

### Задание 3. Анализ данных. Описательная статистика

1. Загрузить исходные данные (см. файл `rus_leader_heights.csv`). В файле представлен список некоторых правителей России (от Древней Руси до настоящего времени) и их рост (в см).

2. Провести описательную статистику исходных данных:

- найти статистические показатели (среднее значение, медиана, дисперсия, стандартное отклонение, мода, минимальное значение, максимальное значение);
- визуализировать распределение данных (построить гистограмму, диаграмму плотности, диаграмму размаха);
- вычислить скошенность и эксцесс данных;
- вычислить значения квартилей.

3. Проанализировать результаты описательной статистики и сделать вывод.

4. По результатам анализа преобразовать исходные данные так, чтобы они стали пригодными для использования при построении моделей машинного обучения.

5. Провести описательную статистику обновленных данных (см. п. 2-3).

*Примечание. При выполнении задания необходимо реализовать функции, где в качестве входных параметров будут использоваться анализируемые массивы данных. В теле функций для вычислений, отрисовки графиков и т.д. нужно использовать библиотеки `numpy`, `scipy`, `matplotlib`, `seaborn` и др. Не дублировать ячейки с кодом!*

6. С помощью функции из библиотеки `numpy` посчитать, сколько правителей:

- имеют рост более 170 см;
- имеют рост более 170 см, но менее 190 см.

7. Реализовать собственную функцию, которая реализует вычисление данных для построения гистограммы (аналог функции `numpy.histogram`). Использовать функции из библиотеки `numpy`: `linspace`, `zeros_like`, `searchsorted`, `add.at`. Протестировать на исходных данных к заданию. Сравнить с построенными гистограммами. Сделать вывод.

8. Написать отчет о проделанной работе.

Отчет должен включать:

- 1) Постановку задачи. Представление исходных данных в табличном виде.
- 2) Описательную статистику исходных данных. Для каждого показателя написать определение и роль в анализе данных. Привести получившиеся значения, графики, диаграммы. Сделать вывод, необходимы ли преобразования исходных данных, какие операции с данными требуется осуществить.

- 3) Описательную статистику обновленных данных. Сделать вывод. Объяснить, достаточно ли сделанных преобразований для использования при построении моделей машинного обучения. Привести аргументы.
- 4) Описание собственных функций.
- 5) Ответы на вопросы.
- 6) Общий вывод по всему заданию.
- 7) Список литературы (в тексте отчета **ОБЯЗАТЕЛЬНО** должны быть ссылки на используемую литературу)