Homework 12 – Big Data

Dave Gagliano

CIS -2532 – NET 001

15 May 2022

Professor Shamsuddin

Outputs

Section 17.2.1: A Books Database

Execute an SQL query and return a DataFrame containing the query's results

```
# execute an SQL query and return a DataFrame containing the query's results
pd.read_sql('SELECT * FROM authors', connection, index_col=['id'])
```

Out[14]:

id	first	last
1	Paul	Deitel
2	Harvey	Deitel
3	Abbey	Deitel
4	Dan	Quirk
5	Alexander	Wald

Use an SQL query to view the titles table and its contents

In [16]:
```
# use an SQL query to view the titles table and its contents
pd.read_sql('SELECT* FROM titles', connection)
```

Out[16]:

	isbn	title	edition	copyright
0	0135404673	Intro to Python for CS and DS	1	2020
1	0132151006	Internet & WWW How to Program	5	2012
2	0134743350	Java How to Program	11	2018
3	0133976890	C How to Program	8	2016
4	0133406954	Visual Basic 2012 How to Program	6	2014
5	0134601548	Visual C# How to Program	6	2017
6	0136151574	Visual C++ How to Program	2	2008
7	0134448235	C++ How to Program	10	2017
8	0134444302	Android How to Program	3	2017
9	0134289366	Android 6 for Programmers	3	2016

Use SQL and pandas to view the five rows from the author_ISBN table

In [31]:
```
# use SQL and pandas to view the five rows from the author_ISBN table
df = pd.read_sql('SELECT * FROM author_ISBN', connection)
df.head()
```

Out[31]:

	id	isbn
0	1	0134289366
1	2	0134289366
2	5	0134289366
3	1	0135404673
4	2	0135404673

Exercise 17.1: Books Database

a) Select all authors' last names from the authors table in descending order

In [133]:
```
# use SQL query to return a dataframe with the query's results and sort in descending order by last name using SQL
pd.read_sql("""
            SELECT *
            FROM authors
            ORDER BY last DESC
            """, connection)
```

Out[133]:

	id	first	last
0	5	Alexander	Wald
1	4	Dan	Quirk
2	1	Paul	Deitel
3	2	Harvey	Deitel
4	3	Abbey	Deitel

b) Select all book titles from the titles table in ascending order

In [120]:
```
# use SQL query to retrun a dataframe with the query's results and sort in ascending order with SQL
pd.read_sql("""
            SELECT * FROM titles
            ORDER BY title ASC
            """, connection)
```

Out[120]:

	isbn	title	edition	copyright
0	0134289366	Android 6 for Programmers	3	2016
1	0134444302	Android How to Program	3	2017
2	0133976890	C How to Program	8	2016
3	0134448235	C++ How to Program	10	2017
4	0132151006	Internet & WWW How to Program	5	2012
5	0135404673	Intro to Python for CS and DS	1	2020
6	0134743350	Java How to Program	11	2018
7	0133406954	Visual Basic 2012 How to Program	6	2014
8	0134601548	Visual C# How to Program	6	2017
9	0136151574	Visual C++ How to Program	2	2008

c) Use an INNER JOIN to select all the books from a specific author. Include the title, copyright year, and ISBN. Order the information alphabetically by title.

In [144]:

```
# use SQL query to retrun a dataframe with the query's results
firstname= 'Paul'
lastname = 'Deitel'

def select_all_books_author(firstname, lastname):

    return pd.read_sql(f"""
                        SELECT
                        authors.first,
                        authors.last,
                        titles.title,
                        titles.copyright,
                        titles.isbn

                        FROM authors

                        INNER JOIN author_ISBN
                            ON authors.id = author_ISBN.id

                        INNER JOIN titles
                            ON author_ISBN.isbn = titles.isbn

                        WHERE authors.first LIKE '{firstname}' AND  authors.last LIKE '{lastname}'

                        ORDER BY title ASC
                        """, connection)

#display output in Jupyter Notebook
select_all_books_author(firstname, lastname)
```

Out[144]:

	first	last	title	copyright	isbn
0	Paul	Deitel	Android 6 for Programmers	2016	0134289366
1	Paul	Deitel	Android Howto Program	2017	0134444302
2	Paul	Deitel	C Howto Program	2016	0133976890
3	Paul	Deitel	C++ Howto Program	2017	0134448235
4	Paul	Deitel	Internet & WWW Howto Program	2012	0132151006
5	Paul	Deitel	Intro to Python for CS and DS	2020	0135404673
6	Paul	Deitel	Java Howto Program	2018	0134743350
7	Paul	Deitel	Visual Basic 2012 Howto Program	2014	0133406954
8	Paul	Deitel	Visual C# Howto Program	2017	0134601548
9	Paul	Deitel	Visual C++ Howto Program	2008	0136151574

d) Insert a new author into the authors table

In [149]:

```
pd.read_sql("""
            SELECT * FROM authors
            """, connection)
```

Out[149]:

	id	first	last
0	1	Paul	Deitel
1	2	Harvey	Deitel
2	3	Abbey	Deitel
3	4	Dan	Quirk
4	5	Alexander	Wald
5	6	Al	Sweigart

e) Insert a new title for an author. Remember that the book must have an entry in the author_ISBN table and an entry in the titles table

In [242]:
```
# insert new entry into titles table
cursor = cursor.execute(f"""
                        INSERT INTO titles(isbn, title, edition, copyright)
                        VALUES('{isbn_book}', '{title}', '{edition}', '{copyright}')
                        """)
```

In [243]:
```
pd.read_sql(f"""
            SELECT * FROM titles
            WHERE isbn LIKE '{isbn_book}'
            """, connection)
```

Out[243]:

	isbn	title	edition	copyright
0	1593275994	Automate the Boring Stuff	1	2015