Python Pandas Practice Questions

1. Load a dataset from a CSV file and display the first 5 rows.

2. Load a dataset from an Excel file and display the last 10 rows.

3. Check the number of rows and columns in the dataset using an appropriate function.

4. Display the column names of the dataset.

5. Retrieve the data type of each column and check for missing values.

6. Display the summary statistics of the numerical columns in the dataset.

7. Extract only the first 3 columns of the dataset using loc.

8. Extract the 5th to 10th rows from the dataset using iloc.

9. Select all rows where a specific column (e.g., 'Price') has a value greater than 1000.

10. Retrieve all rows where a specific column contains missing values.

11. Find the minimum and maximum values for a numerical column.

12. Calculate the sum and mean of a particular numerical column.

13. Determine the count, median, and mode of a selected column.

14. Use the agg() function to compute the sum, mean, and standard deviation of a numerical column in one step.

15. Identify the column with the highest sum in the dataset.

16. Identify all missing values in the dataset.

17. Remove all rows with missing values and check the shape of the dataset before and after.

18. Fill missing values in a specific column with the column's mean.

19. Replace all NaN values in a dataset with a fixed value (e.g., 0).

20. Drop duplicate rows from the dataset and verify the shape before and after.

21. Sort the dataset by a numerical column in descending order.

22. Sort the dataset by multiple columns (e.g., first by 'Category' and then by 'Price').

23. Group the dataset by a categorical column and find the mean of another column for each group.

24. Group the dataset by two categorical columns and find the sum of a numerical column.

25. Find the count of unique values in a specific column using groupby().

26. Replace all occurrences of a specific value (e.g., replace 'Male' with 'M' in a gender column).

27. Rename the column names to be more readable (e.g., 'emp_id' to 'Employee ID').

28. Change the index of the dataset to a specific column and reset it back.

29. Extract all rows where a specific column's value is within a given range (e.g., 'Age' between 20 and 30).

30. Drop a specific column from the dataset.