ОНЛАЙН-УНИВЕРСИТЕТ URBAN

**ДИПЛОМНАЯ РАБОТА**

на тему:

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

**Выполнил:**

Махонин Андрей Николаевич

Петропавловск-Камчатский

2024

**Содержание**

Введение 3

Особенности библиотек 4

Критерии сравнения 5

Описание данных 5

Заключение 7

Список использованных источников 8

Приложение 9

**Введение**

Визуализация данных - это процесс создания графиков, диаграмм и других визуальных представлений данных. Она используется для отображения данных в удобном и понятном виде, что позволяет лучше понять и интерпретировать данные. Визуализация данных также помогает выявлять закономерности и тренды в данных, которые могут быть скрыты в голых числах.

В Python есть множество библиотек для визуализации данных: Matplotlib, Seaborn, Plotly, Bokeh, Altair и многие другие. Каждая из этих библиотек предоставляет свои уникальные возможности для создания различных типов графиков и диаграмм. В работе на примерах использования первых трёх библиотек Matplotlib, Seaborn, Plotly будут показаны их некоторые особенности и возможности.

**Особенности библиотек**

Matplotlib - это одна из самых популярных библиотек визуализации данных на Python. Она предоставляет широкий набор функций для создания различных типов графиков, включая линии, столбцы, круговые диаграммы и много больше. Matplotlib также предоставляет гибкость и контроль над каждым аспектом графика. Она особенно полезна для простых и стандартных задач визуализации данных.

Основные особенности Matplotlib включают:

1. Гибкость и контроль: Matplotlib предоставляет множество функций для настройки и контроля над каждым аспектом графика, включая цвет, стиль, размер и положение элементов графика.

2. Широкий набор функций: Matplotlib предоставляет широкий набор функций для создания различных типов графиков, включая линии, столбцы, круговые диаграммы и много больше.

3. Интеграция с другими библиотеками: Matplotlib хорошо интегрируется с другими библиотеками Python, такими как NumPy, SciPy и Scikit-learn.

4. Поддержка векторного и растрового графики: Matplotlib поддерживает как векторный, так и растровый графики, что позволяет создавать графики высокого качества.

5. Поддержка множества форматов файлов: Matplotlib поддерживает множество форматов файлов, включая PNG, SVG и PDF, что позволяет сохранять графики в различных форматах.

Seaborn - это библиотека визуализации данных, которая предоставляет высокоуровневые API для создания красивых и интуитивно понятных графиков. Она предоставляет набор стандартных графиков, таких как графики рассеяния, графики ящика и графики тепловых карт.

Основные особенности Seaborn включают:

1. Красивые и интуитивно понятные графики: Seaborn предоставляет набор стандартных графиков, которые легко интерпретировать и понять.

2. Гибкость и контроль: Seaborn предоставляет гибкость и контроль над каждым аспектом графика, включая цвет, стиль, размер и положение элементов графика.

3. Интеграция с другими библиотеками: Seaborn хорошо интегрируется с другими библиотеками Python, такими как Matplotlib, NumPy и SciPy.

4. Поддержка множества форматов файлов: Seaborn поддерживает множество форматов файлов, включая PNG, SVG и PDF, что позволяет сохранять графики в различных форматах.

5. Поддержка множества типов данных: Seaborn поддерживает множество типов данных, включая числовые, категориальные и текстовые данные.

Plotly - это библиотека визуализации данных, которая предоставляет интерактивные графики в браузере. Она предоставляет широкий набор функций для создания различных типов графиков, включая линии, столбцы, круговые диаграммы и много больше.

Основные особенности Plotly включают:

1. Интерактивные графики: Plotly предоставляет интерактивные графики, которые можно легко обновить и настроить.

2. Широкий набор функций: Plotly предоставляет широкий набор функций для создания различных типов графиков, включая линии, столбцы, круговые диаграммы и много больше.

3. Интеграция с другими библиотеками: Plotly хорошо интегрируется с другими библиотеками Python, такими как NumPy, SciPy и Matplotlib.

4. Поддержка множества форматов файлов: Plotly поддерживает множество форматов файлов, включая PNG, SVG и PDF, что позволяет сохранять графики в различных форматах.

5. Поддержка множества типов данных: Plotly поддерживает множество типов данных, включая числовые, категориальные и текстовые данные.

**Критерии сравнения**

Чтобы сравнить различные библиотеки, можно использовать много критериев, но я остановлюсь на следующих:

1. Функциональность: сравнение функциональности каждой библиотеки. Какие функции и возможности предоставляет каждая библиотека.

2. Удобство использования: оценить удобство использования каждой библиотеки. Определить, насколько легко можно начать работать с каждой библиотекой и насколько легко можно найти нужную информацию в документации.

3. Поддержка: оценить поддержку каждой библиотеки. Определить, насколько хорошо каждая библиотека поддерживается сообществом разработчиков и сколько документации и примеров доступно.

Во многих случаях этого достаточно, чтобы определить, какая библиотека наиболее подходит для конкретных потребностей проекта.

В этой работе на каждой из библиотек будет написано несколько примеров их использования, для визуализации данных и таким образом проявятся вышеназванные критерии.

**Описание данных**

Созданы четыре функции, каждая из которых возвращает данные, которые могут быть использованы для визуализации.

Функция get\_lissajous\_figure генерирует данные для фигур Лиссажу. Это траектории, прочерчиваемые точкой, совершающей одновременно два гармонических колебания в двух взаимно перпендикулярных направлениях.

Функция get\_pandas\_dataframe загружает данные из CSV-файла и возвращает их в виде объекта DataFrame из библиотеки pandas.

Функция get\_github\_list\_data получает данные о количестве коммитов в репозитории GitHub за последние несколько недель.

Функция get\_3d\_data генерирует данные для 3D-графики. Она создает сетку точек в трехмерном пространстве и возвращает их в виде трех массивов: x\_grid, y\_grid и z\_grid. Получающаяся графика похожа на бутон цветка.

**Заключение**

Каждый из этих инструментов имеет свои особенности и предназначения, поэтому выбор между ними зависит от конкретных потребностей проекта. Если нужно создать простой и стандартный график, Matplotlib может быть хорошим выбором. Он предоставляет гибкость и контроль над каждым аспектом графика, что позволяет создавать сложные и уникальные графики. Seaborn более заточена под анализ данных и для этого код на ней может быть более лаконичным. Если нужно создать интерактивный график в браузере, Plotly может быть лучшим выбором. Он предоставляет широкий набор функций для создания различных типов графиков и позволяет легко обновить и настроить графики.

**Список использованных источников**

1. https:/www.youtube.com
2. https://matplotlib.org
3. https://seaborn.pydata.org
4. https://plotly.com/python

**Приложение**

Исходный программный код проекта доступен по интернет адресу:

https://github.com/Andrey-Ing/urban-university\_homework/tree/master/Диплом