

# Project 1 - Exploring Titanic Database

## Instructions

1. Answer the below question in the boxes.
2. Please submit the assignment through TalentLabs Learning System.

## Project Goal

In this project, you are going to learn about the Data Exploration framework, and practise the framework with the open-sourced titanic database (the most popular database in data science world).

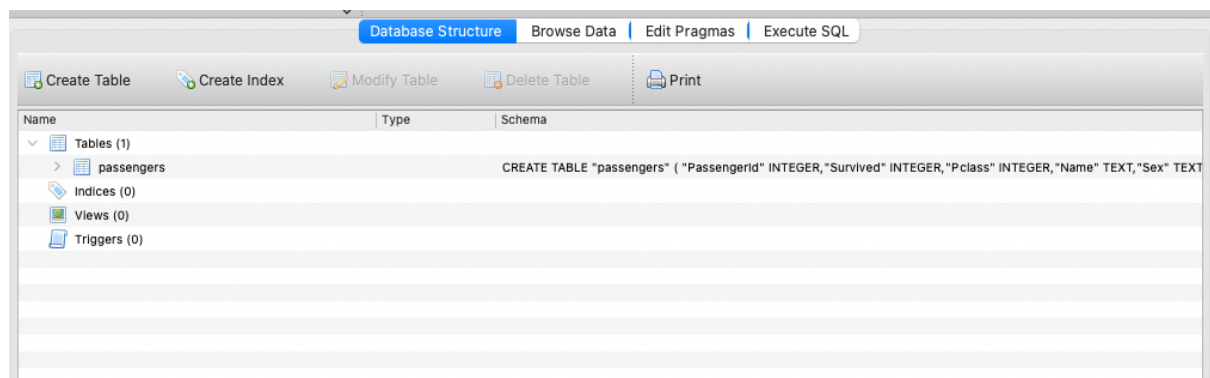
The end goal of this project is to create a data exploration report which answers a few questions that you came up with, in a data-driven way.

## Part 1 - Data Exploration Framework

In this part, you are going to do some read-up on the data exploration framework and understand how to perform data exploration for a novel dataset. Make sure you read through the “Project 1 Reading - Data Exploration Framework” before you proceed.

## Part 2 - Download and Open Up the Titanic Database

You should be able to find the titanic database included in the project folder. Follow the step illustrated in Chapter 3 to open the database using DB Browser for SQLite. You should see only 1 table in the database.



## Part 3 - Data Exploration

The task for this part is to write up a report for each steps in Data Exploration framework (no specific format, could be a powerpoint/slides, or word document)

In the final part of free exploration, you can try with the guiding questions below, or you can think of questions that you want to answer with data (or a mix of both). At least explore **3 questions**. In the report, you should include the steps and SQL you used to explore that question, as well as the conclusion you have.

### Data Schema

You can find the table and field definition here: <https://www.kaggle.com/c/titanic/data>

### Guiding questions for free exploration:

- In the movie, children, elderlies and females can get onboarded to rescue boat first, so
  - Are children and elderlies have a higher survival rate in this accident?
  - Are females more likely to survive in this incident?
- Are rich people have a higher survival rate because they can get onboard to the rescue boat sooner (like what is shown in the movie)?

Feel free to come up with other questions that you want to answer with the data!

**- End of Project -**