

Банк «Метанпром» — Сегментация пользователей по потреблению

Сроки: с 07.02.2022 по 10.02.2022

Автор: Александр Смирнов

Контакт для связи: alex-alex@yandex.ru

Оглавление

1. Общая информация об исследовании
2. Значения, упоминаемые в презентации
3. Сегментация пользователей по количеству потребляемых продуктов
4. Основания для разделения (churn)
5. Верхний уровень сегментации (группы)
6. Рекомендации для группы 1-2 продукта
7. Рекомендации для группы 3-4 продукта
8. Нижний уровень сегментации (подгруппы 1 и 2)
9. Нижний уровень сегментации (подгруппы 3 и 4)
10. Проверка гипотезы 1
11. Проверка гипотезы 2
12. Общий вывод

Общая информация об исследовании

Цели исследования. Сегментация пользователей банка по количеству потребляемых продуктов и проверка гипотез.

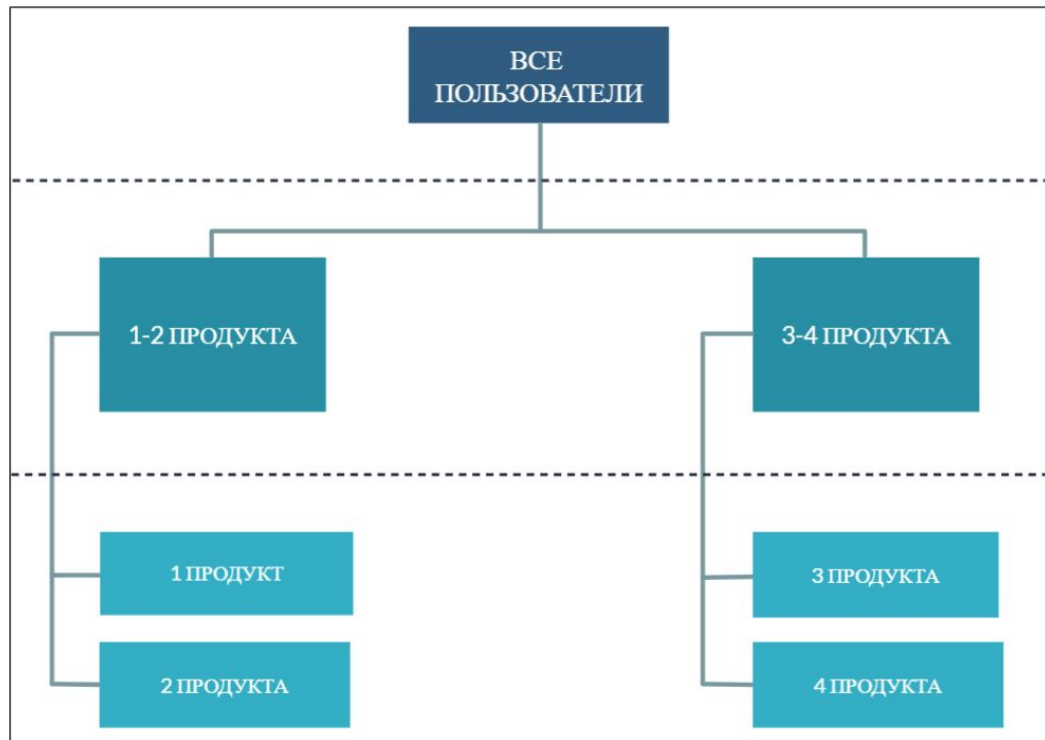
Результат. Сегментация, рекомендации по группам и проверка гипотез.

Источник данных. Для исследования был использован файл *bank_dataset.csv*, полученный из базы данных.

Значения, упоминаемые в презентации

Название поля	Расшифровка
<i>gender_binary</i>	доля женщин
<i>age</i>	средний возраст
<i>credit_card</i>	доля клиентов с кредитной картой
<i>loyalty</i>	доля активных клиентов
<i>churn</i>	доля клиентов, ушедших в отток
<i>objects</i>	среднее количество объектов недвижимости
<i>score</i>	средний кредитный рейтинг
<i>estimated_salary</i>	средняя заработная плата

Сегментация пользователей по количеству потребляемых продуктов



Группы:

- 1-2 продукта
- 3-4 продукта

Подгруппы:

- 1 продукт
- 2 продукта

-
- 3 продукта
 - 4 продукта

Основания для разделения (churn)

Количество продуктов		Процент оттока
	1	27.71%
	2	7.58%
	3	82.71%
	4	100.00%

Основной признак для разделения - **доля оттока**.

Для группы 1-2 продукта он составляет *в среднем 18.2%*.

Для группы 3-4 продукта он составляет *в среднем 85.9%*.

Верхний уровень сегментации (группы)

	1_2_products	3_4_products
gender_binary	0.450279	0.57362
age	38.7624	43.6534
credit_card	0.705396	0.708589
loyalty	0.51778	0.435583
churn	0.181621	0.858896
objects	5.01127	5.05828
score	650.576	649.113
estimated_salary	99945	104400

В таблице **цветом** выделены ячейки, на основании значений которых было проведено разделение.

Рекомендации для группы 1-2 продукта

- Следует поощрять активность клиентов, внедряя систему кэшбеков и скидок по дебетовым и кредитным картам;
- Так как клиенты этой группы более молодые, для них было бы актуальна покупка жилья с использованием материнского капитала. Прием материнского капитала в качестве первоначального взноса позволит увеличить заинтересованность в услугах этой группы клиентов;
- Для этой группы наиболее актуально предложение "Приведи друга": клиент приглашает знакомого и они оба получают бонус в денежной или иной форме.

Рекомендации для группы 3-4 продукта

- Провести интервью по телефону и/или по интернету с клиентами из этой группы и выяснить причину их ухода;
- В случае, если их уход связан с более выгодными условиями по кредиту в другом банке, рассмотреть возможность уменьшения нашей ставки по кредиту;
- В случае, если их уход не связан с вышеизложенной причиной и не связан с неразрешимой проблемой, предложить возобновить сотрудничество с льготными условиями на первое время. Это может быть, например, повышенный кэшбек или бесплатное обслуживание карты.

Нижний уровень сегментации (подгруппы 1 и 2)

	1_product	2_products
gender_binary	0.451613	0.448802
age	39.6731	37.7536
credit_card	0.703777	0.70719
loyalty	0.504131	0.532898
churn	0.277144	0.075817
objects	4.97463	5.05185
score	649.121	652.189
estimated_salary	99487.3	100452

В таблице **цветом** выделены ячейки, на основании значений которых было проведено разделение.

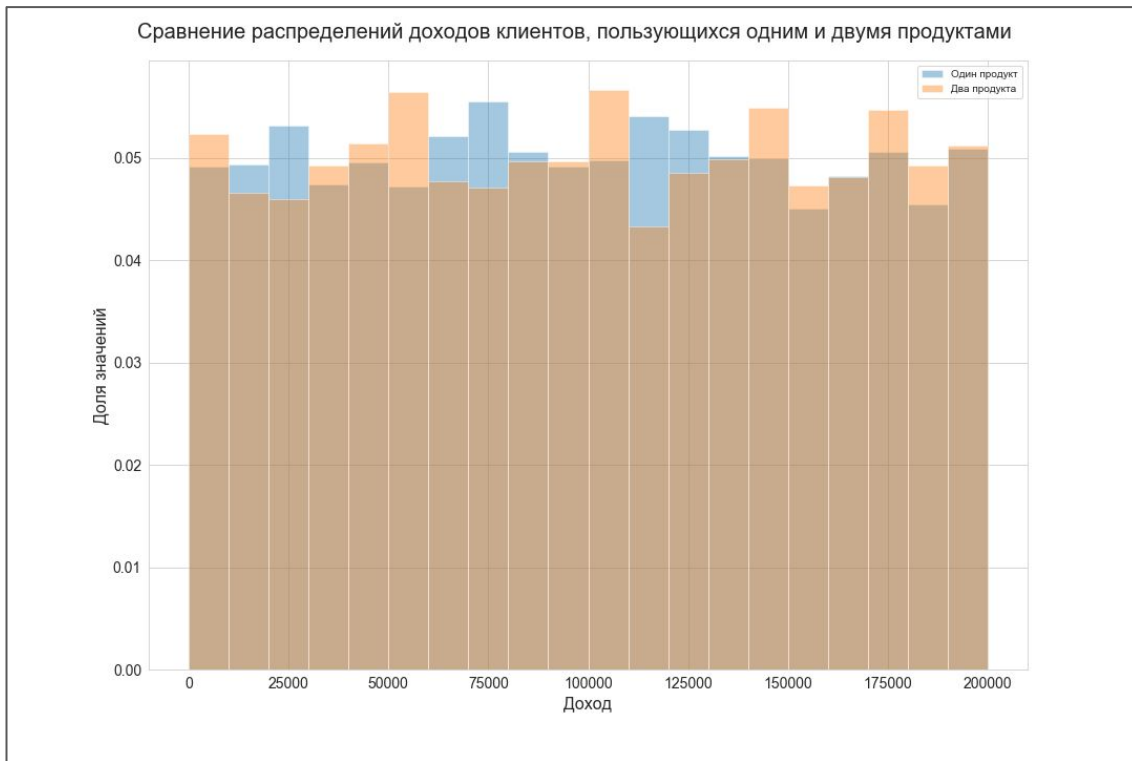
Нижний уровень сегментации (подгруппы 3 и 4)

	3_products	4_products
gender_binary	0.56015	0.633333
age	43.1955	45.6833
credit_card	0.714286	0.683333
loyalty	0.424812	0.483333
churn	0.827068	1
objects	5.00376	5.3
score	648.105	653.583
estimated_salary	104318	104764

В таблице **цветом** выделены ячейки, на основании значений которых было проведено разделение.

С подробными рекомендациями по подгруппам можно ознакомиться в файле Jupyter Notebook.

Проверка гипотезы 1



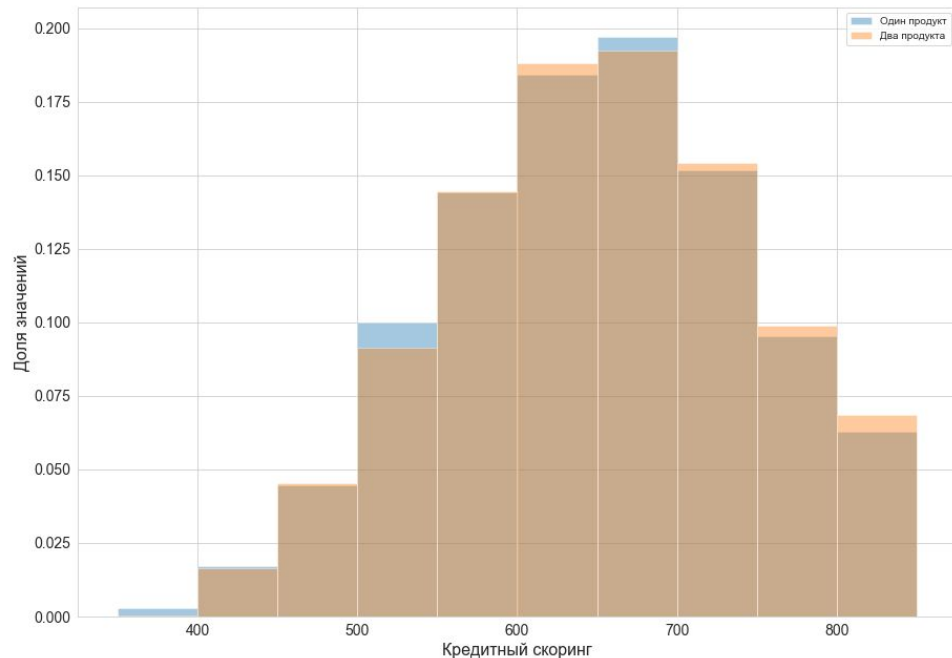
«Средний доход клиентов, которые пользуются одним продуктом банка, равен среднему доходу клиентов, которые пользуются двумя продуктами банка»

Гипотезу не удалось отвергнуть.

p-value = 0.0786

Проверка гипотезы 2

Сравнение распределений кредитного скоринга клиентов, пользующихся одним и двумя продуктами



«Среднее значение кредитного скоринга клиентов, которые пользуются одним продуктом, равно среднему значению кредитного скоринга клиентов, которые пользуются двумя продуктами»

Гипотезу не удалось отвергнуть.

p-value = 0.0786

Общий вывод

Сегментация: в результате проведения сегментации разделили клиентов на две группы: *1-2 продукта* и *3-4 продукта*. Внутри этих групп также имеются отличия, поэтому требуется дополнительное разделение на подгруппы (*1 продукт, 2 продукта и 3 продукта, 4 продукта*).

Проверка гипотез:

1. Средний доход клиентов с одним продуктом равен среднему доходу клиентов с двумя продуктами.
2. Среднее значение кредитного скоринга клиентов с одним продуктом равно среднему значению кредитного скоринга клиентов с двумя продуктами.