Task 7

We are Using Sales Data

Answers Of the given questions.

1. How did you connect R to a database?

We used the RSQLite package to connect R with an SQLite database:

r

CopyEdit

```
library(RSQLite)
conn <- dbConnect(SQLite(), dbname = "sales_data.db")
```

This creates or connects to a local SQLite database file named sales_data.db.

2. What SQL query did you run?

We ran the following SQL query to compute total quantity and revenue by product:

sql

CopyEdit

```
SELECT Description AS product,
    SUM(Quantity) AS total_qty,
    SUM(Quantity * UnitPrice) AS revenue
FROM sales
WHERE Quantity > 0 AND UnitPrice > 0
GROUP BY Description
ORDER BY revenue DESC
LIMIT 10
```

3. What does GROUP BY do?

GROUP BY groups rows with the same value in a specific column. In our case:

sql

CopyEdit

```
GROUP BY Description
```

groups all rows by product name, so we can summarize sales per product.

4. How did you calculate revenue?

Revenue is calculated using:

sql

CopyEdit

SUM(Quantity * UnitPrice) AS revenue

It multiplies Quantity and UnitPrice for each sale and then adds them up for each product.

5. How did you visualize the result?

We used ggplot2 to plot a bar chart of the top 10 products by revenue:

r

CopyEdit

```
ggplot(sales_summary, aes(x = reorder(product, revenue), y = revenue, fill = product)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  labs(title = "Top 10 Products by Revenue", x = "Product", y = "Revenue") +
  theme_minimal()
```

6. What does pandas do in your code?

In R, we don't use pandas (Python-specific). Instead, we use:

- dplyr for data manipulation
- tibble or data.frame for storing tabular data

This is how we pull data from SQLite:

r

CopyEdit

```
library(dplyr)
sales_summary <- dbGetQuery(conn, query)
```

7. What's the benefit of using SQL inside R?

- Efficient querying of large datasets
- Combines SQL's data handling with R's analysis/visualization tools
- Cleaner code using standard SQL
- Avoids memory overload by filtering before loading
- Seamless workflow for data science

8. Could you run the same SQL query in DB Browser for SQLite?

Yes! You can open sales_data.db in **DB Browser for SQLite**, go to the **Execute SQL** tab, and paste the same query:

```
sql

CopyEdit

SELECT Description AS product,

    SUM(Quantity) AS total_qty,

    SUM(Quantity * UnitPrice) AS revenue

FROM sales

WHERE Quantity > 0 AND UnitPrice > 0

GROUP BY Description

ORDER BY revenue DESC

LIMIT 10;
```