**PandasAssignment**

Q1. How do you load a CSV file into a Pandas DataFrame?

Ans1. Importing a CSV file into the DataFrame. Pandas read\_csv() function imports a CSV file to DataFrame format.

Q2. How do you check the data type of a column in a Pandas DataFrame?

Ans2. To check the data type in pandas DataFrame we can use the “dtype” attribute.

Q3. How do you select rows from a Pandas DataFrame based on a condition?

Ans3. Methods for Selecting Rows of the DataFrame by Condition · DataFrame.query() method. · DataFrame.loc() method. · DataFrame.isin() method. · DataFrame.filter() ...

Q4. How do you rename columns in a Pandas DataFrame?

Ans4. One way of **renaming** the **columns in a Pandas Dataframe** is by using the **rename**() function.

Q5. How do you drop columns in a Pandas DataFrame?

Ans5. The .**drop**() method is a built-in function in **Pandas** that allows you to remove one or more rows or **columns** from a **DataFrame**.

Q6. How do you find the unique values in a column of a Pandas DataFrame?

Ans6. The easiest way to obtain a list of **unique values** in a **pandas DataFrame column** is to use the unique() function.

Q7. How do you find the number of missing values in each column of a Pandas DataFrame?

Ans7. We can use the following syntax to count NaN values in Pandas DataFrame:

1. (1) Count NaN values under a single DataFrame column: df['column name'].isna().sum()
2. (2) Count NaN values under an entire DataFrame: df.isna().sum().sum()
3. (3) Count NaN values across a single DataFrame row: df.loc[[index value]].isna().sum().sum()

Q8. How do you fill missing values in a Pandas DataFrame with a specific value?

Ans8. Filling missing values using fillna(), replace() and interpolate() In order to fill null values in a datasets, we use fillna(), replace() and interpolate() function these function replace NaN values with some value of their own. All these function help in filling a null values in datasets of a DataFrame.

Q9. How do you concatenate two Pandas DataFrames?

Ans9. Use **pandas**.**concat**() to **concatenate**/merge **two** or multiple **pandas DataFrames** across rows or columns.

Q10. How do you merge two Pandas DataFrames on a specific column?

Ans10. We can **merge two Pandas DataFrames** on **certain columns** using the **merge** function by simply specifying the **certain columns** for **merge**.

Q11. How do you group data in a Pandas DataFrame by a specific column and apply an aggregation function?

Ans11. Use **pandas DataFrame**.**aggregate**() **function** to calculate any aggregations on the selected **columns** of DataFrame and **apply** multiple aggregations ...

Q12. How do you pivot a Pandas DataFrame?

Ans12. Create Your Own Pandas Pivot Table in 3 Steps

1. Download or import the data that you want to use.
2. In the pivot\_table function, specify the DataFrame you are summarizing, along with the names for the indexes, columns and values.
3. Specify the type of calculation you want to use, such as the mean.

Q13. How do you change the data type of a column in a Pandas DataFrame?

Ans13. **Change column type** into string object using **DataFrame**.astype() ... **DataFrame**.astype() method is used to cast **pandas** object to a specified dtype.

Q14. How do you sort a Pandas DataFrame by a specific column?

Ans14. To sort the DataFrame based on the values in a single column, you'll use . sort\_values() . By default, this will return a new DataFrame sorted in ascending order. It does not modify the original DataFrame.

Q15. How do you create a copy of a Pandas DataFrame?

Ans15. The copy() method returns a copy of the DataFrame. By default, the copy is a "deep copy" meaning that any changes made in the original DataFrame will NOT be reflected in the copy.

Q16. How do you filter rows of a Pandas DataFrame by multiple conditions?

Ans16. To **filter Pandas DataFrame by multiple conditions** use DataFrame.loc[] property along with the **conditions**.

Q17. How do you calculate the mean of a column in a Pandas DataFrame?

Ans17. To calculate the mean of whole columns in the DataFrame, use pandas.Series.mean() with a list of DataFrame columns.

Q18. How do you calculate the standard deviation of a column in a Pandas DataFrame?

Ans18. simply append the std() method to the DataFrame object. It returns a Series object with the standard deviation of each column.

Q19. How do you calculate the correlation between two columns in a Pandas DataFrame?

Ans19. Initialize two variables, col1 and col2, and assign them the columns that you want to find the correlation of. Find the correlation between col1 and col2 by using df[col1]. corr(df[col2]) and save the correlation value in a variable, corr. Print the correlation value, corr.

Q20. How do you select specific columns in a DataFrame using their labels?

Ans20. To access **specific columns** of a Pandas **DataFrame** with **their columns labels**, directly use **DataFrame**[~] or use the **DataFrame**.loc property.

Q21. How do you select specific rows in a DataFrame using their indexes?

Ans21. You can select a single row from pandas DataFrame by integer index using df. iloc[n] . Replace n with a position you wanted to select.

Q22. How do you sort a DataFrame by a specific column?

Ans22. Sorting Your DataFrame on a Single Column. To sort the DataFrame based on the values in a single column, you'll use . sort\_values() . By default, this will return a new DataFrame sorted in ascending order.

Q23. How do you create a new column in a DataFrame based on the values of another column?

Ans23. You can add/append a new column to the DataFrame based on the values of another column using df. assign() , df. apply() , and, np. where() functions and return a new Dataframe after adding a new column.

Q24. How do you remove duplicates from a DataFrame?

Ans24. The drop\_duplicates() method removes duplicate rows. Use the subset parameter if only some specified columns should be considered when looking for duplicates.

Q25. What is the difference between .loc and .iloc in Pandas?

Ans25. The function . loc is typically used for label indexing and can access multiple columns, while . iloc is used for integer indexing.