

## Визуализация данных с помощью Plotly или Altair

*Анишкова Анастасия Сергеевна*

*Приамурский государственный университет имени Шолом-Алейхема*

*Студент*

### Аннотация

Цель исследования - реализовать и сравнить возможности библиотек Plotly и Altair для построения интерактивных и статических визуализаций данных на примере датасета «penguins» в интерактивной среде Google Colab. Google Colab — это бесплатная среда для разработки и выполнения программного кода в облаке. Результат может быть применён в обучающих целях.

**Ключевые слова:** Визуализация данных, Plotly, Altair, Google Colab.

## Data visualization using Plotly or Altair

*Anishkova Anastasia Sergeevna*

*Sholom-Aleichem Priamurskiy State University*

*Student*

### Abstract

Implement and compare the capabilities of the Plotly and Altair libraries for building interactive and static data visualizations using the «penguins» dataset in the Google Colab interactive environment. Google Colab is a free environment for developing and executing software code in the cloud. The result can be used for educational purposes.

**Key words:** Data visualization, Plotly, Altair, Google Colab.

## 1 Введение

### 1.1 Актуальность

Визуализация данных играет важную роль в анализе и представлении информации, особенно с увеличением объёмов данных. Использование современных инструментов, таких как Plotly и Altair, позволяет наглядно отображать данные, выявлять закономерности и эффективно передавать результаты исследования. Эти библиотеки предоставляют широкие возможности для построения интерактивных и статических графиков, что делает их востребованными в научной и практической деятельности.

### 1.2 Обзор исследований

П. А. Пылов, А. В. Протоdjяконов описали программные библиотеки языка программирования Python для реализации алгоритмов визуализации данных [1]. Сравнительный анализ библиотек визуализации данных для задачи обработки образовательной информации провели О. А. Шутова,

Д.А.Грамаков [2]. А. Оразгулыев, С. А. Гараджаева визуализировали данные и представили результаты исследования [3]. Библиотеки Python для визуализации научных концепций в образовательных целях применили Н.К.Аркабаев, Б. У. Доолотбек [4]. Е. В. Филина смоделировала и визуализировала данные на языке программирования Python с помощью библиотеки Plotly в различных областях знаний [5].

### 1.3 Цель исследования

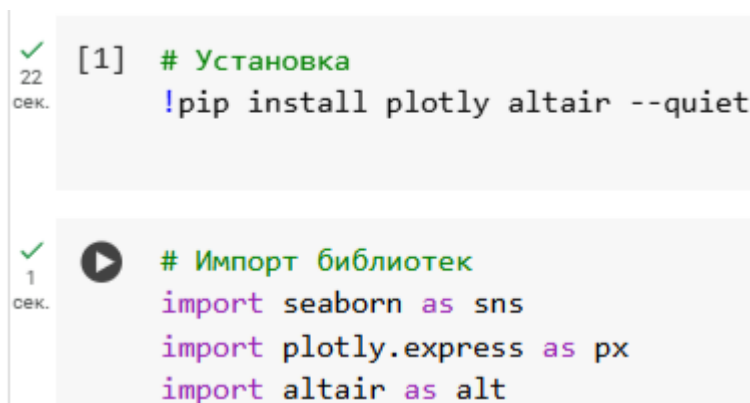
Цель исследования - реализовать и сравнить возможности библиотек Plotly и Altair для построения интерактивных и статических визуализаций данных на примере датасета «penguins» в интерактивной среде Google Colab.

## 2 Материалы и методы

Использовались библиотеки Plotly и Altair для визуализации данных на основе датасета «penguins».

## 3 Результаты

Установим и подключим библиотеки (рис.1) seaborn, plotly и altair для работы с визуализацией данных, позволяя строить разнообразные графики и диаграммы в интерактивной среде Google Colab.



```
[1] # Установка
!pip install plotly altair --quiet

# Импорт библиотек
import seaborn as sns
import plotly.express as px
import altair as alt
```

Рисунок 1. Установка и импорт библиотек

Загрузим датасет «penguins» из библиотеки seaborn и выводим первые несколько строк таблицы для ознакомления с данными (рис.2).

```

✓ [3] # Загрузка датасета
0 сек. df = sns.load_dataset("penguins")

✓ # Просмотр первых строк
0 сек. df.head()

```

|   | species | island    | bill_length_mm | bill_depth_mm | flipper_length_mm | body_mass_g | sex    |
|---|---------|-----------|----------------|---------------|-------------------|-------------|--------|
| 0 | Adelie  | Torgersen | 39.1           | 18.7          | 181.0             | 3750.0      | Male   |
| 1 | Adelie  | Torgersen | 39.5           | 17.4          | 186.0             | 3800.0      | Female |
| 2 | Adelie  | Torgersen | 40.3           | 18.0          | 195.0             | 3250.0      | Female |
| 3 | Adelie  | Torgersen | NaN            | NaN           | NaN               | NaN         | NaN    |
| 4 | Adelie  | Torgersen | 36.7           | 19.3          | 193.0             | 3450.0      | Female |

Рисунок 2. Загрузка датасета

Построим точечную диаграмму с помощью Plotly (рис.3), где по осям отложены длина клюва и масса тела пингвинов, а цвет точек соответствует виду пингвина. График позволяет визуальнo оценить зависимость между этими признаками для разных видов.

```

#Точечная диаграмма: длина клюва в зависимости от массы тела, цвет — вид пингвина
fig = px.scatter(df,
                 x="bill_length_mm",
                 y="body_mass_g",
                 color="species",
                 title="Длина клюва vs масса тела (Plotly)",
                 labels={"bill_length_mm": "Длина клюва (мм)", "body_mass_g": "Масса тела (г)})
fig.show()

```



Рисунок 3. Точечная диаграмма

Построим ящичковую диаграмму с помощью Plotly, показывающую распределение массы тела пингвинов по полу, что позволяет сравнить диапазоны и средние значения массы между самцами и самками (рис.4).

```

#Ящичковая диаграмма: масса тела по полу
fig = px.box(df, x="sex", y="body_mass_g", color="sex",
             title="Распределение массы тела по полу (Plotly)")
fig.show()

```

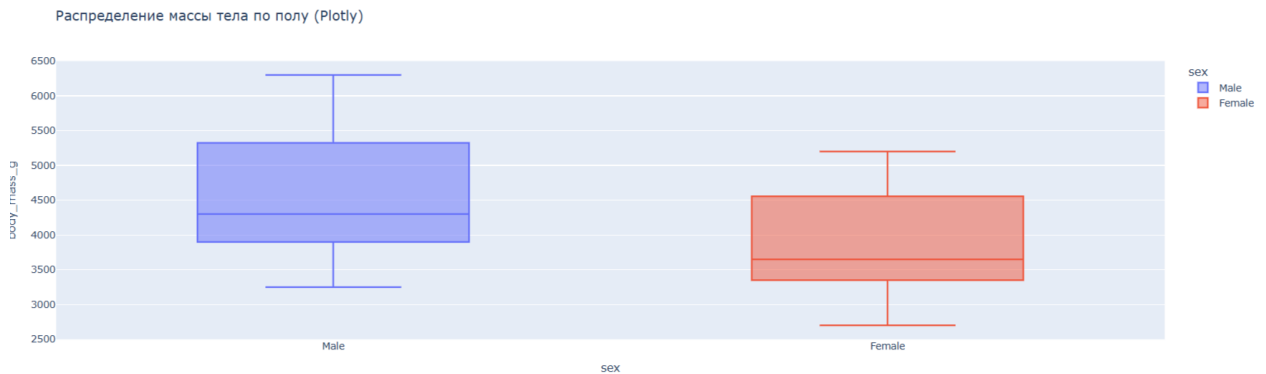


Рисунок 4. Ящичковая диаграмма

Построим гистограмму с помощью Plotly (рис.5), отображающую распределение длины клюва пингвинов по интервалам, что позволяет оценить частоту встречаемости разных значений длины клюва в датасете.

```
#Гистограмма: распределение длины клюва
fig = px.histogram(df, x="bill_length_mm", nbins=30,
                  title="Распределение длины клюва (Plotly)",
                  labels={"bill_length_mm": "Длина клюва (мм)"})
fig.show()
```

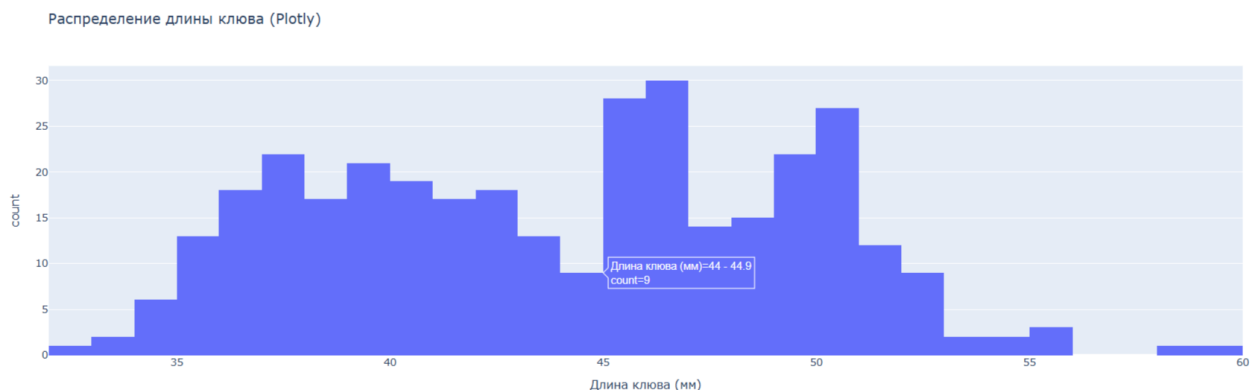


Рисунок 5. Гистограмма

Построим линейный график с помощью Plotly (рис.6), отображающий среднюю массу тела пингвинов по островам, что позволяет визуальнo оценить различия в массе между пингвинами, живущими на разных островах.

```
#Линейный график: средняя масса тела по островам
df_grouped = df.groupby('island')['body_mass_g'].mean().reset_index()

fig = px.line(df_grouped, x="island", y="body_mass_g", markers=True,
              title="Средняя масса тела по островам (Plotly)",
              labels={"body_mass_g": "Средняя масса (г)", "island": "Остров"})
fig.show()
```

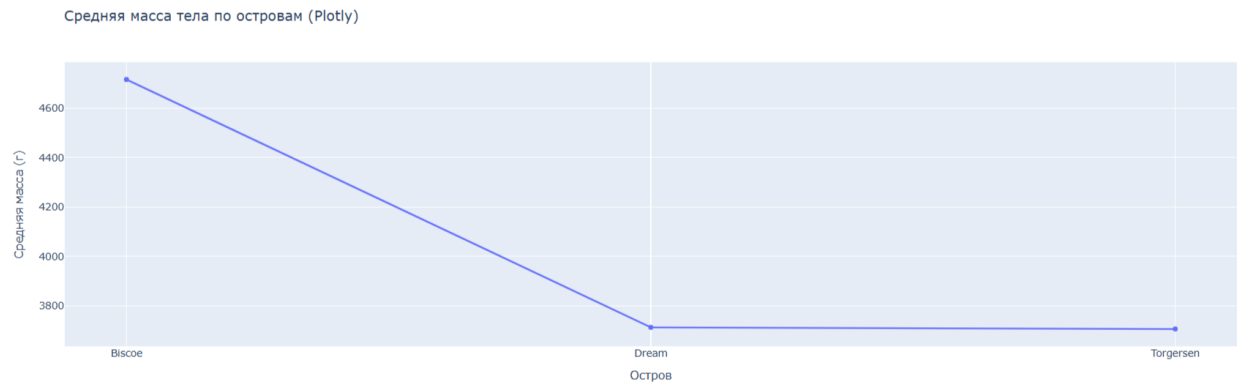


Рисунок 6. Линейный график

Построим столбчатую диаграмму с помощью Altair, отображающую количество пингвинов на каждом острове, что позволяет быстро сравнить численность пингвинов по их месту обитания (рис.7).

```
#Гистограмма: количество пингвинов на каждом острове
chart = alt.Chart(df).mark_bar().encode(
    x='count()',
    y='island',
    color='island',
    tooltip=['island', 'count()']
).properties(title="Количество пингвинов по островам (Altair)")
chart.display()
```



Рисунок 7. Столбчатая диаграмма

Построим групповую столбчатую диаграмму с помощью Altair, показывающую распределение самцов и самок пингвинов по островам, что позволяет сравнить численность разных полов на каждом из островов (рис.8).

```
#Групповая гистограмма: количество самцов и самок по островам
chart = alt.Chart(df.dropna(subset=['sex'])).mark_bar().encode(
    x='count()',
    y='island',
    color='sex',
    row='sex'
).properties(title="Пол пингвинов по островам (Altair)")
chart.display()
```

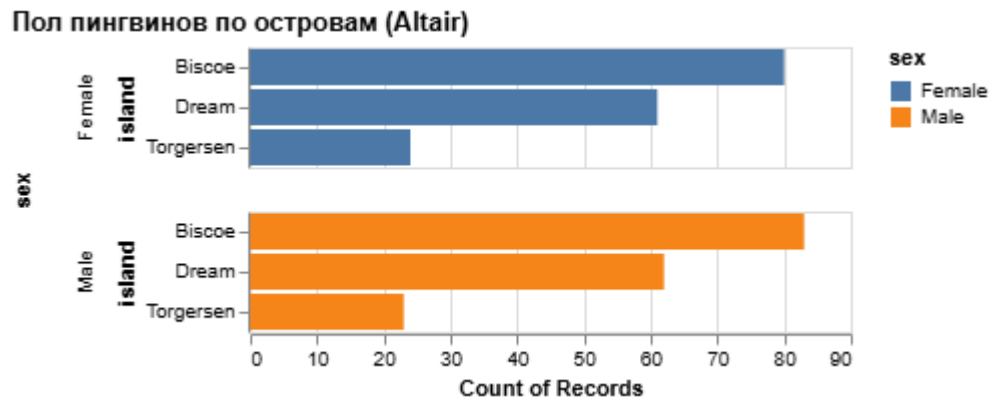


Рисунок 8. Столбчатая диаграмма

Построим фасетный точечный график (рис.9) с помощью Altair, отображающий зависимость длины клюва от его глубины для каждого вида пингвинов отдельно, что позволяет сравнить форму клюва между разными видами.

```
#Фасетный точечный график: длина клюва в сравнении с глубиной клюва, разделённое по видам
chart = alt.Chart(df).mark_circle(size=60).encode(
    x='bill_length_mm',
    y='bill_depth_mm',
    color='species',
    tooltip=['species', 'bill_length_mm', 'bill_depth_mm']
).facet(
    facet='species',
    columns=2
).properties(title="Длина vs глубина клюва по видам (Altair)")
chart.display()
```



Рисунок 9. Точечный график

Построим линейный график с точками, показывающий среднюю длину клюва пингвинов по островам и видам, что позволяет сравнить различия между видами на разных островах (рис.10).

```
#Линия + точка: среднее значение длины клюва по островам и видам
df_avg = df.groupby(['island', 'species'])['bill_length_mm'].mean().reset_index()

chart = alt.Chart(df_avg).mark_line(point=True).encode(
    x='island:N',
    y='bill_length_mm:Q',
    color='species:N'
).properties(title="Средняя длина клюва по островам и видам (Altair)")
chart.display()
```

Средняя длина клюва по островам и видам (Altair)

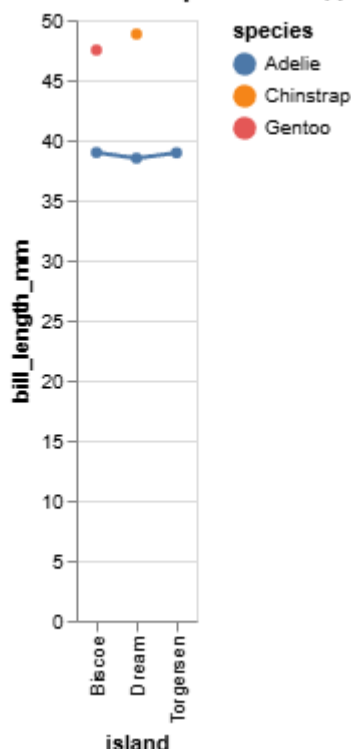


Рисунок 10. Линейный график

#### 4 Вывод

В ходе работы были изучены и применены на практике возможности библиотек Plotly и Altair для визуализации данных в Google Colab. Обе библиотеки позволяют строить различные типы графиков и наглядно представлять данные, но имеют свои особенности. С помощью Plotly были созданы точечные диаграммы, ящичковые диаграммы, гистограммы и линейные графики, которые отличаются высокой интерактивностью и наглядностью. Эта библиотека хорошо подходит для создания динамических визуализаций и информационных панелей. С помощью Altair были созданы столбчатые, групповые и фасетные диаграммы, а также линейные графики с точками. Altair предлагает более простой и декларативный способ построения графиков, что делает его удобным для быстрого анализа данных и создания чистых и структурированных визуализаций.

Выбор между Plotly и Altair зависит от целей исследования: если требуется интерактивность и динамика, предпочтительнее использовать Plotly, а если важна простота и чёткость построения графиков, оптимальным выбором будет Altair. Обе библиотеки являются мощными инструментами для работы с данными и могут эффективно использоваться для анализа и представления информации.

#### Библиографический список

1. Пылов П. А., Протоdjяконов А. В. Программные библиотеки языка программирования Python для реализации алгоритмов визуализации

- данных // Инновации. Наука. Образование. 2020. №. 23. С. 233-239.
2. Шутова О. А., Грамаков Д. А. Сравнительный анализ библиотек визуализации данных для задачи обработки образовательной информации //Проблемы теории и практики инновационного развития и интеграции современной науки и образования. 2021. С. 190-194.
  3. Оразгулыев А., Гараджаева С. А. Визуализация данных. Представление результатов исследования //Наука и мировоззрение. 2025. Т. 1. №. 40. С. 42-48.
  4. Аркабаев Н. К., Доолотбек Б. У. Применение библиотек python для визуализации научных концепций в образовательных целях //Science and innovation. 2024. Т. 3. №. Special Issue 50. С. 398-406.
  5. Филина Е. В. Моделирование и визуализация данных на языке программирования Python с помощью библиотеки Plotly в различных областях знаний //Вестник Саратовского областного института развития образования. 2020. №. 1. С. 81-88.