



DATA ANALYTICS

HOW WOULD YOU DEFINE DATA ANALYTICS?

WHAT IS IT USEFUL FOR?

WHAT IS THE PURPOSE OF IT?



DATA ANALYTICS

NATE: DRIVING DECISION MAKING WITH DATA, ESPECIALLY VISUALIZATION AND CLASSICAL STATISTICS

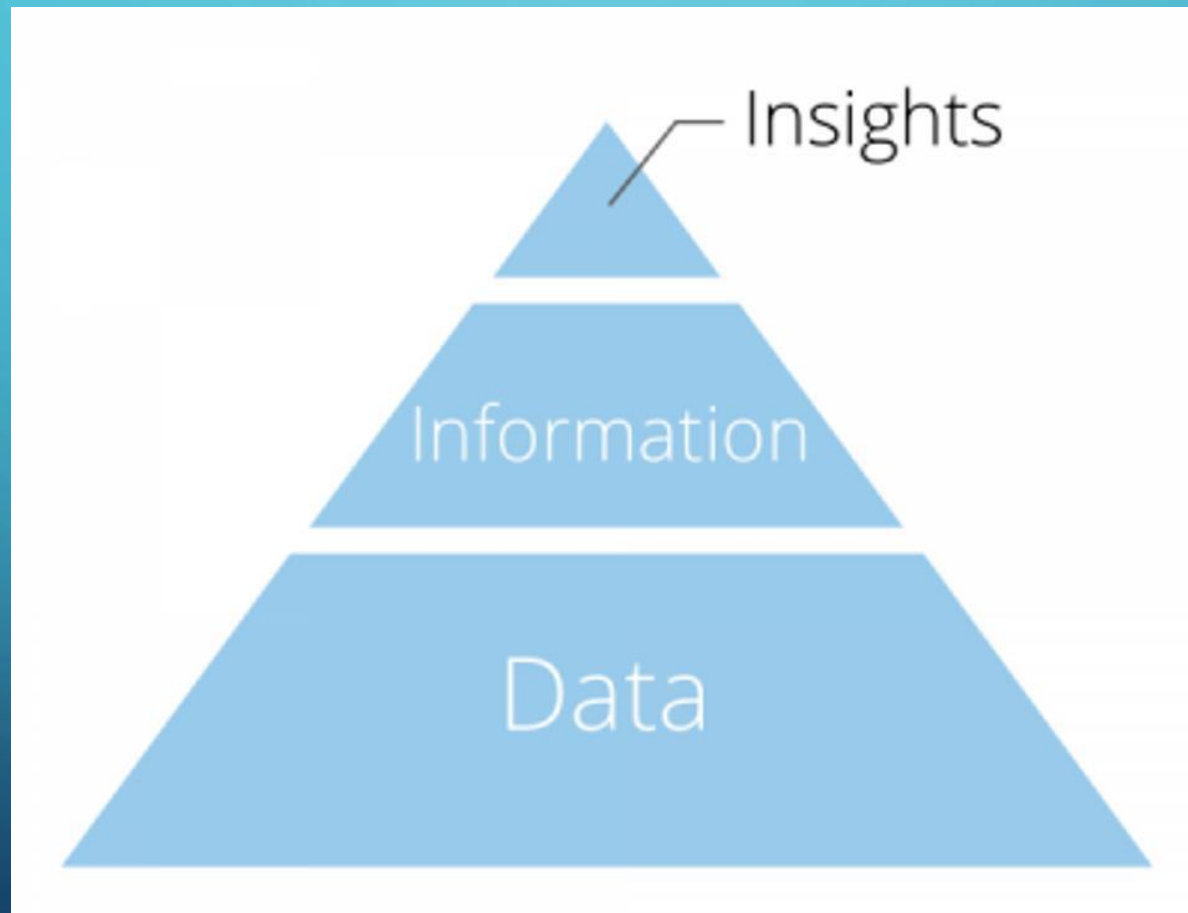
GEENA: TOOLS AND METHODS FOR ANALYZING DATA TO MAKE ACTIONABLE INSIGHTS FROM DATA

WHAT IS DATA ANALYTICS?

- From the reading: 'the scientific process of transforming data into insight for making better decisions'
- Other definitions:
 - **Data analytics** refers to qualitative and quantitative techniques and processes used to enhance productivity and business gain. (techopedia.com)
 - SAS: Big data analytics examines large amounts of data to uncover hidden patterns, correlations and other insights.

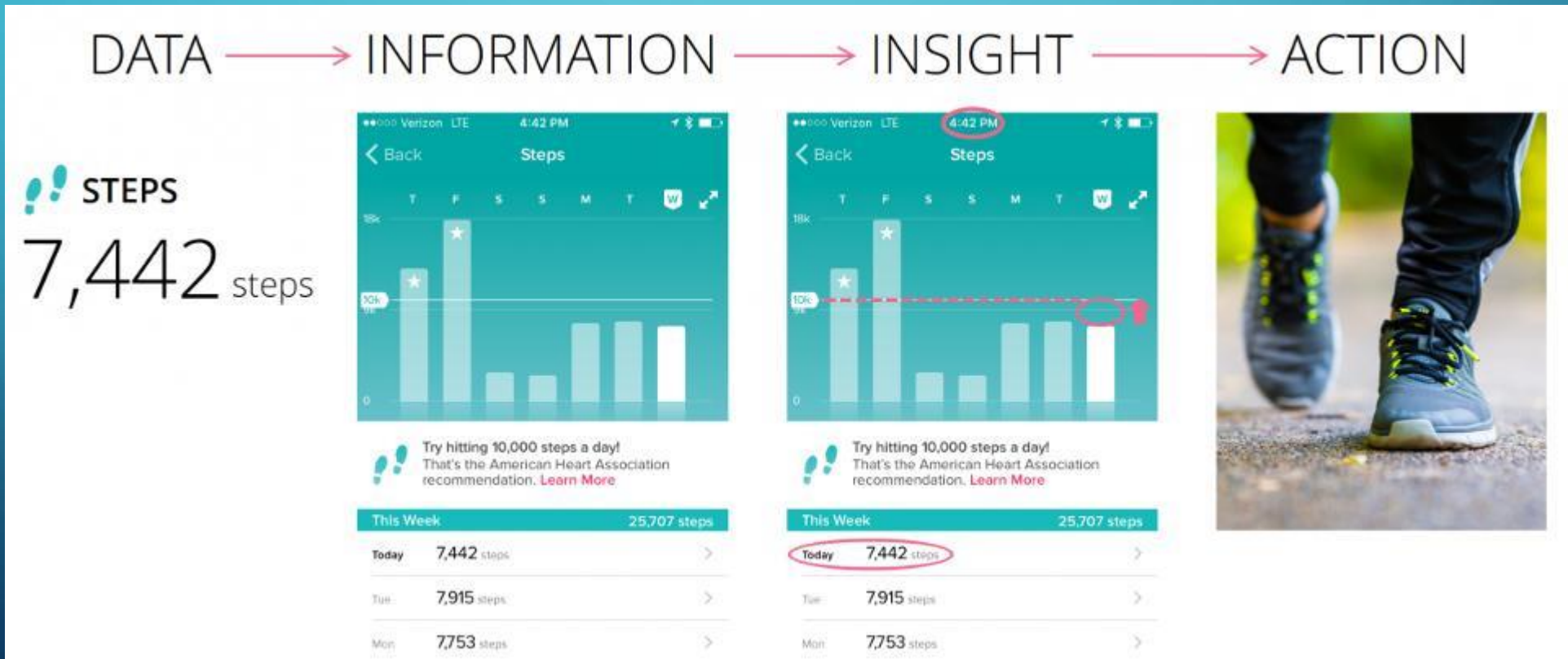
WHAT IS DATA ANALYTICS?

- Turning data into actionable insights



WHAT IS DATA ANALYTICS?

- Fitbit example
 - Decide to walk more based on graphs and statistics of number of steps



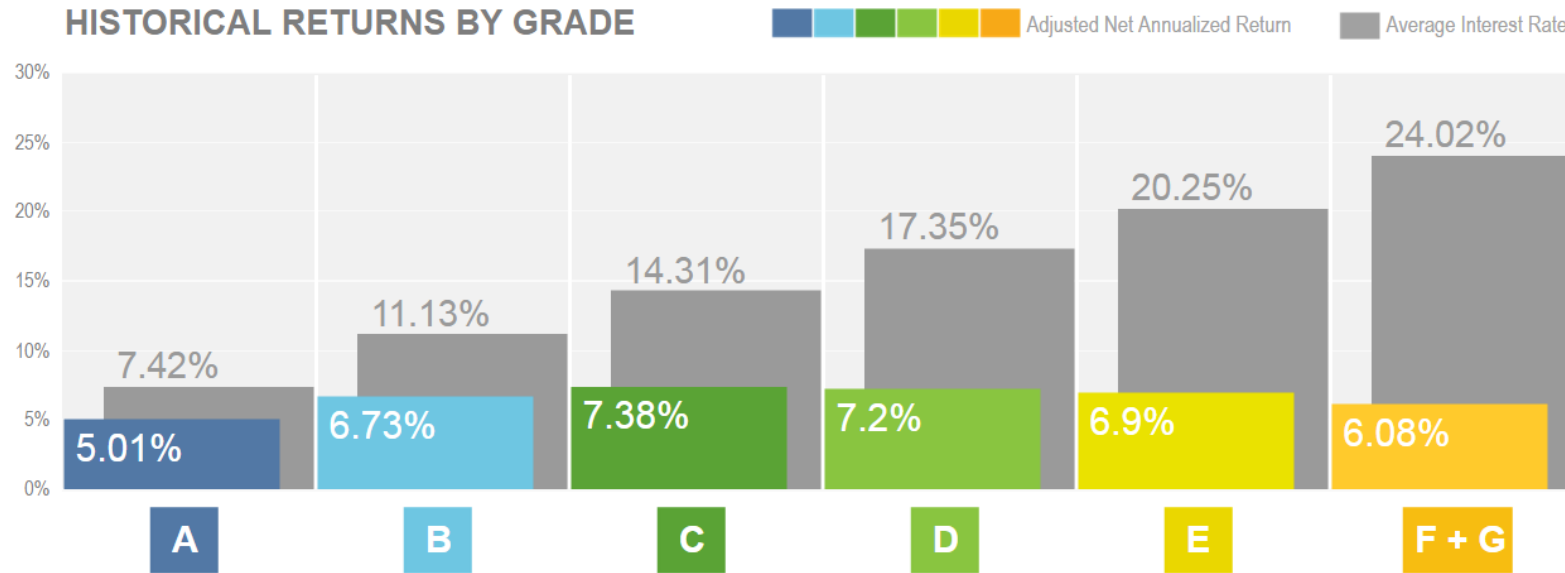
LENDING CLUB EXAMPLE

- <https://www.lendingclub.com/info/demand-and-credit-profile.action>

LAST QUARTER AVERAGE INTEREST RATE

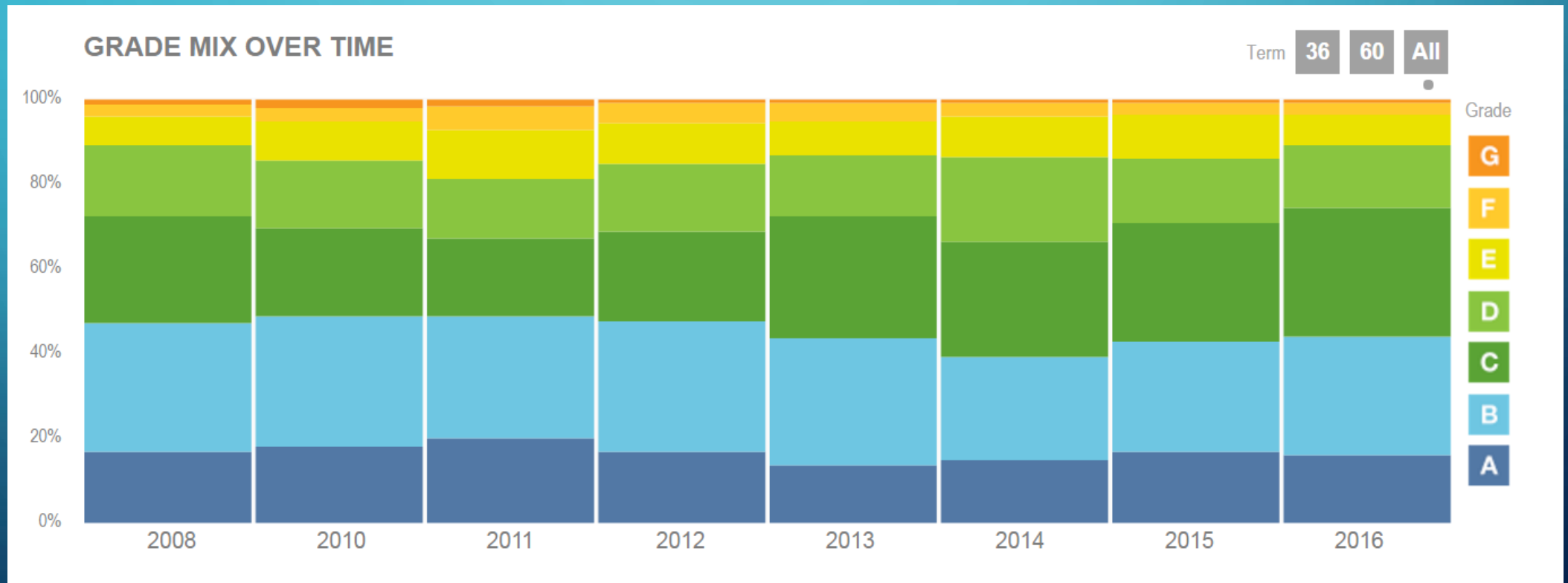
36-Month Loans: **12.61%** 60-Month Loans: **17.28%** All Loan Terms: **14.21%**

HISTORICAL RETURNS BY GRADE



BASED ON THESE TWO PLOTS, WHICH LOAN GRADE SEEMS BEST FOR THE AVERAGE INVESTOR?

- <https://www.lendingclub.com/info/demand-and-credit-profile.action>



SOME INTERESTING SITES FOR ANALYTICS

- <http://flowingdata.com/>
- <http://infosthetics.com/>
- [Kaggle.com/datasets](https://www.kaggle.com/datasets)
- Some analytics, but also regular articles:
- <https://www.priceonomics.com>
- [Fivethirtyeight.com](https://fivethirtyeight.com)

MOST/LEAST RELIABLE CARS

- Article written by a friend of mine (Alan Derk):
- <https://priceconomics.com/the-lemon-index-which-cars-have-the-highest/>
- What kind of decisions might we make with this information?

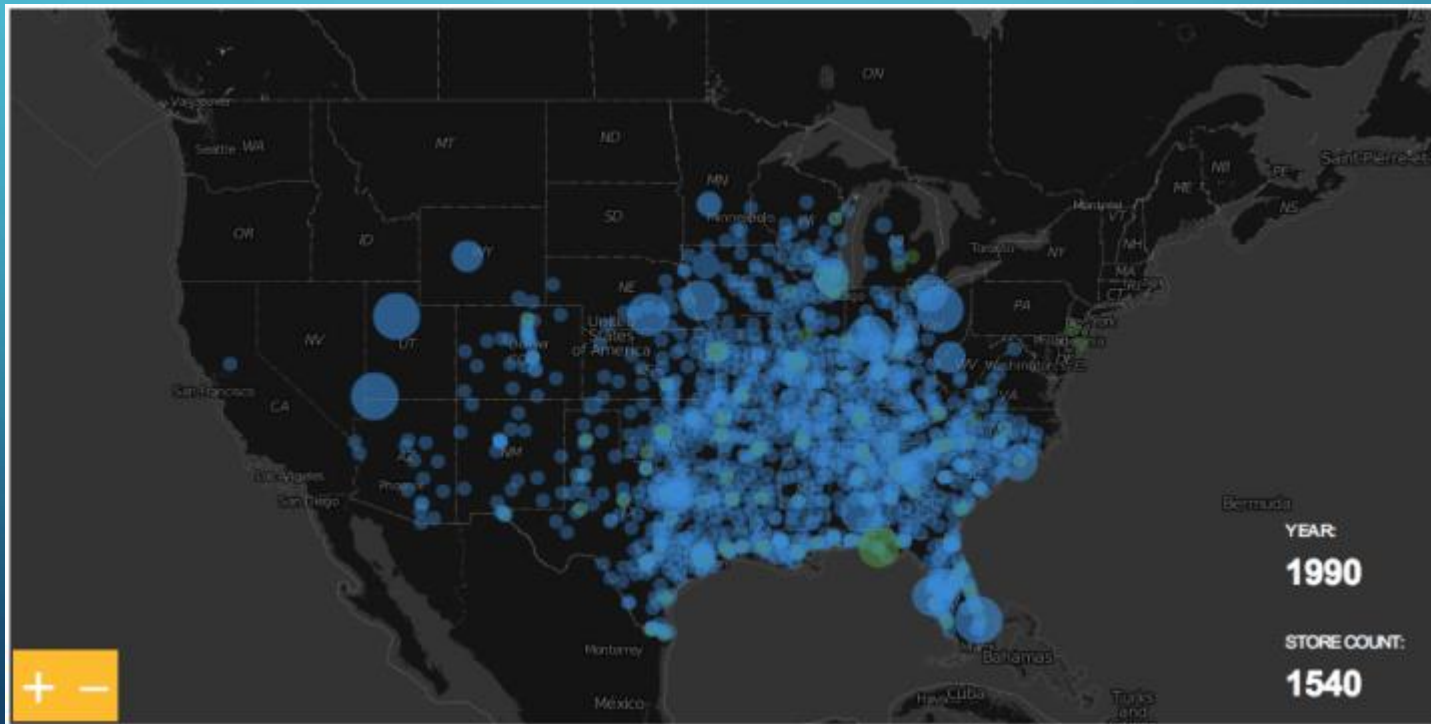
Which Car Models Have the Lowest Maintenance Cost?

Based On Total Maintenance Costs Over 10 Years

Rank	Model	Cost
1	Toyota Prius	\$4,300
2	Kia Soul	\$4,700
3	Toyota Camry	\$5,200
4	Honda Fit	\$5,500
5	Toyota Tacoma	\$5,800
6	Toyota Corolla	\$5,800
7	Nissan Versa	\$5,900
8	Toyota Yaris	\$6,100
9	Scion xB	\$6,300
10	Kia Optima	\$6,400
11	Lexus IS250	\$6,500
12	Nissan Rogue	\$6,500
13	Toyota Highlander	\$6,600
14	Honda Civic	\$6,600
15	Honda Accord	\$6,600
16	Volkswagen Jetta	\$6,800
17	Lexus RX350	\$6,900
18	Ford Fusion	\$7,000
19	Nissan Sentra	\$7,200
20	Subaru Impreza	\$7,500

FLOWINGDATA WALMART TIME LAPSE

- <http://projects.floodingdata.com/walmart/>
- If we are a store (i.e. Walmart), what sort of decisions might we make with this information? What are other types of data we should seek to better inform our decision?



FLOWINGDATA WALMART TIME LAPSE

- <http://projects.floodingdata.com/walmart/>
- I would say we could decide where to build new stores based on this data. We might want to also collect data on:
 - Store revenues
 - Demographics of nearby residents
 - Traffic patterns
 - Legal data (taxes)
 - Population trends

BUSINESS INTELLIGENCE TOOLS

- Anyone currently using any BI tools at work?
- Many BI tools, out there, for visualization, Tableau is probably top
- Some include machine learning (Alteryx, SAS) but are expensive (\$5k/yr +)
- There are dozens out there and new ones every year:
- <http://www.softwareadvice.com/bi/?more=true#more>
- <http://www.docurated.com/all-things-productivity/50-best-business-intelligence-tools>

LOWER-LEVEL TOOLS

- Much more customizable (can incorporate in a web page easily), but much more work
- Javascript libraries: D3, CartoDB, etc <http://selection.datavisualization.ch/>
- R: ggplot, googleVis, [etc](#)
- Python: plot.ly, bokeh, [etc](#)

JUPYTER NOTEBOOKS

- Interactive in-browser way to run code and see the output
- Originally for Python, now has many kernels (languages)
- Why is it nice? We can run a bunch of code, save the output – including plots, and easily show it to someone else later.
- Be careful of the order you run cells in!

INSTALL JUPYTER AND R KERNEL

- Install jupyter notebook if it doesn't already work (comes with Anaconda distro -- try `jupyter notebook` in a terminal or command console)
- You can also do `pip install jupyter` (`sudo pip install jupyter` in Linux)
- Watch out for bugs with plugins in chrome (Lastpass doesn't play well with jupyter)
- Install R kernel:
- <https://irkernel.github.io/installation/>

INSTALL JUPYTER AND R KERNEL

- Install jupyter notebook if it doesn't already work (comes with Anaconda distro -- try `jupyter notebook` in a terminal or command console)
- You can also do `pip install jupyter` (`sudo pip install jupyter` in Linux)
- Watch out for bugs with plugins in chrome (Lastpass doesn't play well with jupyter)
- Install R kernel:
- <https://irkernel.github.io/installation/>