OMSA Central Analysis

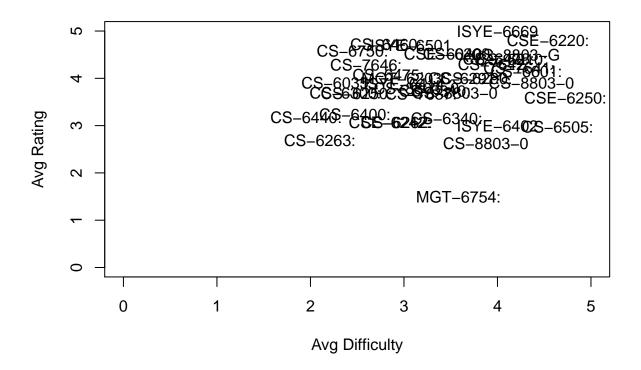
Dylan Peters

As a new student in Georgia Tech's Online Masters in Analytics (OMSA), I wanted to get an idea of the classes offered and how difficult they are and the students' satisfaction with the classes. Fortunately, there's an active community for the online students at omscentral.com, originally for the online Masters in Computer Science program. To compare classes, I copied the course overview from here courses and pasted it into Excel, then saved the file as a CSV. From there it was a simple matter to import it into R and examine the data.

Here I load the data and reset the column names to be more meaningful, then filter out the data without any ratings. Then I plot a simple chart of difficulty versus rating (each on a 1-5 scale).

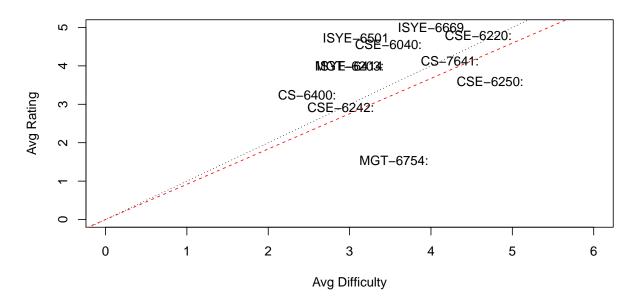
```
oms <- read.csv("OMSCentral.csv")
colnames(oms) <- c("ClassName", "Enrolled", "Reviews", "AvgWork", "AvgDiff", "AvgRating", "AboveB", "BelowC", "
oms <- oms[oms["Reviews"] > 0 & oms["AvgRating"] > 0 , ]
rownames(oms) <- oms$ClassName
oms["ClassNameShort"] <- substr(oms$ClassName,1,9)

plot(oms[,"AvgDiff"], oms[,"AvgRating"], type="n", xlab="Avg Difficulty", ylab="Avg Rating", xlim=c(0,5)
text(oms[,"AvgDiff"], oms[,"AvgRating"], oms[,"ClassNameShort"])</pre>
```



As you can see it looks pretty crowded. We can filter to only look at OMSA related classes

```
OMSA.Classes <- c("CSE-6040: Computing for Data Analytics",
                  "MGT-6203: Data Analytics in Business",
                  "CSE-6220: Intro to High-Performance Computing",
                  "CSE-6242: Data & Visual Analytics",
                  "CS-6400: Database Systems Concepts & Design",
                  "ISYE-6414: Regression Analysis",
                  "ISYE-6501: Introduction to Analytics Modeling",
                  "ISYE-6669: Deterministic Optimization",
                  "MGT-6754: Business Fundamentals for Analytics",
                  "CS-7641: Machine Learning",
                  "CSE-6250: Big Data for Health Informatics")
omsa <- oms[oms$ClassName %in% OMSA.Classes,]
plot(omsa[,"AvgDiff"], omsa[,"AvgRating"], type="n", xlab="Avg Difficulty", ylab="Avg Rating", xlim=c(0
text(omsa[,"AvgDiff"], omsa[,"AvgRating"], omsa[,"ClassNameShort"])
abline(0,1, lty=3)
abline(0,mean(omsa[,"AvgDiff"]) / mean(omsa[,"AvgRating"]), lty=2, col='red')
```



(The two items that are on top of each other are MGT-6203 and ISYE-6414)

In this chart, courses that are higher have a higher rating, and courses that are more to the right are more difficult. The black dotted line is the theoretical line where the difficulty equals the rating. Anything higher than that has a higher ratio of rating to difficulty. The red dashed line represents the average actual ratio—it indicates that on average the rating is less than the difficulty by a ratio of 0.92.

In general the rating and difficulty of the courses are somewhat related, with a correlation of 0.2682538. There are some outliers, specifically MGT-6754 which has a difficulty of 3.545 and a low rating of 1.545.

I think this is a nice way to examine the spread of the courses and get an idea of which ones require more course work as well as which ones are most popular. You might want to avoid taking two classes that are far to the right at the same time. Likewise, if you are trying to decide between two equally difficult courses, you might decide on the one that's higher on the chart.