Mastering Chemistry Class Analysis

Timothy Wanless

April 30, 2020

The purpose of this report is to help provide the instructors of Mastering Chemistry with insights into how to structure their course and orientate their efforts in educating students.

1. **This section shows how the data was imported and edited for analysis.**

This sets the working directory.

setwd("~/MSAE/MSAE-Spring2020/EFR 535")

This installs packages

library(psych)

## Warning: package 'psych' was built under R version 3.5.3

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.5.3

##   
## Attaching package: 'ggplot2'

## The following objects are masked from 'package:psych':  
##   
## %+%, alpha

library(tidyverse)

## Warning: package 'tidyverse' was built under R version 3.5.3

## -- Attaching packages ------------------------------------------- tidyverse 1.3.0 --

## v tibble 2.1.3 v dplyr 0.8.5  
## v tidyr 1.0.2 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.4.0  
## v purrr 0.3.3

## Warning: package 'tibble' was built under R version 3.5.3

## Warning: package 'tidyr' was built under R version 3.5.3

## Warning: package 'purrr' was built under R version 3.5.3

## Warning: package 'dplyr' was built under R version 3.5.3

## Warning: package 'stringr' was built under R version 3.5.3

## Warning: package 'forcats' was built under R version 3.5.3

## -- Conflicts ---------------------------------------------- tidyverse\_conflicts() --  
## x ggplot2::%+%() masks psych::%+%()  
## x ggplot2::alpha() masks psych::alpha()  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(waffle)

## Warning: package 'waffle' was built under R version 3.5.3

library(ggExtra)

## Warning: package 'ggExtra' was built under R version 3.5.3

This section describes how I cleaned the data before writing the code.

I renamed some variable names to be cleaner and removed ID numbers. I also replaced all similar values of categorical variables with common responses i.e. “freshman” to “Freshman”. I added a percentage variable to numerous the main total graded variables with point values. I changed some of classifications of variables later as needed for certain graphs. Notes were made on these changes when they occur in code.

This code loads data.

library(readxl)

## Warning: package 'readxl' was built under R version 3.5.3

Chem <- read\_excel("Final/Chem\_Data\_Edit.xlsx")

## New names:  
## \* `` -> ...11  
## \* `` -> ...16  
## \* `` -> ...20  
## \* `` -> ...24  
## \* `` -> ...33  
## \* ... and 5 more problems

View(Chem)  
attach(Chem)

This creates a new dataset with no missing values.

Chem\_NoMISS <- Chem [complete.cases(Chem), ]

This part creates summary statistics for the data set.

I included the percentage values for variousl parts in the class in order to see how students performed overall on various activities and see the variance in points awarded.

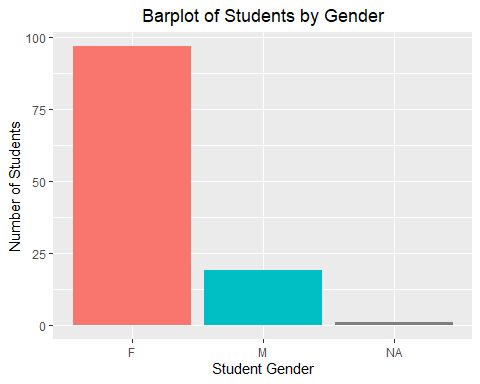
summary <- Chem %>%  
 select(`TotalGrade%`,`AdjFinal%`,`Exam1%`,`Exam2%`,`Exam3%`,`Exam4%`,`MC\_Mod%`,`MC\_HW%`,`MC\_Quiz%`,`TeamReportTot%`,`ExamRev&PeerEval%`)%>%  
 psych::describe(quant=c(.25,.75)) %>%  
 as\_tibble(rownames="rowname") %>%  
 print()

## # A tibble: 11 x 16  
## rowname vars n mean sd median trimmed mad min max  
## <chr> <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 TotalG~ 1 117 0.790 0.123 0.797 0.797 0.0988 0.448 1.01  
## 2 AdjFin~ 2 117 73.9 16.1 71.7 73.4 14.8 0 117.   
## 3 Exam1% 3 117 77.0 14.0 80 77.9 12.7 40 100   
## 4 Exam2% 4 117 74.3 14.7 74.2 75.2 14.4 29.0 100   
## 5 Exam3% 5 116 75.3 16.6 74.2 75.9 14.4 29 106.   
## 6 Exam4% 6 116 73.3 16.1 73.5 74.1 17.4 35.3 103.   
## 7 MC\_Mod% 7 117 0.872 0.193 1 0.913 0 0.0908 1   
## 8 MC\_HW% 8 117 0.845 0.192 0.916 0.871 0.154 0.0969 1.07  
## 9 MC\_Qui~ 9 117 0.696 0.198 0.738 0.705 0.171 0.110 1.09  
## 10 TeamRe~ 10 117 0.951 0.0966 0.969 0.966 0.0741 0.560 1.06  
## 11 ExamRe~ 11 117 0.893 0.159 0.985 0.927 0.0217 0.232 1   
## # ... with 6 more variables: range <dbl>, skew <dbl>, kurtosis <dbl>, se <dbl>,  
## # Q0.25 <dbl>, Q0.75 <dbl>

1. **This section describes characteristics of the student body such as: gender, level, and program.**

This is a bargraph of gender.

ggplot(Chem, aes(x = (Gender), fill=Gender)) +  
 geom\_bar(show.legend = FALSE)+  
 labs(title="Barplot of Students by Gender", x="Student Gender", y="Number of Students")+  
 theme(plot.title = element\_text(hjust = 0.5))



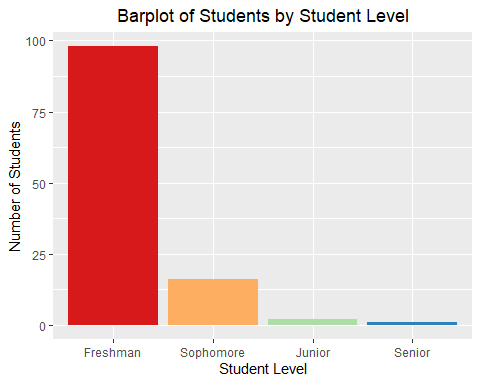
The graph shows that the vast majority of students are female(F) and a minority of students male(M) with some students unknown(NA) in Mastering Chemistry.

This reorders the students levels.

Chem$Level<- factor(Chem$Level, levels = c("Freshman","Sophomore","Junior","Senior"))

This is a barplot of students levels

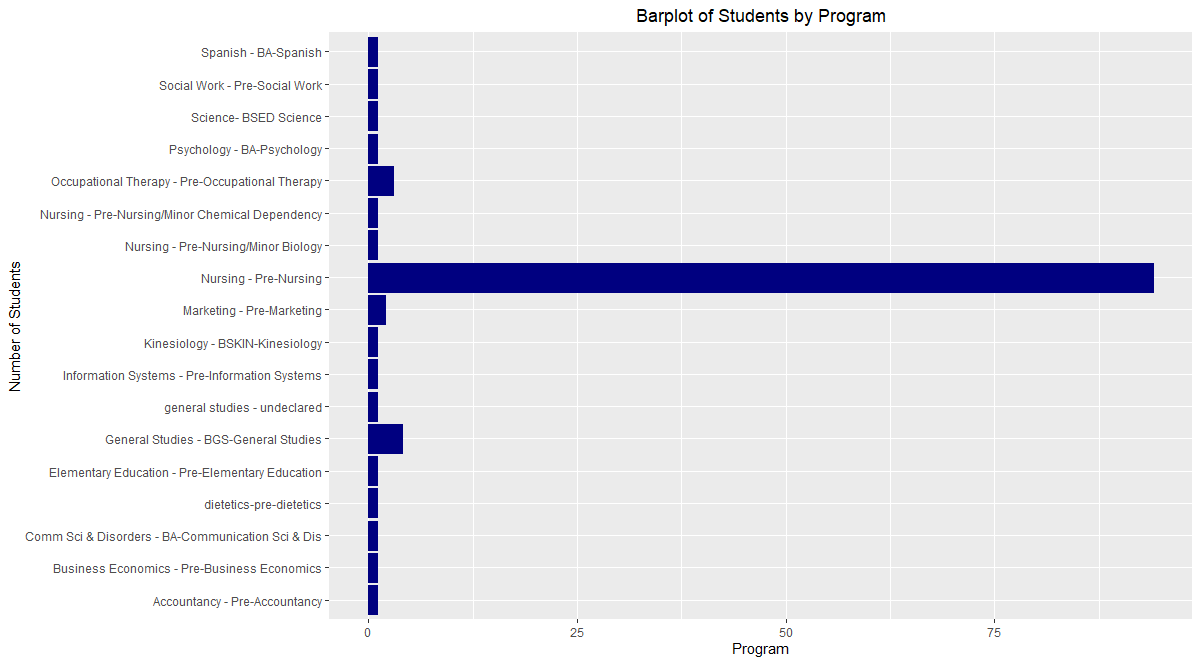
ggplot(Chem, aes(x = Level, fill=Level)) +  
 geom\_bar(show.legend = FALSE) +  
 scale\_fill\_brewer(palette = "Spectral")+  
 labs(title="Barplot of Students by Student Level", x="Student Level", y="Number of Students")+  
 theme(plot.title = element\_text(hjust = 0.5))



This barplot shows that the vast majority of students are freshman, and the number of students decreases as level increases.

This is a bargraph of grade program.

ggplot(Chem, aes(x = Program)) +  
 geom\_bar(fill="navy",color="navy") +  
 coord\_flip()+  
 labs(title = "Barplot of Students by Program", x="Number of Students",y="Program")+  
 theme(plot.title = element\_text(hjust = 0.5))



This shows that the majority of students that are in Mastering Chemistry are in the Pre-Nursing Program.

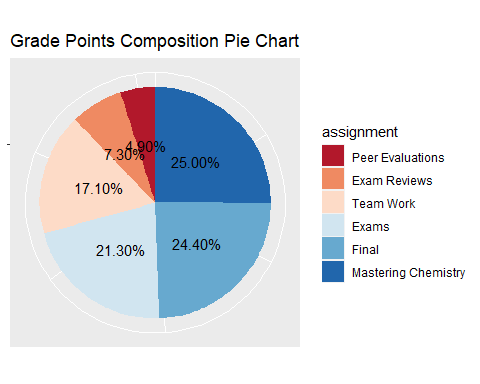
1. **This section analyzes how students were evaluated in the class.**

This creates a dataframe to show how students’ grades are evaluated in the class. Mastering Chemistry contains module, homework, and quiz points. Team Work contains reports and quizzes of the course. Exams cointains Exams 1-4.

GradeComp<- data.frame(  
 assignment=c("Peer Evaluations","Exam Reviews","Team Work","Exams","Final","Mastering Chemistry"),  
 points=c(30, 45,105,131,150,154)  
)  
  
GradeComp$perc<-GradeComp$points/sum(GradeComp$points)  
  
GradeComp$assignment <- factor(GradeComp$assignment, levels = c("Peer Evaluations","Exam Reviews","Team Work","Exams","Final","Mastering Chemistry"))

This creates a pie chart showing the composition of the Mastering Chemistry class.

attach(GradeComp)  
ggplot(GradeComp, aes(x="",y=points, fill= assignment))+  
 geom\_bar(stat = "identity")+  
 coord\_polar(theta = "y", start = 0)+  
 labs(x=NULL, y=NULL, title="Grade Points Composition Pie Chart")+  
 scale\_fill\_brewer(palette = "RdBu")+  
 theme(axis.line = element\_blank(),  
 axis.text = element\_blank(),  
 plot.title = element\_text(hjust = 0.5))+  
 geom\_text(aes(label=scales::percent(round(perc,3))), position = position\_stack(vjust = 0.5))

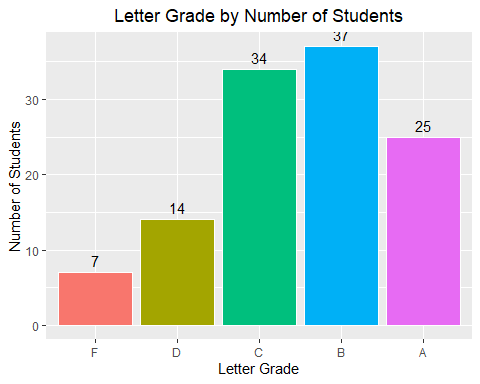


This shows that the majority of the grade comes from exams, Mastering Chemistry coursework and quizzes, and the final.

1. **This section analyzes how students performed and the distribution of grades.**

This shows the proportion of letter grades in a histogram.

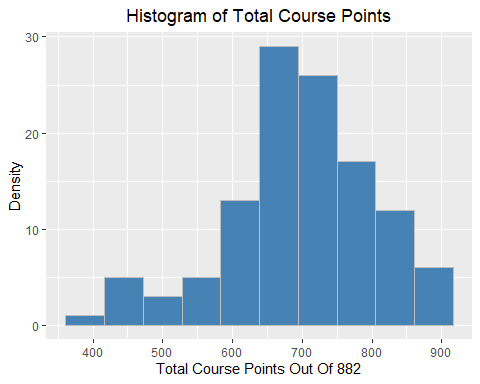
Chem$GRADE<- factor(Chem$GRADE, levels = c("F","D","C","B","A"))  
  
ggplot(data = Chem, mapping = aes(x = GRADE, fill=GRADE))+  
 geom\_bar( color="white",show.legend = FALSE)+  
 geom\_text(stat='count', aes(label=..count..), vjust=-0.5)+  
 ggtitle("Letter Grade by Number of Students")+  
 theme(plot.title = element\_text(hjust = 0.5))+  
 labs(x="Letter Grade", y="Number of Students")



There tend to be more students in passing grade bins rather than failing grade categories.

This creates a histogram for total course points. The plot is adjusted for 10 bins to better show the distribution to the right devoid of noise.

ggplot(data = Chem, mapping = aes(x = TotalPts\_Of882))+  
 geom\_histogram(bins = 10,color="grey",fill="steelblue")+  
 ggtitle("Histogram of Total Course Points")+  
 theme(plot.title = element\_text(hjust = 0.5))+  
 labs(x="Total Course Points Out Of 882", y="Density")



The distribution of total course points is slightly skewed to the right. The may indicate the majority of students prefer to not have a sunk cost of paying for a class they don’t pass.

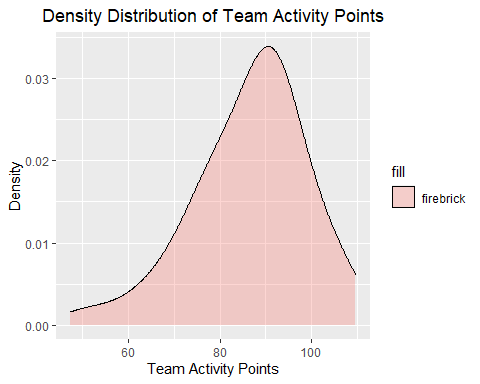
1. **This section analyzes how much students participated in Team activities.**

This creates a Team variable that combines Team Reports and Team Quizzes and another than combines Exams 1-4.

Chem$Teams<-Chem$TeamQuizTot\_OfApprox47+Chem$TeamReportTot\_OfApprox58  
Chem$Exams<-Chem$Exam1\_Of35+Chem$Exam2\_Of31+Chem$Exam3\_Of31+Chem$Exam4\_Of34

This creates a density plot of team activities.

ggplot(data =Chem, mapping = aes(x = Teams,fill="firebrick"))+  
 geom\_density(adjust=1.5, alpha=0.3)+  
 ggtitle("Density Distribution of Team Activity Points")+  
 theme(plot.title = element\_text(hjust = 0.5))+  
 labs(x="Team Activity Points", y="Density")



The distribution is skewed to the right and highly concentrated around 90 points.

This is a scatterplot and trendline for Teamwork and Exams

ggplot(Chem, aes(x=Teams,y=Exams))+  
 geom\_point()+  
 geom\_smooth(method = lm)+  
 labs(title = "Team Work Relation to Exam Performance Scatterplot and Trendline", x="Team Work Points", y="Exam Points")+  
 theme(plot.title = element\_text(hjust = 0.5))

## `geom\_smooth()` using formula 'y ~ x'

## Warning: Removed 2 rows containing non-finite values (stat\_smooth).

## Warning: Removed 2 rows containing missing values (geom\_point).

A close up of a map

Description automatically generated

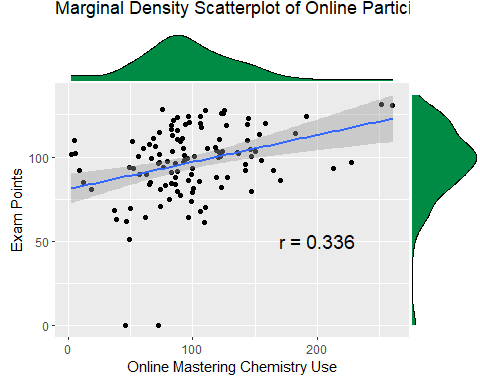
The Team Work performance is skewed right and concentrated on 90 points in the density distribution plot. There is a positive correlation between team work points and points earned on Exams.

1. **This section visualizes how students interact online.**

This creates a marginal density scatterplot of MC performance on exams.

Chem$`BB Mastering Chemistry`[is.na(Chem$`BB Mastering Chemistry`)]<-0  
Chem$Exams[is.na(Chem$Exams)]<-0  
  
a <- ggplot(Chem, aes(y=Exams,x=`BB Mastering Chemistry`))+  
 geom\_point(position = "jitter")+  
 labs(y="Exam Points", x ="Online Mastering Chemistry Use", title = "Marginal Density Scatterplot of Online Participation and Exam Performance")+  
 geom\_smooth(method=lm)+  
 annotate(x=200,y=50,  
 label=paste("r =", round(cor(Chem$Exams,Chem$`BB Mastering Chemistry`),3)),  
 geom="text",size=5  
 )  
  
ggMarginal(a, type = "density", fill = "springgreen4")

## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'

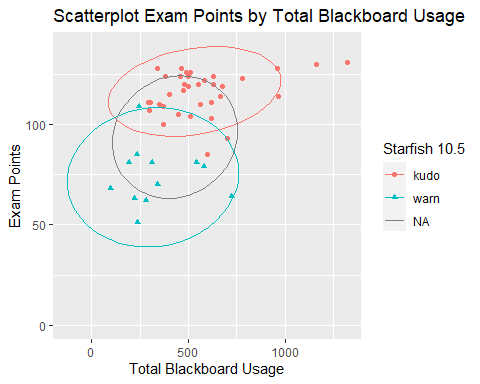


This shows that there is a positive relationship of 0.336 between Exam Points and Online Use. The Blackboard Online Mastering Chemistry is skewed to the left.

1. **This section analyzes Blackboard activity by students.**

ggplot(Chem, aes(x = `BB Total Usage`, y = Exams, color = `Starfish 10.5`,shape=`Starfish 10.5`))+  
 geom\_point()+  
 stat\_ellipse()+  
 labs(x="Total Blackboard Usage", y="Exam Points", title = "Scatterplot Exam Points by Total Blackboard Usage")

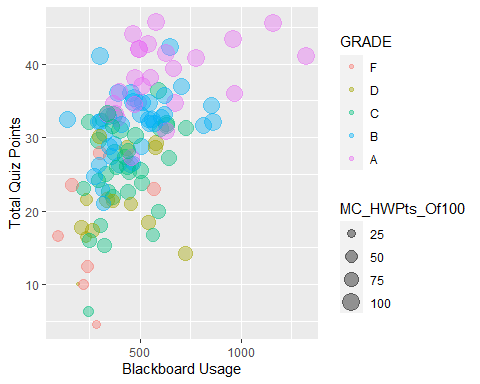
## Warning: Removed 70 rows containing missing values (geom\_point).



Students with kudo messages performed better on tests and students with warn message performed less than others.

This creates a bubble chart showing how quizzes,homework, and final grade are related to blackboard activity.

ggplot(Chem, aes(x=`BB Total Usage`,y=MC\_QuizTot\_Of42,size=MC\_HWPts\_Of100,color=GRADE))+  
 geom\_point(alpha=0.4)+  
 labs(x="Blackboard Usage", y="Total Quiz Points", "Bubblechart of Points Relation to Blackboard Usage")



Here we see that there is a tendency for increase blackboard usage to lead to greater points and grades.

1. **This section checks correlation between the class outcomes and its factors.**

This code creates subsets of data for use in a correlogram.

Chem\_cor <- Chem[, c("AdjFinal%","ExamRev&PeerEval%","TeamQuizTot%","TeamReportTot%","MC\_Quiz%","MC\_HW%","MC\_Mod%")]  
attach(Chem\_cor)

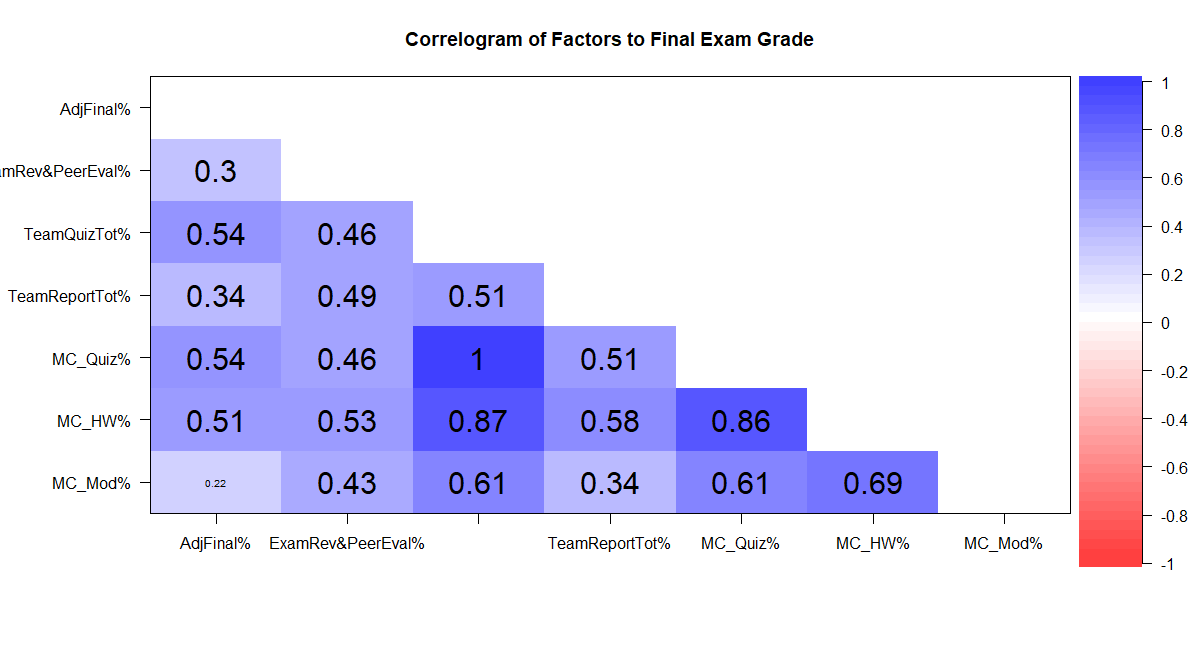
## The following objects are masked from Chem:  
##   
## AdjFinal%, ExamRev&PeerEval%, MC\_HW%, MC\_Mod%, MC\_Quiz%,  
## TeamQuizTot%, TeamReportTot%

Chem\_cor2 <-Chem[,c("Exam1%","Exam2%","Exam3%","Exam4%","AdjFinal%")]  
attach(Chem\_cor2)

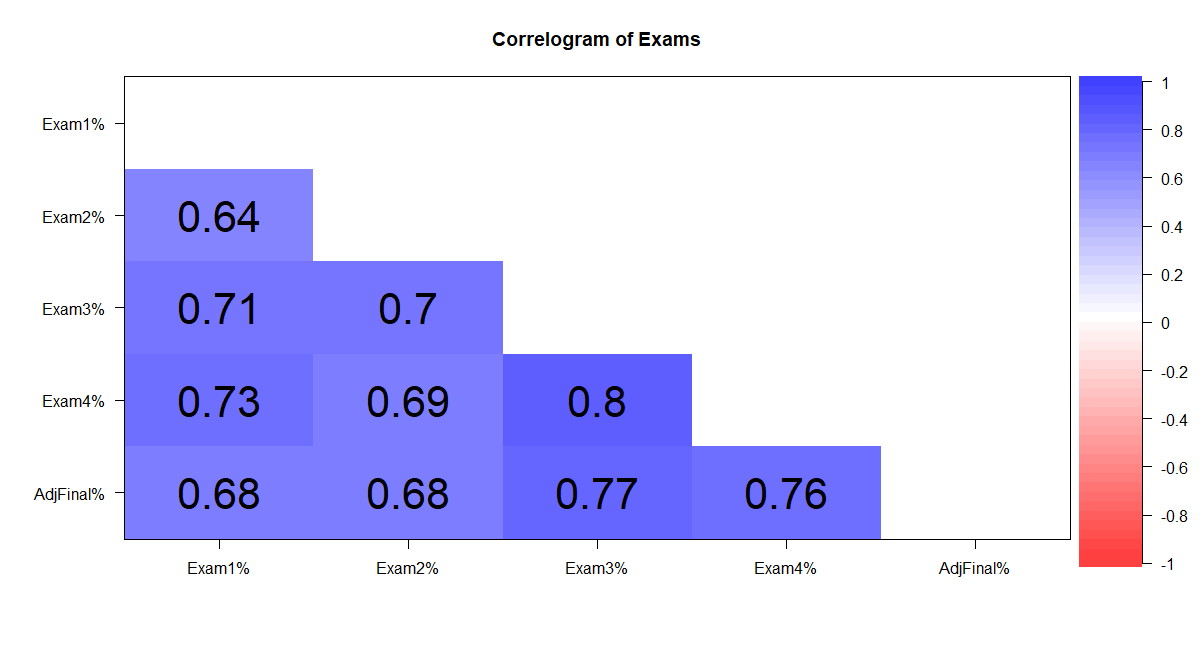
## The following object is masked from Chem\_cor:  
##   
## AdjFinal%

## The following objects are masked from Chem:  
##   
## AdjFinal%, Exam1%, Exam2%, Exam3%, Exam4%

This code creates a correlogram using online MC performance, in class team performance, and exam review sessions.

This correlogram shows that their is a fairly strong relation to quiz and homework performance and less so for reviews and module interaction.

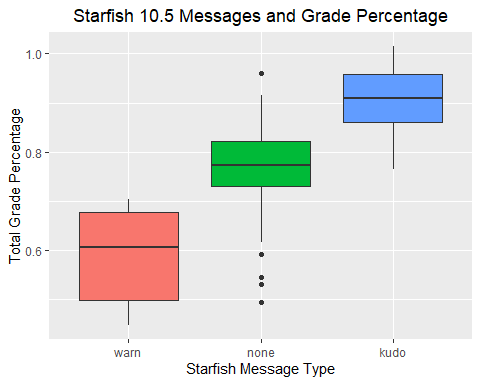
This code shows how other exam outcomes are related to final exam performance.

This shows that Exams are positively correlated to each other. The better someone does on one exam greatly increases their likelihood to do well on another.

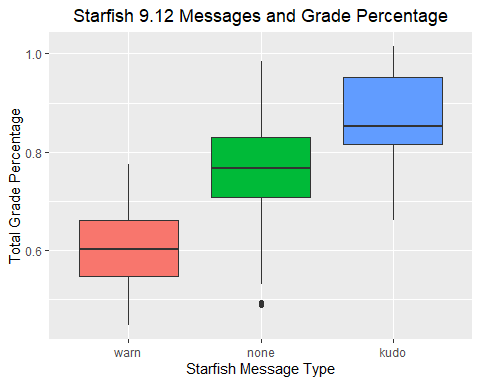
1. **This section looks at the Starfish messages and their effect.**

This creates boxplots to show how starfish messages related to students performance.

Chem$`Starfish 10.5`[is.na(Chem$`Starfish 10.5`)]<-"none"  
Chem$`StarFish 9.12`[is.na(Chem$`StarFish 9.12`)]<-"none"  
Chem$`StarFish 9.12`<- factor(Chem$`StarFish 9.12`, levels = c("warn","none","kudo"))  
Chem$`Starfish 10.5`<- factor(Chem$`Starfish 10.5`, levels = c("warn","none","kudo"))  
  
ggplot(Chem, aes(x=`Starfish 10.5`, y=`TotalGrade%`, fill=`Starfish 10.5`))+  
 geom\_boxplot(show.legend = FALSE)+  
 labs(title = "Starfish 10.5 Messages and Grade Percentage",x="Starfish Message Type", y="Total Grade Percentage")+  
 theme(plot.title = element\_text(hjust = 0.5))



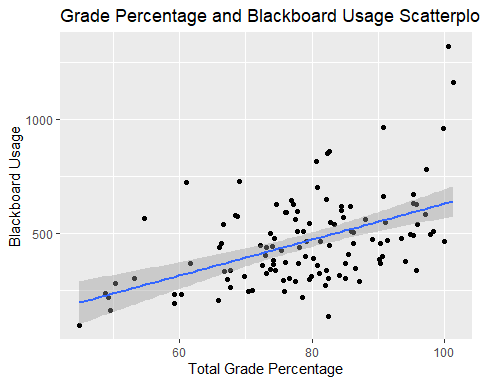
ggplot(Chem, aes(x=`StarFish 9.12`, y=`TotalGrade%`, fill=`StarFish 9.12`))+  
 geom\_boxplot(show.legend = FALSE)+  
 labs(title = "Starfish 9.12 Messages and Grade Percentage",x="Starfish Message Type", y="Total Grade Percentage")+  
 theme(plot.title = element\_text(hjust = 0.5))



This shows that students that receive a warn message perform below average and that students that receive a kudo message perform above average.

ggplot(Chem, aes(x=`TotalGrade%`\*100,y=`BB Total Usage`))+  
 geom\_point()+  
 geom\_smooth(method = lm)+  
 labs(title = "Grade Percentage and Blackboard Usage Scatterplot", x="Total Grade Percentage", y="Blackboard Usage")

## `geom\_smooth()` using formula 'y ~ x'



annotate(x=70,y=1000,  
 label=paste("r =", round(cor(Chem$`TotalGrade%`,`BB Total Usage`),3)),  
 geom="text",size=5  
 )

## mapping: x = ~x, y = ~y   
## geom\_text: na.rm = FALSE  
## stat\_identity: na.rm = FALSE  
## position\_identity

Professors could look at Starfish messages or Blackboard Usage in order to see which students are at risk of failing the class.

1. **This section looks at the change in the class over time.**

This gets that averages of each exam grade including the final. I replace missing values with a grade of 0.

mean(`Exam1%`)

## [1] 77.02145

mean(`Exam2%`)

## [1] 74.3041

`Exam3%`[is.na(`Exam3%`)]<-0  
mean(`Exam3%` )

## [1] 74.65641

`Exam4%`[is.na(`Exam4%`)]<-0  
mean(`Exam4%`)

## [1] 72.69984

mean(`AdjFinal%`)

## [1] 73.90256

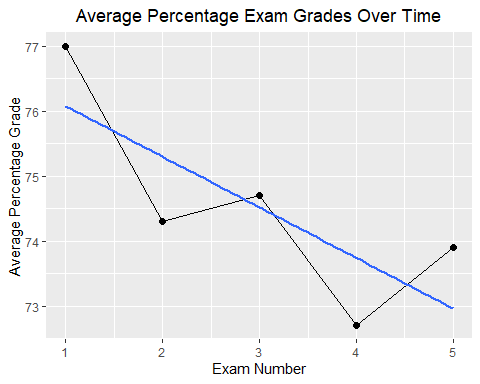
This sets up the Exam data frame.

Exams<- data.frame(  
 Number= c(1,2,3,4,5),  
 ExamAvg= c(77.0,74.3,74.7,72.7,73.9))  
attach(Exams)

This creates a line graph with a blue trendline.

ggplot(Exams, aes(x=Number, y=ExamAvg))+  
 geom\_line()+  
 geom\_smooth(method=lm, se=FALSE)+  
 geom\_point(size=2,color="black")+  
 labs(x="Exam Number", y="Average Percentage Grade", title = "Average Percentage Exam Grades Over Time")+  
 theme(plot.title = element\_text(hjust = 0.5))

## `geom\_smooth()` using formula 'y ~ x'



The average percentage grade for exams appears to slightly decrease from around 77 to 73.

This finds the averages of MC Quiz percentage grades.

summary(Chem[69:98])

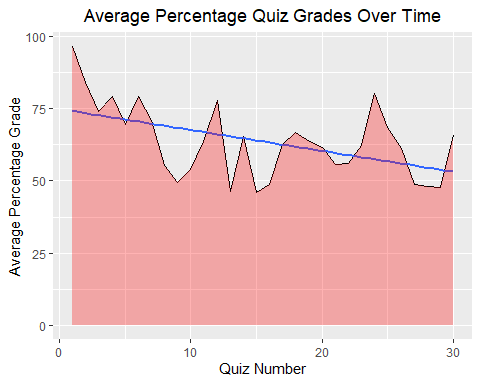
## Quiz 27\_Of1.5 Quiz 26\_Of1.5 Quiz 25\_Of1.5 Quiz 24\_Of1.5   
## Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000   
## 1st Qu.:0.5000 1st Qu.:0.0000 1st Qu.:0.5000 1st Qu.:0.5000   
## Median :1.5000 Median :0.5000 Median :1.0000 Median :0.5000   
## Mean :0.9872 Mean :0.7137 Mean :0.7222 Mean :0.7308   
## 3rd Qu.:1.5000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000   
## Max. :1.5000 Max. :1.5000 Max. :1.5000 Max. :1.5000   
## Quiz 23\_Of1.5 Quiz 22\_Of1.5 Quiz 21\_Of1.5 Quiz 20\_Of1.5   
## Min. :0.0000 Min. :0.000 Min. :0.000 Min. :0.0000   
## 1st Qu.:0.5000 1st Qu.:0.500 1st Qu.:1.000 1st Qu.:0.5000   
## Median :1.0000 Median :1.500 Median :1.500 Median :1.0000   
## Mean :0.9188 Mean :1.026 Mean :1.205 Mean :0.9316   
## 3rd Qu.:1.5000 3rd Qu.:1.500 3rd Qu.:1.500 3rd Qu.:1.5000   
## Max. :1.5000 Max. :1.500 Max. :1.500 Max. :1.5000   
## Quiz 19\_Of1.5 Quiz 18\_Of1.5 Quiz 17\_Of1.5 Quiz 16\_Of1.5   
## Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.000   
## 1st Qu.:0.5000 1st Qu.:0.5000 1st Qu.:0.5000 1st Qu.:0.500   
## Median :1.0000 Median :1.0000 Median :1.0000 Median :1.000   
## Mean :0.8419 Mean :0.8333 Mean :0.9231 Mean :0.953   
## 3rd Qu.:1.5000 3rd Qu.:1.5000 3rd Qu.:1.5000 3rd Qu.:1.500   
## Max. :1.5000 Max. :1.5000 Max. :1.5000 Max. :1.500   
## Quiz 15\_Of1.5 Quiz 14\_Of1.5 Quiz 13b\_Of1.5 Quiz 13\_Of1.5   
## Min. :0.0 Min. :0.0000 Min. :0.0000 Min. :0.000   
## 1st Qu.:0.5 1st Qu.:0.5000 1st Qu.:0.5000 1st Qu.:0.000   
## Median :1.0 Median :1.0000 Median :1.0000 Median :0.500   
## Mean :1.0 Mean :0.9359 Mean :0.7308 Mean :0.688   
## 3rd Qu.:1.5 3rd Qu.:1.5000 3rd Qu.:1.0000 3rd Qu.:1.000   
## Max. :1.5 Max. :1.5000 Max. :1.5000 Max. :1.500   
## Quiz 12\_Of1.5 Quiz 11\_Of1.5 Quiz 10b\_Of2 Quiz 10\_Of1.5   
## Min. :0.0000 Min. :0.0000 Min. :0.000 Min. :0.000   
## 1st Qu.:0.5000 1st Qu.:0.0000 1st Qu.:1.400 1st Qu.:0.670   
## Median :1.0000 Median :1.0000 Median :1.670 Median :1.000   
## Mean :0.9786 Mean :0.6923 Mean :1.553 Mean :0.954   
## 3rd Qu.:1.5000 3rd Qu.:1.0000 3rd Qu.:2.000 3rd Qu.:1.500   
## Max. :1.5000 Max. :1.5000 Max. :2.000 Max. :1.500   
## Quiz 9\_Of2 Quiz 8\_Of1.5 Quiz 7\_Of2 Quiz 6\_Of1.5   
## Min. :0.000 Min. :0.0000 Min. :0.000 Min. :0.00   
## 1st Qu.:0.670 1st Qu.:0.5000 1st Qu.:0.500 1st Qu.:1.00   
## Median :1.170 Median :0.5000 Median :1.500 Median :1.00   
## Mean :1.079 Mean :0.7393 Mean :1.107 Mean :1.06   
## 3rd Qu.:1.500 3rd Qu.:1.0000 3rd Qu.:1.500 3rd Qu.:1.50   
## Max. :2.000 Max. :1.5000 Max. :2.000 Max. :1.50   
## Quiz 5\_Of1.5 Quiz 4a\_Of1.5 Quiz 4\_Of1.5 Quiz 3\_Of1.5   
## Min. :0.000 Min. :0.000 Min. :0.000 Min. :0.000   
## 1st Qu.:1.000 1st Qu.:0.750 1st Qu.:1.000 1st Qu.:1.000   
## Median :1.500 Median :1.000 Median :1.500 Median :1.500   
## Mean :1.188 Mean :1.043 Mean :1.188 Mean :1.111   
## 3rd Qu.:1.500 3rd Qu.:1.500 3rd Qu.:1.500 3rd Qu.:1.500   
## Max. :1.500 Max. :1.500 Max. :1.500 Max. :1.500   
## Quiz 2\_Of1.5 Quiz 1\_Of1.5   
## Min. :0.000 Min. :0.000   
## 1st Qu.:1.250 1st Qu.:1.000   
## Median :1.500 Median :1.500   
## Mean :1.261 Mean :1.145   
## 3rd Qu.:1.500 3rd Qu.:1.500   
## Max. :1.500 Max. :1.500

This creates a dataframe for MC quiz values

MCQuiz <- data.frame(  
 Num=c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30),  
 QuizAvg=c(96.4,84.1,74.1,79.2,69.5,79.2,70.7,55.3,49.3,54.0,63.6,77.6,46.2,65.2,45.9,48.7,62.4,66.7,63.5,61.5,55.6,56.1,62.1,80.3,68.4,61.3,48.7,48.1,47.6,65.8)  
)

ggplot(MCQuiz, aes(x=Num, y=QuizAvg))+  
 geom\_line()+  
 geom\_smooth(method=lm, se=FALSE)+  
 geom\_area(fill="red",alpha=0.3)+  
 labs(x="Quiz Number", y="Average Percentage Grade", title = "Average Percentage Quiz Grades Over Time")+  
 theme(plot.title = element\_text(hjust = 0.5))

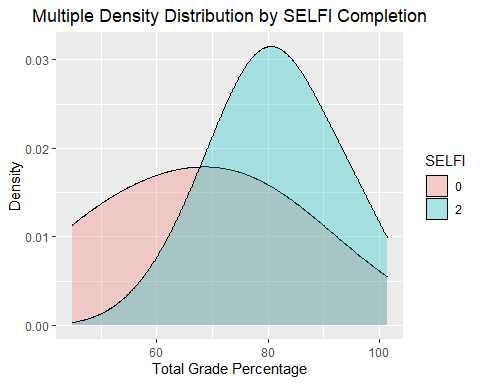
## `geom\_smooth()` using formula 'y ~ x'



**Fun Bonus!**

This shows whether students who completed SELFI,2, performed better or worse than those who did not,0.

Chem$SELFI[is.na(Chem$SELFI)]<-0  
Chem$SELFI<- as.character(Chem$SELFI)  
ggplot(data =Chem, mapping = aes(x = `TotalGrade%`\*100, fill=SELFI))+  
 geom\_density(adjust=2, alpha=0.3)+  
 ggtitle("Multiple Density Distribution by SELFI Completion")+  
 theme(plot.title = element\_text(hjust = 0.5))+  
 labs(x="Total Grade Percentage", y="Density")



This shows that students that completed SELFI were much more likely to obtain a higher grade than students who did not.