

```
In [2]: import sqlite3
```

```
In [3]: # Establish a connection to the database
connection = sqlite3.connect('data/rexon_metals.db')
```

```
In [4]: # Create a cursor object using the connection
cursor = connection.cursor()
```

```
In [5]: # This example retrieves all entries from the PRODUCT table
cursor.execute("SELECT * FROM PRODUCT")

results = cursor.fetchall()

results
```

```
Out[5]: [(1, 'Copper', 7.51),
(2, 'Aluminum', 2.58),
(3, 'Silver', 15),
(4, 'Steel', 12.31),
(5, 'Bronze', 4),
(6, 'Duralumin', 7.6),
(7, 'Solder', 14.16),
(8, 'Stellite', 13.31),
(9, 'Brass', 4.75)]
```

```
In [6]: # Using the placeholder '?' to query with variable names
product_id = 3
cursor.execute("SELECT * FROM product WHERE product_id = ?", (product_id,))
print(cursor.fetchone()) # fetchone gives us a single row

(3, 'Silver', 15)
```

```
In [7]: # Define your query with placeholders
query = "SELECT * FROM product WHERE price BETWEEN ? AND ? AND description I

# Parameters for the placeholders
min_price = 10
max_price = 14
description_pattern = '%S%'

# Execute the query with the parameters
cursor.execute(query, (min_price, max_price, description_pattern))

# Fetch and print the results
results = cursor.fetchall()
for result in results:
    print(result)

(4, 'Steel', 12.31)
(8, 'Stellite', 13.31)
```

```
In [8]: import pandas as pd
```

```
In [9]: # Querying data and loading directly into a DataFrame
df = pd.read_sql_query("SELECT * FROM product", connection)
print(df.head())
```

	PRODUCT_ID	DESCRIPTION	PRICE
0	1	Copper	7.51
1	2	Aluminum	2.58
2	3	Silver	15.00
3	4	Steel	12.31
4	5	Bronze	4.00

```
In [10]: # Make the query a (multi-line) string
query = """
SELECT c.name, co.order_qty * p.price AS total_price
FROM customer_order co
JOIN customer c ON co.customer_id = c.customer_id
JOIN product p ON co.product_id = p.product_id;
"""

df = pd.read_sql_query(query, connection)
print(df)
```

	NAME	total_price
0	LITE Industrial	3379.50
1	Re-Barre Construction	1548.00
2	Re-Barre Construction	1200.00
3	Marsh Lane Metal Works	4616.25
4	Re-Barre Construction	1290.00

```
In [12]: df.sort_values(by = 'total_price', ascending = False)
```

```
Out[12]:
```

	NAME	total_price
3	Marsh Lane Metal Works	4616.25
0	LITE Industrial	3379.50
1	Re-Barre Construction	1548.00
4	Re-Barre Construction	1290.00
2	Re-Barre Construction	1200.00

```
In [13]: # important cleanup steps
cursor.close()
connection.close()
```

---

## Exercise

```
In [14]: # establish connection to the weather stations database
w_connection = sqlite3.connect("data/weather_stations.db")
```

```
In [16]: # Create a cursor object using the connection
w_cursor = w_connection.cursor()
```

```
In [20]: # Querying data and loading directly into a DataFrame
ws = pd.read_sql_query("SELECT * FROM station_data", w_connection)
ws.head()
```

```
Out[20]:
```

	station_number	report_code	year	month	day	dew_point	station_pressure	visibility
0	143080	34DDA7	2002	12	21	33.8	987.4	3.4
1	766440	39537B	1998	10	1	72.7	1014.6	5.9
2	176010	C3C6D5	2001	5	18	55.7	NaN	7.3
3	125600	145150	2007	10	14	33.0	NaN	6.9
4	470160	EF616A	1967	7	29	65.6	NaN	9.2

```
In [34]: import matplotlib.pyplot as plt

x = ws['temperature']
y = ws['precipitation']

plt.scatter(x,y);
plt.ylabel("Precipitation")
plt.xlabel("Temperature")
plt.title("Temperature vs. Precipitation");
```

