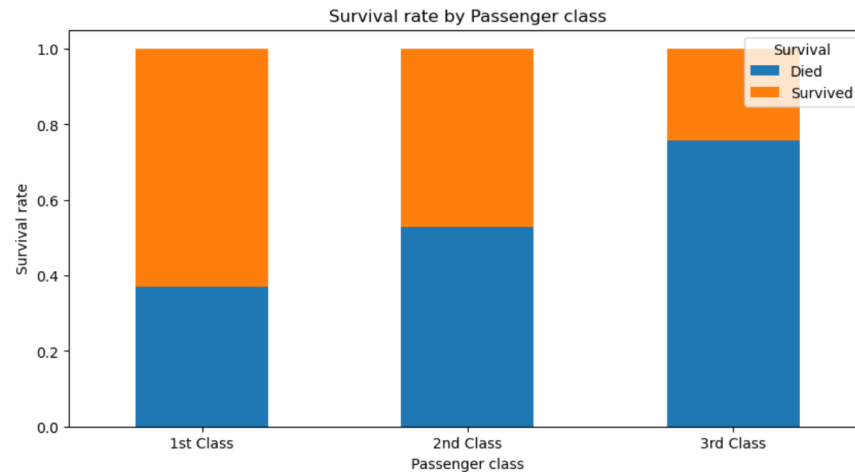# Assignment 3 - part 2

Name of dataset: Titanic Data Set   Source of dataset: Kaggle

1. **Determine if the survival rate is associated to the class of passenger.**
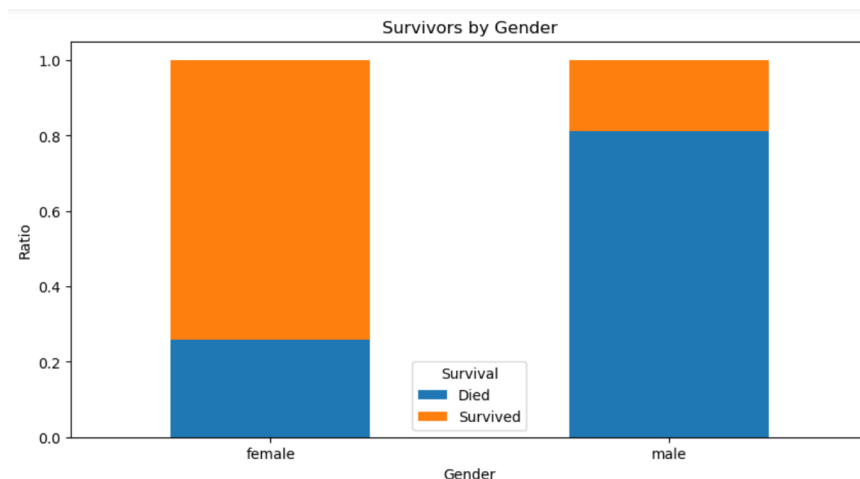

Survival rate by Passenger class

⇨ Passengers in 1st class had higher survival rates than of lower class.

Chi-square test:

- **Null hypothesis**: Passenger class is independent with Survival Rate. In other words, there is no association between Passenger class and Survival rate.
- **Alternative hypothesis**: Passenger class is not independent with Survival Rate. In other words, there is an association between Passenger class and Survival rate.
- ⇨ P-value = 4.549e-23 < α (=0.05) → Reject H0 → There is an association between passenger class and their survival rate

2. **Determine if the survival rate is associated to the gender.**

Survivors by Gender

Phan Hoang An Nguyen
PDD Data Analytics
Langara College

⇨ Female passengers had much higher survival rate than their male counterparts.

Chi-square test:

- **Null hypothesis**: Gender is independent with Survival Rate. In other words, there is no association between Gender and Survival rate.
- **Alternative hypothesis**: Gender is not independent with Survival Rate. In other words, there is an association between Gender and Survival rate.

⇨ P-value = 1.19e-58 < α (=0.05) → Reject H0 → There is an association between genders and their survival rate.

3. **Determine the survival rate is associated to the age.**
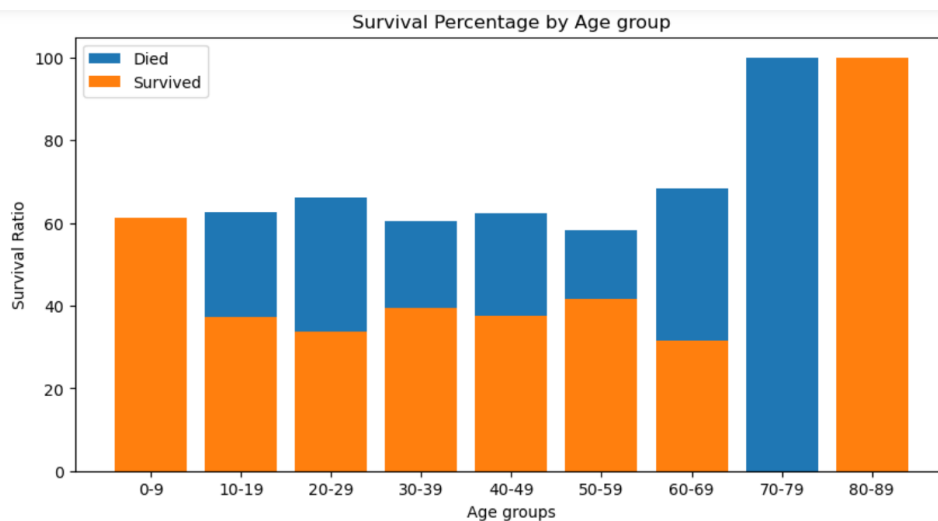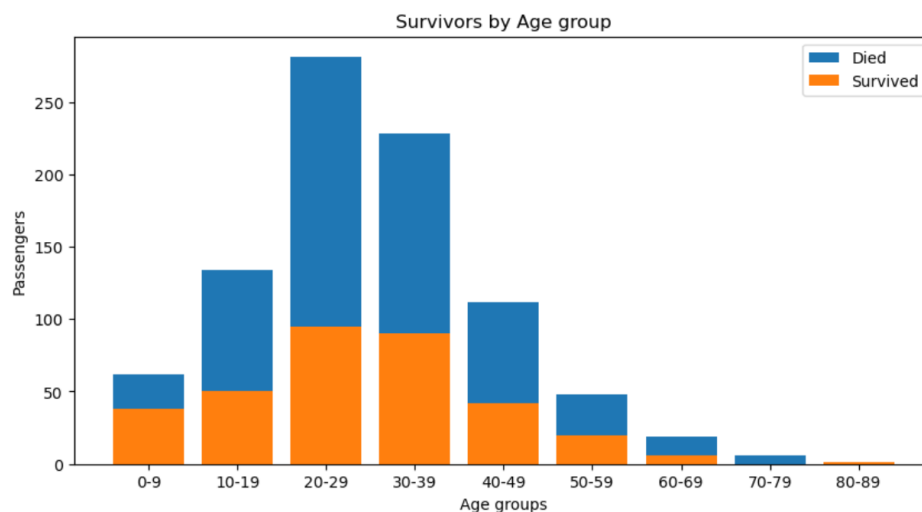
⇨ There is an association between age groups and survival rate because old people and children had higher survival ratio. Male passengers in general had higher age & lower survival rate than female passengers.