Practical Questions on Data Cleaning and Pre-processing

1. **Question**: Implement a Python script to read a CSV file and handle missing data by filling with the mean value of each column.
    - **Hint**: Use the `fillna` function in Pandas.
2. **Question**: Write a Python script to remove duplicate rows from a dataset. Verify the results by checking the number of rows before and after removal.
    - **Hint**: Use the `drop_duplicates` function in Pandas.
3. **Question**: Implement a Python script to transform a column of a dataset using a log transformation. Plot the histogram of the original and transformed data.
    - **Hint**: Use the `apply` function in Pandas and `log` function from NumPy.
4. **Question**: Write a Python script to replace all occurrences of a specific value in a column with a new value. Verify the replacement by displaying the unique values of the column before and after.
    - **Hint**: Use the `replace` function in Pandas.
5. **Question**: Implement a Python script to detect and filter outliers in a dataset using the Z-score method. Plot the dataset before and after filtering outliers.
    - **Hint**: Use the `zscore` function from the SciPy library.
6. **Question**: Write a Python script to perform string manipulation on a column of text data by converting all text to lowercase and removing punctuation.
    - **Hint**: Use the `str.lower` and `str.replace` functions in Pandas.

Practical Questions on Data Visualization

7. **Question**: Implement a Python script to create a line plot for a given dataset. Customize the plot by adding titles, labels, and a legend.
    - **Hint**: Use the `plot` function in Pandas and `title`, `xlabel`, `ylabel`, and `legend` functions in Matplotlib.

8. **Question**: Write a Python script to create a bar plot showing the distribution of a categorical variable in a dataset. Customize the plot with different colors and labels.
    - **Hint**: Use the `bar` function in Pandas.
9. **Question**: Implement a Python script to create a histogram and density plot for a numeric column in a dataset. Compare the results by plotting both on the same graph.
    - **Hint**: Use the `hist` and `kde` functions in Pandas.
10. **Question**: Write a Python script to create scatter plots for pairs of numeric variables in a dataset. Customize the plot by adding regression lines.
    - **Hint**: Use the `scatter` function in Pandas and `regplot` function in Seaborn.
11. **Question**: Implement a Python script to create an animated line plot showing the trend of a time series data over time.
    - **Hint**: Use the `FuncAnimation` class in Matplotlib.
12. **Question**: Write a Python script to create an interactive plot using Plotly that allows the user to zoom in and out on the data points.
    - **Hint**: Use the Plotly library.

Practical Questions on Exploratory Data Analysis (EDA)

13. **Question**: Implement a Python script to calculate and display descriptive statistics (mean, median, mode, standard deviation, variance) for numeric columns in a dataset.
    - **Hint**: Use the `describe` function in Pandas and `mode` function in SciPy.
14. **Question**: Write a Python script to create boxplots for numeric columns in a dataset. Identify and annotate any outliers on the plot.
    - **Hint**: Use the `boxplot` function in Pandas and `annotate` function in Matplotlib.

15. **Question**: Implement a Python script to create a contingency table for two categorical variables in a dataset and perform a chi-square test of independence.
    - **Hint**: Use the `crosstab` function in Pandas and `chi2_contingency` function in SciPy.
16. **Question**: Write a Python script to create parallel coordinates plots for visualizing multivariate data. Customize the plot with different colors for each class.
    - **Hint**: Use the `parallel_coordinates` function in Pandas plotting.

Practical Questions on Data Sources in Artificial Intelligence (AI)

17. **Question**: Implement a Python script to collect data from an online API and store it in a Pandas DataFrame. Process the data by converting it to a suitable format for analysis.
    - **Hint**: Use the `requests` library to fetch data from the API.
18. **Question**: Write a Python script to read and process unstructured text data from a file. Perform basic text preprocessing steps such as tokenization and stopword removal.
    - **Hint**: Use the NLTK library.
19. **Question**: Implement a Python script to download and process images from a URL. Perform basic image preprocessing steps such as resizing and conversion to grayscale.
    - **Hint**: Use the `requests` library and PIL (Pillow) library.
20. **Question**: Write a Python script to read and process historical data from a CSV file. Perform basic time series analysis, including plotting the data and decomposing it into trend, seasonal, and residual components.
    - **Hint**: Use the `seasonal_decompose` function from the Statsmodels library.