

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 its 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 its 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 passenger's 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
4   (PASSENGERID NUMBER(3,0),
5    SURVIVED NUMBER(2,0),
6    PCLASS NUMBER(2,0),
7    NAME VARCHAR2(100),
8    SEX VARCHAR2(6),
9    AGE FLOAT(20),
10   SIBSP NUMBER(2,0),
11   PARCH NUMBER(2,0),
12   TICKET VARCHAR2(50),
13   FARE FLOAT(20),
14   CABIN VARCHAR2(50),
15   EMBARKED VARCHAR2(10));
```

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

```
912 describe TITANIC;
```

DLAWSO10.TITANIC (TABLE) Messages								
COLUMN_NAME	DATA_TYPE	PK	NULLABLE	DEFAULT	AUTOINCREMENT	COMPUTED	REMARKS	POSITION
PASSENGERID	NUMBER(3)	NO	YES		NO	NO		1
SURVIVED	NUMBER(2)	NO	YES		NO	NO		2
PCLASS	NUMBER(2)	NO	YES		NO	NO		3
NAME	VARCHAR2(100 Byte)	NO	YES		NO	NO		4
SEX	VARCHAR2(6 Byte)	NO	YES		NO	NO		5
AGE	FLOAT(20)	NO	YES		NO	NO		6
SIBSP	NUMBER(2)	NO	YES		NO	NO		7
PARCH	NUMBER(2)	NO	YES		NO	NO		8
TICKET	VARCHAR2(50 Byte)	NO	YES		NO	NO		9
FARE	FLOAT(20)	NO	YES		NO	NO		10
CABIN	VARCHAR2(50 Byte)	NO	YES		NO	NO		11
EMBARKED	VARCHAR2(10 Byte)	NO	YES		NO	NO		12

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

Result 1 Messages											
PASSENGERID	SURVIVED	PCLASS	NAME	SEX	AGE	SIBSP	PARCH	TICKET	FARE	CABIN	EMBARKED
103	0	1	White, Mr. Richard Frasar	male	21	0	0	135281	77.29	D26	S
104	0	3	Johansson, Mr. Gustaf Joel	male	33	0	0	07540	8.65		S
105	0	3	Gustafsson, Mr. Anders Vilhelm	male	37	2	0	03101276	7.93		S
106	0	3	Mionoff, Mr. Stoytcho	male	28	0	0	0349207	7.90		S
107	1	3	Salkjelsvik, Miss. Anna Kristine	female	21	0	0	0343120	7.65		S
108	1	3	Moss, Mr. Albert Johan	male		0	0	0312991	7.78		S
109	0	3	Relic, Mr. Tido	male	38	0	0	0349249	7.90		S
110	1	3	Moran, Miss. Bertha	female		1	0	0371110	24.15		Q
111	0	1	Porter, Mr. Walter Chamberlain	male	47	0	0	0110465		52 C110	S
112	0	3	Zabour, Miss. Hileni	female	14.5	1	0	02665	14.45		C
113	0	3	Barton, Mr. David John	male	22	0	0	0324669	8.05		S
114	0	3	Jussila, Miss. Katrina	female	20	1	0	04136	9.83		S
115	0	3	Attalah, Miss. Malake	female	17	0	0	02627	14.46		C
116	0	3	Pekoniemi, Mr. Edvard	male	21	0	0	05TON/O 2. 3101294	7.93		S
117	0	3	Connors, Mr. Patrick	male	70.5	0	0	0370369	7.75		Q
118	0	2	Turpin, Mr. William John Robert	male	29	1	0	011668	21		S
119	0	1	Baxter, Mr. Quigg Edmond	male	24	0	0	1PC 17558	247.52	B58 B60	C
120	0	3	Andersson, Miss. Ellis Anna Maria	female	2	4	0	2347082	31.28		S

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

916--5.a Total Passengers	919--5.a 0 = No, did not survive	923--5.a 1 = Yes, did survive
917 select Count(PassengerID)	920 select Count(PassengerID)	924 select Count(PassengerID)
918 from TITANIC;	921 from TITANIC	925 from TITANIC
	922 where Survived = 0;	926 where Survived = 1;
Result 1 Messages	Result 1 Messages	Result 1 Messages
COUNT(PASSENGERID)	COUNT(PASSENGERID)	COUNT(PASSENGERID)
891	549	342

b. how many men and how many women survived?

934-- 5.b.	928-- 5.b.
935--1 = Yes, did survive and is female	929-- 1 = Yes, did survive and is male
936select Count(PassengerID)	930select Count(PassengerID)
937from TITANIC	931from TITANIC
938where Survived = 1 and Sex = 'female';	932where Survived = 1 and Sex = 'male';
< [SQL Editor]	< [SQL Editor]
Result 1 Messages	Result 1 Messages
COUNT(PASSENGERID)	COUNT(PASSENGERID)
233	109

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

938-- 5.c Average age of all passengers	942-- 5.c Average age of survivors
939select AVG(Age)	943select AVG(Age)
940from TITANIC;	944from TITANIC
	945where Survived = 1;
	946
< [SQL Editor]	< [SQL Editor]
Result 1 Messages	Result 1 Messages
AVG(AGE)	AVG(AGE)
29.70	28.34

d. what is the survival rate by Passenger Class?

948select pclass, sum(Survived) as Survivors, count(PassengerId) as Class_Passengers,	
949 (sum(Survived)/count(PassengerID)) as Survival_Rate	
950from TITANIC	
951group by pclass;	
952	
< [SQL Editor]	
Result 1 Messages	
PCLASS	SURVIVORS
1	136
2	87
3	119
CLASS_PASSENGERS	SURVIVAL_RATE
216	0.63
184	0.47
491	0.24