# Лабораторная работа 03 (проверка качества данных)

## Задание 1: Проверка целостности данных.

Используйте набор данных о продажах и проверьте, заполнены ли все необходимые поля. Если обнаружены пропущенные значения, определите, какой процент они составляют от общего количества данных.

Набор данных для использования: "Online Retail" на Kaggle. Этот набор данных представляет собой транзакции, произошедшие в период с декабря 2010 по октябре 2011 года и зарегистрированные розничной компанией из Великобритании.

Ссылка: https://www.kaggle.com/mashlyn/online-retail-ii-uci

## Задание 2: Проверка консистентности данных.

Рассмотрите набор данных о климате и проверьте, соблюдаются ли форматы дат во всех записях.

Набор данных для использования: "Climate Change: Earth Surface Temperature Data" на Kaggle.

Ссылка: https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data

## Задание 3: Проверка актуальности данных.

Анализируя набор данных о твитах, проверьте даты публикации твитов и определите, какие из них актуальны на текущий момент.

Набор данных для использования: "COVID19 Tweets" на Kaggle, включающий твиты о пандемии COVID-19.

Ссылка: https://www.kaggle.com/gpreda/covid19-tweets

## Задание 4: Проверка полноты данных.

Рассмотрите набор данных о заболеваемости раком, проанализируйте, приведены ли данные по всем странам или некоторые из них отсутствуют.

Набор данных для использования: "Global Cancer Statistics 2018 (GLOBOCAN 2018)" на Kaggle.

Ссылка: https://www.kaggle.com/ashkhagan/global-cancer-statistics-2018-globocan-2018