Question 1: What is Pandas, and why is it commonly used in data cleaning tasks?

Ans=[Pandas is a popular Python library that provides fast and flexible data structure’s for the data analysis and data manipulation](https://realpython.com/python-data-cleaning-numpy-pandas/). It is mainly used for data cleaning task as it offers various function to handle missing and duplicate values, and to sort messy data.

Question 2; Given a DataFrame with missing values, how would you check for missing values in each column and count the total number of missing values?

Ans= to check for the missing values in each column of a DataFrame and to count the total number of missing values is to use the isnull() and sum() methods of Pandas. For ex data.isnull().sum()

Question 3; How can you remove duplicates from a DataFrame while retaining the first occurrence of each unique row?

Ans= [To remove the duplicates from the DataFrame while keepingg the first occurrence of each unique row is to use the drop\_duplicates() method.](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.drop_duplicates.html)

Question 4: If you have a DataFrame with a column containing string values, how can you convert all the values in that column to lowercase?

Ans= To convert all the string values in a column to lowercase is to use the .str.lower() method of Pandas.

Question 5: How do you replace missing values in a DataFrame with a specific value, like 0, for a particular column?

Ans= to replace missing values in a DataFrame with a specific value, like 0 we can use fillna() function of panda.

Question 6: If you have a DataFrame with a datetime column, how can you extract the year, month, and day into separate columns?

Ans= to get the year, month, and day into separate columns from the DataFrame with a datetime column is to use the .dt function of pandas

● Question 7: How can you filter rows in a DataFrame where a specific column's values meet a certain condition (e.g., all rows where 'age' is greater than 30)?

Ans= one can use the .loc function of Pandas with a boolean expression To filter rows in a DataFrame where specific column’s values meet a certain conditions. For example, if df is a DataFrame with a column named age, then the following code will select all the rows where age is greater than 30.

Question 8: What is the purpose of the .apply() function in Pandas, and how would you use it to create a new column based on values from existing columns?

Ans= The purpose of the .apply() function is to apply a function on a particular axis of the DataFrame. The function can be a user-defined function or a built-in function. [The .apply() function can be used to create a new column based on values from existing columns by passing the function and the axis as arguments](https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.apply.html).

Question 9: Suppose you want to merge two DataFrames, 'df1' and 'df2,' on a common column 'key.' How would you perform this merge operation in Pandas

Ans= To merge two DataFrames, ‘df1’ and ‘df2,’ on a common column ‘key,’ you can use the merge() function of Pandas.

● Question 10: You have a DataFrame with a column containing messy text data. How can you clean and standardize the text data (e.g., remove punctuation and convert to lowercase) in that column

Ans= To clean the text data in a column of a DataFrame, we can use the .str function of the Pandas, which allows us to apply various string methods to the column elements.if df is a DataFrame with a column named text, then the following code will remove punctuation and convert to lowercase the values in that column.