```
from google.colab import files
uploaded = files.upload()
```

```
Choose Files deepseek_vs_chatgpt.csv
• deepseek_vs_chatgpt.csv(text/csv) - 2416971 bytes, last modified: 4/6/2025 - 100% done
Saving deepseek_vs_chatgpt.csv to deepseek_vs_chatgpt.csv
```

```
import pandas as pd

df = pd.read_csv("deepseek_vs_chatgpt.csv")
df.columns = df.columns.str.strip()

```

```
#Concatenation
# Split the DataFrame into two parts and concatenate
df1 = df.iloc[:100]
df2 = df.iloc[100:200]

df_concat = pd.concat([df1, df2], axis=0)
print("Concatenated DataFrame shape:", df_concat.shape)
```

```
Concatenated DataFrame shape: (200, 28)
```

```
#Merge
feedback_df = pd.DataFrame({
    "User_ID": df["User_ID"].head(10),
    "Feedback": ["Excellent", "Good", "Average", "Poor", "Great", "Bad", "Average", "Excellent", "Fair", "Good"]
})

merged_df = pd.merge(df, feedback_df, on="User_ID", how="left")
print(merged_df[["User_ID", "AI_Platform", "Feedback"]].head())

```

```
                              User_ID AI_Platform   Feedback
0  c878a177-2da9-4224-8cf8-1d56a1c6a755     ChatGPT  Excellent
1  7096d0f1-d0dc-4333-a5d0-de9dfd1b99fa     ChatGPT       Good
2  e690c254-582f-49c1-89c3-0b4dd8ee59be     ChatGPT    Average
3  0b6a010d-9d03-44c4-bf7f-7f8d2cc461e2     ChatGPT       Poor
4  ffa90616-1fa9-48ff-842d-e84e193c64f4    DeepSeek      Great
```

```
# Right Join Example
right_joined = pd.merge(df, feedback_df, on="User_ID", how="right")
print("Right join result:\n", right_joined)

# Outer Join Example
outer_joined = pd.merge(df, feedback_df, on="User_ID", how="outer")
print("Outer join result:\n", outer_joined)
```

```
9998                0.75                Instant                0
9999                1.98                   Fast                0

      User_Return_Frequency  Customer_Support_Interactions        Region  \
0                         6                              0       Algeria
1                         8                              1         Tonga
2                         1                              1      Kiribati
3                         8                              2      Barbados
4                         8                              0     Indonesia
...                     ...                            ...           ...
9995                      1                              1      Colombia
9996                     10                              0         Spain
9997                      4                              1       Denmark
9998                      5                              3  Saudi Arabia
9999                      3                              1  Saudi Arabia

     Feedback
0         NaN
1         NaN
2         NaN
3         NaN
4         NaN
...       ...
9995      NaN
9996      NaN
9997      NaN
9998      NaN
9999      NaN

[10000 rows x 29 columns]
```

```
#Aggregation and Grouping
agg_result = df.groupby("AI_Platform")["Response_Accuracy"].mean()
print("Average accuracy per platform:\n", agg_result)

agg_multiple = df.groupby("AI_Platform").agg({
    "User_Experience_Score": "mean",
    "Response_Accuracy": ["mean", "max"]
})
print("Grouped stats:\n", agg_multiple)
```

```
Average accuracy per platform:
 AI_Platform
ChatGPT     0.802574
DeepSeek    0.899684
Name: Response_Accuracy, dtype: float64
Grouped stats:
             User_Experience_Score Response_Accuracy
                              mean              mean     max
AI_Platform
ChatGPT                  1.230971          0.802574  0.9467
DeepSeek                 2.034657          0.899684  0.9972
```

```
#Pivot Table
pivot = df.pivot_table(
    values="User_Experience_Score",
    index="AI_Platform",
    columns="Language",
    aggfunc="mean"
)
print("Pivot Table:\n", pivot)
```

```
Pivot Table:
 Language           de        en        es        fr        zh
AI_Platform
ChatGPT      1.221259  1.240488  1.227722  1.232822  1.232149
DeepSeek     2.035496  2.031093  2.043996  2.031644  2.031172
```