



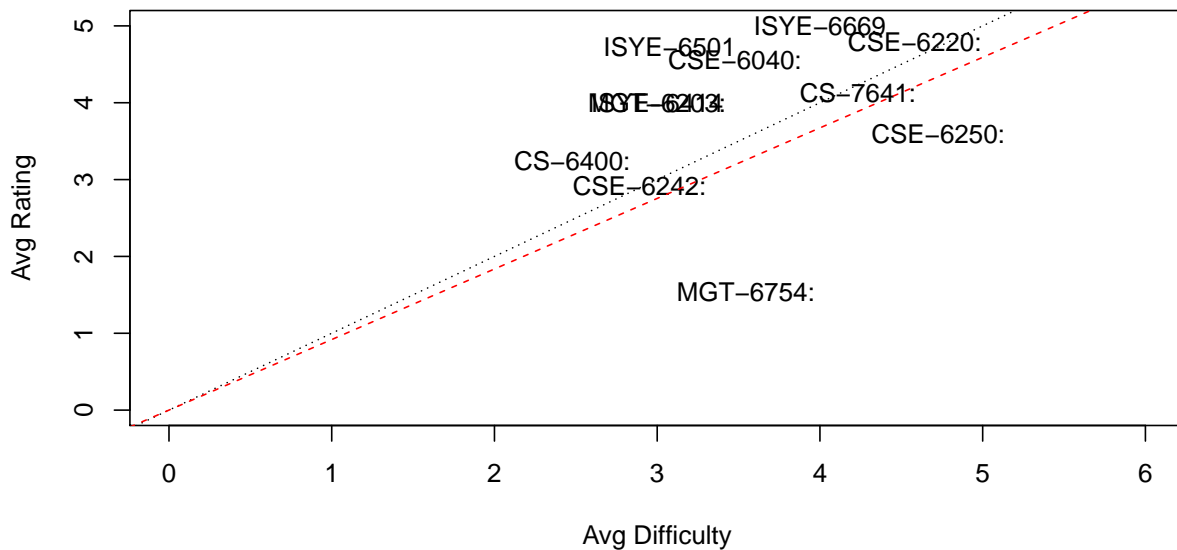
```

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(0,6), ylim=c(0,5))
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.