DSE5002 MODULE 5 LAB

Peter Gyorda April 19, 2025

Beaulieu's book, Exercise 6-2, 6-3, 8-1, 8-2, 10-1

For these questions, I will use the bank database

```
In [3]:  # Set Up and Connect
```

```
In [19]: # Libaries

         import sqlalchemy

         # we will want Pandas for the data frame structure

         import pandas as pd
```

```
In [21]: # Connect to the database
         # Alter this to reflect your username and password,   this is for postgres on the s

         engine=sqlalchemy.create_engine('postgresql://todd:password@localhost:5432/bank')
```

```
In [ ]:  # Exercise 6-2
         Write a compound query that finds the first and last names of all individual custom
```

```
In [41]: import pandas as pd
         import sqlalchemy

         def fetch_customer_employee_names_bankdb(engine):
             """
             Fetches first and last names of customers and employees.

             Args:
                 engine (sqlalchemy.engine.base.Engine): The SQLAlchemy engine connected
                                                         to the bank's PostgreSQL database.

             Returns:
                 pandas.DataFrame: A DataFrame with columns 'fname', 'lname', and 'Type'
                                   containing the names of customers and employees.
             """
             # Query for customers
             customer_query = """
                 SELECT fname, lname, 'Customer' AS Type
                 FROM individual
             """
             customers_df = pd.read_sql(customer_query, engine)

             # Query for employees
             employee_query = """
```

```
        SELECT fname, lname, 'Employee' AS Type
        FROM employee
    """
    employees_df = pd.read_sql(employee_query, engine)

    # Combine the DataFrames
    all_names_df = pd.concat([customers_df, employees_df], ignore_index=True)
    return all_names_df

if __name__ == "__main__":
    # PostgreSQL connection details for the 'bank' database
    db_user = 'todd'
    db_password = 'password'
    db_host = 'localhost'
    db_port = '5432'
    db_name = 'bank'  # Assuming your database is named 'bank'

    engine = sqlalchemy.create_engine(f'postgresql://{db_user}:{db_password}@{db_ho

    names_df = fetch_customer_employee_names_bankdb(engine)

    if not names_df.empty:
        print("First and Last Names of Customers and Employees:")
        print(names_df)
    else:
        print("Could not retrieve names from the bank database.")

    # Dispose of the engine
    engine.dispose()
```

```
First and Last Names of Customers and Employees:
       fname      lname      type
0      James     Hadley  Customer
1      Susan    Tingley  Customer
2      Frank     Tucker  Customer
3       John    Hayward  Customer
4    Charles    Frasier  Customer
5       John    Spencer  Customer
6   Margaret      Young  Customer
7      Louis      Blake  Customer
8    Richard     Farley  Customer
9    Michael      Smith  Employee
10     Susan     Barker  Employee
11    Robert      Tyler  Employee
12     Susan  Hawthorne  Employee
13      John    Gooding  Employee
14     Helen    Fleming  Employee
15     Chris     Tucker  Employee
16     Sarah     Parker  Employee
17      Jane   Grossman  Employee
18     Paula    Roberts  Employee
19    Thomas    Ziegler  Employee
20  Samantha    Jameson  Employee
21      John      Blake  Employee
22     Cindy      Mason  Employee
23     Frank    Portman  Employee
24   Theresa    Markham  Employee
25      Beth     Fowler  Employee
26      Rick     Tulman  Employee
```

In [ ]:

```
#Exercise 6-3

Sort the results from Exercise 6-2 by the lname column.
```

In [43]:

```
import pandas as pd
import sqlalchemy

def fetch_customer_employee_names_bankdb(engine):
    """
    Fetches first and last names of customers and employees.

    Args:
        engine (sqlalchemy.engine.base.Engine): The SQLAlchemy engine connected
                                                to the bank's PostgreSQL database.

    Returns:
        pandas.DataFrame: A DataFrame with columns 'fname', 'lname', and 'Type'
                          containing the names of customers and employees,
                          sorted by last name.
    """
    # Query for customers
    customer_query = """
        SELECT fname, lname, 'Customer' AS Type
        FROM individual
    """
    customers_df = pd.read_sql(customer_query, engine)
    # Query for employees
    employee_query = """
        SELECT fname, lname, 'Employee' AS Type
        FROM employee
    """
    employees_df = pd.read_sql(employee_query, engine)

    # Combine the DataFrames
    all_names_df = pd.concat([customers_df, employees_df], ignore_index=True)

    # Sort by last name
    all_names_df = all_names_df.sort_values(by='lname')

    return all_names_df

if __name__ == "__main__":
    # PostgreSQL connection details for the 'bank' database
    db_user = 'todd'
    db_password = 'password'
    db_host = 'localhost'
    db_port = '5432'
    db_name = 'bank'  # Assuming your database is named 'bank'

    engine = sqlalchemy.create_engine(f'postgresql://{db_user}:{db_password}@{db_ho

    names_df = fetch_customer_employee_names_bankdb(engine)

    if not names_df.empty:
        print("First and Last Names of Customers and Employees (sorted by last name
        print(names_df)
    else:
        print("Could not retrieve names from the bank database.")

    # Dispose of the engine
    engine.dispose()
```

```
First and Last Names of Customers and Employees (sorted by last name):
       fname      lname      type
10     Susan     Barker  Employee
21      John      Blake  Employee
7      Louis      Blake  Customer
8    Richard     Farley  Customer
14     Helen    Fleming  Employee
25      Beth     Fowler  Employee
4    Charles    Frasier  Customer
13      John    Gooding  Employee
17      Jane   Grossman  Employee
0      James     Hadley  Customer
12     Susan  Hawthorne  Employee
3       John    Hayward  Customer
20  Samantha    Jameson  Employee
24   Theresa    Markham  Employee
22     Cindy      Mason  Employee
16     Sarah     Parker  Employee
23     Frank    Portman  Employee
18     Paula    Roberts  Employee
9    Michael      Smith  Employee
5       John    Spencer  Customer
1      Susan    Tingley  Customer
15     Chris     Tucker  Employee
2      Frank     Tucker  Customer
26      Rick     Tulman  Employee
11    Robert      Tyler  Employee
6   Margaret      Young  Customer
19    Thomas    Ziegler  Employee
```

In [ ]:
```
#Exercise 8-1

Construct a query that counts the number of rows in the account table.
```

In [45]:
```
import pandas as pd
import sqlalchemy

def count_accounts(engine):
    """
    Counts the number of rows in the account table.

    Args:
        engine (sqlalchemy.engine.base.Engine): The SQLAlchemy engine connected
                                                to the bank's PostgreSQL database.

    Returns:
        int: The number of accounts in the database.
    """
    # Query to count rows in the account table
    count_query = """
        SELECT COUNT(*) as account_count
        FROM account
    """
    result_df = pd.read_sql(count_query, engine)

    # Extract the count value
    account_count = result_df['account_count'].iloc[0]

    return account_count

if __name__ == "__main__":
    # PostgreSQL connection details for the 'bank' database
    db_user = 'todd'
    db_password = 'password'
    db_host = 'localhost'
    db_port = '5432'
    db_name = 'bank'  # Assuming your database is named 'bank'

    engine = sqlalchemy.create_engine(f'postgresql://{db_user}:{db_password}@{db_ho

    account_count = count_accounts(engine)

    print(f"Total number of accounts in the database: {account_count}")

    # Dispose of the engine
    engine.dispose()
```

```
Total number of accounts in the database: 24
```

In [ ]:
```
# Exercise 8-2

Modify your query from Exercise 8-1 to count the number of accounts held by each cu
```

In [49]:
```
import pandas as pd
import sqlalchemy

def count_customer_accounts(engine):
    """
    Counts the number of accounts held by each customer.

    Args:
        engine (sqlalchemy.engine.base.Engine): The SQLAlchemy engine connected
                                                to the bank's PostgreSQL database.

    Returns:
        pandas.DataFrame: A DataFrame with columns 'cust_id' and 'account_count'
                          showing the number of accounts for each customer.
    """
    # Query to count accounts per customer
    count_query = """
        SELECT cust_id, COUNT(*) as account_count
        FROM account
        GROUP BY cust_id
        ORDER BY account_count DESC
    """
    # Changed 'customer_id' to 'cust_id' to match the actual column name in the dat
    result_df = pd.read_sql(count_query, engine)

    return result_df

if __name__ == "__main__":
    # PostgreSQL connection details for the 'bank' database
    db_user = 'todd'
    db_password = 'password'
    db_host = 'localhost'
    db_port = '5432'
    db_name = 'bank'  # Assuming your database is named 'bank'

    engine = sqlalchemy.create_engine(f'postgresql://{db_user}:{db_password}@{db_ho

    customer_accounts_df = count_customer_accounts(engine)

    if not customer_accounts_df.empty:
        print("Number of accounts held by each customer:")
        print(customer_accounts_df)
    else:
        print("Could not retrieve account information from the database.")

    # Dispose of the engine
    engine.dispose()
```

```
Number of accounts held by each customer:
    cust_id  account_count
0         4              3
1         9              3
2         1              3
3        10              2
4         8              2
5         2              2
6         6              2
7         3              2
8        12              1
9        11              1
10        7              1
11       13              1
12        5              1
```

In [ ]:

```
# Exercise 10-1
Write a query that returns all product names along with the accounts based on that
```

In [53]:

```
# SQL query to return all product names along with accounts based on that product
query = """
SELECT p.product_cd, p.name AS product_name, a.account_id, a.cust_id
FROM product p
LEFT JOIN account a ON p.product_cd = a.product_cd
ORDER BY p.product_cd, a.account_id
"""

# To execute this query using pandas and SQLAlchemy:
import pandas as pd
import sqlalchemy

# Assuming you have already set up your engine connection
# engine = sqlalchemy.create_engine('your_connection_string')

result_df = pd.read_sql(query, engine)
print(result_df)
```

```python
# To close the database connection
engine.dispose()
```

```
  product_cd               product_name  account_id  cust_id
0        AUT                  auto loan         NaN      NaN
1        BUS  business line of credit        25.0     10.0
2        BUS  business line of credit        27.0     11.0
3         CD   certificate of deposit         3.0      1.0
4         CD   certificate of deposit        15.0      6.0
5         CD   certificate of deposit        17.0      7.0
6         CD   certificate of deposit        23.0      9.0
7        CHK         checking account         1.0      1.0
8        CHK         checking account         4.0      2.0
9        CHK         checking account         7.0      3.0
10       CHK         checking account        10.0      4.0
11       CHK         checking account        13.0      5.0
12       CHK         checking account        14.0      6.0
13       CHK         checking account        18.0      8.0
14       CHK         checking account        21.0      9.0
15       CHK         checking account        24.0     10.0
16       CHK         checking account        28.0     12.0
17        MM     money market account         8.0      3.0
18        MM     money market account        12.0      4.0
19        MM     money market account        22.0      9.0
20       MRT            home mortgage         NaN      NaN
21       SAV          savings account         2.0      1.0
22       SAV          savings account         5.0      2.0
23       SAV          savings account        11.0      4.0
24       SAV          savings account        19.0      8.0
25       SBL      small business loan        29.0     13.0
```

In [ ]: