David Lawson Assignment 3 AIT 580 Summer 2017

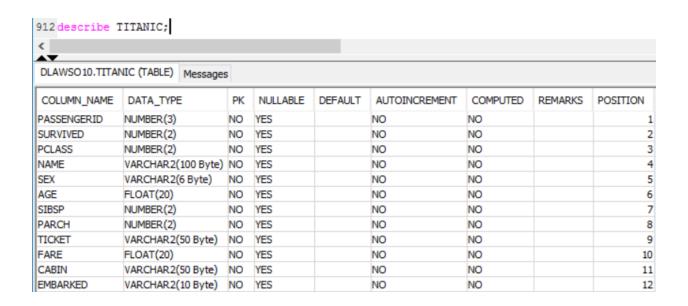
1. Describe the contents of the dataset. Briefly explain/speculate what knowledge can be gained by analyzing the data.

The dataset describes the passengers aboard the Titanic, once the largest ship in the world, and who survived it's infamous sinking. The dataset contains attributes on the passengers including age, sex, and socio-economic status. The Titanic notoriously had too few lifeboats to account for all of its passengers and many passengers did not survive it's sinking. This lack of lifeboats led to new laws and regulations around boating safety and the number of lifeboats required for each ship. By analyzing this dataset we can draw insights into how each passengers personal attributes contributed to their chance of surviving and if lifeboat priority was given to any particular grouping of passengers.

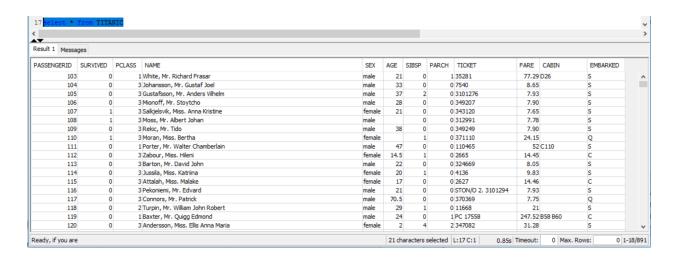
2. For each variable in the dataset, describe its type and how to indicate its data type in SQL.

```
3 CREATE TABLE TITANIC
   (PASSENGERID NUMBER (3,0),
   SURVIVED NUMBER (2,0),
   PCLASS NUMBER (2,0),
   NAME VARCHAR2 (100),
   SEX VARCHAR2 (6),
   AGE FLOAT (20),
0
   SIBSP NUMBER (2,0),
1
   PARCH NUMBER (2,0),
2
   TICKET VARCHAR2 (50),
3
   FARE FLOAT (20),
4
   CABIN VARCHAR2 (50),
5
   EMBARKED VARCHAR2 (10));
```

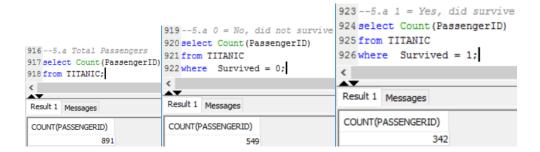
3. Create an SQL table to contain the data; use the SQL "DESCRIBE" command to display the table characteristics.



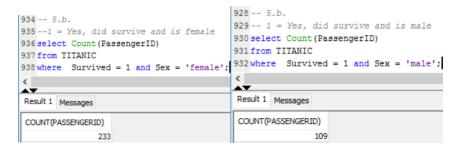
4. Insert rows of data into the table; display what you entered using the SQL "SELECT" command.



- 5. Write SQL commands that find the following:
  - a. how many total passengers are in the data set? how many died/survived?



b. how many men and how many women survived?



c. what is the average age of all passengers and of the survivors?



491

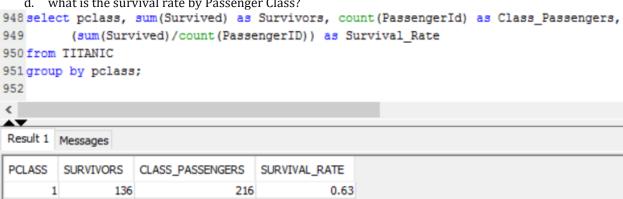
d. what is the survival rate by Passenger Class?

2

3

87

119



0.47

0.24