Private Debt Portfolios & Customer Behavior Analysis in Banking and Finance 🏦

Business Understanding

⇩

Problem Statement

⇩

Data Collection

⇩

Data Cleansing

⇩

Data Analysis

⇩

Recommendations

Business Understanding

Introduction

A-Credit, an *international non-bank financial institution* founded and headquartered in 2000 in the Czech Republic and headquartered in the Netherlands, operates in 7 countries.

Business Understanding

Business Model

A-Credit is primarily focused on:

1. Centered around providing financial assistance to an **underserved segment** of the population.

2. Offering loans to those who may **not have access to traditional banking services** due to their limited credit history.

3. Lending to individuals with little or **no credit history.**

Business Understanding

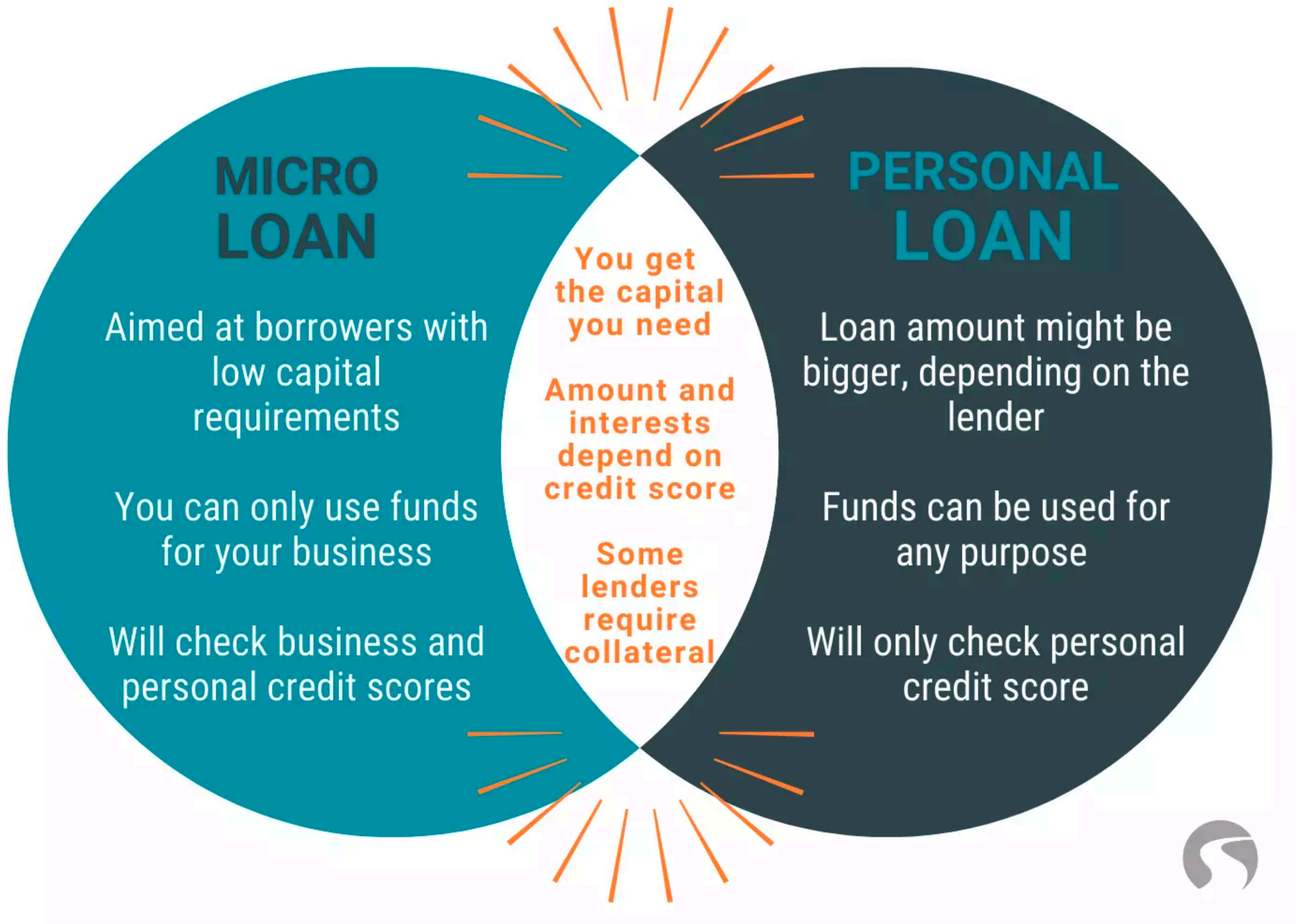
Business Model

Target Customer Segment: A-Credit primarily targets individuals with limited or no credit history, including young professionals, recent immigrants, and others who may not meet the typical credit requirements of traditional banks.

Business Understanding

Business Model

Loan Products: The company offers a range of loan products, such as **personal loans**, **microloans**, and **other financial instruments tailored** to meet the diverse financial needs of its customers.



Business Understanding

Business Model

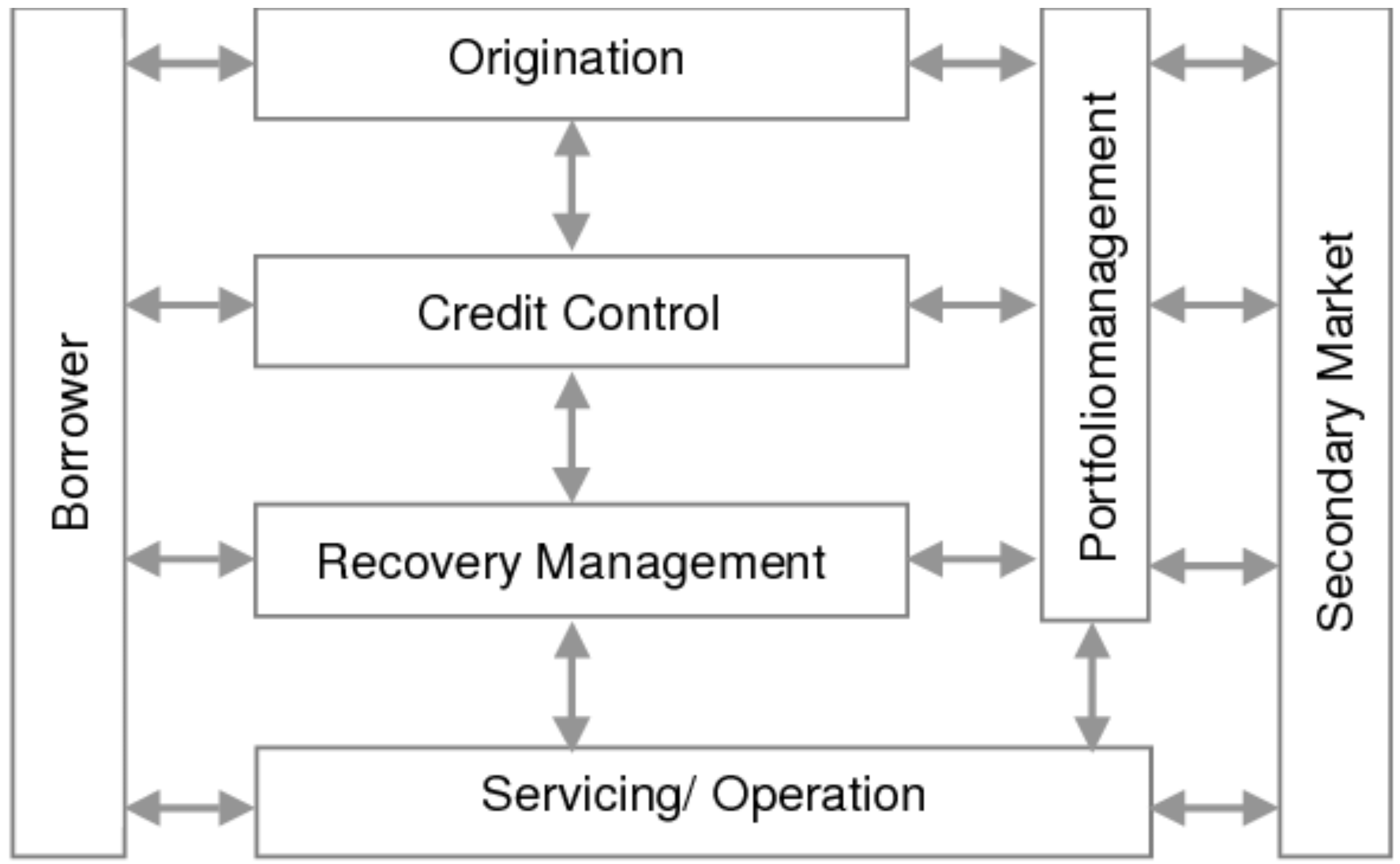
Risk Assessment: Given the unique customer segment, A-Credit employs a **specialized risk assessment process** to **determine creditworthiness**. This involves considering alternative data sources and proprietary **scoring models** to evaluate applicants.

Problem Statement

Context Analysis

Underwriting Department requires the Data Analyst team to observe focused Private Debt Portfolio & Customer Behavior Analysis within 18 months (here assumed from March 2018 to March 2020).

To understand the context of the problem that needs to be analyzed and to outline the requirements of the business stakeholders, it's essential to consider the following aspects: (1) Credit Risk Assessment, (2) Customer Segmentation, (3) Payment Behavior Analysis & (4) Data-Driven Decision Making.



Data Collection

SEANNY · UPDATED 4 YEARS AGO

725 New Notebook Download (6 MB)

Credit Card Approval Prediction

A Credit Card Dataset for Machine Learning

Data Card Code (131) Discussion (11)

About Dataset

A Credit Card Dataset for Machine Learning!

Don't ask me where this data come from, the answer is I don't know!

Context

Credit score cards are a common risk control method in the financial industry. It uses personal information and data submitted by credit card applicants to predict the probability of future defaults and credit card borrowings. The bank is able to decide whether to issue a credit card to the applicant. Credit scores can objectively quantify the magnitude of risk.

Generally speaking, credit score cards are based on historical data. Once encountering large economic fluctuations. Past models may lose their original predictive power. Logistic model is a common method for credit scoring. Because Logistic is suitable for binary classification tasks and can calculate the coefficients of each feature. In order to facilitate understanding and operation, the score card will multiply the logistic regression coefficient by a certain value (such as 100) and round it.

Usability
10.00

License
CC0: Public Domain

Expected update frequency
Annually

Tags
Business
Computer Science
Finance
Beginner

Credit Card Approval Prediction (by Seanny, 2020) [URL]

Data Collection

Column Name	Data Type	Description
ID	int	A unique identifier for each client. This column distinguishes each applicant from others and allows for tracking and referencing client data.
CODE_GENDER	nvarchar(50)	This column specifies the gender of the applicant, with "M" indicating male and "F" indicating female.
FLAG_OWN_CAR	bit	Indicates whether the applicant owns a car, with "Y" for Yes and "N" for No.
FLAG_OWN_REALTY	bit	Specifies whether the applicant owns real estate property, with "Y" for Yes and "N" for No.
CNT_CHILDREN	tinyint	Represents the count of children that the applicant has, providing insight into their family situation.
AMT_INCOME_TOTAL	float	Denotes the total annual income of the applicant, which is an important factor in assessing creditworthiness.
NAME_INCOME_TYPE	nvarchar(50)	Describes the category or source of the applicant's income, such as "Working," "Pensioner," or other income types.
NAME_EDUCATION_TYPE	nvarchar(50)	Indicates the highest level of education attained by the applicant, providing insight into their educational background.
NAME_FAMILY_STATUS	nvarchar(50)	Specifies the marital status of the applicant, such as "Single," "Married," or other statuses.
NAME_HOUSING_TYPE	nvarchar(50)	Provides information about the applicant's housing situation, including whether they rent or own their residence.
DAYS_BIRTH	smallint	Represents the applicant's age by counting backward from the current day (0 is the current day, -1 means yesterday).
DAYS_EMPLOYED	int	Indicates the start date of the applicant's employment, counting backward from the current day (0). If positive, it means the person is currently unemployed.
FLAG_MOBIL	tinyint	Specifies whether the applicant has a mobile phone, with "1" for Yes and "0" for No.
FLAG_WORK_PHONE	bit	Indicates whether the applicant has a work phone, with "1" for Yes and "0" for No.
FLAG_PHONE	bit	Specifies whether the applicant has a personal phone, with "1" for Yes and "0" for No.
FLAG_EMAIL	bit	Indicates whether the applicant has an email address, with "1" for Yes and "0" for No.
OCCUPATION_TYPE	nvarchar(50)	Describes the applicant's occupation or job type, providing information about their employment.
CNT_FAM_MEMBERS	float	Represents the count of family members in the applicant's household, which is relevant for assessing their financial responsibilities.
ID	int	A unique identifier for each client, linking their credit history to their application and personal details in the "application_record" dataset.
MONTHS_BALANCE	smallint	Indicates the month of the extracted data as the starting point, counting backward. In this column, "0" represents the current month, "-1" is the previous month, and so on. This allows for tracking credit behavior over time.
STATUS	nvarchar(50)	Describes the client's credit status for a specific month. The column uses codes to represent different credit statuses: 0: Indicates on-time payments made within 1-29 days after the payment due date. 1: Signifies payments made within 30-59 days after the payment due date. 2: Represents payments made within 60-89 days after the payment due date. 3: Denotes payments made within 90-119 days after the payment due date. 4: Refers to payments made within 120-149 days after the payment due date. 5: Indicates overdue or bad debt payments, including write-offs for over 150 days. C: Suggests that clients pay off their credit card debt in full before or on the payment due date. X: Implies that there was no credit card usage during the month, resulting in a debt amount of zero.

Credit Card Approval Prediction's **Data Dictionary**

Data Cleansing

Run Cancel Disconnect Change Database: GROUP2 Estimated Plan Enable Actual Plan Parse Enable SQLCMD To Notebook

```sql
SELECT COUNT(DAYS_EMPLOYED)
FROM application_record
where DAYS_EMPLOYED = 365243

UPDATE application_record
SET DAYS_EMPLOYED = null
where DAYS_EMPLOYED = 365243

ALTER TABLE application_record
ADD CLIENT_AGE int

UPDATE application_record
SET CLIENT_AGE = REPLACE(DAYS_BIRTH , '-', '')


ALTER TABLE application_record
ADD CLIENT_EMPLOYED int

UPDATE application_record
SET CLIENT_EMPLOYED = REPLACE(DAYS_EMPLOYED , '-', '')


UPDATE application_record
set CLIENT_AGE = CLIENT_AGE/365

UPDATE application_record
set CLIENT_EMPLOYED = CLIENT_EMPLOYED/365
```

Results Messages

DAYS_BIRTH	DAYS_EMPLOYED	FLAG_MOBIL	FLAG_WORK_PHONE	FLAG_PHONE	FLAG_EMAIL	OCCUPATION_TYPE	CNT_FAM_MEMBE
-21688	NULL	1	0	1	0	NULL	1
-21688	NULL	1	0	1	0	NULL	1
-21688	NULL	1	0	1	0	NULL	1
-21688	NULL	1	0	1	0	NULL	1
-23299	NULL	1	0	1	0	NULL	1
-23299	NULL	1	0	1	0	NULL	1
-23299	NULL	1	0	1	0	NULL	1
-23299	NULL	1	0	1	0	NULL	1

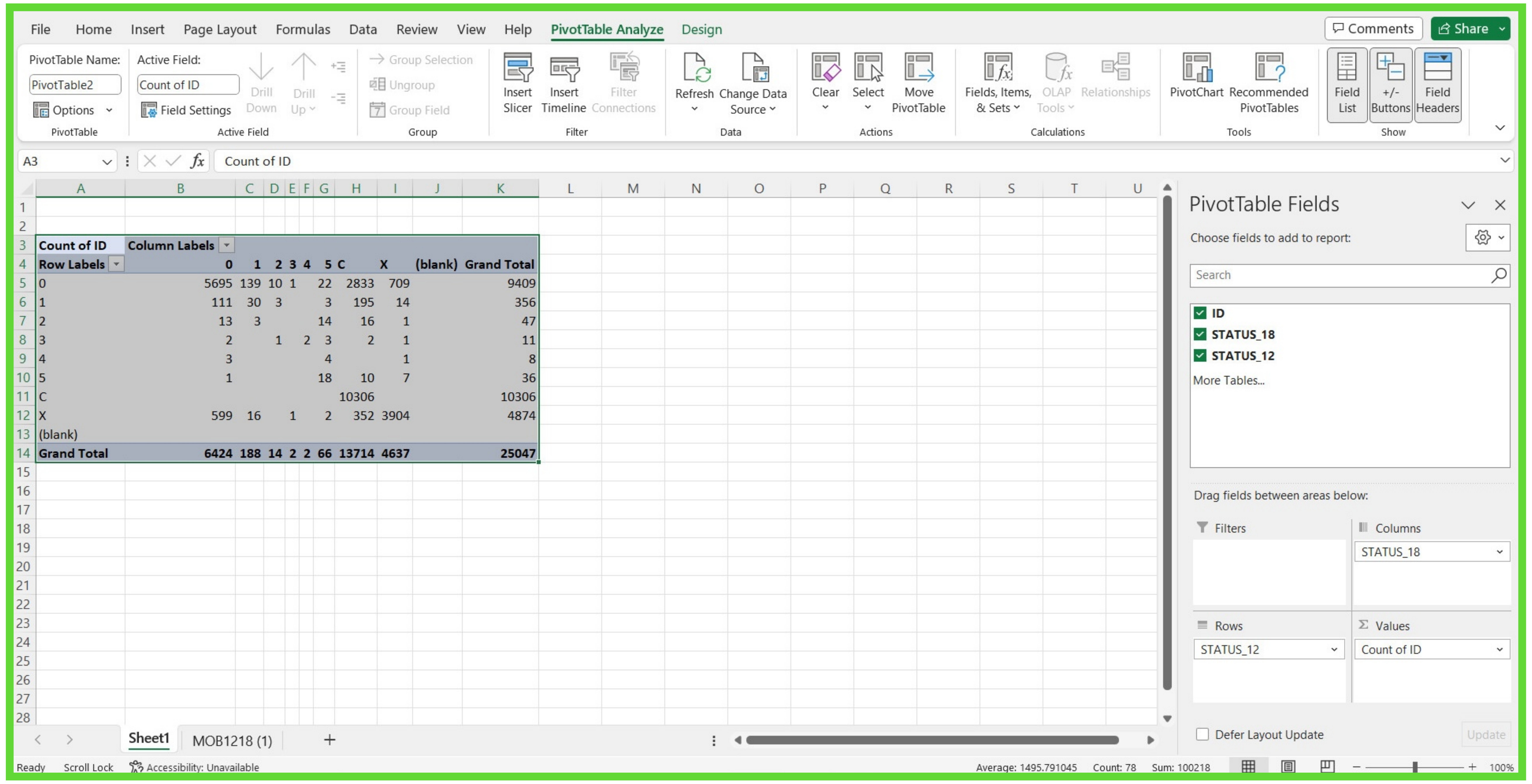
O	P	Q	R	S	T
OCCUPATION_TYPE	CNT_FAM_MEMBERS	DAYS_BIRTH	DAYS_EMPLOYED	CLIENT_AGE	CLIENT_EMPLOYED
NULL	1	-21688	NULL	59	NULL
NULL	1	-21688	NULL	59	NULL
NULL	1	-21688	NULL	59	NULL
NULL	1	-21688	NULL	59	NULL
NULL	1	-23299	NULL	63	NULL
NULL	1	-23299	NULL	63	NULL
NULL	1	-23299	NULL	63	NULL
NULL	1	-23299	NULL	63	NULL
NULL	1	-23299	NULL	63	NULL
Managers	2	-12999	-1560	35	4
Managers	2	-12999	-1560	35	4
Managers	2	-12999	-1560	35	4
Managers	2	-12999	-1560	35	4
Managers	2	-12999	-1560	35	4
Managers	2	-12999	-1560	35	4
Managers	2	-12999	-1560	35	4
NULL	2	-20863	NULL	57	NULL
NULL	2	-20863	NULL	57	NULL
NULL	2	-20863	NULL	57	NULL
NULL	2	-20863	NULL	57	NULL
NULL	2	-20863	NULL	57	NULL
Managers	2	-11644	-344	31	0
Managers	2	-11644	-344	31	0
Managers	2	-11644	-344	31	0
Managers	2	-11644	-344	31	0
Sales staff	2	-11102	-433	30	1
Sales staff	2	-11102	-433	30	1
NULL	1	-9380	-1713	25	4
NULL	1	-9380	-1713	25	4

1. Replace irrelevant data with 'NULL' values.
2. Manipulate 'DAYS_BIRTH' column (in the 'application_record' table) to 'CLIENT_AGE' column.
3. Manipulate 'DAYS_EMPLOYED' column (in the 'application_record' table) to 'CLIENT_EMPLOYED' column.

Data Cleansing

ID	MONTHS_BALANCE	STATUS	STATUS_DEFINE
5001711	0	X	No loan for the month
5001711	-1	0	1-29 days past due
5001711	-2	0	1-29 days past due
5001711	-3	0	1-29 days past due
5001712	0	C	Paid off that month
5001712	-1	C	Paid off that month
5001712	-2	C	Paid off that month
5001712	-3	C	Paid off that month
5001712	-4	C	Paid off that month
5001712	-5	C	Paid off that month
5001712	-6	C	Paid off that month
5001712	-7	C	Paid off that month
5001712	-8	C	Paid off that month
5001712	-9	0	1-29 days past due
5001712	-10	0	1-29 days past due
5001712	-11	0	1-29 days past due
5001712	-12	0	1-29 days past due
5001712	-13	0	1-29 days past due
5001712	-14	0	1-29 days past due
5001712	-15	0	1-29 days past due
5001712	-16	0	1-29 days past due
5001712	-17	0	1-29 days past due
5001712	-18	0	1-29 days past due
5001713	0	X	No loan for the month
5001713	-1	X	No loan for the month
5001713	-2	X	No loan for the month

ID	STATUS_18	STATUS_12
5001712	C	C
5001713	X	X
5001715	X	X
5001717	C	0
5001718	X	0
5001719	C	C
5001720	0	0
5001723	X	X
5001724	C	C
5001726	C	C
5001730	C	C
5001732	X	X
5001735	C	C
5001736	C	C
5001737	C	0
5001738	C	0
5001739	C	0
5001742	C	C
5001743	C	X
5001744	C	C
5001747	C	0
5001748	C	C
5001756	C	C
5001757	1	X
5001758	0	0
5001759	C	0



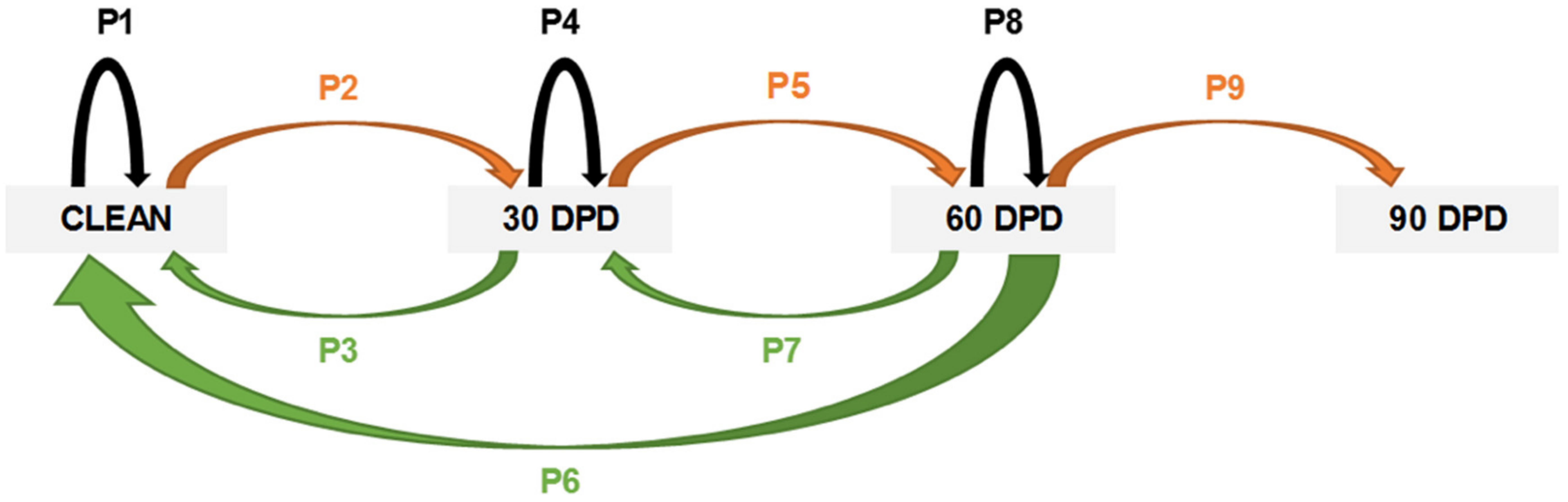
Count of ID	Column Labels									
Row Labels	0	1	2	3	4	5	C	X	(blank)	Grand Total
0	5695	139	10	1		22	2833	709		9409
1	111	30	3			3	195	14		356
2	13	3				14	16	1		47
3	2		1		2	3	2	1		11
4	3					4		1		8
5	1					18	10	7		36
C							10306			10306
X	599	16		1		2	352	3904		4874
(blank)										
Grand Total	6424	188	14	2	2	66	13714	4637		25047

4. Manipulate 'STATUS' column (in the 'credit_record' table) to 'STATUS_DEFINE' column.
5. Pivot 'MONTHS_BALANCE' of Customer to a MOB12 & MOB18 Table.
6. Pivot 'MOB12 & MOB18' table to a 'Roll Rate Analysis' table.

Data Analysis via Roll Rate Analysis

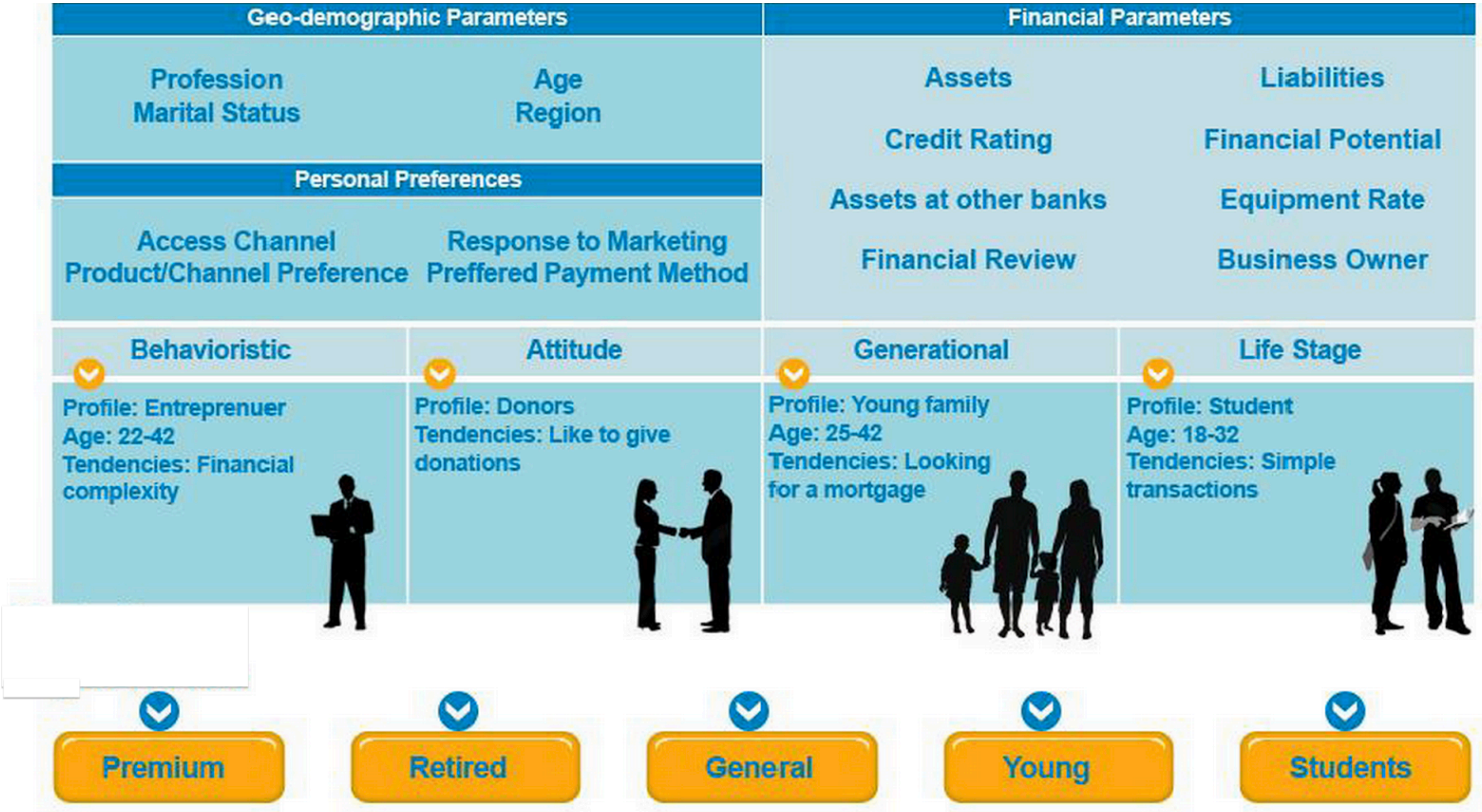
		After 12 Months on Book (MoB)								
	Delinquency	No Due	0-29 DPD	30-59 DPD	60-89 DPD	90-119 DPD	120-149 DPD	150+ DPD	No Loan	Grand Total
After 12 Months on Book (MoB)	No Due	10306								10306
	0-29 DPD	2833	5695	139	10	1		22	709	9409
	30-59 DPD	195	111	30	3			3	14	356
	60-89 DPD	16	13	3				14	1	47
	90-119 DPD	2	2		1		2	3	1	11
	120-149 DPD		3					4	1	8
	150+ DPD	10	1					18	7	36
	No Loan	352	599	16		1		2	3904	4874

		After 12 Months on Book (MoB)									
	Delinquency	No Due	0-29 DPD	30-59 DPD	60-89 DPD	90-119 DPD	120-149 DPD	150+ DPD	No Loan	Roll Back	Roll Forward
After 12 Months on Book (MoB)	No Due	100%	0%	0%	0%	0%	0%	0%	0%		
	0-29 DPD	30%	61%	1%	0%	0%	0%	0%	8%	30%	9%
	30-59 DPD	55%	31%	8%	1%	0%	0%	1%	4%	86%	6%
	60-89 DPD	34%	28%	6%	0%	0%	0%	30%	2%	68%	32%
	90-119 DPD	18%	18%	0%	9%	0%	18%	27%	9%	45%	54%
	120-149 DPD	0%	38%	0%	0%	0%	0%	50%	13%	38%	63%
	150+ DPD	28%	3%	0%	0%	0%	0%	50%	19%	31%	19%
	No Loan	7%	12%	0%	0%	0%	0%	0%	80%	19%	0%



Data Analysis via Visualizations

Recommendations



Recommendations

Internal

Debt collection often starts with the company you owe. Their internal collection team (if applicable) will contact you about the debt, leading up to the charge-off stage.

Third-Party

Third-party collectors will take over the account on behalf of the original creditor. This group represents "traditional" collection as most of us think of it. This group is regulated by the FDCPA.

Debt Buyers

Debt buyers purchase old accounts in portfolios and even in online auctions at pennies on the dollar. They continue to collect the debt, sometimes even after the statute of limitations has passed.

Thanks for Listening

Have a great weekend!