

```
import sqlite3

# DB connection
conn = sqlite3.connect("retail_orders.db")
cursor = conn.cursor()
```

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



Choose files Walmart_Sales.csv

- **Walmart_Sales.csv**(text/csv) - 363732 bytes, last modified: 05/09/2025 - 100% done
Saving Walmart_Sales.csv to Walmart_Sales.csv

```
import pandas as pd

df = pd.read_csv("Walmart_Sales.csv")
df.head()
```



1 to 5 of 5 entries

Filter



index	Store	Date	Weekly_Sales	Holiday_Flag	Temperature	Fuel_Price	CPI	Unemployrn
0	1	05-02-2010	1643690.9	0	42.31	2.572	211.0963582	8
1	1	12-02-2010	1641957.44	1	38.51	2.548	211.2421698	8
2	1	19-02-2010	1611968.17	0	39.93	2.514	211.2891429	8
3	1	26-02-2010	1409727.59	0	46.63	2.561	211.3196429	8

Next steps:

[Generate code with df](#)

[View recommended plots](#)

[New interactive sheet](#)

```
import sqlite3

conn = sqlite3.connect("retail_orders.db")
df.to_sql("orders", conn, if_exists="replace", index=False)

print("Data loaded into SQL table 'orders' successfully!")
```



Data loaded into SQL table 'orders' successfully!

```
pd.read_sql("SELECT Store, Date, Weekly_Sales FROM orders LIMIT 10;", conn)
```



```
pd.read_sql("SELECT * FROM orders ORDER BY Weekly_Sales DESC LIMIT 10;", conn)
```



	Store	Date	Weekly_Sales	Holiday_Flag	Temperature	Fuel_Price	CPI	Unemp
0	14	24-12-2010	3818686.45	0	30.59	3.141	182.544590	
1	20	24-12-2010	3766687.43	0	25.17	3.141	204.637673	
2	10	24-12-2010	3749057.69	0	57.06	3.236	126.983581	
3	4	23-12-2011	3676388.98	0	35.92	3.103	129.984548	
4	13	24-12-2010	3595903.20	0	34.90	2.846	126.983581	

```
pd.read_sql("""
SELECT Store, SUM(Weekly_Sales) AS Total_Sales, AVG(Weekly_Sales) AS Avg_Sales
FROM orders
GROUP BY Store
ORDER BY Total_Sales DESC;
""", conn)
```




	Store	Total_Sales	Avg_Sales
0	20	3.013978e+08	2.107677e+06
1	4	2.995440e+08	2.094713e+06
2	14	2.889999e+08	2.020978e+06
3	13	2.865177e+08	2.003620e+06
4	2	2.753824e+08	1.925751e+06
5	10	2.716177e+08	1.899425e+06
6	27	2.538559e+08	1.775216e+06
7	6	2.237561e+08	1.564728e+06
8	1	2.224028e+08	1.555264e+06
9	39	2.074455e+08	1.450668e+06
10	19	2.066349e+08	1.444999e+06
11	31	1.996139e+08	1.395901e+06
12	23	1.987506e+08	1.389864e+06
13	24	1.940160e+08	1.356755e+06
14	11	1.939628e+08	1.356383e+06
15	28	1.892637e+08	1.323522e+06
16	41	1.813419e+08	1.268125e+06
17	32	1.668192e+08	1.166568e+06
18	18	1.551147e+08	1.084718e+06
19	22	1.470756e+08	1.028501e+06
20	12	1.442872e+08	1.009002e+06
21	26	1.434164e+08	1.002912e+06
22	34	1.382498e+08	9.667816e+05
23	40	1.378703e+08	9.641280e+05
24	35	1.315207e+08	9.197250e+05
25	8	1.299512e+08	9.087495e+05
26	17	1.277821e+08	8.935814e+05
27	45	1.123953e+08	7.859814e+05
28	21	1.081179e+08	7.560691e+05
29	25	1.010612e+08	7.067215e+05
30	43	9.056544e+07	6.333247e+05
31	15	8.913368e+07	6.233125e+05



32	7	8.159828e+07	5.706173e+05
33	42	7.956575e+07	5.564039e+05
34	9	7.778922e+07	5.439806e+05
35	29	7.714155e+07	5.394514e+05
36	16	7.425243e+07	5.192477e+05
37	37	7.420274e+07	5.189003e+05
38	30	6.271689e+07	4.385796e+05
39	3	5.758674e+07	4.027044e+05
40	38	5.515963e+07	3.857317e+05
41	36	5.341221e+07	3.735120e+05
42	5	4.547569e+07	3.180118e+05
43	44	4.329309e+07	3.027489e+05
44	33	3.716022e+07	2.598617e+05

```
df['Date'] = pd.to_datetime(df['Date'], errors='coerce')
```

```
df.to_sql("orders", conn, if_exists="replace", index=False)
```

 6435

```
pd.read_sql("""
SELECT strftime('%m', Date) AS Month, SUM(Weekly_Sales) AS Monthly_Sales
FROM orders
GROUP BY Month
ORDER BY Monthly_Sales DESC;
""", conn)
```



	Month	Monthly_Sales	
0	None	4.018647e+09	
1	10	2.463838e+08	
2	09	2.450696e+08	
3	06	2.447777e+08	
4	02	2.409695e+08	
5	03	2.394468e+08	
6	04	2.391020e+08	
7	05	2.379863e+08	
8	11	2.351902e+08	
9	12	2.327650e+08	
10	08	1.885009e+08	
11	07	1.868208e+08	
12	01	1.815590e+08	

```
pd.read_sql("SELECT * FROM orders WHERE Holiday_Flag = 1 ORDER BY Weekly_Sales DESC LIMIT
```



	Store	Date	Weekly_Sales	Holiday_Flag	Temperature	Fuel_Price	CPI	Unem
0	4	None	3004702.33	1	47.96	3.225	129.836400	
1	10	None	2950198.64	1	60.68	3.760	129.836400	
2	10	None	2939946.38	1	55.33	3.162	126.669267	
3	14	None	2921709.71	1	46.15	3.039	182.783277	
4	20	None	2906233.25	1	46.38	3.492	211.412076	
5	13	None	2864170.61	1	38.89	3.445	129.836400	
6	20	None	2811634.04	1	46.66	3.039	204.962100	
7	4	None	2789469.45	1	48.08	2.752	126.669267	
8	13	None	2766400.05	1	28.22	2.830	126.669267	
9	14	None	2685351.81	1	48.71	3.492	188.350400	

```
pd.read_sql("SELECT Store, Weekly_Sales AS Revenue FROM orders LIMIT 5;", conn)
```



	Store	Revenue	
0	1	1643690.90	
1	1	1641957.44	
2	1	1611968.17	
3	1	1409727.59	
4	1	1554806.68	

```
df['Date'] = pd.to_datetime(df['Date'], errors='coerce')
df.to_sql("orders", conn, if_exists="replace", index=False)
```



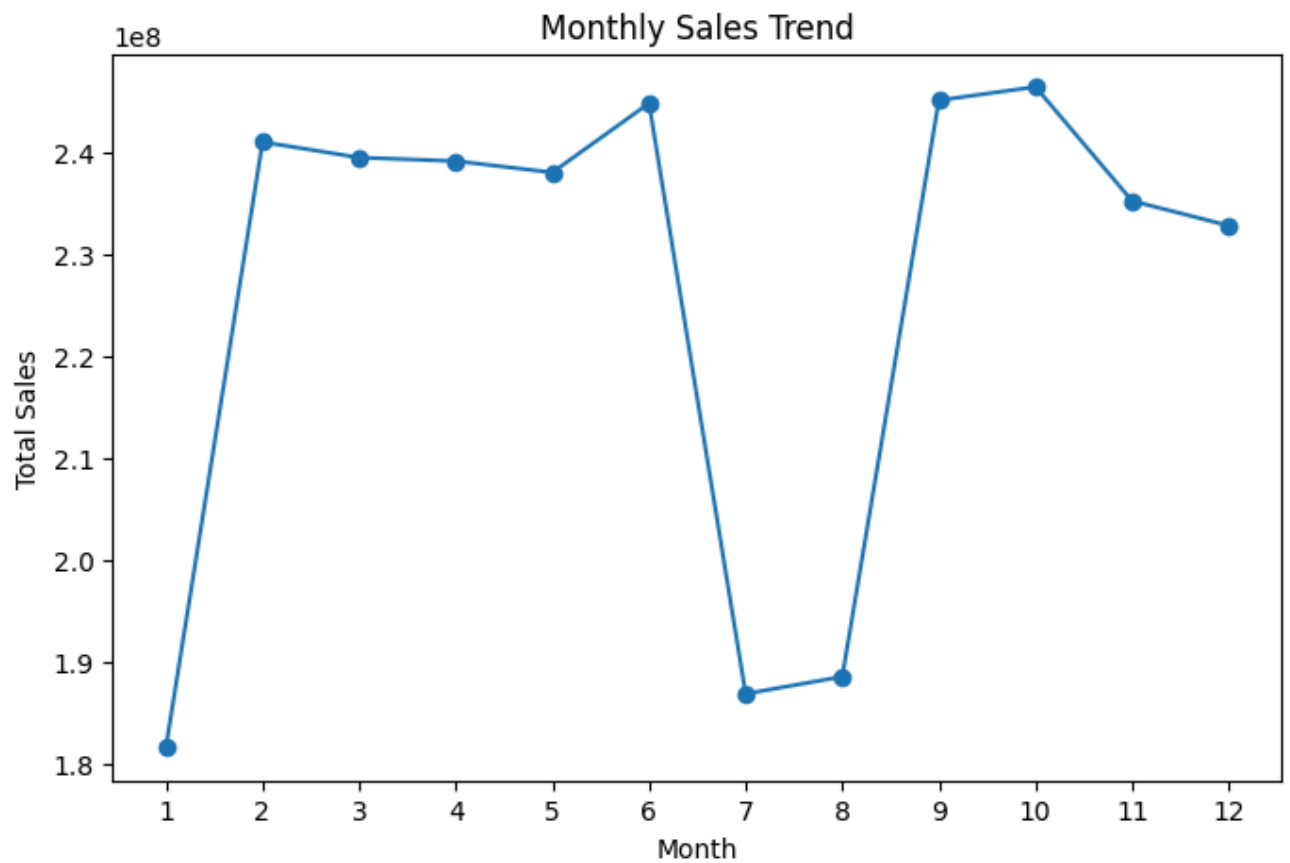
6435

```
monthly_sales = pd.read_sql("""
SELECT strftime('%m', Date) AS Month, SUM(Weekly_Sales) AS Monthly_Sales
FROM orders
WHERE Date IS NOT NULL
GROUP BY Month
ORDER BY Month;
""", conn)
```

```
monthly_sales = monthly_sales.dropna()
monthly_sales['Month'] = monthly_sales['Month'].astype(int)
```

```
import matplotlib.pyplot as plt

plt.figure(figsize=(8,5))
plt.plot(monthly_sales['Month'], monthly_sales['Monthly_Sales'], marker='o')
plt.title("Monthly Sales Trend")
plt.xlabel("Month")
plt.ylabel("Total Sales")
plt.xticks(range(1,13))
plt.show()
```



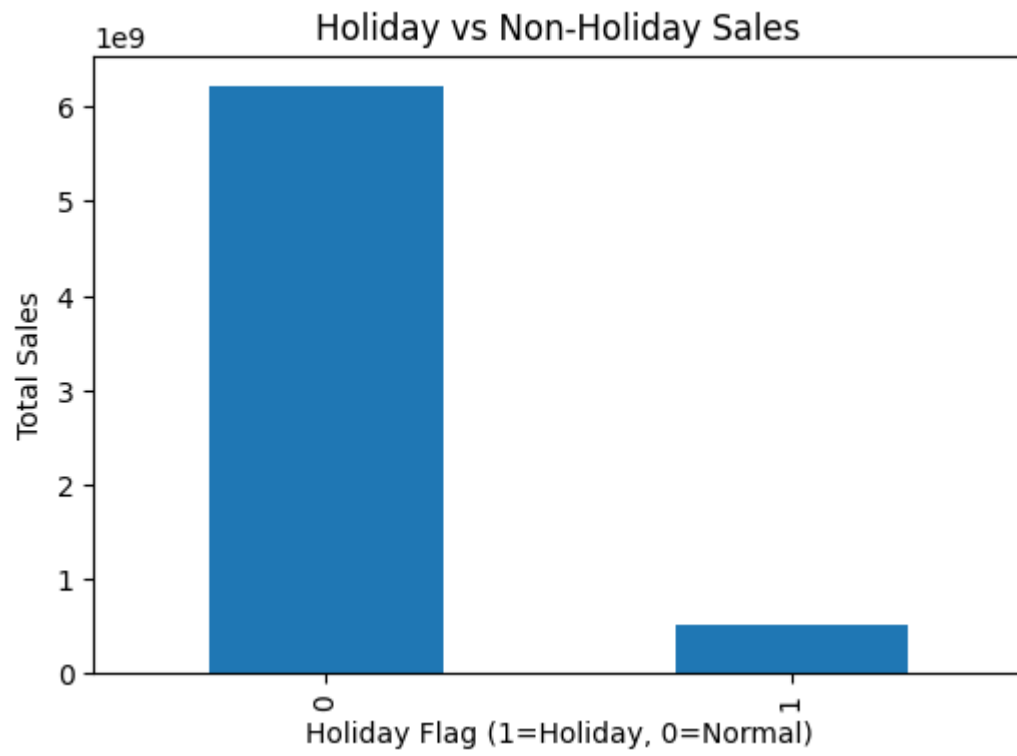
```
store_sales = pd.read_sql("""
SELECT Store, SUM(Weekly_Sales) AS Total_Sales
FROM orders
GROUP BY Store
ORDER BY Total_Sales DESC
LIMIT 10;
""", conn)
```

```
store_sales.plot(kind="bar", x="Store", y="Total_Sales", figsize=(8,5), legend=False)
plt.title("Top 10 Stores by Total Sales")
plt.xlabel("Store")
plt.ylabel("Total Sales")
plt.show()
```




```
holiday_sales = pd.read_sql("""
SELECT Holiday_Flag, SUM(Weekly_Sales) AS Total_Sales
FROM orders
GROUP BY Holiday_Flag;
""", conn)
```

```
holiday_sales.plot(kind="bar", x="Holiday_Flag", y="Total_Sales", figsize=(6,4), legend=F
plt.title("Holiday vs Non-Holiday Sales")
plt.xlabel("Holiday Flag (1=Holiday, 0=Normal)")
plt.ylabel("Total Sales")
plt.show()
```



```
pd.read_sql("SELECT DISTINCT Store FROM orders;", conn)
```



Store



0	1
1	2
2	3
3	4
4	5
5	6
6	7
7	8
8	9
9	10
10	11
11	12
12	13
13	14
14	15
15	16
16	17
17	18
18	19
19	20
20	21
21	22
22	23
23	24
24	25
25	26
26	27
27	28
28	29
29	30
30	31
31	32

32 33
33 34
34 35
35 36
36 37
37 38
38 39
39 40
40 41
41 42
42 43
43 44
44 45

```
pd.read_sql("SELECT * FROM orders WHERE Weekly_Sales BETWEEN 50000 AND 100000;", conn)
```



Store	Date	Weekly_Sales	Holiday_Flag	Temperature	Fuel_Price	CPI	Unemployment
-------	------	--------------	--------------	-------------	------------	-----	--------------

```
pd.read_sql("""
SELECT Store, SUM(Weekly_Sales) AS Total_Sales
FROM orders
GROUP BY Store
HAVING Total_Sales > 10000000
ORDER BY Total_Sales DESC;
""", conn)
```



	Store	Total_Sales
0	20	3.013978e+08
1	1	2.005440e+08

