**Denise Dodd**

**Github Repository**

**https://github.com/ddodd24/Data-Science**

I have selected the below projects to highlight in my Github Repository because they showcase a diverse range of projects demonstrates my versatility and proficiency across different platforms and technologies

Python Projects:

1. Olympics (Data Cleaning and Visualizations) - The first half of this project highlights extensive data cleaning, data manipulation, and joining tables. The second half of this project highlights a large variety of data visualization options. This project also highlights obtaining Ing data from a flat file, a web file, and an API.

2. Classifying Images (Convolutional Neural Network (CNN)) - Used functions of the Keras module to train a CNN Classifier to identify pictures of numbers. Analyzed results using an accuracy score and confusion matrix.

3. Solar Energy (Trends, Correlations, Regressions) - Explores the usage of three sectors (wind, solar, hydro) of clean energy both by country and over time. Correlations are found between the three sectors of clean energy. The “Country” variable is transformed into dummy variables so Linear and Ridge Regressions can be performed predicting which countries will be using the most clean energy in the future. R-Squared, RMSE, and MAE metrics are analyzed.

4. Miami Housing Price Prediction (Random Forest Regression with Scaling and Hypertuning) - Predicts the sales price of Miami homes based on distance to important landmarks. I scaled numerical data and ran training and testing sets through several regression models using RMSE, R-Squared, and Regression plot to determine Random Forest was the best fit. I used grid search to find the best hyperparameters for the Random Forest model. Created presentation informing Miami Real Estate company how this regression will be beneficial for their clients.

5. St. Louis Crime (A Study of Trends and Predictive Analytics) – Uses a dataset of St. Louis crime reported in the year 2020. Performs a variety of filtering, grouping and visualizing techniques to analyze trends in crime across St. Louis. Also utilizes a Holt-Winters time series predictive model to predict future crime volume in St. Louis. A presentation is included which can be presented to the St. Louis Police Department.

6. Project #2 – Possibly something with a car dataset either predicting prices or classifying cars. This is a previous project that attempted and ended up pivoting from because it was too difficult, but I believe I am better equipped now to handle dummy variables and Boolean variables.

7. Project # 3 – Possibly a recommender using a dataset of books.

Non-Python Projects:

8. Aviation Dashboard (Tableau) - This project shows my abilities to import a dataset, create visualizations, and combine the visualizations into a dashboard using the Tableau platform.

9. Finances in Education (R, Visualization, Correlation, Quadratic Model, Markup) - This project uses R coding to explore how state education expenditure, teacher pay, and average income affects a student's obtainment of a bachelor's degree.

10. Pet Database (Python Functions and SQL Database) - In this project, I join tables within a SQL database. I use SQL queries to isolate and count entries and calculate metrics.