---|
| 1 | Client_id | Client identifier | Original from client table |
| 2 | Client_district_name | District name of the client | Link district_id from client table to A2(district name) of district table |
| 3 | Cilent_Region_Bohemia | Region of the client address is Bohemia | Link district_id from client table to A3(region name) of district table |
| 4 | Cilent_Region_Moravia | Region of the client address is Moravia | Link district_id from client table to A3(region name) of district table |
| 5 | Cilent_Region_Prague | Region of the client address is Prague | Link district_id from client table to A3(region name) of district table |
| 6 | Birth_year | Year of client born | Created from birth_number of client table |
| 7 | Birth_day | Day of month that the client was born | Created from birth_number of client table |
| 8 | Age | Age of the client | Created from birth_number of client table |
| 9 | Birth_month | Month of year that the client was born | Created from birth_number of client table |
| 10 | Is_Male | Gender of the client. 0 is female. 1 is male. | Created from birth_number of client table |
| 11 | disp_id | Record identifier of disposition of the client | Original from disposition table |
| 12 | account_id | Account identification of the client | Original from disposition table |
| 13 | Is_Disponent | Type of disposition. 1 means the client is the user of the account * only owner can issue permanent orders and ask for a loan | Created from the disposition table |
| 14 | Is_Owner | Type of disposition. 1 means the client is the owner of the account *only owner can issue permanent orders and ask for a loan | Create from the disposition table |
| 15 | Account_created_date | The date the client open the account | Formatted from account table |
| 16 | Account_district_name | The district name of the branch office where the client open the account | Link district_id from client table to A2(district name) of district table |
| 17 | Issuance_monthly | Frequency of issuance of statements | The account issues monthly |

| No. | Variable Name | Values | Source table |
|-----|----------------------------|--|--|
| 18 | Issuance_after_transaction | Frequency of issuance of statements | The account issues after transaction |
| 19 | Issuance_weekly | Frequency of issuance of statements | The account issuance weekly |
| 20 | Account_Region_Bohemia | Region of the branch office is Bohemia | Link district_id from account table to A3(region name) of district table |
| 21 | Account_Region_Moravia | Region of the branch office is Moravia | Link district_id from account table to A3(region name) of district table |
| 22 | Account_Region_Prague | Region of the branch office is Prague | Link district_id from account table to A3(region name) of district table |
| 23 | card_id | Record identifier of credit card | Original from Credit card table |
| 24 | Card_issued_date | The date the client issued the credit card | Original from Credit card table |
| 25 | Classic_card | Card type is classic | Created from Credit card table |
| 26 | Gold_card | Card type is Gold | Created from Credit card table |
| 27 | Junior_card | Card type is Junior | Created from Credit card table |
| 28 | loan_id | Record identifier of loan | Original from Loan table |
| 29 | Loan_granted_date | The date the loan was granted | Formatted from Loan table |
| 30 | Loan_amount | The amount of the loan granted | Transferred from Loan table |
| 31 | Loan_duration_by_months | Number of months that the loan lasts | Original from Loan table |
| 32 | Monthly_payments | Loan payment amount per month | Original from Loan table |
| 33 | Loan_Finished_OK | Loan-pay-off status: Contract finished, no problems | Transferred from Loan table |
| 34 | Loan_Finished_not_paid | Loan-pay-off status: Contract finished, loan not payed yet | Transferred from Loan table |
| 35 | Loan_Running_OK | Loan-pay-off status: Contract running, OK so far | Transferred from Loan table |
| 36 | Loan_Running_in_debt | Loan-pay-off status: Contract running, client in debt | Transferred from Loan table |
| 37 | order_id | Record identifier of order | Original from permanent order table |
| 38 | Order_Recipient_Bank_AB | The amount of order to recipient bank AB | Created from permanent order table |
| 39 | Order_Recipient_Bank_CD | The amount of order to recipient bank CD | Created from permanent order table |
| 40 | Order_Recipient_Bank_EF | The amount of order to recipient bank EF | Created from permanent order table |
| 41 | Order_Recipient_Bank_GH | The amount of order to recipient bank GH | Created from permanent order table |
| 42 | Order_Recipient_Bank_IJ | The amount of order to recipient bank IJ | Created from permanent order table |
| 43 | Order_Recipient_Bank_KL | The amount of order to recipient bank KL | Created from permanent order table |
| 44 | Order_Recipient_Bank_MN | The amount of order to recipient bank MN | Created from permanent order table |
| 45 | Order_Recipient_Bank_OP | The amount of order to recipient bank OP | Created from permanent order table |
| 46 | Order_Recipient_Bank_QR | The amount of order to recipient bank QR | Created from permanent order table |
| 47 | Order_Recipient_Bank_ST | The amount of order to recipient bank ST | Created from permanent order table |

| No. | Variable Name | Values | Source table |
|-----|------------------------------|---|------------------------------------|
| 48 | Order_Recipient_Bank_UV | The amount of order to recipient bank UV | Created from permanent order table |
| 49 | Order_Recipient_Bank_WX | The amount of order to recipient bank WX | Created from permanent order table |
| 50 | Order_Recipient_Bank_YZ | The amount of order to recipient bank YZ | Created from permanent order table |
| No. | Variable Name | Values | Source table |
| 51 | Order_Purpose_Others | The order is for other payment (excluding Household, Insurance, Leasing and Loan) | Created from permanent order table |
| 52 | Order_Purpose_HouseHold | The order is for household payment | Created from permanent order table |
| 53 | Order_Purpose_Insurance | The order is for insurance payment | Created from permanent order table |
| 54 | Order_Purpose_Leasing | The order is for leasing payment | Created from permanent order table |
| 55 | Order_Purpose_Loan | The order is for loan payment | Created from permanent order table |
| 56 | Total_Trans_in_1993 | Total amount of transactions the client made in 1993 | Created from Transaction table |
| 57 | Total_Trans_in_1994 | Total amount of transactions the client made in 1994 | Created from Transaction table |
| 58 | Total_Trans_in_1995 | Total amount of transactions the client made in 1995 | Created from Transaction table |
| 59 | Total_Trans_in_1996 | Total amount of transactions the client made in 1996 | Created from Transaction table |
| 60 | Total_Trans_in_1997 | Total amount of transactions the client made in 1997 | Created from Transaction table |
| 61 | Total_Trans_in_1998 | Total amount of transactions the client made in 1998 | Created from Transaction table |
| 62 | Credit_card_withdrawal | Transaction amount of withdrawal by credit card | Created from Transaction table |
| 63 | Credit_in_cash | Transaction amount of credit in cash | Created from Transaction table |
| 64 | Transfer_in | Transaction amount of transfer in | Created from Transaction table |
| 65 | Transfer_out | Transaction amount of transfer out | Created from Transaction table |
| 66 | Withdrawal_in_cash | Transaction amount of withdrawal in cash | Created from Transaction table |
| 67 | Total_Trans_of_Credit | Total transaction amount of credit | Created from Transaction table |
| 68 | Total_Trans_of_Withdrawal | Total transaction amount of withdrawal | Created from Transaction table |
| 69 | Trans_Type_HouseHold | Transaction for household payment | Created from Transaction table |
| 70 | Trans_Type_Insurance_Payment | Transaction for insurance payment | Created from Transaction table |
| 71 | Interest_credited | Transaction amount of interest credited | Created from Transaction table |
| 72 | Trans_Type_Loan_Payment | Transaction for loan payment | Created from Transaction table |
| 73 | Trans_Type_Old_Age_Pension | Transaction of old-age pension | Created from Transaction table |
| 74 | Trans_Type_Others | Transaction for other payment | Created from Transaction table |
| 75 | Trans_Type_Statement | Transaction for payment for statement | Created from Transaction table |

| No. | Variable Name | Values | Source table |
|-----|---|---|--|
| 76 | Trans_Type_interest_if_negative_balance | Transaction for insurance payment | Created from Transaction table |
| 77 | Trans_Partner_Bank_AB | The amount of transaction to partner bank AB | Created from Transaction table |
| 78 | Trans_Partner_Bank_CD | The amount of transaction to partner bank CD | Created from Transaction table |
| 79 | Trans_Partner_Bank_EF | The amount of transaction to partner bank EF | Created from Transaction table |
| 80 | Trans_Partner_Bank_GH | The amount of transaction to partner bank GH | Created from Transaction table |
| 81 | Trans_Partner_Bank_IJ | The amount of transaction to partner bank IJ | Created from Transaction table |
| 82 | Trans_Partner_Bank_KL | The amount of transaction to partner bank KL | Created from Transaction table |
| 83 | Trans_Partner_Bank_MN | The amount of transaction to partner bank MN | Created from Transaction table |
| 84 | Trans_Partner_Bank_OP | The amount of transaction to partner bank OP | Created from Transaction table |
| 85 | Trans_Partner_Bank_QR | The amount of transaction to partner bank QR | Created from Transaction table |
| 86 | Trans_Partner_Bank_ST | The amount of transaction to partner bank ST | Created from Transaction table |
| 87 | Trans_Partner_Bank_UV | The amount of transaction to partner bank UV | Created from Transaction table |
| 88 | Trans_Partner_Bank_WX | The amount of transaction to partner bank WX | Created from Transaction table |
| 89 | Trans_Partner_Bank_YZ | The amount of transaction to partner bank YZ | Created from Transaction table |
| 90 | num_of_inhabitants | Number of inhabitants of the client region | Original from Demographic data |
| 91 | num_of_municipalities_with_inhabitants_lt_499 | Number of municipalities with inhabitants < 499 of the region | Original from Demographic data |
| 92 | num_of_municipalities_with_inhabitants_bt_500_and_1999 | Number of municipalities with inhabitants 500-1999 of the region | Original from Demographic data |
| 93 | num_of_municipalities_with_inhabitants_bt_2000_and_9999 | Number of municipalities with inhabitants 2000-9999 of the region | Number of municipalities with inhabitants 500-1999 |
| 94 | num_of_municipalities_with_inhabitants_gt_10000 | Number of municipalities with inhabitants >10000 of the region | Original from Demographic data |
| 95 | num_of_cities | Number of cities of the region | Original from Demographic data |
| 96 | ratio_of_urban_inhabitants | Ratio of urban inhabitants of the region | Original from Demographic data |
| 97 | average_salary_of_region | Average salary of the region | Original from Demographic data |
| 98 | unemployment_rate_of_1995 | Unemployment rate of 1995 of the region | Original from Demographic data |
| 99 | unemployment_rate_of_1996 | Unemployment rate of 1996 of the region | Original from Demographic data |
| 100 | num_of_entrepreneurs_per_1000_inhabitants | Number of entrepreneurs per 1000 inhabitants | Original from Demographic data |
| 101 | num_of_committed_crimes_of_1995 | Number of committed crimes of 1995 | Original from Demographic data |

| No. | Variable Name | Values | Source table |
|-----|-----------------------------------|--|---|
| 102 | num_of_committed_crimes_of_1996 | Number of committed crimes of 1996 | Original from Demographic data |
| 103 | Different_client_account_district | Client open the account at the district outside his/her address district. 0 is no. 1 is yes. | Created from the client table and account table |

2. Summary

The data mart can be summarized to gather and visualize different observations or characteristics. Some of sample observations have been highlighted below.

2.1 Popularity of card type among different age groups

This graph below shows the distribution of the type of credit cards used by clients across different age groups. This data can be used to design specific marketing strategies focused on these age groups. It also gives an idea about the spending behavior according to age groups as Credit card limit varies (along with other facilities across different card types).

Popularity of Cards among different age groups

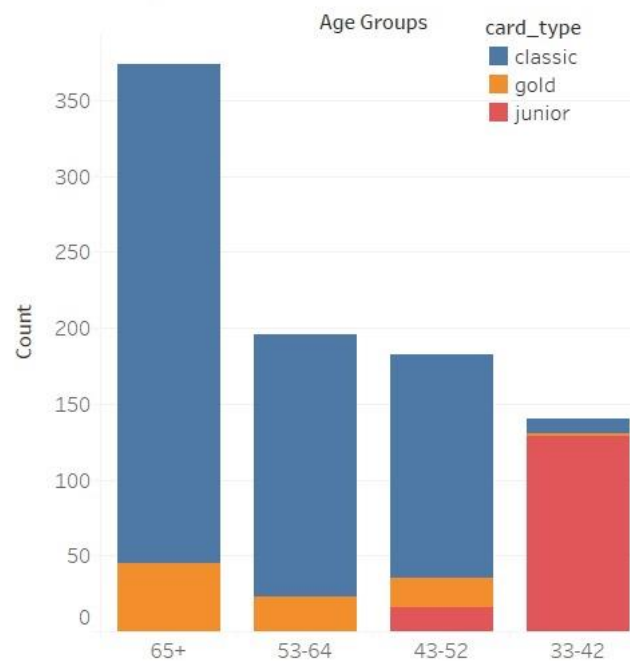


Figure-1

Figure-1 shows: for different age groups have their own preference credit card type:

- Clients with age over 65 prefer to open the classic cards.
- Clients with age over 52 do not open the junior cards.
- Most of the gold cards own by the clients over 65 years old.

2.2 Client region and age

The following figure allows us to understand the breakup of clients on the basis of their age group across the three main regions of the banks operation. This information can be used to customize facilities at the banks offices in these regions to handle such customers. Some key take away points are highlighted below the figure.



Figure-2

Figure-2 shows the age distribution of the three regions:

- Most of the clients are above 52 years old.
- Most of the clients live in Bohemia among all age groups.

2.3 Loan status of region

This is a very important visualization which shows the status of Loan accounts across three different regions. It allows for the bank to take preventive measures regarding loan accounts which are degrading and have a high probability of defaulting or becoming Non Performing Assets (NPAs).

It also gives us information regarding regions with maximum defaults. Hence there should be better scrutiny and inspection before loan sanctioning of clients in these regions. It also gives clear understanding about the percentage of loan accounts which are performing as expected and generating income in terms of interest compared to accounts which may default or not performing well which may lead to loss for the bank.

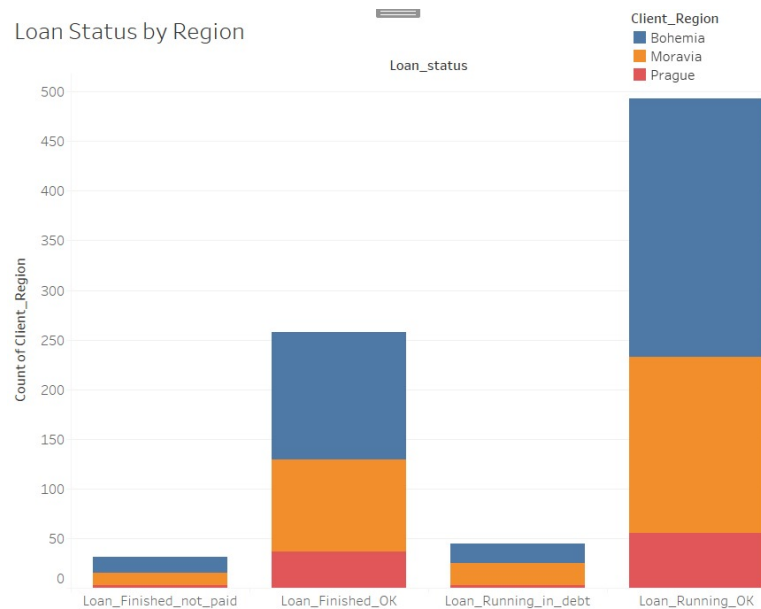


Figure-3

Figure-3 shows different loan contact status with regions:

- Most of the clients have loan contact running and their loan payment are undergoing but not so much to concern about.
- Most of the clients live in Bohemia among all different loan status.
- For region Moravia, most of the clients having their loan contract running and doing OK.

2.4 District difference between clients' addresses and account

The following graph shows the top 5 districts from where clients are opening accounts in a different districts. Further investigation may reveal opportunities which the bank might target to get more clients. For example: if clients are holding accounts in a different district because they are not satisfied the service in their home district, then the bank can take necessary actions to correct this as there might be many churners because of this reason.

TOP 5 Districts with clients having accounts outside their home district

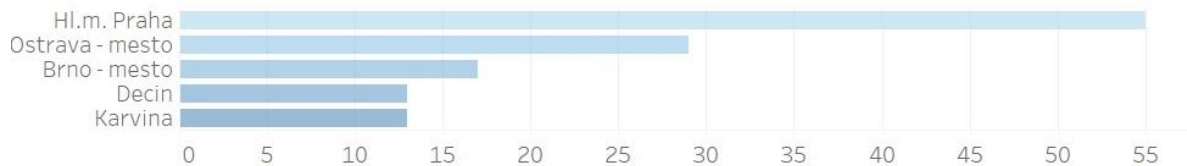


Figure-4

Figure-4 shows on each district, how many clients open their account at the other district:

Hl.m.Praha is the most popular district where the account office is that the clients open their accounts there instead of their address district.

2.5 Unemployment rate among districts in 1995 and 1996

The following graph shows the variation in the unemployment rate across districts in the year 1995-96. This data can help the bank to identify districts where the rate of unemployment is increasing or decreasing as this may impact the clients of those districts and also can be a source of income leakage for the bank.

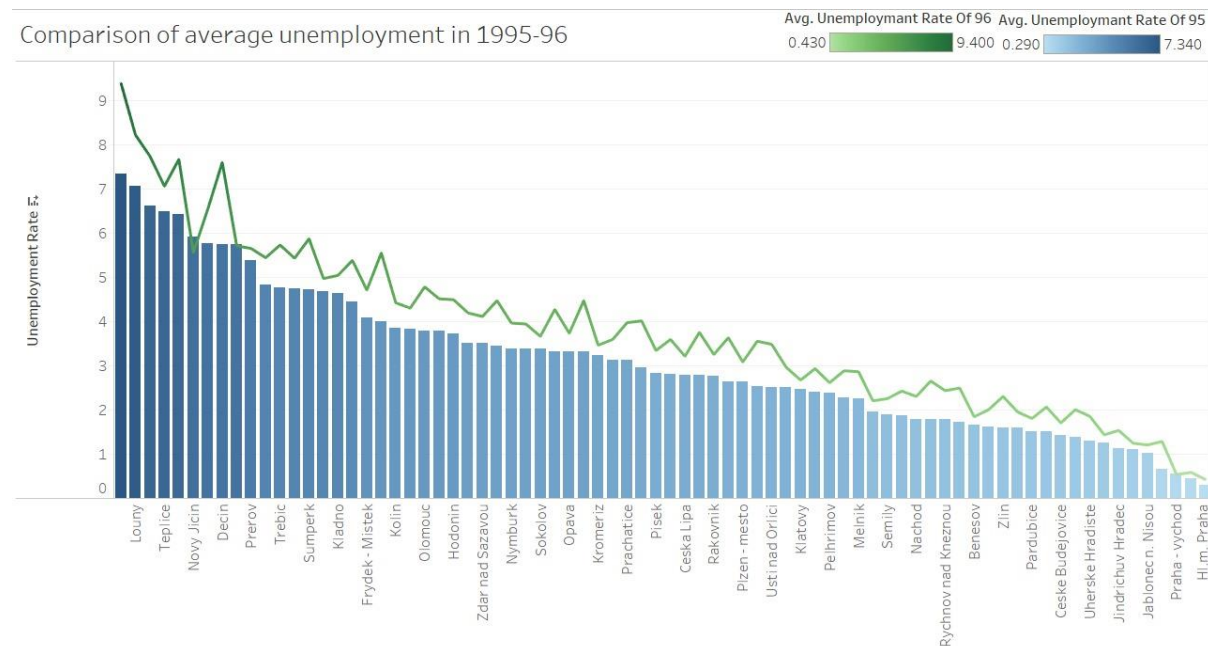


Figure-5

Figure-5 shows the unemployment rate of different districts between 1995 and 1996:

Hl.m.Praha has the lowest unemployment rate among all districts in 1995 and 1996, and that could be related to why Hl.m.Praha is the district outranks other districts that the clients open their account there instead of their living district.

3. Conclusion

We can conclude that having a well-structured data mart provides a competitive advantage to obtain several insights to improve the performance of the brand in the analyzed region. For example, we can add value by leveraging the information to train a model that will be able to determine the likelihood of default of loans based on the different information related to each client. Also, maybe targeting specific profiles to increase the activity related to credit cards and the transaction activity.