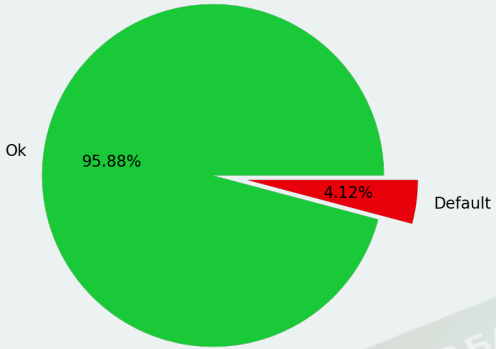
Skoltech

Skolkovo Institute of Science and Technology

СБЕР

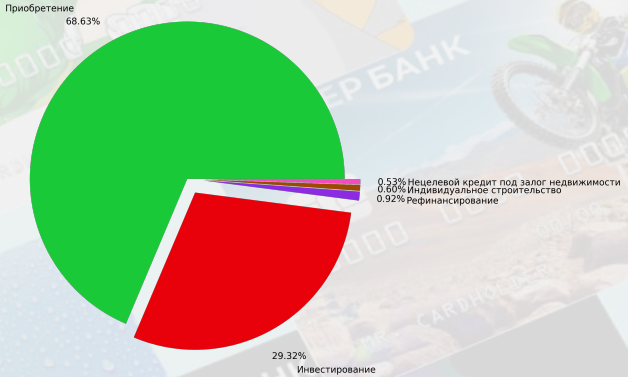
First Glance

Target ratio

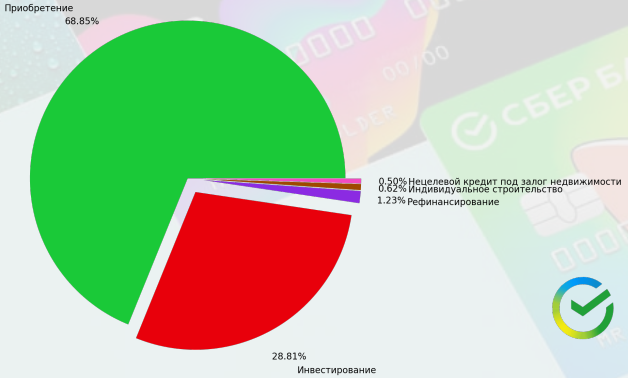


Train
1.5 M
Samples

Goals ratio train

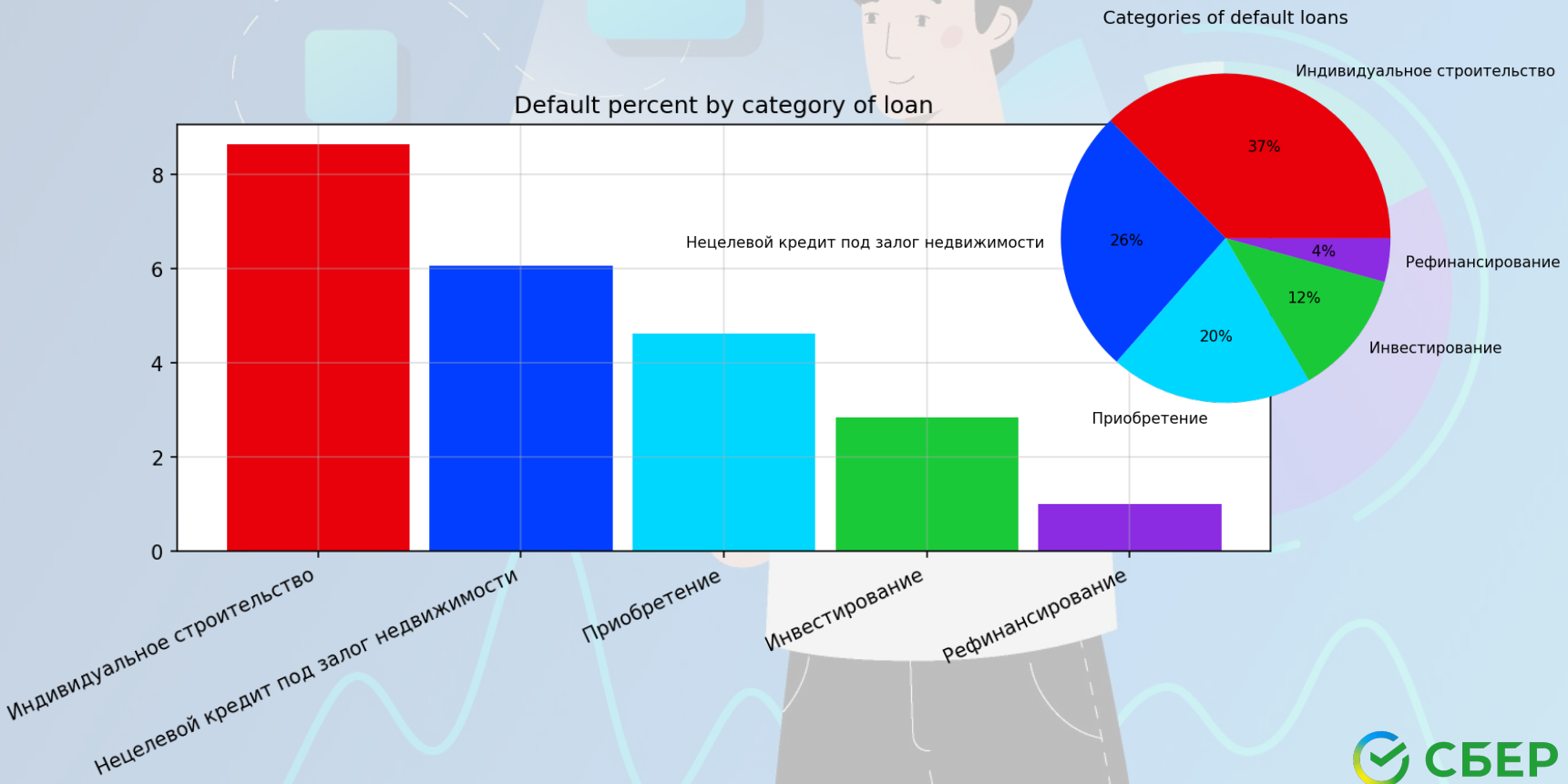


Goals ratio test



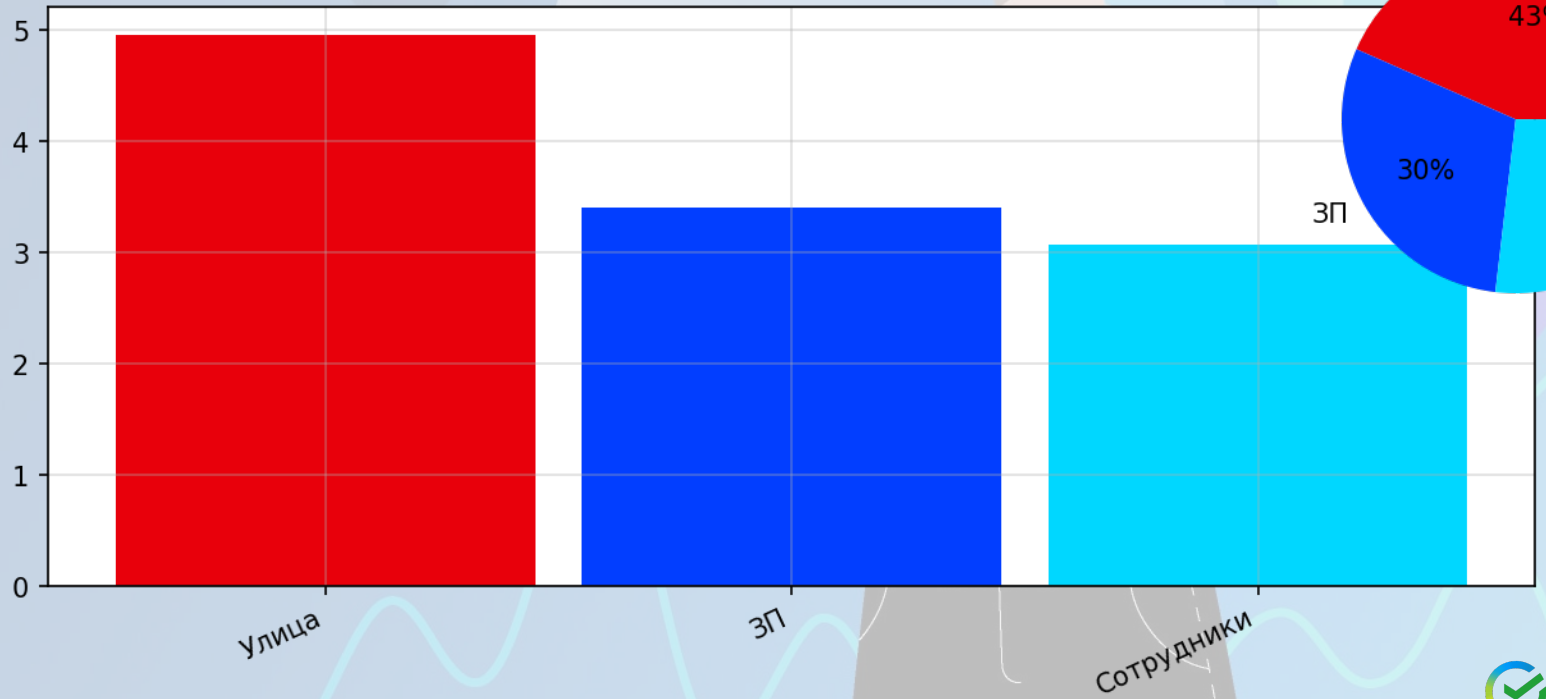
СБЕР

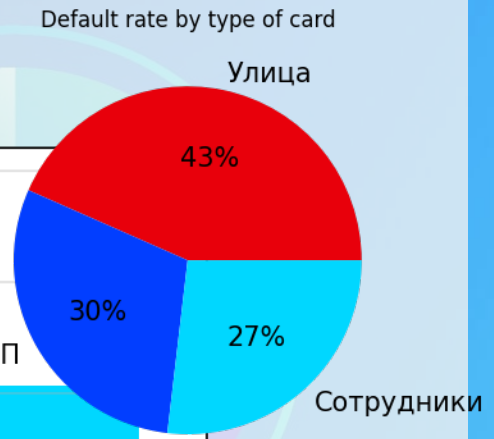
Category of loan vs default rate (x_21)



СБЕР

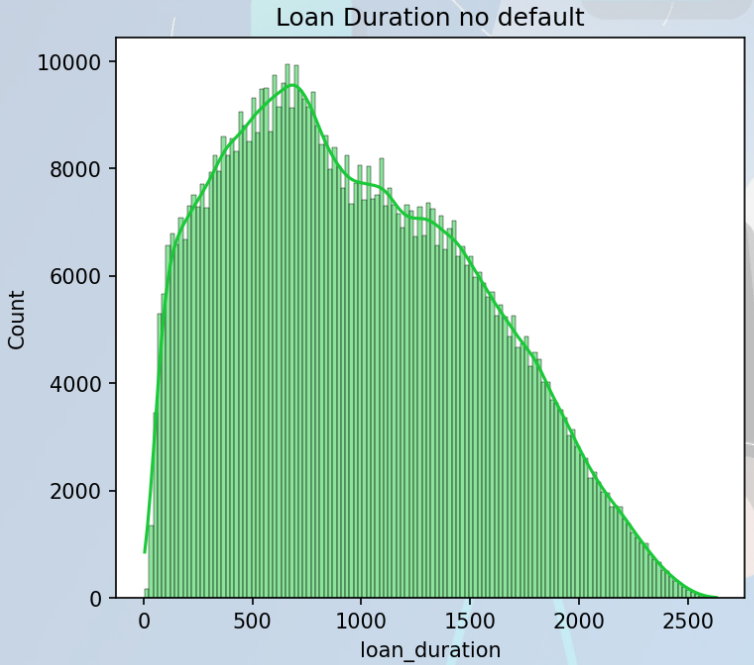
Default rate vs type of card (x_628)

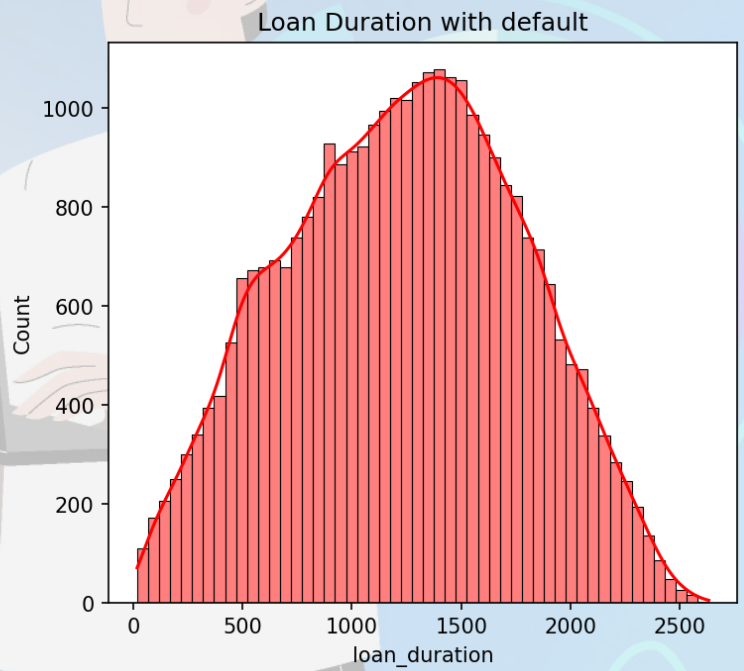




СБЕР

Duration of loan vs default rate

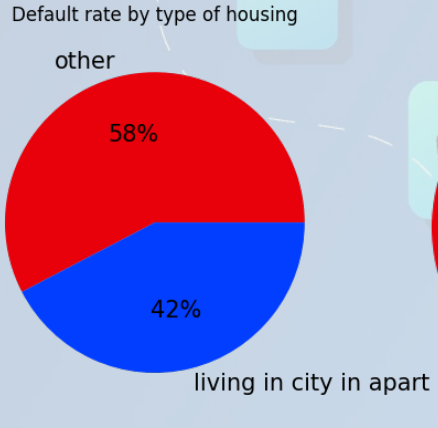


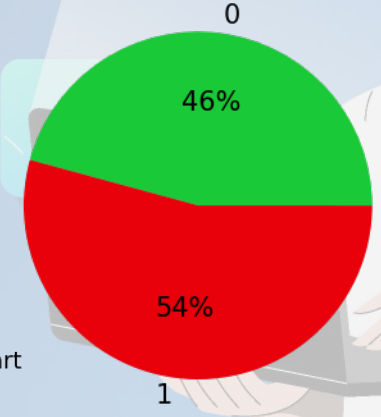


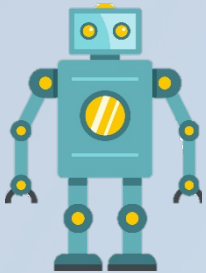
loan_duration = REPORT_DT - X_9

СБЕР

EDA Conclusion







Conclusion

The most important features ones are:
x21 - which demonstrates the goal of the particular loan in this case category of 'Individual building' is the riskiest, with a huge default rate of **8.6%** compared to ~4.5% of the mean default rate.
x_628- which shows which type of card the borrower has. In this case, the riskiest category is people who a not a cardholder of Sber with a default rate **~5%** and obviously the least risky category is Sber employees.
x_625 - which shows whether the borrower leave in the city also shows a quite significant difference of **4.94%** for borrowers who live somewhere else compare to **3.63%** for borrowers living in the city.

СБЕР

Initial Data Preparation

Train1
(763801, 647)

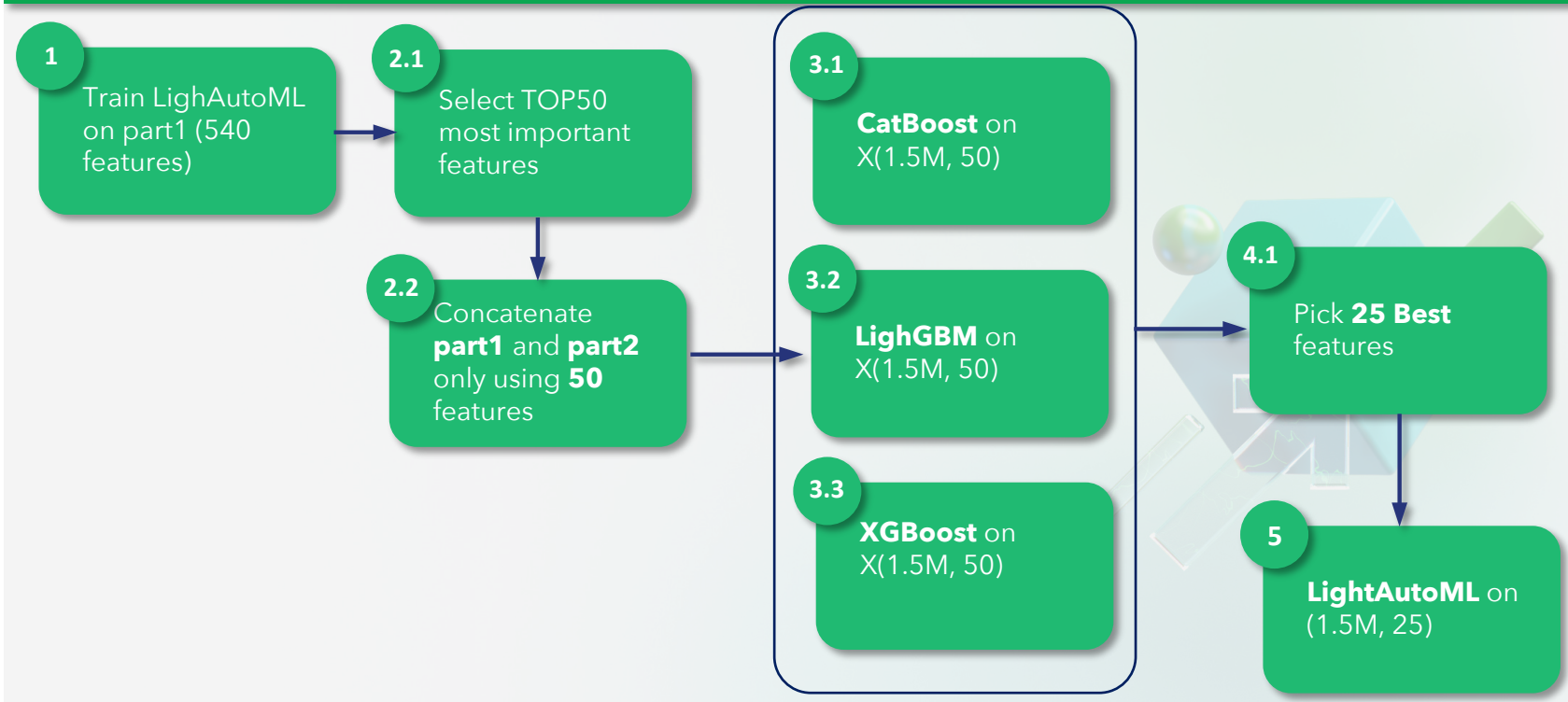
Train2
(763802, 647)

Drop:
1. cols with >80% of **NaNs**
2. Const cols
3. Equal cols

X = (1527598, 540)
Use 5 fold crossvalidation for model evaluation

СБЕР

Pipline



СБЕР

Final results

#	Team Name	Notebook	Team Members	Score	Entries	Last
1	KPM			0.82915	1	6h
2	poBEDA			0.81646	8	9m
3	Sk Learn			0.81611	8	2h
4	404 found			0.81222	3	25m
5	Dmitry Volkov			0.80232	11	21m

Your Best Entry

Your submission scored 0.80232, which is an improvement of your previous score of 0.78302. Great job! Tweet this!

#	Team Name	Notebook	Team Members	Score	Entries	Last
6	Qwerty			0.79805	1	7h
7	SerAi			0.78336	6	14h

baseline → **1st checkpoint** → **2nd checkpoint** → **Final model**

Yandex CatBoost (48 features) → LogReg (55 features) → LightGBM (25 features)

0.763 → 0.783 → 0.802 → 0.791

СБЕР