Extra Credit Problem Set

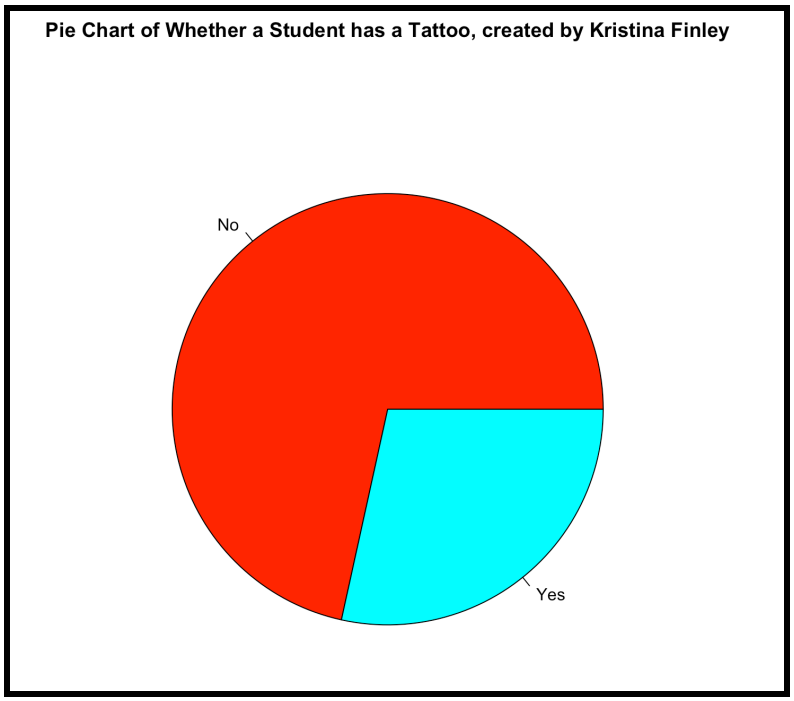
Kristina Finley
STAT 100, SECTION 0221

PROBLEM #1

```
> #1 (a) - extract a variable from a data frame, by Kristina Finley
> Tattoos <- Course_Data_Set$Tatoos
> #create a frequency table, by Kristina Finley
> Tattoos_Freq_Table <- table(Tattoos)
> Tattoos_Freq_Table
Tattoos
 No Yes
986 392
> # create proportion table, by Kristina Finley
> Tattoos_Prop_Table <- prop.table(Tattoos_Freq_Table)
> Tattoos_Prop_Table
Tattoos
       No       Yes
0.7155298 0.2844702
>
```

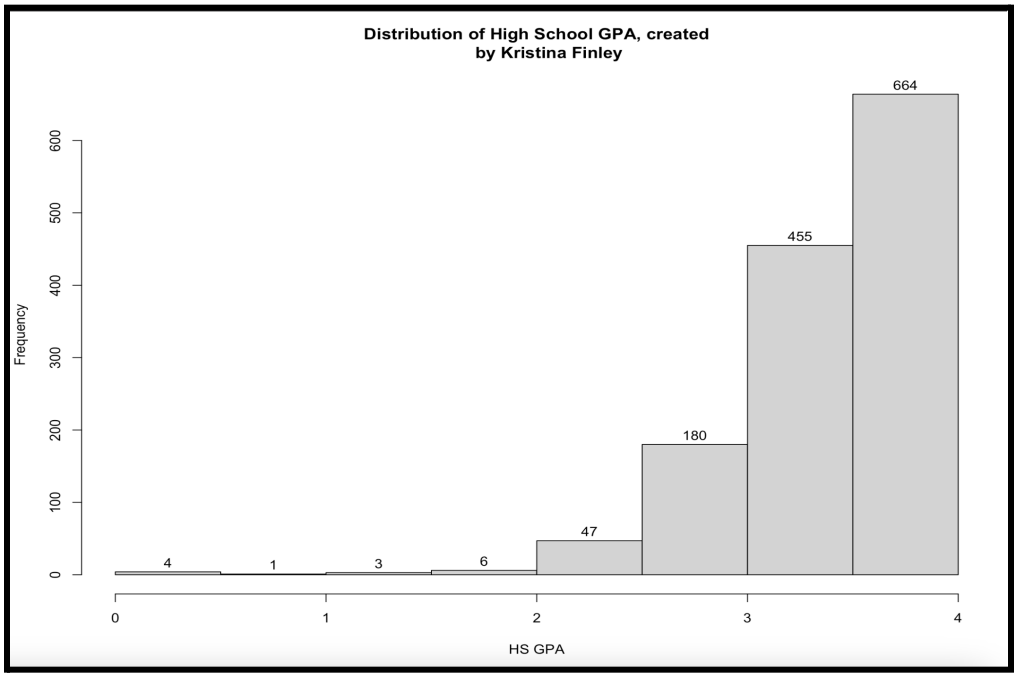
B.

```
> #1 (b) - create a pie chart by a frequency table, by Kristina Finley
> pie(x= Tattoos_Freq_Table, col=rainbow(length(Tattoos_Freq_Table)), main=
+       "Pie Chart of Whether a Student has a Tattoo, created by Kristina Finley")
```



PROBLEM #2

```
> #2(a) - extract variable from a data frame, by Kristina Finley
> HS_GPA <- Course_Data_Set$HS_GPA
> #generate a histogram for variable HS_GPA with default number of bins, by Kristina Finley
> hist(x=HS_GPA,right= FALSE, labels= TRUE, main= "Distribution of High School GPA, created
+      by Kristina Finley", xlab= "HS GPA")
```



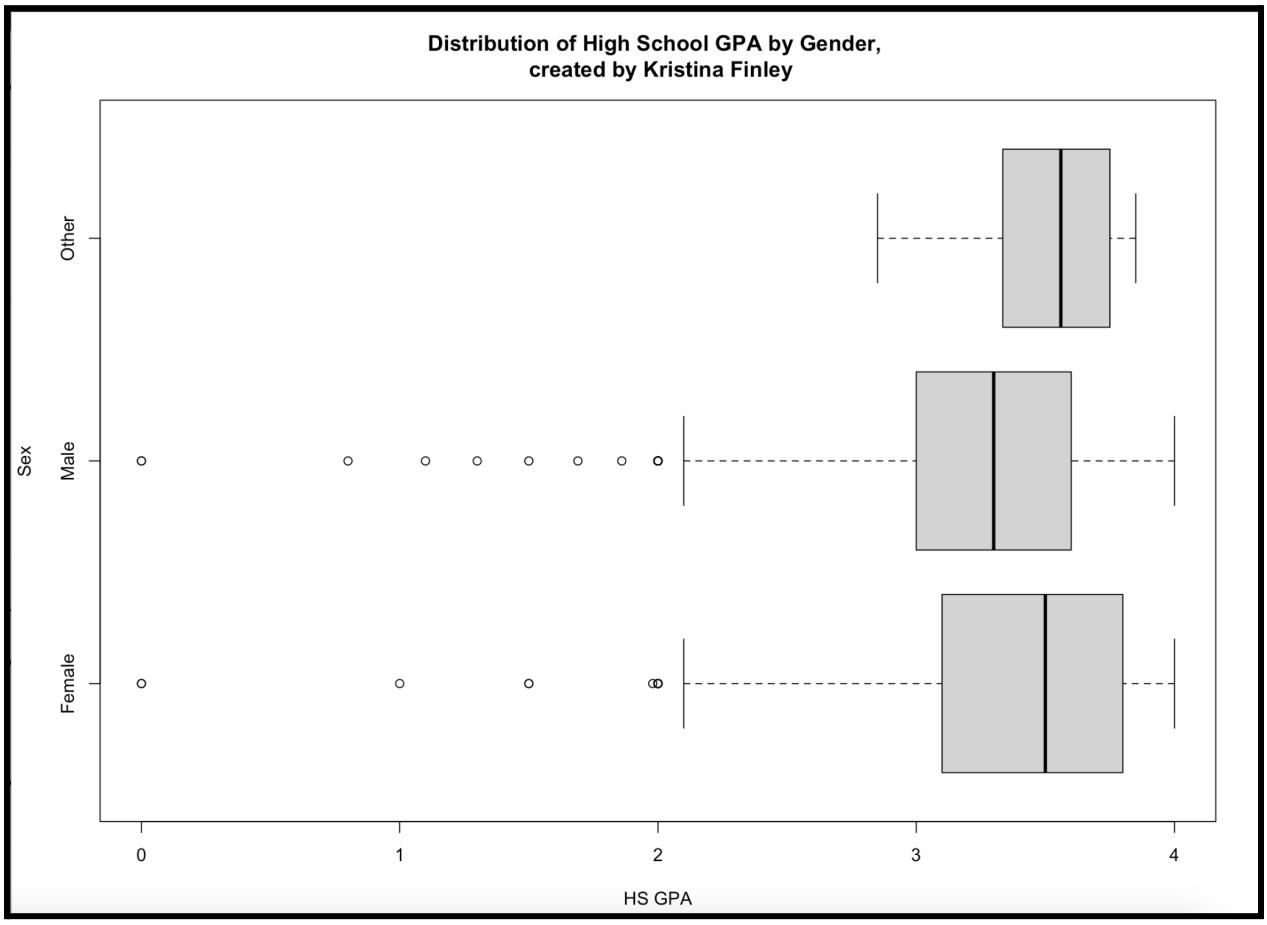
B.

4+1+3+6+47+180+455+664 = 680

680 falls between 3rd and 4th interval on the histogram.

PROBLEM #3

```
> #3(a) - extract variable from a data frame, by Kristina Finley
> Sex <- Course_Data_Set$Sex
> HS_GPA <- Course_Data_Set$HS_GPA
> #create the side-by-side boxplots of response variable by predictor variable,
> #by Kristina Finley
> boxplot(formula = HS_GPA ~ Sex, main= "Distribution of High School GPA by Gender, \n created by Kristina Finley",
+         xlab= "HS GPA", ylab="Sex", horizontal = TRUE)
>
```



B.

Looking at the side-by-side boxplots, the male median is about 3.25, and the female median is about 3.5.

PROBLEM #4

```
> #4(a) - define a variable vector with all possible values, by Kristina Finley
> X_values <- c(0,3,6,9,12)
> X_values
[1]  0  3  6  9 12
> # calculate the corresponding probabilities by using the formula, by Kristina Finley
> X_probs <- (27 - X_values) / 105
> X_probs
[1] 0.2571429 0.2285714 0.2000000 0.1714286 0.1428571
> # create probability distribution table, by Kristina Finley
> X_table <- data.frame(X_values,X_probs)
> names(X_table) <- c("x", "Pr(X = x)")
> print(X_table, row.names=FALSE)
  x Pr(X = x)
  0 0.2571429
  3 0.2285714
  6 0.2000000
  9 0.1714286
 12 0.1428571
>
```

B.

Using the numbers for the probability distribution table

0(0.257) + 3(0.229) + 6(0.200) + 9(0.171) + 12(0.143) = **5.142**

PROBLEM #5

```
> #5(a) - fdetermine probability/percentile for a normally distributed variable, by Kristina Finley
> round(100 * (pnorm(670,mean = 528, sd = 120, lower.tail = TRUE)),0)
[1] 88
>
```

B.
P(X < 510)
Z-Score: (510-528) / 120 = -0.15
P(Z < -0.15) = 0.4404 → **44%**
A score of 510 would be in the 44th percentile.