Data Visualization (Hands-on 1)

Problem Statement:
Visualizing data can be quite insightful for further analysis and narrowing on target problems. You are given sample data, after doing the basic analysis using pandas, visualize the outcomes to find insights and patterns in the data.

Use the cars dataset for the following questions that contains the following-
The dataset contains information about 260 cars that include horsepower, cubic inches, time to 60, brand, make year, weight, cylinders, mpg, etc.

Tasks To Be Performed:

1. Create scatter plots for the following data, make sure all the plots appear in the same plane with x-labels and y-labels evenly spaced.
    a. Create scatter plots with between (mpg-hp, mpg-weight), and (hp-mpg, hp-time-to-60) with respect to each brand.
2. Create bar plots for the following, make sure all the plots appear in the same plane with x-labels and y-labels evenly spaced.
    a. Create a bar plot that shows the visual representation of hp, mpg, weight, time-to-60 with respect to the number of cylinders for each of the brands.
    b. Create a bar plot that shows the visual representation of hp, mpg, weight and time-to-60 with respect to the years, for each brand.
3. Create box plots for the following:
    a. Create box plots for the columns in the step1 with respect to the number of cylinders and brand respectively.
4. Create pair plots for the entire data to study various patterns in the data.
    a. Create pair plots with respect to brand, number of cylinders, year, etc.
5. Create a heatmap for the entire data to study correlation between each of the columns.
6. Create a histogram for the following:
    a. Create histograms for hp, mpg, time-to-60, weight lbs, cubic inches, year, with respect to brand.
7. Create line plots for the following:
    a. Line plot between hp-mpg, hp-(time-to-60), mpg-(time-to-60), hp-weight lbs, mpg-weight lbs, cubic inches-weight lbs.
8. Create a pie chart for the following columns and their distribution in the data:
    a. Brands
9. Create violin plots for the following:
    a. hp, mpg, time-to-60, with respect to the number of cylinders for each brand.
10. Create categorical plots for the following:
    a. Hp, mpg, time-to-60, weight, cubic inches, with respect to each brand for various numbers of cylinders.
11. Create the area charts for the following:
    a. Mpg with respect to each brand of the car
12. Create Doughnut plots to represent the following:
    a. The cars are based on various counts of cylinders.

For the next set of questions, make the visualizations on the toyota car dataset that contains the following:

Price, age, KM, HP, Metcolor, Automatic, CC, Weight as columns. Load the dataset, handle the missing null values and execute the following.

Tasks to be performed:

1. Plot the scatterplots, and deduce the columns that contain a linear relationship between the price column.
2. Using the histograms, Deduce the frequency distribution of the Kilometers traveled by the cars.
3. Plot the count of cars with respect to fuel type in the toyota dataset.
4. Plot the count of cars with respect to fuel type and their type(automatic or not).
5. Plot the best fit line using the variables price and age from the toyota dataset.
6. Plot the frequency distribution of the age column with the kernel density estimate.