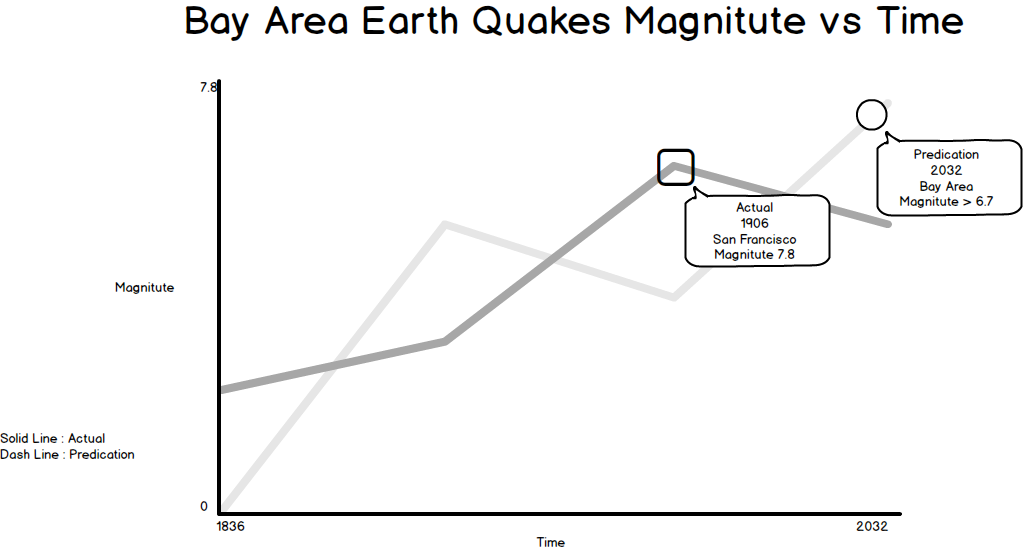
**Data Visualization Home Work 1**

Leonardo Ji

1. **Bay Area Earth Quakes**

This is a graph of earth quakes from year 1850 to 2000. I do not think this graph is a good visualization because of the following reasons:

1. The example graph used gradient colors from red to yellow to show frequency, and blue to white to show magnitudes. The gradient colors are confusing and harder to see the difference.
2. The example graph used 3D visualization for the time line and smaller the user of smaller box for each earth quake is confusing and unnecessary.
3. The time axis does not show actual beginning of the time. There are two earth quakes before 1850.
4. There is no title. There is no clear label for x axis (assuming y axis is time). Which one is first for the earth quakes occurred on the same year?
5. There are a lot of earth quakes between years 1850 to 1906. It was hard to distinguish them.
6. The scientist predicted 6.7 or greater earth quake by 2032. There is no clear indicator on the chart.



This is my redesign using Balsamiq Mockup. It has two lines but it can also be one. The future earth quake prediction as a dash line extension. 2D visualization is sufficient and less confusing. The use of tooltip also gives more detailed information instead trying to fit so many data in one area (year 1850-1906). Clearly labeled y axis helps to understand the difference in magnitudes. We can use square shape or round shape to denote frequency of earth quake and avoid overcrowding. The future prediction can be in dotted lines. It is easy to see exact year and frequency.

1. **Obama/Bush/Clinton/Regan second term job approval rating**

This graph shows four US presidents’ 2nd term approval ratings over four years period. I think this graph also has some deficiencies.

1. The x-axis does not have clear indicator of what are those numbers. Is it number of days?
2. The y-axis also missed label. Is it percentage?
3. The y-axis is truncated. It made the actual approval rating difference appear bigger than it actual is.
4. The change of US House seats by president’s party table on the top right hand side is confusing and easy to miss. If add a second bar chart under it would be more visible.
5. The colors of lines were red and darker red, blue and darker blue. We want to use more distinctive colors.

This is my resigned graph. I think either separate them into two graphs or have the bar chart under the approval rating chart is much easy to see than a small table on the top of right hand side. The approval y-axis is not truncated. The y-axis unit can be in weeks since poll number is only available weekly. For brevity I only entered approval ratings for each year here. I also added the title and axis labels for both charts. The color selections is more distinct, the same colors as the line chart.

1. **Diamonds were girls’ best friend**

This chart is designed by Nigel Holmes. The graphics artist designed it to be eye caching and entertaining. Chart like this is also called Chart Junk. Chart junk blend art and data visualization. According to Dirk Mjk [1] Cart Junk succeed in attention grapping but not very useful to understand the data. It is useful for magazine or tabloid publish a chart junk to grab attention but it is not suitable for professional business reporting. The use of chart junk here also misled actual reason why diamond price went up and down. During 1979-1981 US experienced hyperinflation and diamond price went up in a bubble. The chart misled the audience that diamond price went down because less consumer demand. However according to Nathan Yao’s “Data Point” book, the visualization has entertainment aspect. In this case a chart junk can grab attention, engage/entertain audience, and people tend to remember it.

[1] <http://dirkmjk.nl/nl/2013/10/chart-junk>

I redesigned it with simpler design for business setting. The chart is easy to understand. The actual reason of diamond price went up and down is explained. Both x-axis and y-axis are clearly labeled. The title explains what is the data measured. This chart is more suitable for business/professional settings.

Diamond price crashed in 1981 after bubble burst at the end of hyperinflation

1. **Bar chart of predicting user errors per error types**

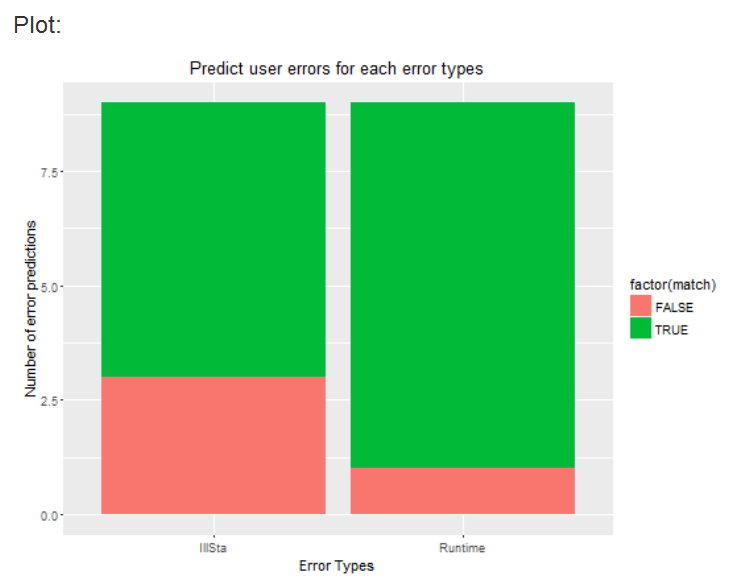
Dataset: Mock1477271919\_170291.csv

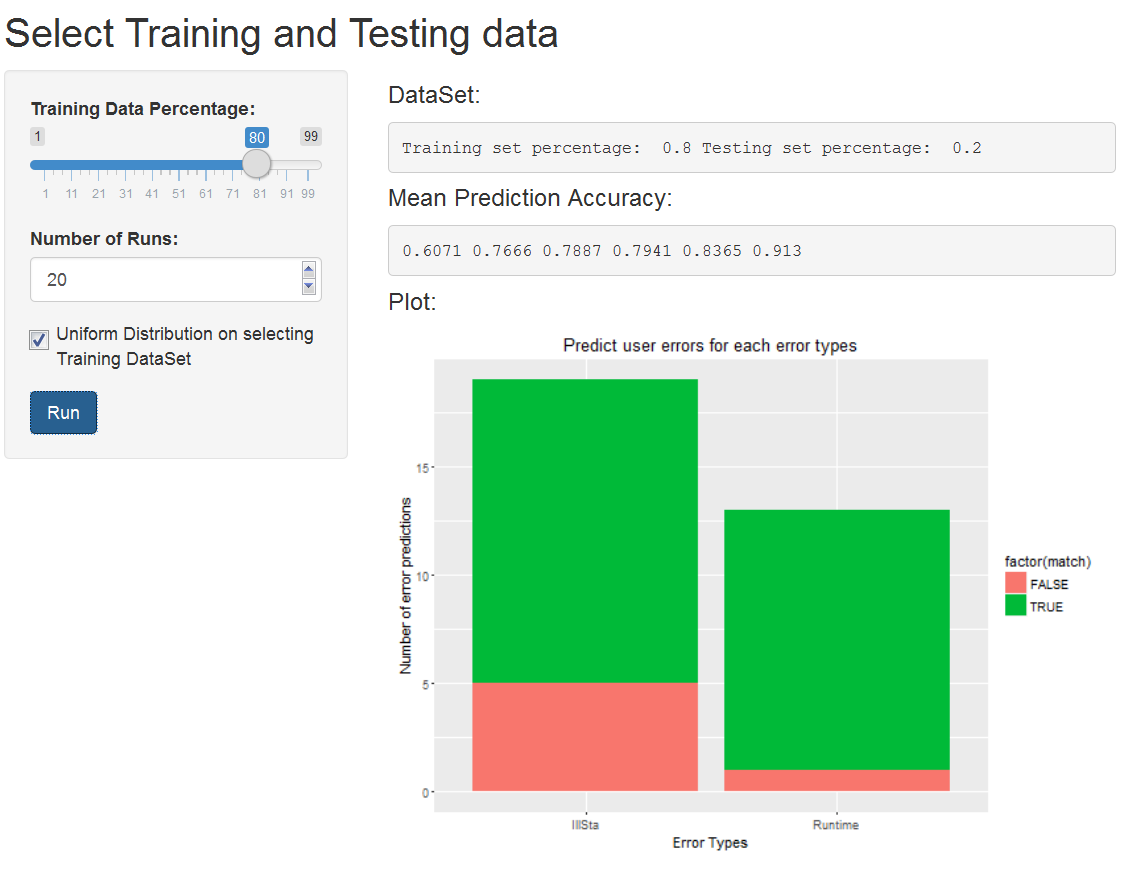
* What the data is:

This is dataset of user errors from server log files from mock domain. I exported the csv file from a mock domain at work. It is not real client data.

* What story your visualization tells:

I created this plot using ggplot2 in R. This is a bar chart for prediction of user errors (y-axis) for each error types (x-axis). The chart shows the number of correct predictions in color green vs the number of wrong predictions in color red. There are 133 rows of data in the csv.

* A jpeg or screen capture of each of your visuals: 



1. **World map of 2015 International trade**

Dataset: comtrade.csv

<http://comtrade.un.org/data>

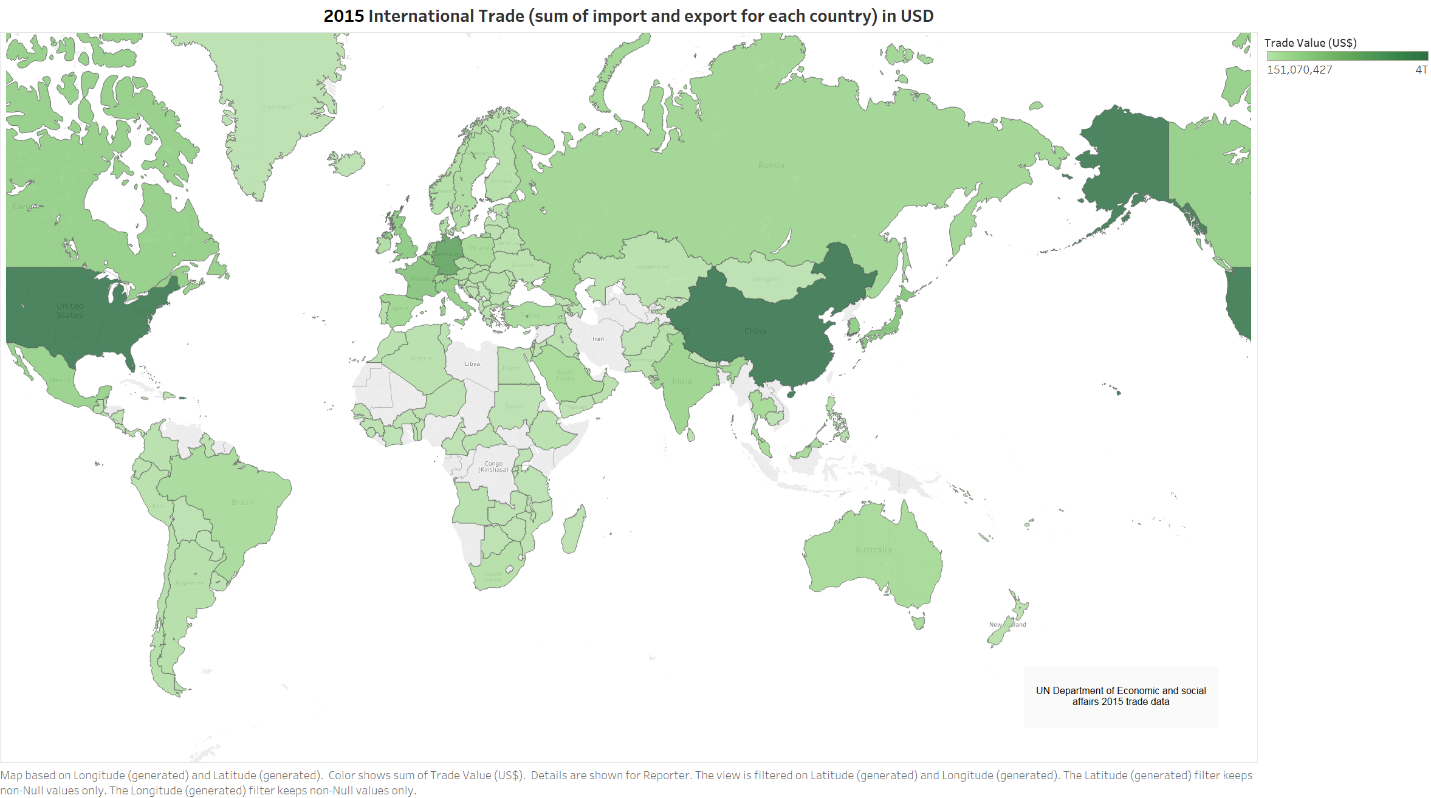
* What the data is:

This is data of international trade data from UN Department of Economic and social affairs. The data has all countries’ 2015 import and export in US dollars. The csv has 361 rows of data.

* What story your visualization tells:

This graph is created in Tableau. In 2015, US is the biggest importer and second largest exporter. China is biggest exporter and second largest importer. In 2015 both US and China have over four trillion dollars’ worth of international trade. This map shows dollar amount of each countries’ international trade (sum of import and export) in the global economy. The darker green country has higher international trade dollars than lighter green country.

* A jpeg or screen capture of each of your visuals:



1. **Heat Map of 2013-2014 NBA players’ Field Goal Percentage**

Dataset: 2013NBAPlayer.xlsx

<https://www.statcrunch.com/app/index.php?dataid=1096769>

* What the data is:

This is a data set of NBA 2013-2014 season players’ statistics. I found this data from statcrunch.com website. The dataset has 343 rows of data. Each row is a player’s statistics for the season.

* What story your visualization tells:

Using Tableau I created a heap map of NBA player field goal percentage (FGP). PFG tells how efficient a player is. I changed the default color to player by teams. The tooltip also shows Players’ Point Per Game.

* A jpeg or screen capture of each of your visuals:

