

**E-commerce —  
Выявление профилей  
потребления интернет-магазина  
товаров для дома и быта  
«Пока все ещё тут»**



Смирнова Е.С.  
Яндекс Практикум, 2024 г.

# Цели исследования. Исходные данные.

- »»» Необходимо провести исследовательский анализ данных транзакций интернет-магазина товаров для дома и быта «Пока все ещё тут», категоризировать товары по содержанию, а также сегментировать покупателей на основе истории их покупок, проверить статистические гипотезы о различиях в среднем чеке и количестве заказов по сегментам покупателей;
- »»» На основании полученных сегментов покупателей выработать рекомендации менеджменту интернет-магазина по разработке персонализированных предложений для покупателей.

## Исходные данные:

Датасет с транзакциями интернет-магазина товаров для дома и быта «Пока все ещё тут» за период с 1 октября 2018 года по 31 января 2020 года.

# Этапы исследования

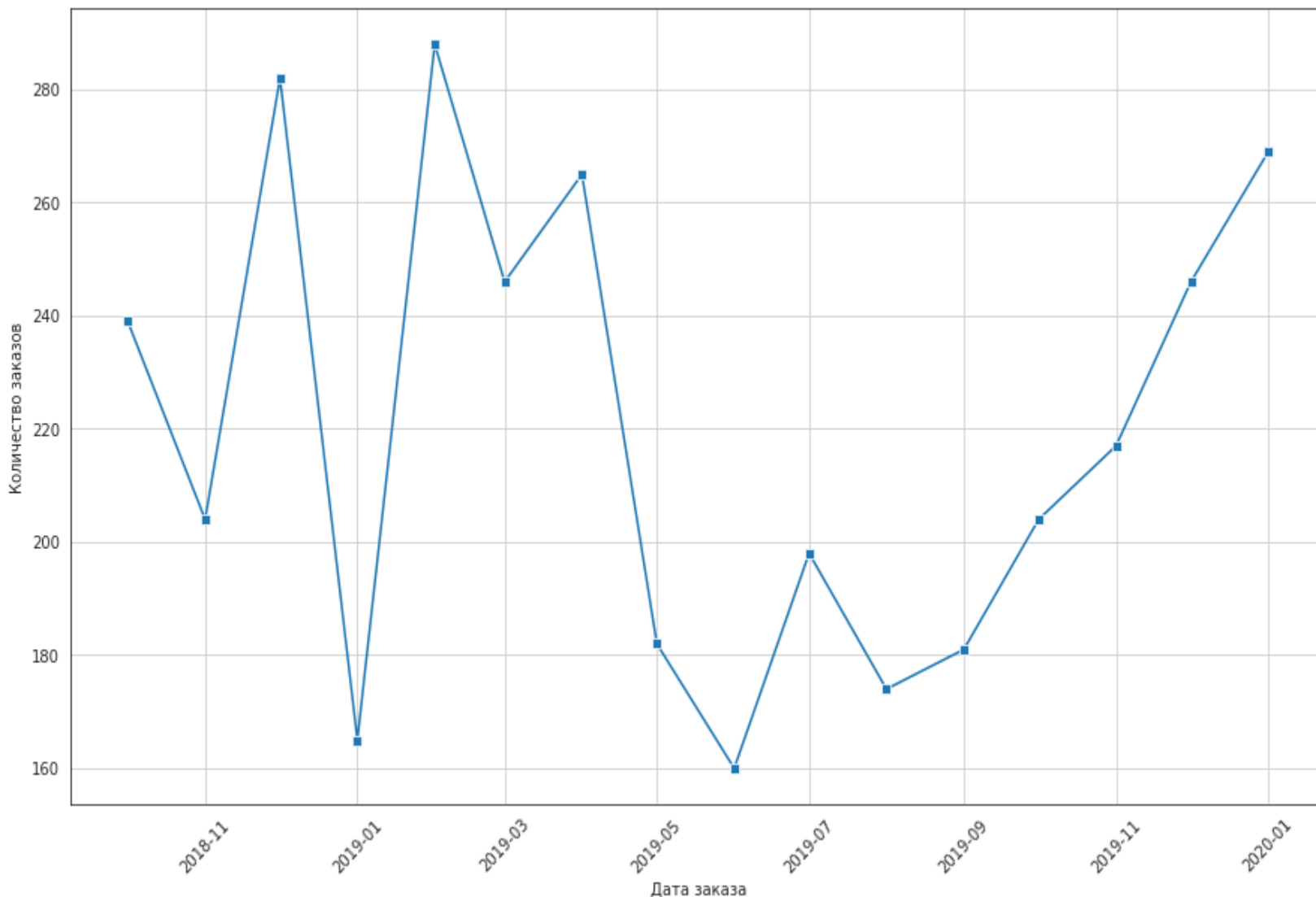


1. Обзор данных;
2. Предобработка данных;
3. Исследовательский анализ данных (EDA);
4. Сегментация покупателей на основе истории покупок (RFM - анализ);
5. Проверка статистических гипотез:
  - Посчитаем статистическую значимость различий в среднем чеке между сегментами покупателей;
  - Посчитаем статистическую значимость различий в количестве заказов между сегментами покупателей;
6. Общий вывод и рекомендации по разработке персонализированных предложений

# ОСНОВНЫЕ ВЫВОДЫ

1. Топ-3 категории (по количеству заказов): "Растения, рассада, семена", "Разные товары для дома" и "Хранение";
2. Топ-3 категории (по выручке): "Хранение ", "Разные товары для дома" и "Текстиль" ;
3. Топ-3 категории (по среднему чеку): "Хранение ", "Инструменты/ремонт" и "Текстиль" ;
4. На основании RFM- анализа выделили 18 сегментов покупателей по истории покупок;
5. Для детального анализа укрупнили ранее полученные в ходе RFM- анализа 18 сегментов в 4 группы:
  - сомневающиеся (506 покупателей);
  - лояльные (768 покупателей);
  - в спячке (280 покупателей);
  - в зоне риска (875 покупателей).
6. Проведенные статистические гипотезы показали различия между парами сегментов покупателей - разбивка на сегменты покупателей на основе истории покупок была проведена корректно.

# Распределение количества заказов по месяцам



Покупатели интернет-магазина предпочитают сезонные закупки впрок:



Наблюдается всплеск количества заказов **в декабре 2018, феврале 2019** года;

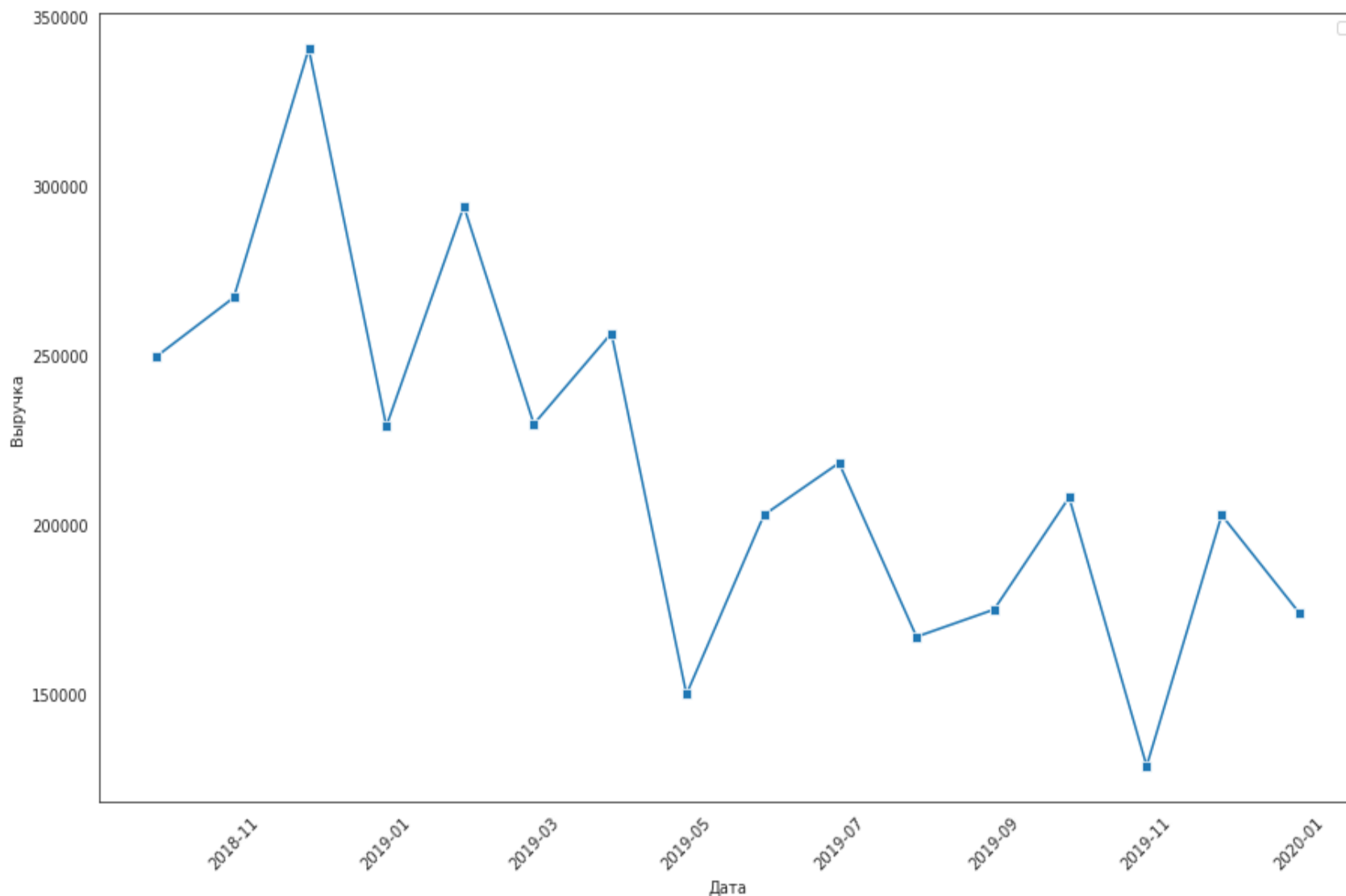


с **апреля 2019** года по **сентябрь 2019** года количество заказов снижается с небольшим увеличением заказов в июле 2019 года;



с **октября** наблюдается постепенный рост вплоть **до января 2020** года.)

# Изменение выручки во времени



В течение всего периода наблюдается зигзагообразное (небольшие всплески через каждые 2-3 месяца) падение выручки с продаж к текущему периоду;

Динамика роста выручки с продаж отрицательная:  
Темп роста продаж =  $(\text{выручка янв 2020г} / \text{выручка янв 2019г}) * 100 = (174036.33 / 229130.00) * 100 = 75,96\%$ .

# Распределение товаров по категориям



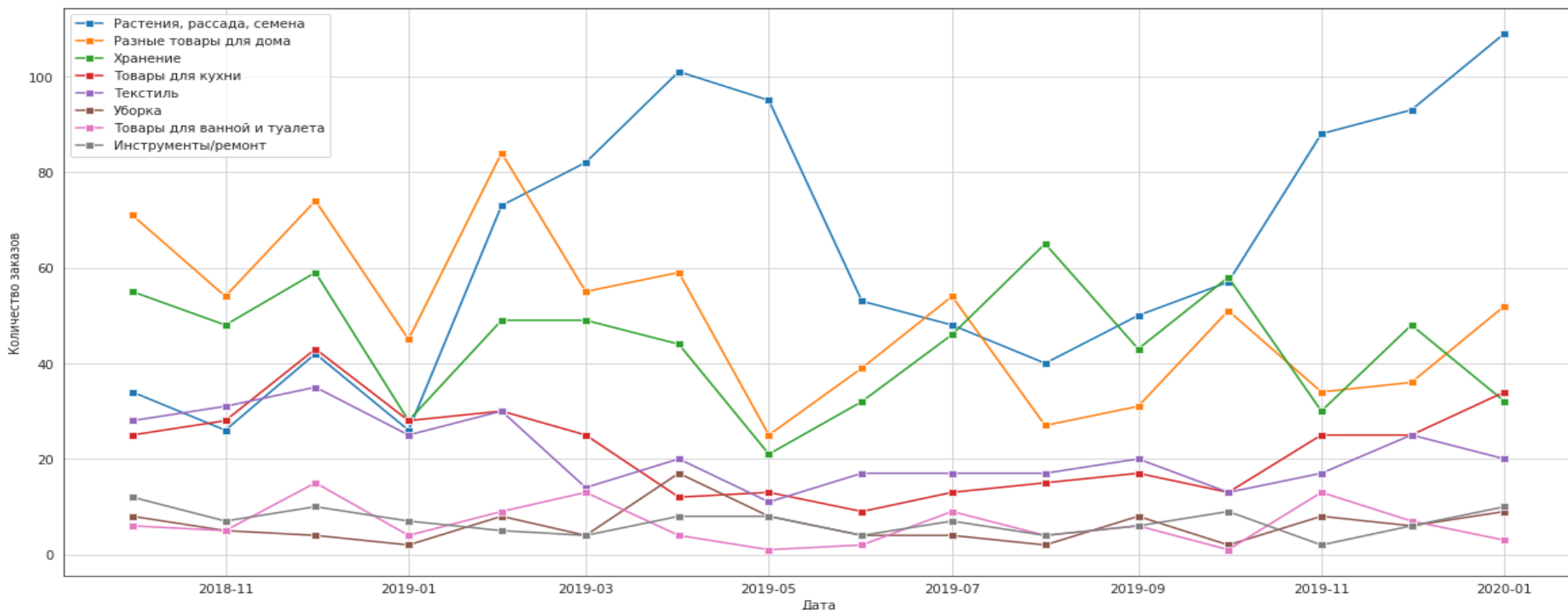
## Топ-3 категории

(по количеству проданного товара):

- Растения, рассада, семена;
- Разные товары для дома;
- Хранение.



# Распределение количества заказов по месяцам по категориям



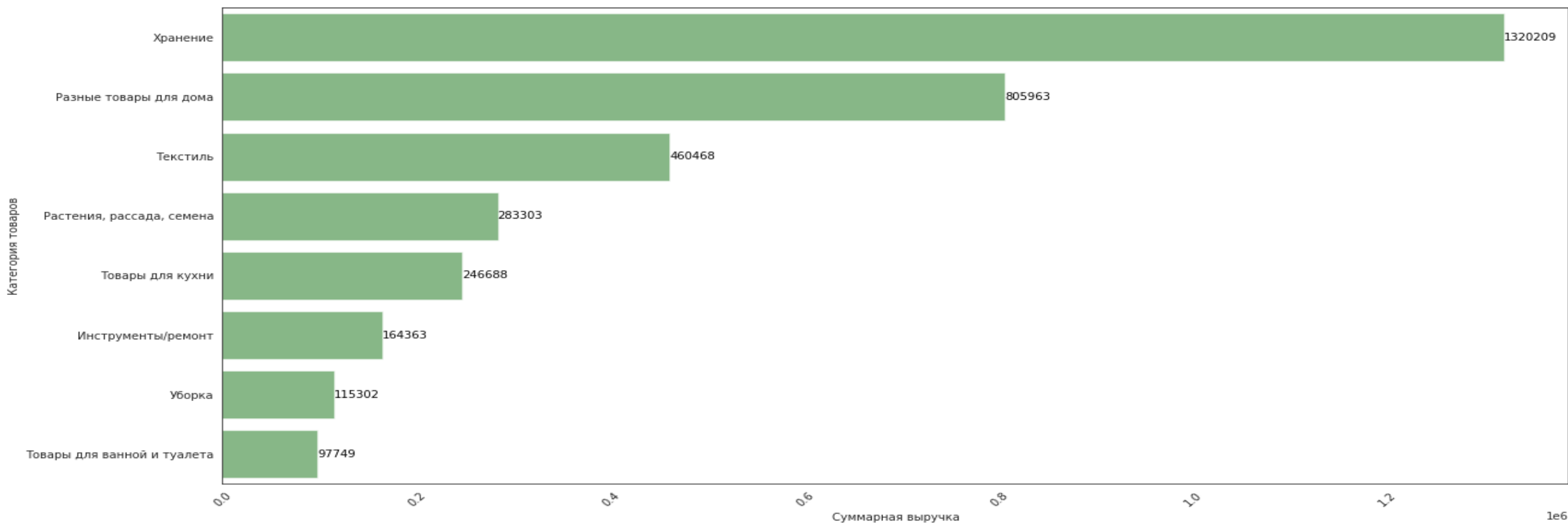
## Топ-3 категории (по количеству заказов):

- Растения, рассада, семена;
- Разные товары для дома;
- Хранение.

- По категории товаров "Разные товары для дома", "Хранение" график скачкообразный - всплески наблюдаются через каждые 2-3 месяца;
- По категории товаров "Растения, рассада, семена" наблюдается пик количества заказов в апреле 2019 г и январе 2020 г. Покупатели закупаются впрок на сезон;
- По остальным категориям, начиная с марта 2019 года, количество заказов приблизительно на одном уровне, нет влияния сезонности



# Выручка по категориям товаров

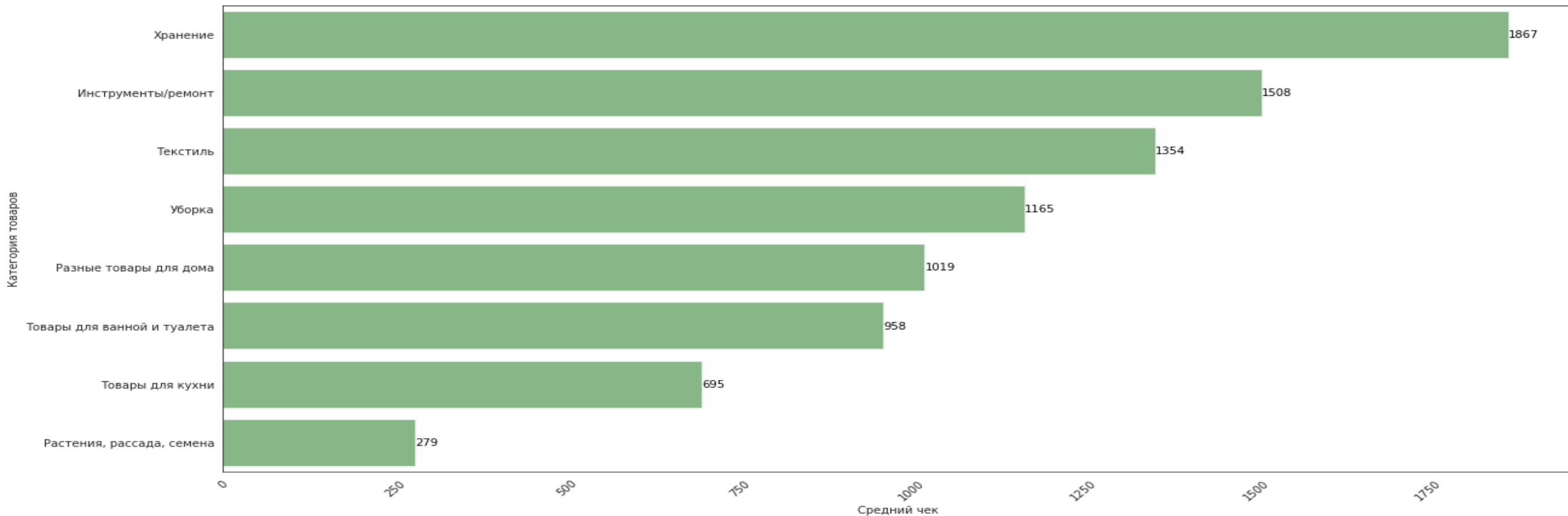


## Топ-3 категории (по выручке):

- "Хранение": 1 320 209, хотя по объему продаж она составляет всего 20%.
- "Разные товары для дома" (805 963, 22% по объему продаж)
- "Текстиль" (460 468, 10% по объему продаж).

Менее всего выручки приносят категории "Товары для ванной и туалета"(97 749) и "Уборка" (115 301), которые также являются наименее продаваемыми (не более 3% по объему продаж).

# Средний чек по категориям товаров



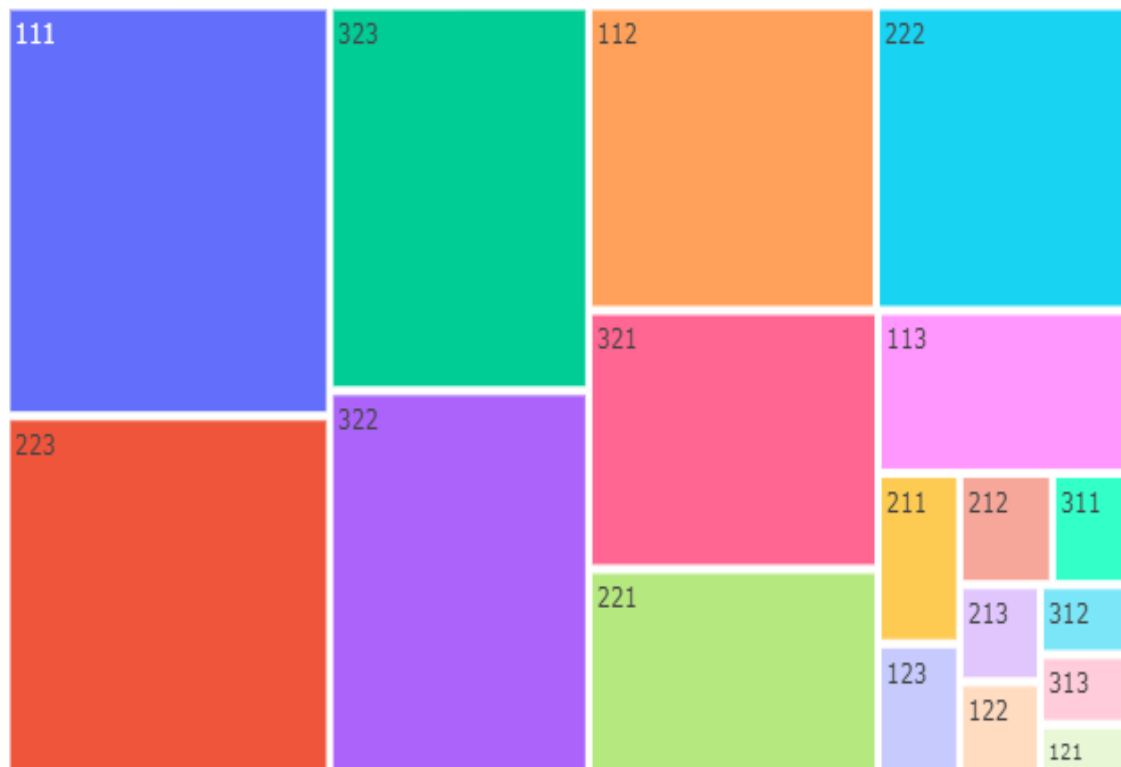
## Топ-3 категории (по среднему чеку):

- Хранение (1867);
- Инструменты/ремонт (1508);
- Текстиль (1354).

Самая продаваемая категория "Растения, рассада, семена" дает самый маленький средний чек (279).

# Распределение покупателей по сегментам (RFM – анализ)

Всего на основании RFM- анализа выделили **18 сегментов покупателей** по истории покупок



## Топ-5 сегментов покупателей :

- 111 - 372 покупателя (покупали часто, не так давно и на большие суммы);
- 223 - 325 покупателя (покупали не так часто, сравнительно давно и на небольшие суммы);
- 323 - 280 покупателей (покупали не так часто, давно и на небольшие суммы);
- 322 - 279 покупателей (покупали не так часто, давно и на средние суммы);
- 112 - 246 покупателей (покупали часто, не так давно и на средние суммы).

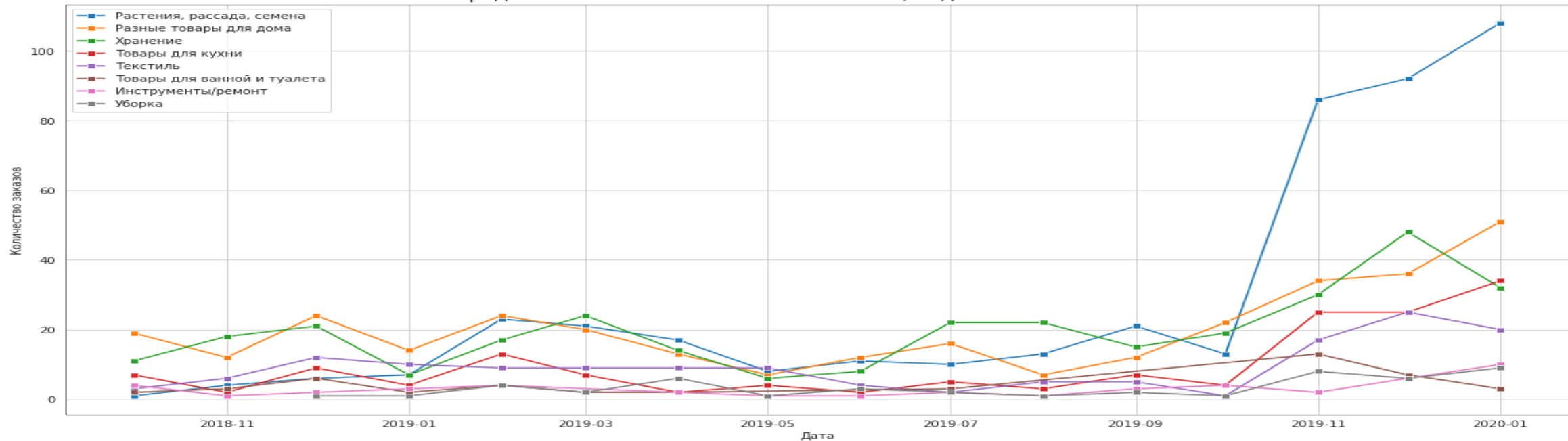
# Обобщение полученных сегментов в группы

|   | profile       | customer_count | R_avg | F_avg | M_avg |
|---|---------------|----------------|-------|-------|-------|
| 0 | лояльные      | 768            | 1.00  | 1.04  | 1.65  |
| 1 | сомневающиеся | 506            | 1.94  | 1.82  | 1.70  |
| 2 | в зоне риска  | 875            | 2.63  | 1.93  | 2.12  |
| 3 | в спячке      | 280            | 3.00  | 2.00  | 3.00  |

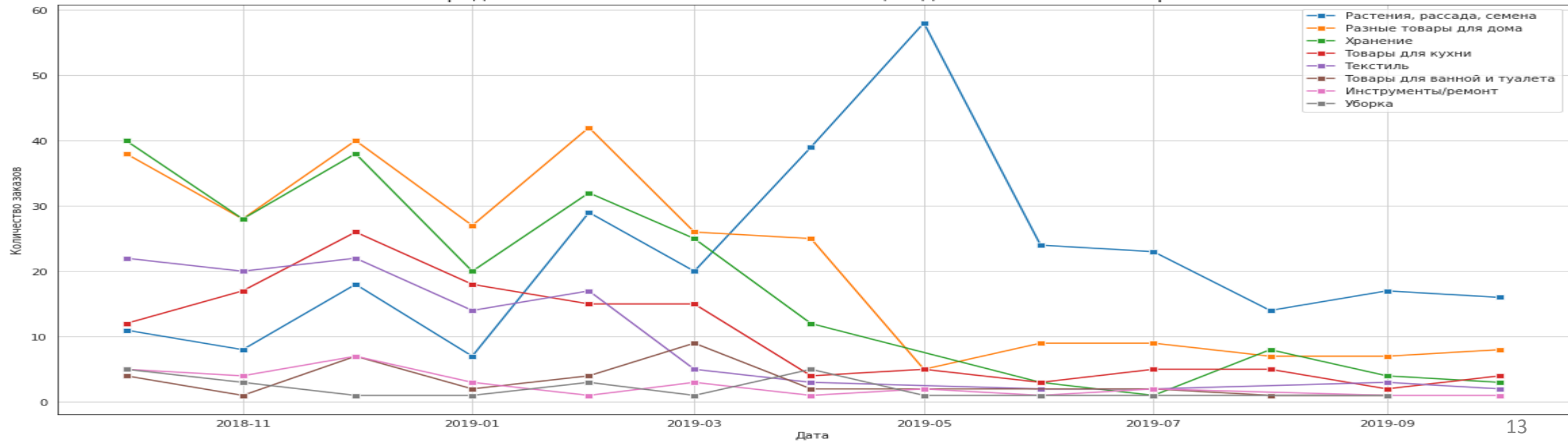
Ранее полученные на основании RFM- анализа 18 сегментов покупателей укрупнили в 4 группы:

1. сомневающиеся (506 покупателей);
2. лояльные (768 покупателей);
3. в спячке (280 покупателей);
4. в зоне риска (875 покупателей).

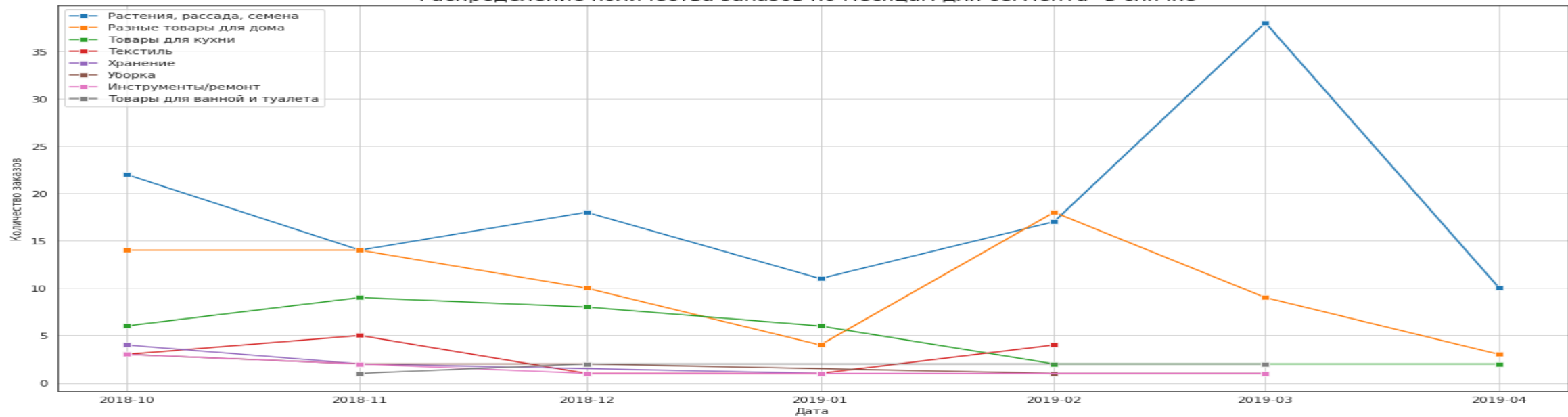
Распределение количества заказов по месяцам для сегмента "лояльные"



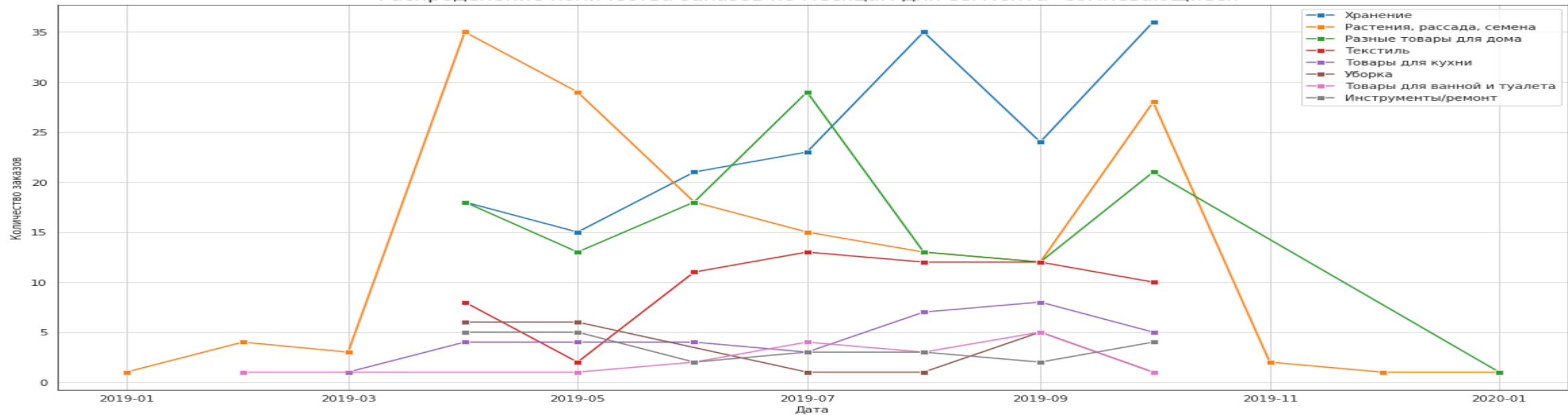
Распределение количества заказов по месяцам для сегмента "в зоне риска"



Распределение количества заказов по месяцам для сегмента "в спячке"



Распределение количества заказов по месяцам для сегмента "сомневающиеся"



# Распределение количества заказов по сегментам

1) По сегменту **"лояльные"** по всем категориям товаров наблюдается незначительное скачкообразное увеличение количества заказов с периодичностью в 2-3 месяца, а, начиная с октября 2019 года наблюдается резкое увеличения спроса по категориям, особенно по категории "Растения, рассада, семена";

2) По сегменту **"в зоне риска"** по всем категориям товаров наблюдается снижение закупок, начиная с апреля-мая 2019 года. При этом следует отметить, что до снижения закупок по категории "Растения, рассада, семена" наблюдается всплеск по закупке в мае 2019 года - спрос на данную категорию носит сезонный характер, покупатели покупаются данной категорией накануне дачного сезона.

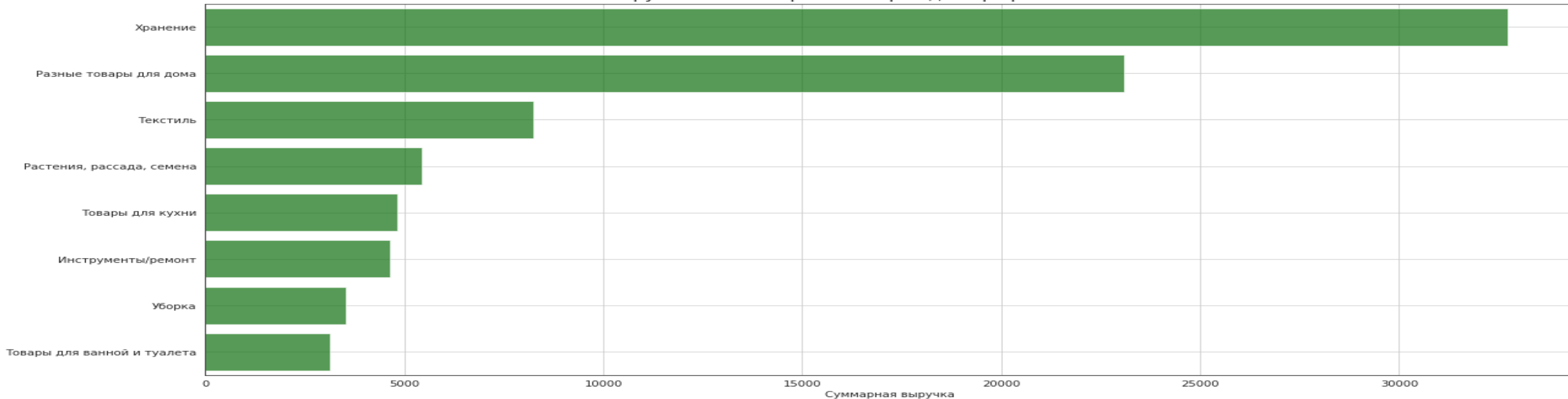
В целом, если посмотреть на данный сегмент покупателей можно отметить хороший спрос до падения закупок по категориям товаров "Растения, рассада, семена", "Хранение", "Разные товары для дома", "Товары для кухни" и "Текстиль";

3) К сегменту **"в спячке"** относятся покупатели, которые совершали покупки давно и на маленькие суммы. По графику видно, что данный сегмент перестал закупаться, начиная с апреля 2019 года, т.е. более 9 месяцев назад. Ранее наблюдались хорошие закупки по категориям "Растения, рассада, семена" и "Разные товары для дома";

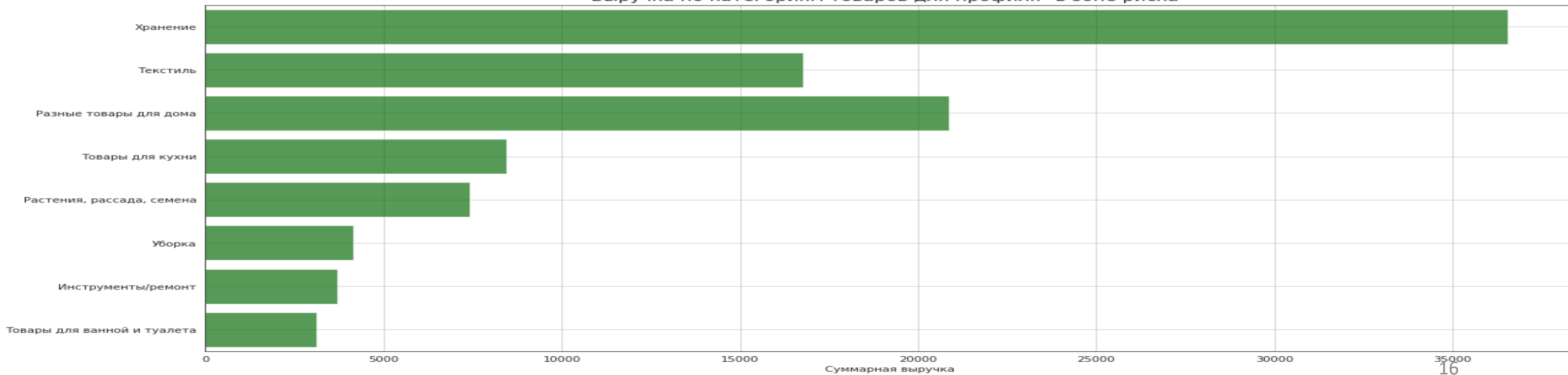
4) **"Сомневающиеся"** являются наиболее перспективным для перехода в "лояльные" сегментом. По нему по всем категориям, кроме категории "Растения, рассада, семена" и "Разные товары для дома" закупка по остальным категориям была последний раз в октябре 2019 года. По "Растения, рассада, семена" и "Разные товары для дома" совершались незначительные покупки в ноябре 2019 года-январе 2020 года.

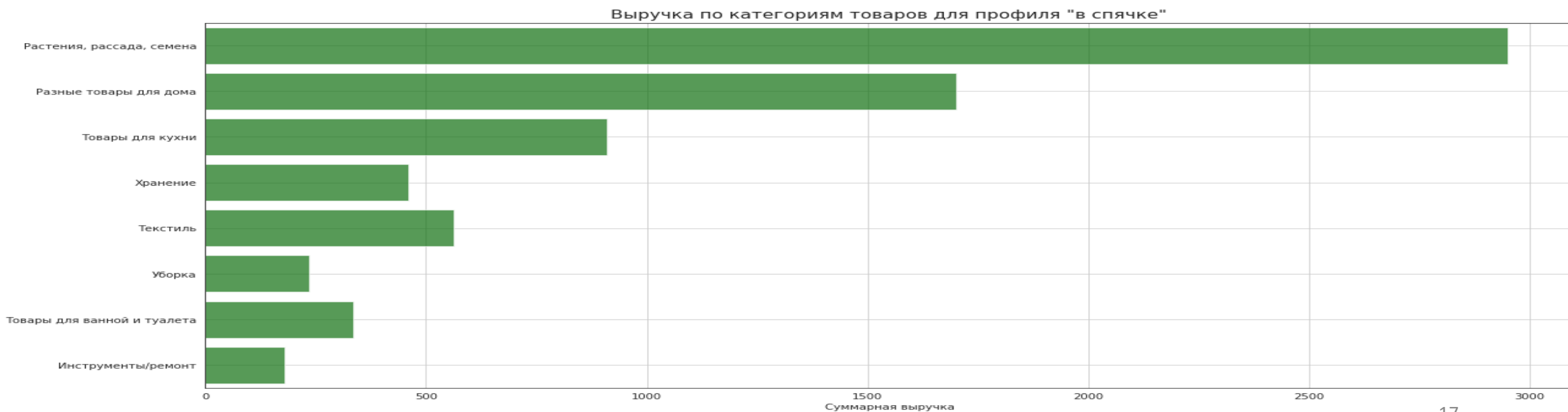
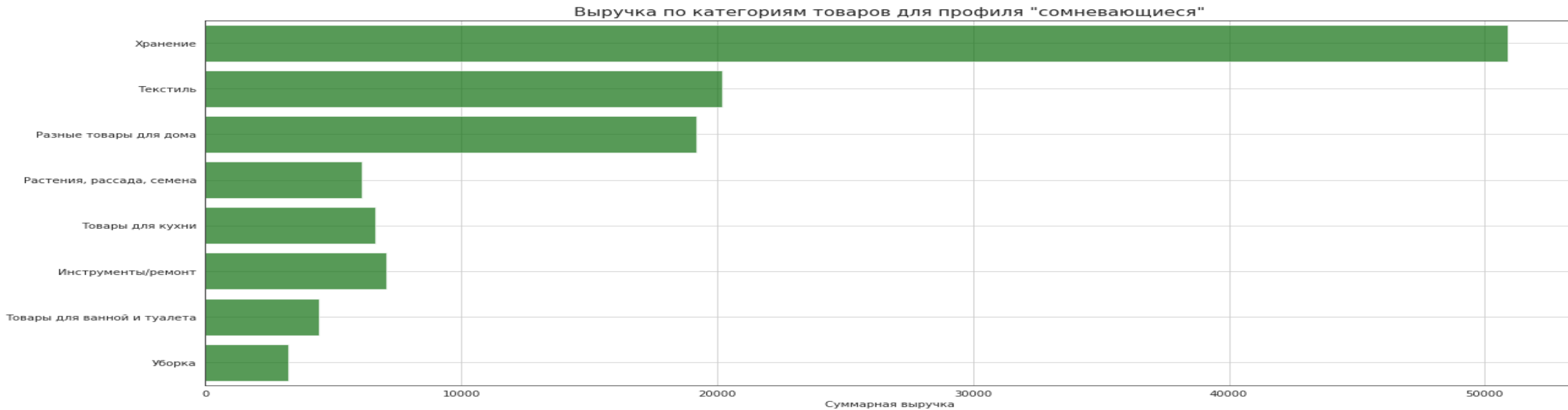


Выручка по категориям товаров для профиля "лояльные"



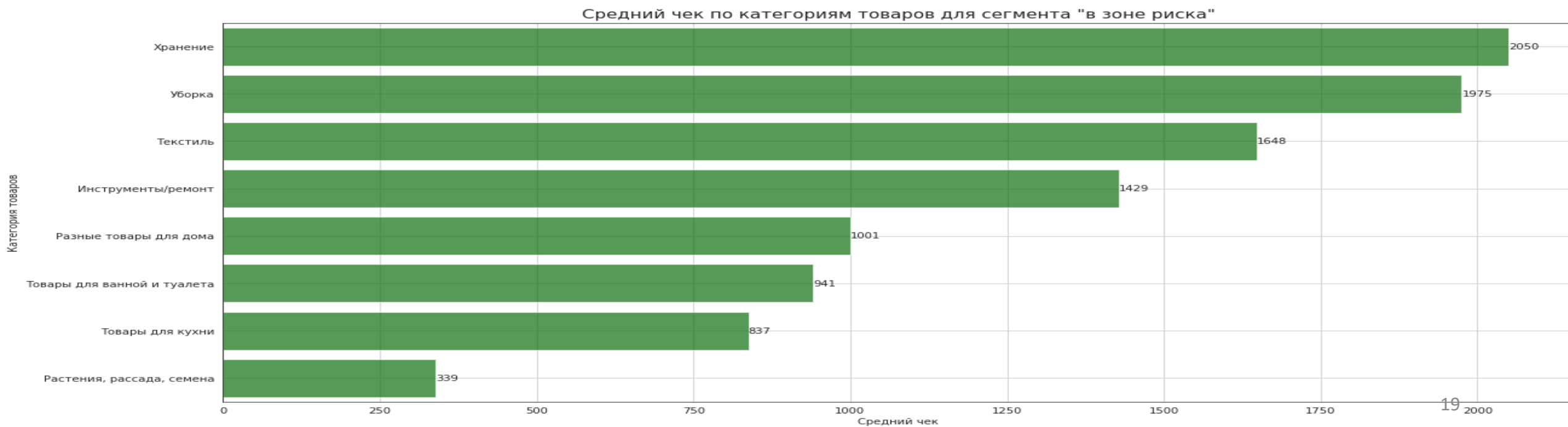
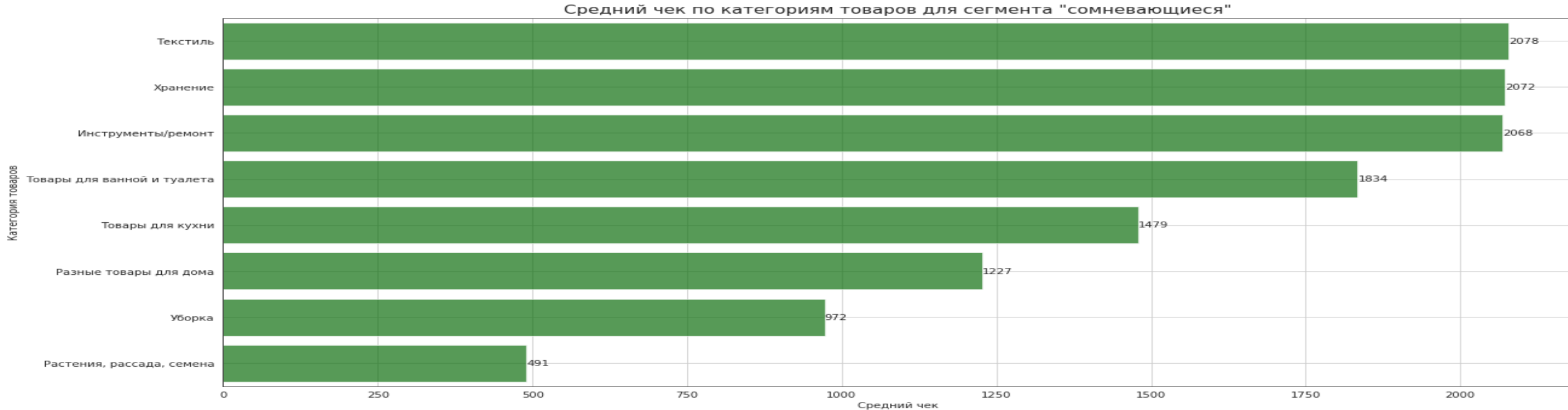
Выручка по категориям товаров для профиля "в зоне риска"



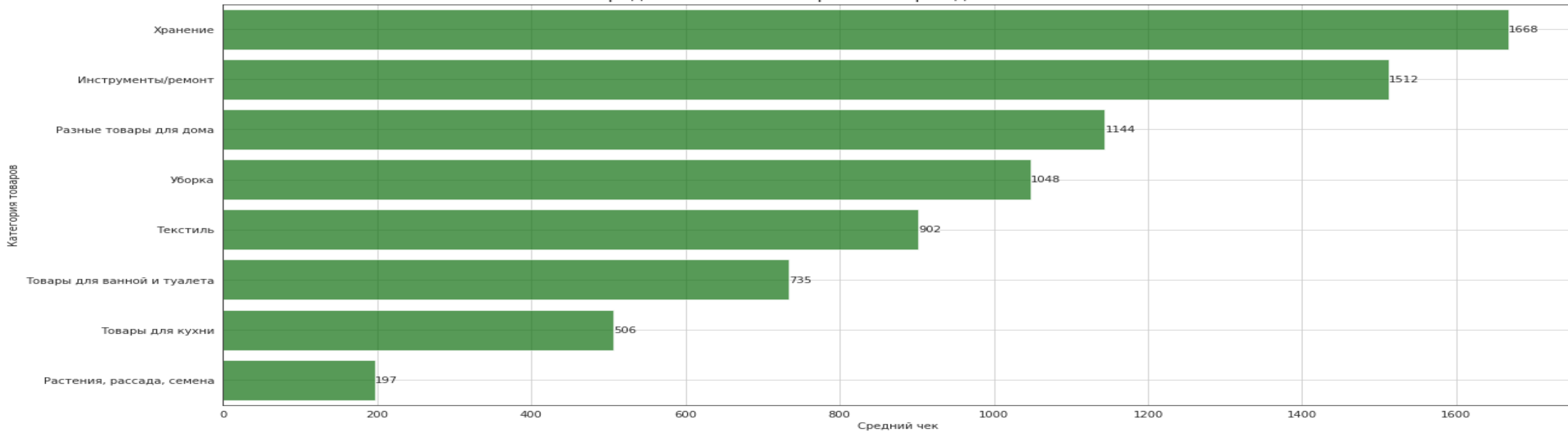


# Выручка по категориям товаров по сегментам

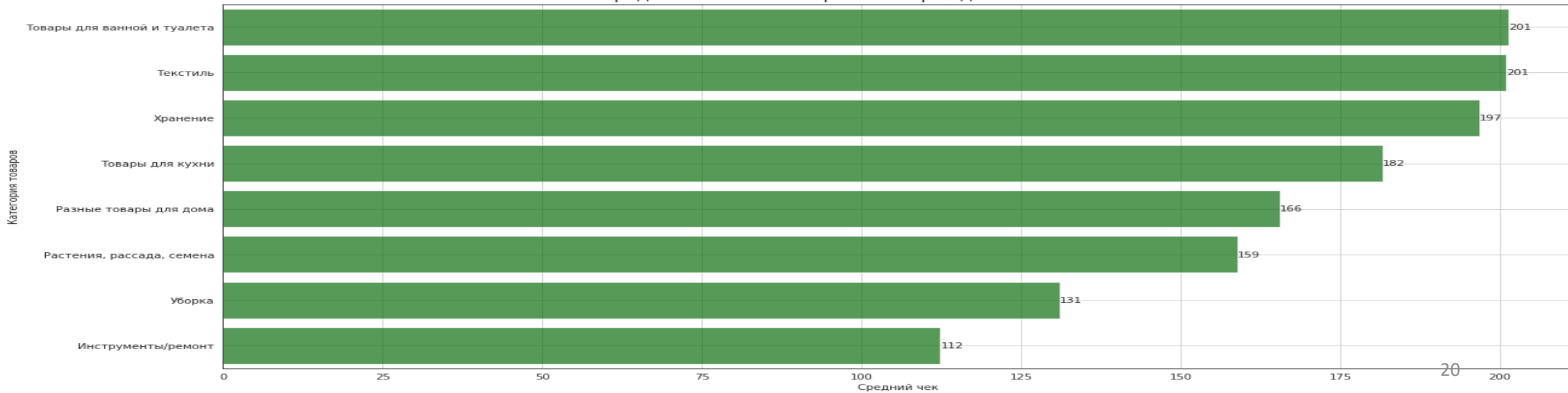
- 1) По сегменту **"лояльные"** самая большая суммарная выручка по категории "Хранение", далее следуют (в порядке убывания) "Разные товары для дома" и "Текстиль". Наименее всего выручки принесла категория "Товары для ванной и туалета";
- 2) Такие же топ-3 категории по выручке, как и в сегменте "лояльные", наблюдаются по сегменту **"в зоне риска"** - "Хранение", "Разные товары для дома" и "Текстиль". И наименее популярная категория "Товары для ванной и туалета" также совпадает;
- 3) По сегменту **"сомневающиеся"** самая большая суммарная выручка по категории "Хранение", далее следуют (в порядке убывания) "Текстиль" и "Разные товары для дома". Наименее всего выручки принесла категория "Уборка";
- 4) По сегменту **"в спячке"** самая большая суммарная выручка по категории "Растения, рассада, семена", далее следуют (в порядке убывания) "Разные товары для дома" и "Товары для кухни". Наименее всего выручки принесла категория "Инструменты/ремонт".



Средний чек по категориям товаров для сегмента "лояльные"



Средний чек по категориям товаров для сегмента "в спячке"



# Средний чек по категориям товаров по сегментам

- 1) По сегменту **"сомневающиеся"** практически одинаковые самые высокие средние чеки по категориям "Текстиль", "Хранение" и "Инструменты/ремонт". Наименьший средний чек у категории "Растения, рассада, семена";
- 2) По сегменту **"в зоне риска"** самый высокий средний чек по категории "Хранение", далее следуют (в порядке убывания) "Уборка" и "Текстиль". Наименьший средний чек у категории "Растения, рассада, семена";
- 3) По сегменту **"лояльные"** самый высокий средний чек по категории "Хранение", далее следуют (в порядке убывания) "Инструменты/ремонт" и "Разные товары для дома". Наименьший средний чек у категории "Растения, рассада, семена";
- 4) По сегменту **"в спячке"** самый высокий средний чек по категории "Товары для ванной и туалета", далее следуют (в порядке убывания) "Текстиль" и "Хранение". Наименьший средний чек у категории "Инструменты/ремонт".

# Проверка статистических гипотез (тестом Уилкоксона-Манна-Уитни)

## 1-я гипотеза:

**Нулевая гипотеза:** Нет различий в среднем чеке между сегментами покупателей

**Альтернативная гипотеза:** Есть различия в среднем чеке между сегментами покупателей

Проверка показала различия в среднем чеке между парами сегментов покупателей:

- 'сомневающиеся' и 'в зоне риска';
- 'сомневающиеся' и 'лояльные';
- 'сомневающиеся' и 'в спячке';
- 'в зоне риска' и 'в спячке';
- 'лояльные' и 'в спячке'.

Между парой 'в зоне риска' и 'лояльные' не получилось отвергнуть нулевую гипотезу, вывод о различии сделать нельзя.

## 2-я гипотеза:

**Нулевая гипотеза:** Нет различий в количестве заказов между сегментами покупателей

**Альтернативная гипотеза:** Есть различия в количестве заказов между сегментами покупателей

Проверка показала различия в количестве заказов между всеми парами сегментов покупателей.



# Рекомендации по разработке персонализированных предложений:

## 1) По сегменту "лояльные" :

*Предпочтительные категории товаров:* "Хранение", "Инструменты/ремонт", "Разные товары для дома" и "Текстиль".  
Учитывая скачкообразный спрос каждые 2-3 месяца проводить рассылку с предложением промокодов, скидок на предпочтительные категории товаров + предлагать скидку при заказе сопутствующего товара (как вариант, по категории "Растения, рассада, семена", как одной из самой покупаемой категории по сегменту);

## 2) По сегменту "в зоне риска" :

*Предпочтительные категории товаров:* "Растения, рассада, семена", "Хранение", "Разные товары для дома", "Товары для кухни" и "Текстиль".  
Спрос на товары по сегменту достаточно ровный, низкий, поэтому рассылки предложений можно проводить, например, ежемесячно или каждые 2 недели;

## 3) По сегменту "в спячке" :

*Предпочтительные категории товаров:* "Растения, рассада, семена" и "Разные товары для дома".  
Можно попробовать реанимировать данный сегмент путем рассылок промокодов, но, скорее всего, не стоит рассчитывать на большую конверсию в покупки, так как прошло уже достаточно много времени с момента последней покупки;

## 4) По сегменту\*"Сомневающиеся" :

*Предпочтительные категории товаров:* "Разные товары для дома" , "Текстиль", "Растения, рассада, семена" и "Хранение".  
Рассылки с промокодами, а также параллельно с информацией с полезным образовательным/информационным контентом проводить не реже 1 раза в месяц.

# Рекомендации по разработке персонализированных предложений (продолжение):



Кроме того, менеджменту магазина рекомендуется провести:

- 1) **Ребрендинг страницы интернет-магазина** для привлечения большего количества покупателей;
- 2) **Массовый запуск рекламы магазина** в различных источниках для целей увеличения пользовательских метрик и , как следствие, увеличении прибыльности бизнеса;
- 3) **Модернизацию** (при наличии) или **разработать новую программу лояльности**, побуждающую покупателей осуществлять покупки большее количество раз;
- 4) Совместно с маркетологами **внедрить такие акции, как "Купи в текущем периоде и в течение 30 дней получи скидку на следующую покупку"** и т.д.

Спасибо за внимание!