**CMSC 636 – Data Visualization**

**Assignment 1 – Part 2 Tutorial**

Faisal Rasheed Khan

VB02734

[vb02734@umbc.edu](mailto:vb02734@umbc.edu)

**Dataset:**

Columns: ‘aiddata\_id’, ‘year’, ‘donor’, ‘recipient’, ‘commitment\_amount\_usd\_constant’, ‘coalesced\_purpose\_code’, ‘coalesced\_purpose\_name’

**Tutorial Steps:**

**Step 0: Install all the required packages.**

* Use the command: pip install #package\_name
* Install pandas, matplotlib, numpy, squarify, scikitlearn

**A screenshot of a computer code

Description automatically generated**

**Step 1: Loading dataset into Python**

* Download the dataset
* Load the dataset in python using pandas library by specifying the path of the dataset

A screenshot of a computer program

Description automatically generated

* Look at the dataset, how it is

A screenshot of a computer

Description automatically generated

**Step 2: Data Preprocessing**

* Processing is the main step for a good visualization which appears good for a user to have information insights.
* First, we need to clean the data, as our data is clean and free of outliers, we proceed for the processing step.

A computer code with text

Description automatically generated with medium confidence

A screenshot of a computer code

Description automatically generated

**Step 3: Visualization for Vis/Task 3**

* We will be plotting here the countries whose aid has been increased over the years.
* To know this information we use linear regression to plot the line for the data.
* With this plot we will be knowing which countries has increased aid over the years.

**A screenshot of a computer code

Description automatically generated**

A graph of the country's average sales

Description automatically generated with medium confidence

**Step 4: Visualization for Vis/Task 4**

* We will be visualizing the plot using matplotlib and will be plotting Treemap.
* The Squarify package provides the Treemap in python.
* Here, created the custom legend to give the description regarding the colors

**A screenshot of a computer code

Description automatically generated**

A blue rectangle with black text

Description automatically generated

A graph showing the amount of a number of individuals

Description automatically generated with medium confidence

**Step 5: Adding Caption**

* Matplotlib provides the caption option, we need to provide customized text and the matplotlib displays that.

****

References:

[1] Tierney, Michael J., Daniel L. Nielson, Darren G. Hawkins, J. Timmons Roberts, Michael G. Findley, Ryan M. Powers, Bradley Parks, Sven E. Wilson, and Robert L. Hicks. 2011. More Dollars than Sense: Refining Our Knowledge of Development Finance Using AidData. World Development 39 (11): 1891-1906

[2] Matplotlib

[Tutorials — Matplotlib 3.8.3 documentation](https://matplotlib.org/stable/tutorials/index.html)

[3] Pandas

[pandas.DataFrame — pandas 2.2.1 documentation (pydata.org)](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.html)

[4] Scikit Learn

[scikit-learn: machine learning in Python — scikit-learn 1.4.1 documentation](https://scikit-learn.org/stable/index.html)

[5] Squarify

[Treemaps in matplotlib with squarify | PYTHON CHARTS (python-charts.com)](https://python-charts.com/part-whole/treemap-matplotlib/)

[6] Numpy

[NumPy -](https://numpy.org/)