Assignment 11: Data Visualization in Python

(60 Points Total)

Data available under Resources>Portland Data.

The **portland_data.gpkg** GeoPackage contains three data layers for Portland, Oregon:

neighborhoods: boundaries of Portland neighborhoods

park_trees: mapped trees in parks

street_trees: mapped trees along streets

The **neighborhood** boundaries were obtained from the City of Portland (https://www.portlandoregon.gov/28130) while the **tree** inventory data were obtained from the City of Portland Office of Parks & Recreation (https://www.portlandoregon.gov/parks/article/433143).

To complete this exercise, you will need to have set up a **conda environment** with the following: (1) **NumPy**, (2) **Pandas**, (3) **GeoPandas**, (4) all dependencies for **GeoPandas**, (5) **matplotlib**, and (6) **seaborn**. In this exercise, you will only work with the **park_trees** dataset. Also, you do not need the spatial information. The spatial information can be removed using the following code:

trees = gpd.read_file("YOUR FOLDER PATH/portland_data.gpkg", layer='park_trees')

trees2 = pd.DataFrame(trees.drop(columns='geometry'))

This exercise will focus on the use of matgplotlib and seaborn for creating data visualizations.

Description of Problem

Produce code to generate the graphs. Deliver code in a Python or plain text file.

Graph 1: Create a **scatterplot** for just trees in the *Ulmus* genus with DBH mapped to X and tree height mapped to Y (Hint: You will need to use the "Genus", "DBH", and "TreeHeight" attributes). **(6 Points)**

Graph 2: Create a **scatterplot** for just trees in the *Ulmus* genus with DBH mapped to X, tree height mapped to Y, and tree species mapped to the hue (Hint: You will need to use the "Genus", "Genus_spec", "DBH", and "TreeHeight" attributes). **(6 Points)**

Graph 3: Create a **boxplot** of DBH for just the *Ulmus* genus. **(6 Points)**

Graph 4: Create a **boxplot** of DBH for just the *Ulmus* genus differentiated by species (or, each species should have its own boxplot). **(6 Points)**

Graph 5: Combine Graphs 1 through 4 into a single figure with 2 rows and two columns. Do not plot a legend for any of the graphs. **(6 Points)**

Graph 6: Refine Graph 4 as follows: **(12 Points)**

- Provide a plot title and increase the font size.
- Define colors to use to fill the boxplots as opposed to using the default colors.
- Remove the legend.
- Change the x-axis title to "Species" and the y-axis title to "DBH". Set a font size for both.
- Set limits, tick intervals, tick interval labels, and a font size for the y-axis.
- Edit the species names on the x-axis. You will need to shorten some of them. Set a font size and italicize the scientific names. Apply a rotation to the labels.
- Add y-axis grid lines.

Graph 7: Create a figure containing Graph 4 and an additional graph of DBH vs. Tree Height for the *Ulmus* trees. In the new graph/subplot, refine as follows: **(18 Points)**

- Provide a plot title and increase the font size.
- Define colors to use to fill the points. These should match the colors used for the same species in the boxplot.
- Remove the legend.
- Change the x-axis title to "DBH" and the y-axis title to "Tree Height". Set a font size for both.
- Set limits, tick intervals, tick interval labels, and a font size for the x-axis.
- Set limits, tick intervals, tick interval labels, and a font size for the y-axis.
- Add y-axis grid lines.

Deliverables

- Code in Python or plain text file.