

Graphs

Alexander Lee

December 6, 2015

Creates the charts for visualization of the data. Each chart shows the frequency of each grade (A+, A, A-, etc.) and lines indicating the average grade for that class. Classes are arranged by teacher.

#re-import the data and whatnot so we can knit the .Rmd

```
library(readr)
setwd("../Data")
```

```
stat2010f <- read_csv("Grades2010f.csv")
stat2011f <- read_csv("Grades2011f.csv")
stat2012f <- read_csv("Grades2012f.csv")
stat2013f <- read_csv("Grades2013f.csv")
stat2014f <- read_csv("Grades2014f.csv")
```

```
stat2010s <- read_csv("Grades2010s.csv")
stat2011s <- read_csv("Grades2011s.csv")
stat2012s <- read_csv("Grades2012s.csv")
stat2013s <- read_csv("Grades2013s.csv")
stat2014s <- read_csv("Grades2014s.csv")
stat2015s <- read_csv("Grades2015s.csv")
```

#fall 2012

```
DeborahNolan2012f <- stat2012f[stat2012f$`Instructor Name` == "Deborah
Nolan",]
FletcherIbser2012f <- stat2012f[stat2012f$`Instructor Name` == "Fletcher
Ibser",]
```

#fall 2013

```
Ingileif2013f <- stat2013f[stat2013f$`Instructor Name` == "Ingileif
Hallgrimsdottir",]
FletcherIbser2013f <- stat2013f[stat2013f$`Instructor Name` == "Fletcher
Ibser",]
```

#fall 2014

```
BenoitDherin2014f <- stat2014f[stat2014f$`Instructor Name` == "Benoit
Dherin",]
FletcherIbser2014f <- stat2014f[stat2014f$`Instructor Name` == "Fletcher
Ibser",]
```

```
DeborahNolan2012s <- stat2012s[stat2012s$`Instructor Name` == "Deborah  
Nolan",]  
FletcherIbser2012s <- stat2012s[stat2012s$`Instructor Name` == "Fletcher  
Ibser",]
```

```
BenoitDherin2014s <- stat2014s[stat2014s$`Instructor Name` == "Benoit Dherin",]  
CariKaufman2014s <- stat2014s[stat2014s$`Instructor Name` == "Cari Kaufman",]
```

```
DeborahNolan2015s <- stat2015s[stat2015s$`Instructor Name` == "Deborah  
Nolan",]  
Ingileif2015s <- stat2015s[stat2015s$`Instructor Name` == "Ingileif  
Hallgrimsdottir",]
```

```
gpa <- c(12, 12, 11.1, 9.9, 9, 8.1, 6.9, 6, 5.1, 3.9, 3, 2.1, 0)
grade_names <- c("A+", "A", "A-", "B+", "B", "B-",
                  "C+", "C", "C-", "D+", "D", "D-", "F")
names(gpa) <- grade_names
gpa <- gpa / 3
```

#loopdy loop - find the total grade distribution for all classes

[illegible]

```
#create a vector with the grade relative frequencies for sampling purposes
grade_samples <- rep(grade_names, times = total_sum)
```

Data Visualization

Cari Kaufman

Cari Kaufman Fall 2010

```
#stat2010f

barplot(height = stat2010f$`Enrollment Cnt`, names.arg = stat2010f$`Grade
Nm`,
        main = "Cari Kaufman Fall 2010 Grade Distribution",
        xlab = "Grade Name", ylab = "Frequency", las = 1)

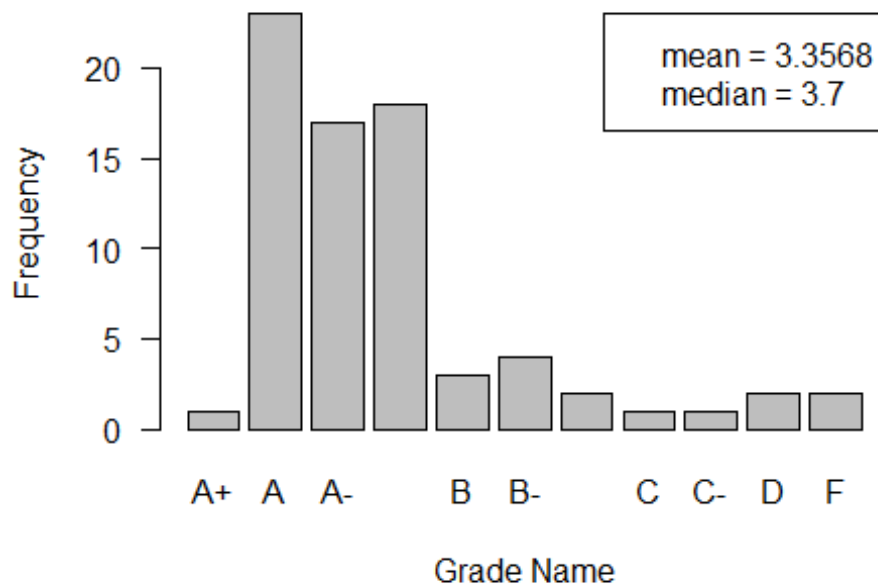
#create the GPA distribution of the class based on its
#`Enrollment Cnt`'s, used to find the mean/median
grade_samples <- rep(stat2010f$`Average Grade`, times = stat2010f$`Enrollment
Cnt`)

#mean and median
average_grade <- mean(grade_samples)
median_grade <- median(grade_samples)

#round the mean and median
average_grade <- round(average_grade, 4)
median_grade <- round(median_grade, 4)

#create a legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                             paste0("median = ",
as.character(median_grade))))
```

Cari Kaufman Fall 2010 Grade Distribution



Cari Kaufman Fall 2011

```
#stat2011f
```

```
barplot(height = stat2011f$`Enrollment Cnt`, names.arg = stat2011f$`Grade  
Nm`,
```

```
      main = "Cari Kaufman Fall 2011 Grade Distribution",  
      xlab = "Grade Name", ylab = "Frequency", las = 1)
```

```
#create the GPA distribution of the class based on its
```

```
#`Enrollment Cnt`'s, used to find the mean/median
```

```
grade_samples <- rep(stat2011f$`Average Grade`, times = stat2011f$`Enrollment  
Cnt`)
```

```
#mean and median
```

```
average_grade <- mean(grade_samples)
```

```
median_grade <- median(grade_samples)
```

```
#round the mean and median
```

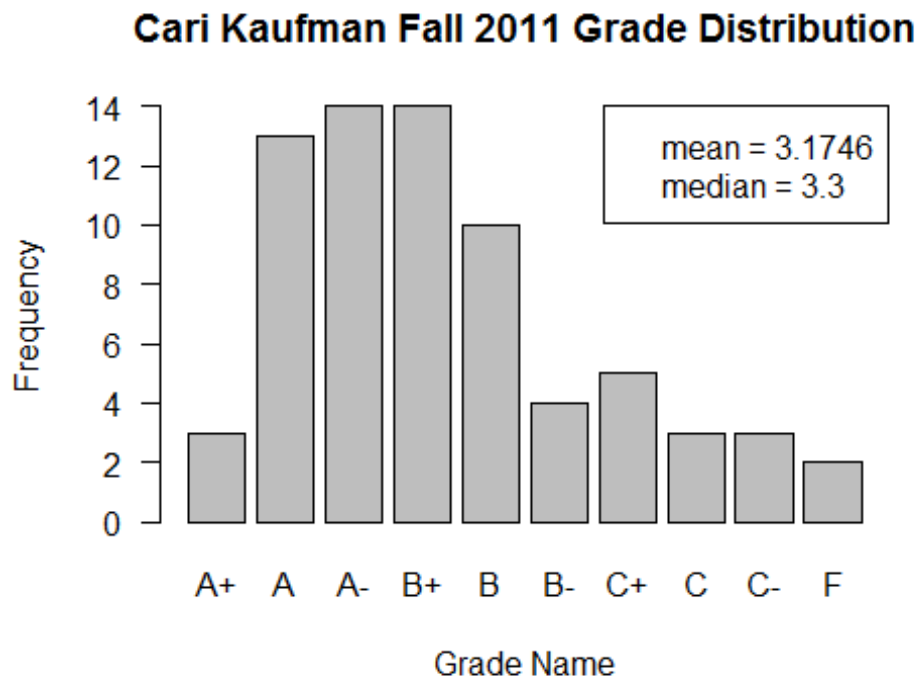
```
average_grade <- round(average_grade, 4)
```

```
median_grade <- round(median_grade, 4)
```

```
#create a legend that says what the lines mean
```

```
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
```

```
paste0("median = ",
as.character(median_grade))))
```



Cari Kaufman Spring 2014

```
#stat2014s
```

```
barplot(height = stat2014s$`Enrollment Cnt`, names.arg = stat2014s$`Grade
Nm`,
        main = "Cari Kaufman Spring 2014 Grade Distribution",
        xlab = "Grade Name", ylab = "Frequency", las = 1)
```

```
#create the GPA distribution of the class based on its
```

```
#`Enrollment Cnt`'s, used to find the mean/median
```

```
grade_samples <- rep(stat2014s$`Average Grade`, times = stat2014s$`Enrollment
Cnt`)
```

```
#mean and median
```

```
average_grade <- mean(grade_samples)
```

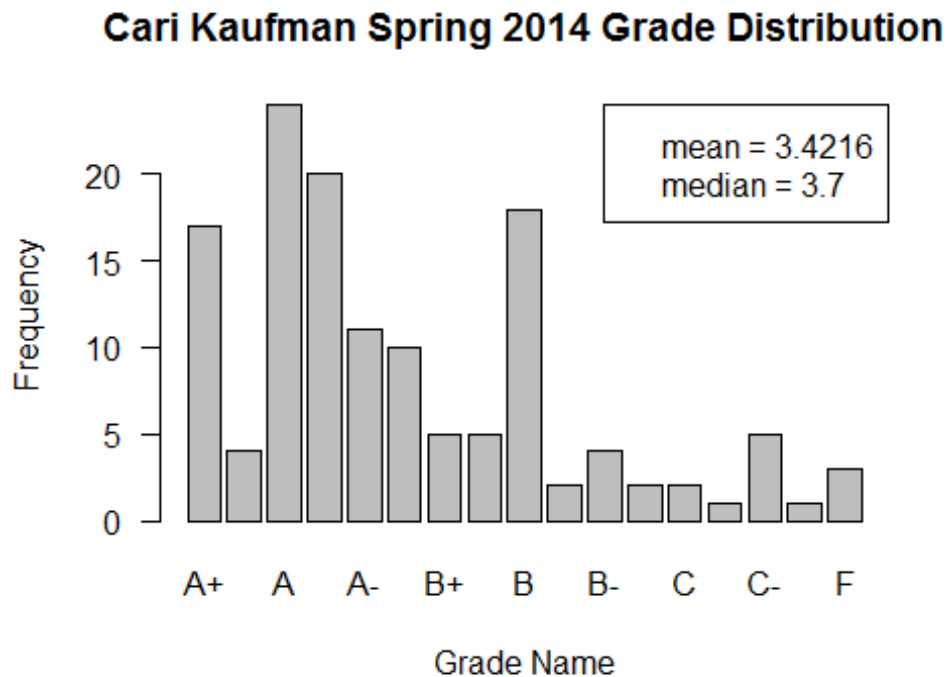
```
median_grade <- median(grade_samples)
```

```
#round the mean and median
```

```
average_grade <- round(average_grade, 4)
```

```
median_grade <- round(median_grade, 4)
```

```
#create a legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                              paste0("median = ",
as.character(median_grade))))
```



Cari Kaufman - Totals for All Years

```
#stat2010f
#stat2011f
#CariKaufman2014s

#combine the grade data for all of
#Cari Kaufman's classes

grades <- list(DeborahNolan2012f, DeborahNolan2012s, DeborahNolan2015s)

total_sum <- double(13)
names(total_sum) <- grade_names

for(i in 1:length(grades)){

  #create a temporary variable for the data
  temp_data <- data.frame(grades[i])

  for(j in 1:13){
```

```

    total_sum[j] <- total_sum[j] + sum(temp_data[temp_data$Grade.Nm ==
grade_names[j],
                                     "Enrollment.Cnt"])
  }
}

total_sum <- total_sum[total_sum != 0]
gpa_values <- unname(gpa[intersect(names(gpa), names(total_sum))])

barplot(height = total_sum, names.arg = names(total_sum),
        main = "Cari Kaufman Grade Distribution for All Classes",
        xlab = "Grade Name", ylab = "Frequency", las = 1)

#create the GPA distribution of the class based on its
#`Enrollment Cnt`s, used to find the mean/median
grade_samples <- rep(gpa_values, times = total_sum)

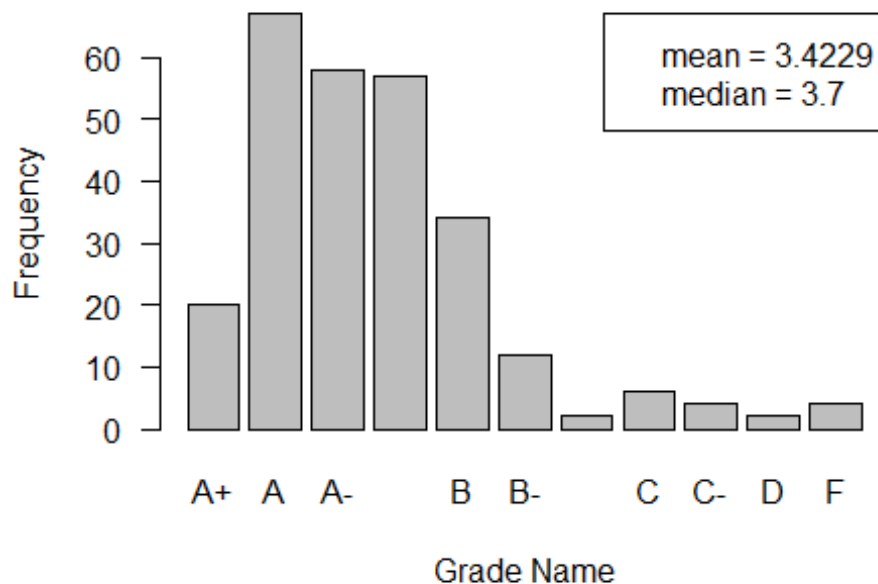
#mean and median
average_grade <- mean(grade_samples)
median_grade <- median(grade_samples)

#round the mean and median
average_grade <- round(average_grade, 4)
median_grade <- round(median_grade, 4)

#create a legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                             paste0("median = ",
as.character(median_grade))))

```

Cari Kaufman Grade Distribution for All Classes



#2013s

```
barplot(height = stat2013s$`Enrollment Cnt`, names.arg = stat2013s$`Grade Nm`,  
        main = "Haiyan Huang Spring 2013 Grade Distribution",  
        xlab = "Grade Name", ylab = "Frequency", las = 1)
```

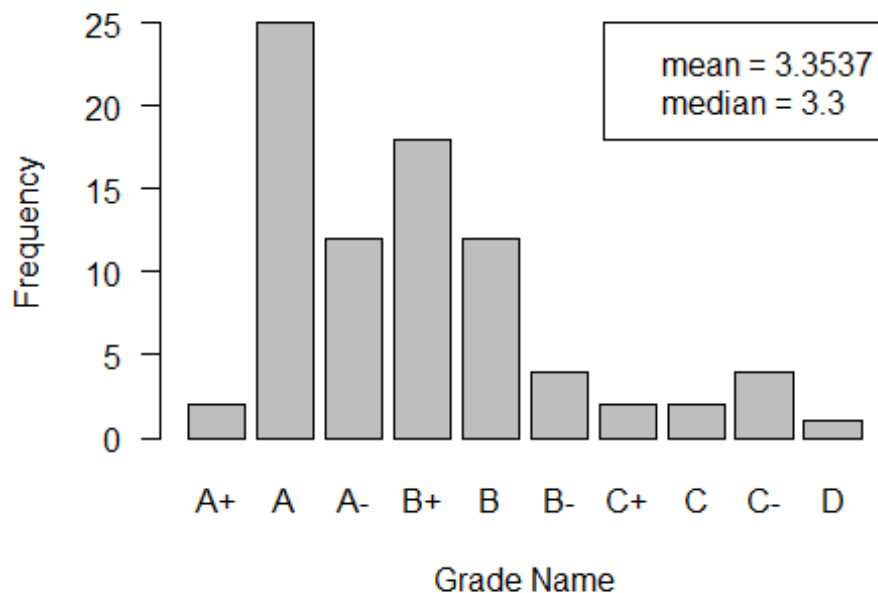
```
#create the GPA distribution of the class based on its  
#`Enrollment Cnt`s, used to find the mean/median  
grade_samples <- rep(stat2013s$`Average Grade`, times = stat2013s$`Enrollment Cnt`)
```

```
#mean and median  
average_grade <- mean(grade_samples)  
median_grade <- median(grade_samples)
```

```
#round the mean and median  
average_grade <- round(average_grade, 4)  
median_grade <- round(median_grade, 4)
```

```
#create a legend that says what the lines mean  
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),  
                             paste0("median = ",  
as.character(median_grade))))
```


Haiyan Huang Spring 2013 Grade Distribution



Deborah Nolan - All

```
grades <- list(DeborahNolan2012f, DeborahNolan2012s, DeborahNolan2015s)

total_sum <- double(13)
names(total_sum) <- grade_names

for(i in 1:length(grades)){

  #create a temporary variable for the data
  temp_data <- data.frame(grades[i])

  for(j in 1:13){

    total_sum[j] <- total_sum[j] + sum(temp_data[temp_data$Grade.Nm ==
grade_names[j],
                                     "Enrollment.Cnt"])

  }

}

total_sum <- total_sum[total_sum != 0]
gpa_values <- unname(gpa[intersect(names(gpa), names(total_sum))])
```

```

barplot(height = total_sum, names.arg = names(total_sum),
        main = "Deborah Nolan Grade Distribution for All Classes",
        xlab = "Grade Name", ylab = "Frequency", las = 1)

#create the GPA distribution of the class based on its
#`Enrollment Cnt`'s, used to find the mean/median
grade_samples <- rep(gpa_values, times = total_sum)

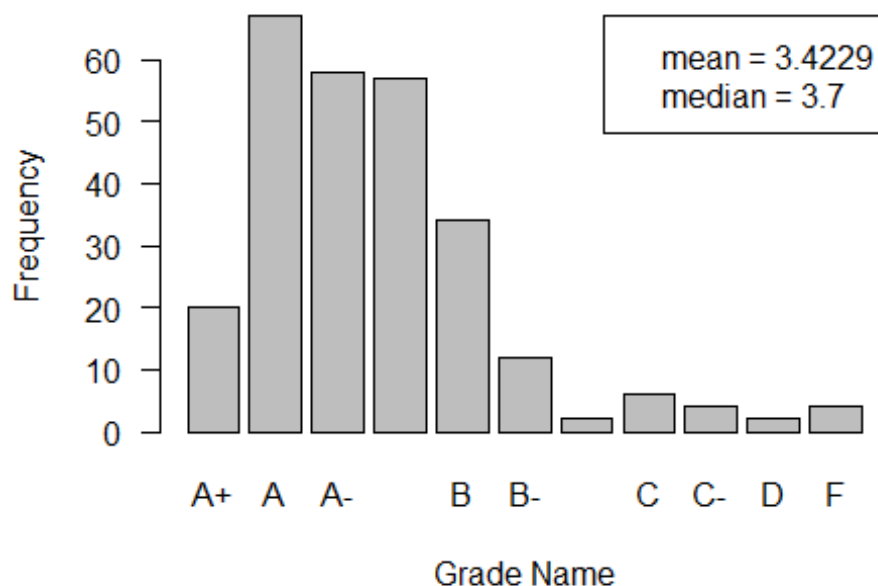
#mean and median
average_grade <- mean(grade_samples)
median_grade <- median(grade_samples)

#round the mean and median
average_grade <- round(average_grade, 4)
median_grade <- round(median_grade, 4)

#create a Legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                             paste0("median = ",
as.character(median_grade))))

```

Deborah Nolan Grade Distribution for All Classes



Fletcher Ibser - All

```

grades <- list(FletcherIbser2012f, FletcherIbser2012s,
               FletcherIbser2013f, FletcherIbser2014f)

```

```
total_sum <- double(13)
names(total_sum) <- grade_names

for(i in 1:length(grades)){

  #create a temporary variable for the data
  temp_data <- data.frame(grades[i])

  for(j in 1:13){

    total_sum[j] <- total_sum[j] + sum(temp_data[temp_data$Grade.Nm ==
grade_names[j],
                                         "Enrollment.Cnt"])

  }

}

total_sum <- total_sum[total_sum != 0]
gpa_values <- unname(gpa[intersect(names(gpa), names(total_sum))])

barplot(height = total_sum, names.arg = names(total_sum),
        main = "Fletcher Ibsen Grade Distribution for All Classes",
        xlab = "Grade Name", ylab = "Frequency", las = 1)

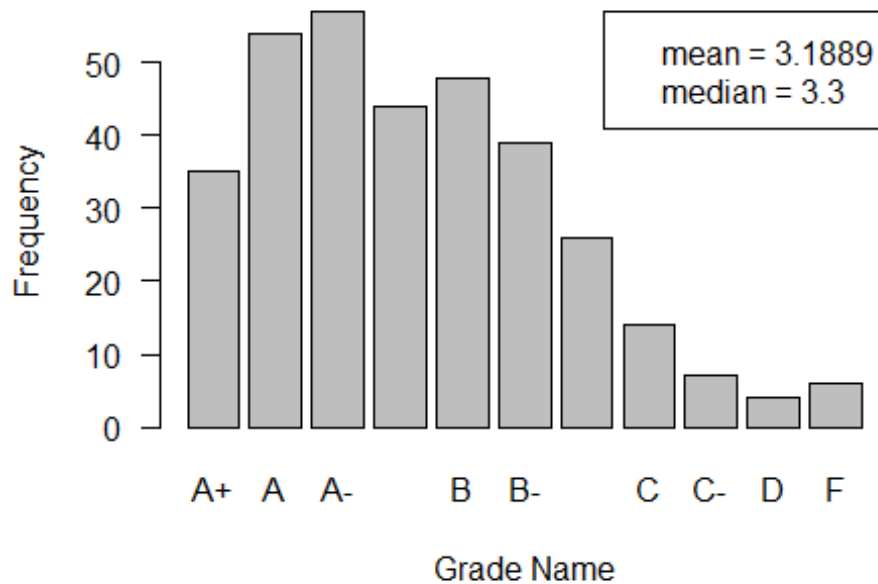
#create the GPA distribution of the class based on its
#`Enrollment Cnt`'s, used to find the mean/median
grade_samples <- rep(gpa_values, times = total_sum)

#mean and median
average_grade <- mean(grade_samples)
median_grade <- median(grade_samples)

#round the mean and median
average_grade <- round(average_grade, 4)
median_grade <- round(median_grade, 4)

#create a legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                             paste0("median = ",
as.character(median_grade))))
```

Fletcher Ibser Grade Distribution for All Classes



Benoit Dherin - All

```
grades <- list(stat2013f, BenoitDherin2014s, BenoitDherin2014f)

total_sum <- double(13)
names(total_sum) <- grade_names

for(i in 1:length(grades)){

  #create a temporary variable for the data
  temp_data <- data.frame(grades[i])

  for(j in 1:13){

    total_sum[j] <- total_sum[j] + sum(temp_data[temp_data$Grade.Nm ==
grade_names[j],
                                         "Enrollment.Cnt"])

  }

}

total_sum <- total_sum[total_sum != 0]
gpa_values <- unname(gpa[intersect(names(gpa), names(total_sum))])
```

```

barplot(height = total_sum, names.arg = names(total_sum),
        main = "Benoit Dherin Grade Distribution for All Classes",
        xlab = "Grade Name", ylab = "Frequency", las = 1)

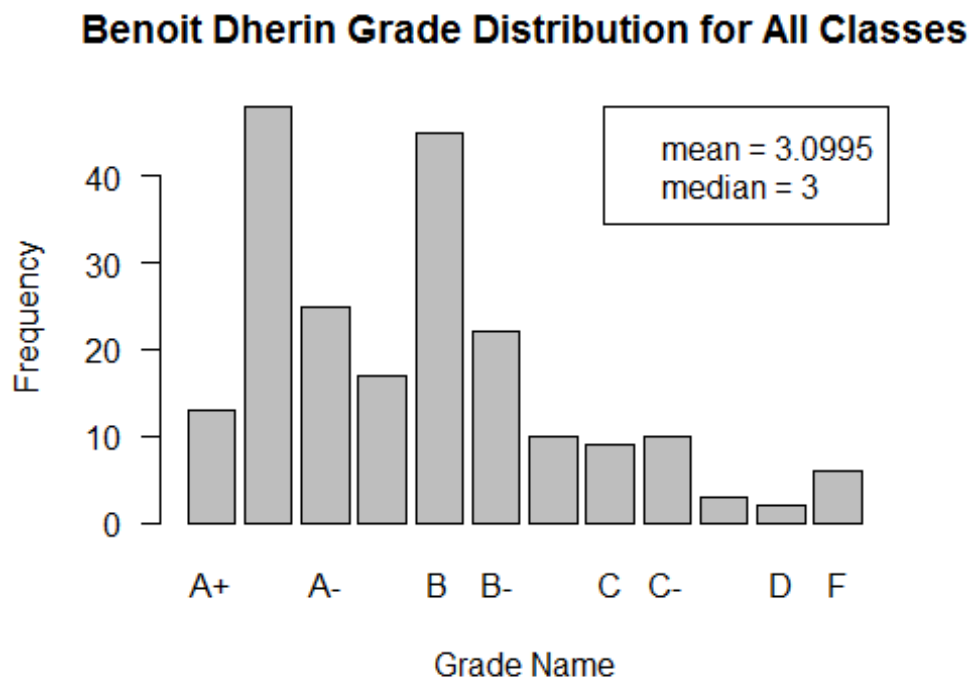
#create the GPA distribution of the class based on its
#`Enrollment Cnt`s, used to find the mean/median
grade_samples <- rep(gpa_values, times = total_sum)

#mean and median
average_grade <- mean(grade_samples)
median_grade <- median(grade_samples)

#round the mean and median
average_grade <- round(average_grade, 4)
median_grade <- round(median_grade, 4)

#create a Legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                             paste0("median = ",
as.character(median_grade))))

```



Ingileif Hallgrimsdottir - All

```
grades <- list(Ingileif2013f, Ingileif2015s)
```

```

total_sum <- double(13)
names(total_sum) <- grade_names

for(i in 1:length(grades)){

  #create a temporary variable for the data
  temp_data <- data.frame(grades[i])

  for(j in 1:13){

    total_sum[j] <- total_sum[j] + sum(temp_data[temp_data$Grade.Nm ==
grade_names[j],
                                     "Enrollment.Cnt"])

  }

}

total_sum <- total_sum[total_sum != 0]
gpa_values <- unname(gpa[intersect(names(gpa), names(total_sum))])

barplot(height = total_sum, names.arg = names(total_sum),
        main = "Ingileif Hallgrimsdottir Grade Distribution for All Classes",
        xlab = "Grade Name", ylab = "Frequency", las = 1)

#create the GPA distribution of the class based on its
#`Enrollment Cnt`s, used to find the mean/median
grade_samples <- rep(gpa_values, times = total_sum)

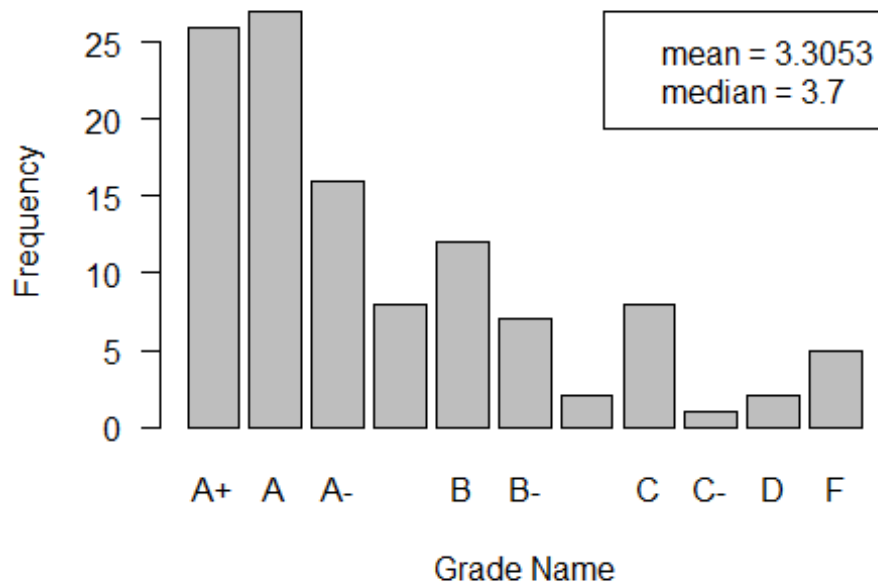
#mean and median
average_grade <- mean(grade_samples)
median_grade <- median(grade_samples)

#round the mean and median
average_grade <- round(average_grade, 4)
median_grade <- round(median_grade, 4)

#create a legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                             paste0("median = ",
as.character(median_grade))))

```

Ingileif Hallgrimsdottir Grade Distribution for All Clas



Philip Spector - All

```
grades <- list(stat2010s, stat2011s)

total_sum <- double(13)
names(total_sum) <- grade_names

for(i in 1:length(grades)){

  #create a temporary variable for the data
  temp_data <- data.frame(grades[i])

  for(j in 1:13){

    total_sum[j] <- total_sum[j] + sum(temp_data[temp_data$Grade.Nm ==
grade_names[j],
                                     "Enrollment.Cnt"])

  }

}
```

```
total_sum <- total_sum[total_sum != 0]
gpa_values <- unname(gpa[intersect(names(gpa), names(total_sum))])
```

```

barplot(height = total_sum, names.arg = names(total_sum),
        main = "Philip Spector Grade Distribution for All Classes",
        xlab = "Grade Name", ylab = "Frequency", las = 1)

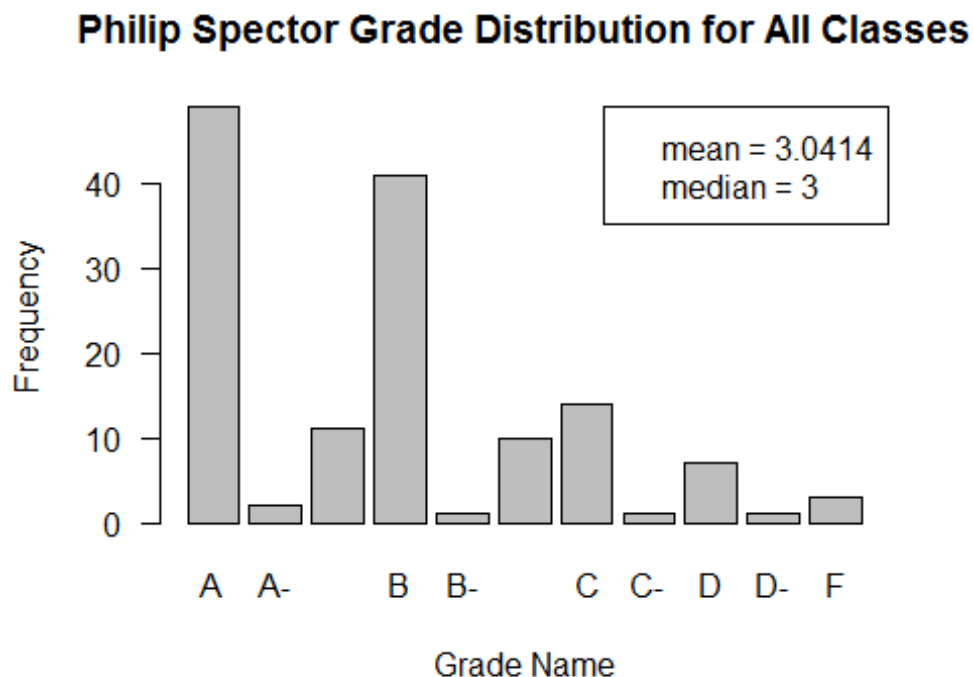
#create the GPA distribution of the class based on its
#`Enrollment Cnt`'s, used to find the mean/median
grade_samples <- rep(gpa_values, times = total_sum)

#mean and median
average_grade <- mean(grade_samples)
median_grade <- median(grade_samples)

#round the mean and median
average_grade <- round(average_grade, 4)
median_grade <- round(median_grade, 4)

#create a Legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                             paste0("median = ",
as.character(median_grade))))

```



All Classes

```

grades <- list(stat2010f, stat2011f, stat2012f,
               stat2013f, stat2014f, stat2010s,

```



```

stat2011s, stat2012s, stat2013s,
stat2014s, stat2015s)

total_sum <- double(13)
names(total_sum) <- grade_names

for(i in 1:length(grades)){

  #create a temporary variable for the data
  temp_data <- data.frame(grades[i])

  for(j in 1:13){

    total_sum[j] <- total_sum[j] + sum(temp_data[temp_data$Grade.Nm ==
grade_names[j],
                                     "Enrollment.Cnt"])

  }
}

total_sum <- total_sum[total_sum != 0]
gpa_values <- unname(gpa[intersect(names(gpa), names(total_sum))])

barplot(height = total_sum, names.arg = names(total_sum),
        main = "Grade Distribution for All Classes for All Years",
        xlab = "Grade Name", ylab = "Frequency", las = 1)

#create the GPA distribution of the class based on its
#`Enrollment Cnt`s, used to find the mean/median
grade_samples <- rep(gpa_values, times = total_sum)

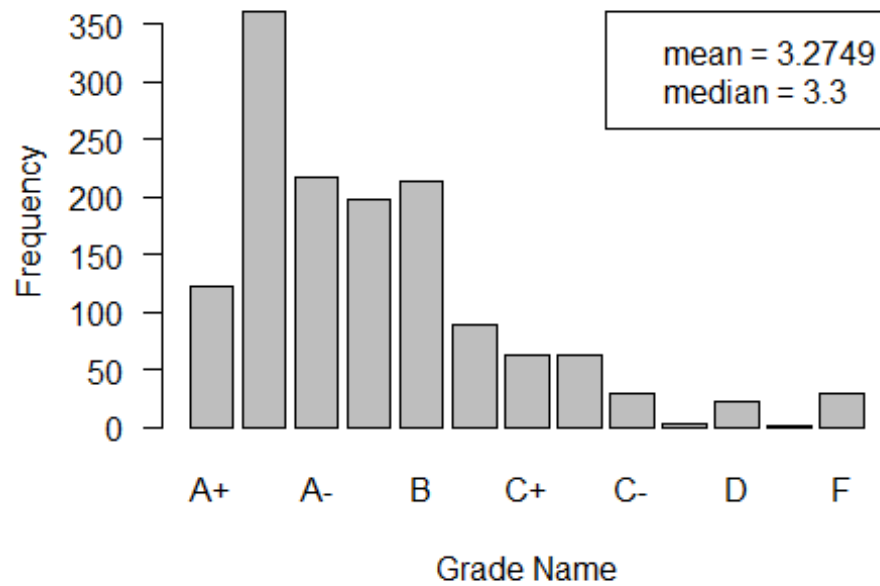
#mean and median
average_grade <- mean(grade_samples)
median_grade <- median(grade_samples)

#round the mean and median
average_grade <- round(average_grade, 4)
median_grade <- round(median_grade, 4)

#create a Legend that says what the lines mean
legend("topright", legend = c(paste0("mean = ", as.character(average_grade)),
                             paste0("median = ",
as.character(median_grade))))

```

Grade Distribution for All Classes for All Years



```
total_prop <- prop.table(total_sum)
sum(head(total_prop, 3))
## [1] 0.497516
sum(total_prop[4:6])
## [1] 0.3548616
```