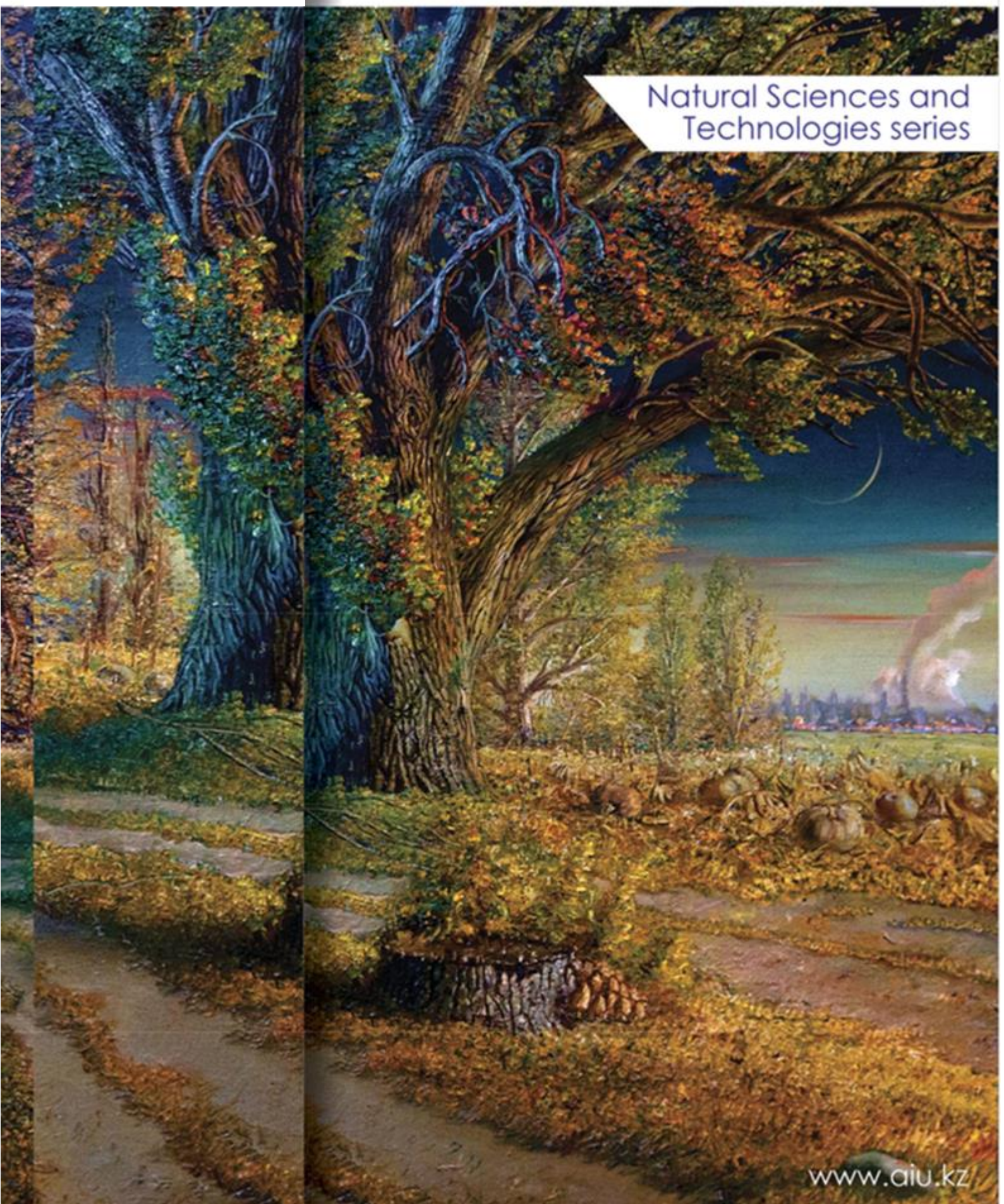
№3 (5) 2024

INTERNATIONAL SCIENCE REVIEWS

AIU

Natural Sciences and Technologies series



www.aiu.kz

ISSN: 2707-4862

ASTANA INTERNATIONAL UNIVERSITY

INTERNATIONAL SCIENCE REVIEWS
Natural Sciences and Technologies series

Has been published since 2020

№3 (5) 2024

Astana

INTERNATIONAL SCIENCE REVIEWS. NATURAL SCIENCES AND TECHNOLOGIES SERIES ЖУРНАЛЫНЫҢ РЕДАКЦИЯСЫ

БАС РЕДАКТОР

Қалимолдаев Мақсат Нұрадилович, техникалық ғылымдар докторы, ҚР ҰҒА академигі, профессор, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты бас директорының кеңесшісі, бас ғылыми қызметкері (*Қазақстан*)

БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ

Мырзағалиева Анар Базаровна, биология ғылымдарының докторы, профессор, бірінші вице-президент, Астана халықаралық университеті (*Қазақстан*);

РЕДАКТОРЛАР:

- **Сейткан Айнур Сейтканқызы**, техника ғылымдарының кандидаты, PhD, жаратылыстану ғылымдары жоғары мектебінің деканы, Астана халықаралық университеті (*Қазақстан*);
- **Муканова Асель Сериковна,** PhD, Ақпараттық технологиялар және инжерения жоғары мектебінің деканы, Астана халықаралық университеті (*Қазақстан*);
- **Абдилдаева Асель Асылбековна,** PhD, қауымдастырылған профессор, Әл-Фараби атындағы ҚазҰУ (Қазақстан);
- **Хлахула Иржи** PhD, профессор, Познаньдағы Адам Мицкевич атындағы университет (Польша);
- **Редферн Саймон А.Т.,** PhD, профессор, Наньян технологиялық университеті (Сингапур);
- **Сяолей Фенг,** PhD, Наньян технологиялық университеті (Сингапур);
- **Шуджаул Мулк Хан,** PhD, профессор, Каид-және-Азам университеті (Пакистан);
- **Базарнова Наталья Григорьевна,** химия ғылымдарының докторы, профессор, Химия және химиялық-фармацевтикалық технологиялар институты (Ресей);
- **Черёмушкина Вера Алексеевна,** биология ғылымдарының докторы, профессор, РҒА СБ Орталық Сібір ботаникалық бағы (Ресей);
- **Тасболатұлы Нұрболат,** PhD, Ақпараттық технологиялар және инжерения жоғары мектебі деканының орынбасары, Астана халықаралық университеті (Қазақстан);
- **Байшоланов Сакен Советович,** география ғылымдарының кандидаты, доцент, Астана халықаралық университеті (*Қазақстан*);
- **Нуркенов Серик Амангельдинович,** PhD, қауымдастырылған профессор, Астана халықаралық университеті (*Қазақстан*).

РЕДАКЦИЯ ЖУРНАЛА INTERNATIONAL SCIENCE REVIEWS. NATURAL SCIENCES AND TECHNOLOGIES SERIES

ГЛАВНЫЙ РЕДАКТОР

Калимолдаев Максат Нурадилович, доктор технических наук, академик НАН РК, профессор, ГНС, советник генерального директора Института информационных и вычислительных технологии КН МНВО РК (*Казахстан*)

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА

Мырзагалиева Анар Базаровна, доктор биологических наук, профессор, первый вице-президент, Международный университет Астана (*Казахстан*)

РЕДАКТОРЫ:

- **Сейткан Айнур Сейтканкызы**, кандидат технических наук, PhD, декан высшей школы естественных наук, Международный университет Астана (Казахстан);
- **Муканова Асель Сериковна,** PhD, декан Высшей школы информационных технологии и инженерии, Международный университет Астана (Казахстан);
- **Абдилдаева Асель Асылбековна,** PhD, ассоциированный профессор, Казахский национальный университет имени Аль-Фараби (Казахстан);
- **Хлахула Иржи** PhD, профессор, Университет имени Адама Мицкевича в Познани (Польша);
- **Редферн Саймон А.Т.,** PhD, профессор, Наньянский технологический университет (Сингапур);
- **Фенг Сяолей,** PhD, Наньянский технологический университет (Сингапур);
- **Шуджаул Мулк Хан,** PhD, профессор, Университет Каид-и Азама (Пакистан);
- **Базарнова Наталья Григорьевна,** доктор химических наук, профессор, Институт химии и химико-фармацевтических технологий (Россия);
- **Черёмушкина Вера Алексеевна,** доктор биологических наук, профессор, Центральный Сибирский Ботанический сад СО РАН (Россия);
- **Тасболатұлы Нұрболат,** PhD, заместитель декана Высшей школы информационных технологии и инженерии, Международный университет Астана (Казахстан);
- **Байшоланов Сакен Советович,** кандидат географических наук, доцент, Международный университет Астана (Казахстан);
- **Нуркенов Серик Амангельдинович,** PhD, ассоциированный профессор, Международный университет Астана (Казахстан);

EDITORIAL TEAM OF THE JOURNAL INTERNATIONAL SCIENCE REVIEWS. NATURAL SCIENCES AND TECHNOLOGIES SERIES

CHIEF EDITOR

Maksat Kalimoldayev, Doctor of Technical Sciences, Academician of NAS RK, Professor, SRF, CEO's councelor «The Institute of Information and Computational Technologies» CS MSHE RK (Kazakhstan)

DEPUTY CHIEF EDITOR

Anar Myrzagaliyeva, Doctor of Biological Sciences, Professor, First Vice-President, Astana International University (Kazakhstan)

EDITORS:

- **Ainur Seitkan**, Candidate of Technical Sciences, PhD, Dean of the Higher School of Natural Sciences, Astana International University (Kazakhstan);
- **Assel Mukanova,** PhD, Dean of the Higher School of Information Technology and Engineering, Astana International University (Kazakhstan);
- **Assel Abdildayeva,** PhD, Associate Professor, of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University (Kazakhstan);
- **Jiri Chlachula,** PhD, Dr.Hab., Full Professor, Adam Mickiewicz University, Poznań (Poland);
- **Simon A.T. Redfern,** PhD, Professor, Nanyang Technological University (Singapore);
- **Xiaolei Feng,** PhD, Nanyang Technological University (Singapore);
- **Khan Shujaul Mulk,** PhD, Professor, Quaid-i-Azam University (Pakistan);
- **Natal'ya Bazarnova,** Doctor of Chemical Sciences, Professor, Institute of Chemistry and Chemical-Pharmaceutical Technologies (Russia);
- **Vera Cheryomushkina,** Doctor of Biological Sciences, Professor, Central Siberian Botanical Garden SB RAS (Russia);
- **Nurbolat Tasbolatuly,** PhD, Deputy Dean of the Higher School of Information Technology and Engineering, Astana International University (Kazakhstan);
- **Saken Baisholanov,** Candidate of Geographical Sciences, Associate Professor, Astana International University (Kazakhstan);
- **Serik Nurkenov,** PhD, Associate Professor, Astana International University (Kazakhstan).

Editorial address: 8, Kabanbay Batyr avenue, of.316, Nur-Sultan,
Kazakhstan, 010000
Tel.: (7172) 24-18-52 (ext. 316)
E-mail: natural-sciences@aiu.kz

International Sciense Reviews NST - 76153
International Science Reviews
Natural Sciences and Technologies series
Owner: Astana International University
Periodicity: quarterly
Circulation: 500 copies

CONTENT

1. С.Е.Базаров, А.М.Касымханов, С.Б.Нигметжанов, А.А.Аманжолов БҰҚТЫРМА СУ ҚОЙМАСЫНДАҒЫ КӘСІПШІЛІК МАҢЫЗЫ БАР ЖЫРТҚЫШ БАЛЫҚ ТҮРЛЕРІНІҢ ЭКОЛОГИЯЛЫҚ-БИОЛОГИЯЛЫҚ СИПАТТАМАСЫ.....7

2. А.М. Касымханов СОВРЕМЕННОЕ ГИДРОХИМИЧЕСКОЕ СОСТОЯНИЕ ОЗЕРА ЖАЙСАН.....12

3. Е.К.Жаксылыкова, С.Б. Абеуова ВНЕДРЕНИЕ ЭКОЛОГИЧЕСКИХ ПРАЗДНИКОВ: ШАГ К ЗЕЛЕНОМУ БУДУЩЕМУ ……19

4. Ж.Балтабай ОРТА МЕКТЕПТЕ БИОЛОГИЯНЫ ОҚЫТУДЫҢ ЗЕРТТЕУ ӘДІСТЕРІ ……..... 24

5. M.Karibayeva ABOUT THE CHOICE OF DRINKING WATER PURIFICATION METHODS ……..……30

6. А.Б.Карабалаева, Т.Абуова, А.Шәлтік, М.Ермек БИОДЕГРАДАЦИЯ (УТИЛИЗАЦИЯ) БЫТОВОГО ПЛАСТИКА МИКРОМИЦЕТАМИ ……....37

7. С.Б. Марзен, А.Е.Сулейменова М ИЗУЧЕНИЕ ОСОБЕННОСТЕЙ ПРИМЕНЕНИЯ БИОЛОГИЧЕСКИ АКТИВНОГО ВЕЩЕСТВА Д3 ……..…44

8. Г.Қ. Ситахметова, А.М. Касымханов, А.А. Аманжолов СОСТАВ И РАСПРЕДЕЛЕНИЕ АКТИВНОЙ МОЛОДИ РЫБ ПО РЕКЕ ЕРТИС ……..….48

9. А. Б. Шуакбаева АУЫЗ СУДЫҢ ФИЗИКА-ХИМИЯЛЫҚ ЖӘНЕ МИКРОБИОЛОГИЯЛЫҚ КӨРСЕТКІШТЕРІ ……54

10. Т.А.Бейсен, О.К.Суйинханов, А.Н.Сұлтанғазиева БІЛІМ БЕРУДЕГІ ЖАСАНДЫ ИНТЕЛЛЕКТ: АРТЫҚШЫЛЫҚТАРЫ МЕН БОЛАШАҒЫ.….....61

11. Т.А.Шалбай, М.Ж.Қалдарова МӘТІН БОЙЫНША АДАМНЫҢ КӨҢІЛ-КҮЙІН ТАНУҒА ӘРТҮРЛІ ТӘСІЛДЕРДІ ТАЛДАУ ……..67

12. Ж.Семейхан, М.Ж.Калдарова ФРЕЙМВОРКИ ДЛЯ FRONTEND-РАЗРАБОТКИ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ .....……...78

13. Тишбаева А.Д., Рыспаева Д.С. ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В МЕДИЦИНЕ КАЗАХСТАНА ……..88

14. Женіс Д.Е., Бурибаева А.К. ЖАСАНДЫ ИНТЕЛЛЕКТ ЖӘНЕ ОНЫҢ ҚОҒАМҒКА ӘСЕРІ: ЖИ ЕНГІЗУДІҢ ЭТИКАЛЫҚ, ЭКОНОМИКАЛЫҚ ЖӘНЕ ӘЛЕУМЕТТІК АСПЕКТІЛЕРІН ЗЕРТТЕУ 93

15. А.Т. Баймуханбетова ПЕРВЫЙ ШАГ В PYTHON ДЛЯ АНАЛИЗА ДАННЫХ: ИСПОЛЬЗОВАНИЕ БИБЛИОТЕК PANDAS И MATPLOTLIB 101

ПЕРВЫЙ ШАГ В PYTHON ДЛЯ АНАЛИЗА ДАННЫХ: ИСПОЛЬЗОВАНИЕ БИБЛИОТЕК PANDAS И MATPLOTLIB

Баймуханбетова Анель Талгатовна

4 course Data Engineering,
e-mail: bajmuhanbetovaanel@gmail.com
Astana International University, Astana, Kazakhstan

Аннотация. В данной статье рассматриваются ключевые моменты в работе с библиотеками Pandas и Matplotlib для анализа данных в языке программирования Python. Обсуждаются основные методы работы с табличными данными внутри которой входит загрузка, обработка и фильтрация данных с использованием Pandas, а также создание различных типов графиков для визуализации данных с помощью Matplotlib. Также были рассмотрены способы работы с отсутствующими значениями, фильтрация данных на основе условий, создание простых и сложных графиков для наглядного представления данных.

Ключевые слова – Python, Pandas, Matplotlib, библиотека, табличные данные, фильтрация, типы графиков, визуализация.

ВВЕДЕНИЕ

Анализ данных - это процесс получения данных, их понимания и оценка имеющей информации с использованием аналитических и логических рассуждений для получения итоговых данных и составление выводов и отчетов. Включает методику, при которой данные собираются и структурируются таким образом, чтобы обеспечить ясное представление и анализ., чтобы из них можно было извлечь полезную информацию. Но анализ данных это не просто анализ информации после ее сбора и получения, это также способ проверки выдвинутых гипотез. Цель любого анализа данных состоит в том , чтобы глубже понять исследуемую информацию в целом.[3]

Целью данной статьи является сведение основных возможностей анализа данных с помощью библиотек Pandas, Matplotlib на языке Python. Актуальность применения этих библиотек – их широкое применение как в научно-исследовательских работах, так и в разных индустриях, что дает возможность быстро и эффективно анализировать и визуализировать большие объемы данных.

Анализ данных часто используется в различных отраслях, например в таких как бизнес, финансы, наука, медицина, продажи и многое другое. Он позволяет принимать разумные решения на основе фактов и данных. А также предсказывать

будущие тенденции и результаты для развития какой либо сферы. Большим плюсом является что анализ данных может помочь выявить проблемы и недостатки в бизнес – процессах и принять меры по их устранению.

Python является одним из самых популярных языков программирования для анализа данных и славится своей простатой, гибкостью и эффективными способами обработки больших объемов информации. Python предоставляет широкий выбор различных инструментов и библиотек для работы с данными, включая Pandas, NumPy, Matplotlib, Seaborn, Scikit – learn и другие. [5]

Pandas — главная библиотека , то есть набор заранее подготовленных методов и функций, в Python. [1] Она быстрая и мощная, в ней можно работать с таблицами, в которых миллионы строк и также использовать уже созданные и протестированные алгоритмы. Само название pandas образовано как от panel data (панельные данные), при-меняемого в эконометрике термина для обозначения многомерных структурированных наборов данных, так и от фразы Python data analysis.[8]

NumPy сокращение от “Numerical Python” представляет собой часто используемой библиотекой в работе с числовыми данными в языке программирования Python.[8] Ее основная работа заключается в том, что обеспечивает эффективные инструменты для выполнения операций над массивами данных. [2]

Matplotlib – это из популярных библиотек для визуализации данных, который предоставляет большие возможности для построения различных графиков. [3] С его помощью можно создавать линейные графики, гистограммы, круговые диаграммы и многие другие типы визуальных графиков. Это позволяет более наглядно понимать и видеть информацию, выявлять тренды и обнаруживать закономерности в данных. Графики интерактивны – можно увеличить масштаб определенного объекта и выполнять панорамирование с помощью панели инструментов.[8]

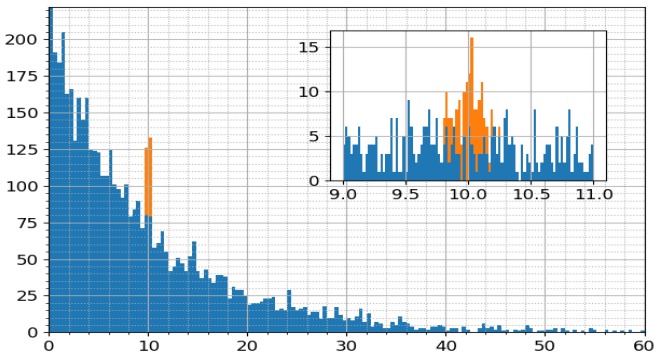


Рисунок 1. – Результат визуализации Matplotlib

МЕТОДЫ И МЕТОДОЛОГИЯ

Seaborn - это мощная и гибкая библиотека для визуализации данных, которая работает выше Matplotlib. Благодаря Seaborn можно построить более качественнее визуализацию данных, чем с предыдущей библиотекой. С данной библиотекой вам удастся легко отображать сложные зависимости между переменными, разными видами графиков (столбчатые и прямоугольные диаграммы, точечные, линейные, тепловые карты). [7]

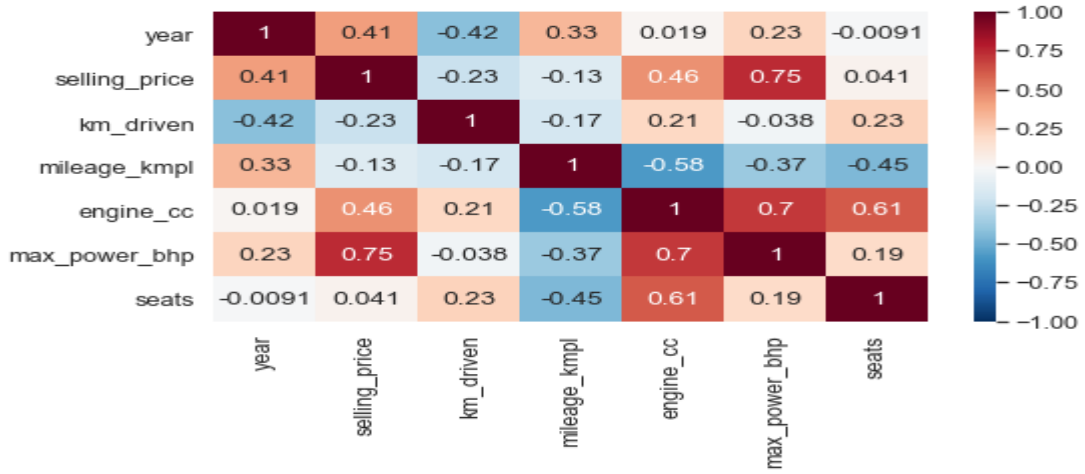


Рисунок 2. – Результат визуализации Seaborn

Scikit – learn - это простая и удобная библиотека для machine learning в языке программирования Python. Благодаря этой библиотеке можно создавать модель, которая способна предсказывать данные и находить закономерности. Также есть готовые алгоритмы для классификации, регрессии, кластеризации и других задач. [4]

В ходе исследования использовались библиотеки Pandas для работы с табличными данными и Matplotlib для визуализации результатов. Данные были взяты из открытых источников в интернете и не проверялись на точность, поэтому их использование направлено на демонстрацию возможностей этих инструментов. В процессе анализа были созданы линейные графики, гистограммы и круговые диаграммы, что позволило наглядно представить и исследовать выбранные наборы данных

Анализ данных в Pandas

Поэтапная инструкция по анализу данных на Python (Google Colab)

В этом разделе мы представим поэтапную инструкцию по анализу данных с использованием Python. Процесс анализа данных можно разбить на несколько этапов, начиная с подготовки среды и загрузки данных и заканчивая их

визуализацией и интерпретацией, каждый этап будет представлен примером и фотографией.

Перед началом работы нужно настроить рабочую среду для всего процесса. Изначально, удостоверится что установлены все необходимые библиотеки, такие как Pandas для работы с табличными данными и Matplotlib для их визуализации. Чтобы установить данные библиотеки вам нужно использовать менеджер пакетов pip: 'pip install pandas matplotlib'

Как только установили данные библиотеки вам необходимо открыть вашу рабочую среду в данном случае (Google Colab) . После создайте новый блокнот.[1]

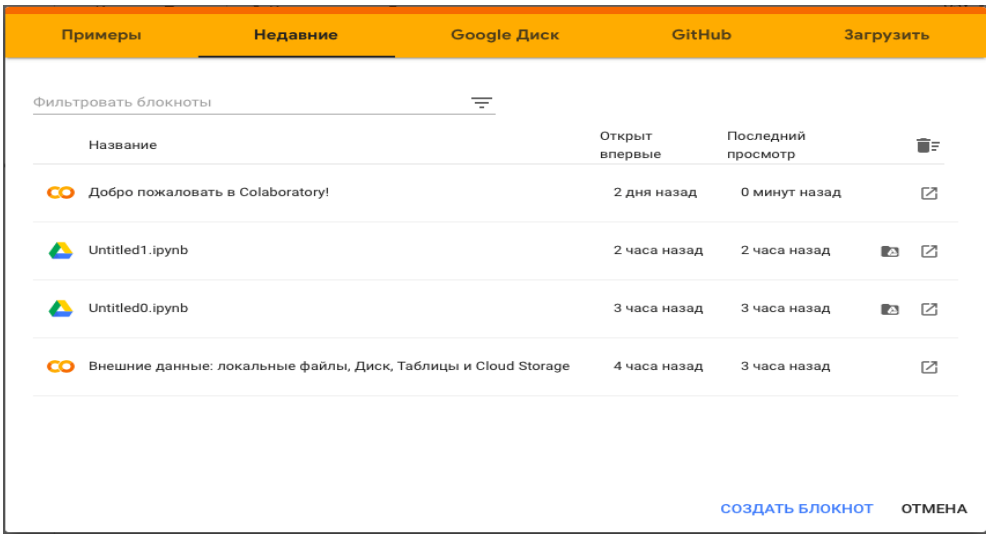


Рисунок 3. - Создание нового блокнота

Импортирование библиотеки

Чтобы начать работу с Pandas вам следует импортировать данные с помощью следующего кода:

Import pandas as pd

**Примечание pd – это сокращенное название библиотеки.*

После импортирования библиотек следующим шагом является загрузка данных, которые будут использоваться для анализа. Обычно данные загружаются из внешних файлов, таких как CSV, Excel или базы данных, и с помощью библиотек, таких как Pandas, этот процесс становится простым и интуитивно понятным. [1]

import pandas as pd

df = pd.read_csv('data.csv')

Создать объект типа DataFrame

Это будет полезно если есть несколько списков и нужно соединить их в одну таблицу или если хотите наглядно оформить небольшой набор данных. Это можно произвести несколькими способами, например через словарь и через преобразование вложенных списков.[2]

Через словарь (1)

my_df = pd.DataFrame({'id': [1, 2, 3], 'name': ['Bob', 'Alice', 'Scott'], 'age': [21, 15, 30]})

Через вложенные списки (2)

df = pd.DataFrame([[1,'Bob', 21], [2,'Alice', 15], [3,'Scott', 30]], columns = ['id','name', 'age'])

Результаты будут одинаковыми.

Просмотр данных

После загрузки данных полезно воспользоваться методом df.head(), который по умолчанию выводит пять первых строк.[1]

	Overall rank	Country or region	Score	GDP per capita	Social support	Healthy life expectancy	Freedom to make life choices	Generosity	Perceptions of corruption
0	1	Finland	7.769	1.340	1.587	0.986	0.596	0.153	0.393
1	2	Denmark	7.600	1.383	1.573	0.996	0.592	0.252	0.410
2	3	Norway	7.554	1.488	1.582	1.028	0.603	0.271	0.341
3	4	Iceland	7.494	1.380	1.624	1.026	0.591	0.354	0.118
4	5	Netherlands	7.488	1.396	1.522	0.999	0.557	0.322	0.298

Рисунок 7. - Просмотр данных с помощью df.head()

Если нужно посмотреть на другое количество строк, оно указывается в скобках, например df.head(12). Также последние строки фрейма выводятся методом .tail().

В свою очередь чтобы полностью красиво отобразить датасет, используется функции display(). В Jupyter Notebook, если переменная указана в последней строке ячейки без использования ключевого слова display, е ё содержимое автоматически отобразится.

display(df) #эквивалентно команде df, если это последняя строка ячейки.

Добавление новых строк и столбцов

В исходный набор данных можно добавлять новые столбцы, формируя так называемые «признаки», что часто используется в машинном обучении. Для примера, создадим столбец «Сумма», в котором будут просуммированы значения

из колонок «ВВП на душу населения» и «Социальная поддержка» (хотя суммирование этих показателей не является практически значимым, сделаем это для учебных целей):df['Сумма'] = df['ВВП на душу населения'] + df['Социальная поддержка'][1]

	Место в рейтинге	Страна или регион	Баллы	ВВП на душу населения	Социальная поддержка	Ожидаемая продолжительность здоровой жизни	Свобода жизненных выборов	Щедрость	Восприятие коррупции	Сумма
0	1	Finland	7.769	1.340	1.587	0.986	0.596	0.153	0.393	2.927
1	2	Denmark	7.600	1.383	1.573	0.996	0.592	0.252	0.410	2.956
2	3	Norway	7.554	1.488	1.582	1.028	0.603	0.271	0.341	3.070
3	4	Iceland	7.494	1.380	1.624	1.026	0.591	0.354	0.118	3.004
4	5	Netherlands	7.488	1.396	1.522	0.999	0.557	0.322	0.298	2.918

Рисунок 8. – добавление новых значений

Удаление строк и столбцов

Чтобы удалить отдельные столбцы вам необходимо использовать метод *drop()* . В Pandas есть несколько способов для обработки пропущенных данных, включая удаление строк с пропусками и заполнение отсутствующих значений:

data = data.dropna() “Удаление строк с пропущенными значениями” data = data.fillna(0) “Заполнение пропусков значением 0”

Визуализация данных с Matplotlib

Визуализация данных одна их самых важных частей в процессе анализа данных, данная библиотека содержит большое количество инструментов для двумерных графиков. Также она помогает лучше понять структуру данных и выявить закономерности и какие либо тренды. Она легка в использовании и позволяет получать графики в высоком качестве.

РЕЗУЛЬТАТЫ

Для начала работы с данной библиотекой необходимо установить и импортировать ее в проект, установку можно произвести также с помощью менеджера пакетов “pip”

pip install matplotlib – установка библиотеки

import matplotlib.pyplot as plt – импортирование ее в код.

Часто matplotlib импортируется с ее упрощенной формой “ plt ”, в дальнейшем будет использоваться данная форма для вызова методов библиотеки. В этой части кода используется модуль (matplotlib.pyplot), который во многом копирует инструменты пакета Matlab. Давайте рассмотрим пример, нарисуем графики функций sin и cos с matplotlib.pyplot.[2]

```
import numpy as np

import matplotlib.pyplot as plt

phi = np.linspace(0, 2.*np.pi, 100)

plt.plot(phi, np.sin(phi))

plt.plot(phi, np.cos(phi))

plt.show()
```

в результате получаем-

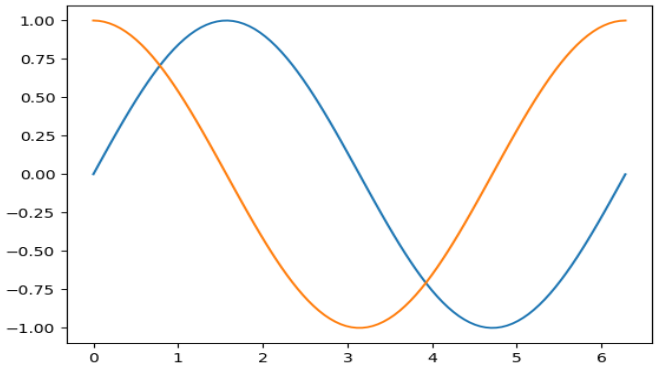


Рисунок 9. - График функции sin и cos[2]

Мы воспользовались функцией plot, передав ей два аргумента — списки значений для осей X и Y. При последовательных вызовах этой функции графики отображаются на одних и тех же осях, а цвета линий автоматически меняются для каждого нового графика. Matplotlib предоставляет множество вариантов графиков, которые могут быть полезны для различных аналитических задач. Разнообразие инструментов позволяет глубже анализировать данные и выявлять важные закономерности, что делает Matplotlib востребованным для анализа информации в самых разных сферах.

Линейные графики - применяются, когда нужно показать, как меняются значения со временем или как одна переменная влияет на другую. Например, они отлично подходят для отображения тенденций роста или падения чисел, таких как доходы за несколько месяцев. Линии соединяют точки данных, что позволяет легко проследить общий тренд и заметить изменения в значениях.[2]

Пример кода линейного графика

```
plt.plot(x, y)
```

Рисунок 10. – Пример линейного графика

Гистограммы - отлично подходят для того, чтобы показать, как данные распределяются по определенным интервалам или категориям. Например, с их помощью можно проанализировать распределение возрастов в группе людей. Каждый столбец на гистограмме отображает количество элементов, попадающих в тот или иной диапазон значений, что позволяет легко понять, в каких пределах сосредоточено большинство данных.

Пример кода гистограммы

```
import matplotlib.pyplot as plt

data = [25, 32, 47, 56, 43, 22, 36, 42, 55, 30]

plt.hist(data)

plt.title('Пример гистограммы')

plt.show()
```

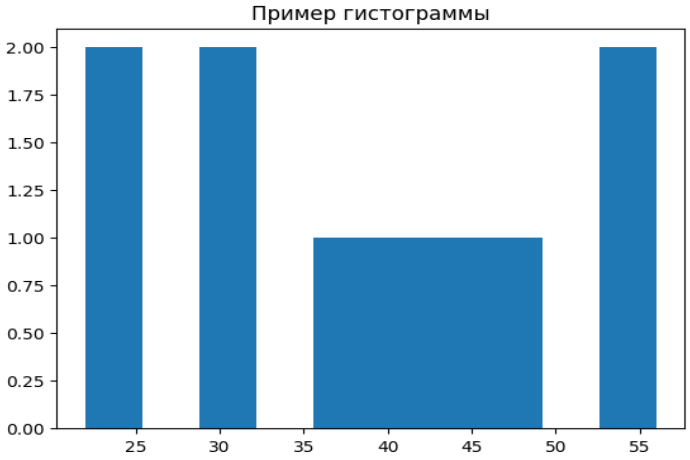


Рисунок 11. – Пример гистограммы

Круговые диаграммы - используются для того, чтобы показать, какую долю каждая категория занимает в общей массе данных. Круг разбивается на сегменты, и размер каждого сегмента зависит от величины категории. Это помогает быстро понять, какая часть данных принадлежит той или иной категории и сравнить их между собой. [2]

Пример кода для круговой диаграммы

```
import matplotlib.pyplot as plt

import pandas as pd

categories = ['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J']

data_series = pd.Series(data, index=categories, name='category')

data_series.value_counts().plot(kind='pie', autopct='%1.1f%%')

plt.title('Доля категорий')

plt.show()
```

Рисунок 12. – Пример круговой диаграммы

Функции plot позволяет задавать тип маркера, тип линии и цвет.

```
fmt = '[marker][line][color]'
```

Пример использвани данной функции

```
x = np.linspace(0, 1, 100)

f1 = 0.25 - (x - 0.5)**2

f2 = x**3

plt.plot(x, f1, ':b')    # пунктирная синяя линия

plt.plot(x, f2, '--r')   # штрихованная красная линия

plt.plot(x, f1+f2, 'k')  # черная непрерывная линия

plt.show()
```

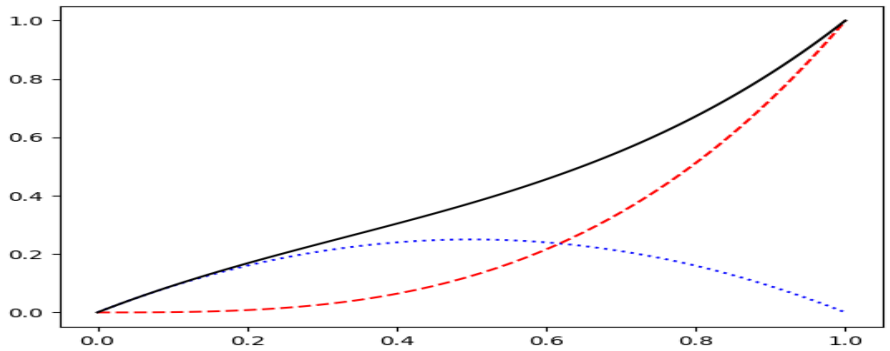


Рисунок 13. – Результат данного кода с использованием функции plot[2]

Более тонкую настройку параметров можно выполнить, передавая различные именованные аргументы, например:

marker: str — тип маркера

markersize: float — размер маркера

linestyle: str — тип линии

linewidth: float — толщина линии

color: str — цвет

Пример использования Matplotlib на практике чтобы показать, как Matplotlib может быть применён для решения реальных задач, построим график, отображающий выручку компании по месяцам на основе реальных данных

```
months = ['Январь', 'Февраль', 'Март', 'Апрель', 'Май', 'Июнь']

revenue = [2000, 3000, 3500, 4000, 4500, 5000]

plt.plot(months, revenue, marker='o', linestyle='-', color='b')
```

```
plt.title('Выручка по месяцам')

plt.xlabel('Месяцы')

plt.ylabel('Выручка')

plt.grid(True)

plt.show()
```

Рисунок 14. – Результат кода “Выручка по месяцам”

Этот график демонстрирует рост выручки компании по месяцам и может помочь выявить тенденции и улучшить стратегию ведения бизнеса.

Анализ данных показал, что с помощью библиотек Pandas и Matplotlib можно эффективно проводить обработку и визуализацию информации. Так, на одном из графиков представлены изменения синусоидальных функций, что иллюстрирует широкие возможности Matplotlib для создания разнообразных графиков. Визуализация данных по выручке компании (см. рисунок 14) продемонстрировала стабильный рост показателей по месяцам, что свидетельствует о положительной динамике в бизнесе. Эти результаты могут служить основой для более глубокого анализа и дальнейшего планирования стратегических решений.

Использование Pandas и Matplotlib продемонстрировало их эффективность в анализе и визуализации данных. Эти библиотеки отлично справляются с задачами обработки и представления информации в графическом виде. Однако при работе с большими объемами данных можно рассмотреть дополнительные инструменты, такие как Dask или Plotly, которые расширяют возможности в работе с массивными наборами данных и интерактивной визуализацией. В дальнейшем возможно интегрировать эти решения для решения более сложных задач анализа.

ЗАКЛЮЧЕНИЕ

Python – это широко используемый язык программирования в области анализа данных. Он предоставляет множество библиотек и инструментов для обработки и анализа данных, включая библиотеку Pandas и Matplotlib. Мы начали с основ Pandas изучив методы загрузки и первичной обработки данных таких как создание новых объектов, добавление и удаление данных . Эти навыки позволят вам эффективно управлять данными и подготавливать их для дальнейшего использования. Далее мы рассмотрели Matplotlib, инструмент для создания визуализаций. Мы научились строить различные типы графиков: линейные графики для отслеживания изменений во времени, гистограммы для анализа распределения данных и круговые диаграммы для отображения долей. Эти графики делают данные более понятными и помогают выявить важные тренды.

Использование Pandas и Matplotlib в совокупности позволяет проводить эффективный анализ данных и представлять результаты в наглядной и доступной форме. Эти навыки необходимы для принятия обоснованных решений и подготовки информативных отчетов в самых разных областях, от бизнеса до научных исследований.

В итоге, библиотеки Pandas и Matplotlib показали себя как мощные инструменты для решения прикладных задач анализа данных. Они позволяют эффективно обрабатывать значительные объемы информации и визуализировать результаты, что упрощает процесс выявления ключевых закономерностей и трендов. В будущем их потенциал можно будет расширить, применяя для более сложных проектов, включая работу с большими данными и задачи машинного обучения. Это откроет новые горизонты как для исследователей, так и для аналитиков, стремящихся к более глубокому пониманию данных.

СПИСОК ЛИТЕРАТУРЫ

1. Мария Жарова, "Как анализировать данные на Python с Pandas: первые шаги", SKILLFACTORY MEDIA, 25 июня 2024
2. "Язык Python: Визуализация данных с Matplotlib", C++ и Python.
3. А. М. Васильченко, "Как проводить анализ данных при помощи Python", Инновации и инвестиции, №5, 2023
4. "Что такое Scikit Learn - гайд по популярной библиотеке Python для начинающих", DataStart.
5. "What is Data Analysis and Why It Matters?", Linkedin, 12 ноября 2023
6. "Анализ данных: определение, цели и задачи", Аналитика ПЛЮС.

7. Мария. Жарова, "Анализ данных на Python с Pandas и Matplotlib", Habr, 12 июля 2022.

8. Маккинни У. Python и анализ данных. – Litres, 2022.