Training Day – 38

Topic:* Combining Datasets with Pandas

- Learned to concatenate and merge datasets.

- Example: Merged two datasets on a common key using pd.merge().set.

Combining Multiple Datasets:

- **Concatenation:**
  Combine along rows or columns.

  ```python
  Copy code
  df1 = pd.DataFrame({'A': [1, 2]})
  df2 = pd.DataFrame({'A': [3, 4]})
  combined = pd.concat([df1, df2])
  ```

- **Merging (Join):**
  Combine based on a common key.

  ```python
  Copy code
  df1 = pd.DataFrame({'ID': [1, 2], 'Name': ['Alice', 'Bob']})
  df2 = pd.DataFrame({'ID': [1, 2], 'Score': [85, 90]})
  merged = pd.merge(df1, df2, on='ID')
  ```

- **Cleaning After Merging:**
  Ensure no duplicate or irrelevant columns remain.

  ```
  Copy code
  merged.dropna(inplace=True)
  ```

Topic: Combining Datasets with Pandas

- Learned to concatenate and merge datasets.

- **Example:**

  ```python
  import pandas as pd
  df1 = pd.DataFrame({"ID": [1, 2], "Value": [10, 20]})
  df2 = pd.DataFrame({"ID": [1, 2], "Description": ["A", "B"]}
  merged_df = pd.merge(df1, df2, on="ID")
  print(merged_df)
  ```

Output:

```
   ID  Value Description
0   1     10           A
1   2     20           B
```