DATABASE SYSTEMS AND CLOUD COMPUTING

1st Homework Assignment

Due on: November 11, 2022

In this homework, you work with the IMDb database, and you are required to answer the following questions running the SQL statement in Python script the way we did in the class:

1. **Find the movies from 70s**.

```
from mysql import connector

my_connection = connector.connect(
  host="localhost",
  user="root",
  password="Secret_123",
  database="imdb"
)

my_cursor = my_connection.cursor()

"""
Find the movies from 70s.
"""
my_cursor.execute("""
  !!!ENTER YOUR SQL STATEMENT HERE!!!
""")

print("{:40} {:4}".format("Movie Title", "Year"))
for row in my_cursor:
  print(f"{row[0]:40} {row[1]:<4}")
```

Sample output:

```
Movie Title                              Year
Le magnifique                            1973
Dog Day Afternoon                        1975
Network                                  1976
The Little Girl Who Lives Down the Lane  1976
Der amerikanische Freund                 1977
The Last Wave                            1977
```

2. **Find the number of movies of each director and order in descending order**.

```
from mysql import connector

my_connection = connector.connect(
    host="localhost",
    user="root",
    password="Secret_123",
    database="imdb"
)

my_cursor = my_connection.cursor()

"""
Find the number of movies of each director.
"""
my_cursor.execute("""
    !!!ENTER YOUR SQL STATEMENT HERE!!!
""")

print("{:24} {:16}".format("Director Name", "Number of movies"))
for row in my_cursor:
    print(f"{row[0]:24} {row[1]:<16}")
```

Sample output:

```
Director Name            Number of movies
Ki-duk Kim               5
Niels Arden Oplev        3
Sidney Lumet             3
Billy Wilder             3
Alfred Hitchcock         3
Mira Nair                2
Grant Heslov             2
Chan-wook Park           2
Lasse Hallström          2
Jae-young Kwak           2
Francis Ford Coppola     2
Peter Jackson            2
```

3. Find the number of genres of each director's movies and order in descending order.

```
from mysql import connector

my_connection = connector.connect(
    host="localhost",
    user="root",
    password="Secret_123",
    database="imdb"
)

my_cursor = my_connection.cursor()

"""
Find the number of genres of each director's movies.
"""
my_cursor.execute("""
    !!!ENTER YOUR SQL STATEMENT HERE!!!
""")

print("{:24} {:16} {:100}".format("Director Name",
                                  "Number of movies", "Genres"))
for row in my_cursor:
    print(f"{row[0]:24} {row[1]:<16} {row[2]:<100}")
```

Sample output:

```
Director Name            Number of movies Genres
Billy Wilder             6                Comedy,Crime,Drama,Film-Noir,Thriller,War
Carlos Saldanha          6                Action,Adventure,Animation,Comedy,Family,Romance
David Yates              6                Action,Adventure,Family,Fantasy,Mystery,Romance
Guy Ritchie              6                Action,Adventure,Crime,Drama,Mystery,Thriller
John Sturges             6                Action,Adventure,Drama,Horror,Thriller,War
Peter Jackson            6                Crime,Drama,Fantasy,Horror,Romance,Thriller
Albert Hughes            5                Action,Adventure,Drama,Thriller,Western
Alfred Hitchcock         5                Crime,Horror,Mystery,Romance,Thriller
Allen Hughes             5                Action,Adventure,Drama,Thriller,Western
Chan-wook Park           5                Drama,Horror,Mystery,Romance,Thriller
Chris Wedge              5                Action,Adventure,Animation,Comedy,Family
Costas Ferris            5                Biography,Drama,History,Music,War
```

4. **Find the list of movies having the genres "Drama" and "Comedy" only.**

```
from mysql import connector

my_connection = connector.connect(
    host="localhost",
    user="root",
    password="Secret_123",
    database="imdb"
)

my_cursor = my_connection.cursor()

"""
Find the list of movies having the genres "Drama" and "Comedy" only.
"""
my_cursor.execute("""
    !!!ENTER YOUR SQL STATEMENT HERE!!!
""")

print("{:32} {:4} {:100}".format("Movie Title", "Year", "Genres"))
for row in my_cursor:
    print(f"{row[0]:32} {row[1]:4} {row[2]:<100}")
```

Sample output:

```
Movie Title                      Year Genres
Stalag 17                        1953 Comedy,Drama
Rembetiko                        1983 Drama,Drama
Bacheha-Ye aseman                1997 Comedy,Drama
Yeopgijeogin geunyeo             2001 Comedy,Drama
Jeux Denfants                    2003 Comedy,Drama
Broken English                   2007 Comedy,Drama
Ex Drummer                       2007 Comedy,Drama
Sonbahar                         2008 Comedy,Drama
Sunshine Cleaning                2008 Comedy,Drama
500 Days Of Summer               2009 Comedy,Drama
A Serious Man                    2009 Comedy,Drama
Away We Go                       2009 Comedy,Drama
```

5. **Find and list the histogram of movies where bins are genres.**

```
from mysql import connector

my_connection = connector.connect(
    host="localhost",
    user="root",
    password="Secret_123",
    database="imdb"
)

my_cursor = my_connection.cursor()

"""
Find and list the histogram of movies where bins are genres.
"""
my_cursor.execute("""
    !!!ENTER YOUR SQL STATEMENT HERE!!!
""")

print("{:32} {:4}".format("Genre", "Count"))
for row in my_cursor:
    print(f"{row[0]:32} {row[1]:<4}")
```

Sample output:

```
Genre                            Count
Drama                            180
Thriller                         70
Comedy                           56
Romance                          56
Crime                            41
Action                           36
Mystery                          33
Adventure                        26
Biography                        20
Horror                           19
War                              18
Sci-Fi                           15
```

IMPORTANT

- Academic dishonesty, including but not limited to cheating, plagiarism, and collaboration, is unacceptable and subject to disciplinary action. Any student found guilty will have a grade of F. Assignments are due in class on the due date. Late assignments will generally not be accepted. Any exception must be approved. Approved late assignments are subject to a grade penalty.