hw1_spp2122

Sharleen Price

Part 1

i)

```
titanic <-read.table("Titanic.txt", header = TRUE, as.is = TRUE)
```

ii)

```
dim(titanic)
```

```
## [1] 891  12
```

iii)

```
titanic$Survived.Word <- ifelse(titanic$Survived == 1, "Survived", "Died")
```

Part 2

i)

```
mat <-titanic[,c("Survived","Age","Fare")]
apply(mat, 2, mean)
```

```
##   Survived        Age       Fare
##  0.3838384         NA 32.2042080
```

The survived mean tells us that only 38% of people survived the Titanic. Age says NA because not all ages of all people on the titanic is known.

ii)

```
survivors = titanic[titanic$Survived == 1,]
females_survivors = survivors[survivors$Sex=="female",]
round(nrow(females_survivors)/nrow(titanic),2)
```

```
## [1] 0.26
```

iii)

```
round(nrow(females_survivors)/nrow(survivors),2)
```

```
## [1] 0.68
```

iv)

```
classes <- sort(unique(titanic$Pclass))
Pclass.Survival <- vector("numeric", length = 3)
names(Pclass.Survival) <- classes
for (i in 1:3) {
 Pclass.Survival[i]<-round((nrow(survivors[survivors$Pclass==i,])/nrow(titanic[titanic$Pclass==i,])),2)

}

Pclass.Survival
```

```
##    1    2    3
## 0.63 0.47 0.24
```

v)

```
classes2 <- sort(unique(titanic$Pclass))
Pclass.Survival2 <- vector("numeric", length = 3)
names(Pclass.Survival2) <- classes

Pclass.Survival2<-round((tapply(titanic$Survived,titanic$Pclass,mean)),2)


Pclass.Survival2
```

```
##    1    2    3
## 0.63 0.47 0.24
```

vi) Yes, there is a relationship between survival rate and class. It seems that the survival rate increases as the class increased and so high class had a higher rate of survival than the lower class.