**МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ**

**УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»**

**(РУТ (МИИТ)**

Институт/факультет «Академия “Высшая Инженерная Школа”, АВИШ»

Специальность/Направление подготовки Информатика и вычислительная техника

**КУРСОВАЯ РАБОТА**

**по дисциплине:** Технологии хранения больших данных

|  |  |
| --- | --- |
| **на тему:** | "Аналитика продаж 3 филиалов сети с помощью методов Big Data" |

|  |  |
| --- | --- |
| **Студент группы ШАД 212** | **(Петров А.Г )** |
| **Научный руководитель** | **(Зуев И. А. )** |

**Москва 2023 г.**

# **Оглавление**

[**Оглавление** 2](#_Toc153474520)

[**Введение** 3](#_Toc153474521)

[**Исследование графиков построенных в ходе работы** 4](#_Toc153474522)

[**Заключение** 10](#_Toc153474523)

[**Список использованных источников и программ** 12](#_Toc153474524)

# **Введение**

Тема анализ продаж в разных филиалах всегда будет актуальной, так как это помогает определить успешность каждого филиала и выявить причины различий в продажах. Целью данной курсовой работы является проведение аналитического исследования данных об 3-ех различных филиалов и сравнении их.

В работе будет осуществлён сравнительный анализ и определение причины различий в продажах и выявить факторы, являющиеся наиболее важными для успешной работы каждого филиала.

Результаты исследования позволят выяснить, какие филиалы сети работают более эффективно, а какие требуют дальнейшей оптимизации и улучшения. Аналитика продаж также поможет определить, какие маркетинговые стратегии и действия приводят к увеличению продаж. Это позволит сети сосредоточить свои усилия на наиболее эффективных мероприятиях и услугах, что приведет к увеличению выручки и повышению общей прибыльности сети.

Полученные выводы будут полезны как для руководства сети, так и для каждого филиала отдельно. Руководство сети сможет определить, какие стратегии и меры способствуют повышению продаж в целом, а также применить успешные практики одного филиала к другим. Филиалы смогут сосредоточить свои усилия на улучшении эффективности продаж и изменении своих стратегий в соответствии с полученными результатами и рекомендациями. В итоге, все заинтересованные стороны смогут повысить успешность бизнеса и улучшить свои финансовые показатели.

# **Исследование графиков построенных в ходе работы**

Для проведения исследования был использован датасет, по итогу которого было произведено построение графиков для понятного отображения результатов анализа. Далее представлено их более детальное описание.

A pie chart with numbers and a few words

Description automatically generated with medium confidence

Рис. (1)

Данная круговая диаграмма предоставляет информацию об объемах продаж 3-ех филиалов. На графике наглядно видно, что хоть и незначительно, но в филиале A самые большие объемы продаж, тогда как в филиале C наименьшие объемы продаж.

A pie chart with text on it

Description automatically generated

Рис. (2)

Данная круговая диаграмма предоставляет информацию о действительных карт лояльности у покупателя во всех трех филиалах. Большинство людей, приходящих в филиалы, имеют карту лояльности.

A screenshot of a graph

Description automatically generated

Столбчатая диаграмма иллюстрирует самые продаваемые товары по категориям. По графику можно наглядно увидеть что категория “Electronic accessories” будет самой продаваемой, а категория “Health and beauty” самая отстающая.

A pie chart with text

Description automatically generated

Рис. (4)

На представленной круговой диаграмме можно сразу заметить клиенты предпочли чаще всего используют электронные кошельки, кредитками пользуются реже.

A blue rectangular object with black lines

Description automatically generated

Рис. (5)

Представленный график показывает распределение оценок товаров от покупателей. Наглядно видно, что средняя оценка – 7. Обычно покупатели оценивают товары от 5, до 9. Остальные оценки встречаются крайне редко.

A comparison of a graph

Description automatically generated with medium confidence

Рис. (6)

Представленные графики показывают интересы покупателей в зависимости от пола. Видно что женщины очень интересуются чаще всего модными аксессуарами и сильно реже интересуются товарами про здоровье. Мужчины же как раз чаще всего выбирают здоровье, и меньше всего им интересны товары про спорт и путешествия. Также можно заметить что сам проявленный интерес к любимому товару у женщин сильно выше чем у мужчин.

A row of colorful squares

Description automatically generated with medium confidence

Представленный график показывает нам наглядно информацию о продажах единиц товара. Заметим что единиц категории Fashion accessories было продано больше всех, а товары категории Health and beauty были проданы меньше всего раз, сразу заметно что эти наблюдения совпадают с графиком интереса у женщин, отсюда можно предположить что женщины покупают чаще мужчин.

A red circle with white text

Description automatically generated

Данный график показывает кто купил больше: мужчины или женщины? И да, сразу видно, что женщины действительно покупают чаще мужчин причем с довольно хорошим отрывом.

A graph of blue and orange bars

Description automatically generated

Рис. (8)

Данный график показывает изменения линейки товаров в зависимости от Гендера покупателя. По графику можно сделать вывод, что когда увеличивается количество женщин, увеличивается линейка модных аксессуаров, а когда увеличивается количество мужчин, увеличивается линейка товаров для здоровья и красоты

A diagram of different colored squares

Description automatically generated

Данный график показывает нам продажи по месяцам в трех разных филиалах, видно что больше всего продаж было в январе.

# **Заключение**

* Средняя оценка клиентов составила 7.
* Хотя большая часть продаж пришлась на филиал А, наибольший валовой доход приходится на филиал С.
* Больше всего продаж было в линейках "Мода и аксессуары" и "Продукты питания и напитки".
* Наибольший валовой доход\* приходился на продуктовую линейку "Продукты питания и напитки", затем "Спорт и путешествия".
* Женщины больше, чем мужчины, интересуются "Модой и аксессуарами", "Едой и напитками" и "Спортом и путешествиями", в то время как мужчины больше всего предпочитают "Здоровье и красоту", "Аксессуары для электроники" и "Дом и стиль жизни".
* Большая часть продаж и валового дохода приходится на январь.
* Было установлено, что предпочтительным способом оплаты для клиентов являются электронный кошелек и наличные.

И подводя итог в результате сравнения 3 филиалов были выполнены все поставленые задачи и выявлено несколько тенденций и выводов, которые будут полезны как для руководства сети, так и для каждого филиала отдельно, руководство сети сможет определить, какие стратегии и меры способствуют повышению продаж в целом, а также применить успешные практики одного филиала к другим.

Филиалы смогут сосредоточить свои усилия на улучшении эффективности продаж и изменении своих стратегий в соответствии с полученными результатами и рекомендациями. В итоге, все заинтересованные стороны смогут повысить успешность бизнеса и улучшить свои финансовые показатели.

# **Список использованных источников и программ**

1. Использовался датасет с сайта Kaggle:

<https://www.kaggle.com/datasets/> aungpyaeap/supermarket-sales/

1. Библиотеки, использованные в изучении датасета: numpy, pandas, plotly, matplotlib, seaborn