

## TASKS TO PERFORMED IN PANDAS

### ◆ 1. Load and Inspect

- Load the CSV file using `pd.read_csv()`.
- Display the first 5 rows.
- Show column names and data types.
- Check for null values.

### ◆ 2. Data Cleaning

- Fill missing values in Doctor Name with 'Unknown'.
- Convert Admission Date and Discharge Date to datetime.
- Fill values where Length of Stay is missing
- Remove rows Treatment Cost is missing.

### ◆ 3. Data Analysis

- Find the average Length of Stay per department.
- Calculate total and average treatment cost by gender.
- Find top 3 most common diagnoses.
- Count patients admitted in each month.

### ◆ 4. Grouping and Aggregation

- Group by Department and calculate:
  - Total cost of treatments
- Group by Doctor Name and count how many patients each treated.

### ◆ 5. Sorting and Filtering

- List patients whose Treatment Cost is above the 90th percentile.
- Sort patients by Length of Stay in descending order.

### ◆ 6. Derived Columns

- Create a new column: `Out-of-Pocket = Treatment Cost - Insurance Covered`
- Create another column: Age Group (e.g., <18, 18-35, 36-60, 60+)

### ◆ 7. Save Cleaned Data

- Save your final cleaned and processed DataFrame to a new file: `cleaned_hospital_records.csv`