# ITD105 Laboratory Exercises #2

# Data Collections

**Name:** Clint Joshua O. Velasquez

**Subject & Section:** ITD105 – IT4D

1. **Public Dataset.** Look for at least 4 published public datasets (2 for classification and 2 for regression) in websites such as Kaggle, UCI Machine Learning Repository, Google Dataset search, and data.gov.

# Output:

|  |  |  |  |
| --- | --- | --- | --- |
| **Title** | **Type (Classification/**  **Regression)** | **Description** | **Source Link** |
| 1. Car Price Prediction (Linear Regression - RFE) | Regression | This model is intended to help the management team gain insights into how car prices are influenced by these independent factors. By analyzing this relationship, the management can make informed decisions regarding car design, business strategies, and other aspects to achieve specific price targets. | https://www.kaggle.com/code/goyalshalini93/car-price-prediction-linear-regression-rfe/notebook |
| 2. Car Price Prediction | Regression | This dataset contains information about used cars.  This data can be used for a lot of purposes such as price prediction to exemplify the use of linear regression in Machine Learning. | https://www.kaggle.com/code/mohaiminul101/car-price-prediction/notebook |
| 3. Statlog (Vehicle Silhouettes) | Classification | The purpose is to classify a given silhouette as one of four types of vehicles, using a set of features extracted from the silhouette. The vehicle may be viewed from one of many different angles. | https://archive.ics.uci.edu/dataset/149/statlog+vehicle+silhouettes |
| 4. Car Evaluation | Classification | Car Evaluation Database was derived from a simple hierarchical decision model originally developed for the demonstration of DEX, M. Bohanec, V. Rajkovic: Expert system for decision making. Sistemica 1(1), pp. 145-157, 1990.). The model evaluates cars according to the a set of concept structures. | https://archive.ics.uci.edu/dataset/19/car+evaluation |

1. **Web Scraping**
   1. Scrape table in <https://pokemondb.net/pokedex/all> to create a dataset.

Instruction link:

<https://drive.google.com/drive/folders/1uLjA2fObzVtl1m0LaN2lKaPWadZfovZO?usp=sharing>

# Note: Use MSEdge driver and follow the updated code.

**Output:**

* + - Video the python code and output.

**Video Link:**

Gdrive: <https://drive.google.com/file/d/1KioZ9JnD6ne04h2xT1mBFYCyaHKcUomA/view?usp=drive_link>

Youtube: <https://youtu.be/iKA7tunKGpQ>

**Source Code:**

<https://github.com/kiyojiii/WebScraping>

* 1. Select a Philippine website. Scrape the data and create a dataset.

# Output:

* + - Video the python code and output.
    - Provide the sample dataset with its features and label in excel.
    - Identify what type of dataset (Classification/ Regression).

**Type of Dataset:** Classification

**Video Link:**

Gdrive: <https://drive.google.com/file/d/11Pg-n-EWLHl_bnVJ3wPmP5zS5VsAuhKZ/view?usp=sharing>

Youtube: <https://youtu.be/GeG4Nbk9WIE>

**Sample Dataset Link:**

<https://docs.google.com/spreadsheets/d/1a6C2LwL50uFNFHbuFlTSgEiqsMsr8uTw/edit?usp=sharing&ouid=108608604768269683518&rtpof=true&sd=true>

**Source Code:**

<https://github.com/kiyojiii/WebScraping>

# Surveys and Questionnaires.

**Output:**

Create a survey questionnaire that collects specific data from people. Provide the sample dataset with its features and label in excel (at least 20 records). Identify what type of dataset (Classification/ Regression).

**Type of Dataset:** Classification

**Survey Questionnaire Link:**

<https://forms.gle/CdV6PYSfAEuvrkTv7>

**Sample Dataset Link:**

<https://docs.google.com/spreadsheets/d/1L_37L37x11tNNE7Ip1Ndb0DmNAPUt3hQb2jKV_tn3kE/edit?usp=sharing>

***Bonus:***

***Using API, collect data programmatically in any websites (Twitter API, Google Maps API, or financial market data APIs).*** Provide the sample dataset with its features and label in excel. Identify what type of dataset (Classification/ Regression).

**Type of Dataset:** Classification

**Sample Dataset Link:**

<https://docs.google.com/spreadsheets/d/1UlZcIW8V5Utox43kN1549DOeTYIOjxYa/edit?usp=sharing&ouid=108608604768269683518&rtpof=true&sd=true>

**Source Code:**

<https://github.com/kiyojiii/WebScrapingUsingAPI>

**Video and Output:**

Gdrive: <https://drive.google.com/file/d/16zmq8kyQ5lMVUAr5XX_6L77vH-Kskvns/view?usp=drive_link>

Youtube: <https://youtu.be/P4nRf3wi2Io>