National Research University Higher School of Economics
Faculty of Computer Science
HSE and University of London Double Degree Programme in
Data Science and Business Analytics

TERM PAPER

Research Project

Construction of a report on factor changes in the provision for losses on bank loans

Prepared by the student: 182, in Year: 3, Krauze Natalia
Term Paper Supervisor: Vice-President, Deputy Director of the IT Development Department, Pervukhin Dmitriy
Advisor: DWH Head of IT Development, Ivan Zipukho

Moscow
2021

Table of contents

1. Introduction

1.1 Abstract
1.2 List of key words
1.3 Instruments
1.4 Main result

2. Review and comparative analysis

2.1 Review of abstracts which were used for project and review of analogues

3. Selection of methods, algorithms, and models for project implementation

3.1 Modelling of simple situation
3.2 Calculating balance on different dates
3.3 Work with small example
3.4 Study the economic model and derive formulas for calculating the reserve change factors
3.5 Study the data table
3.6 Develop a calculation methodology and algorithm, draw up a technical requirement for implementation
3.7 Develop calculations and report on the technical requirement

4. Results and conclusion

4.1 Analyze the report and make economic conclusions

5. List of sources

1. Introduction:

1.1 Abstract

If your company has a large amount of trade debtors, then there are likely to be some bad debts hidden among the “good” debts and you should make a provision for them.

When a person or entity deposits money in a bank, a certain percentage of this amount must be transferred to a special account at the Central Bank as a reserve and kept there until the money is withdrawn from the bank. Financial institutions are allowed to use their clients' money to make loans. From the available reserves, the Central Bank provides loans to maintain liquidity. They are thereby used in a situation where there is an urgent need to repay the money.

The CBR is one of the monetary policy instruments used by the Central Bank to regulate the amount of money in circulation. In Russia, reserve requirements in accordance with Federal Law No. 86-FZ of July 10, 2002 are determined by the Bank of Russia.

The purpose of this project is to develop an algorithm laying down the changes in the loan loss provisions for the influencing factors: changes in change in accounts payable, changes in currency exchange rates, changes in the rate of the rate of the allowance for possible losses, and other factors.

1.2 List of key words

- Mandatory provision for losses of commercial banks - are funds that credit institutions must hold as a mandatory reserve in a correspondent account with central banks.
- Basel 2 - The "International Convergence of Capital Measurement and Capital Standards: New Approaches," published by the Basel Committee on Banking Supervision in 2004, offers guidelines in the field of banking regulation.
- IFRS 9 - International Financial Reporting Standard, an international financial reporting standard that establishes financial reporting procedures for financial assets and financial liabilities, recognition and measurement requirements, impairment, derecognition of general hedging arrangements, principles for recognizing expected credit losses over the life of a financial instrument, effective 01.01.2018.
- Assets - is an economically valuable property that a person, company, or country owns or manages with the hope of future profit.
- Warranties - is a warranty given by a retailer or other third party about the quality of a commodity. It also applies to the conditions and circumstances in which the device can be repaired or exchanged if it does not work as defined or planned.
- Liabilities - a debt owed by an individual or a corporation, normally in the form of income. Liabilities are resolved over time by exchanging economic benefits such as capital, commodities, or services.
- The loan portfolio - is the aggregate of the outstanding principal balances of active credit transactions at a given date.
- Expected credit loss (ECL) - is a probability-weighted measure of collateral defaults over the projected life of a Financial Instrument (i.e., the current value of all cash shortfalls). In the sense of IFRS 9, the definition is especially relevant.

1.3 Instruments

1) Dataset on credit data on individuals (The Bank provides six Excel files with data on provisions for possible losses (for six reporting dates)
2) Python, Anaconda, Google Collaboratory
3) Excel
4) Credit calculator in excel
5) CBR parser for currencies

1.4 Main result

The aim of this project is to create an algorithm that calculates adjustments in loan loss conditions based on influencing variables such as changes in accounts payable, currency exchange rates, and the rate of the allowance for future defaults, among other things.
There are also plans to create a report on the factors that have influenced the change in the Bank's loan loss provisions.

2. Review and comparative analysis

2.1 Review of abstracts which were used for project and review of analogues

When cash is deposited with a bank, only a portion of it is kept as a balance, while the rest will be used to make a loan (or spent by the bank to buy securities). Money that is released or expended in this manner is then deposited with another bank, resulting in fresh deposits and the ability to lend it out again. The money supply in a nation is increased by lending, redepositing, and re-lending money. The money supply in most countries is several times greater than the monetary base provided by the central bank due to widespread partial banking reservations. The mandatory reserve ratio or other financial ratios recommended by financial regulators limit this multiplication (known as the money multiplier).

There are several approaches based on basic Central Bank law around the world. In Russia, commercial banks can use independent examples to calculate the provisioning rate and the factors affecting provisioning.

Methods are based on the IFRS 9 document which states:

Whether the trade receivables have a large borrowing aspect (which they shouldn't if the trade debtors are all due within one year), IFRS 9 allows you to pay for bad debts on a “expected loss” basis using the simplistic method.

Idea is to multiply each segment of trade debtors by its default rate to calculate the bad debt clause. Base the default rate on past credit defaults, with forward-looking details such as the economic crisis accompanying the coronavirus factored in.

Also, methods are regulated by the Basel II document which states:

Basel II maintains the concept of regulatory capital and an 8 percent minimum coefficient for regulatory capital over risk-weighted assets, as well as providing guidance for calculating minimum regulatory capital ratios. A bank's qualifying regulatory capital is divided into three levels under

Basel II. The lower the tier, the less subordinated securities a bank is permitted to hold. Each tier must account for a certain percentage of overall regulatory capital and is used as the numerator in regulatory capital ratio calculations.

The second pillar of Basel II is regulatory oversight, which establishes a mechanism for national regulatory authorities to deal with a variety of challenges, such as financial risk, liquidity risk, and legal risks. The market discipline cornerstone establishes a number of transparency criteria for banks' risk profiles, risk management procedures, and capital adequacy, all of which are beneficial to financial market participants.

It is possible to count in various ways how factors influence the formation of the reserve. In general, there are 4 ways to do this, which I will explain below.

3. Selection of methods, algorithms, and models for project implementation

3.1 Modelling of simple situation

We started with a simplified example of how provisioning works. banks and banking in general - as the subject of finance and banking is new to me.

Imagine that we have opened a "nonfinancial" organization: we will give loans to people we know.

Capital: 1000 rubles.

	Статья	Сумма
1 января	**Актив**	**1000**
	Касса	1000
	Расходы	0
	Пассив	**1000**
	Доходы	0
	Капитал	1000
	Прибыль	0

Fig. 1 there is a balance built here

Assets **ALWAYS** equal Liabilities. Income and Profit are in Liabilities. Expenses are in the Assets. Profit = Income - Expenses. This item is not calculated every day, but once a year, in the so called "financial year".
The capital invested in the organization is, in our case, all in Cash, i.e. in cash.
Loans can be issued.

We then made three potential loans:

Loan1, in the amount of 160 rubles, for 4 weeks, repayable weekly at 40 rubles. Interest weekly, 16 rubles per week.

Loan2, RUB 800, repayment after one week, interest at the end of the term, RUB 80.

Loan3, worth 40 rubles, for 4 weeks, weekly repayment of 10 rubles. Interest rate weekly, 8 rubles per week.

When a loan is issued, there is still the risk of non-payment, which means that the bank cannot guarantee that the interest will be returned in full and on schedule at the time of the transaction and during the loan's life cycle.

As a result, the bank is accumulating a chance of non-repayment by making a provision (the so-called "credit risk"). As a result, this reserve creates a more predictable operating climate for the bank, allowing it to prevent fluctuations in the volume of earnings associated with loan defaults.

The deductions due to the bank's operations are the basis of the clause. In other words, in the bank's accounting statements, the production of provisions is recorded as a cost, while the recovery, as a result of loan repayment or a reduction in the provision amount, is recorded as revenue.

We don't know yet how the loan repayments will go, but we are already starting to take income into account

We perform an interest calculation operation on the loan prior to the interest and half of the principal maturity deadline. It is up to you if you want to accrue interest every day or every month. In our business, we quantify it on a weekly basis.

The interest object is what we believe we can win but have not yet won. We want to get the Loans item back, but we have not received it yet. We also make Provisions for Interest - there is a possibility of not getting it.

Статья	Сумма
Актив	**1104**
Касса	0
Кредиты	1000
Проценты	104
Расходы	0
Пассив	**1104**
Доходы	0
Капитал	1000
Резервы (кредиты)	500
Резервы (проценты)	52
Прибыль	-448

Fig. 2 there is new balance built here

3.2 Calculating balance on different dates

Modelling the situation - not all loans are repayable
Borrower behavior:

Loan1, the interest will be paid, the loan is repaid at 50% of the schedule each week.

Loan2, interest will be paid, loan will be repaid in full on time.

Loan3, full default, interest will not be paid, loan will not be repaid.

The repayment day came, and as we saw, on loan1 the borrower only paid the interest and 50% of the scheduled debt, on loan 2 the borrower paid everything, both the interest and the debt, and on loan 3 no payments were made.

Let us build two balance sheet options for 10.01, with a 10.01 loan of RUB 800 for one week. Reserve rate 20% or 50%:

Баланс за 10/01			
Статья	**Сумма**	**20%**	**50%**
Актив	**1292**	116+950+30+8+0+348=**1452**	116+950+30+0+8+558=**1692**
Касса	916	916-800=**116**	**116**
Кредиты	150	800+150=**950**	800+150=**950**
Кредиты (просроч)	30	30	30
Проценты	0	0	0
Проценты (просроч)	8	8	8
Расходы	188	188+(160)=**348**	188+(400(50 процентов от 800))=**588**
Пассив	**1292**	104+1000+180+160+8+0=**1452**	104+1000+580+8+0=**1692**
Доходы	104	104	104
Капитал	1000	1000	1000
Резервы (кредиты)	180	180+(160(20 процентов от 800))=**340**	180+400=**580**
Резервы (проценты)	8	8	8
Прибыль	0	0	0

Fig. 3 there is a balance for different provision rates

And calculating tables with the formation of reserves for the two cases:

№	Показатель	01.янв	02.янв	08.янв	09.янв	10.янв	
1	**Формирование резерва, в том числе**	**0**	**500**	**52**	**94**	**160(20% от 800)**	
1.1.	выдача новых ссуд	0	500	0	0	160(20% от 800)	
1.2.	увеличение ставки резервирования	0	0	0	552-916/2=**94** => 188-94	0	
1.3.	изменения курса валюты	0	0	0	0	0	Баланс сходится: 0+500+52+94-458+160=**348**
1.4.	прочие причины(резерв по процентам)	0	0	52	0	0	
2	**Восстановление резерва, в том числе:**	**0**	**0**	**0**	**916/2=458**	**0**	
2.1.	погашение ссуд	0	0	0	916/2=**458**	0	
2.2.	уменьшение ставки резервирования	0	0	0	0	0	
2.3.	изменения курса валюты	0	0	0	0	0	
2.4.	прочие причины	0	0	0	0	0	
					вернули кредит		

Fig. 4 making provisions for 20%

№	Показатель	01.янв	02.янв	08.янв	09.янв	10.янв	
1	**Формирование резерва, в том числе**	**0**	**500**	**52**	**94**	**400(50% от 800)**	
1.1.	выдача новых ссуд	0	500	0	0	400(50% от 800)	
1.2.	увеличение ставки резервирования	0	0	0	552-916/2=**94** => 188-94	0	Баланс сходится: 0+500+52+94-458+400=**588**
1.3.	изменения курса валюты	0	0	0	0	0	
1.4.	прочие причины(резерв по процентам)	0	0	52	0	0	
2	**Восстановление резерва, в том числе:**	**0**	**0**	**0**	**916/2=458**	**0**	
2.1.	погашение ссуд	0	0	0	916/2=**458**	0	
2.2.	уменьшение ставки резервирования	0	0	0	0	0	
2.3.	изменения курса валюты	0	0	0	0	0	
2.4.	прочие причины	0	0	0	0	0	
					вернули кредит		

Fig. 5 making provisions for 50%

Conclusion: We should have seen that with a reduced capital burden we have more spare cash that we can put to work.

We add the impact of the exchange rate to our model - now the reservation rate will affect the exchange rate and our model depends on its value according to the Central Bank. We have free

money again. And, suddenly, on January 18 we were asked for 10€ for 2 weeks. We didn't get confused, went to exchange them at 85₽ and gave them away at 5% (weekly, as usual), but we decided to make a reserve of 80% because it's the first time we've seen this man, and it's also we don't know if he's coming back from this Europe or not... How should we take this into account? Every day we have to revalue foreign currency loans according to the exchange rate, as well as the reserves against them. Revaluation may be positive (the exchange rate went up) or negative (the rate has fallen). Positive revaluation of Assets is recorded against Income, a negative revaluation of Assets is posted to Expenses. The opposite is true for Liabilities.

*Let's take a look (**Fig. 6**) at how the balance behaves from the 18th to the 22nd of January. The only movements during these days are changes in the € exchange rate.*

Дата	Курс
18.янв	85,00 ₽
19.янв	86,50 ₽
20.янв	86,00 ₽
21.янв	87,50 ₽
22.янв	87,00 ₽
23.янв	88,50 ₽
24.янв	88,00 ₽

Статья	18.янв	19.янв	20.янв	21.янв	22.янв
Актив	**2062,00**	**2089,00**	**2089,00**	**2116,00**	**2156,00**
Касса	180,00	180,00	180,00	180,00	180,00
Кредиты	1010,00	1025,00	1020,00	1035,00	1030,00
Кредиты	*950,00*	*965,00*	*960,00*	*975,00*	*970,00*
Кредиты (просроч)	*60,00*	*60,00*	*60,00*	*60,00*	*60,00*
Проценты	16,00	16,00	16,00	16,00	36,00
Проценты	*0,00*	*0,00*	*0,00*	*0,00*	*20,00*
Проценты (просроч)	*16,00*	*16,00*	*16,00*	*16,00*	*16,00*
Расходы	856,00	868,00	873,00	885,00	910,00
Пассив	**2062,00**	**2089,00**	**2089,00**	**2116,00**	**2156,00**
Доходы	206,00	221,00	225,00	240,00	264,00
Капитал	1000,00	1000,00	1000,00	1000,00	1000,00
Резервы	856,00	868,00	864,00	876,00	892,00
Резервы (кредиты)	*840,00*	*852,00*	*848,00*	*860,00*	*856,00*
Резервы (проценты)	*16,00*	*16,00*	*16,00*	*16,00*	*36,00*
Прибыль	0,00	0,00	0,00	0,00	0,00

Fig. 6 changing of balance

SIMULATING THE SITUATION UNTIL THE END OF THE MONTH
Borrower behavior:
Loan1, continued to pay interest and 50% of the schedule, but decided on 31 January that they could repay the whole delinquency.
Loan3, they never paid back anything, so we decided to write off the entire overdue amount against on account of the reserves.
Loan5, we decided to pay back the entire amount early on January 24 (the first scheduled payment date). Thanks to him for not having to reassess the loan every day, but the interest he didn't pay us for the second week (we should have asked for an early repayment fee!).

№	Показатель	15.янв	16.янв	17.янв
1	**Формирование**	**22**	**16**	**0**
1,1	Выдача новых ссуд	0	0	0
1,2	увелечение ставки резервирвоания	0	0	0
1,3	изменения курса валюты	0	0	0
1,4	прочие причины(резерв по процентам)	22	16	0
2	**Восстановление**	**0**	**34**	**176**
2,1	погашение ссуд	0	20	160
2,2	уменьшение ставки резервирования	0	0	0
2,3	изменения курса валюты	0	0	0
2,4	прочие причины (вернули проценты %)	0	14	16

Fig. 7 making new provisions

3.3 Study the economic model and derive formulas for calculating the provision change factors

The formula for calculating the loss allowance is as follows
P = A×C×R, where

P - value of the allowance for possible losses
A - the asset (debt balance as at the calculation date) to be reserved (i.e. net of collateral)
C - currency rate of the asset as at the calculation date
R - reservation rate as at the calculation date

So, we have a function in the form of a product of three factors. In order to isolate the impact of each factor, we will take the function value at point 1 and point 2 (i.e., we will take the provisioning rate at date1 and date2):

$P_1 = A_1 \times C_1 \times R_1$
$P_2 = A_2 \times C_2 \times R_2 = (A_1 + \Delta A_1) \times (C_1 + \Delta C_1) \times (R_1 + \Delta R_1)$

Next, calculate the difference in values of the function P:
$P_2 - P_1 = (A_1 + \Delta A_1) \times (C_1 + \Delta C_1) \times (R_1 + \Delta R_1) - A_1 \times C_1 \times R_1$
Open the brackets and omit the "multiply" sign for simplicity (A×C = AC):
$\Delta P = (A_1C_1 + \Delta A_1C_1 + A_1\Delta C_1 + \Delta A_1\Delta C_1) \times (R_1 + \Delta R_1) - A_1 \times C_1 \times R_1$
$\Delta P = \cancel{A_1C_1R_1} + A_1C_1\Delta R_1 + \Delta A_1C_1R_1 + \Delta A_1C_1\Delta R_1 + A_1\Delta C_1R_1 + A_1\Delta C_1\Delta R_1 + \Delta A_1\Delta C_1R_1 + \Delta A_1\Delta C_1\Delta R_1 \cancel{- A_1C_1R_1}$

Let's regroup the summands:
$\Delta P = \Delta A_1C_1R_1 + A_1\Delta C_1R_1 + \Delta A_1\Delta C_1R_1 + A_1C_1\Delta R_1 + \Delta A_1C_1\Delta R_1 + A_1\Delta C_1\Delta R_1 + \Delta A_1\Delta C_1\Delta R_1$
$\Delta P = \Delta A_1C_1R_1 + \Delta C_1(A_1R_1 + \Delta A_1R_1) + \Delta R_1(A_1C_1 + \Delta A_1C_1 + A_1\Delta C_1 + \Delta A_1\Delta C_1)$
$\Delta P = \Delta A_1C_1R_1 + \Delta C_1(R_1(A_1 + \Delta A_1)) + \Delta R_1(C_1(A_1 + \Delta A_1) + \Delta C_1(A_1 + \Delta A_1))$
$\Delta P = \Delta A_1C_1R_1 + \Delta C_1R_1A_2 + \Delta R_1(C_1A_2 + \Delta C_1A_2)$
$\Delta P = \Delta A_1C_1R_1 + \Delta C_1R_1A_2 + \Delta R_1(A_2(C_1 + \Delta C_1))$
$\Delta P = \Delta A_1C_1R_1 + \Delta C_1R_1A_2 + \Delta R_1A_2C_2$

Total:
Asset Change Factor: ΔA = ΔA1C1R1
Currency rate change factor: ΔC = ΔC1R1A2
Reserve Rate Factor: ΔR = ΔR1A2C2

However, there are other methods to calculate the loss allowance (4 methods including ours):

1) An example of Factor Analysis. The method of absolute differences.

The change in the resultant indicator due to each factor by method of differences is defined as the product of the deviation of the factor under study by the baseline or reporting value of another factor depending on the chosen substitution sequence:

$y_0 = a_0b_0c_0$ – first date
$y_1 = a_1b_1c_1$ - second date

Thus, we get that

$$\Delta y_a = \Delta a b_0 c_0$$
$$\Delta y_b = \Delta b a_1 c_0$$
$$\Delta y_c = \Delta c a_1 b_1$$
$$\Delta y = \Delta y_a + \Delta y_b + \Delta y_c$$

2) An example of Factor Analysis. The method of Deterministic models.

1. $y = x_1 \cdot x_2$.
Влияние факторов:

$$\Delta y_1 = x_2^0 \cdot \Delta x_1 + {}^1/_2 \Delta x_1 \cdot x_2;$$
$$\Delta y_2 = x_1^0 \cdot \Delta x_2 + {}^1/_2 \Delta x_1 \cdot x_2.$$

2. $y = x_1 \cdot x_2 \cdot x_3$.
Влияние факторов:

$$\Delta y_1 = {}^1/_2 \Delta x_1 (x_2^0 \cdot x_3^1 + x_2^1 \cdot x_3^0) + {}^1/_3 \Delta x_1 \cdot \Delta x_2 \cdot \Delta x_3;$$
$$\Delta y_2 = {}^1/_2 \Delta x_2 (x_1^0 \cdot x_3^1 + x_1^1 \cdot x_3^0) + {}^1/_3 \Delta x_1 \cdot \Delta x_2 \cdot \Delta x_3;$$
$$\Delta y_3 = {}^1/_2 \Delta x_3 (x_1^0 \cdot x_2^1 + x_1^1 \cdot x_2^0) + {}^1/_3 \Delta x_1 \cdot \Delta x_2 \cdot \Delta x_3.$$

3. $y = x_1 / x_2$.
Влияние факторов:

$$\Delta y_1 = \frac{\Delta x_1}{\Delta x_2} \cdot \ln\left(\left|\frac{x_2^0}{x_1^0}\right|\right);$$
$$\Delta y_2 = \Delta y - \Delta y_1.$$

4. $y = \frac{x_1}{x_2 + x_3}$.
Влияние факторов:

$$\Delta y_1 = \frac{\Delta x_1}{\Delta x_2 + \Delta x_3} \cdot \ln\left(\left|\frac{x_2^1 + x_3^1}{x_2^0 + x_3^0}\right|\right);$$

$\Delta y_2 = \frac{\Delta y - \Delta y_1}{\Delta x_2 + \Delta x_3} \cdot \Delta x_2$; $\Delta y_3 = \frac{\Delta y - \Delta y_1}{\Delta x_2 + \Delta x_3} \cdot \Delta x_3$.

3) Method of averaging factors

This method averages the influence of each factor - that is, it takes all factors and considers them to have the same effect on the factors.

Conclusion: we can see that all of these methods are suitable for calculating factor change. The only difference is that the impact of each factor is counted differently in the moment. But it is most important to consider the fact that the factors influence the formation of the reserve and also to look at the percentage deviation of each factor. The most important thing for the bank is the dynamics and not the moment-by-moment calculation of the change of each factor (provisioning rate, exchange rate or asset).

We decided with the supervisor to choose first method - as it is the most interpretable and gives an excellent indication of the day

Asset Change Factor: ΔA = ΔA1C1R1
Currency rate change factor: ΔC = ΔC1R1A2
Reserve Rate Factor: ΔR = ΔR1A2C2

3.4 Study the data table

We were given a dataset - the Bank provides six Excel files with data on loss provisions (for six reporting dates). we only considered the file for physical individuals.

While studying the dataset I noticed that loans are issued in several currencies - dollars, rubles and euros (**Fig. 8**).

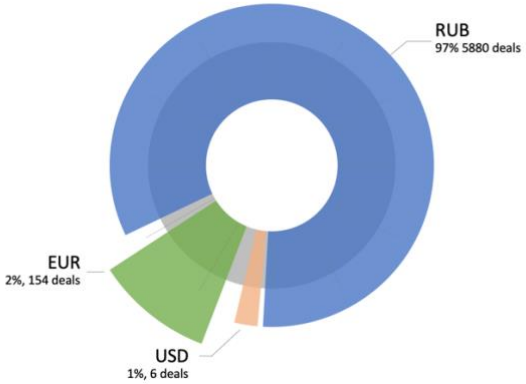


Fig. 8 distribution of currency in deals

In order to calculate the formation of reserves in currency equivalent – it is important to know the values of the currency at the moment.

With a function that parses the CBR collecting currency information knowing the date – I motivated the dataset and added a new column with currencies.

I used python parser forex_python.converter and pycbrf.toolbox to get the currency value for each date. This is how we get a new dataset with exchange rates (**Fig. 9**).

PROVISION_INTEREST_AMT	PROVISION_OVERDUEINTEREST_AMT	rate
NaN	27423484,280	85.7389
NaN	25809748,750	80.0488
209118,380	25326143,220	78.5489
NaN	25780983,340	78.6812
NaN	28491682,310	86.2532
NaN	29550985,410	88.7448

Fig. 9 example of new dataset

There is no need to do any special work on the dataset, as it is already fully operational from the beginning. The dataset is artificially simulated.

3.5 Develop a calculation methodology and algorithm, draw up a technical requirement for implementation

I write a function that works through the derived formula and compute the change in provisions automatically.
There were given real data set from the bank.

The next step after the getting rates and adding this column to dataset. Based on the dataset I calculated new dataset:

Общий резерв	dProvison	dA	dC	dR	dPA	dPR	dPC	dXR
25 653,93 ₽	25 653,93 ₽	66 005,68 ₽	-	- ₽	330,03 ₽	- ₽	- ₽	25 323,90 ₽
24 071,90 ₽	- 1 582,03 ₽	- 672,28 ₽	-4,04 ₽	- ₽	- 261,29 ₽	- ₽	- 1 320,75 ₽	0,01 ₽
22 877,57 ₽	- 1 194,33 ₽	- 663,64 ₽	-2,94 ₽	- ₽	- 244,52 ₽	- ₽	- 949,80 ₽	- 0,01 ₽
22 380,86 ₽	- 496,71 ₽	- 679,94 ₽	-0,80 ₽	- ₽	- 240,54 ₽	- ₽	- 256,18 ₽	0,01 ₽
23 226,16 ₽	845,30 ₽	- 671,61 ₽	3,41 ₽	- ₽	- 234,90 ₽	- ₽	1 080,21 ₽	- 0,01 ₽
23 377,66 ₽	151,50 ₽	- 675,59 ₽	1,27 ₽	- ₽	- 247,82 ₽	- ₽	399,32 ₽	0,00 ₽

Fig. 10 example of dataset-result for one unique deal

The columns here are computed in a way:

Provision: sum(SUMRESERVEMAINRPBU SUMRESERVEPRMNRPBU SUMRESERVEPRMN SUMRESERVEPRPR) для всех строк

Delta_provision: First row for each deal is 1(=provision in that row), next values are computed as n(provision(deltas(n - (n-1)))

dA: 1(MAINRESTCUR + PRMNRESTCUR), next values are computed as 2((MAINRESTCUR + PRMNRESTCUR)n - ((MAINRESTCUR + PRMNRESTCUR)n-1))

dC: First row for each deal is 1(=0), next values are computed as n(CUR_RATE(deltas(n - (n-1)))

dR: First row for each deal is 1(=0), next values are computed as n(RESERVERATERPBU(deltas(n - (n-1)))

dPA: (dA * RESERVERATERPBU * CUR_RATE)/ 100 for all values in rows

dPR: First row for each deal is 1(=0), next values are computed as n((MAINRESTCUR + PRMNRESTCUR) * (dR / 100) * CUR_RATE)

dPC: First row for each deal is 1(=0), next values are computed as n((dC * RESERVERATERPBU * (MAINRESTCUR + PRMNRESTCUR)) /100)

dXR: For all values like **Delta_provision** - dPA - dPR - dPC

There was a complication, as there are different numbers of dates for transactions in the databank (some transactions lasted 6 months, some 5, some 3 and so on) - it was just not possible to calculate. We had to write a function that would calculate for each transaction separately and then combine the transactions in a single dataset. Because of that the calculations took a lot of time and I had to convert the function that would calculate it several times.
It was quicker in terms of time and calculations to do it in Excel, so I did both so that I could compare the results and check myself.

3.6 Develop calculations and report on the technical requirement

The next step is to write the functional and non-functional requirements for the project. Requirement's analysis is a part of the software development process that involves collecting software requirements (software), systematizing them, identifying relationships, and documenting them. It is part of the general engineering discipline of requirements engineering.

The TOR should be drafted for the implementation of a function that will consider the changes in the factors influencing the formation of the reserve.

Why is it important to document the entire work process in the form of technical documentation?

1) The technical requirement spells out the agreements between the contractor and the client, which are difficult to express in a contract due to the use of specific IT terminology.

2) It will save time on communication: recorded technical solutions will save a lot of retellings, confirmations, confusions in indications.
3) The document will allow a clear division of responsibilities between the parties to the project.
4) The technical requirement provides an opportunity to analyze the future project and identify problems at the planning stage.
5) A properly drafted technical requirement will make the behavior of all participants predictable and avoid numerous misunderstandings.
6) From a legal point of view, the existence of this document will make it easier for the parties to resolve disputes.
7) The terms of reference make financial planning possible, which is the key to a successful business. The customer will be able to see in advance what his money is being spent on.

Functional requirements:

A functional requirement describes what the software system must do, while non-functional requirements impose restrictions on how the system will do it.
In my case it is:

Downloading and processing data

A dataset with data for the last 6 months in csv format is accepted as input.

Read out the file and process it for possible errors:

Type of error in the data	Problem solution
Missing values	1) for numerical values, this means in our case that the number is 0 - and this counts towards the total, so we do not delete it (see last 4 columns data) ***SUMRESERVEMAINRPBU*** *(Amount of allowance for possible main debt losses)* ***SUMRESERVEPRMNRPBU*** *(amount of allowance for possible losses of overdue principal)* ***SUMRESERVEPRMN*** *(Amount of allowance for principal debt interest)* ***SUMRESERVEPRPR*** *(Sum of Provision for Interest on Overdue Principal Debt)* 2) For categorical variables, replace with the most popular value
Checking for outliers	Remove outlays from the table so that they do not affect the reserve

Functions

1) Function get_currency

Since we have currency loans - you need to search the Central Bank website https://www.cbr.ru/currency_base/ and pull up the currency data as of a certain date.

We do this using the pycbrf.toolbox library in python.

Code	Output
''' c = CurrencyRates() amount = 1 date = datetime(2021,4,6) print(c.convert('USD','RUB', amount, date)) '''	76.6232644768

Check result with
https://www.cbr.ru/currency_base/daily/?UniDbQuery.Posted=True&UniDbQuery.To=31.03.2020

Resources

Table	Description
Loanphys.csv	Loan portfolio of individuals

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period
2	currency	The value of the currency we want

Working procedure

The function goes through the dataset, looks at the value of the currency on a certain date and pairs the Central Bank site to find out the currency, if the currency is in Rubles, the value is set to 1.

2) Breakdown by number of credits

As the dataset can store a different number of credit dates, it is important to correctly count the values for each number of credits. so I use pandas to split the dataset and sort it by number of credits. Deals with the same number are stored in a separate sheet for later use.

Result

Output	Description
credits_dfs[0]	Client transactions with 1 date only
credits_dfs[1]	Client transactions with 2 dates only
credits_dfs[2]	Client transactions with 3 dates only
credits_dfs[3]	Client transactions with 4 dates only
credits_dfs[4]	Client transactions with 5 dates only
credits_dfs[5]	Client transactions with 6 dates only

3) Function that counts provision

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
Value is the sum of the reserve fields	'SUMRESERVEMAINRPBU' + SUMRESERVEPRMNRPBU' + 'SUMRESERVEPRMN' + 'SUMRESERVEPRPR'

4) Function that counts delta_provision

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
First value always is the same as the provision on that row	deltas = [tmp_df.iloc[0]['PROVISION']]
All subsequent values are the amount of provision in that term minus the previous one	tmp_df.iloc[i-1] init_row = tmp_df.iloc[i] deltas.append(init_row['PROVISION'] - prev_row['PROVISION'])

5) Function that counts dA

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
First value always is the sum of 2 values in that row	**MAINRESTCUR + PRMNRESTCUR**
All subsequent values are the sum of the current MAINRESTCUR + PRMNRESTCUR values (in	**(MAINRESTCUR_(n+1) + PRMNRESTCUR_(n+1)) – (MAINRESTCUR_n + PRMNRESTCUR_n)**

this row) minus the previous sum of the same values	

6) Function that counts dC

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
First value always is the same as the currency on that row	tmp_df = df[df['DEAL_ID'] == credit_id] tmp_df.reset_index().drop(columns = ['index']) **dC_s = [0]**
All subsequent values are the amount of currency in that term minus the currency one	tmp_df.iloc[i-1] tmp_df.iloc[i-1] init_row = tmp_df.iloc[i] dR_s.append(init_row[' CUR_RATE'] - prev_row['CUR_RATE'])

7) Function that counts dR

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
First value always is the same as the rate on that row	tmp_df = df[df['DEAL_ID'] == credit_id] tmp_df.reset_index().drop(columns = ['index']) **dR_s = [0]**
All subsequent values are the amount of rate in that term minus the rate one	tmp_df.iloc[i-1] init_row = tmp_df.iloc[i] dR_s.append(init_row['RESERVERATERPBU'] - prev_row['RESERVERATERPBU'])

8) Function that counts dPA

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
This value for each date is calculated separately - that is, the value in this row is multiplied by dA by the value of the currency and by the RESERVERATERPBU rate and everything is divided by 100 to get rid of percentage	next_row = tmp_df.iloc[i+1] init_row = tmp_df.iloc[i] dPA_s.append(next_row['dA'] * (init_row['RESERVERATERPBU'] / 100) * init_row['CUR_RATE'])

9) Function that counts dPR

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
First value always is 0	tmp_df = df[df['DEAL_ID'] == credit_id] tmp_df.reset_index().drop(columns = ['index']) dPR_s = [0]
the value equals the sum of MAINRESTCUR and PRMNRESTCUR for that date in the row, all multiplied by the currencies on that date according to the CB and multiplied by the rate change on that day and the previous day, to get rid of the percentage	init_row = tmp_df.iloc[i] dPR_s.append((init_row['MAINRESTCUR'] + init_row['PRMNRESTCUR']) * (init_row['dR'] / 100) * init_row['CUR_RATE'])

10) Function that counts dPC

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value

First value always is 0	tmp_df = df[df['DEAL_ID'] == credit_id] tmp_df.reset_index().drop(columns = ['index']) dPC_s = [0]
the value equals the currency change on that day and the previous day, multiplied by RESERVERATERPBU on that date, multiplied by the sum of MAINRESTCUR and PRMNRESTCUR for that day and divided by 100, to get rid of the percentage	init_row = tmp_df.iloc[i] dPC_s.append(((init_row['MAINRESTCUR'] + init_row['PRMNRESTCUR']) * init_row['RESERVERATERPBU'] * init_row['dC']) / 100)] * init_row['dC']) / 100)

11) Function that counts dXR

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
The value is counted as of the specified date in the row for each separately cumulative total from the previous data frame by subtracting values such as dPA, dPR and dPC from dProvison	for column in ['dPA', 'dPR', 'dPC']: df['dXR'] -= df[column]

12) Function that counts deltas

Parameters of the function

	Parameter	Parameter value formula
1	df	Dataset for the period

Working procedure

Description of the fields	Value
This is a summary function, which in turn calls the functions above in the correct order to obtain the correct values for each trade.	df = count_provision(df) df = count_delta_provision(df) df = count_dA(df) df = count_dC(df) df = count_dR(df) df = count_dPA(df)

	df = count_dPR(df) df = count_dPC(df) df = count_dXR(df)

In this function we learn for each data frame (we have 6 for each number of days), so we call it 6 times and get 6 calculated data frames with the necessary values to calculate the factors of the reserve change and fill in the summary table.

These datasets can then be merged into one common dataset by grouping by date for each deal.

Non-functional requirements:

Describe the characteristics of the system and its environment, not the behavior of the system. A list of constraints on the actions and functions performed by the system can also be given here. These include time constraints, constraints on the system development process, standards, etc.
In my case it is:

Requirements for regulations

The provisioning table is calculated monthly, the calculation for the previous 6 months is made on the 1st of the following month. The table contains monthly data cuts starting from 31.03.2020.

The report does not include information for the calendar month not completed.

Data quality requirements

csv / xslx table with separator ';'

4. Results and conclusion

4.1 Analyze the report and make economic conclusions

Finally, we get the table for all credit IDs:

```
cols = ['PROVISION','DELTA_PROVISION','dA','dC','dR','dPA','dPR','dPC','dXR','Month']
agg = res_df[cols].groupby(by = ['Month']).sum()
```

PROVISION	DELTA_PROVISION	dA	dC	dR	dPA	dPR	dPC	dXR
1171931679.5299997	1171931679.5299997	1733960968.9100003	0.0	0.0	799656440.5589797	0.0	0.0	372275238.9710212
1119945150.55	-27389613.339999998	-21688482.530000024	-106.76759999999989	202.64999999999995	-10693451.366862496	-7356428.6860449985	-6506948.830143247	-17545641.829040032
1111648464.87	-8163845.769999999	-6090028.470000002	-74.934900000000027	-84.70000000000005	-3195096.6111265398	-923727.61755	-4033822.0289478227	-11199.512375635328
1100931322.7299995	-5977087.529999999	-21652178.61	-19.885199999998	116.1500000000001	-5449336.0700696	-2392531.985468281	-957691.5248955048	-1962591.9205031702
1106686398.9299989	-8510163.209999992	-22830241.509999998	-92.871999999998	211.39999999999992	-9483671.10306812	-5001066.434691152	-6206302.358900791	-6786465.519476162
1107883335.3299994	-1376498.5400000005	-14446408.890000006	-34.364100000000005	22.299999999999997	-4196794.28942671	-290970.3453309801	-2227962.1018519206	-3636301.0729057733

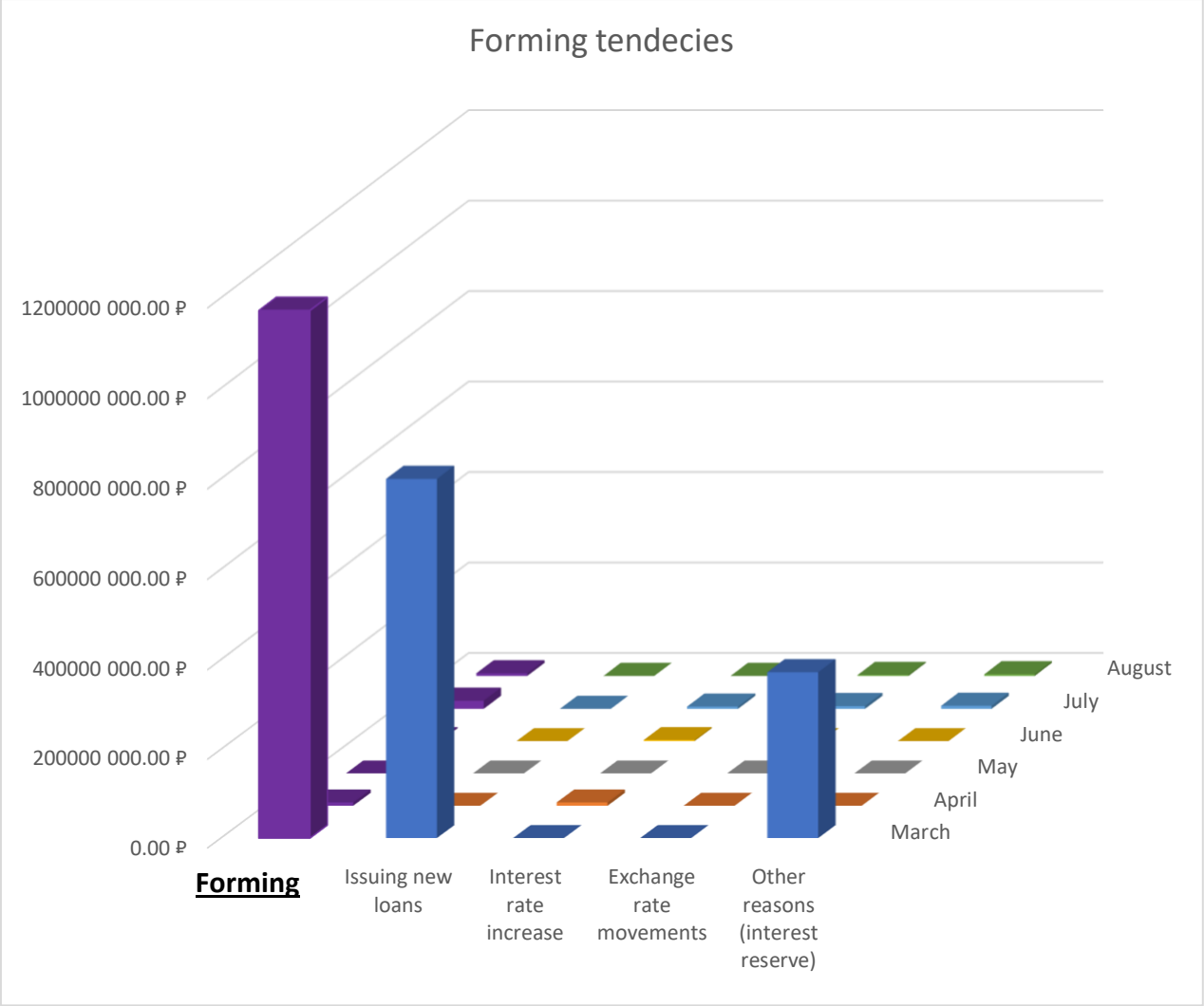
To make this table more accurate I transformed it with Excel:

Provision	dProvison	dA	dC	dR	dPA	dPR	dPC	dXR
1 171 931 679,53 ₽	1 171 931 679,53 ₽	1 733 960 968,91 ₽	0,00 ₽	0,00 ₽	799 656 440,56 ₽	0,00 ₽	0,00 ₽	372 275 238,97 ₽
1 119 945 150,55 ₽	-27 389 613,34 ₽	-21 688 482,53 ₽	-106,77 ₽	202,65 ₽	-10 693 451,37 ₽	7 356 428,69 ₽	-6 506 948,83 ₽	-17 545 641,83 ₽
1 111 648 464,87 ₽	-8 163 845,77 ₽	-6 090 028,47 ₽	-74,93 ₽	-84,70 ₽	-3 195 096,61 ₽	-923 727,62 ₽	-4 033 822,03 ₽	-11 199,51 ₽
1 100 931 322,73 ₽	-5 977 087,53 ₽	-21 652 178,61 ₽	-19,89 ₽	116,15 ₽	-5 449 336,07 ₽	2 392 531,99 ₽	-957 691,52 ₽	-1 962 591,92 ₽
1 106 686 398,93 ₽	8 510 163,21 ₽	-22 830 241,51 ₽	92,87 ₽	211,40 ₽	-9 483 671,10 ₽	5 001 066,43 ₽	6 206 302,36 ₽	6 786 465,52 ₽
1 107 883 335,33 ₽	1 376 498,54 ₽	-14 446 408,89 ₽	34,36 ₽	22,30 ₽	-4 196 794,29 ₽	-290 970,35 ₽	2 227 962,10 ₽	3 636 301,07 ₽

After that with Excel formulas I create the Final table with recovery and forming of provisions with several factors for which we calculated formulas in the beginning of the work

№	Indicator	March	April	May	June	July	August
1	**Forming**	**1 171 931 679,53 ₽**	**7 356 428,69 ₽**	**0,00 ₽**	**2 392 531,99 ₽**	**17 993 834,31 ₽**	**5 864 263,17 ₽**
1.1	Issuing new loans	799 656 440,56 ₽					
1.2	Interest rate increase		7 356 428,69 ₽		2 392 531,99 ₽	5 001 066,43 ₽	
1.3	Exchange rate movements					6 206 302,36 ₽	2 227 962,10 ₽
1.4	Other reasons (interest reserve)	372 275 238,97 ₽				6 786 465,52 ₽	3 636 301,07 ₽
2	**Recovering**	**0,00 ₽**	**34 746 042,03 ₽**	**8 163 845,77 ₽**	**8 369 619,52 ₽**	**9 483 671,10 ₽**	**4 487 764,63 ₽**
2.1	Loan repayment		10 693 451,37 ₽	3 195 096,61 ₽	5 449 336,07 ₽	9 483 671,10 ₽	4 196 794,29 ₽
2.2	Interest rate reduction			923 727,62 ₽			290 970,35 ₽
2.3	Exchange rate movements		6 506 948,83 ₽	4 033 822,03 ₽	957 691,52 ₽		
2.4	Other reasons (returned interest)		17 545 641,83 ₽	11 199,51 ₽	1 962 591,92 ₽		

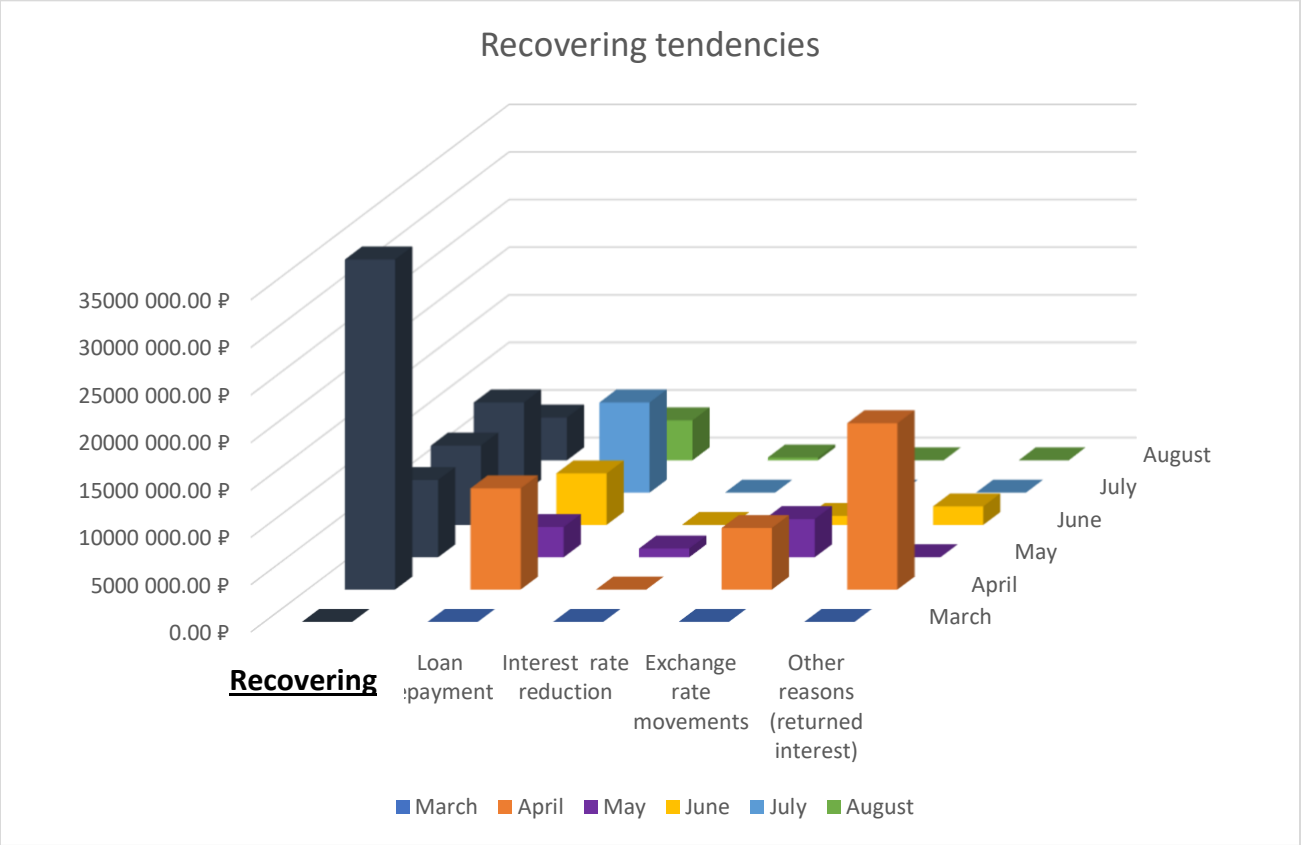
Then, I obtained the picture that represents the tendencies in the bank:



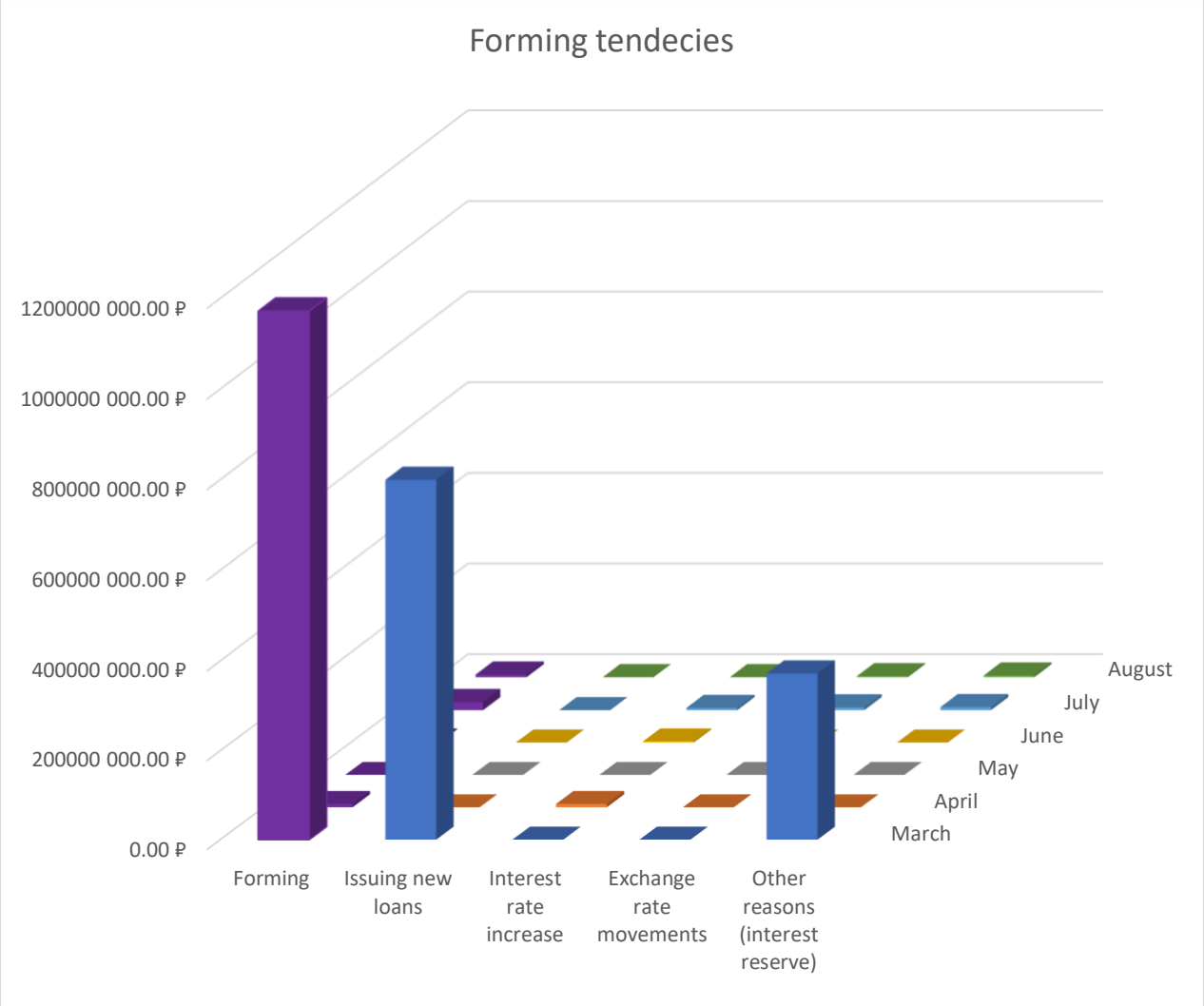
As the result we see that provisions are forming mostly with increase in interest rate, exchange rates movements and other reasons. We will measure those changes related to the provisioning rate are most important in April - it is logical, because April 2020 is the beginning of provisioning (the Central Bank reduced rates to record lows, everyone knew that the economic crisis was starting - and one of the methods - monetary policy - to reduce the key rate, so that in February the rate was 6% and in late April to 5,5%) to make loans cheaper and more accessible. Since July 2020 and November 2020, the devaluation of the ruble has increased due to the situation with geopolitical situations (Belarus, the situation with the opposition in Russia). Investors are withdrawing Russian assets as geopolitical factors have increased. As the ruble depends on capital outflows-inflows, commodity prices and therefore money was withdrawn, the exchange rate changed a lot. investors were scared of sanctions. Thus, we can see that from April to August, the greatest number of reserves was formed due to changes in exchange rates. The rapid spread of the disease forced the authorities to introduce quarantine measures, which brought the entire world economy to a temporary standstill. In a panic, investors started to withdraw their money from so-called risky assets, which included the ruble. In early spring the number of cases began to grow rapidly worldwide, and in mid-March the World Health Organization announced the beginning of a pandemic. At the time, the exchange rates of the dollar and the euro rose above 75 and 84 rubles. In addition to investor panic, the worsening situation on the energy market played against the Russian currency. The imposition of lockdowns and border closures led to a sharp drop in global fuel consumption. Amid falling demand, the price of benchmark Brent crude fell from $71-72 per barrel in January to $49-50 by early March.

However, already in autumn geopolitical factors started to put pressure on the ruble. Many investors were concerned that the military confrontation in the region and the political crisis in neighboring Belarus could create a negative background for the key emerging market currencies, including the ruble. Against this backdrop, the dollar and euro began to rise steadily again in September. In addition, as the autumn approached, the US presidential election influenced the dynamics of the Russian currency.

I also built the same graph to look at the recovery of reserves.



We can see that the recovery is being made mainly by loan repayments - there was a key rate cut to 5,5% in April and in July - and so the rate on other loans (including mortgages - and we have the most lending on mortgages in these months. Mortgage rates are very low, banks are lowering mortgage rates. And now you can refinance, go to another bank and find another rate that is more favorable. People have started refinancing mortgages, taking a new mortgage from a bank. So, in these months the most recovery was through repayment of loans and paying interest. The rate cuts were therefore just in April and July. So, the bottom line is that formation does come from repayment of loans throughout this period.



Now if we look at March, it is considered as if all loans were issued in March (because the first date on our dataset is March), which means that in March loans are encouraged due to clients taking loans. And because they take out loans, the interest reserve also plays a role - so there is a very large build-up factor about it too.

Optimization of the bank's operations is achieved by making the right management decisions, which requires a comprehensive analysis of the bank's performance. The result of the analysis should be information describing the mechanisms of the bank's operation and showing opportunities for adjustments to the banking process. But most methods do not allow to determine reasons of occurrence of dependences and thus do not give possibilities to find such administrative impacts, which could provide the most effective development of dynamic processes, do not allow to give recommendations to each bank on how to build optimal conditions of bank management strategy and investments. For the decision of these tasks, it is offered to use methods of factor analysis. Construction of the time factor model allows to identify those initial characteristics, which weakly manifested themselves in the past, but significantly influenced during the last years. With the help of this model, it becomes possible to obtain additional information about the economic process under study by identifying the attributes that have weakened or strengthened their influence in recent years, as well as the attributes that significantly influence the factor during the whole period of time under study.

Thus, two objectives of factor analysis can be distinguished:

- Determining the relationships between the variables, classifying them, i.e., "objective R-classification";
- Reducing the number of variables.

This analysis can help investors assess how reserves are built up and rebuilt at a particular bank in order to make an investment decision. If reserves are very dependent on the provisioning rate and foreign exchange rate - an investor will always have to monitor these indicators and be on the lookout for bank failures. Also, building up a reserve in line with the risk groups of the customers who have chewed out can help the bank avoid unnecessary expenses.

6. List of sources

https://www.investopedia.com/terms/b/baselii.asp
https://www.cpdbox.com/ifrs-9-expected-credit-loss/
https://www.cpdbox.com/how-to-calculate-bad-debt-provision-under-ifrs-9/
https://www.mytipsandadvice.co.uk/2020-09/how-to-calculate-a-bad-debt-provision-under-ifrs-9-UKTAFCAR_EU12110301
https://studfile.net/preview/2152457/page:4/
https://en.wikipedia.org/wiki/Non-functional_requirement
Presentation and materials from Банк открытие from our supervisors