# Project Title: Titanic Dataset Analysis and Case Study
**Description:**

Utilizing SQL, this project offers a deep dive into the Titanic dataset, uncovering insights on passenger demographics and survival rates. The analysis includes data exploration, visualization, and a detailed case study to provide a comprehensive understanding of the factors influencing survival.

```sql
1   SELECT * FROM titanic_project.titanic;
2   use titanic_project
3   -- 1) show the first 10 rows of the dataset
4   SELECT * from titanic LIMIT 10
5   SELECT * from titanic order by PassengerId desc LIMIT 10
6   -- 2) find the num of passengers who survived
7   SELECT * from titanic
8   SELECT count(Survived) as "Passengers_Survived" from titanic where Survived = 1
9   -- 3) find the average age of all passengers.
10  SELECT avg(Age) as "Avg_Age" from titanic
11  -- 4) find the num of passengers in each class
12  SELECT Pclass,count(*) as "Num of passengers" from titanic group by Pclass
13  -- 5) find the first 10 rows of the dataset sorted by passenger class in descending order
14  SELECT * from titanic order by Pclass desc limit 10
15  -- 6) find the num of passengers in each class sorted by class in ascending order
16  SELECT * from titanic
17  SELECT Pclass ,count(*) from titanic group by Pclass order by Pclass asc
18  -- 7) find the avg fare paid by passenger in each class
19  SELECT Pclass,avg(Fare) as "Avg_Fare" from titanic group by Pclass
20  -- 8) find the name of the passenger with the highest fare.
21  SELECT name,Fare from titanic where fare = (select max(Fare) from titanic)
22  SELECT max(Fare) from titanic
```

```sql
23  -- 9) find the name of the passenger who had the highest age among the survivors.
24  SELECT name,age from titanic where age = (SELECT max(age) from titanic where survived = 1)
25  and Survived = 1
26  --SELECT max(age) from titanic where survived = 1
27  -- 10) Find the num of passengers who paid more than the avg fare
28  SELECT * from titanic
29  SELECT count(*) as num_Passengers from titanic where Fare > (select avg(Fare) from titanic); -- 172
30  -- 11) Find the name of the passengers woh had most parents or children on board
31  SELECT Name from titanic where Parch = (SELECT max(Parch) from titanic);
32  -- SELECT max(Parch) from titanic
33  -- 12) Find the num of Male and Female passengers who survived
34  -- and order the results by sex in ascending order
35  SELECT Sex,count(*) as num_survived from titanic where Survived = 1
36  group by Sex
37  order by Sex asc;
38  -- 13) Find the name,age and fare of the oldest passenger who survived
39  SELECT Name,Age,Fare from titanic where Age = (SELECT  max(Age) from titanic where Survived = 1)
40  and Survived = 1
41  -- SELECT  max(Age) from titanic where Survived = 1
42  -- 14) Find the name and age of the youngest female passenger who survived in 3rd class
43  SELECT name,Age from titanic where Sex = "female" and Pclass = 3 and Survived = 1
44  and age = (SELECT min(Age) from titanic where Sex = "female" and Pclass = 3 and Survived = 1);
45  -- SELECT min(Age) from titanic where Sex = "female" and Pclass = 3 and Survived = 1
46  -- 15) Find the Number of male and female passengers
47  SELECT
48  sum(case when Sex = "male" then 1 else 0 end) as num_male,
49  sum(case when Sex = "female" then 1 else 0 end) as num_female
50  from titanic
51  -- 16) Select all passengers who traveled ina cabin that was not shared  by other passengers
52  SELECT * from titanic where Cabin not in (SELECT Cabin from titanic group by Cabin having count(*) > 1)
```

Source Code:

```sql
SELECT * FROM titanic_project.titanic;

use titanic_project

-- 1) show the first 10 rows of the dataset

SELECT * from titanic LIMIT 10

SELECT * from titanic order by PassengerId desc LIMIT 10

 -- 2) find the num of passengers who survived

 SELECT * from titanic

 SELECT count(Survived) as "Passengers_Survived" from titanic where Survived = 1

 -- 3) find the average age of all passengers.

 SELECT avg(Age) as "Avg_Age" from titanic

 -- 4) find the num of passengers in each class

 SELECT Pclass,count(*) as "Num of passengers" from titanic group by Pclass

 -- 5) find the first 10 rows of the dataset sorted by passenger class in descending order

 SELECT * from titanic order by Pclass desc limit 10

 -- 6) find the num of passengers in each class sorted by class in ascending order

 SELECT * from titanic

 SELECT Pclass ,count(*) from titanic group by Pclass order by Pclass asc

 -- 7) find the avg fare paid by passenger in each class

 SELECT Pclass,avg(Fare) as "Avg_Fare" from titanic group by Pclass

 -- 8) find the name of the passenger with the highest fare.

 SELECT name,Fare from titanic where fare = (select max(Fare) from titanic)

 SELECT max(Fare) from titanic

 -- 9) find the name of the passenger who had the highest age among the survivors.

 SELECT name,age from titanic where age = (SELECT max(age) from titanic where
survived = 1)

 and Survived = 1

 --SELECT max(age) from titanic where survived = 1

 -- 10) Find the num of passengers who paid more than the avg fare

 SELECT * from titanic
```

```sql
 SELECT count(*) as num_Passengers from titanic where Fare > (select avg(Fare) from titanic); -- 172
 -- 11) Find the name of the passengers woh had most parents or children on board
SELECT Name from titanic where Parch = (SELECT max(Parch) from titanic);
-- SELECT max(Parch) from titanic
-- 12) Find the num of Male and Female passengers who survived
-- and order the results by sex in ascending order
SELECT Sex,count(*) as num_survived from titanic where Survived = 1
group by Sex
order by Sex asc;
-- 13) Find the name,age and fare of the oldest passenger who survived
SELECT Name,Age,Fare from titanic where Age = (SELECT  max(Age) from titanic where Survived = 1)
and Survived = 1
-- SELECT  max(Age) from titanic where Survived = 1
-- 14) Find the name and age of the youngest female passenger who survived in 3rd class
SELECT name,Age from titanic where Sex = "female" and Pclass = 3 and Survived = 1
and age = (SELECT min(Age) from titanic where Sex = "female" and Pclass = 3 and Survived = 1);
-- SELECT min(Age) from titanic where Sex = "female" and Pclass = 3 and Survived = 1
-- 15) Find the Number of male and female passengers
SELECT
sum(case when Sex = "male" then 1 else 0 end) as num_male,
sum(case when Sex = "female" then 1 else 0 end) as num_female
from titanic
-- 16) Select all passengers who traveled ina cabin that was not shared  by other passengers
 SELECT * from titanic where Cabin not in (SELECT Cabin from titanic group by Cabin having count(*) > 1)
-- Cabin from titanic group by Cabin having count(*) > 1
```