Interview Questions task5:

1.What is Pandas used for?

Pandas is a **Python library** used for **data manipulation and analysis**.
It provides powerful tools to **read, clean, transform, and analyze structured data** (like CSV, Excel, SQL, etc.).

2.What's a DataFrame?

A **DataFrame** is a **2D labeled data structure** (like a table in Excel or SQL).
It has **rows and columns**, where each column can hold different data types (int, float, string, etc.).

3.How do you read a CSV file?

import pandas as pd

df = pd.read_csv('filename.csv')import pandas as pd

df = pd.read_csv('filename.csv')

4.What is groupby()?

he `groupby()` function is used to **group data based on one or more columns**, then perform **aggregations** like `sum()`, `mean()`, `count()`, etc.

df.groupby('Category')['Sales'].sum()

5.How do you filter rows?

You can filter rows using **conditional selection**:

df[df['Sales'] > 1000]

This returns only the rows where the `Sales` column is greater than 1000.

6.Difference between loc[] and iloc[]?

Feature	`loc[]`	`iloc[]`
Access by	Labels (row/column names)	Integer positions (index numbers)
Example	`df.loc[2, 'Name']`	`df.iloc[2, 1]`

7.What does .head() do?

The `.head()` method returns the **first 5 rows** of the DataFrame by default.

df.head()

You can also specify the number of rows: `df.head(10)`

8.How can you create a bar chart?

You can use the built-in plotting:

df['Sales'].plot(kind='bar')

or with Matplotlib:

import matplotlib.pyplot as plt

df.plot.bar(x='Category', y='Sales')

plt.show()

9.What’s the shape of a DataFrame?

he `df.shape` attribute returns a **tuple (rows, columns)**:

df.shape

Example output: (100, 5)

10.What is NaN?

NaN stands for **"Not a Number"**, and it represents **missing or undefined values** in a dataset.