SKCT | EDA System Dataset for Colab | Untitled10.ipynb - Colab

colab.research.google.com/drive/1IF033K-ibn83bNOiVREaBOn6rtSbVhNW#scrollTo=qiF1BLZcQqQt

School

Sharing placement.skct.edu.in to placement.skct.edu.in | Stop sharing | Share this tab instead

Untitled10.ipynb

File Edit View Insert Runtime Tools Help

Share

Commands + Code + Text Run all

RAM Disk

[4] ✓ 13s

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

# 2. Import libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import sqlite3

# 3. Load dataset
df = pd.read_csv("customer_transactions.csv")

# 4. Basic understanding
print("DATA INFO")
df.info()

print("\nMISSING VALUES")
print(df.isnull().sum())

print("\nSTATISTICAL SUMMARY")
print(df.describe())

# 5. Data cleaning & transformation
df.drop_duplicates(inplace=True)
```

Variables Terminal

6:16 PM Python 3

31°C Partly sunny

Search

18:17 08-02-2026

SKCT | EDA System Dataset for Colab | Untitled10.ipynb - Colab

colab.research.google.com/drive/1IF033K-ibn83bNOiVREaBOn6rtSbVhNW#scrollTo=qiF1BLZcQqQt

School

Sharing placement.skct.edu.in to placement.skct.edu.in | Stop sharing | Share this tab instead

Untitled10.ipynb

File Edit View Insert Runtime Tools Help

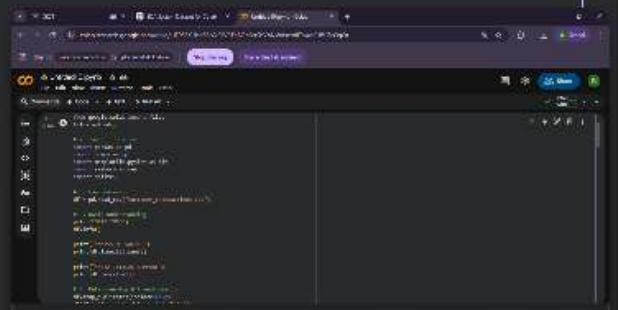
Share

Commands + Code + Text Run all

RAM Disk

[4] 13s

```
plt.title("Region-wise Sales Distribution")
plt.ylabel("")
plt.show()

        # Correlation heatmap
sns.heatmap(correlation, annot=True, cmap="coolwarm")
plt.title("Correlation Matrix")
plt.show()

        # 9. SQL Integration
conn = sqlite3.connect("customer.db")
df.to_sql("transactions", conn, if_exists="replace", index=False)

query = """
        SELECT Region, SUM(TotalAmount) AS TotalSales
        FROM transactions
        GROUP BY Region
        """

sql_result = pd.read_sql(query, conn)

print("\nSQL QUERY RESULT")
print(sql_result)

        # 10. Business Insights
```



Snipping Tool

Screenshot copied to clipboard
Automatically saved to screenshots folder.

Mark-up and share

Variables Terminal

31°C Partly sunny

Search

18:17
08-02-2026

SKCT | EDA System Dataset for Colab | Untitled10.ipynb - Colab

colab.research.google.com/drive/1IF033K-ibn83bNOiVREaBOn6rtSbVhNW#scrollTo=qiF1BLZcQqQt

School

Sharing placement.skct.edu.in to placement.skct.edu.in | Stop sharing | Share this tab instead

Untitled10.ipynb

File Edit View Insert Runtime Tools Help

Share

Commands + Code + Text Run all

RAM Disk

```
Data columns (total 9 columns):
 #   Column           Non-Null Count  Dtype 
---  ------           --------------  ----- 
 0   TransactionID    12 non-null     object
 1   CustomerID       12 non-null     object
 2   Date             12 non-null     object
 3   Region           12 non-null     object
 4   ProductCategory  12 non-null     object
 5   ProductName      12 non-null     object
 6   Quantity         12 non-null     int64 
 7   UnitPrice        12 non-null     int64 
 8   PaymentMode      12 non-null     object
dtypes: int64(2), object(7)
memory usage: 996.0+ bytes

MISSING VALUES
TransactionID     0
CustomerID        0
Date              0
Region            0
ProductCategory   0
ProductName       0
Quantity          0
UnitPrice         0
PaymentMode       0
dtype: int64
```

Variables Terminal

6:16 PM Python 3

31°C Partly sunny

Search

18:17 08-02-2026

SKCT | EDA System Dataset for Colab | Untitled10.ipynb - Colab

colab.research.google.com/drive/1IF033K-ibn83bNOiVREaBOn6rtSbVhNW#scrollTo=qiF1BLZcQqQt

School

Sharing placement.skct.edu.in to placement.skct.edu.in | Stop sharing | Share this tab instead

Untitled10.ipynb

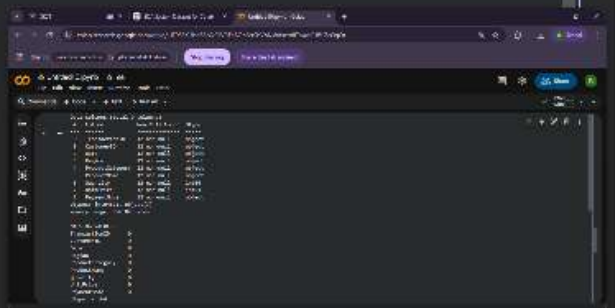
File Edit View Insert Runtime Tools Help

Share

Commands + Code + Text Run all

RAM Disk

```
STATISTICAL SUMMARY
       Quantity    UnitPrice
count  12.000000    12.000000
mean    2.000000  1616.666667
std     1.206045  1330.812010
min     1.000000   150.000000
25%     1.000000   725.000000
50%     2.000000  1350.000000
75%     2.250000  2125.000000
max     5.000000  5000.000000

MEAN SALES: 2250.0
VARIANCE OF SALES: 1498636.3636363635

CORRELATION MATRIX
             Quantity  UnitPrice  TotalAmount
Quantity     1.000000  -0.668358    -0.461803
UnitPrice   -0.668358   1.000000     0.868265
TotalAmount -0.461803   0.868265     1.000000

OUTLIERS
  TransactionID CustomerID       Date Region ProductCategory  ProductName  \
5          T006       C105 2023-01-10  North     Electronics  Smart Watch

   Quantity  UnitPrice PaymentMode  TotalAmount
5         1       5000         UPI         5000
```



Snipping Tool

Screenshot copied to clipboard
Automatically saved to screenshots folder.

Mark-up and share

Variables Terminal

31°C Partly sunny

Search

18:17 08-02-2026

SKCT | EDA System Dataset for Colab | Untitled10.ipynb - Colab

colab.research.google.com/drive/1IF033K-ibn83bNOiVREaBOn6rtSbVhNW#scrollTo=qiF1BLZcQqQt

Sharing placement.skct.edu.in to placement.skct.edu.in | Stop sharing | Share this tab instead

Untitled10.ipynb

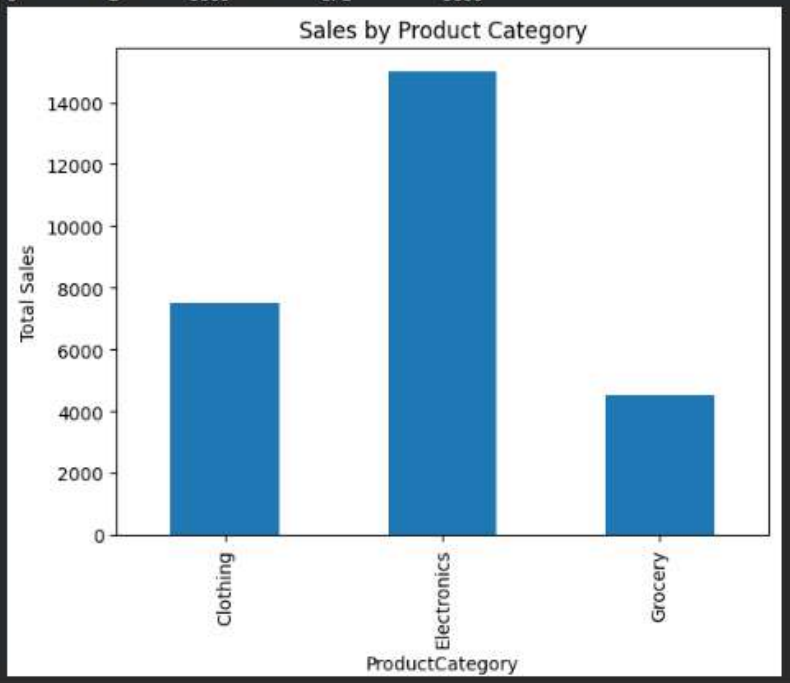
File Edit View Insert Runtime Tools Help

Commands + Code + Text Run all

RAM Disk

Share

```
   Quantity  UnitPrice PaymentMode  TotalAmount
5         1       5000         UPI         5000
```



Variables Terminal

6:16 PM Python 3

31°C Partly sunny

Search

18:17 08-02-2026

SKCT | EDA System Dataset for Colab | Untitled10.ipynb - Colab

colab.research.google.com/drive/1IF033K-ibn83bNOiVREaBOn6rtSbVhNW#scrollTo=qiF1BLZcQqQt

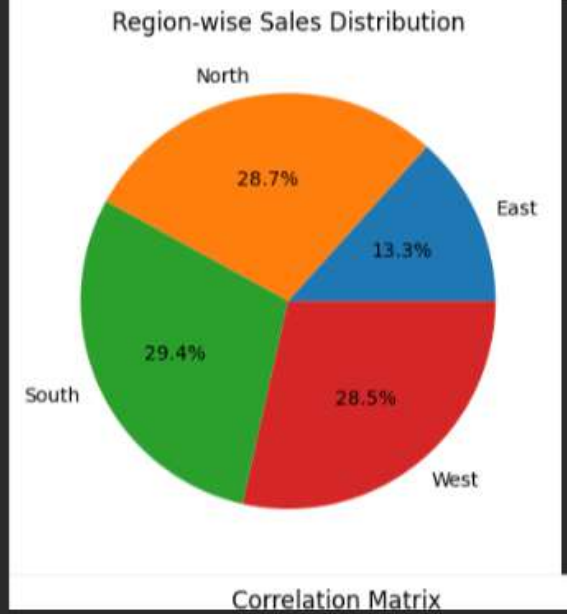
School

Sharing placement.skct.edu.in to placement.skct.edu.in | Stop sharing | Share this tab instead

File Edit View Insert Runtime Tools Help

Commands + Code + Text Run all

RAM Disk

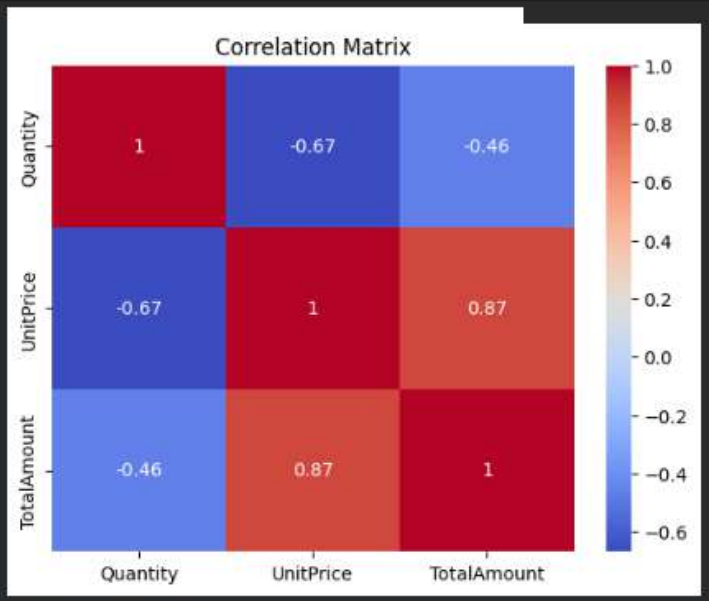


Correlation Matrix

31°C Partly sunny

Search

18:18 08-02-2026




```
SQL QUERY RESULT
  Region  TotalSales
0   East        3600
1  North        7750
2  South        7950
```