

# SQL Project: Analysis of Titanic Dataset

Presented by Anurag Tripathi

# Introduction

---

This project presents an in-depth analysis of the Titanic dataset using Structured Query Language (SQL). The Titanic dataset is one of the most well-known datasets in data science and contains information about the passengers who were on board the Titanic during its tragic voyage in 1912.

The goal of this project is to:

- Extract meaningful insights from the data
- Perform data cleaning, filtering, and aggregation
- Understand survival patterns based on gender, age, class, and other factors
- Demonstrate the use of SQL queries in real-world data analysis

Through this project, we explore how SQL can help uncover trends and patterns in large datasets and support data-driven decision-making.

# How many passengers were in each class (Pclass)?

---

1

```
4 • select Pclass,  
5           count(Pclass) as Total_Passanger  
6           from sql_project.`titanic-dataset`  
7           group by Pclass ;  
8
```

2

Result Grid | Filter Rows:

	Pclass	Total_Passanger
▶	3	355
	1	186
	2	173



# How many passengers were on the Titanic?

---

1

```
select Count(*)  
      As Total_passanger_on_Titanic  
from sql_project.`titanic-dataset` ;
```

2

	Total_passanger_on_Titanic
▶	714

# What is the average age of the passengers

1

Select

```
    AVG(age) AS TOTAL_AVERAGE_OF AGE  
from sql_project.`titanic-dataset` ;
```

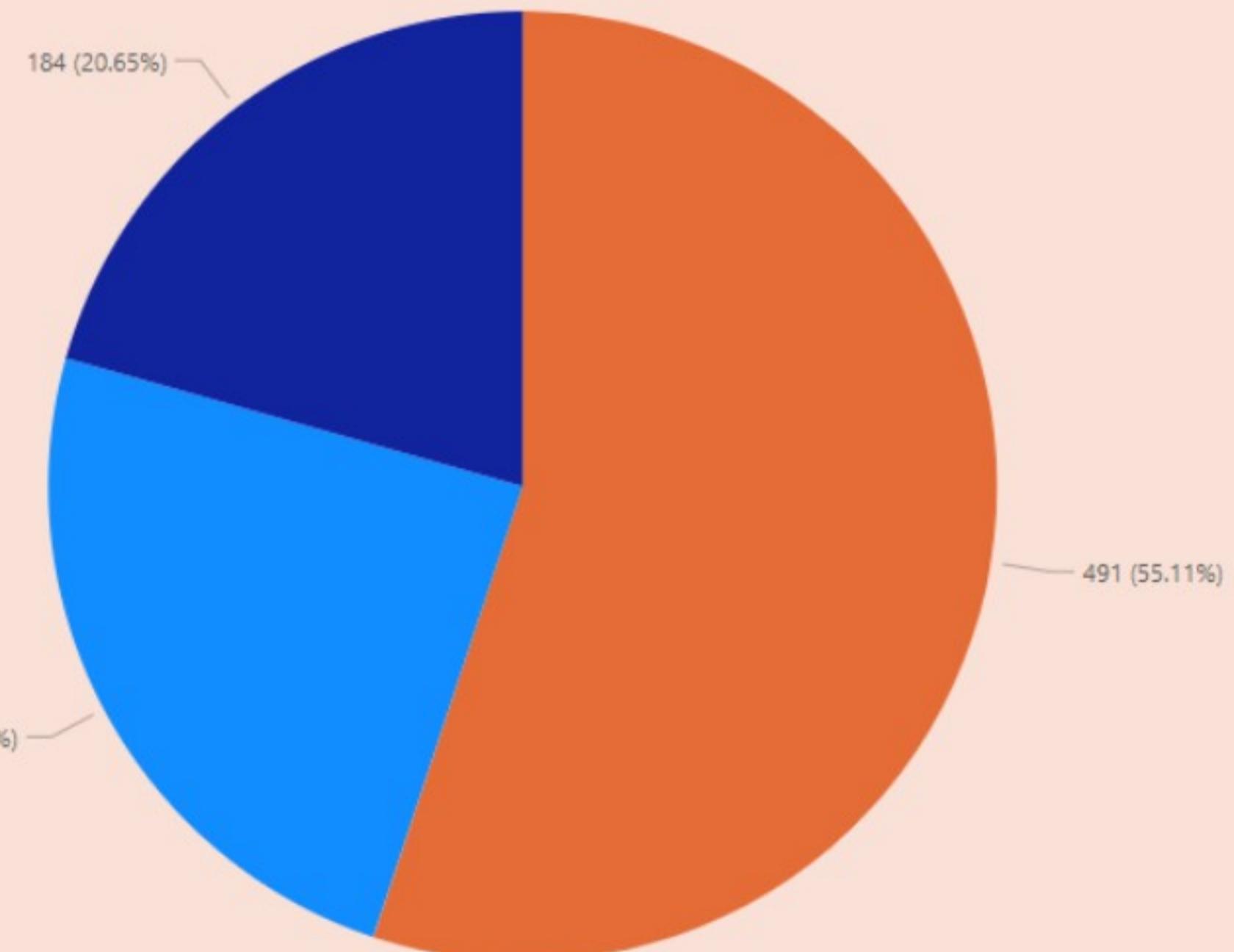
2

| Result Grid | Filter Row

	TOTAL_AVERAGE_OF AGE
▶	29.7129

# Visualizations

## Passenger Count by Class (Pclass)



891

Count of PassengerId

29.70

Average of Age

# Thank you!

