Data 607
Week Two - R and SQL

Part 1: Build Table
- Choose six recent popular movies.
- Ask at least five people that you know (friends, family, classmates, imaginary friends if necessary) to rate each of these movies that they have seen on a scale of 1 to 5.

Part 2: Store data in SQL database
- Take the results (observations) and store them in the class MySQL database:
  - Server name: cunydata607sql.mysql.database.azure.com
  - Username / password: will be given to you in an email

Note: it is good practice to change your password. To do so, use this SQL command:
SET PASSWORD = '<your new password here>';

Part 3: Transfer data from SQL database to R dataframe
- Load the information from the SQL database into an R dataframe.

Part 4: Missing data strategy
- Implement an approach to missing data
- Explain why you decided to take the chosen approach

Note: consider that later in the course you will revisit this information you have collected and will use it to implement a Recommender.

Bonus Challenge Questions:
You're encouraged to optionally find other ways to make your solution better. For example, consider incorporating one or more of the following suggestions into your solution:
- Use survey software to gather the information.
- Are you able to use a password without having to share the password with people who are viewing your code?
  - There are a lot of interesting approaches that you can uncover with a little bit of research.
- While it's acceptable to create a single SQL table, can you create a normalized set of tables that corresponds to the relationship between your movie viewing friends and the movies being rated?
- Is there any benefit in standardizing ratings? How might you approach this?