

Отчет по заданию Kaggle.

Выполнила Абдраманова Яна, БПМ-17-2.

Набор данных:

Представленный в работе набор данных показывает сведения о клиентах банка и состоит из 10 000 клиентов, указывающих свой возраст, зарплату, семейное положение, лимит кредитной карты, категорию кредитной карты и т. д.

Цель работы:

Визуализировать наиболее релевантные переменные, влияющие на клиентов.

Основная часть:

Импортируем данные и проверяем корректность работы.

Таблица содержит:

- CLIENTNUM - Номер клиента. Уникальный идентификатор клиента, владеющего счетом;
- Attrition_Flag - Активность клиента - если счет закрыт, то 1 иначе 0;
- Customer_Age — возраст клиента;
- Dependent_count - число иждивенцев;
- Education_Level - образовательная квалификация владельца счета (пример: выпускник средней школы, колледжа и т. д.);
- Gender — пол клиента (M = мужчина, F = женщина);
- Marital_Status - Женат, Холост, Разведен, Неизвестен;
- Income_Category - категория годового дохода владельца счета;
- Card_Category - тип карты (Синяя, серебряная, золотая, Платиновая);
- Months_on_book - период взаимоотношений с банком.

подключаем библиотеки:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import sys
```

Загружаем данные, преобразовываем, заменяем лексические значения цифрами и удаляем NaN.

Функция `correct_indexing(dataset)` корректирует индексы для дальнейшей работы.

Функция `catagory_churn_unchurn(type_of_card, data)` производит набор данных на основе типа кредитной карты клиента.

Функция `plotter(column, Group)` строит графики данного столбца 1 против столбца 2 и для 5 наборов данных.

1. полный набор

2. данных синие держатели кредитных карт
3. серебряные держатели кредитных карт
4. золотые держатели кредитных карт
5. платиновые держатели кредитных карт

Графики находятся в этом конкретном порядке. Результат работы программы:

```
plotter('Contacts_Count_12_mon', 'Attrition_Flag')
```

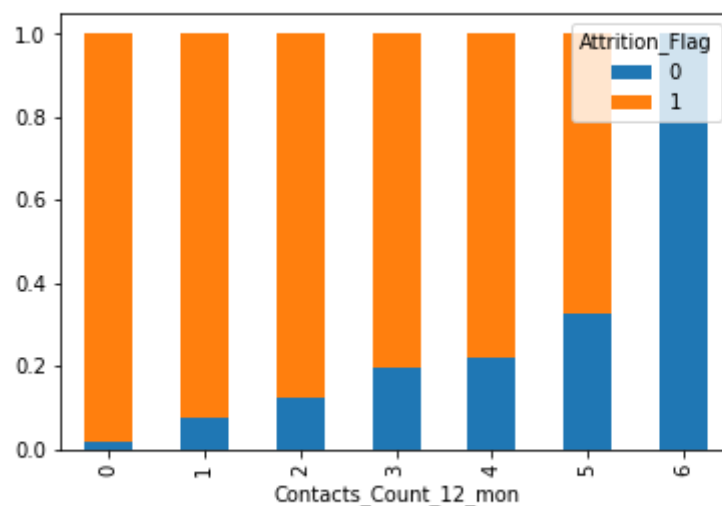
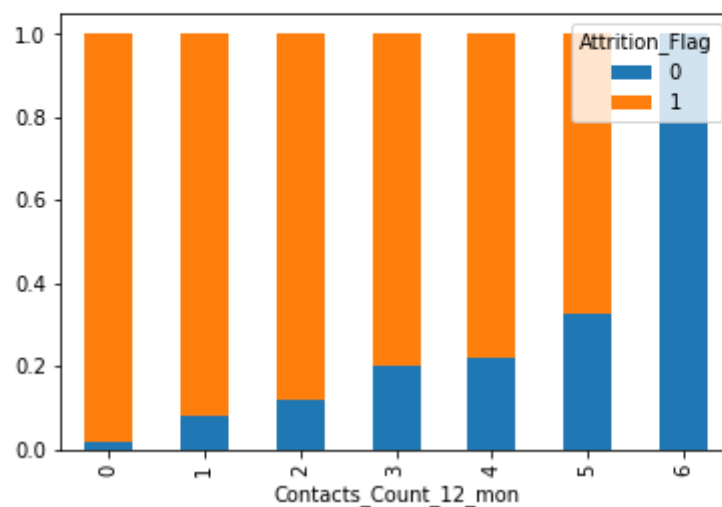


Рис. 1: Результат работы функции plotter.

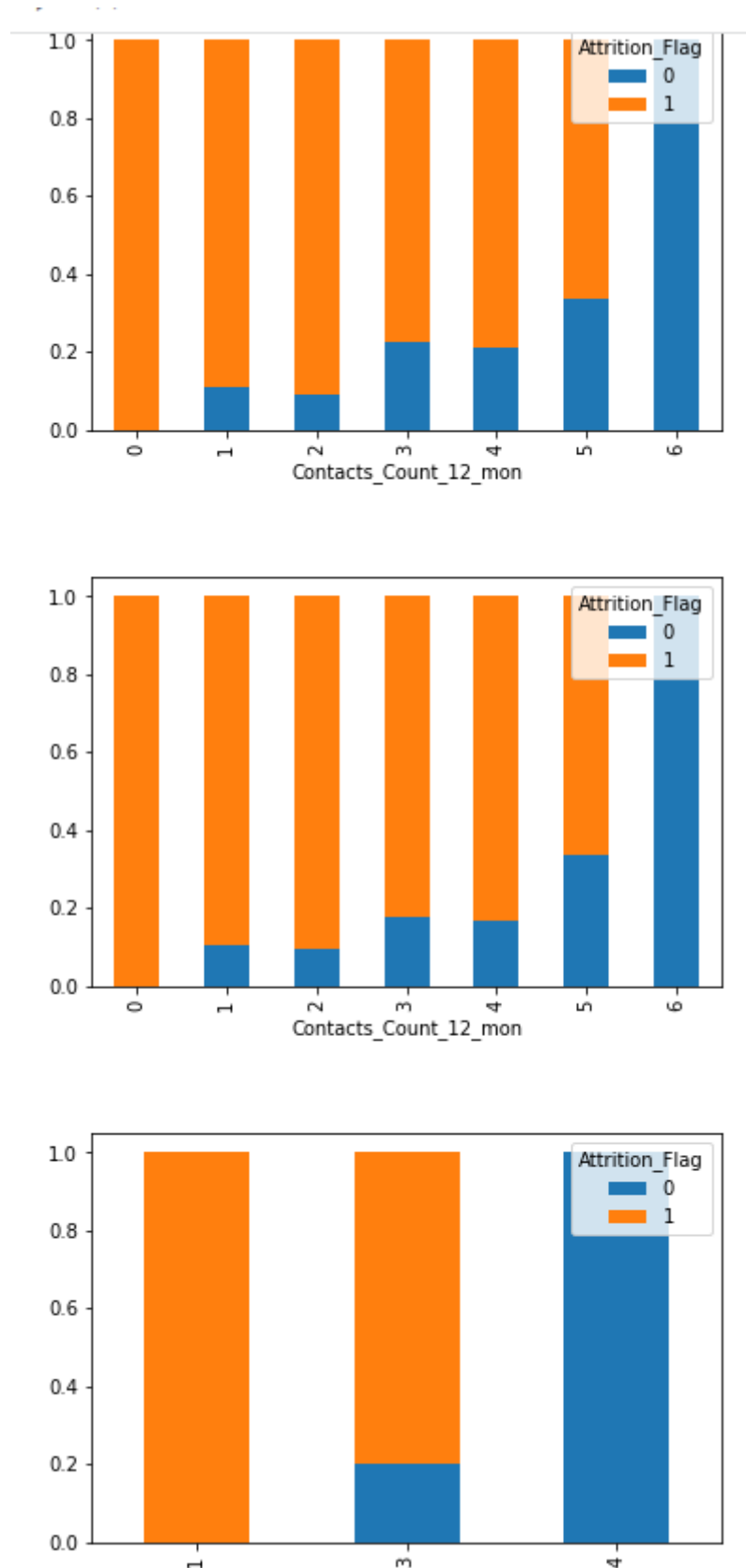


Рис. 2: Contacts_Counts_12_mon показывает слабую отрицательную корреляцию с Attrition_flag клиента, но очень последовательна, и эта связь может быть использована для прогнозирования будущих вспенивающих клиентов.

```
:
plotter('Total_Revolving_Bal', 'Attrition_Flag')
```

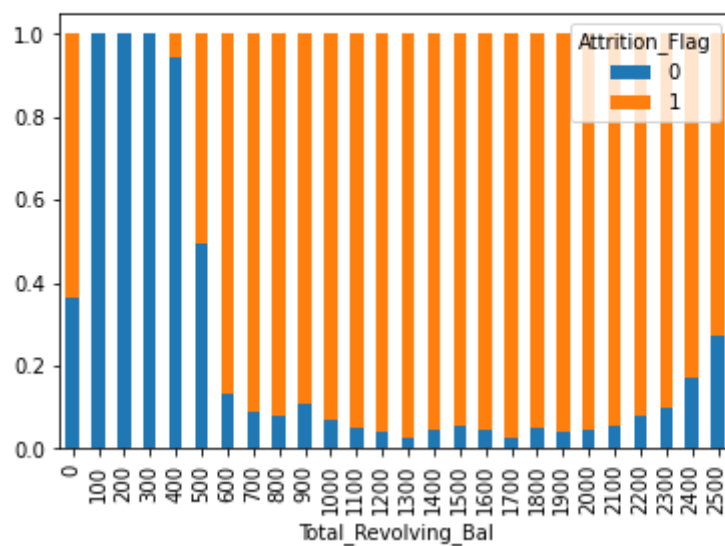
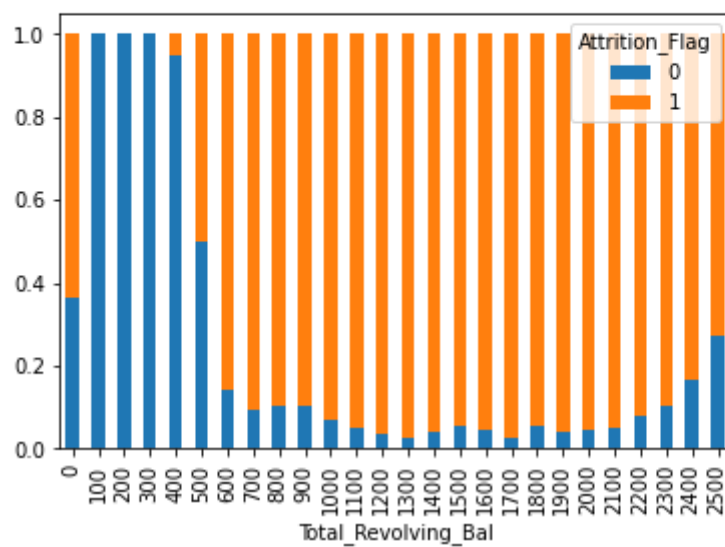


Рис. 3: Результат работы функции plotter.

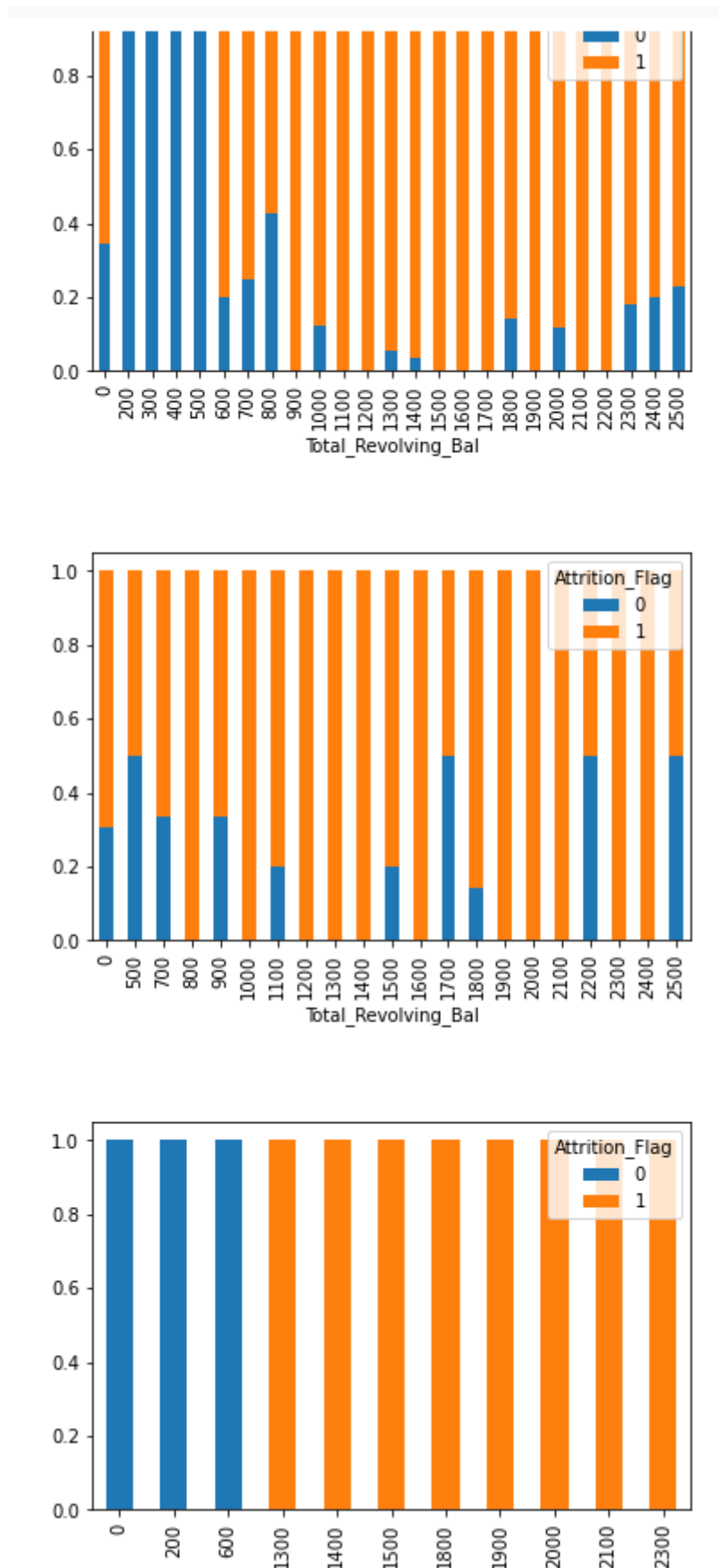


Рис. 4: Total_Revolving_Bal показывает сильную положительную корреляцию с Attrition_flag клиента и очень последовательна, и эта связь может быть использована для прогнозирования будущих вспенивающихся клиентов

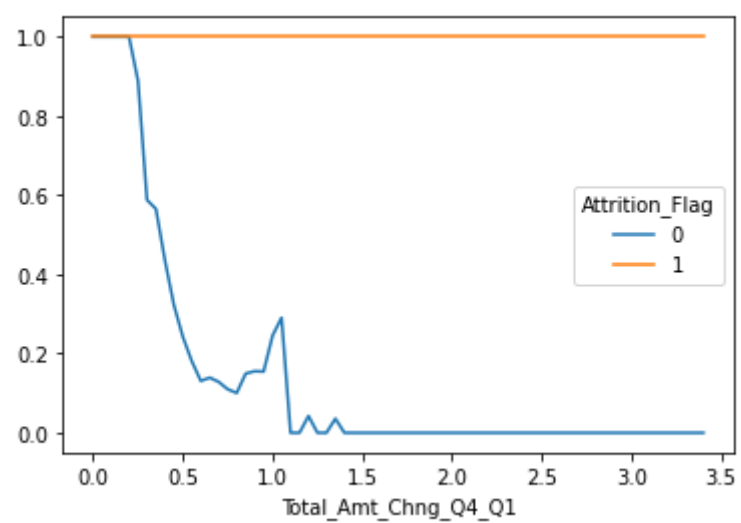
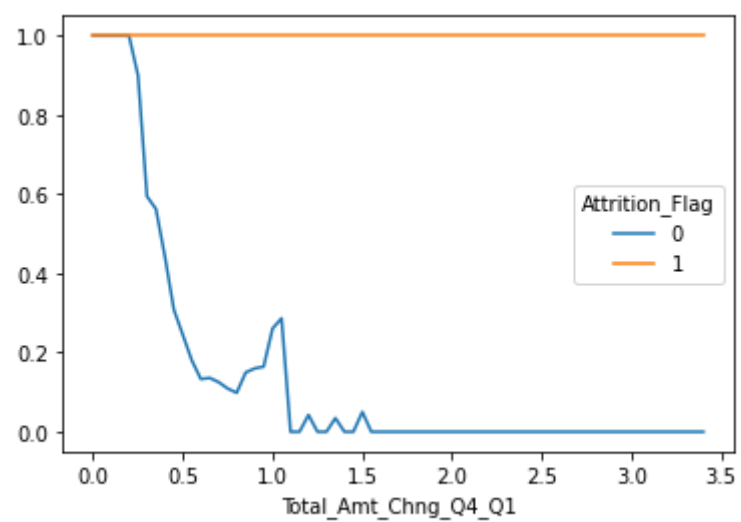


Рис. 5: Результат работы функции plotter.

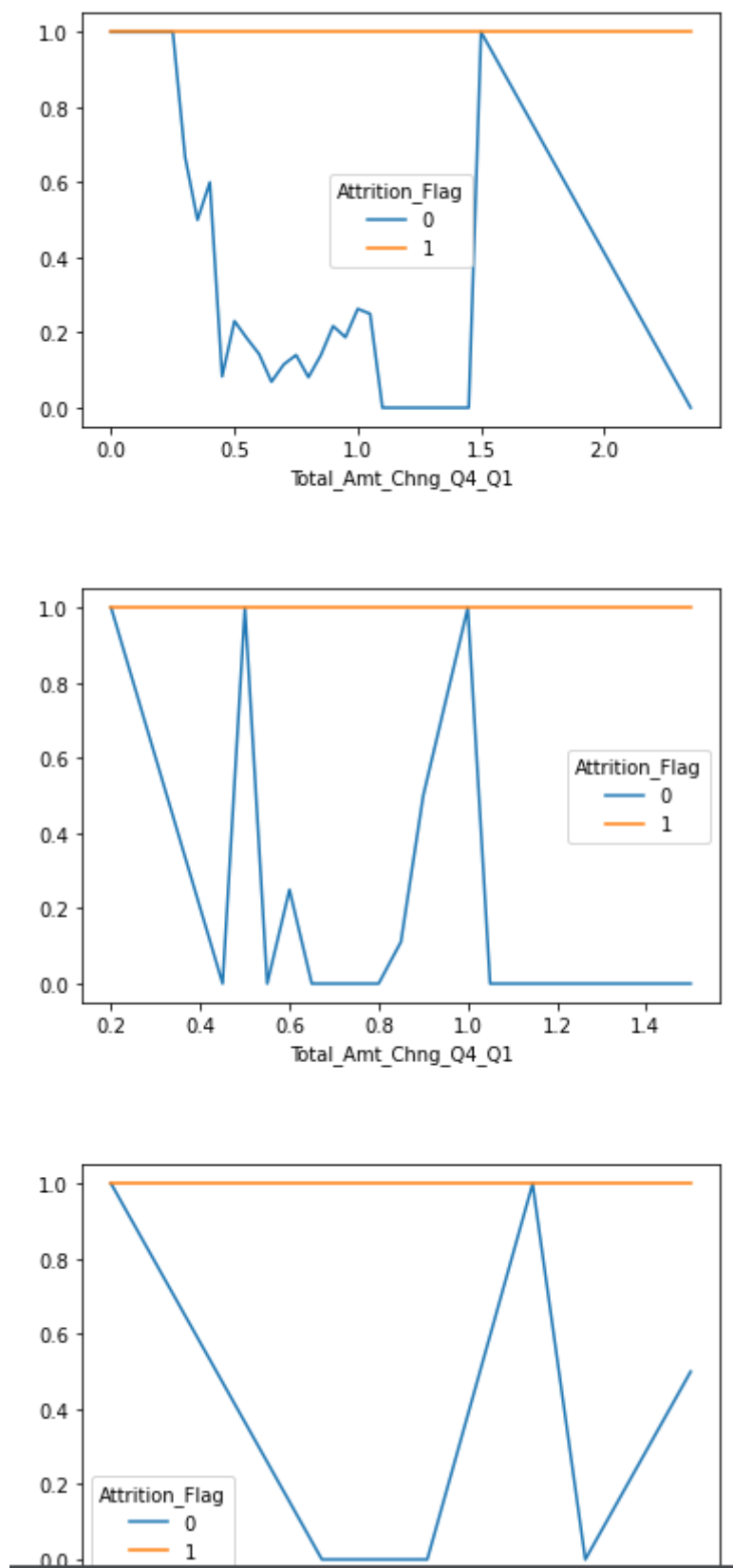


Рис. 6: Total_Amt_Chng_Q4_Q1 показывает сильную положительную корреляцию с Attrition_flag клиента и очень последовательна, и эта связь может быть использована для прогнозирования будущих вспенивающихся клиентов