NumPy Problems

1. **Basic Array Operations**
   Convert the mpg column into a NumPy array and calculate:
   - The mean, median, and standard deviation of mpg.
   - The number of cars with mpg greater than 25.
2. **Filtering**
   Using NumPy, filter all cars with more than 6 cylinders.
   Return the corresponding car_name as a list.
3. **Statistical Analysis**
   Compute the 25th, 50th, and 75th percentiles of the weight column using NumPy.
4. **Array Manipulation**
   Convert the acceleration column into a NumPy array and normalize its values (scale between 0 and 1).
5. **Broadcasting**
   Increase all horsepower values by 10% and store the updated values in a new NumPy array. Handle missing data (if any) by replacing it with the mean of the column before applying the increase.
6. **Boolean Indexing**
   Find the average displacement of cars with an origin of 2 (Europe) using NumPy indexing.
7. **Matrix Operations**
   Create a 2D NumPy array containing the columns mpg, horsepower, and weight.
   Compute the dot product of this matrix with a given vector [1, 0.5, -0.2].
8. **Sorting**
   Use NumPy to sort the cars by model_year in descending order and display the first five car names.
9. **Correlation**
   Compute the Pearson correlation coefficient between mpg and weight using NumPy.
10. **Conditional Aggregates**
    Calculate the mean mpg for cars grouped by the number of cylinders using NumPy techniques.

Pandas Problems

1. **Basic Exploration**
   Load the dataset into a Pandas DataFrame. Display:
   - The first 10 rows
   - The total number of rows and columns
   - Summary statistics for numerical columns
2. **Filtering and Indexing**
   Find all cars manufactured in 1975 with a weight less than 3000. Return the DataFrame with selected columns: car_name, weight, and mpg.
3. **Handling Missing Data**
   Identify if there are any missing values in the dataset. Replace missing values in the horsepower column with the column's median.
4. **Data Transformation**
   Add a new column power_to_weight_ratio, calculated as horsepower / weight.
5. **Group By**
   Group the cars by origin and calculate the mean mpg for each group.
6. **Sorting**
   Sort the DataFrame by mpg in descending order and display the top 10 cars with the highest mpg.

   Apply Function
   Create a new column performance_score using a custom function:

```
def performance_score(row):
  return row['mpg'] * row['acceleration'] / row['weight']
```

7. Apply this function to each row and store the result in the new column.
8. **Visualization Preparation**
   Generate a summary DataFrame with:
   - Average mpg, weight, and horsepower for each model_year.
9. **Exporting Data**
   Save a subset of the data containing only mpg, cylinders, horsepower, and weight for cars with mpg > 30 into a CSV file named high_mpg_cars.csv.
10. **Finding Anomalies**
    Identify potential outliers in the mpg column using the **Interquartile Range (IQR)** method. Specifically:
    - Calculate the IQR for mpg.
    - Define outliers as values less than Q1 - 1.5 * IQR or greater than Q3 + 1.5 * IQR.

- Create a DataFrame of cars classified as outliers, displaying car_name, mpg, and model_year.

Matplotlib Problems

1. **What is the distribution of miles per gallon (mpg) in the dataset?**
   Plot a histogram of mpg values.

2. **How does mpg vary with the number of cylinders?**
   Use a boxplot to compare mpg across different cylinders.

3. **Is there a relationship between horsepower and mpg? Summarize your observation**
   Plot a scatter plot of horsepower vs. mpg.

4. **How does car weight influence mpg?**
   Plot a scatter plot with a trend line for weight vs. mpg.

5. **What is the trend of average mpg across model years?**
   Plot a line chart of average mpg per model year.

6. **How is the count of cars distributed by origin?**
   Use a bar chart to show the number of cars for each origin.

7. **How do acceleration values vary across different cylinders?**
   Use a boxplot of acceleration grouped by cylinders.

8. **Which year had the most number of car entries?**
   Plot a histogram or bar chart of car counts by model year.

9. **Is there a clustering pattern among weight, horsepower, and mpg?**
   Create a 3D scatter plot of these three variables.

10. **Which 10 cars have the best fuel efficiency?**
    Plot a horizontal bar chart showing the top 10 car names with the highest mpg.