

# Project sequence

## ML project sequence

### Data Collection

- Survey
- Collect from websites (web scraping)
- Find ready-made datasets (GitHub, Kaggle)
- From client

### Familiarization with Data

- Place files and datasets in the same location (side by side)
- Install Pandas
- Import Pandas
- Load data correctly
- View first/last rows (`head / tail`)
- Input / Output operations
- Remove unnecessary data

## ML project sequence (Part 2)

**Data Preprocessing** (Preparing the data)

**Model Selection** (Choosing the model)

**Model Training** (Training the model) — 70%

**Model Prediction** (Making predictions)

**Model Evaluation** (Assessing performance) — 15%

**Model Testing** (Testing the model) — 15%

**Model Deployment** (Deploying the model)

**Model Maintenance / Monitoring** (Maintaining and monitoring the model)

## Data Preprocessing

- **Mean**
- **Mode**
- **Median**
- **Fixed**

## Libraries

Pandas

- **Data Entry** (Inputting data)
- **Familiarization with Data** (Getting to know the data)
- **Data Transformation and Cleaning** (Modifying and cleaning the data)
- **Data Analysis** (Analyzing the data)

**Numpy**

- **Statistics and Mathematical Management**
- **Mathematical Analysis and Calculations**

**Matplotlib**

- **Graphs / Charts**
- **Designs**

**Seaborn**

- **Displaying statistical data in various ways (in a visual format)**

**Scikit-learn**

- **ML Training**
- **Executing the stages of an ML project**