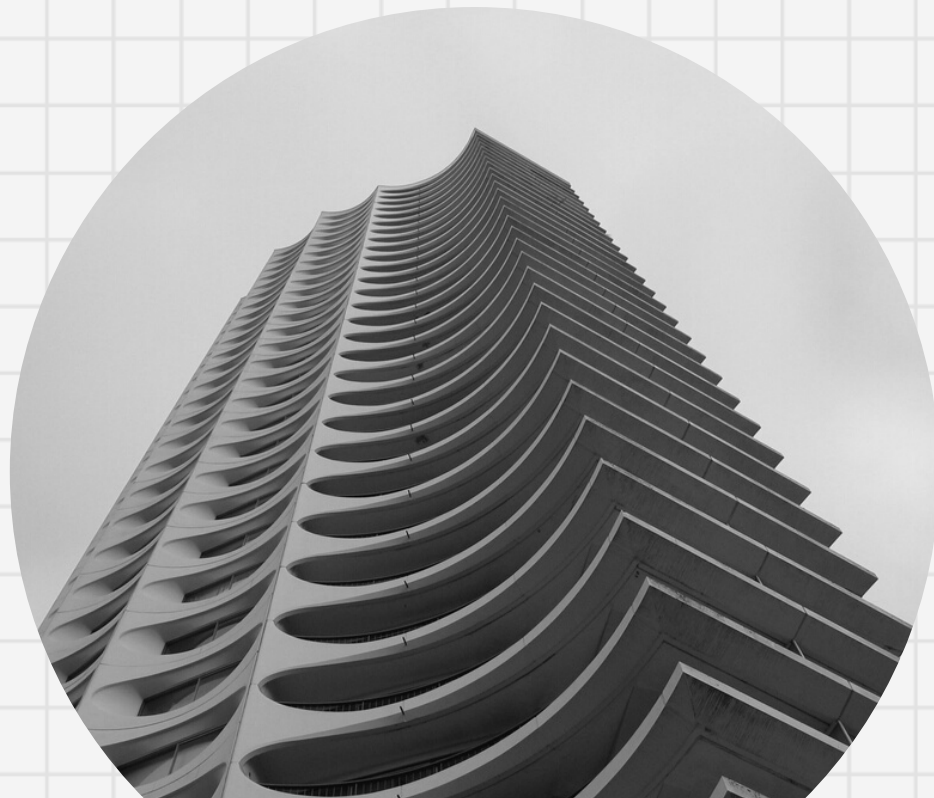


**DEVKIT PRO**

# **2021 ANALYSIS REPORT**

**An annual report by Artur Akimov**



Этот набор данных случайным образом собирается из базы данных иранской телекоммуникационной компании в течение 12 месяцев. В общей сложности 3150 строк данных.

# Описание набора данных

1. Call Failures - количество неудачных вызовов
2. Complains - Жалобы бинарные (0: нет, 1: 1 или больше)
3. Subscription Length - продол-ность подписки
4. Seconds of Use - общее количество секунд разговоров
5. Frequency of use - общее количество звонков
6. Frequency of SMS - общее количество текстовых сообщений
7. Distinct Called Numbers - общее количество различных телефонных звонков
8. Churn - Отток бинарные (1: да, 0: нет)

## Датасет без None значений

	Call Failure	Complains	Subscription Length	Charge Amount	Seconds of Use	Frequency of use	Frequency of SMS	Distinct Called Numbers	Churn
count	706.000000	706.000000	706.000000	706.000000	706.000000	706.000000	706.000000	706.000000	706.000000
mean	7.678470	0.069405	32.855524	0.885269	4282.563739	68.262040	74.912181	23.011331	0.169972
std	6.857001	0.254322	8.260928	1.351493	4009.338768	56.037102	118.570400	16.923018	0.375874
min	0.000000	0.000000	3.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	1.000000	0.000000	30.000000	0.000000	1360.750000	27.000000	7.000000	10.000000	0.000000
50%	7.000000	0.000000	35.000000	0.000000	3039.000000	52.500000	21.000000	21.000000	0.000000
75%	12.000000	0.000000	38.000000	1.000000	6298.000000	93.000000	81.750000	33.000000	0.000000
max	34.000000	1.000000	47.000000	10.000000	16980.000000	254.000000	498.000000	97.000000	1.000000

## Датасет с None значениями

	Call Failure	Complains	Subscription Length	Charge Amount	Seconds of Use	Frequency of use	Frequency of SMS	Distinct Called Numbers	Churn
count	2524.000000	2558.000000	2633.000000	3150.000000	2561.000000	2479.000000	2488.000000	2514.000000	3150.000000
mean	7.731379	0.076622	32.499810	0.942857	4457.024209	69.052844	73.308682	23.361575	0.157143
std	7.286990	0.266043	8.540095	1.521072	4167.652501	56.910760	112.651426	17.275861	0.363993
min	0.000000	0.000000	3.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	1.000000	0.000000	29.000000	0.000000	1390.000000	27.000000	6.000000	10.000000	0.000000
50%	6.000000	0.000000	35.000000	0.000000	3018.000000	54.000000	21.000000	21.000000	0.000000
75%	12.000000	0.000000	38.000000	1.000000	6458.000000	94.000000	87.250000	33.000000	0.000000
max	36.000000	1.000000	47.000000	10.000000	17090.000000	255.000000	522.000000	97.000000	1.000000

# Выбор наиболее важных параметров **feature importance**

Q4

Q1

**Call Failures:  
количество  
неудачных  
вызовов**

Модель определила  
важность данного  
параметра при выборе  
более чем на 24%

Q2

**Frequency of use:  
общее  
количество  
звонков**

Модель определила  
важность данного  
параметра при выборе  
более чем на 16%

Q3

**Seconds of Use:  
общее  
количество  
секунд  
разговоров**

Модель определила  
важность данного  
параметра при выборе  
более чем на 14%

**Complains:**

**Жалобы  
бинарные  
(0: нет, 1: один  
или больше  
одного)**

Модель определила  
важность данного  
параметра при выборе  
более чем на 12%

Заключение

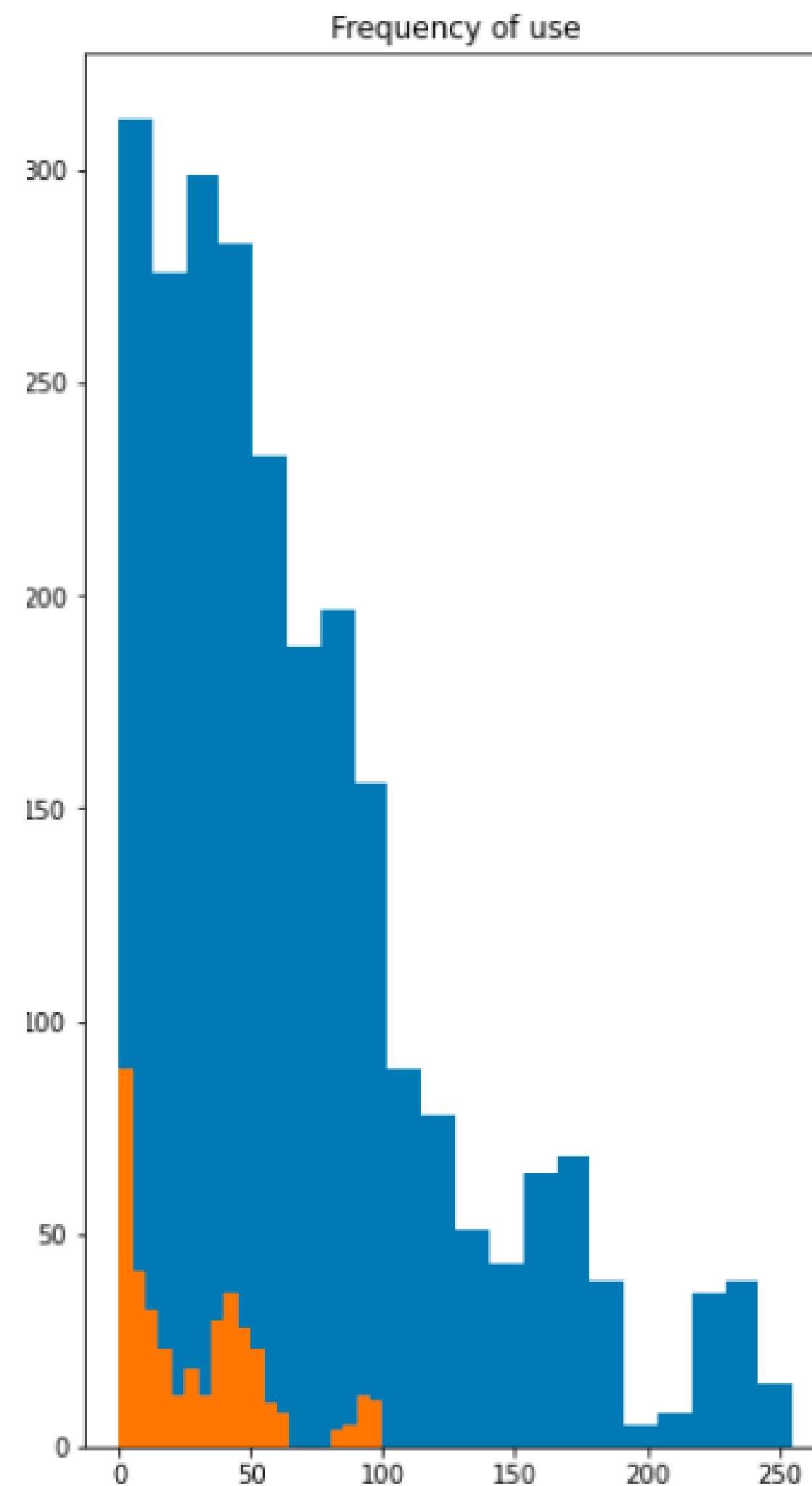
**Эти четыре  
параметра  
являются  
ключевыми при  
выборе решения**

Сумма их важности  
составляет 66%

# Гистограмма общего количество ЗВОНКОВ

Датасет с None значениями

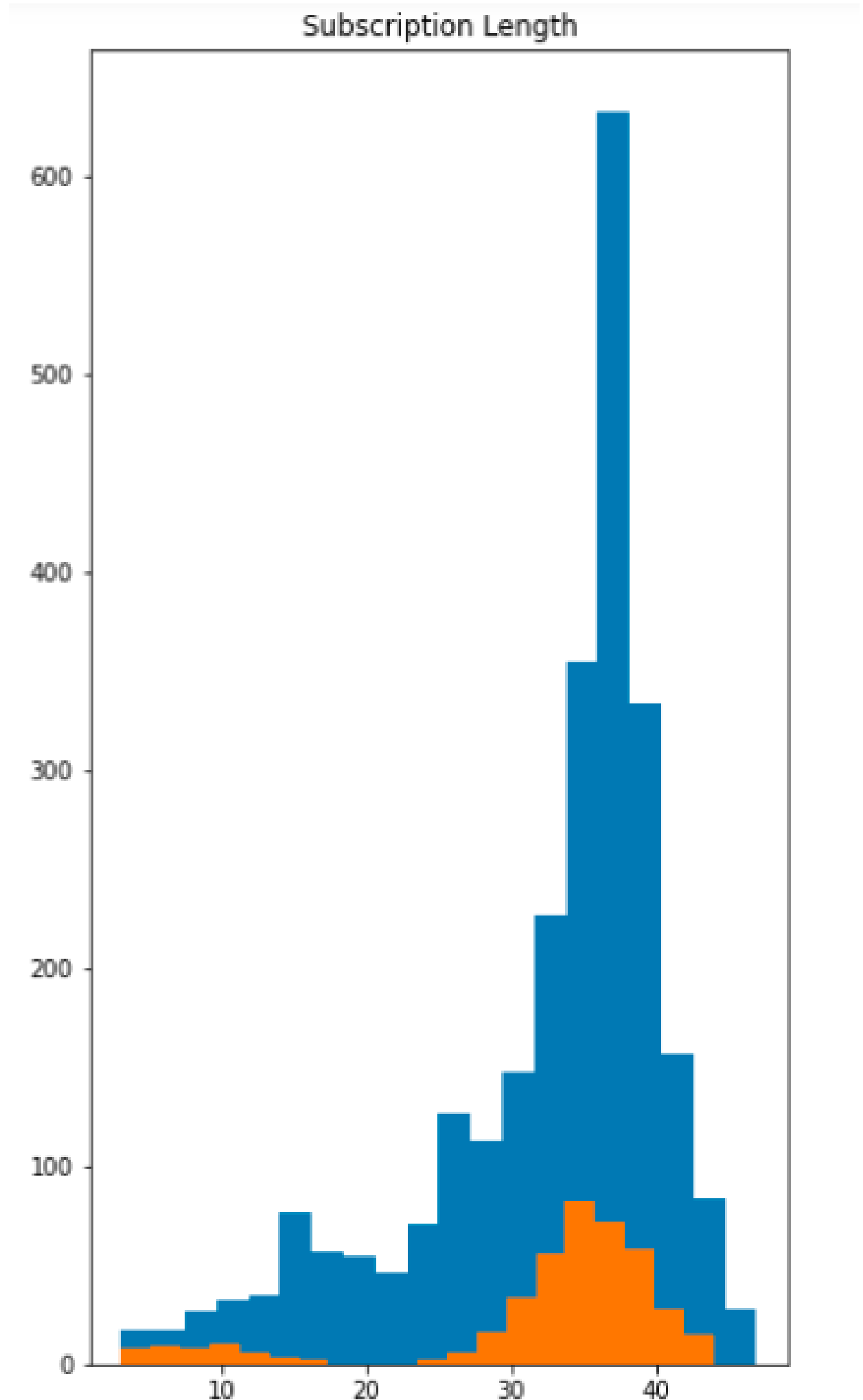
**Здесь отчётливо видно что основной  
отток происходит у людей с частотой  
использования меньше 50 ти вызовов**

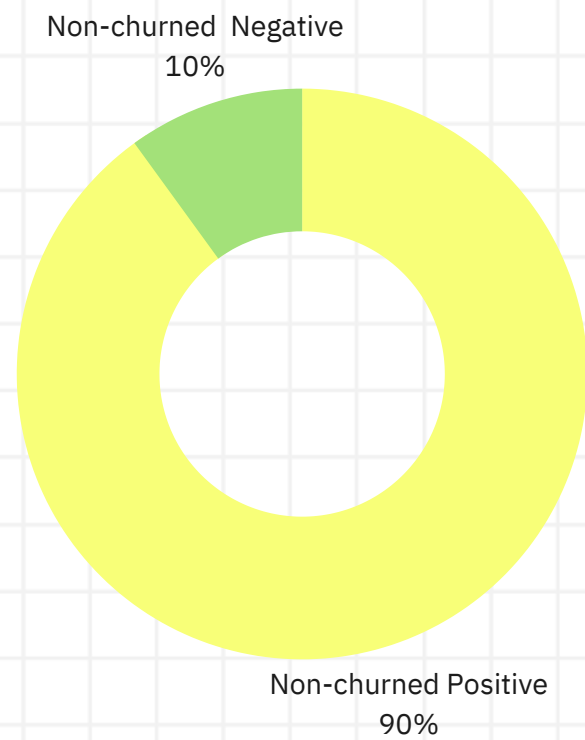


# Гистограмма длительности подписки

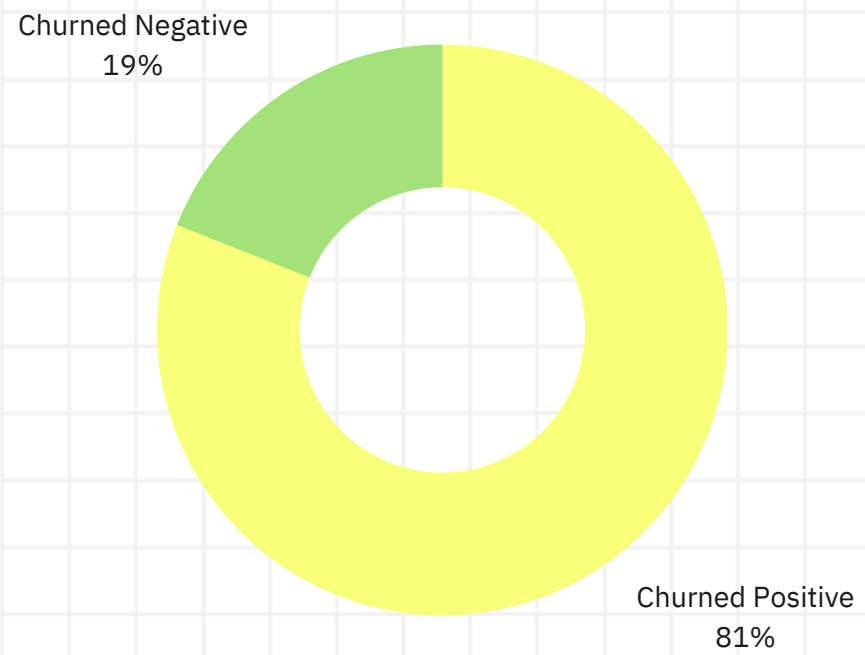
Датасет с None значениями

**Здесь отчётливо видно что основной  
отток людей происходит у людей с  
подпиской от 28 до 42 месяцев**

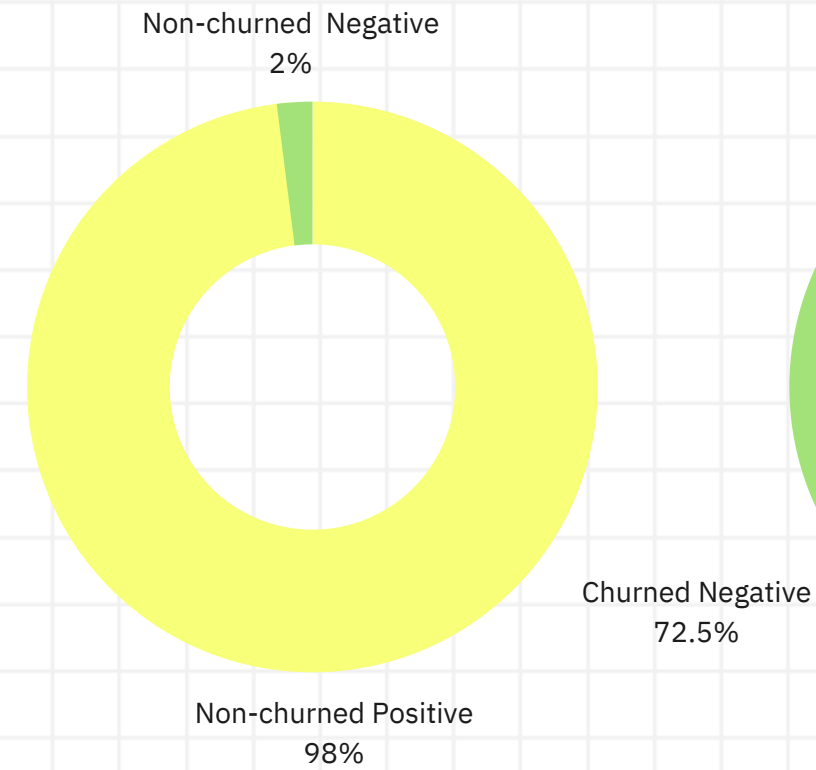




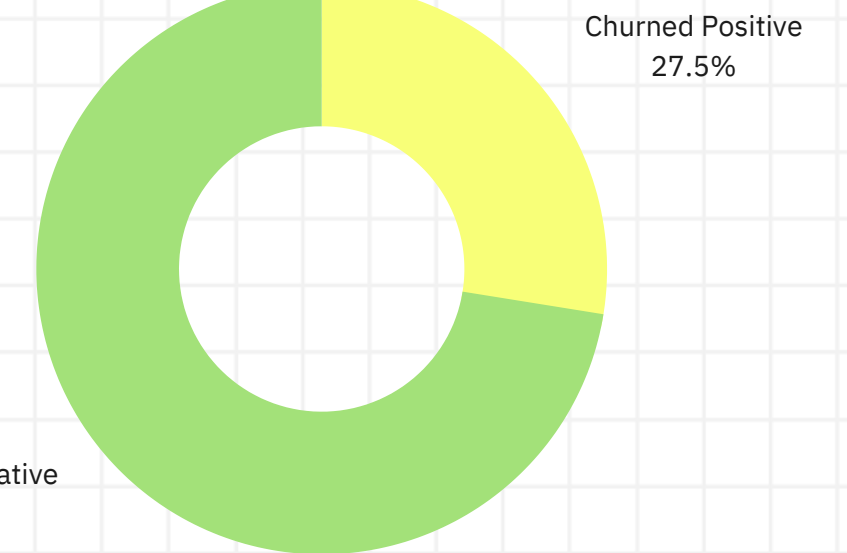
**precision**  
Logistic Regression



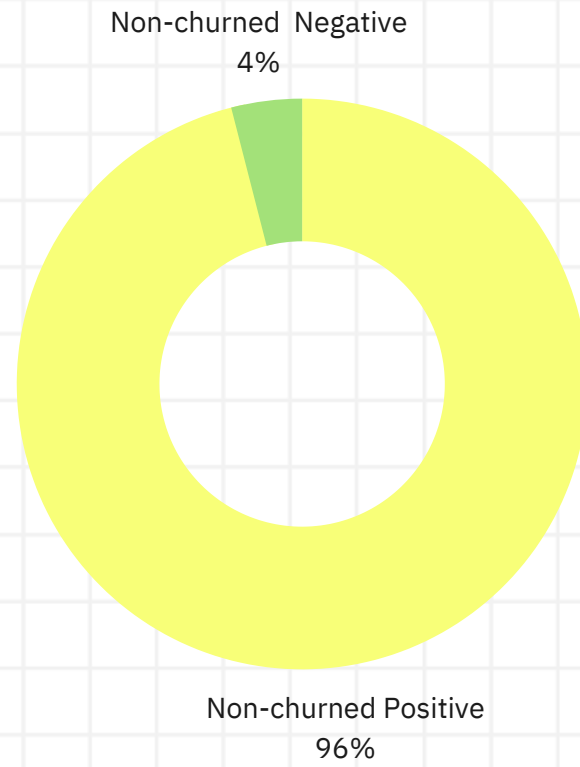
**precision**  
Logistic Regression



**recall**  
Logistic Regression

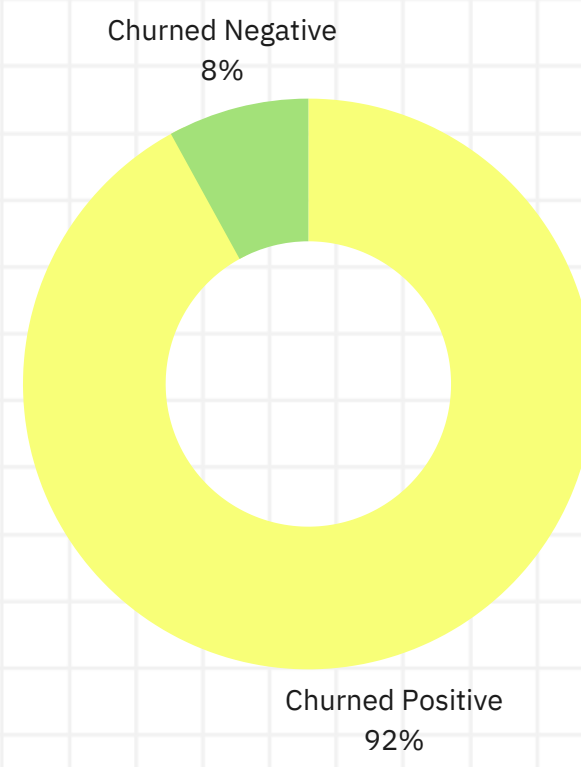


**recall**  
Logistic Regression



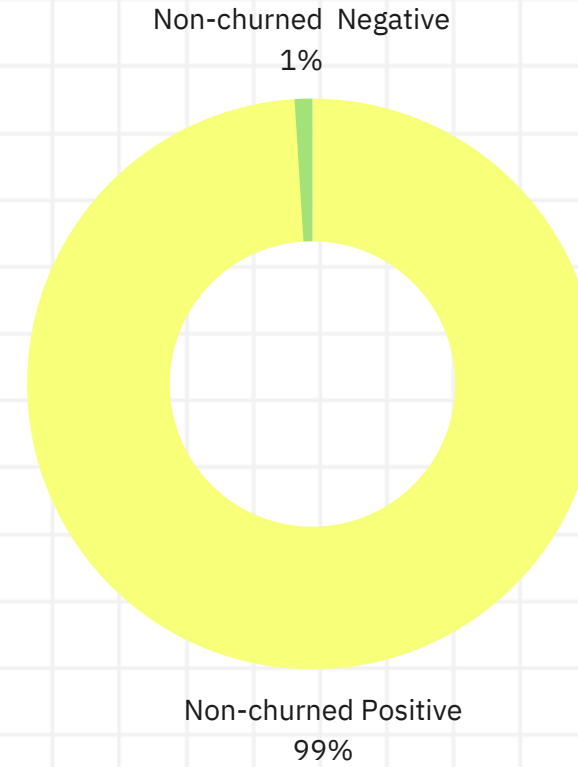
**precision**

Random Forest Classifier



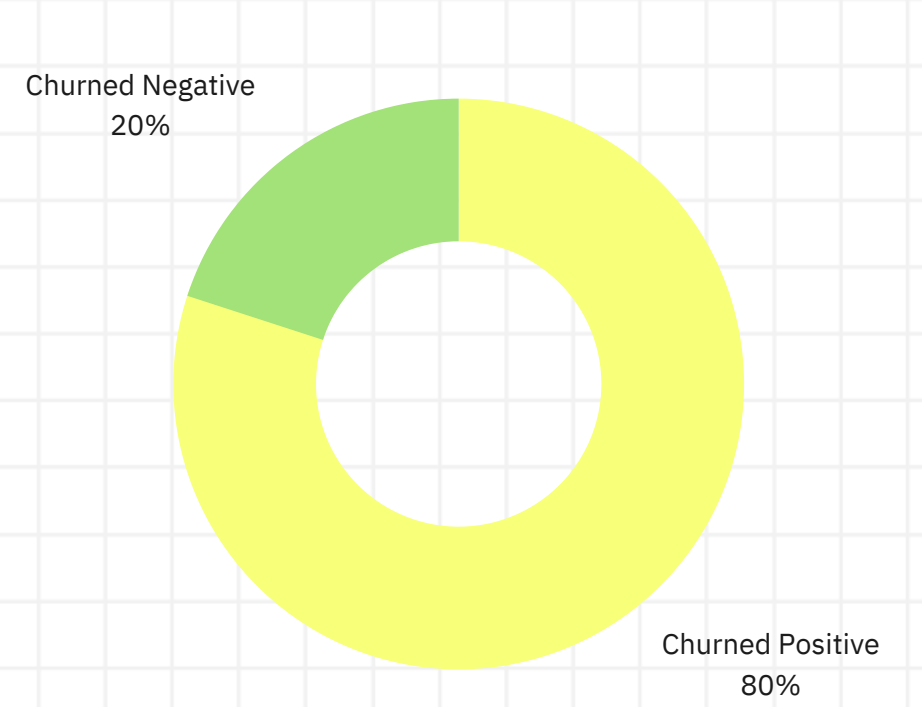
**precision**

Random Forest Classifier



**recall**

Random Forest Classifier



**recall**

Random Forest Classifier



# Заключение