

В рамках задачи необходимо провести анализ **одного** из источников данных, используя Python и соответствующие библиотеки для анализа данных.

Список Наборов данных (   ):

- COVID trials:  
<https://www.kaggle.com/parulpandey/covid19-clinical-trials-dataset>
- Список высокооплачиваемых спортсменов Forbes:  
<https://www.kaggle.com/datasets/parulpandey/forbes-highest-paid-athletes-19902019>
- Одобрения по кредитным картам :  
<https://www.kaggle.com/rikdifos/credit-card-approval-prediction>
- Анализ спроса на резервирование отелей:  
<https://www.interviewquery.com/blog-free-datasets/#:~:text=Hotel%20Booking%20Demand%20Data>
- Анализ маркетинговых показателей:  
<https://www.kaggle.com/jackdaoud/marketing-data>
- Статистика разводов:  
<https://www.interviewquery.com/blog-free-datasets/#:~:text=Divorce%20Predictors%20Dataset>

**Задание:**

1. Ознакомьтесь с выбранным набором данных. Посчитайте количество строк и столбцов
2. Проведите EDA(Exploratory Data Analysis), то есть
  - a. для каждой числовой переменной посчитайте :
    - i. Долю пропусков
    - ii. Максимальное и минимальное значение
    - iii. Среднее значение
    - iv. Медиану
    - v. Дисперсию
    - vi. Квантиль 0.1 и 0.9
    - vii. Квартиль 1 и 3
  - b. для каждой категориальной переменной посчитайте:
    - i. Долю пропусков
    - ii. Количество уникальных значений
    - iii. Моду
3. Постройте минимум 2 информативных графика, которые будут полезны для анализа

4. Сформулируйте 2 гипотезы и проверьте их с помощью статистических критериев

**Форма предоставления результата -**

Файл ipynb (jupyter ноутбук) или ссылка на google colab