Week 4 blackboard

http://www.forbes.com/sites/joshbersin/2013/10/07/big-data-in-human-resources-a-world-of-haves-and-have-nots/

This article discusses how some companies are using data analytics to improve operational performances. It discusses how companies that are investing in data analytics are reaping benefits in market returns and are 4 times more likely to be respected in their industry. Their transformation was to use their research technology for 2 years on 480 large organizations.

In one example, they have found that by using analytics, top performing companies can retain top performers by giving these employees higher raises. They infer that retaining top performers is a key to a top performing company. Without any additional information the article states that top performers out deliver mid-level performers by a wide margins. This statement is unreliable since a wide margin means different things to different people. If I were making this argument I would back it up with real numbers, specific performance metrics in different industries and compare top level performers to mid-level performers. This data will help justify the evidence and counter this specific rebuttal.

Week 6

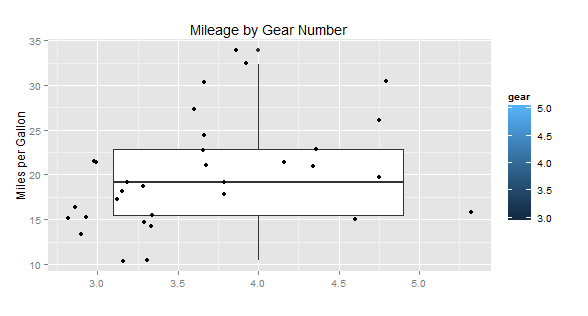
With many great options for graphs in R, having one simple plotting function for testing will be helpful in saving time to review data quickly. Qplot is an alternative ggplot when you just want to run through the command line. Both are part of the GGPLOT2 package. One difference is GGPlot uses data.frame and qplot uses vector, so you must convert to the a vector data type to run the code.

Reference: <http://www.statmethods.net/advgraphs/ggplot2.html>

qplot(gear, mpg, data=mtcars, geom=c("boxplot", "jitter"),

+ fill=gear, main="Mileage by Gear Number",

+ xlab="", ylab="Miles per Gallon")



References: <http://stackoverflow.com/questions/5322836/choosing-between-qplot-and-ggplot-in-ggplot2>

Week 7

The link below describes a process built with R to gather market sentiment and correlate to stock prices.

<http://www.rinfinance.com/agenda/2012/talk/Nagar+Hahsler.pdf>

The idea is to pull data from various sources via several R programming packages, merge the data, and run analysis. It takes pieces of sentences from Google, Yahoo, Twitter and rates the sentiment in the tags. It then tries to correlate market sentiment to the price of a particular stock.

The question is how can we differentiate sentiment from the market vs sentiment specific to the company? Many times stocks will go down based on world events that are out of the scope to the company itself. Can one model necessarily work on different companies that are in different industries? If not, why or how can we build a model that is universal to all sentiments

Week 8

Reproducible markdown tools can help in many areas. In my area we work with several systems tools. One area where we have been adding more resource to is a central database that brings all global system data sources together. I think a reproducible statistical markdown tool can help at the presentation layer by bringing everything together to build reports to track and forecast performance bottlenecks before they occur. The markdown’s ability to include scripts and graphs can help visualize a global environment from a birds-eye view. With its multiple document formatting options, we may also be able to include some drill down capabilities to get more detailed information.

QUIZ 8

1)

Beamer is greater math support it is built for statistics

You can use LaTex without knowing to much about it

2) Packrat does project isolation by recording and packaging other package versions that it depends on into one package. That way when you install your project it also installs any dependencies and correct package versions.

3) To keep code from displaying in R Markdown you use the following code:

{r, echo=FALSE}

4) To keep the results from displaying in R Markdown you use the following code:

{r, eval=FALSE)

5) You can use python as the language by using the following code:

{r, engine=’python’}

Quiz 9

Your task is to:

Identify a web site that has an API that you find interesting.[Please provide the URLs for both the web site home page, and the API main/ instruction page]

This url below has worldbank data that includes time series indicators, projects for worldbank operations and financial market data.

<http://data.worldbank.org/node/9>

The url for instructions and to build RestFul queries is:

<http://data.worldbank.org/node/9>

Using basic drop downs I was able to create this RestFul url query:

http://api.worldbank.org/countries/USA/indicators/9.0.Employee.T60?per\_page=100&date=2000:2014

How might you use some or all of the data to create value in an application?[One or two sentences]

This data would be good for financial applications for market trends. It can also be a good source to keep a pulse on world event.

What problems might you encounter building an application that used this web site’s data? [One or two sentences]

One problem may be that it has a plethora of data. Unless you know exactly what you are looking for and have a specific scope you may get lost in the data.

Write and share out a small R program that successfully pulls data from the web site using the provided API.A very short “beachhead” program will receive 100% credit here!

dats <- html("http://api.worldbank.org/v2/en/country/usa")

require(XML)

xmlChildren(dats)

Include thoughts on your initial experience pulling data from this site.  If you were in charge of the site’s API, what would you have done differently.

The data may need more categorization. Seems to be a wide variety of data and takes more time to decide on data then to use the API.

WEEK 12

* 1) **Describe the application** and how the data would be helpful in solving it.  This can be an existing application that you find interesting, or an interesting idea that you might want to build out.  *Due end of day Thursday, November 13th.*

Be sure to consider some of the following prompts:

* + Where would the data be sourced from?
  + Who would use the application?
  + In what way(s) does the [data in the] application create value for its intended users?
  + What would be some of the challenges in getting the data that the application needs?

Wearable Devices:

There is a trend that is happening with wearable devices and health. I brought one of these devices last year and it has become as much of a part of my daily life as my mobile phone. When I purchased the device last year the well-known products in the market at the time were, Up 24 by Jawbone and Fitbit. Both products offered a way to track your activity, log your meals and log your sleep patterns. I chose UP because of it’s ease of use, the Bluetooth sync capability and the detailed reporting. Apple has also recently jumped in to the wearable device market with their new apple watch.

The idea is to have this data sync to a secure location in the cloud where doctors can retrieve to review physical activities, diet and sleep pattern. This will help doctors provide better and more efficient services to their patients.

Some challenges will be data security and Health Care Privacy (HIIPA) concerns. The best way to deal with these concerns would be to just give patients the option to give this information or a select set of information voluntarily when they choose. It would be an open available cloud area where patients can place data as they choose. My theory is as patients begin to see the value of efficiencies and services doctors provide with this data, they will provide more data willingly.

Week 13

Brights spots

In current economic times, many companies can benefit from looking at internal and external data to build more efficient processes. While some companies are focused on cutting cost without investing in the bright spots in the organizations, companies like Google and Apple continue to invest in their employees to stay ahead of the market. All of this is as a result of looking at the bright spots internally while keeping an eye out on market trends.

Week 14

The 1999 article below is about a crash of a 125 million dollar orbiter mission to Mars. The mission failed because of a mismatch in the metrics used between the orbiter and the base station. The orbiter came 21 miles closer than expected and the mission failed.

In this scenario, although technology was used to build and monitor, the collaboration between humans failed the mission. In the end of the day, a program is only as good as the skill of the programmer that builds it.

http://www.cnn.com/TECH/space/9909/30/mars.metric.02/

As for the reading article for this assignment, I believe that although technology does cause loss of jobs at first, eventually jobs evolve. I don’t agree that technology makes us dumber. If this was the case, we would never have advanced or evolved as humans. It’s not advances of technology that makes us dumber, it’s our ability to accept that change and evolving is a part of life and this will determine whether we will remain ignorant or knowledgeable and be part of the change. The best thing is to accept and embrace it. How many grandparents do you see using their IPhone to lookup things they would have never known if it wasn’t for mobile devices?

As technology changes, jobs will change. Instead of supporting systems, you may instead be part of a team that builds automation tools to automatically and proactively recover systems when they are about to breakdown. Instead of going to the doctor for minor aches and pain you may now go online and talk to a doctor about your symptoms for a monthly or one time charge.

As humans we are always reaching, grasping, wondering, imagining, and we will continue to evolve and advance until the history book of our species is closed.