PandasAssignment

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

Ans- By using read\_csv function eg. pd.read\_csv(“filepath”)

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

Ans- we can check the data type by using “dtype”

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

Ans-

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

Ans- We can rename columns by using rename() function

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

Ans- we can drop column by using drop()

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

Ans-we can find unique values in column by using unique()

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

Ans- df.isnull().sum()

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

Ans- df.fillna(value)

Q9. How do you concatenate two Pandas DataFrames?

Ans- pd.concat([df1,df2])

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

Ans-pd.merge(df1,df2,on=”id”)—merge is based on common column

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

Ans- we can group data by using groupby() then we can use aggrate function on that such as count(),sum()

Q12. How do you pivot a Pandas DataFrame?

Ans- we can pivot by using pivot()

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

Ans- By using df.astype(str)

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

Ans-df.sortvalues(by = ‘columnname’)

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

Ans-by using df.copy()

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

Ans-by using loc()

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

Ans-df.mean()

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

Ans-df.std()

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

Ans-df.corr()

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

Ans-result=df.[“column\_name”] or we can use loc() eg.df.loc[“columname”]

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

Ans-result=df.iloc[2]

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

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

Ans- We can use apply() method to create new column

Q24. How do you remove duplicates from a DataFrame?

Ans-df.drop\_duplicates()

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

Ans-.loc - The loc() function is label based data selecting method which means that we have to pass the name of the row or column which we want to select.

.iloc - The iloc() function is an indexed-based selecting method which means that we have to pass an integer index in the method to select a specific row/column.