Contents

[1. What is pandas? 4](#_Toc171156448)

[2. What is a DataFrame in pandas? 4](#_Toc171156449)

[3. What is a Series in pandas? 4](#_Toc171156450)

[4. How do you read a CSV file into a pandas DataFrame? 4](#_Toc171156451)

[5. How do you write a DataFrame to a CSV file? 4](#_Toc171156452)

[6. How do you handle missing data in a pandas DataFrame? 4](#_Toc171156453)

[7. How do you filter rows in a DataFrame based on a condition? 4](#_Toc171156454)

[8. How do you select a subset of columns in a DataFrame? 4](#_Toc171156455)

[9. What is the difference between loc and iloc in pandas? 5](#_Toc171156456)

[10. How do you add a new column to a DataFrame? 5](#_Toc171156457)

[11. How do you remove a column from a DataFrame? 5](#_Toc171156458)

[12. How do you rename columns in a DataFrame? 5](#_Toc171156459)

[13. What is the purpose of the groupby method? 5](#_Toc171156460)

[14. How do you concatenate two DataFrames? 5](#_Toc171156461)

[15. How do you merge two DataFrames? 5](#_Toc171156462)

[16. What is the difference between merge and join in pandas? 5](#_Toc171156463)

[17. How do you sort a DataFrame by a specific column? 6](#_Toc171156464)

[18. How do you apply a function to each element in a DataFrame? 6](#_Toc171156465)

[19. How do you apply a function to each column or row in a DataFrame? 6](#_Toc171156466)

[20. How do you find the unique values in a column? 6](#_Toc171156467)

[21. How do you get the number of unique values in a column? 6](#_Toc171156468)

[22. How do you check for duplicate rows in a DataFrame? 6](#_Toc171156469)

[23. How do you remove duplicate rows in a DataFrame? 6](#_Toc171156470)

[24. What is the purpose of the pivot\_table function? 6](#_Toc171156471)

[25. How do you reshape a DataFrame using melt? 6](#_Toc171156472)

[26. What is the purpose of the crosstab function? 7](#_Toc171156473)

[27. How do you change the index of a DataFrame? 7](#_Toc171156474)

[28. How do you reset the index of a DataFrame? 7](#_Toc171156475)

[29. How do you get the summary statistics of a DataFrame? 7](#_Toc171156476)

[30. How do you handle large datasets in pandas? 7](#_Toc171156477)

[31. How do you perform date and time operations in pandas? 7](#_Toc171156478)

[32. How do you convert a column to a categorical type? 7](#_Toc171156479)

[33. How do you create a DataFrame from a dictionary? 7](#_Toc171156480)

[34. How do you create a DataFrame from a list of dictionaries? 8](#_Toc171156481)

[35. What are the advantages of using pandas over other data manipulation tools? 8](#_Toc171156482)

[36. How do you perform data cleaning in pandas? 8](#_Toc171156483)

[37. How do you plot data from a pandas DataFrame? 8](#_Toc171156484)

[38. How do you handle string data in pandas? 8](#_Toc171156485)

[39. How do you handle categorical data in pandas? 8](#_Toc171156486)

[40. What is the pd.read\_html() function used for? 8](#_Toc171156487)

[41. How do you convert a DataFrame to a NumPy array? 8](#_Toc171156488)

[42. How do you find the correlation between columns in a DataFrame? 9](#_Toc171156489)

[43. How do you create a DataFrame with a MultiIndex? 9](#_Toc171156490)

[44. How do you stack and unstack a DataFrame? 9](#_Toc171156491)

[45. How do you perform element-wise operations on a DataFrame? 9](#_Toc171156492)

[46. What is the difference between map, apply, and applymap? 9](#_Toc171156493)

[47. How do you create a pivot table in pandas? 9](#_Toc171156494)

[48. How do you handle duplicates in a DataFrame? 9](#_Toc171156495)

[49. How do you filter a DataFrame using the query method? 9](#_Toc171156496)

[50. How do you calculate the rolling mean of a column in a DataFrame? 10](#_Toc171156497)

[51. How do you shift the values in a DataFrame? 10](#_Toc171156498)

[52. How do you fill missing values with the mean of a column? 10](#_Toc171156499)

[53. How do you calculate the cumulative sum of a column? 10](#_Toc171156500)

[54. How do you drop rows with any missing values? 10](#_Toc171156501)

[55. How do you calculate the percentile rank of values in a column? 10](#_Toc171156502)

[56. How do you filter a DataFrame based on the values in a list? 10](#_Toc171156503)

[57. How do you concatenate DataFrames along the columns? 10](#_Toc171156504)

[58. How do you convert a DataFrame to a dictionary? 10](#_Toc171156505)

[59. How do you read data from an Excel file into a DataFrame? 11](#_Toc171156506)

[60. How do you write a DataFrame to an Excel file? 11](#_Toc171156507)

## What is pandas?

Answer: Pandas is an open-source data analysis and manipulation library for Python. It provides data structures like DataFrame and Series to efficiently handle and analyze structured data.

## What is a DataFrame in pandas?

Answer: A DataFrame is a two-dimensional, size-mutable, and heterogeneous tabular data structure with labeled axes (rows and columns).

## What is a Series in pandas?

Answer: A Series is a one-dimensional array-like object containing an array of data and an associated array of data labels, called its index.

## How do you read a CSV file into a pandas DataFrame?

Answer: You can use the pd.read\_csv('file\_path') function to read a CSV file into a pandas DataFrame.

## How do you write a DataFrame to a CSV file?

Answer: You can use the DataFrame.to\_csv('file\_path') method to write a DataFrame to a CSV file.

## How do you handle missing data in a pandas DataFrame?

Answer: You can use methods like dropna() to remove missing values and fillna() to fill missing values with a specified value.

## How do you filter rows in a DataFrame based on a condition?

Answer: You can use boolean indexing, e.g., df[df['column'] > value], to filter rows based on a condition.

## How do you select a subset of columns in a DataFrame?

Answer: You can use df[['column1', 'column2']] to select a subset of columns in a DataFrame.

## What is the difference between loc and iloc in pandas?

Answer: loc is label-based indexing, while iloc is integer position-based indexing.

## How do you add a new column to a DataFrame?

Answer: You can add a new column by assigning a value to a new column name, e.g., df['new\_column'] = value.

## How do you remove a column from a DataFrame?

Answer: You can use the drop('column\_name', axis=1) method to remove a column from a DataFrame.

## How do you rename columns in a DataFrame?

Answer: You can use the rename(columns={'old\_name': 'new\_name'}) method to rename columns.

## What is the purpose of the groupby method?

Answer: The groupby method is used to split the data into groups based on some criteria and apply a function to each group independently.

## How do you concatenate two DataFrames?

Answer: You can use the pd.concat([df1, df2]) function to concatenate two DataFrames.

## How do you merge two DataFrames?

Answer: You can use the pd.merge(df1, df2, on='key') function to merge two DataFrames based on a key column.

## What is the difference between merge and join in pandas?

Answer: merge is used to combine DataFrames based on a common column(s), while join is primarily used to join on the index.

## How do you sort a DataFrame by a specific column?

Answer: You can use the sort\_values('column\_name') method to sort a DataFrame by a specific column.

## How do you apply a function to each element in a DataFrame?

Answer: You can use the applymap(function) method to apply a function to each element in a DataFrame.

## How do you apply a function to each column or row in a DataFrame?

Answer: You can use the apply(function, axis=0) method to apply a function to each column and apply(function, axis=1) to apply a function to each row.

## How do you find the unique values in a column?

Answer: You can use the unique() method, e.g., df['column\_name'].unique().

## How do you get the number of unique values in a column?

Answer: You can use the nunique() method, e.g., df['column\_name'].nunique().

## How do you check for duplicate rows in a DataFrame?

Answer: You can use the duplicated() method to check for duplicate rows.

## How do you remove duplicate rows in a DataFrame?

Answer: You can use the drop\_duplicates() method to remove duplicate rows.

## What is the purpose of the pivot\_table function?

Answer: The pivot\_table function is used to create a spreadsheet-style pivot table as a DataFrame.

## How do you reshape a DataFrame using melt?

Answer: You can use the melt(id\_vars, value\_vars) function to unpivot a DataFrame from wide to long format.

## What is the purpose of the crosstab function?

Answer: The crosstab function is used to compute a simple cross-tabulation of two or more factors.

## How do you change the index of a DataFrame?

Answer: You can use the set\_index('column\_name') method to change the index of a DataFrame.

## How do you reset the index of a DataFrame?

Answer: You can use the reset\_index() method to reset the index of a DataFrame.

## How do you get the summary statistics of a DataFrame?

Answer: You can use the describe() method to get the summary statistics of a DataFrame.

## How do you handle large datasets in pandas?

Answer: You can handle large datasets by using chunking, optimizing data types, and leveraging Dask or Vaex for out-of-core computations.

## How do you perform date and time operations in pandas?

Answer: You can use the pd.to\_datetime() function and the dt accessor to perform date and time operations.

## How do you convert a column to a categorical type?

Answer: You can use the astype('category') method to convert a column to a categorical type.

## How do you create a DataFrame from a dictionary?

Answer: You can use the pd.DataFrame(dictionary) function to create a DataFrame from a dictionary.

## How do you create a DataFrame from a list of dictionaries?

Answer: You can use the pd.DataFrame(list\_of\_dicts) function to create a DataFrame from a list of dictionaries.

## What are the advantages of using pandas over other data manipulation tools?

Answer: Pandas provides intuitive and flexible data structures, powerful data manipulation capabilities, integration with other libraries, and efficient handling of missing data.

## How do you perform data cleaning in pandas?

Answer: Data cleaning in pandas can be performed using methods like dropna(), fillna(), replace(), astype(), and string methods for text data.

## How do you plot data from a pandas DataFrame?

Answer: You can use the plot() method for basic plotting or leverage libraries like Matplotlib and Seaborn for more advanced visualizations.

## How do you handle string data in pandas?

Answer: You can use the str accessor with methods like str.contains(), str.replace(), str.extract(), etc., to handle string data in pandas.

## How do you handle categorical data in pandas?

Answer: You can use the astype('category') method and the pd.get\_dummies() function for handling categorical data.

## What is the pd.read\_html() function used for?

Answer: The pd.read\_html() function is used to read HTML tables into a list of DataFrames.

## How do you convert a DataFrame to a NumPy array?

Answer: You can use the to\_numpy() method to convert a DataFrame to a NumPy array.

## How do you find the correlation between columns in a DataFrame?

Answer: You can use the corr() method to find the correlation between columns in a DataFrame.

## How do you create a DataFrame with a MultiIndex?

Answer: You can create a DataFrame with a MultiIndex using the set\_index([list\_of\_columns]) method or directly with the pd.MultiIndex.from\_arrays() method.

## How do you stack and unstack a DataFrame?

Answer: You can use the stack() method to pivot the columns into rows and the unstack() method to pivot the rows into columns.

## How do you perform element-wise operations on a DataFrame?

Answer: You can use the applymap() method to perform element-wise operations on a DataFrame.

## What is the difference between map, apply, and applymap?

Answer: map is used for element-wise transformations in a Series, apply is used for row/column-wise transformations in a DataFrame, and applymap is used for element-wise transformations in a DataFrame.

## How do you create a pivot table in pandas?

Answer: You can use the pivot\_table(index, columns, values, aggfunc) method to create a pivot table.

## How do you handle duplicates in a DataFrame?

Answer: You can use the duplicated() method to find duplicates and the drop\_duplicates() method to remove them.

## How do you filter a DataFrame using the query method?

Answer: You can use the query('expression') method to filter a DataFrame based on an expression.

## How do you calculate the rolling mean of a column in a DataFrame?

Answer: You can use the rolling(window).mean() method to calculate the rolling mean.

## How do you shift the values in a DataFrame?

Answer: You can use the shift(periods) method to shift the values in a DataFrame.

## How do you fill missing values with the mean of a column?

Answer: You can use the fillna(df['column'].mean()) method to fill missing values with the mean of a column.

## How do you calculate the cumulative sum of a column?

Answer: You can use the cumsum() method to calculate the cumulative sum of a column.

## How do you drop rows with any missing values?

Answer: You can use the dropna() method to drop rows with any missing values.

## How do you calculate the percentile rank of values in a column?

Answer: You can use the rank(pct=True) method to calculate the percentile rank of values in a column.

## How do you filter a DataFrame based on the values in a list?

Answer: You can use the isin(list\_of\_values) method to filter a DataFrame based on the values in a list.

## How do you concatenate DataFrames along the columns?

Answer: You can use the pd.concat([df1, df2], axis=1) function to concatenate DataFrames along the columns.

## How do you convert a DataFrame to a dictionary?

Answer: You can use the to\_dict() method to convert a DataFrame to a dictionary.

## How do you read data from an Excel file into a DataFrame?

Answer: You can use the pd.read\_excel('file\_path') function to read data from an Excel file into a DataFrame.

## How do you write a DataFrame to an Excel file?

Answer: You can use the to\_excel('file\_path') method to write a DataFrame to an Excel file.