✅ PANDAS — FULL PRACTICAL MASTERY TEST (100% Hands-On Data Analyst Version)

Total Tasks: 40 (Every task requires actual coding & real-world thinking)

After completing this set, you will be *job-ready* for Pandas as a Data Analyst.

🔹 SECTION 1 — DATA LOADING & DATAFRAME BASICS (8 Practical Tasks)

T1. Load a CSV file named `employees.csv` into a DataFrame and display the top 12 rows.

T2. Display:

- row count
- column count
- data types
- missing value count per column

T3. Convert column `JoinDate` to datetime format.

T4. Convert column `Department` to category dtype.

T5. Keep only columns: `Name`, `Age`, `Department`, `Salary`.

T6. Remove duplicate rows based on `EmployeeID`.

T7. Replace all NaN values in `Salary` with the column's median.

T8. Create a new column `SalaryLakh` = Salary / 100000.

🔹 SECTION 2 — FILTERING, INDEXING & SELECTION (8 Practical Tasks)

T9. Select all employees with Salary > ₹80,000.

T10. Select employees in `Finance` OR `IT` department using `.isin()`.

T11. Select rows where Name starts with letter "S".

T12. Extract only columns from "Age" to "Salary" using `.loc`.

T13. Select rows 20 till 35 using `.iloc`.

T14. Add a new column `HighSalary` = True if Salary > 1,00,000 else False.

T15. Remove all employees younger than 25 years old.

T16. Randomly select 10% of DataFrame rows.

□ SECTION 3 — GROUPBY, AGGREGATION & PIVOT TABLES (8 Practical Tasks)

Use a dataset: **sales_data.csv**

T17. Find total revenue per region.

T18. Find: avg revenue, max revenue, min revenue per product category (one groupby).

T19. For each city, calculate: number of orders and total revenue.

T20. Group by `CustomerID` and list top 10 customers by total revenue.

T21. Create a pivot table:

Rows = City, Columns = Category → values = sum of Quantity

T22. Calculate monthly total sales using groupby (not resample).

T23. Sort categories by highest average revenue.

T24. After grouping by Category, convert the output to a DataFrame with proper column names.

□ SECTION 4 — MERGE, JOIN & CONCAT (6 Practical Tasks)

You have two DataFrames: **orders.csv** & **customers.csv**.

T25. Perform INNER JOIN on `CustomerID`.

T26. Perform LEFT JOIN and keep unmatched customers also.

T27. Concatenate two DataFrames vertically.

T28. Concatenate two DataFrames horizontally.

T29. Merge orders & customers on two keys: `City` & `State`.

T30. After merging, drop all duplicate rows based on `OrderID`.

□ SECTION 5 — DATA CLEANING & TRANSFORMATION (6 Practical Tasks)

T31. Identify numeric outliers in column `Amount` using IQR and remove them.

T32. Split column `FullName` into `FirstName` and `LastName`.

T33. Create a new column:

```
Profit = SellingPrice - CostPrice
```

T34. Apply a lambda to clean all text columns by stripping spaces.

T35. Rename multiple columns at once using a dictionary.

T36. Scale all numeric columns using (value - min) / (max - min).

□ SECTION 6 — TIME SERIES OPERATIONS (4 Practical Tasks)

Using a DataFrame with datetime column `OrderDate`:

T37. Extract Year, Month, Week, Day from OrderDate.

T38. Filter all orders placed in year 2023.

T39. Resample sales monthly and compute total revenue.

T40. Calculate days difference between consecutive transactions for each customer.

☆ SECTION 7 — FULL REAL-WORLD ANALYSIS PROJECT

T41. Case Study: E-Commerce Dataset (Master Level)

Dataset: **ecommerce.csv**
Columns: OrderID, CustomerID, Category, SubCategory, Quantity, Amount, Discount, Date, City, State, PaymentMode

Do the following:

1. Load dataset
2. Handle missing values appropriately
3. Remove outliers from `Amount`
4. Create `NetAmount = Amount - Discount`
5. Create city-wise total sales table
6. Create monthly sales trend
7. Create category-wise: total sales, avg sales, total quantity
8. Find top 10 customers by total spend
9. Create pivot table: State vs Category → sum of NetAmount
10. Filter customers spending more than ₹50,000
11. Find most profitable category
12. Identify the month with highest orders
13. Find average discount per category
14. Calculate revenue contribution % of each state
15. Export final cleaned dataset to `output_clean.csv`

This single task makes you completely ready for any Data Analyst job.
