

# **Презентация аналитической работы для фирмы спортивной атрибутики**

**Автор: Теребиленко Матвей**

# Бизнес-проблема

Все этапы работы связаны аналитикой данных в целях того, чтобы увеличить выручку компании. Задачами были построить модель предсказания покупки, проанализировать результаты эксперимента, найти предрасположенность клиентов к товарам, этими методами можно значительно увеличить выручку компании. Без значительного роста выручки развитие компании невозможно, поэтому проблема заключается в отсутствии значимого роста выручки

# Этапы работы

- Выгрузка данных и восстановление поврежденных данных с помощью модели бинарной классификации
- Создание единой витрины данных
- Проведения A/B-тестирования и интерпретация результатов
- Кластеризация данных
- Ассоциативный анализ
- Построение модели склонности клиента к покупке

# Создание единой витрины данных

- Для восстановления поврежденных данных использовалась модель логистической регрессии
- Пропуски были в цвете и гендерной принадлежности товара
- Пропуски в цвете были заполнены - Unknown
- Пропуски в гендерной принадлежности товара были заполнены на основании гендера покупателя
- Размер итогового датафрейм: 790000 строк

id	0
gender	0
age	0
education	0
city	0
country	0
lbt_coef	0
ac_coef	0
sm_coef	0
personal_coef	0
product	0
colour	119524
cost	0
product_sex	314712
base_sale	0
dt	0



# Рекомендация по хранению данных

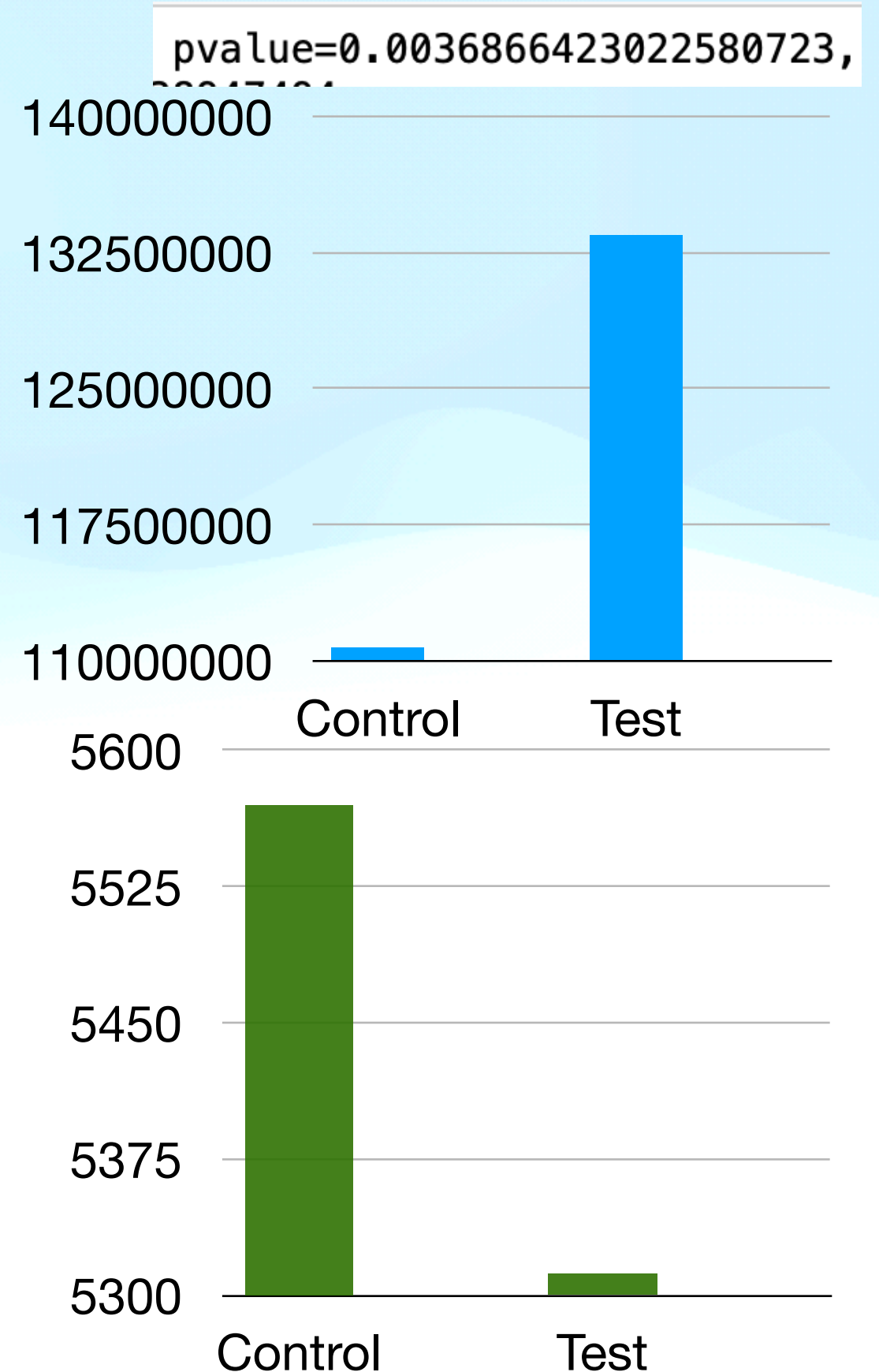
Так как все данные используются моделью для предсказания конверсии в покупку, стоит улучшить этап первичного вноса данных о товаре для минимализации количества пропусков

Вид итоговой витрины данных:

id	gender	age	education	city	country	lbt_coef	ac_coef	sm_coef	personal_coef	product	colour	cost	product_sex	base_sale	dt
0	0	36	среднее	1201	32	5.078678	-0.307147	0.959027	0.5072	Стол Outventure	зелёный	1499	NaN	0	37
0	0	36	среднее	1201	32	5.078678	-0.307147	0.959027	0.5072	Набор Outventure: стол + 4 стула	бежевый	4799	NaN	0	37
4	0	35	среднее	1134	32	4.569378	0.063693	0.820892	0.5072	Роликовые коньки детские Roces Orlando III	черный/ розовый	9199	NaN	0	0
4	0	35	среднее	1134	32	4.569378	0.063693	0.820892	0.5072	Шлем FILA NRK FUN	белый	4999	NaN	0	0
4	0	35	среднее	1134	32	4.569378	0.063693	0.820892	0.5072	Роликовые коньки детские Roces Yuma TIF	мультицвет	6999	NaN	0	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
16355	1	32	среднее	1188	32	7.084163	-0.132395	0.509138	0.4304	Велосипед горный Trek Marlin 6 29", 2022	черный	69999	NaN	0	10
16355	1	32	среднее	1188	32	7.084163	-0.132395	0.509138	0.4304	Костюм-сауна Everlast	чёрный	4299	NaN	0	23
15885	1	33	среднее	1190	32	7.321749	0.470877	0.699928	0.4304	Бейсболка Columbia Mesh Snap Back - High	серый	1999	NaN	0	3
15885	1	33	среднее	1190	32	7.321749	0.470877	0.699928	0.4304	Массажер Demix	тёмно- зелёный	399	NaN	0	52
15885	1	33	среднее	1190	32	7.321749	0.470877	0.699928	0.4304	Тренажер гирскопический KETTLER	голубой	1799	NaN	0	52

# Интерпретация результатов A/B теста

- Средняя выручка у контрольной группы больше на 4%
- Общая выручка больше у тестовой группы на 19%
- Количество клиентов в тестовой группе больше на 25%
- Эксперимент можно разворачивать

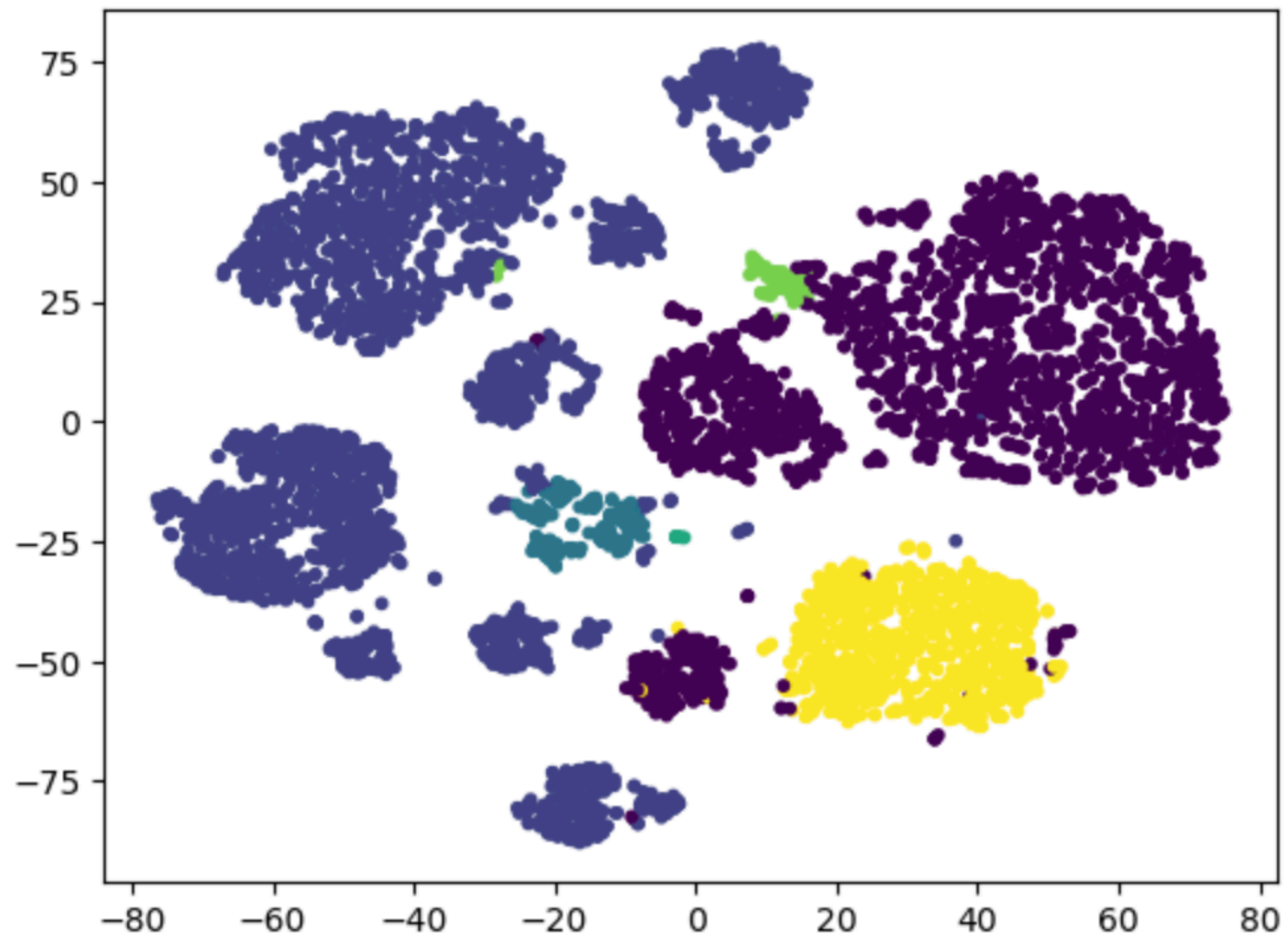
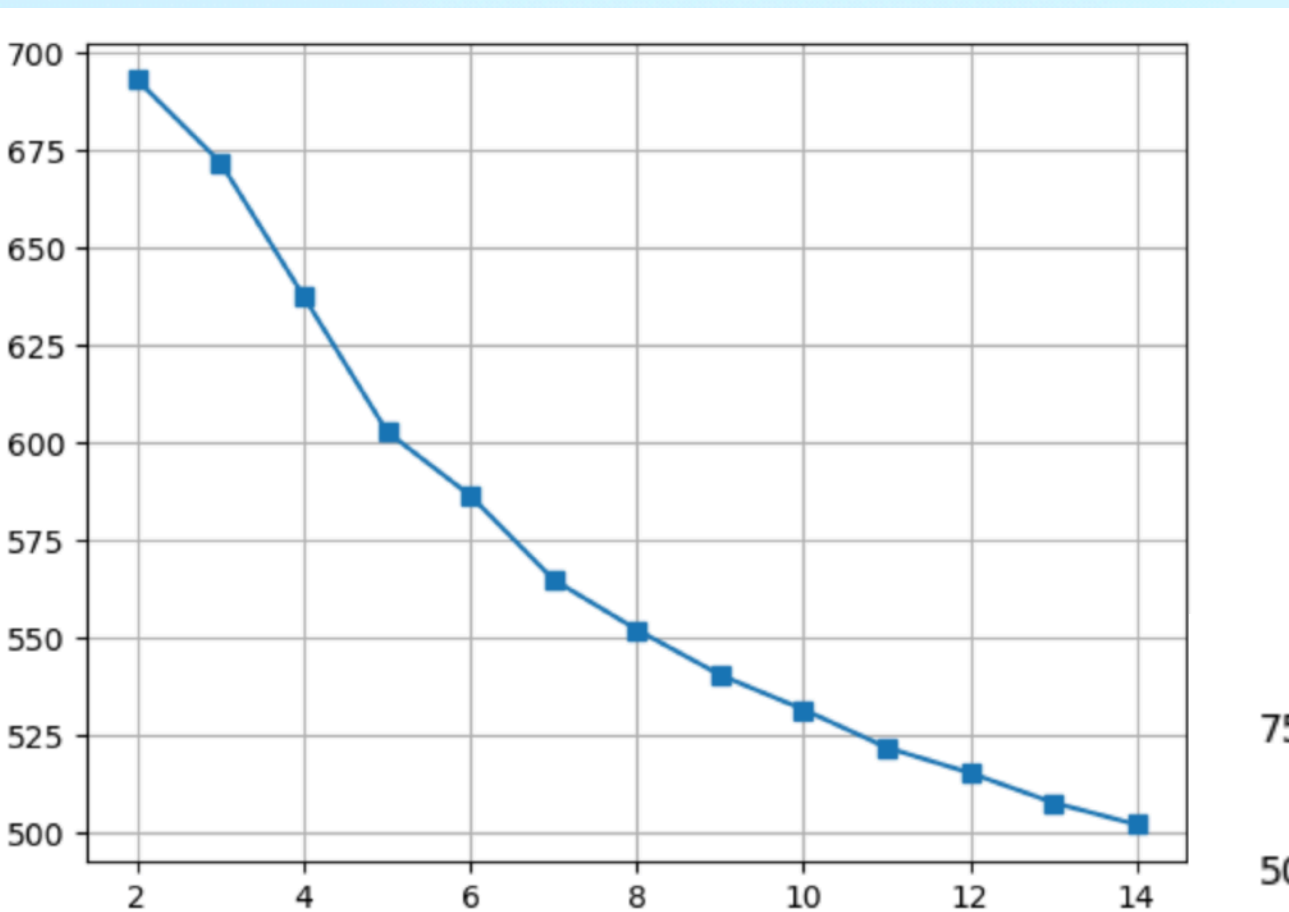


# **Рекомендация по результатам эксперимента**

**Можно применить скидку ко всей выборке клиентов, это даст прирост в общей выручке, но следует увеличить метрику средней выручки, рекомендую провести A/B тест для проверки следующей гипотезы — Если предложить клиентам скидку меньше, чем в предыдущем эксперименте, средняя выручка вырастет на 5%, при проведении эксперимента обратить внимание так же на метрику количества клиентов.**



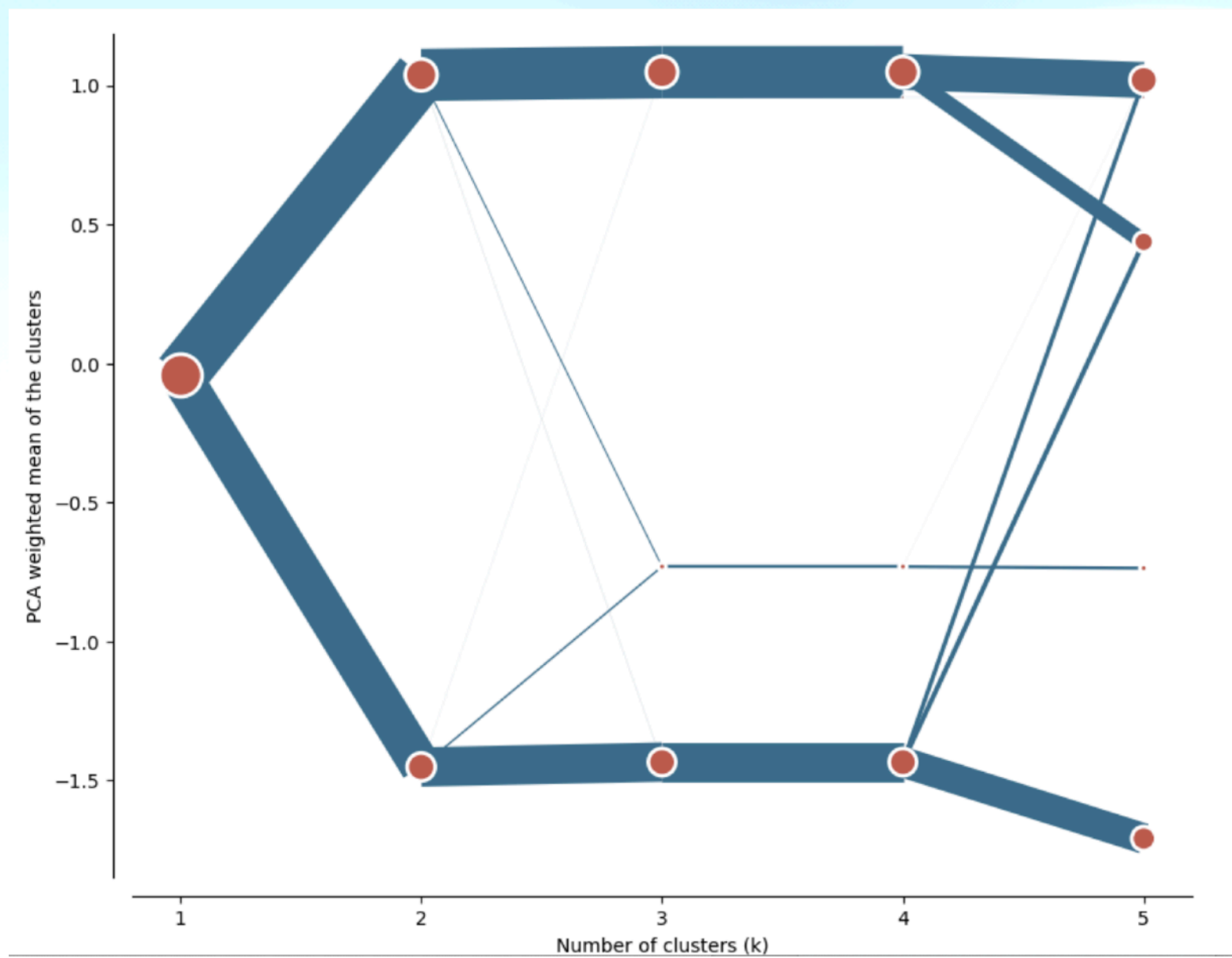
# Кластерный анализ





# Рекомендация по результатам кластеризации

Вследствие проведения кластерного анализа, можно сказать, что идеальное количество разбиений на кластеры это 3, то есть при следующих экспериментах делите выборку на 3 кластера.



# Ассоциативный анализ

- Порог значимости был установлен 0.04
- Был создан столбец - была ли совершена покупка
- Выявил товары с большей конверсией в покупку у всех клиентов
- Выявил товары с большей конверсией в покупку у клиентов города 1188
- Связал, что люди с средним образованием чаще не переходят к покупке

```
{no} -> {среднее},  
{Брюки женские Outventure} -> {yes},  
{Брюки мужские Demix} -> {yes},  
{Брюки мужские FILA} -> {yes},  
{Брюки мужские Outventure} -> {yes},  
{Купальник женский Joss} -> {yes},  
{Лиф женский Joss} -> {yes},  
{Плавки женские Joss} -> {yes},  
{Футболка женская Demix} -> {yes},  
{Футболка мужская Demix} -> {yes},  
{Шорты мужские Demix} -> {yes},  
{высшее} -> {yes},  
{среднее} -> {yes},
```

---

```
{no} -> {среднее},  
{Брюки женские Outventure} -> {yes},  
{Брюки мужские Demix} -> {yes},  
{Брюки мужские Outventure} -> {yes},  
{Купальник женский Joss} -> {yes},  
{Лиф женский Joss} -> {yes},  
{Плавки женские Joss} -> {yes},  
{Футболка женская Demix} -> {yes},  
{Футболка мужская Demix} -> {yes},  
{Шорты мужские Demix} -> {yes},  
{высшее} -> {yes},
```

# Товары с высокой конверсией в покупку(все пользователи)

- Брюки женские Outventure
- Брюки мужские Demix
- Брюки мужские FILA
- Брюки мужские Outventure
- Купальник женский Joss
- Лиф женский Joss
- Плавки женские Joss
- Футболка женская/мужская Demix



# Товары с высокой конверсией в покупку(город 1188)

- Брюки женские Outventure
- Брюки мужские Demix
- Брюки мужские Outventure
- Купальник женский Joss
- Лиф женский Joss
- Плавки женские Joss
- Футболка женская/мужская Demix
- шорты Demix

# Модель склонности клиента к покупке

- Фичи: id, gender, age, city, education, country, personal\_coef, product, colour, cost, product\_sex, base\_sale, dt
- Предсказываемая переменная - buy(будет ли совершена покупка)
- Используется модель логистической регрессии
- Предсказывает на кросс-валидации с результатом 0.99

# **Рекомендация по использованию модели**

**Основываясь на большем разнообразии информации можно предсказать покупку, это повысит лояльность клиента и успешность маркетинговых компаний, поэтому применять модель стоит для определения склонности к покупке и при прогнозировании успешности дальнейших экспериментов.**



**Спасибо за внимание!**