# 

# 

# Titanic data exploratory

Institution:

Affiliation:

Name:

Instructor:

Date:

1. How many male passengers were on the Titanic? Female passengers? Please show the R code.

**code**:glimpse(titanic)

table(etitanic$sex)

**Result**: female male

388 658

2. What was the survival rate for male passengers? Female passengers? Please show the R code.

**Code** :ServeR <- table(etitanic$survived, etitanic$sex)

prop.table(ServeR)

**Result** :female male

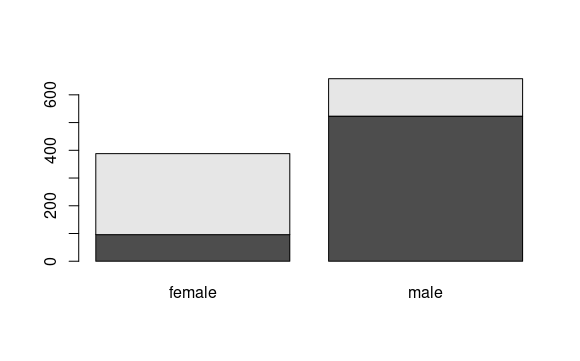
0 0.0917782 0.5000000

1 0.2791587 0.1290631

3. Visualize survival rates by gender. Please show the R code.

**code**:barplot(ServeR)

**Result**:



4. The above are two competing models, without any test data can you comment on which model is better? Use model params/output metrics to support

**I think model 1 is better because it provides us with detailed results about our data set which means this model can be valid to be relied on when being used because it performs modeling for all the variables rather than model 2 with performs modeling for 2 .The first model does not specify the variable on which the model is supposed to train whereas the second model specifies variable Sex to perform training on .from the results we can see differences for example model 1 has rank 7 which means performed modeling on 7 variables whereas model two has rank two because sex has only male and female**