

# Tinkoff Data Science Challenge



## Задача 1:

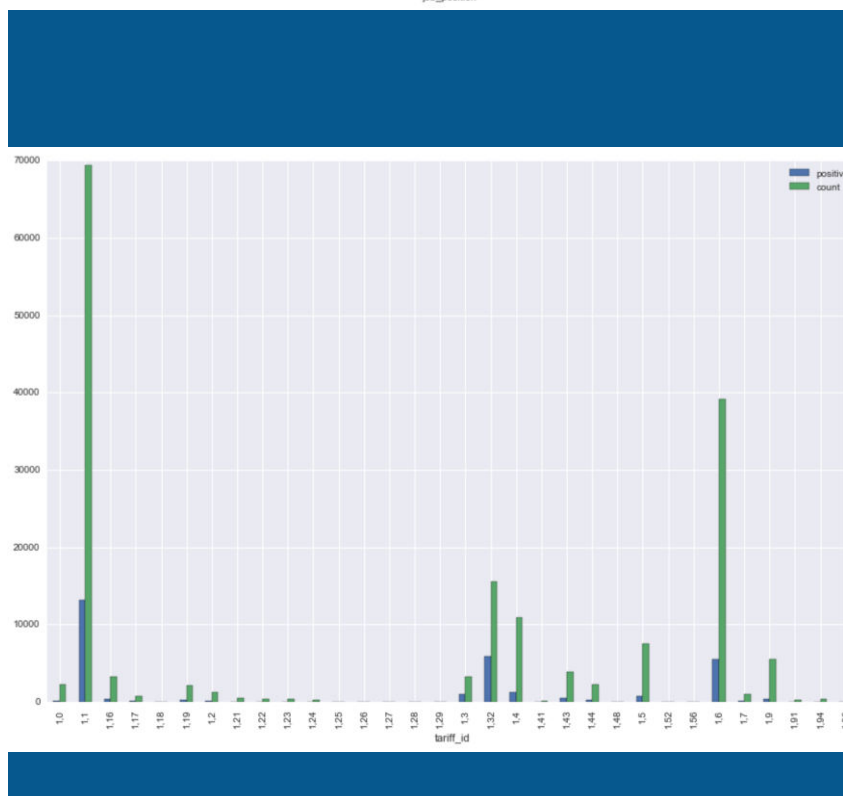
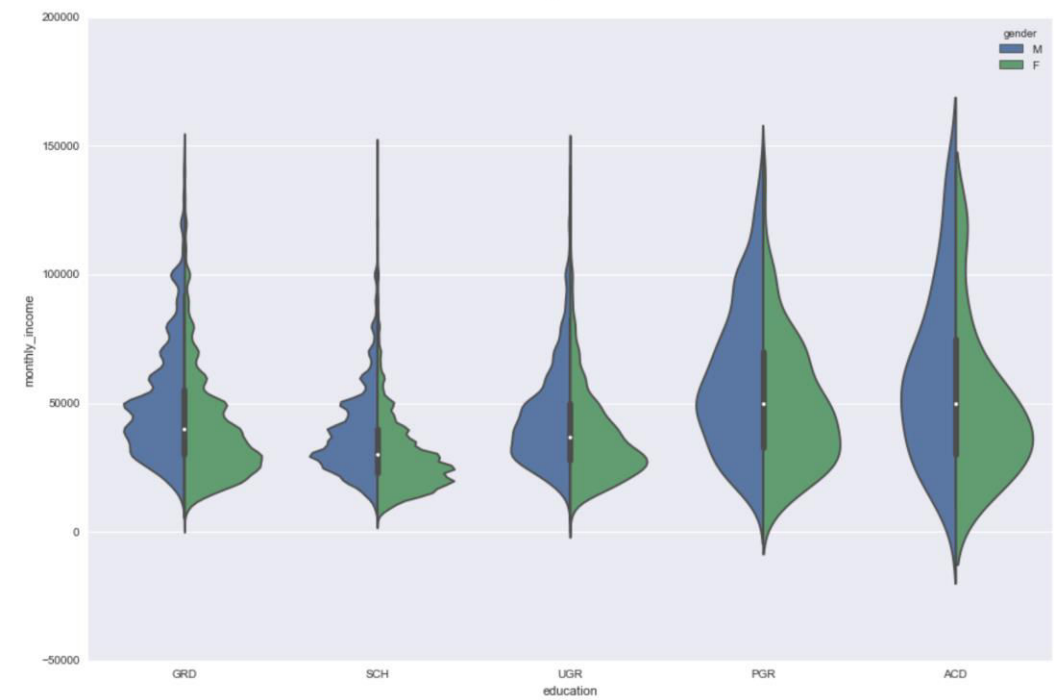
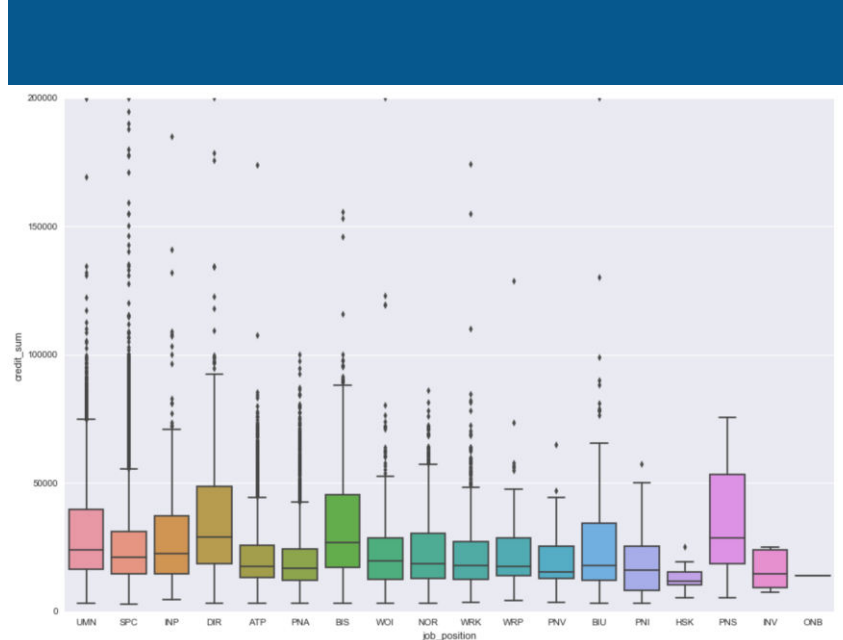
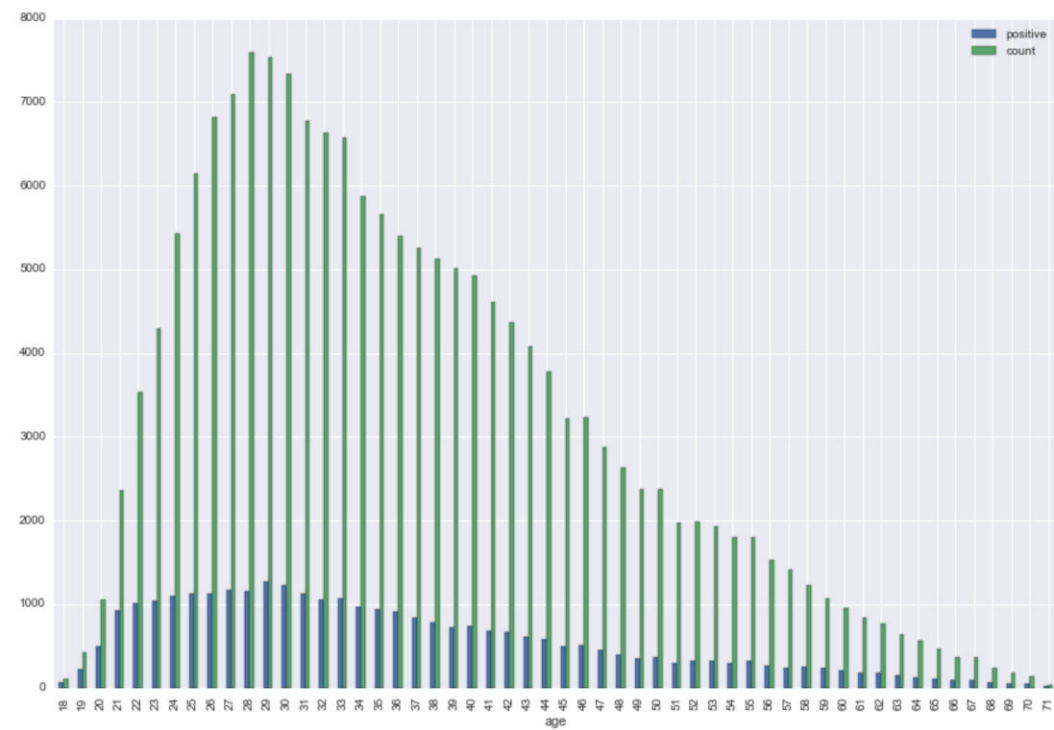
Датасет содержит данные о кредитах на покупку электроники, которые были одобрены Tinkoff.ru. Необходимо предсказать, выберет ли покупатель кредит от Tinkoff.ru

### Данные:

170 000 – train, средняя вероятность отклика  $p=0.175$   
91 000 - test

**Метрика:** ROC-AUC

Several white lines of varying lengths and angles are drawn in the bottom right corner of the slide, creating a modern, abstract graphic element.

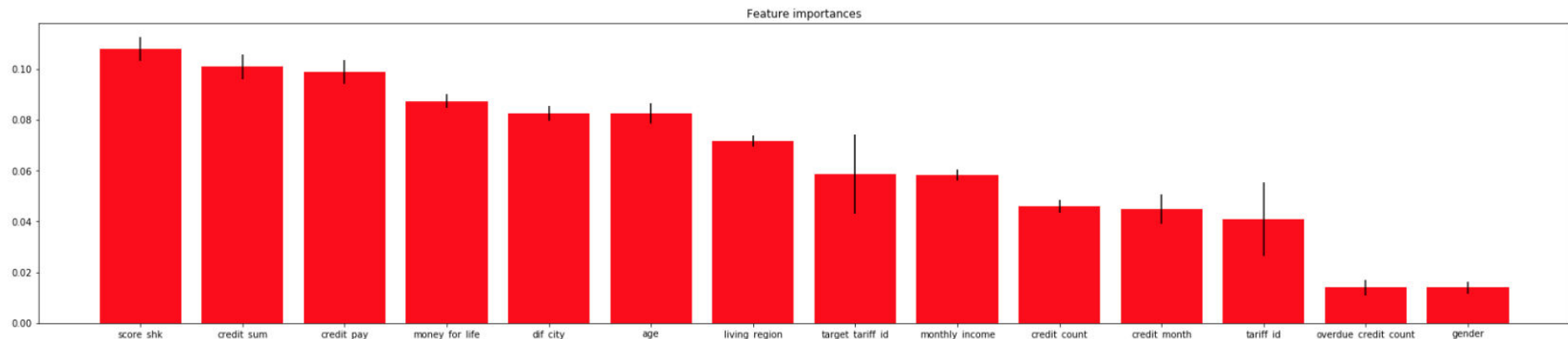


## Решения

1. 'living\_region' содержал 381 уникальное значение.  
После обработки → 84
2. Добавить признаки, связанные с кредитными рисками:
  - $\text{score\_shk} / (\text{mean}(\text{score\_shk}))$
  - $\text{credit\_sum} / \text{credit\_month}$
  - зарплата – плата по кредиту
  - разница зарплаты и средней по региону
  - Null → медиана



0.766 на LB



Нетипичные условия кредита → повышение отклика

Хорошие признаки:

- частота появления такой же суммы кредиты, срока кредита
- квантили по сумме кредита, score\_shk, частота суммы кредита

Взять несколько алгоритмов и сделать стэкинг  
(используем rank( ) от вероятности прогноза)