Dataset used in all questions:

Name	Age	City	Salary	Department
Amit	25	Delhi	50000	IT
Riya	30	Mumbai	85000	HR
Kunal	22	Noida	40000	IT
Sara	28	Pune	62000	Finance
John	35	Delhi	90000	IT
Meena	27	Mumbai	70000	HR

⭐ **SLIDE 1 — Introduction to DataFrames**

Q1: What is a DataFrame in pandas?

A: A DataFrame is a 2D table (rows + columns) used to store structured data.

Q2: How do you load a DataFrame?

A:

```
import pandas as pd

df = pd.read_csv("data.csv")
```

Q3: How do you view the first 5 rows?

A:

```
df.head()
```

Q4: How do you check the structure of the DataFrame?

A:

```
df.info()
```

Q5: How do you describe numerical columns?

A:

```
df.describe()
```

⭐ **SLIDE 2 — Viewing, Selecting, Filtering Rows**

Q1: How do you select only the Name and Salary columns?

A:

df[['Name', 'Salary']]

Q2: How do you filter employees with salary > 60,000?

A:

df[df['Salary'] > 60000]

Q3: How do you filter HR department employees?

A:

df[df['Department'] == "HR"]

Q4: How do you filter employees from Mumbai?

A:

df[df['City'] == "Mumbai"]

Q5: How do you apply multiple conditions (salary > 60K AND city = Mumbai)?

A:

df[(df['Salary'] > 60000) & (df['City'] == "Mumbai")]

⭐ **SLIDE 3 — Sorting Data**

Q1: How do you sort employees by salary (descending)?

A:

df.sort_values("Salary", ascending=False)

Q2: How do you sort by Age (ascending)?

A:

df.sort_values("Age")

Q3: How do you sort by Department, then by Salary?

A:

df.sort_values(["Department", "Salary"])

⭐ **SLIDE 4 — Handling Missing Data**

Q1: How do you check how many missing values each column has?

A:

df.isnull().sum()

Q2: How do you remove rows with missing values?

A:

df.dropna()

Q3: How do you replace missing salary values with 0?

A:

df['Salary'].fillna(0)

Q4: How do you fill missing cities with “Unknown”?

A:

df['City'].fillna("Unknown")

⭐ **SLIDE 5 — Handling Duplicate Data**

Q1: How do you check if there are duplicate rows?

A:

df.duplicated()

Q2: How do you remove duplicate rows?

A:

df.drop_duplicates()

⭐ **SLIDE 6 — Grouping Data**

Q1: How do you find the average salary by department?

A:

df.groupby('Department')['Salary'].mean()

Q2: How many employees are there in each city?

A:

df.groupby('City')['Name'].count()

Q3: How do you calculate min, max, mean salary by city?

A:

df.groupby('City')['Salary'].agg(['min', 'max', 'mean'])

⭐ **SLIDE 7 — Merging DataFrames**

(Example assumes df_dept contains department descriptions)

Q1: How do you merge employee info with department info?

A:

pd.merge(df, df_dept, on="Department")

Q2: What join types are available in pandas?

A:
A: inner, left, right, outer

⭐ **SLIDE 8 — Reshaping Data**

Q1: How do you convert rows into columns using pivot?

A:

df.pivot(index='City', columns='Department', values='Salary')

Q2: How do you convert wide data to long format?

A:

df.melt()

⭐ **SLIDE 9 — Creating New Columns**

Q1: How do you create a new column Bonus = 10% of Salary?

A:

df['Bonus'] = df['Salary'] * 0.10

Q2: How do you categorize employees as "Senior" if Age > 30 using lambda?

A:

df['Category'] = df['Age'].apply(lambda x: "Senior" if x > 30 else "Junior")

Q3: How do you calculate Total Compensation = Salary + Bonus?

A:

df['TotalComp'] = df['Salary'] + df['Bonus']

⭐ **SLIDE 10 — Descriptive Statistics**

Q1: How do you calculate the mean salary?

A:

df['Salary'].mean()

Q2: How do you calculate the median age?

A:

df['Age'].median()

Q3: How do you calculate the standard deviation of salary?

A:

df['Salary'].std()

Q4: How do you find correlation between Age and Salary?

A:

df[['Age', 'Salary']].corr()
