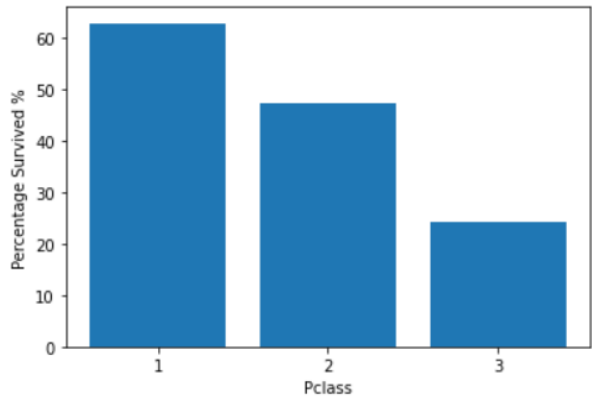
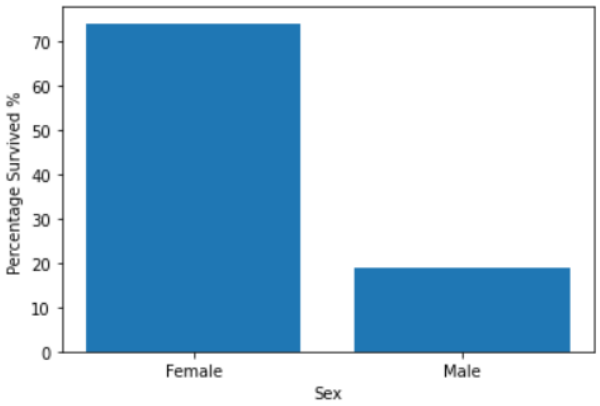
Titanic Dataset Hypothesis

If survival rate is associated with the *class of passengers*.



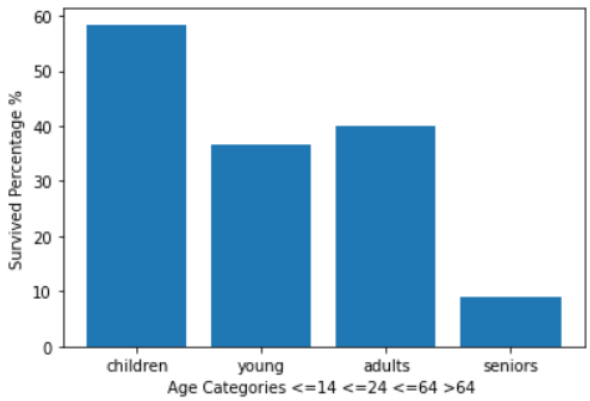
The percentage of those survived is dropping with the ticket class. To identify if there is dependency between the two categorical variables (Survived , Pclass), chi- square test is used. Chi-square test result returned a p-value of 4.549251711298793e-23 , which is less then 0.05. This indicates that there is dependency between Pclass and survival. Statistically, There is association between the two.

If survival rate is associated with *gender* :



Female had a very high survival rate compared to male. Chi-square test result returned a p-value of 1.1973570627755645e-58, which is less then 0.05. This indicates that null hypothesis can be rejected. Statistically, The survival rate is associated with gender.

If Survival rate is associated with *age*:



Divided the age into 4 categories (<=14, <=24 , <=64 , >64) . Children had a high survival rate of ~60%, whereas seniors (>64) had a very survival rate of ~10%. Chi-square test result returned a p-value of 0.0010709216002096906 which is less then 0.05. This indicates that the null hypothesis can be rejected. Statistically, Age has association with survival rate.