**Pandas Interview Questions & Answers**

**1.What is Pandas used for?**

**Answer:**  
Pandas is a **Python library for data analysis and manipulation**.  
It helps handle **structured data** (like tables) easily — you can **clean, filter, analyze, and visualize** datasets efficiently using Pandas.

**2. What’s a DataFrame?**

**Answer:**  
A **DataFrame** is a **2-dimensional labeled data structure** (like an Excel spreadsheet).  
It consists of **rows and columns**, where each column can hold different data types (integers, strings, floats, etc.).  
Example:

import pandas as pd

data = {'Name': ['A', 'B'], 'Age': [20, 25]}

df = pd.DataFrame(data)

**3.How do you read a CSV file?**

**Answer:**  
You can load a CSV file into a DataFrame using:

df = pd.read\_csv('filename.csv')

You can also specify separators, headers, or encoding if needed.

**4. What is groupby()?**

**Answer:**  
groupby() is used to **split data into groups**, perform **aggregations**, and **combine** the results.  
Example:

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

→ This groups data by Product and calculates the total Sales for each group.

**5.How do you filter rows?**

**Answer:**  
You can filter rows using **conditions**.  
Example:

df[df['Sales'] > 1000]

→ Returns all rows where the Sales value is greater than 1000.

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

**Answer:**

| **Function** | **Description** | **Example** |
| --- | --- | --- |
| loc[] | Access rows/columns **by labels (names)** | df.loc[0:5, ['Product', 'Sales']] |
| iloc[] | Access rows/columns **by integer positions** | df.iloc[0:5, [1, 4]] |

**7. What does .head() do?**

**Answer:**  
.head() displays the **first few rows** (default 5) of the DataFrame.  
Example:

df.head()

→ Quickly checks the top part of your dataset.

**8. How can you create a bar chart?**

**Answer:**  
You can use Matplotlib directly or via Pandas built-in plotting:

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

plt.show()

This creates a simple bar chart showing Sales values.

**9.What’s the shape of a DataFrame?**

**Answer:**  
The .shape attribute returns a **tuple** showing (number of rows, number of columns).  
Example:

df.shape

If output is (100, 5) → the DataFrame has 100 rows and 5 columns.

**10. What is NaN?**

**Answer:**  
NaN stands for **“Not a Number”**.  
It represents **missing or undefined data** in Pandas.  
Example:

df.isnull().sum()

→ Shows how many NaN values are present in each column.