**Eastvantage- Assignment – Solution – Data Engineer**

To achieve the objectives of extracting total quantities of each item bought per customer aged 18-35 from the provided SQLite3 database and storing the results in a CSV file we can use both SQL and Pandas.

Both solutions will connect to the SQLite3 database, execute the SQL query to extract the required data, and then store the results in a CSV file with a semicolon (';') delimiter.

**Pandas Solution:**

import pandas as pd

import sqlite3

# Connect to the SQLite3 database

conn = sqlite3.connect('your\_database.db')

# SQL query to extract the desired data

sql\_query = '''

SELECT c.customer\_id, c.age, i.item\_name, SUM(o.quantity) AS total\_quantity

FROM Customers AS c

JOIN Sales AS s ON c.customer\_id = s.customer\_id

JOIN Orders AS o ON s.sales\_id = o.sales\_id

JOIN Items AS i ON o.item\_id = i.item\_id

WHERE c.age >= 18 AND c.age <= 35

GROUP BY c.customer\_id, i.item\_name

HAVING total\_quantity > 0;

'''

# Use Pandas to read the query results directly into a DataFrame

df = pd.read\_sql\_query(sql\_query, conn)

# Close the database connection

conn.close()

# Write the DataFrame to a CSV file

df.to\_csv('output.csv', sep=';', index=False)

**SQL SOLUTION:**

import sqlite3

import csv

# Connect to the SQLite3 database

conn = sqlite3.connect('your\_database.db')

cursor = conn.cursor()

# SQL query to extract the desired data

sql\_query = '''

SELECT c.customer\_id, c.age, i.item\_name, SUM(o.quantity) AS total\_quantity

FROM Customers AS c

JOIN Sales AS s ON c.customer\_id = s.customer\_id

JOIN Orders AS o ON s.sales\_id = o.sales\_id

JOIN Items AS i ON o.item\_id = i.item\_id

WHERE c.age >= 18 AND c.age <= 35

GROUP BY c.customer\_id, i.item\_name

HAVING total\_quantity > 0;

'''

# Execute the query and fetch the results

cursor.execute(sql\_query)

results = cursor.fetchall()

# Close the database connection

conn.close()

# Write the results to a CSV file

with open('output.csv', 'w', newline='') as csvfile:

csv\_writer = csv.writer(csvfile, delimiter=';')

csv\_writer.writerow(['customer', 'Age', 'Item', 'Quantity'])

for row in results:

csv\_writer.writerow(row)