Task 4\task4.py

```python
import pandas as pd
import sqlite3

# Connect to the database
conn = sqlite3.connect('Task 4/shipment_database.db')
cursor = conn.cursor()

# Load data
shipping_data0 = pd.read_csv('Task 4/data/shipping_data_0.csv')
shipping_data1 = pd.read_csv('Task 4/data/shipping_data_1.csv')
shipping_data2 = pd.read_csv('Task 4/data/shipping_data_2.csv')

# Merge shipping data
merged_shipping_data = shipping_data1.merge(shipping_data2, on='shipment_identifier')

# Group by shipment_identifier and product to count quantities
grouped_data = merged_shipping_data.groupby(['shipment_identifier', 'product',
'origin_warehouse', 'destination_store']).size().reset_index(name='quantity')

# Insert products into the database, ensuring no duplicates
all_products = pd.concat([shipping_data0['product'],
shipping_data1['product']]).drop_duplicates()

for product_name in all_products:
    cursor.execute('INSERT OR IGNORE INTO product (name) VALUES (?)', (product_name,))

# Insert shipments into the database
for index, row in grouped_data.iterrows():
    product_name = row['product']
    quantity = row['quantity']
    origin = row['origin_warehouse']
    destination = row['destination_store']

    cursor.execute('SELECT id FROM product WHERE name = ?', (product_name,))
    product_id_result = cursor.fetchone()

    if product_id_result:
        product_id = product_id_result[0]
        cursor.execute('''
            INSERT INTO shipment (product_id, quantity, origin, destination)
            VALUES (?, ?, ?, ?)
        ''', (product_id, quantity, origin, destination))
    else:
        print(f"Product '{product_name}' not found in product table!")

# Commit the transaction and close the connection
conn.commit()
conn.close()
```

```
47
48
```