Pandas Interview Assignment for AI & Data Science

Instructions:

- Solve all questions using Pandas.
- Write well-commented code.
- Use efficient Pandas functions where possible.

---

1. Data Selection & Filtering

1. Load a CSV file (`data.csv`) into a Pandas DataFrame and display the first 10 rows.
2. Select all rows where `Salary > 50,000`.
3. Select only the `Name` and `Department` columns for employees in the "IT" department.
4. Find all employees whose names **start with 'A'**.

---

2. Data Cleaning & Transformation

5. Convert the `Date_Of_Joining` column to datetime format.
6. Replace all missing values in the `Salary` column with the **median salary**.
7. Remove all duplicate rows from the dataset.
8. Standardize the `City` column to **lowercase**.

---

3. GroupBy & Aggregation

9. Find the **total salary per department**.
10. Find the **average salary per experience level**.
11. Find the **min and max salary per department**.
12. Find the **total number of employees per department & experience level**.
13. Compute the **standard deviation of salaries per department**.
14. Rank employees within each department by salary (highest salary = Rank 1).
15. Find the **top 3 highest-paid employees** per department.

---

4. Merging & Joining

16. Merge two DataFrames:

- `df_employees` (Employee details)
- `df_departments` (Department info) using the `Department_ID` column.

17. Perform a **left join** on `df1` and `df2` based on the `Employee_ID` column.

---

5. Time Series Handling

18. Extract the **year** from the `Date_Of_Joining` column.
19. Compute the **total sales per month** from a dataset containing daily sales.

---

6. Performance Optimization

20. Convert a DataFrame column from `object` to `category` for memory optimization.
21. Use `apply()` and `map()` functions to efficiently modify a column.
22. Find the column with the **highest number of missing values**.

Submission Guidelines:

- Solve at least **15 questions**.
- Use Jupyter Notebook with **well-commented code**.