Лабораторная работа №2

**Тема:** Корреляция. Статистические гипотезы и тесты на значимость.

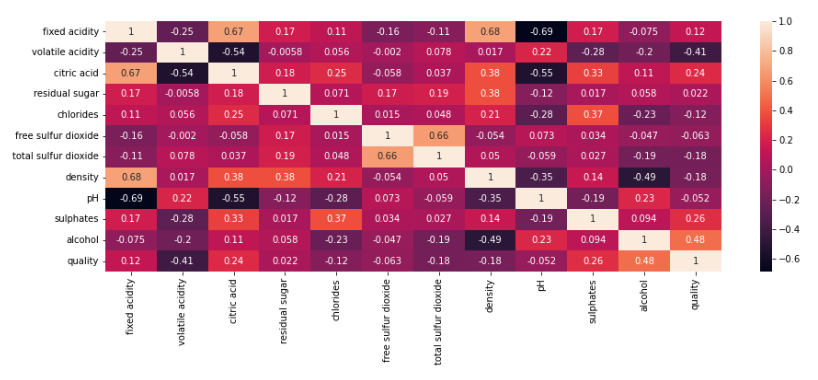
**Цель работы:** продолжение разведочного анализа данных с помощью Python.

**Ход работы:** открыть Google Colab либо Jupyter Notebook и загрузить датасет **согласно варианту**.

**Столбцы** датафрейма принято называть признаками или переменными

Посмотреть на распределение данных каждого признака.

Вывести матрицу корреляции и оценить ее. Сделать выводы касательно коррелирующих между собой признаков (сильная, слабая корреляция). Оценить мультиколлинеарность признаков. Использовать для этого можно как сам pandas, так и сторонние библиотеки seaborn или matplotlib.



Придумать две статистические гипотезы на свой датасет, принять или опровергнуть их, согласно проведенному статистическому тесту (критерий Стьюдента или t-тест, Хи-квадрат, Точный тест Фишера).

Для проведения статистических тестов можно использовать scipy stats.

При проверке статистических гипотез подробно описать процесс в каждой отдельной ячейке Google Colab или Jupyter Notebook. Сделать выводы о значимости того или иного признака.