ТИТУЛЬНЫЙ ЛИСТ

Содержание

[1. Постановка задачи 3](#_Toc184560590)

[2. Структура проекта 4](#_Toc184560591)

[2.1 Построение графиков средствами Matplotlib 4](#_Toc184560592)

[2.2 Построение графиков средствами Seaborn 4](#_Toc184560593)

[2.3 Построение графиков средствами Plotly 5](#_Toc184560594)

[3. Сравнение используемых библиотек 6](#_Toc184560595)

[3.1 Библиотека Matplotlib 7](#_Toc184560596)

[3.2 Библиотека Seaborn 8](#_Toc184560597)

[3.3 Библиотека Plotly 9](#_Toc184560598)

[Заключение 10](#_Toc184560599)

[Приложение 1. Структура файла 11](#_Toc184560600)

[Приложение 2. Список команд для установки библиотек 15](#_Toc184560601)

# Постановка задачи

Сравнение различных библиотек для визуализации данных: Matplotlib, Seaborn и Plotly:

Создать набор визуализаций с использованием Matplotlib, Seaborn и Plotly, сравнить их функциональность и удобство использования.

Для выполнения поставленной задачи необходимо разработать проект средствами языка программирования python использую библиотеки: Matplotlib, Seaborn и Plotly для построения графиков по данным из файла: «visualization\_comparison\_data.csv». В исходном файле 100 записей.

Структура файла представлена на рисунке 1.

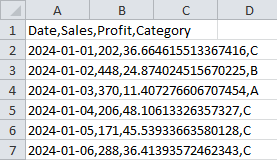


Рисунок 1 – Структура файла данных

# Структура проекта

Проект состоит из нескольких ключевых компонентов, реализованных средствами различных библиотек.

## 2.1 Построение графиков средствами Matplotlib

На рисунке 2 представлен пример отображения данных из файла средствами библиотеки Matplotlib в виде пяти графиков:

* Линейный график продаж;
* Гистограмма прибыли;
* Точечный график;
* Сравнение продаж по категориям;
* Круговая диаграмма категорий.



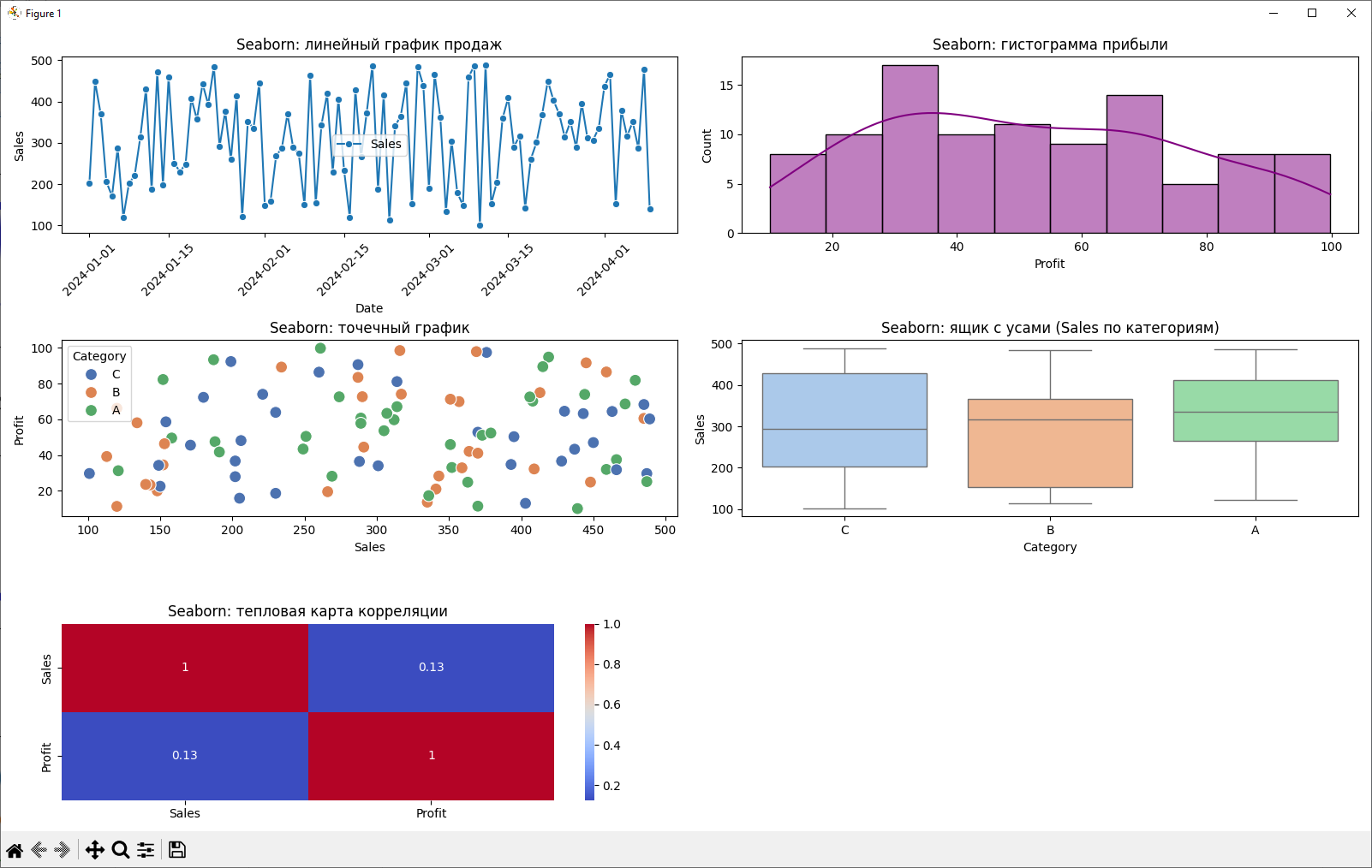
Рисунок 2 – Реализация графиков средствами библиотеки «Matplotlib»

## 2.2 Построение графиков средствами Seaborn

На рисунке 3 представлен пример отображения данных из файла средствами библиотеки Seaborn в виде трех графиков:

* Линейный график продаж;
* Гистограмма прибыли;
* Точечный график;
* Ящик с усиками по категориям;
* Тепловая карта корреляции.

Реализация примерно похожа, с небольшими отличиями.

Рисунок 3 – Реализация графиков средствами библиотеки «Seaborn»

## 2.3 Построение графиков средствами Plotly

На рисунке 4 представлен пример отображения данных из файла средствами библиотеки Plotly в виде точечного графика. Реализация выполнена в веб-странице.

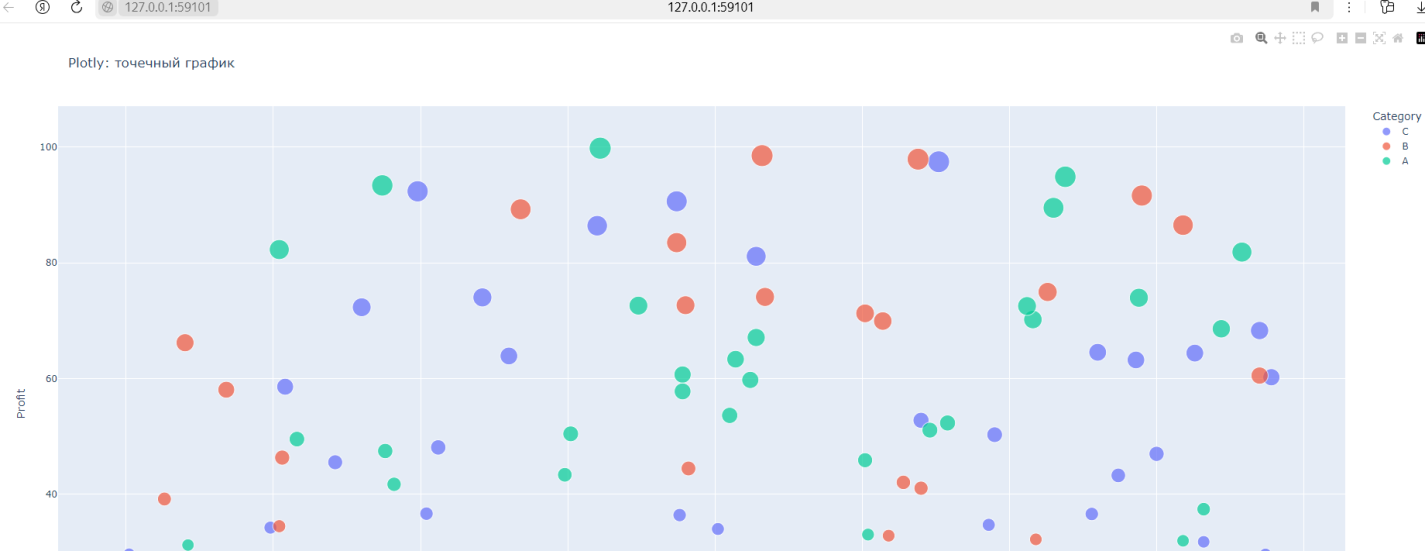


Рисунок 4 – Реализация точечного графика средствами библиотеки «Plotly»

На рисунке 5 представлен пример отображения данных из файла средствами библиотеки Plotly в виде графика гистограммы прибыли. Реализация выполнена в веб-странице.

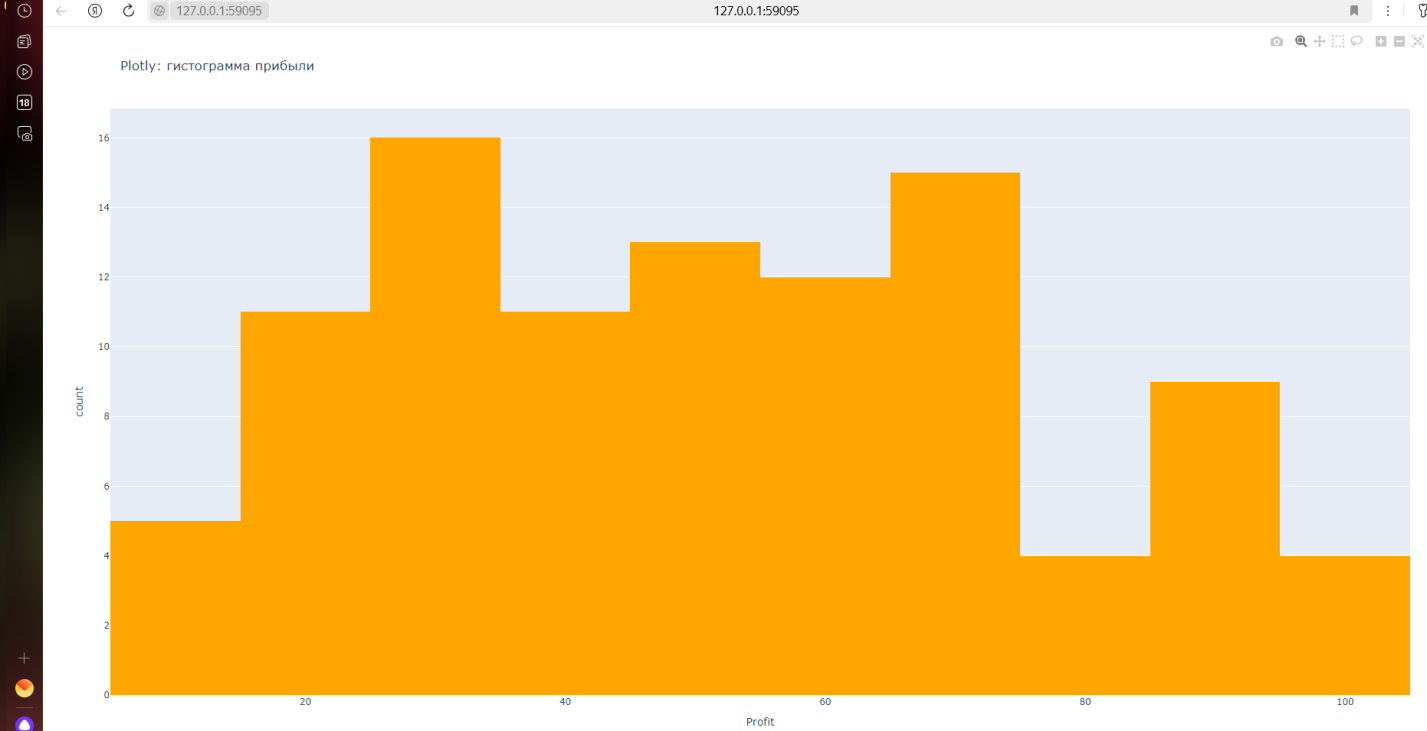


Рисунок 5 – Реализация гистограммы прибыли средствами библиотеки «Plotly»

На рисунке 6 представлен пример отображения данных из файла средствами библиотеки Plotly в виде линейного графика. Реализация выполнена в веб-странице.

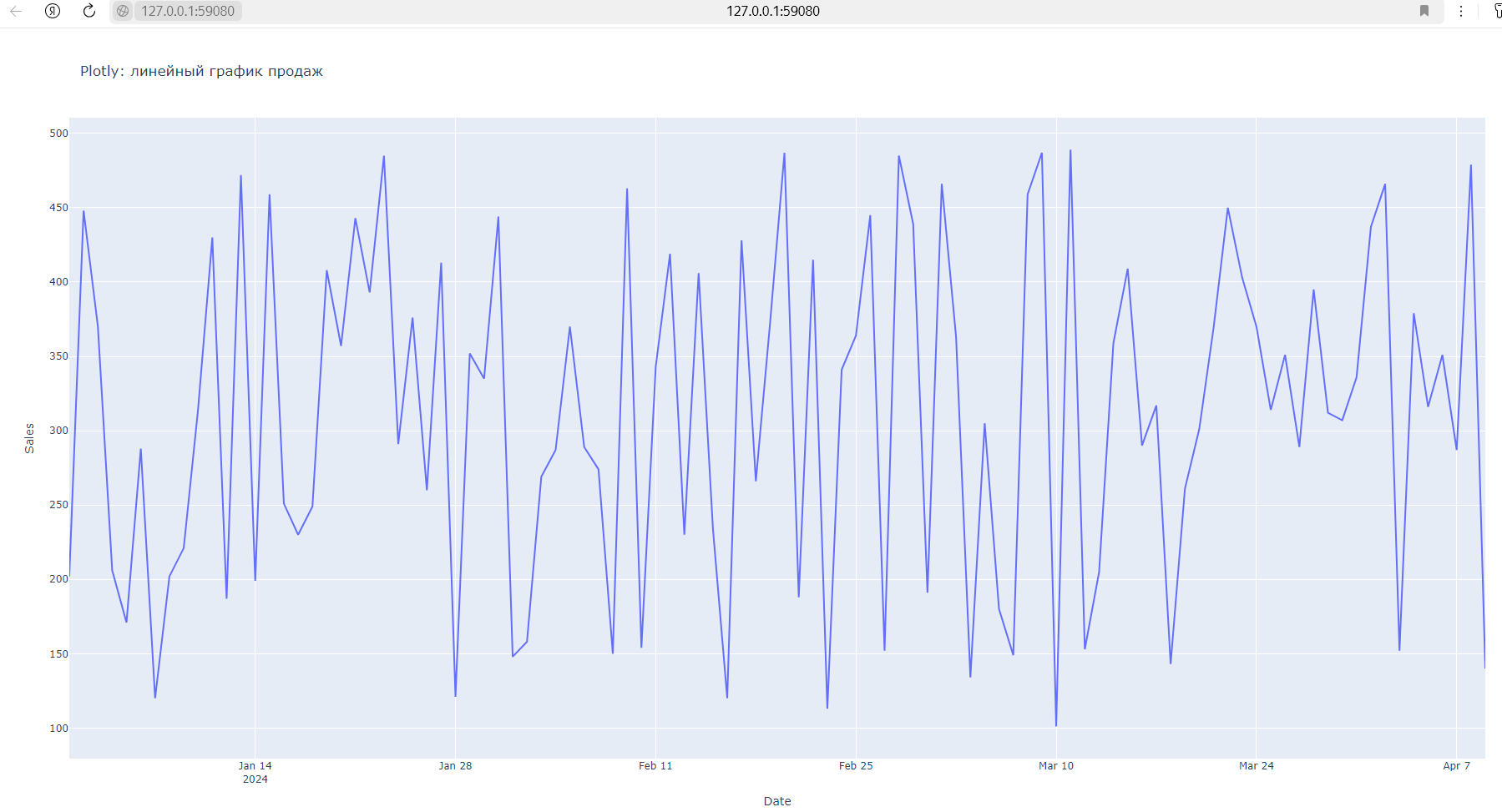


Рисунок 6 – Реализация линейного графика продаж средствами библиотеки «Plotly»

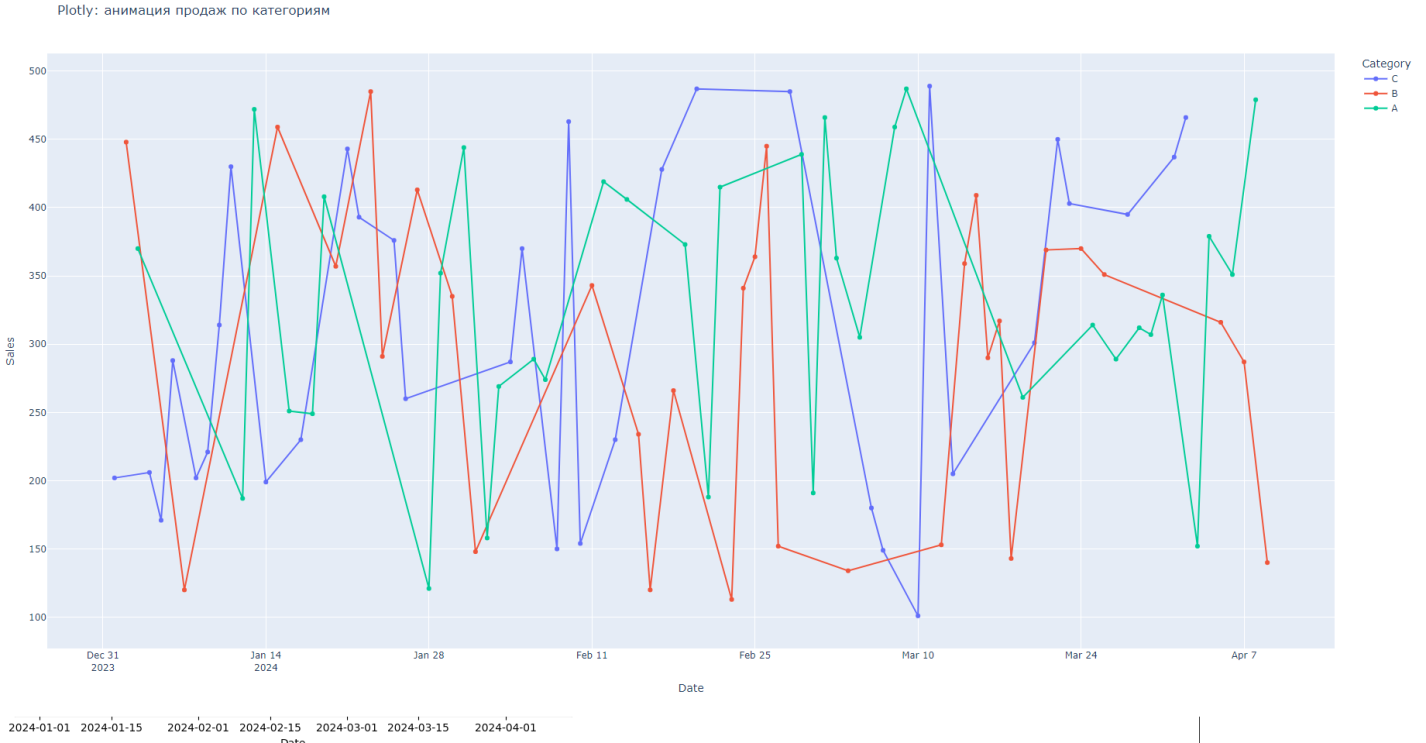


Рисунок 7 – Анимация продаж по категориям библиотеки «Plotly»

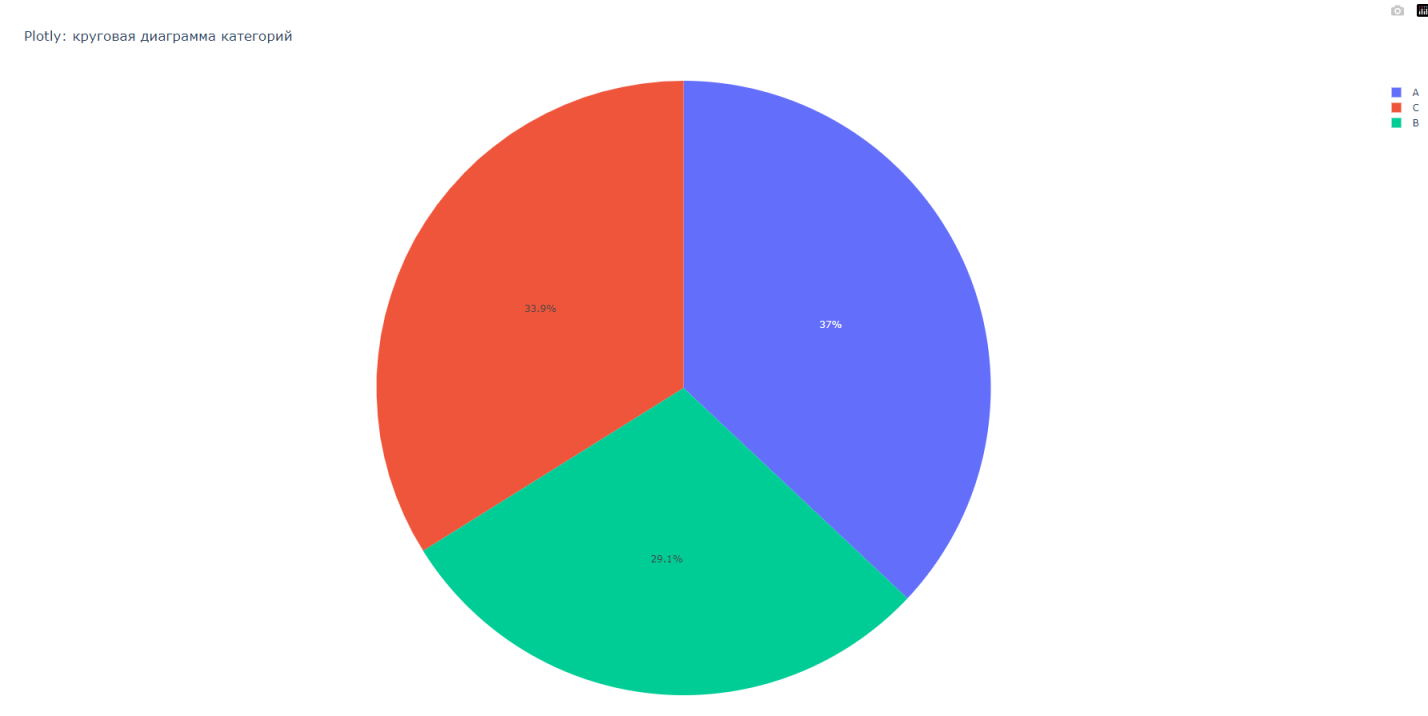


Рисунок 8 – Круговая диаграмма категорий библиотеки «Plotly»

# Сравнение используемых библиотек

Каждая из библиотека выбирается под различные задачи. Если необходимо выбрать одну библиотеку, стоит учитывать одну из задач:

* Аналитик – лучше подойдет Seaborn
* Интерактивные дашборды – лучше подойдет Plotly
* Исследования и сложные настройки – лучше подойдет Matplotlib.

| Критерий | Matplotlib | Seaborn | Plotly |
| --- | --- | --- | --- |
| **Удобство использования** | Низкое: требует много кода | Высокое: простота синтаксиса | Среднее: интерактивность, но сложнее настройки |
| **Эстетика** | Умеренная | Высокая | Высокая |
| **Интерактивность** | Отсутствует | Отсутствует | Присутствует |
| **Гибкость** | Высокая | Средняя | Низкая |
| **Типы графиков** | Большой выбор, но базовые | Специализация на аналитике | Широкий выбор, включая сложные |
| **Производительность** | Высокая | Умеренная | Средняя |
| **Поддержка больших данных** | Высокая | Низкая | Средняя |

В следующих разделах будут рассмотрены достоинства и недостатки используемых библиотек.

## 3.1 Библиотека Matplotlib

К достоинствам данной библиотеки относятся:

1. Гибкость и контроль:

* Полный контроль над графиками: цвета, шрифты, метки, оси и т.д.
* Подходит для создания сложных и кастомизированных графиков.

1. Популярность:

* Широкое сообщество, множество обучающих материалов и документации.
* Часто используется в комбинации с другими библиотеками, например, Seaborn.

1. Базовая библиотека:

* Основа многих других библиотек визуализации (например, Seaborn).

1. Работа с большими данными:

* Эффективен для работы с большими объемами данных.

К недостаткам данной библиотеки относятся:

1. Сложность для начинающих:

* Требует больше кода для создания простых графиков.
* Настройка графиков может быть громоздкой и трудоемкой.

1. Эстетика:

* По умолчанию графики выглядят менее привлекательно по сравнению с Seaborn и Plotly.

## 3.2 Библиотека Seaborn

К достоинствам данной библиотеки относятся:

1. Удобство использования:

* Простота построения сложных графиков за счет высокоуровневого API.
* Минимум кода для получения профессионально выглядящих графиков.

1. Эстетика:

* Эстетически приятные графики по умолчанию, включая цветовые палитры и стили.

1. Интеграция с Pandas:

* Легко работает с DataFrame, упрощая обработку и отображение данных.

1. Дополнительная функциональность:

* Поддержка регрессионного анализа, heatmap и pairplot для визуализации отношений между переменными.

К недостаткам данной библиотеки относятся:

1. Ограниченная гибкость:

* Менее гибкий по сравнению с Matplotlib, так как основан на нем.
* Для сложных графиков иногда приходится дописывать код на Matplotlib.

1. Производительность:

* При работе с большими объемами данных производительность может снижаться.

## 3.3 Библиотека Plotly

К достоинствам данной библиотеки относятся:

* 1. Интерактивность:
* Поддержка интерактивных графиков: масштабирование, навигация, всплывающие подсказки.
* Удобен для представления данных на веб-платформах и в дашбордах.
  1. Эстетика и функциональность:
* Графики выглядят профессионально без дополнительных настроек.
* Легкость в создании сложных графиков (географические карты, 3D-графики).
  1. Широкие возможности визуализации:
* Поддержка большого количества типов графиков, включая тепловые карты, географические карты, треллис-плоты и другие.
  1. Поддержка нескольких языков:
* Помимо Python, доступна интеграция с R, JavaScript и другими языками.

К недостаткам данной библиотеки относятся:

1. Сложность настройки:

* Графики сложнее настраивать на низком уровне по сравнению с Matplotlib.

1. Производительность:

* Может быть менее эффективным при обработке огромных объемов данных.

1. Зависимость от браузера:

* Для отображения графиков часто требуется использование браузера.

# Заключение

В ходе выполнения работы были исследованы возможности трех популярных библиотек для визуализации данных: **Matplotlib**, **Seaborn** и **Plotly**. Каждая из библиотек была применена для создания набора визуализаций, включающих линейные графики, гистограммы и точечные графики. Это позволило продемонстрировать их функциональные особенности, преимущества и области применения.

**Matplotlib** показал себя как мощный инструмент для создания базовых и кастомных графиков. Его гибкость позволяет управлять мельчайшими деталями графиков, но требует от пользователя большего количества кода для достижения сложных визуализаций. Эта библиотека отлично подходит для научных исследований и проектов, где требуется высокий уровень контроля над графическими элементами.

**Seaborn** предложил более удобный интерфейс для создания графиков, автоматизируя ряд рутинных задач. Это позволило создавать визуализации с минимальными усилиями и с эстетически привлекательным дизайном по умолчанию. Seaborn хорошо справляется с задачами анализа данных, где необходимо быстро получать наглядные и информативные графики, особенно при работе с pandas DataFrame.

**Plotly** продемонстрировал свои уникальные возможности интерактивной визуализации. Эта библиотека позволила пользователю исследовать данные непосредственно на графиках, что особенно важно для презентаций и веб-приложений. Однако для базовых визуализаций Plotly может быть избыточным, так как требует большего объема вычислительных ресурсов.

Сравнение библиотек показало, что выбор инструмента для визуализации данных зависит от контекста задачи. **Matplotlib** лучше всего подходит для пользователей, которым нужен детальный контроль, **Seaborn** упрощает создание визуализаций для аналитиков данных, а **Plotly** предоставляет превосходные возможности для интерактивной работы с графиками.

Работа продемонстрировала практическое применение каждой библиотеки, их достоинства и ограничения. Полученные результаты помогут выбирать оптимальный инструмент для визуализации данных в зависимости от конкретных требований и условий.

# Приложение 1. Структура файла

Date,Sales,Profit,Category

2024-01-01,202,36.664615513367416,C

2024-01-02,448,24.874024515670225,B

2024-01-03,370,11.407276606707454,A

2024-01-04,206,48.10613326357327,C

2024-01-05,171,45.53933663580128,C

2024-01-06,288,36.41393572462343,C

2024-01-07,120,11.2671840443576,B

2024-01-08,202,27.895816367992463,C

2024-01-09,221,74.0207757473785,C

2024-01-10,314,81.1157986478085,C

2024-01-11,430,64.53639773029101,C

2024-01-12,187,93.36707906620141,A

2024-01-13,472,68.59693229517501,A

2024-01-14,199,92.34637079894027,C

2024-01-15,459,86.50347200108193,B

2024-01-16,251,50.45056067243831,A

2024-01-17,230,18.586910484137018,C

2024-01-18,249,43.37364269784397,A

2024-01-19,408,70.19571273972466,A

2024-01-20,357,69.9330120955747,B

2024-01-21,443,63.21680089369544,C

2024-01-22,393,34.72496136910577,C

2024-01-23,485,60.5119083262931,B

2024-01-24,291,44.46341872784109,B

2024-01-25,376,97.45408858501933,C

2024-01-26,260,86.40224418394754,C

2024-01-27,413,74.95565690483859,B

2024-01-28,121,31.238642777406017,A

2024-01-29,352,33.046149048519155,A

2024-01-30,335,13.639023058458822,B

2024-01-31,444,73.95966007172086,A

2024-02-01,148,19.98017387306482,B

2024-02-02,158,49.54028516791931,A

2024-02-03,269,28.15472821018566,A

2024-02-04,287,90.61872361061674,C

2024-02-05,370,52.783320086390056,C

2024-02-06,289,60.69480147787453,A

2024-02-07,274,72.59644777835148,A

2024-02-08,150,22.539830896528812,C

2024-02-09,463,64.39756413500355,C

2024-02-10,154,58.58569821715058,C

2024-02-11,343,28.275510226129246,B

2024-02-12,419,94.8568213502183,A

2024-02-13,230,63.89789198396824,C

2024-02-14,406,72.5306439735734,A

2024-02-15,234,89.2421055113732,B

2024-02-16,120,66.19186433204139,B

2024-02-17,428,36.60703172539426,C

2024-02-18,266,19.494483384724354,B

2024-02-19,373,51.08811134346192,A

2024-02-20,487,29.659639349515025,C

2024-02-21,188,47.48589530833296,A

2024-02-22,415,89.49522330269815,A

2024-02-23,113,39.191051890474654,B

2024-02-24,341,20.9879159230606,B

2024-02-25,364,42.06680542692774,B

2024-02-26,445,91.61455973911787,B

2024-02-27,152,34.49190244461718,B

2024-02-28,485,68.29211084872261,C

2024-02-29,439,10.046833929578424,A

2024-03-01,191,41.73119707007521,A

2024-03-02,466,37.43031323422261,A

2024-03-03,363,24.819026782864753,A

2024-03-04,134,58.068047743789776,B

2024-03-05,305,53.63469742230849,A

2024-03-06,180,72.31924296012433,C

2024-03-07,149,34.24711004186693,C

2024-03-08,459,31.971297002299675,A

2024-03-09,487,25.14619379556375,A

2024-03-10,101,29.68877976157632,C

2024-03-11,489,60.22918018156071,C

2024-03-12,153,46.345255395223674,B

2024-03-13,205,15.840302239808342,C

2024-03-14,359,32.85238725409102,B

2024-03-15,409,32.21884565547411,B

2024-03-16,290,72.66738455558095,B

2024-03-17,317,74.10435309319978,B

2024-03-18,143,23.32782369580599,B

2024-03-19,261,99.79664365440476,A

2024-03-20,301,34.01029128477565,C

2024-03-21,369,97.89534602493876,B

2024-03-22,450,46.99333119864082,C

2024-03-23,403,12.974565961049354,C

2024-03-24,370,41.05641232240147,B

2024-03-25,314,67.09162102312274,A

2024-03-26,351,71.26349063992902,B

2024-03-27,289,57.78411249854228,A

2024-03-28,395,50.30048481157825,C

2024-03-29,312,59.760378016419516,A

2024-03-30,307,63.34270514914542,A

2024-03-31,336,17.27679936994437,A

2024-04-01,437,43.2689010455264,C

2024-04-02,466,31.79439444496833,C

2024-04-03,152,82.28257807419062,A

2024-04-04,379,52.327057100143456,A

2024-04-05,316,98.50808268053586,B

2024-04-06,351,45.894199820009774,A

2024-04-07,287,83.47886858974455,B

2024-04-08,479,81.8510612486096,A

2024-04-09,140,23.564578956888653,B

# Приложение 2. Список команд для установки библиотек

pip install pandas

pip install matplotlib

pip install seaborn

pip install plotly