PROCESS BOOK

Week 9: Project Proposal

1-2 paragraph abstract. The abstract should include your **motivation** for doing this project, **what you are trying to do** and **which goal you want to achieve**. Also include a description about your **planned datasets** and **where you will acquire them**.

Due 3/21/2016 -- submitted at https://docs.google.com/forms/d/1owRoNNi37_jcOtgVSJWBgT7RKXZU0pvLG7HSOFLvAWw/formResponse?edit2=2_ABaOnuexff7t2o2f-uhWvwASyRuZz6Ab9viHC5pC0fLDd-jhUZn8oigxBsOsZLI

Let's extract this to another word doc, finalize the edits and use that as the deliverable!

I'd like to examine the relationships between US parties (President, Congress) in office and economic well-being (income, income growth, unemployment) and macro indicators (like gdp, gdp growth, deficit, deficit as % of gdp, deficit growth). The goal is to separate fact from fiction when it comes to claims that two parties make vs empirical data from their record -- statements about small government, job creation, growth of the economy, etc. that, I feel, get over-abused.

E.g.
Select (Median) **Income (Growth)**, and then create a bar graph with all the presidents and how they stack against each other (sort high-low, chronologically)

I think given it's an election year that could be very interesting to the general public. I am not certain whether the scope will be sufficient so that's a risk but we can certainly expand when we start brainstorming...

Bar graphs and Charts will

GDP for the world -- time-lapse

Data Sets:

Most of the data is available from census.gov, e.g.

- Median and Mean **Income** from **Table P-4. Race and Hispanic Origin of People (Both Sexes Combined) by Median and Mean Income [XLS - 58k] -- this is what we need, and we can calculate growth/decline as percentages, starts from 1947 to 2014**
  - Real Median Household Income in the United States -- https://research.stlouisfed.org/fred2/series/MEHOINUSA672N
- **Unemployment** -- numbers, percentages and percentages growth -- a lot to choose from at http://www.bls.gov/data/#api
  - Percentage is available at http://data.bls.gov/pdq/SurveyOutputServlet from **1947 to 2016**
  - **Labor Force Statistics from the Current Population Survey**
  - **Employment-Population Ratio**

- **Budget** Deficit/Surplus -- numbers and percentages of GDP, -- https://www.whitehouse.gov/sites/default/files/omb/budget/fy2017/assets/hist01z2.xls Table 1.2—Summary of Receipts, Outlays, and Surpluses or Deficits (-) as Percentages of GDP: 1930–2021 41 K XLS
- **Federal Debt: Total Public Debt as Percent of Gross Domestic Product**
- **Consumer Price Index for All Urban Consumers: All Items** https://research.stlouisfed.org/fred2/series/CPIAUCSL#

- **GDP**, GDP Growth
  - https://research.stlouisfed.org/fred2/series/GDP/downloaddata has nominal numbers (units) from 1947 to 2015; also percent change, etc.
  - There's more from the same source -- **inflation**, unemployment, consumer price index

President, Congress Control (question: do we want to include that?): that can be acquired from many places or just created, i.e. we don't need to get it from some place -- it's public record, and we should probably limit from 1947's or whenever the datasets start from.

https://en.wikipedia.org/wiki/Consumer_price_index#/media/File:US_Consumer_Price_Index_Graph.svg

Few inspirations:
http://www.usnews.com/news/blogs/data-mine/2015/10/28/which-presidents-have-been-best-for-the-economy -- but it's not interactive and the conclusions are drawn out by the authors instead of let the data speak to the audience

http://us-presidents.insidegov.com/compare/2-3-4-7-13-21-26-39/Barack-Obama-vs-Gerald-Ford-vs-Jimmy-Carter-vs-Richard-Nixon-vs-Ronald-Reagan-vs-George-H-W-Bush-vs-Bill-Clinton-vs-George-W-Bush -- does a nice job to tell a story but i find the visualizations (charts) lacking and in need of better interactivity.

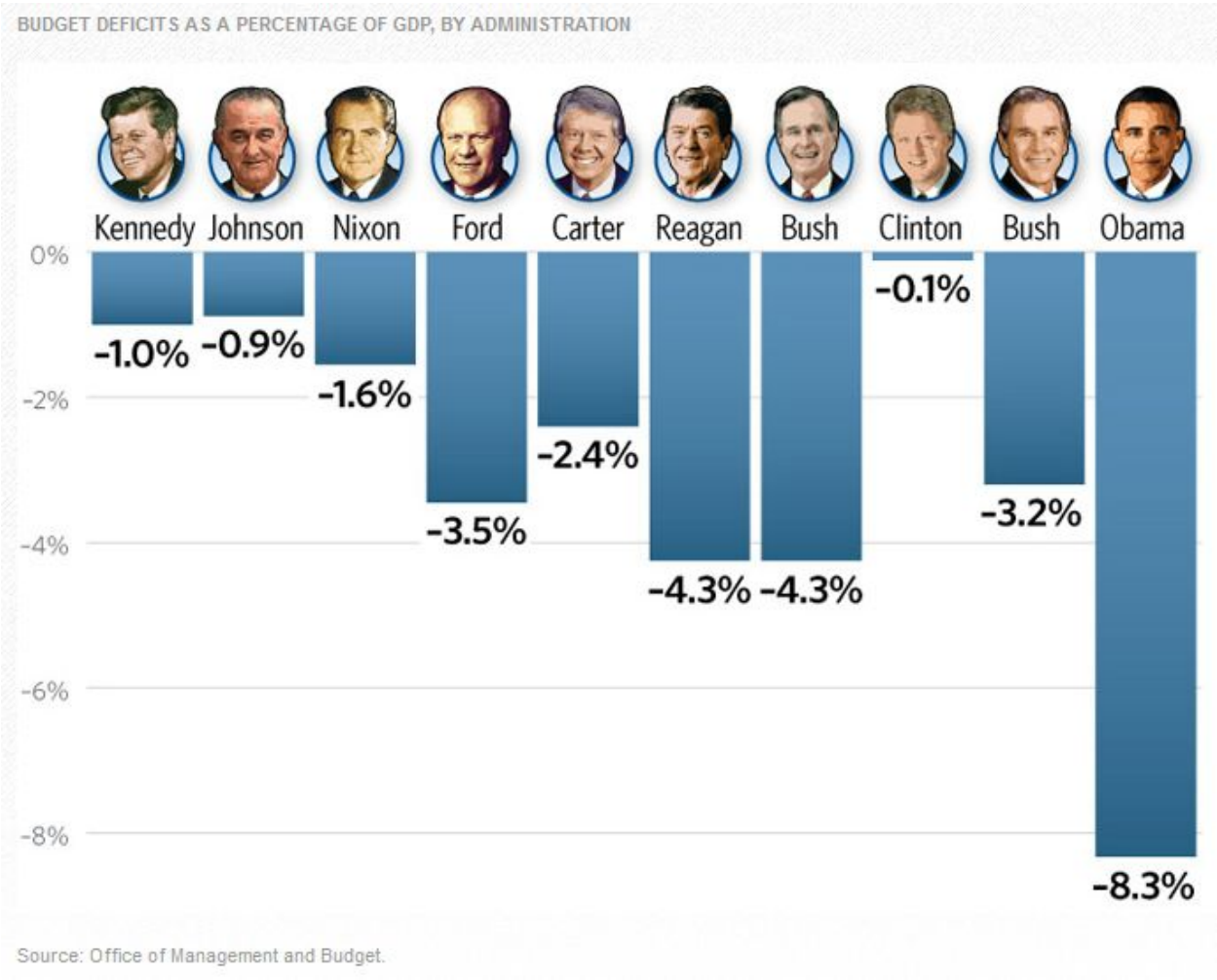
http://www.dailykos.com/story/2012/9/2/1127055/-Which-party-is-best-for-the-economy-It-s-not-even-close#one

Time Availability:

Chris: except for Mon-Tuesday, generally available...
Phone 765 479 2439
Email and google hangouts - chris.m.servin@gmail.com
Anytime with at least 2 hours advance notice. Prefered dates Wed thru Sun - 765-479-2439 or chris.m.servin@gmail.com

Ghaith: mg@takrity.com

Vlad: I'm CET too (like Chris) but still on ET working-wise; Available almost daily between 12pm and 5pm ET (except on 3/22/2016), available on:
gchat/email: vgeorgiev@g.harvard.edu,
skype: vielgigeorgiev
Cell: 646.373.4931
==

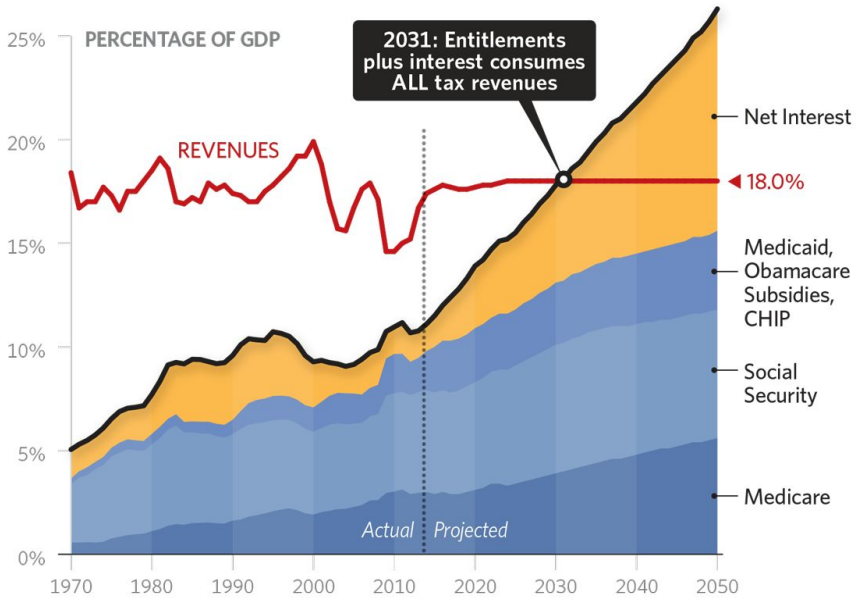


http://www.zerohedge.com/sites/default/files/images/user3303/imageroot/2012/05/20120503_FedBudgDebt4.png

CHART 4

All Tax Revenue Will Go Toward Entitlements and Net Interest by 2031

In less than two decades, all projected tax revenues would be consumed by three federal programs (Medicare, Social Security, and Medicaid, which includes CHIP and Obamacare) and interest on the debt.



Sources: U.S. Office of Management and Budget, *Budget of the United States Government, Fiscal Year 2015: Historical Tables*, 2014, Tables 8.4, 8.5, and 10.1, http://www.whitehouse.gov/omb/budget/Historicals (accessed January 16, 2015), and Congressional Budget Office, *The 2014 Long-Term Budget Outlook*, July 14, 2014, http://www.cbo.gov/publication/45308 (accessed January 16, 2015).

From http://budgetbook.heritage.org/wp-content/uploads/2015/01/SR-budget-book-2015-chart-4-1600.png

Data sets are readily available from a number of government and non-government sources.

For our project we will do few data manipulations so we can use all data sets when we need them. The first manipulation will be filtering the dates. We will establish starting and end points of our domain (years), and filter out any information outside those years.

The next data manipulation will be sorting the economic indicators by presidency term. As mentioned in our goal statement, we want to show key relationships between US parties in office and economic indicators. Most of our data sets contain those key indicators separately. For example the World Bank is a source where we could get the gross domestic product (GDP) for the USA in the last 50 years, and the US Bureau of Labor is a source where we could obtain the unemployment rates in

the past 50 years. We will need to sort both data sets by presidency periods and group them together, so we can have GDP and unemployment in the same data sets, sorted by period.

Week 10: Assignment Instructions

Due Monday Mar 28 by 11:59pm!!

This is a team assignment, please submit only once! The submission should be done by the assigned person in your team who is doing the submissions.

For this submission please prepare and upload a **single zip file** of the following:

- Scan in a signed copy of a '**team expectations agreement**'. The team expectations agreement should be a single piece of paper and include your names, and list the rules and expectations you agree as a team to adopt (e.g., preparation for and attendance at group meetings, making sure everyone understands all designs and code, communicating frankly but with respect when conflicts arise,...). Each team member should sign the sheet, indicating the acceptance of these expectations and intention to fulfill them.
- Definition of **goals and tasks** of the final project **(1-2 pages)**
- At least **3 sketches** of visualization ideas for your project
  - 2-3 each (Sat 12pm)
  - Discuss and select final 3
- A **description of your data** and where you will get the data from (at least concrete ideas on where to acquire the data)
- Start a **process book**, and submit your entries for the first week
- Prepare slides for a **3 minute presentation** of goals/tasks/sketches/data for peer feedback in your studio and include the presentation slides in

your submission (DCE students: please prepare a 3 minute screencast instead)

Team Expectations Agreement

VG's expectations:

- **Communication** -- convergence of platform where we do the work (e.g. google docs) and email (google docs has email alerting). I find writing emails explaining what we mean in a separate document superfluous -- it can be avoided! Same goes for our code -- let's use whatever version control makes sense that will have email alerting capabilities -- e.g. GIT does. The two platforms have by definition timestamps and who/what contributes -- that keeps everyone on their best behavior.
    - Prefer to over-communicate than make assumptions about expected contributions, deadlines (luckily those are externally imposed for the milestones but we still need to manage our contributions' task executions (each must be as detailed as possible, and also have a deadline, including who's it assigned to)
- Regular **meetings** -- at least twice weekly we need to meet. Other asynchronous communication/ad-hoc meetings.
    - I prefer **Google Hangouts** because we can instantly share our screens and not just talk about what we mean -- showing works better
    - If possible I'd like to have our initial meetings with **video conference** -- helps get to know each other and consequently work better together!
- Fair **contribution** from each team member -- set expectations including what happens when we don't contribute.
    - Preparation for meetings
- **Conflicts** -- we're all adults and professionals at arguably the most reputable academic institution in the world so I'm sure that we can handle the assignment (we're properly motivated, correct?) and any conflicts that may arise. Respectful communication and making sure that everyone is clear what's expected of them should ensure we don't run into any problems in my humble opinion!

My understanding is that although each one of us can do the entire project individually working in a team environment is another equally important 'project' that can teach us even more valuable lessons than d3, visualization, etc. So I am looking forward to the opportunity to learn as much as possible from you individually as well as from working as a team together. (See my response at the end of the document -- full disclosure and all that...)

Chris's Expectations

Completing all team assignments at least 1 day before the due date - to allow for any changes or revisions.

Regular meetings - at least twice a week.

Ghaith Takrity

Goals and Tasks

1. First Team Meeting:
    a. VG: I can do any afternoon from 1-6pm ET
    b. Chris: Same as VG - any afternoon from 1 - 6pm ET
    c.

Define the Goal -

The goal of this project is to help people understand the impact that US parties have had in the periods they have governed the country. We will visually examine **key relationships between US parties in office (**President, and possibly Congress (Senate/House of Representatives)) and various **indicators** (**individual** economic well-being (income (nominal (1000, 950, 990...) vs income growth (-5%, 5%,...) unemployment) and various other **national** macro indicators (like gdp, gdp growth, jobs, budget deficit/surplus, deficit as % of gdp, deficit growth)). The goal is to present the audience with visualizations to help separate fact from fiction when it comes to claims the two major US parties make -- statements about the economy, job growth, small government, job creation, etc. -- statements, that in our minds, are over-abused and need to be approached critically by all concerned and continually at that.

Disclaimer -- keep our biases in checks; balance

Tasks - (update March 27 - Chris)

For our project we will do few data manipulations so we can use all data sets when we need them. The first manipulation will be filtering the dates. As mentioned in the goal, we will use various indicators provided by US agencies. The data sources of these indicators provide historic data based on the time they were established (e.g. US Bureau of Labor - 1884) or the date they started gathering information (e.g. Consumer Price index - 1913) and of course most these start dates vary in all data sets. We will establish starting and end points of our domain (years), and filter out any information outside those years.

The next data manipulation will be sorting the economic indicators by presidency term. As mentioned in our goal statement, we want to show key relationships between US parties in office and economic indicators. Most of our data sets contain those key indicators separately. For example the World Bank is a source where we could get the gross domestic product (GDP) for the USA in the last 50 years, and the US Bureau of Labor is a source where we could obtain the unemployment rates in the past 50 years. We will need to sort both data sets by presidency periods and group them together, so we can have GDP and unemployment in the same data sets, sorted by period.

As far as visual manipulations, one of the main techniques we will use is selection and navigation. As stated in our goal, we want to separate facts from fiction, therefore our visualizations will be interactive to allow users select different time frames or select specific presidents for a side-by-side comparison. Therefore we need to allow users select bar charts, data points in a line chart or states in a map to find more information about the charts (e.g. tooltips). We will also allow users to navigate the visualization by using check boxes, radio buttons, and time lines.

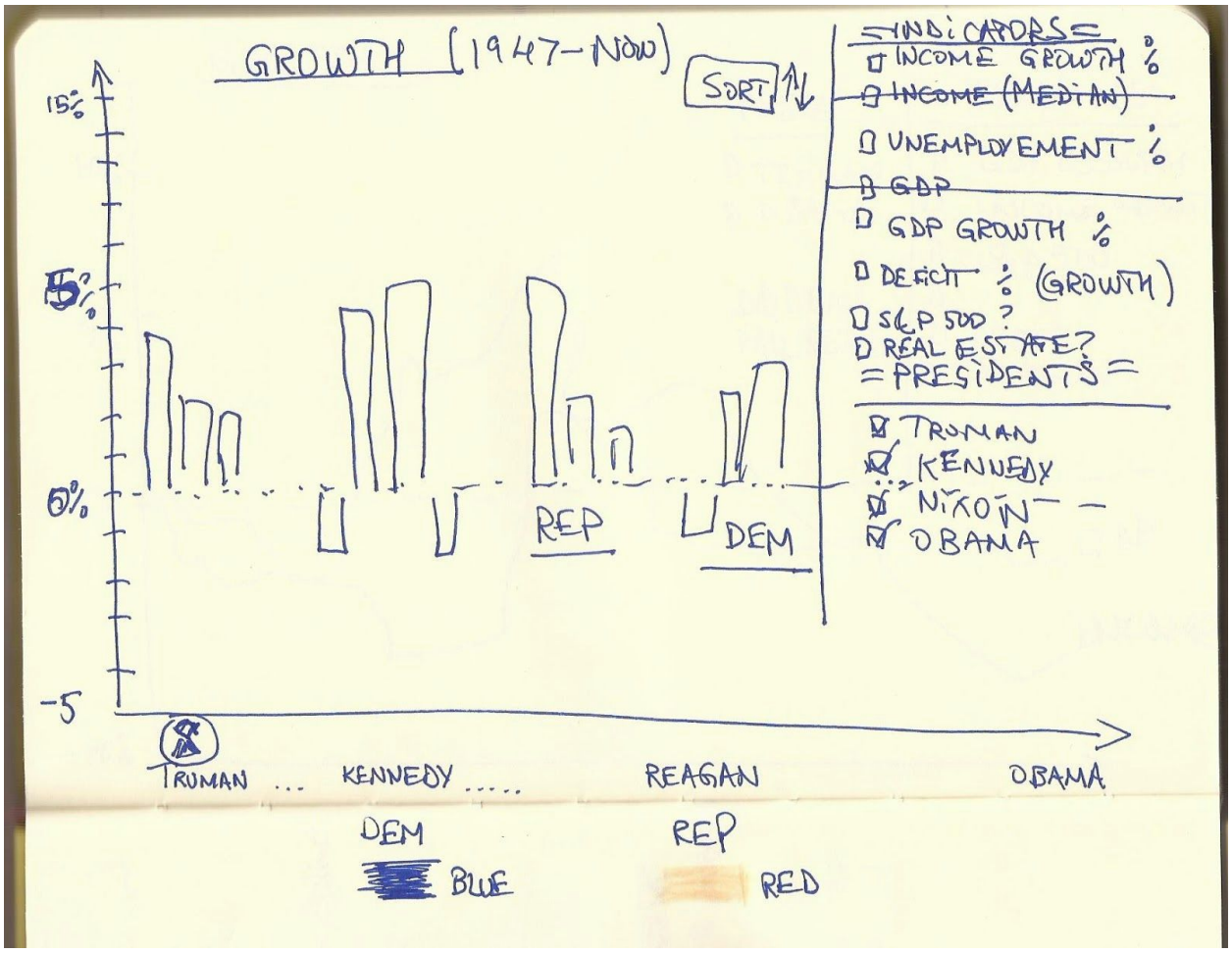
Tasks - Project Proposal

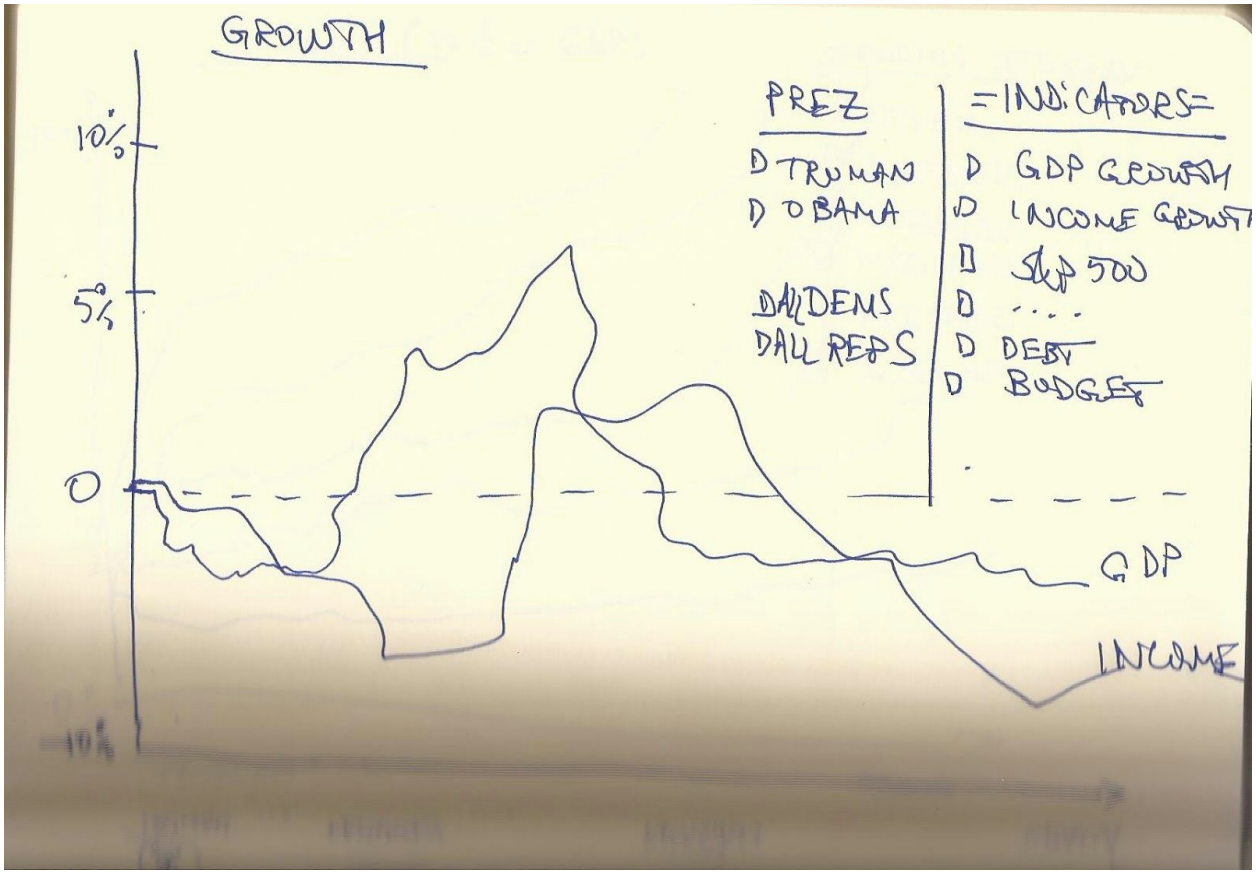
1. Team Formation -- including setting our ground rules and expectations agreement
2. Project Plan
3. Create a process book and document all research, sketches, note meetings, presentations, prototypes and final designs.
4. Gather data
5. Analyse the data
6. Define a concept for the website that will present the gathered data to the audience.
7. Gather information that would help explain our concept
8. Sketch potential visualizations.
9. Get feedback from peers and instructors
10. Prototype the site and visualizations
    a. We can use https://pages.github.com/ for publishing too

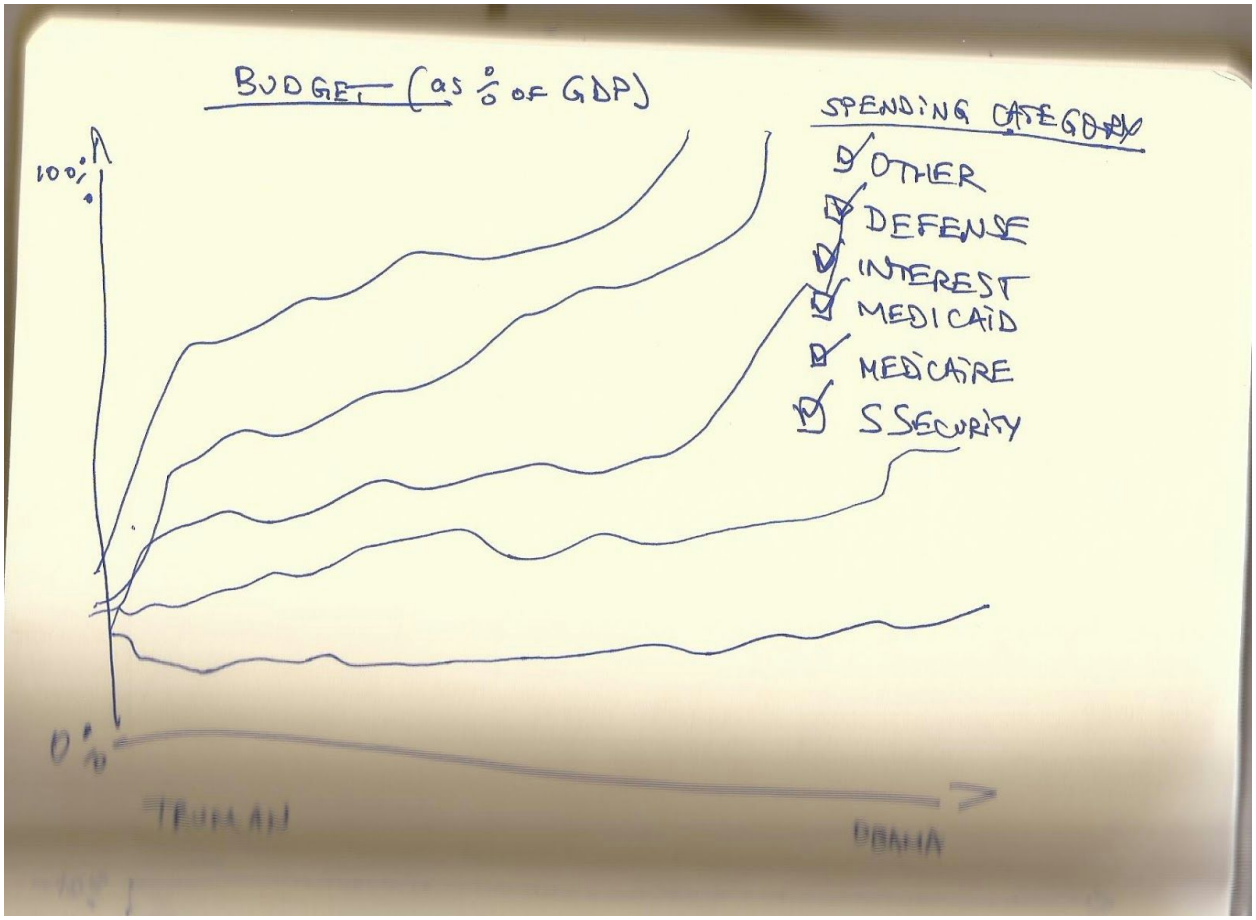
11. Debug visualizations.
12. Proofread the texts that will be included in the site.
13. Narrow the final texts that will be used in the site.
14. Finalize version 1 of the project
15. Get feedback from peer and instructors
16. Complete the project.

3 Sketches

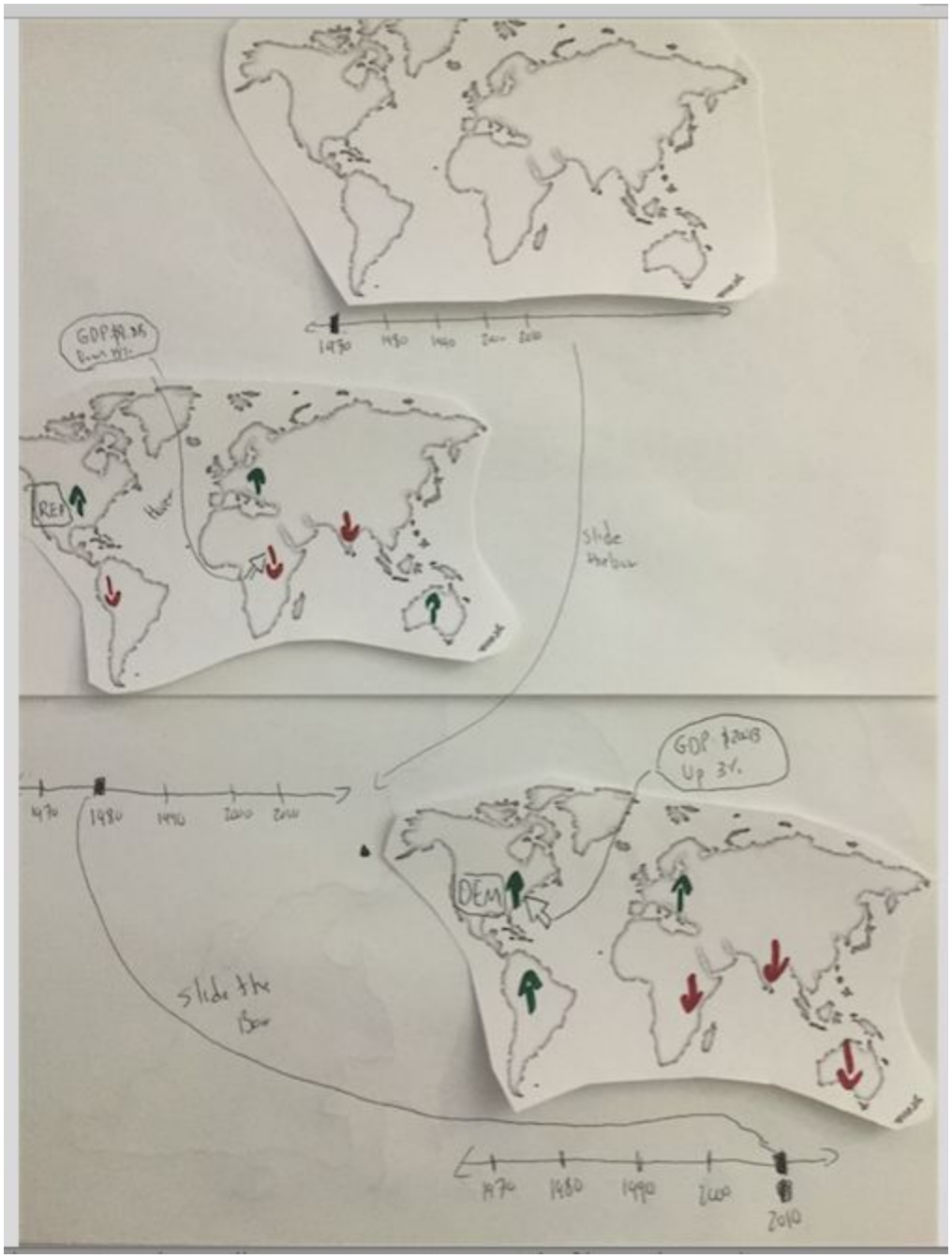
VG:




GROWTH
10%
5%
0
-10%
PREZ
D TRUMAN
D OBAMA
DALDEMS
DALL REPS
=INDICATORS=
D GDP GROWTH
D INCOME GROWTH
D S&P 500
D ....
D DEBT
D BUDGET
GDP
INCOME

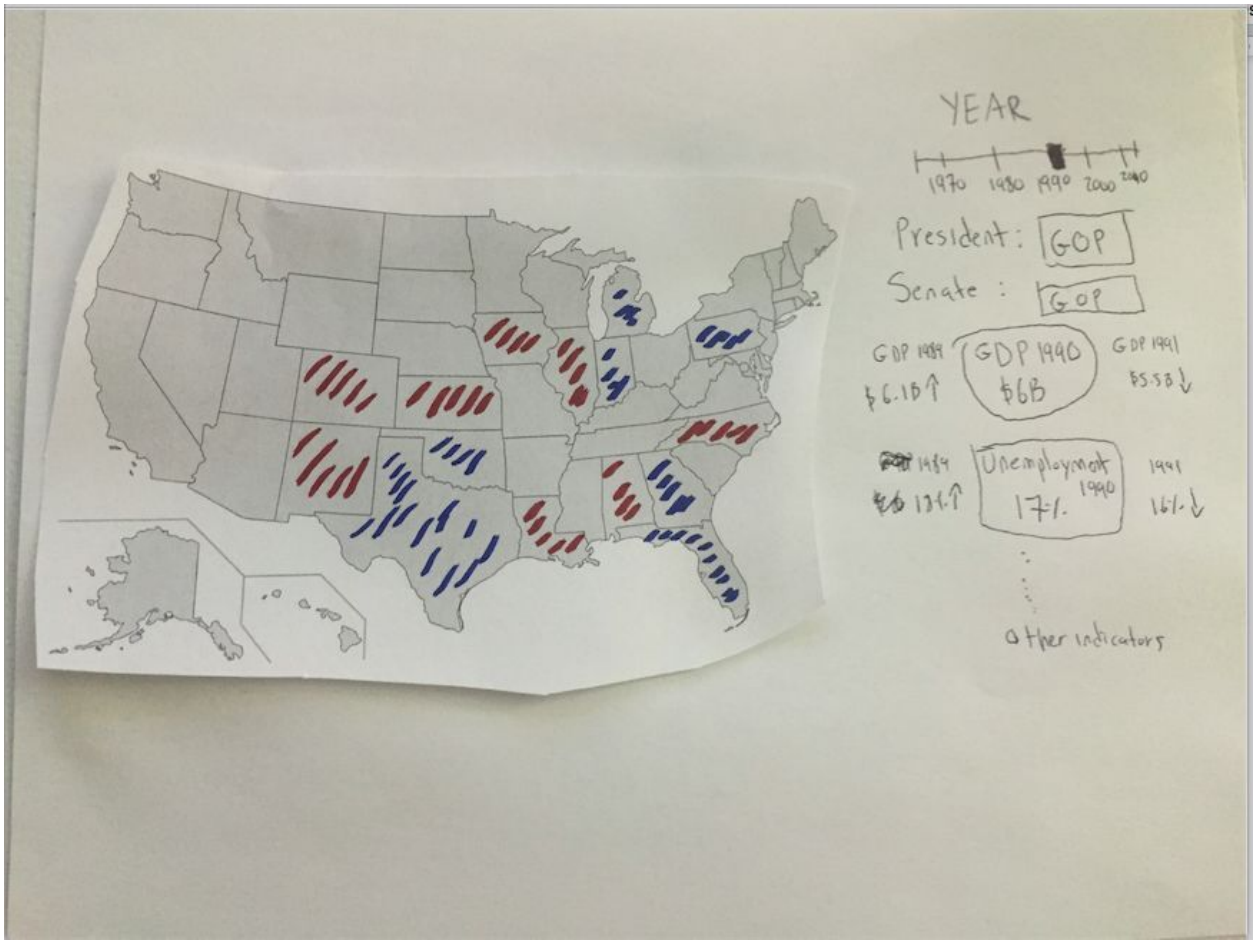

BUDGET (as % of GDP)
SPENDING CATEGORY
OTHER
DEFENSE
INTEREST
MEDICAID
MEDICARE
S SECURITY
100%
0%
TRUMAN
OBAMA

Chris Viz



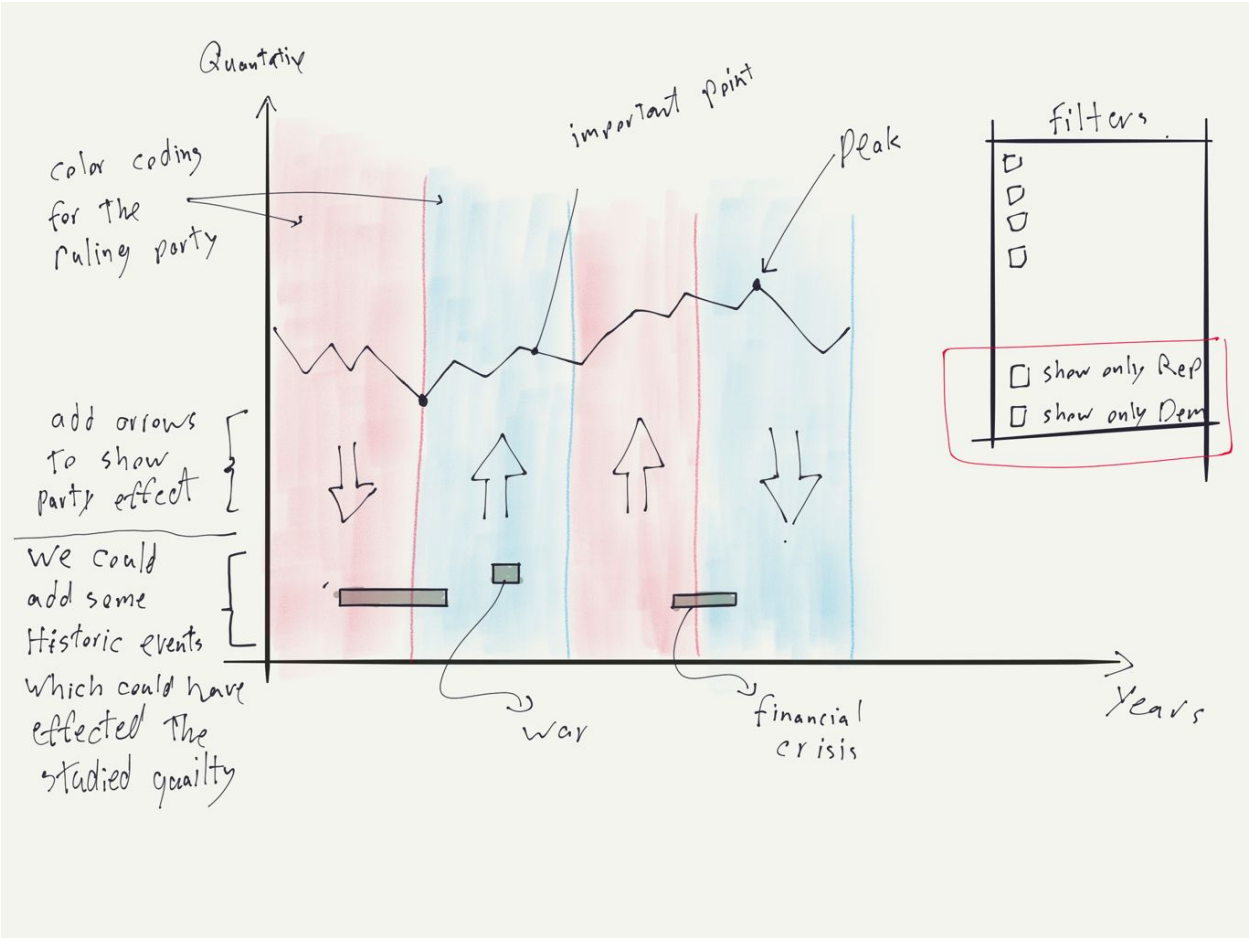
-- final thoughts -- there are **of course** multiple causes and here consider the world gdp

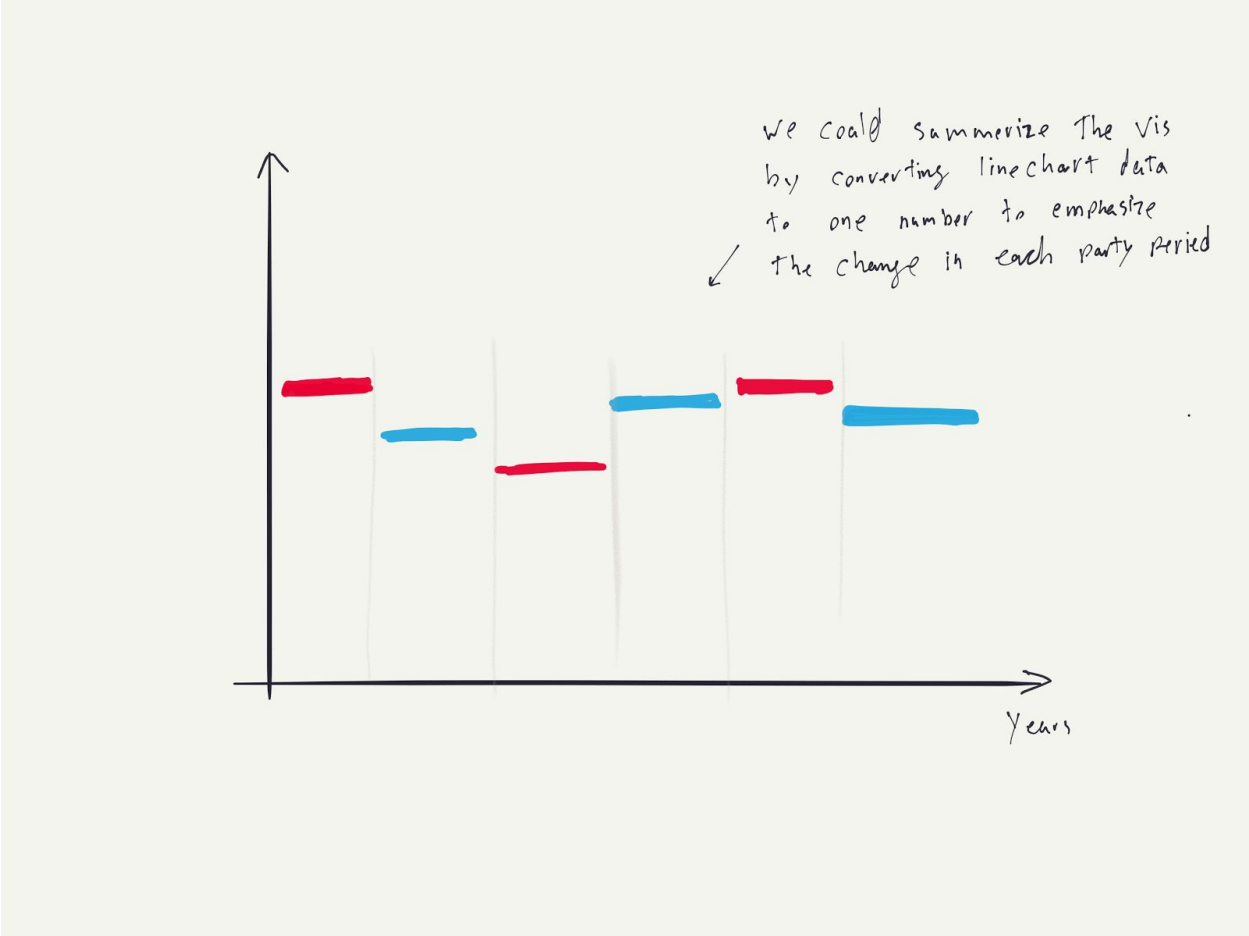
-



*The relationship between the two offices of power (congress vs prez in any given year) -- we need to show that somehow.

Ghaith vis


Quantity
important point
peak
filters
color coding for the ruling party
show only Rep
show only Dem
add arrows to show party effect
we could add some Historic events which could have effected the studied quality
war
financial crisis
years



the first one is build on VG sketches but added some more details we could use..
The second one is an alternative solution to show the differences in a summarized way.

Description of Data

- Upload on git/shared repository ( ) - https://github.com/vielgi/d3/research data
- **Data Yearly Range**: Ideally data should be from April 1945 (**Truman** took over) onward, maybe till Feb 2016 (unemployment rate source has it)
- Presidents' and Congress' **Start and End in office**: We should be consider either Jan or Feb till either Dec or Jan, correspondingly.

Leaning towards Feb -- i.e. they had one month in office and any change may be attributed to them (btw, among myriad other reasons, ha...). But what to do when we have yearly data?

- **Just a thought:** maybe we focus on full terms only so i.e. from Truman's second term to Obama's first, 1949 to 2012
- **Consistent** number format -- We need to calculate to the same decimal point, e.g. talking about percentages should be 0.41% or 0.419%
- CSV format

1. **Parties in Power --**
    - Congress -- 2 year terms
    - President -- 4 year terms
    - Years
2. **Unemployment Rate** - excel spreadsheet - 1948 to 2016 (uploaded to git) Bureau of Labor Statics http://data.bls.gov/pdq/SurveyOutputServlet
    - **Monthly data**
    - **Unemployment Rate** -- **unemployed** as % from **total labor force**
    - **Year-on-Year Change** -- we need to come up with a way to calculate this with the most precision.
    - Min: 2.5 Max: 10.8 Avg:5.8% -- for the unemployment
    - Ideally our data would be like this:

Month-Year	Prez	Congress	Unemployed #	Total Labor Market	Unemployment Rate	Change %

				#		
Jan-2000			5	100	5.00%	
Feb-2000			6	101	5.94%	15.83%

3. **Consumer Price Index** - excel spreadsheet (uploaded to git)
   https://research.stlouisfed.org/fred2/series/CPIAUCSL#
   - Monthly data
   - Ideally our data would be

Month-Year	CPI	% Change
Jan-2000	169.300	
Feb-2000	170.000	0.41%

4. **GDP**, **Budget**, etc - 1930 to 2015 - excel spreadsheet (uploaded to git)
   https://www.whitehouse.gov/sites/default/files/omb/budget/fy2017/assets/hist01z2.xls
   - Yearly data of GDP (in billions of dollars)
   - We can calculate **change of gdp %**
   - Ideally our data would be

Month-Year	CPI	% Change
Jan-2000	169.300	
Feb-2000	170.000	0.41%

5. **World's GDP** in percentage - 1960 to 2015 - excel spreadsheet (uploaded to git) http://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG
6. Median Income -- **Race and Hispanic Origin of People (Both Sexes Combined) by Median and Mean Income** [XLS - 58k] -- this is what we need, and we can calculate growth/decline as percentages,
    - starts from 1947 to 2014
    - Yearly!!
    - In 2014 dollars
    - Ideally our data would be

Month-Year	Median Income	% Change
2000	$29,578	
2001	$29,330	0.41%

7. ~~Real Estate Prices -- can't find from the 1940s~~
8. **S&P 500** -- from okfn.org ; http://data.okfn.org/data/core/s-and-p-500 → https://raw.githubusercontent.com/datasets/s-and-p-500/master/data/data.csv
    - http://data.okfn.org/data/core/s-and-p-500/r/data.json

VG: to do -- detail the data dimensions; the link, put on github

Process Book

- Let's' investigate what this is exactly

Your process book should include:
1. Cover Page that includes: your name, class, semester, and professor
2. Table of Contents
3. Project introductory pages
4. Research

5. Brainstorming
6. Thumbnails (pencil/ink)
7. Roughs (pencil/ink or digital)
8. Final design

I can distill from the few examples (fame wall --> previous years) if there's need for more detail. I think that can be used to inform more of the tasks too:

- http://itisaasta.com/nycs/process.html (btw, I like how this person embedded the pdf in the page
- http://teamdatahub.github.io/process.html -- this one maybe even better because it is an html page (with navigation)

3 Minute Presentation/Screencast

- Sunday evening --

----https://www.youtube.com/watch?v=FVmmS3KRZnk&feature=youtu.be

VG's response to Coping with hitchhikers,etc.:

There's no way to deal effectively with a couch potatoes or hitchhikers in an academic team setting -- there's simply no proper reward/punishment mechanism/strategy. Having said that, I did appreciate the advice from the article and already implemented -- since we didn't hear anything from one of our team members we didn't include him in our team. (We understand that there could be a number of extenuation circumstances, e.g. genuine medical emergency,etc. and we'll be open to let them join our team later on!) . I've been on a number of teams where everyone in the team was getting the same grade so I like that feedback from all team members about the rest will be taken -- that seems to be another good strategy to ensure that everyone will be on their behavior.

There's no way that any of these stereotypical characters will make me feel guilty (as the article suggest) about reminding them that everyone needs to chip in . The way to ensure is to communicate well, including about one's expectations from the rest. However, ultimately, since people have different expectations about their educational experiences, it is hard to make everyone contribute and do quality work. Luckily we're counting that since everyone is here voluntarily that they'd like to learn and contribute in a fair manner as no one (apart from the stereotypes that the article discusses) would like to be "that" team member.

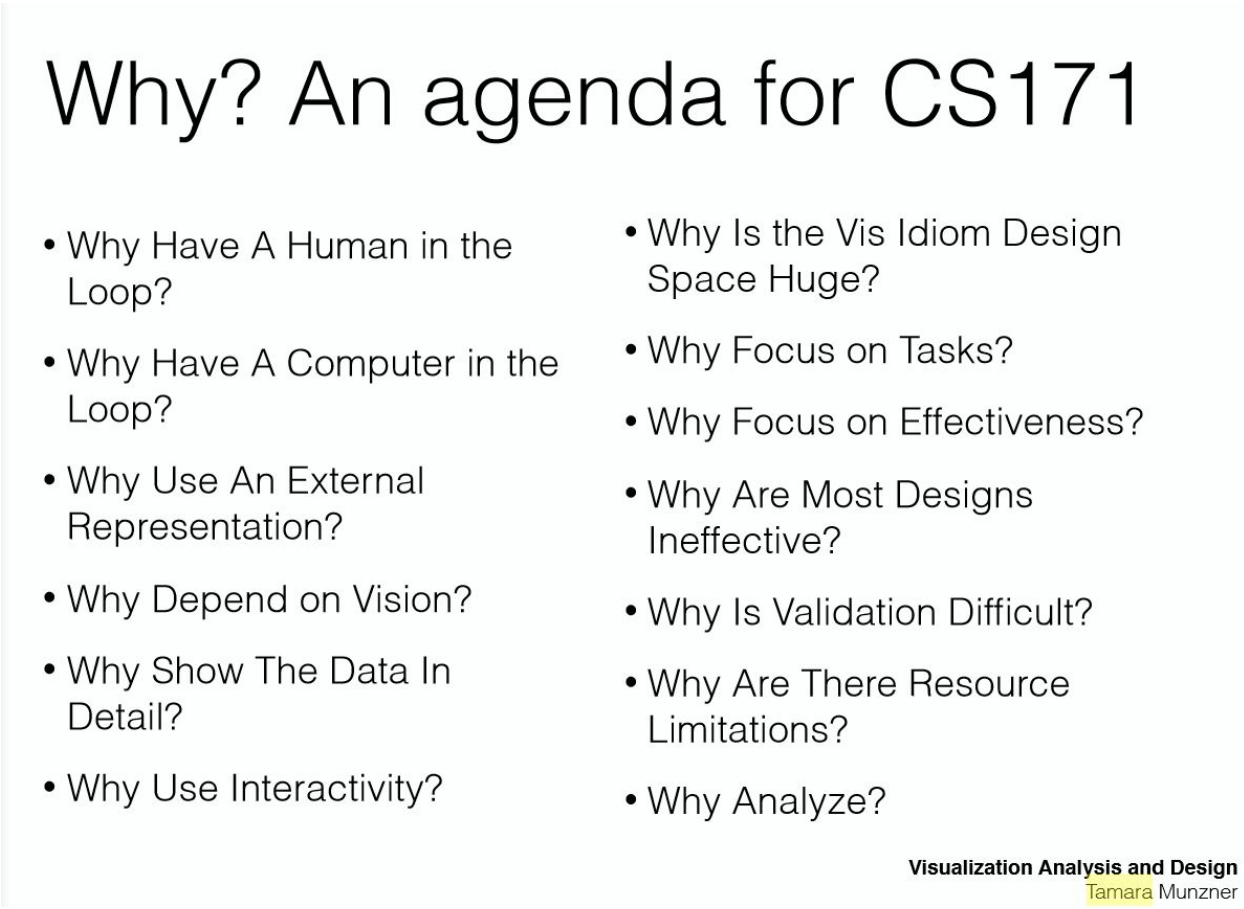
Week 10: Feedback from Alain:

Hello Chris+Ghith+Vladimir,

I have revised your project proposal. This is a very interesting topic, especially now that we are in the time of elections.

1. Team brainstorm then group decision: For your next iteration, I would like you to **state your domain and high level tasks** that your visualizations will be supporting (see Tamara Munzner's reading).

VG: I am a bit at a loss here...Any ideas?



For your next submission of your project plan, I would like you to do the following:

1. Team brainstorm then group decision: Break out in an own section, **the questions that you are trying to answer** with the visualizations

   VG:

   1.1. Under which Party has the [Economy/Median Income/CPI/etc.] fared the best?

   1.2. How do different Terms stack against each other?

   1.3. Can either of the Parties decisively claim they're better for the [Economy/Median Income/CPI/etc.]

   1.4. What president/party has had the higher unemployment rates?

   1.5. (using the visualization map I suggested) - How has the US GDP been affected by other economies?

      1.5.1. What presidents have been in office?

      1.5.2. For how long have presidents been in office?

      1.5.3. What party does the president in given year belong to?

   Additional in greater detail

   * can we talk of relationship without invoking causality? How?

   - **Median Income** -- because GDP can grow and Median Income can still go down, it is important even more important to look at that measurement. Bottom line this is what people end up taking home so they can take care of their family. GDP can be misleading because it is a mean measurement and not a median one, i.e. in an uneven income distribution where most of the gains go to the top it will pull the average but not necessarily the median up.
   - **Consumer Price Index** -- the GDP grows but what happens to the cost of goods and services that people buy.
   - **Economy** -- as indicated by GDP, more specifically Change from Year Before in %. Positive it would mean growth, negative -- decline.
   - **Budget** -- Both parties want to balance the budget but when has it been done.
     - Who spends more (as a % of GDP)?
   - **S&P 500** -- have the markets fared better or worse during each term?

2. INdividual assign: Write down the **list of features** that you will use to accomplish your high level tasks (e.g. brushing+linking, zooming, filtering, etc.)
   a. VG: select, filter, sort

List of features:

Select - The select features will allow users choose the indicators they want to see in the bar and line charts.

Hovering - Additionally, all line and bar charts we will have tooltips to provide additional information.

Filtering - The map visualization with global GDPs will have a filtering bar to allow users select different years

Brush + Link - The stack chart with budget percentages (see below) will have a brush+link feature to allow the user select individual line areas or time periods in the graph

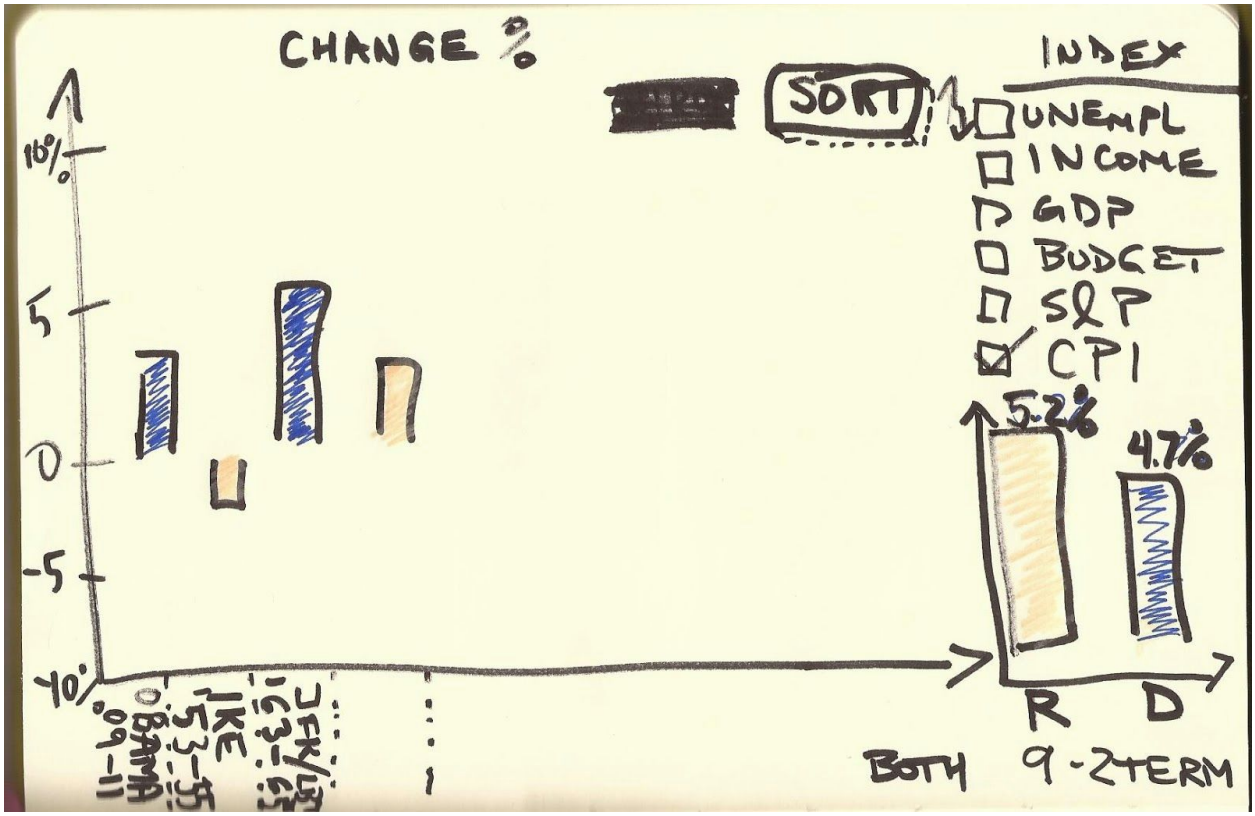
3. **Solidify what visualizations** you will need to support your goals
    a. VG: line graph, bar graph, choropleth - DONE, SEE BELOW FOR VISUALIZATIONS.
4. **Team: Concoct a solid structure for your storytelling**. Think about logical flow and visual flow of information.Keep in mind that you will need to implement some coordinated views for your final product. Your project is a good candidate for this as you are measuring presidents against several metrics. You might want to entertain the idea of including a scatterplot matrix for the different indices.

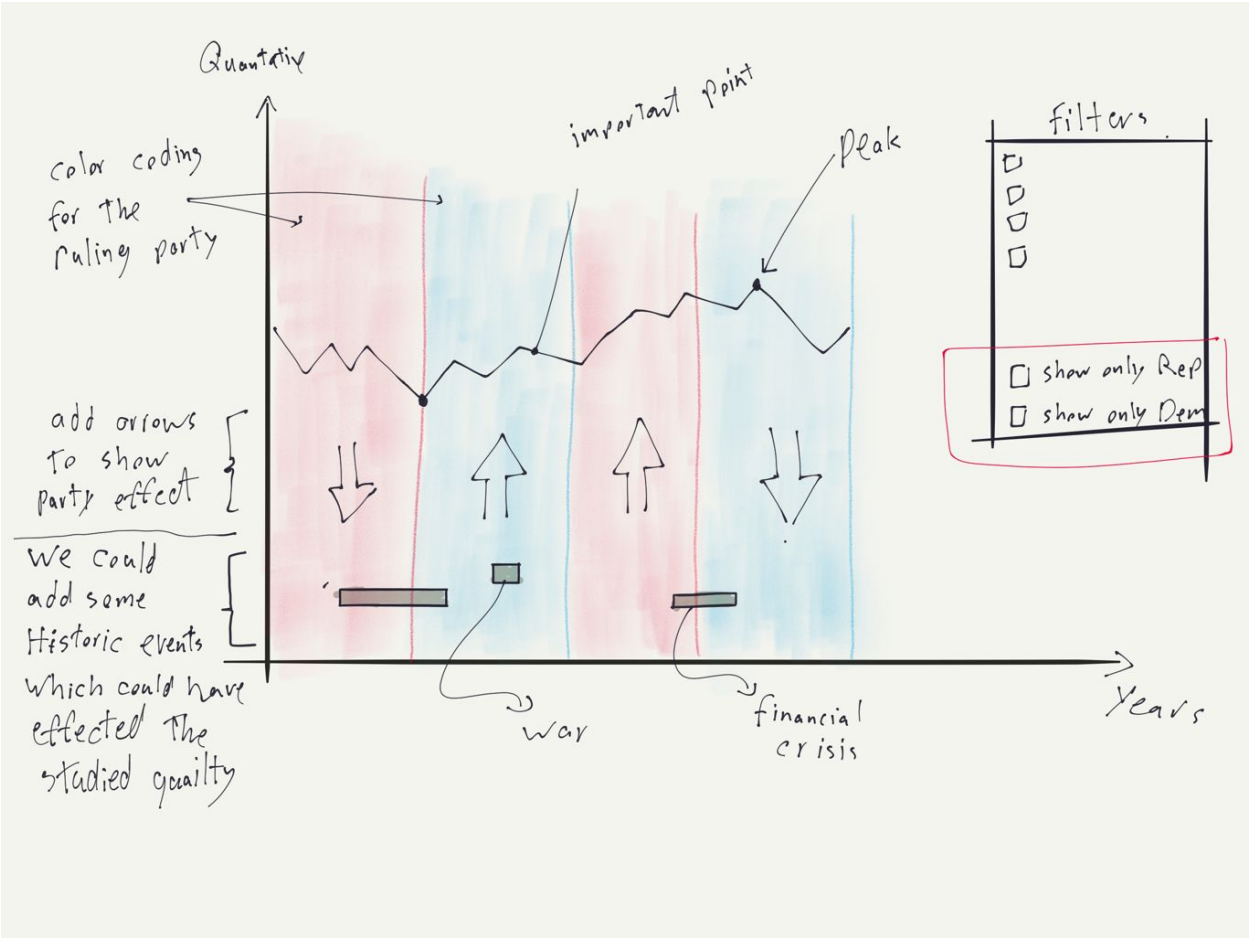
DONE - SEE BELOW

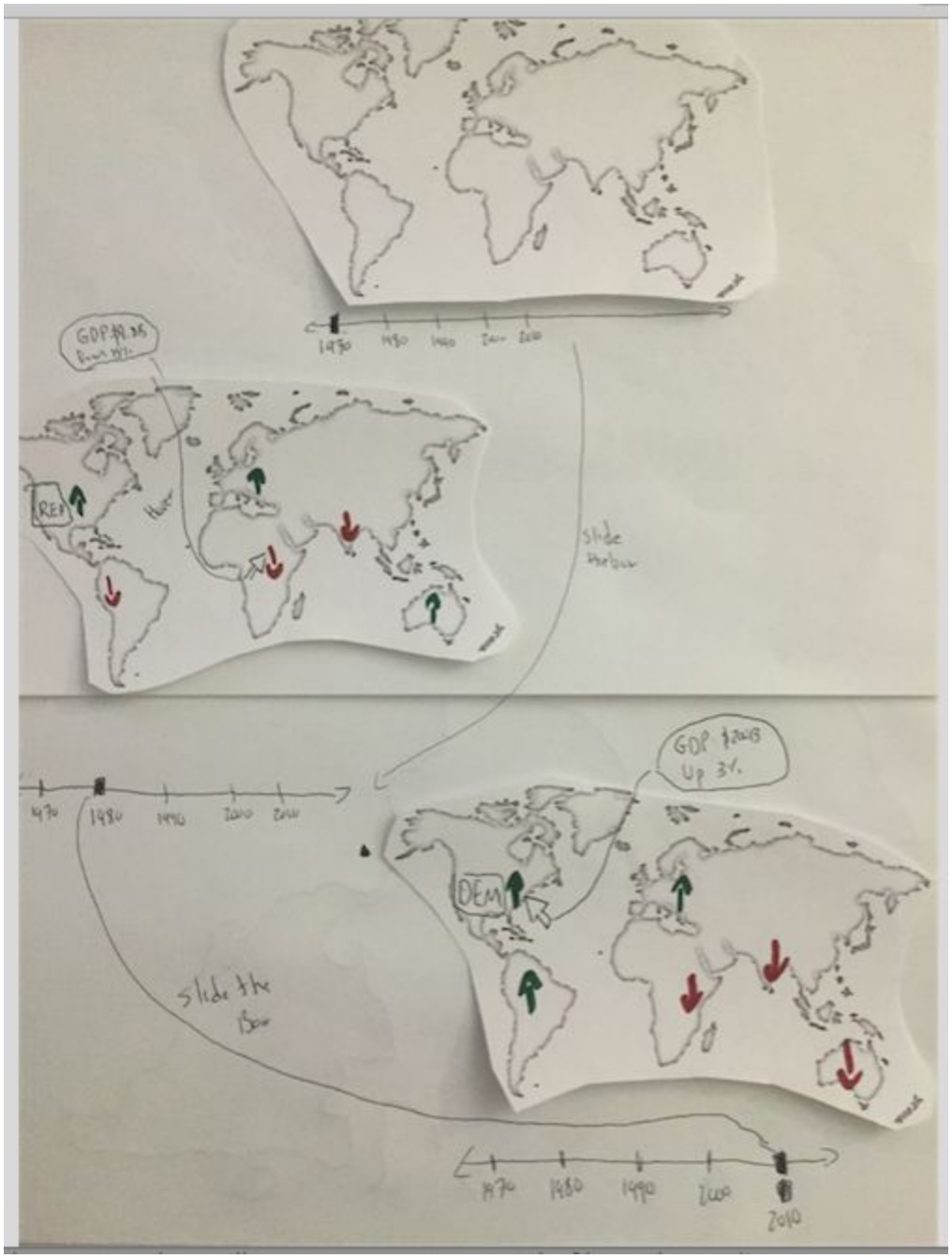
5. Contact Alain with questions

Project Plan Instructions:

1. 2nd iteration on sketches, including new visualizations you might have added after the poster session -- **individual 1 each:**
    a. vg


CHANGE %
SORT
INDEX
UNEMPL
INCOME
GDP
BUDGET
S&P
CPI
10%
5
0
-5
-10%
5.2%
4.7%
R
D
BOTH
9-2TERM


Quantity
important point
peak
filters
color coding for the ruling party
show only Rep
show only Dem
add arrows to show party effect
we could add some Historic events which could have effected the studied quality
war
financial crisis
years



2. **description of team roles**
    a. **Ghaith -- visual guru; d3 everyone**

Visual Specialists - In charge of verifying we are using visualization principles correctly.

D3 programmer - Programming using d3 library

Quality control proposals - Check final written deliveries

b. VG:organizer; d3

Organizer - setting up times for team meetings. Delegating project activities.

Liaison between team and TF - Final submissions

D3 programmer - Programming using d3 library

Resource investigator - Make sure the data sources are in line with our project goal.

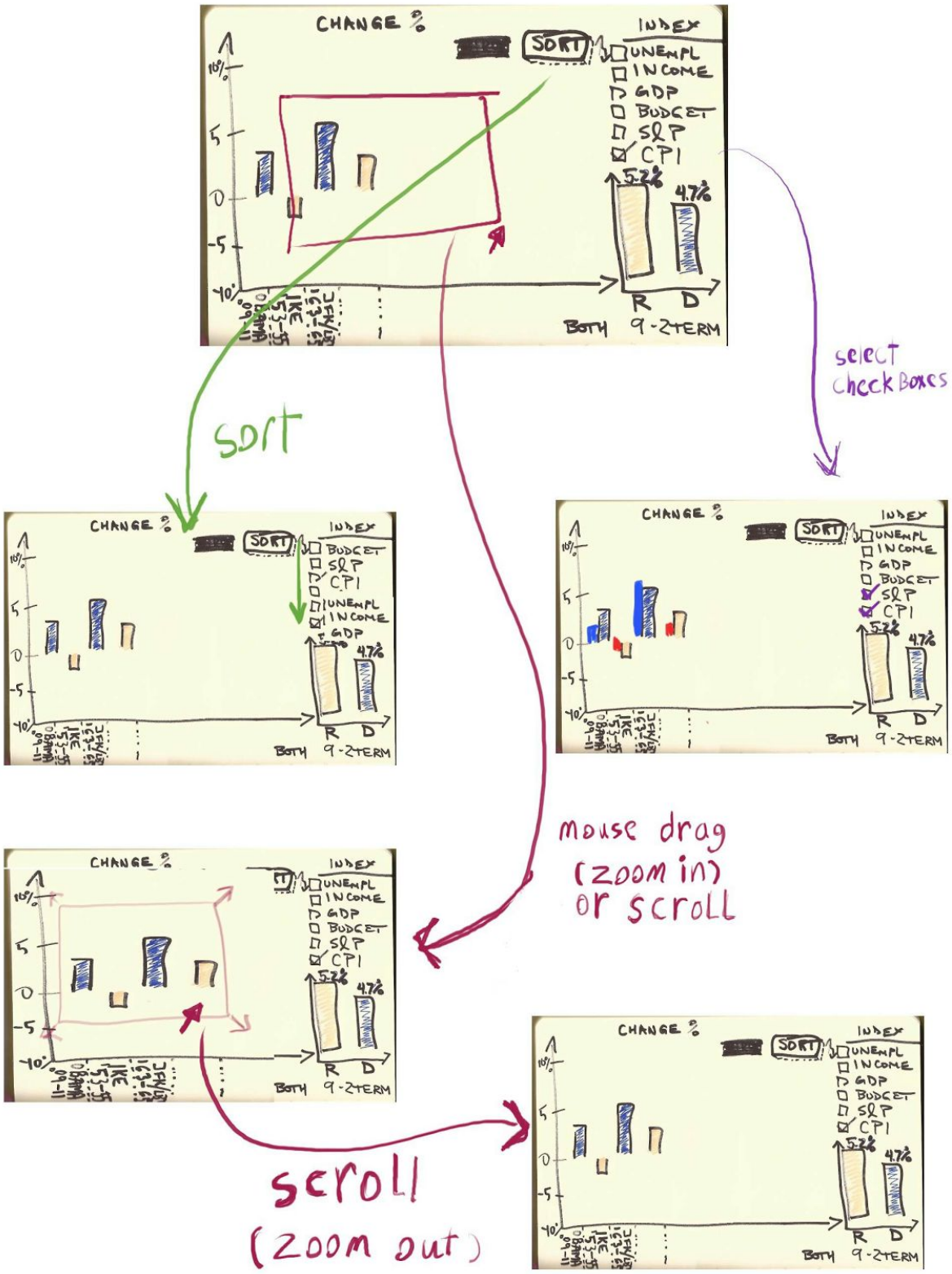
c. Chris -- js guru; d3

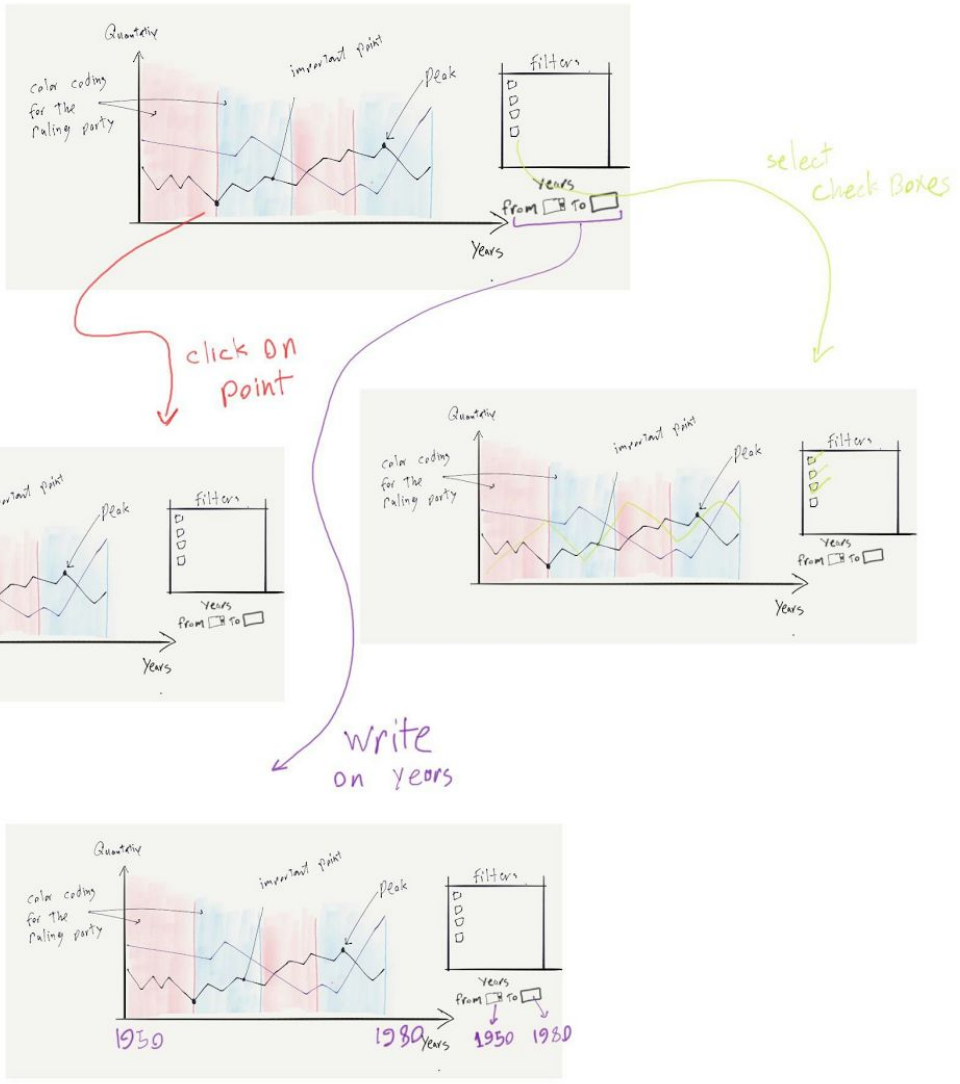
JavaScript coder - Coding the website with JS

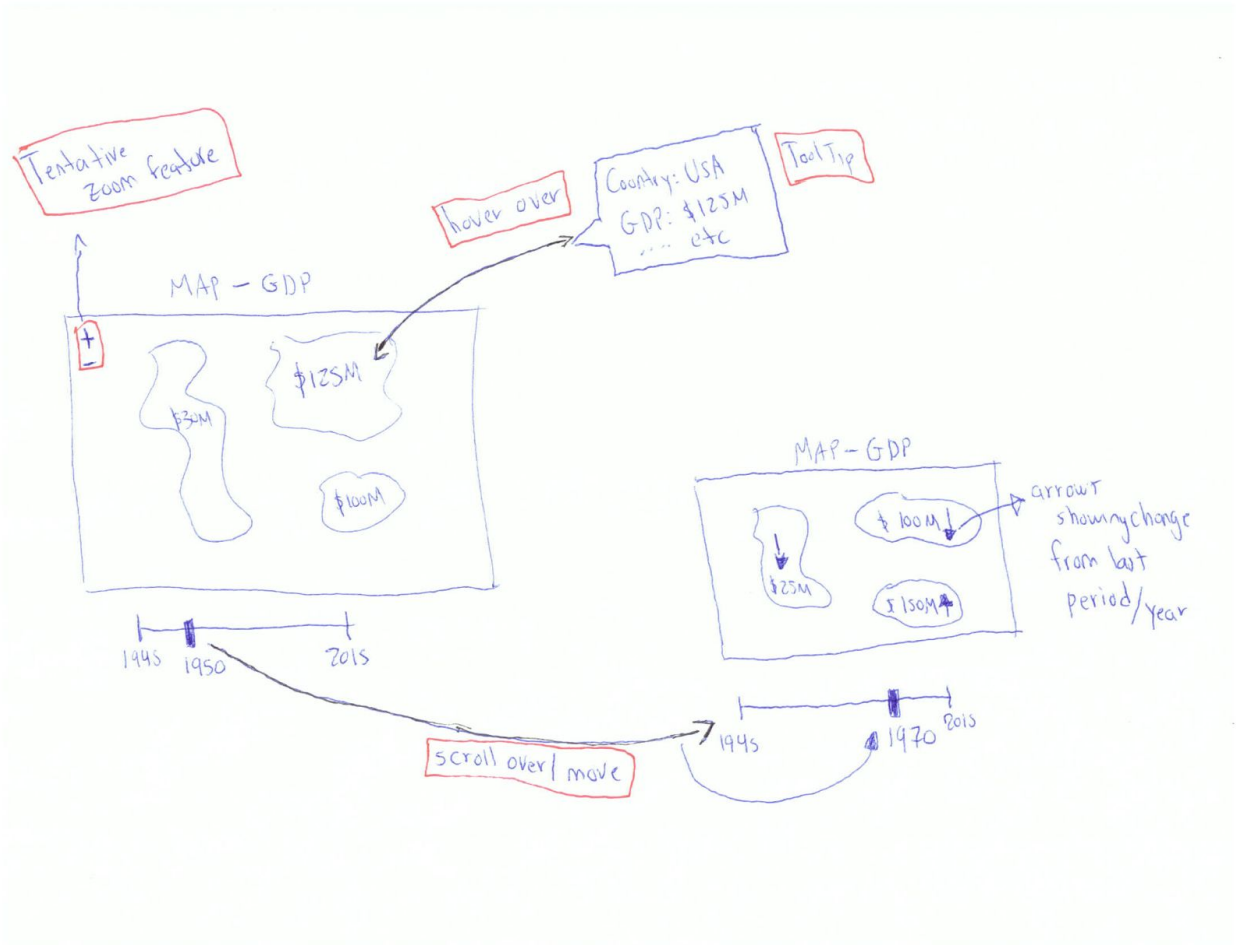
D3 programmer - Programming using d3 library

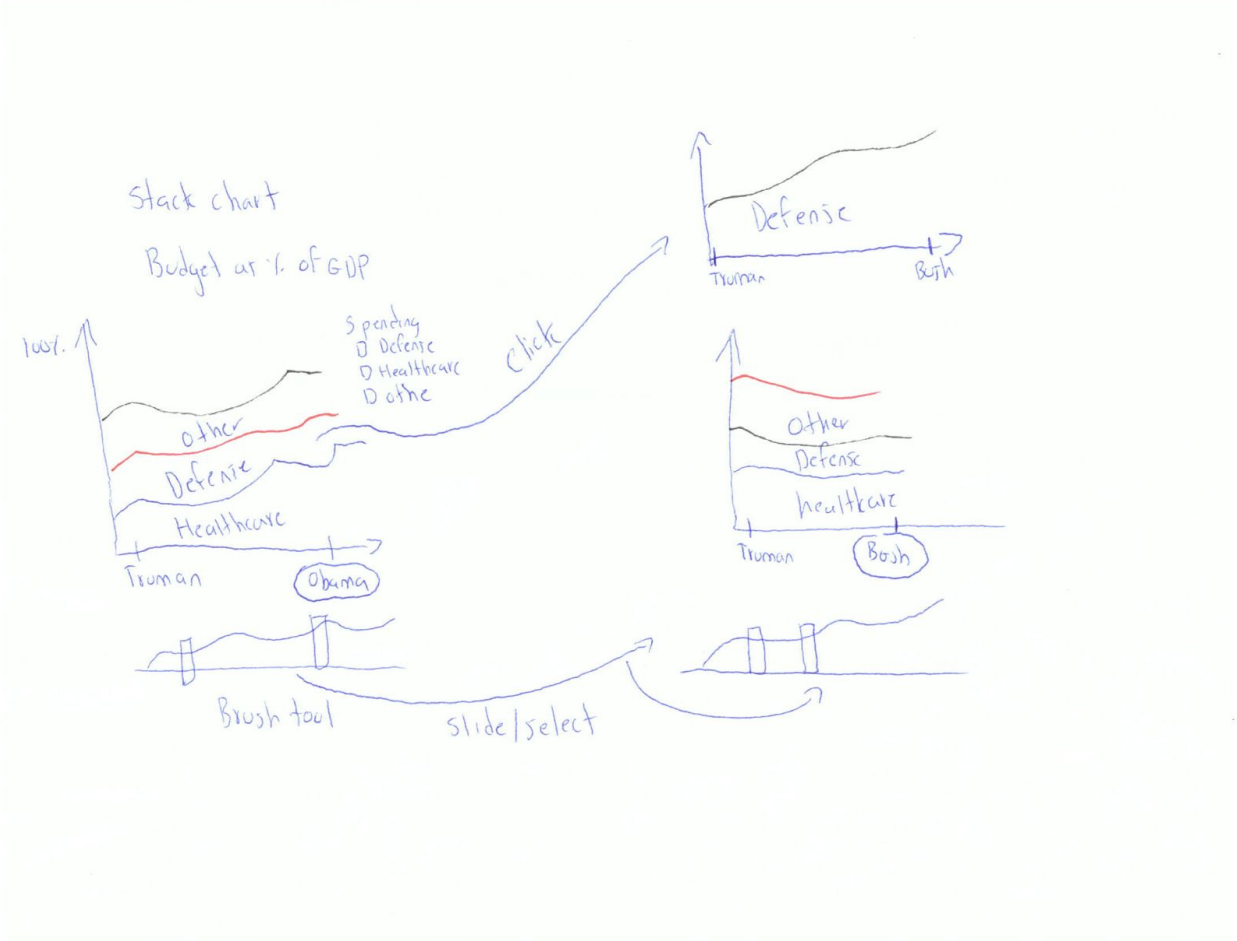
Code implementer - Making sure the code from team members follows best practices learned in class. Implementation of the website.

3. story board for interaction -- Ghaith, then all review

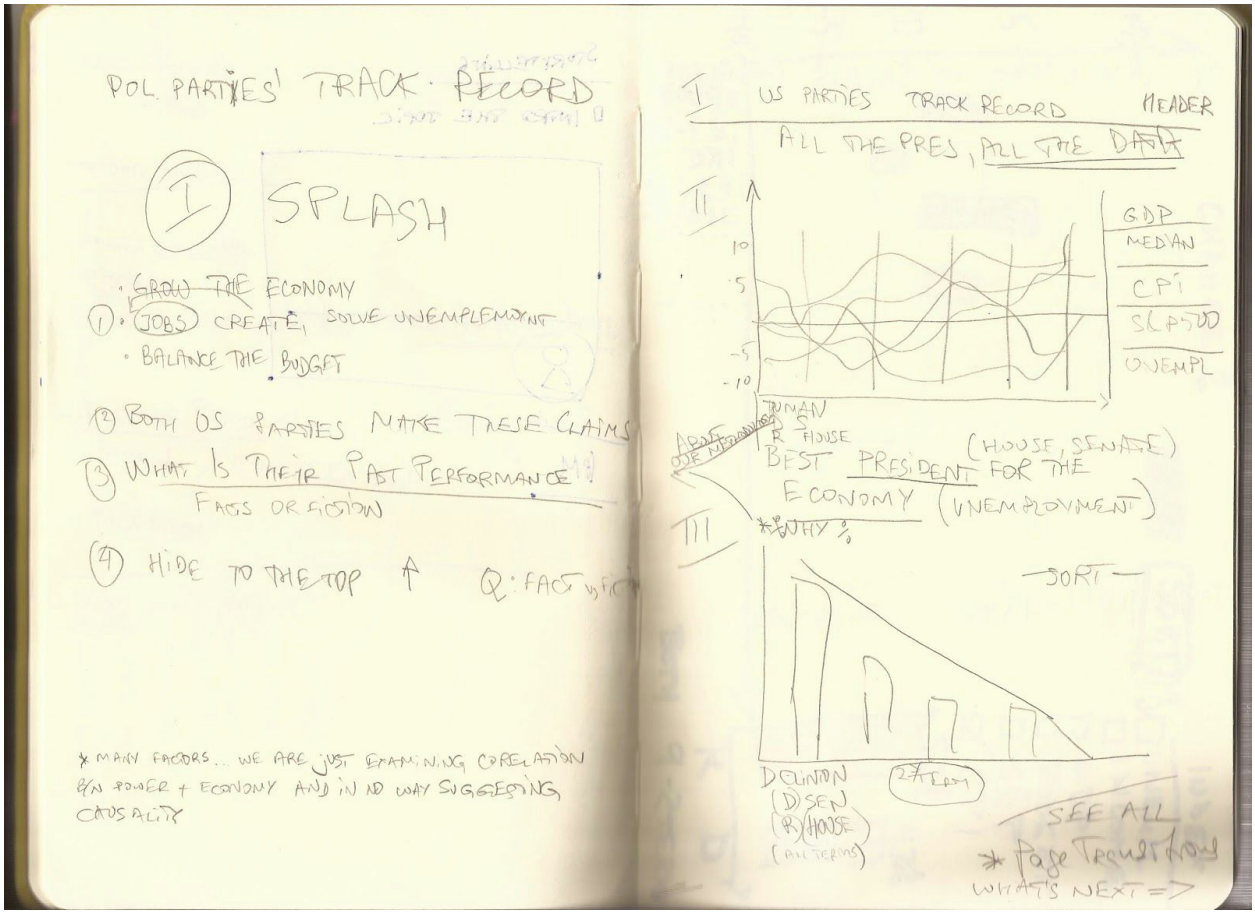

CHANGE %
SORT
INDEX
UNEMPL
INCOME
GDP
BUDGET
S&P
CPI
5.2%
4.7%
R
D
BOTH
9-2TERM
select
check Boxes
sort
mouse drag
(zoom in)
or scroll
scroll
(zoom out)


Quantity
color coding for the ruling party
important point
peak
filters
Years
from To
Years
select check boxes
click on point
show tooltip
write on years
1950
1980
1950 1980


Tentative zoom feature
hover over
Country: USA
GDP: $125M
... etc
Tool Tip
MAP - GDP
+
-
$125M
$32M
$100M
1945
1950
2015
MAP - GDP
$100M
$25M
$150M
arrow showing change from last period/year
1945
1970
2015
scroll over/move



4. webpage layout/storytelling --
    a. vg, set up the basic bootstrap
    b. All team members brainstorm
    c. vg:

1. 

5. **project timeline (with milestones when you are planning to finish which feature)**
    a. **Chris**

	03/22 – 03/28	03/29 – 04/04	04/05 – 04/11	04/12 – 04/18	04/19 – 04/25	04/26 – 05/02	05/02 – 05/04	05/05 -05/08
Project Proposal								
Detailed project plan								
Final Project Plan								
Project re-design								

Project Prototype V1								
Project Prototype V2								
Project Review in Studio Session								
TF/peer Feedback								
Final project screencast								
Final Project submission								
Group peer evaluation, and project demos								

Additionally, you should also include the following:

6. **Acquire your complete dataset**, and add a description/screenshot/etc. to the project plan. You should have the dataset by the time of submission of this homework!
    a. Vg: complete what was started last week
7. Please submit Photos of your Vis Exploration Poster Session Questionnaire (one for each team member)
    a. individual

CS 171

Vis Exploration Poster Session

Please submit this questionnaire - one sheet for each team member - and a photo of your poster (one per team) with your homework

The timetable gives you an overview which expert group has to present and which expert group will go around and explore visualization techniques. Please stand next to your poster when your expert group is presenting.

		Presenting	Exploring
2:45 - 3:00	A – Geographical Data / Maps	A	B C D
3:00 - 3:15	B – Trees & Networks	B	A C D
3:15 - 3:30	C – Visualization for Text	C	A B D
3:30 - 3:45	D – Vis for High Dimensional Data	D	A B C
3:45 - 4:00	Vote for your best poster		

Best Poster Award: Please put your sticky dots on the posters that you like most. You can put multiple dots on the same poster if you wish to do so.
One Minute Paper: Please do not forget to submit the one minute papers! You have until tomorrow.

Name: VLADIMIR GEORGIEV

Group name: GROUP 4

List **three** visualization techniques you found most commonly on the posters about Geographical Data

CHOROPLETH
HEAT MAP
FLOW MAP

CS 171

List **three** visualization techniques you found most commonly on the posters about Trees and Networks

NODE LINK
TREE MAP
ADJECENY MATRIX

List **three** visualization techniques you found most commonly on the posters about Text Visualization

TAG CLOUDS / WORD CLOUD / FREQUENCY TEXT
WORD TREE

List **three** visualization techniques you found most commonly on the posters about High Dimensional Data

high dimensional clustering
Multi-Coordinate Plot
SCATTERPLOT

Next lecture: Evaluation & Innovation with D3

CS 171

Vis Exploration Poster Session

Please submit this questionnaire - one sheet for each team member - and a photo of your poster (one per team) with your homework

The timetable gives you an overview which expert group has to present and which expert group will go around and explore visualization techniques. Please stand next to your poster when your expert group is presenting.

		Presenting	Exploring
2:45 - 3:00	A — Geographical Data / Maps	A	B C D
3:00 - 3:15	B — Trees & Networks	B	A C D
3:15 - 3:30	C — Visualization for Text	C	A B D
3:30 - 3:45	D — Vis for High Dimensional Data	D	A B C
3:45 - 4:00	Vote for your best poster		

Best Poster Award: Please put your sticky dots on the posters that you like most. You can put multiple dots on the same poster if you wish to do so.
One Minute Paper: Please do not forget to submit the one minute papers! You have until tomorrow.

Name: M Ghaith Takriti

Group name: Group 4

List **three** visualization techniques you found most commonly on the posters about Geographical Data

Heat map
Flow map
Choropleth

CS 171

List **three** visualization techniques you found most commonly on the posters about Trees and Networks

Treemaps
Node link - node gardens -
Sunburst plot

List **three** visualization techniques you found most commonly on the posters about Text Visualization

Plain text
Word cloud
Word tree

List **three** visualization techniques you found most commonly on the posters about High Dimensional Data

Scatterplot
Small multiples
Star maps

Next lecture: Evaluation & Innovation with D3

Complete Data Sets:

Description of Data

Major data dimensions from 1949 to 2012:

1. Political Parties in Power
2. Unemployment
3. GDP
4. CPI
5. S&P500
6. Median Income
7. Unemployment
8. Budget Deficit
9. Select Countries GDP

    - Upload on git/shared repository ( ) - https://github.com/vielgi/d3/research data
    - **Data Yearly Range**: Ideally data should be from April 1945 (**Truman** took over) onward, maybe till Feb 2016 (unemployment rate source has it); we would have a parity then
        - **Just a thought:** maybe we focus on full terms only so i.e. from Truman's second term to Obama's first, 1949 to 2012
        - Final thought: if we extend to include two years before (Truman 47-49) and two extra years for Obama (13-15), then we will have equal number of terms (2 years) of Dem and Rep Presidencies
    - Presidents' and Congress' **Start and End in office**: We should be consider either Jan or Feb till either Dec or Jan, correspondingly. Leaning towards Feb -- i.e. they had one month in office and any change may be attributed to them (btw, among myriad other reasons, ha...). But what to do when we have yearly data?
    - We focus on the President's party affiliations instead of the individuals -- see my note about JFK and Nixon

- **Consistent** number format -- We need to calculate to the same decimal point, e.g. talking about percentages should be 0.4%, 0.41% or 0.419%
- CSV format

1. **Parties in Power --**
    - Congress -- 2 year terms
    - President -- 4 year terms
    - Years -- 1949 to 2012 (including)
    - We need to be careful about a couple of blips in our data -- JFK's assassination and Nixon's resignation means that they didn't serve full terms; however the president's that took over were still of the same party so maybe we just focus on that vs the individuals.
2. **Unemployment Rate** - excel spreadsheet - 1948 to 2016 (uploaded to git) Bureau of Labor Statics http://data.bls.gov/pdq/SurveyOutputServlet
    - **Monthly data**
    - **Unemployment Rate** -- **unemployed** as % from **total labor force**
    - **Year-on-Year Change** -- we need to come up with a way to calculate this with the most precision.
    - Min: 2.5 Max: 10.8 Avg:5.8% -- for the unemployment
    - Ideally our data would be like this:

Month-Year	Prez	Congress	Unemployed #	Total Labor Market	Unemployment Rate	Change %

				#		
Jan-2000			5	100	5.00%	
Feb-2000			6	101	5.94%	15.83%

3. **Consumer Price Index** - excel spreadsheet (uploaded to git) https://research.stlouisfed.org/fred2/series/CPIAUCSL#
    - Monthly data -- must convert to annual
    - Ideally our data would be

Month-Year	CPI	% Change
Jan-2000	169.300	
Feb-2000	170.000	0.41%

4. **GDP**,etc - 1930 to 2015 - https://www.whitehouse.gov/sites/default/files/omb/budget/fy2017/assets/hist01z2.xls
    - Yearly data of GDP (in billions of dollars)
    - We can calculate **change of gdp %**
    - **Yearly**
    - Ideally our data would be

Year	GDP Nominal	**GDP Change in %**
2000		6.71%
2001		4.10%

5. **Budget**

- From https://research.stlouisfed.org/fred2
- Federal Reserve Bank of St. Louis and US. Office of Management and Budget, Federal Surplus or Deficit [-] as Percent of Gross Domestic Product [FYFSGDA188S], retrieved from FRED, Federal Reserve Bank of St. Louis https://research.stlouisfed.org/fred2/series/FYFSGDA188S, April 2, 2016.
- Yearly
- Additional Data

- Data

Year	**Budget as Percentage of GDP**	**Budget as Percentage of GDP**
2000		6.71%
2001		4.10%

6. **World's GDP Per Capita** in percentage - 1960 to 2015 - excel spreadsheet (uploaded to git) http://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG
   - There's no good data set so far -- I'm even re-considering it
   - I found the following data that has GDP Per Capita for Italy United Kingdom United States Japan China India Chile South Africa South Korea Australia Sweden World

Average

But only till 2008, from 1944

- Data

Year	**GDP Nominal**	**GDP Growth**
2000		6.71%
2001		4.10%

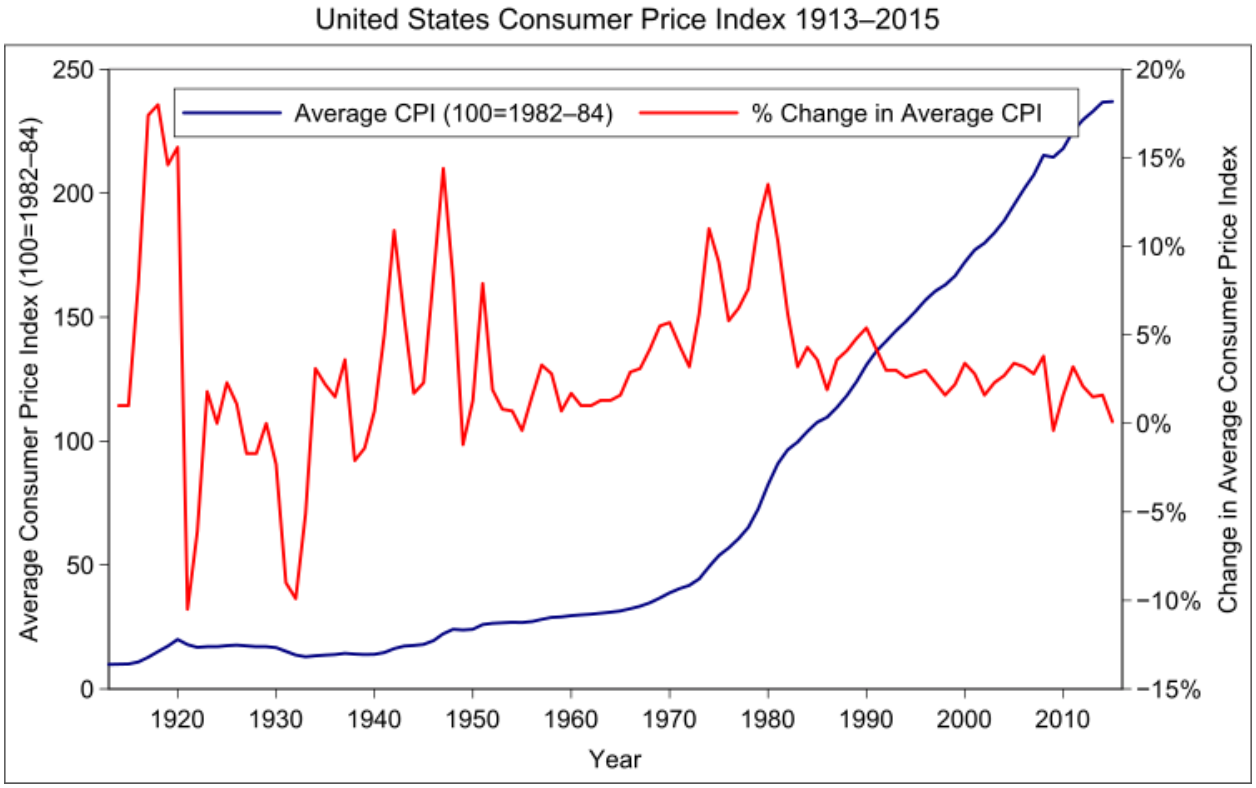
7. Median Income -- **Race and Hispanic Origin of People (Both Sexes Combined) by Median and Mean Income** [XLS - 58k] -- this is what we need, and we can calculate growth/decline as percentages,
   - starts from 1947 to 2014
   - Yearly!!
   - In 2014 dollars
   - Ideally our data would be

Month-Year	# Population with Income	Median Income	% Change
2000		$29,578	
2001		$29,330	0.41%

8. **S&P 500** -- from okfn.org ; http://data.okfn.org/data/core/s-and-p-500 → https://raw.githubusercontent.com/datasets/s-and-p-500/master/data/data.csv
   - http://data.okfn.org/data/core/s-and-p-500/r/data.json
   - Monthly -- must convert to annual
   - Our data is:

Month-Year	Nominal	Change % from Year before
2000		
2001		

Why we're doing % and not nominal data:



I propose we explain our methodology including a few of the gotchas and disclaimers.

Ideas for the demo-site -- presentation (jquery effects)

Wow, so many great presentation techniques:

- http://tympanus.net/Tutorials/FullscreenBookBlock/
- http://tympanus.net/Tutorials/ExpandingOverlayEffect/

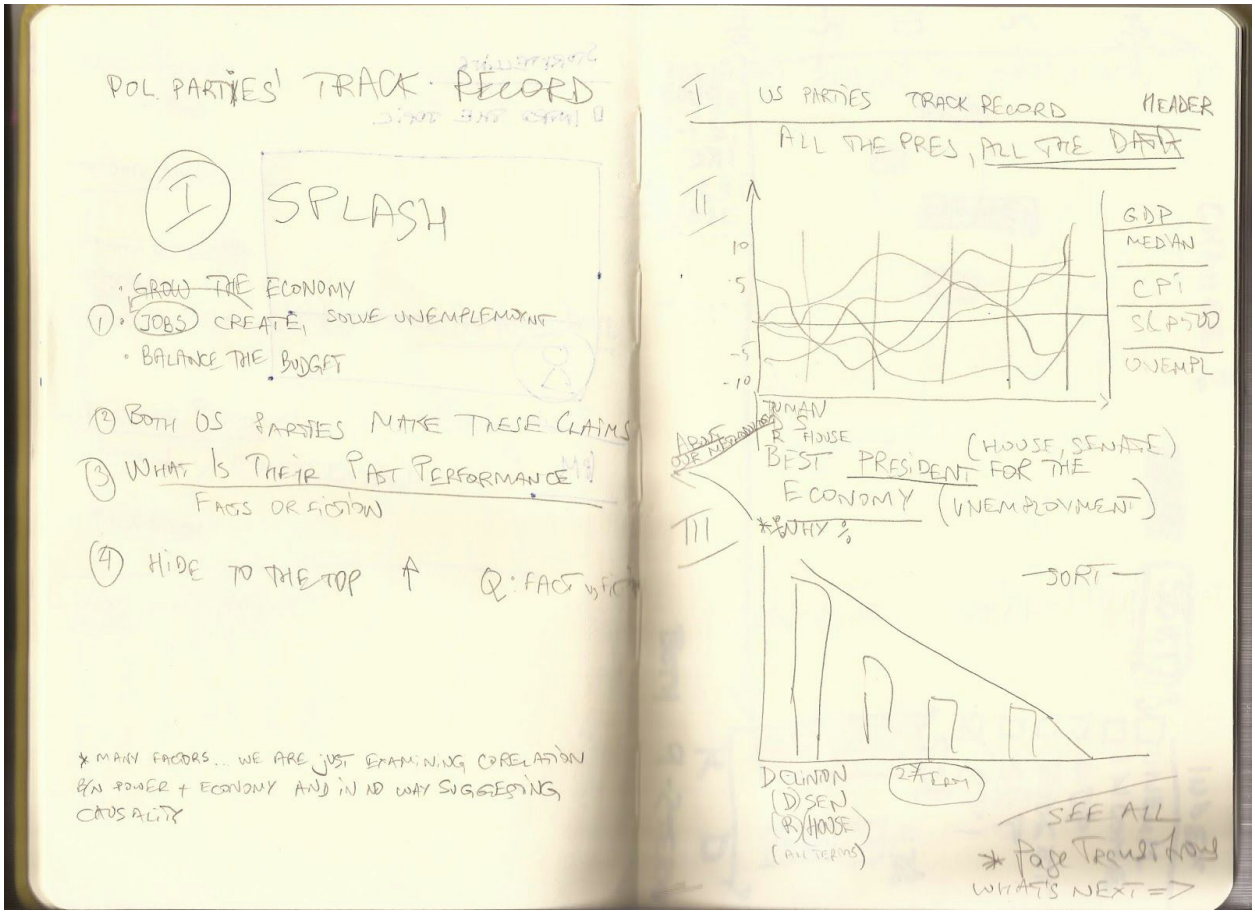
- http://designdrizzle.com/15-amazing-page-transitions-effects-tutorials-in-jquery-and-css3/ -- 20 links
- http://ninodezign.com/20-jquery-plugin-for-page-transition-effects-with-css3/ -- 20
- http://git.blivesta.com/animsition/zoom/
- https://codyhouse.co/demo/animated-page-transition/index.html

Mg: I have checked the links and liked these:

- http://tympanus.net/Development/PageTransitions/
- http://tympanus.net/Tutorials/ExpandingOverlayEffect/
- http://tympanus.net/Development/FullscreenLayoutPageTransitions/
- http://tympanus.net/Tutorials/VerticalShowcaseSlider/

Site Structure

1. Problem -- introduce, splash on page, the questions arising from the indiscriminate statements that the parties make (“grow the economy”,etc.), 5 (+story)
2. The line graph is then slowly rolled out -- implicit instructions (+story)
3. “Which TErm has been the Best for the [econ/unemployment/etc.] (+story)
4. “Final Thoughts” Disclaimer --
    1. world-map Compare GDP with other countries world map
    2. Stacked line
        2.1. Stacked line chart with Brush tool -- tool budget breakdown -- as % of GDP
        2.2. Stacked line chart with Brush tool -- tool budget breakdown -- as % of GDP Change in %
5. *Methodology -- should also be accessible but only on demand; i like the show/hide effect -- let the user be in command of what they’re shown...

2. 

Week of 4/4/2016

Meet Agenda of 4/8/2016

1. Sat: 1-3 pm (1 hr) -- two teams conferences
    a. read the instructions: All
    b. email the two teams : V
2. submission -- innovation/redesign -- discuss more tomorrow Sat
3. working prototype 4/18th
    a. For Sat/Tomorrow: interactive site features (based on the links in the process books); look at those interactive site features -- the jquery links and think of what will work best for our storytelling: All
    b. Will quickly 'draft' the site structure and publish: V: published at http://vielgi.github.io/ (committed to both repositories)

Tasks (As hosts):

Prepare a scenario and two tasks for your evaluators to try to "solve" using your user interfaces--this will give them a chance to put themselves in the role of your intended users and to get to know the UIs.

1. How did various terms compare against the others in terms of the [Growth of the Economy/Unemployment/etc] ?
2. Who was the **best** [President/Senate/House] for the [Growth of the Economy/Unemployment/etc]

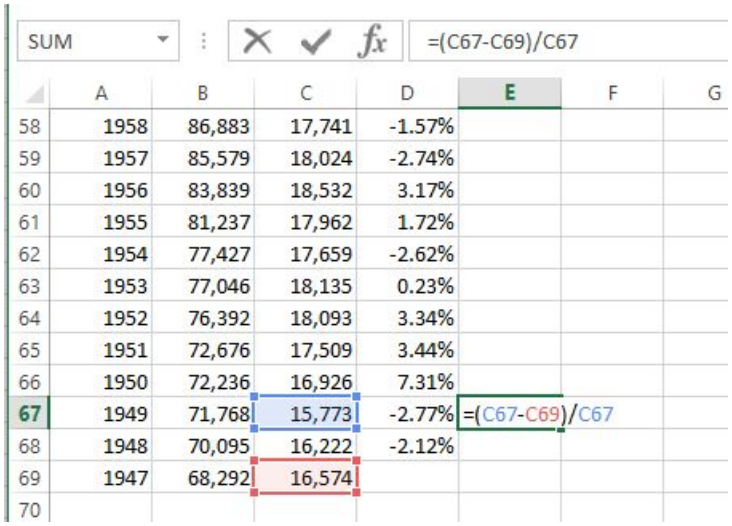
Specific Questions:

1. How unemployment rates compare between the terms of President Gerald Ford and President Jimmy Carter?

==

VG Questions:

1. Methodology for calculating change for each 2-term (e.g. 1947-1948):



	A	B	C	D	E	F	G
58	1958	86,883	17,741	-1.57%			
59	1957	85,579	18,024	-2.74%			
60	1956	83,839	18,532	3.17%			
61	1955	81,237	17,962	1.72%			
62	1954	77,427	17,659	-2.62%			
63	1953	77,046	18,135	0.23%			
64	1952	76,392	18,093	3.34%			
65	1951	72,676	17,509	3.44%			
66	1950	72,236	16,926	7.31%			
67	1949	71,768	15,773	-2.77%	=(C67-C69)/C67		
68	1948	70,095	16,222	-2.12%			
69	1947	68,292	16,574				
70							

You can see on github d3/research data/MedianIncome4142016.csv and doublecheck

2. Data structure is at d3/data/bars.csv in this format:

term,president,president_party,senate,house,gdp_change,income_change(etc)

1947,Truman,D,R,R,0.021948421,-0.050782984

1949,Truman,D,D,D,0.164046537,0.099149009

1951,Truman,D,D,D,0.055791795,0.034518886

1953,Eisenhower,R,R,R,0.086665708,-0.009631444

1955,Eisenhower,R,D,D,0.023046655,0.003439858

1993,Clinton,D,D,D,0.061053184,0.03853877

1995,Clinton,D,R,R,0.082879138,0.061535116

1997,Clinton,D,R,R,0.090316717,0.056841098

1999,Clinton,D,R,R,0.030066193,0.002488919

2001,Bush,R,D,R,0.060863202,-0.005037179

2003,Bush,R,R,R,0.05876828,0.01034319

2005,Bush,R,R,R,0.041249216,0.030031907

2007,Bush,R,D,D,-0.030938186,-0.054235878

2009,Obama,D,D,D,0.042685135,-0.030334429

2011,Obama,D,D,R,0.036240206,0.011269297

2013,Obama,D,R,R,0.043052469,0.032270404

3. Check out the root index page with integrated bar charts for income and gdp change (Chris, didn't have a chance to integrate yours on index as well..)

Submission Requirements

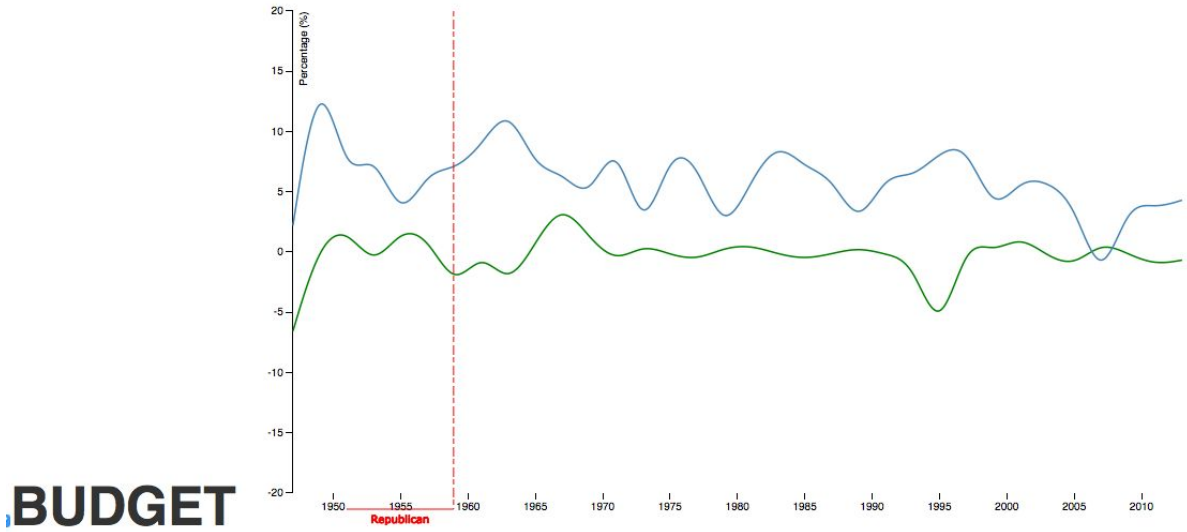
- Using the real datasets
  - Functional D3 visualizations
  - All D3 visualizations except the innovative view have to be working already
  - Rough webpage design and structure has to be done and implemented
  - Interactions (e.g., filtering, brushing, etc.) have to be working
  - Up to date process book

Remaining Tasks:

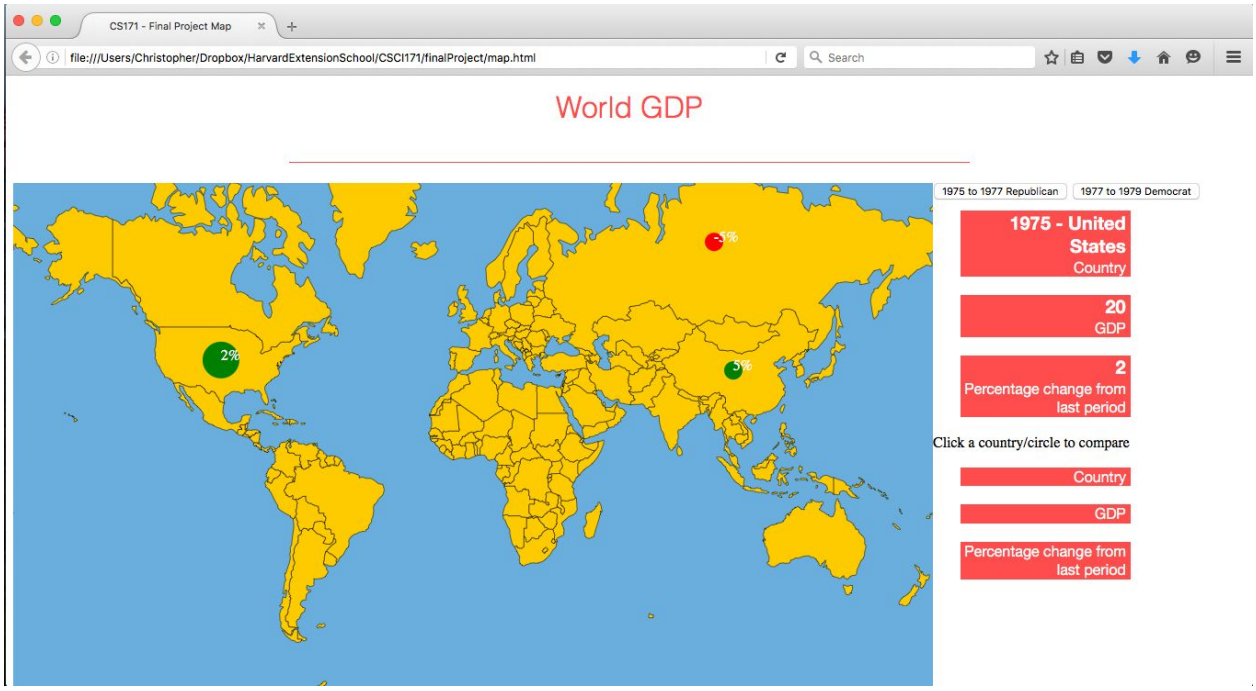
1. Data
   1.1. check of data/bars.csv -- Chris

1.2. Format “resaerch data/countriesGdpGrowth.csv” -- VG; just submitted to github

2. Line viz -- Chris
   2.1. Missing this file ui-bg_flat_75_ffffff_40x100.png
   2.2. Review
   2.3. Tool tip for each line
   2.4. Line Scale for the lines



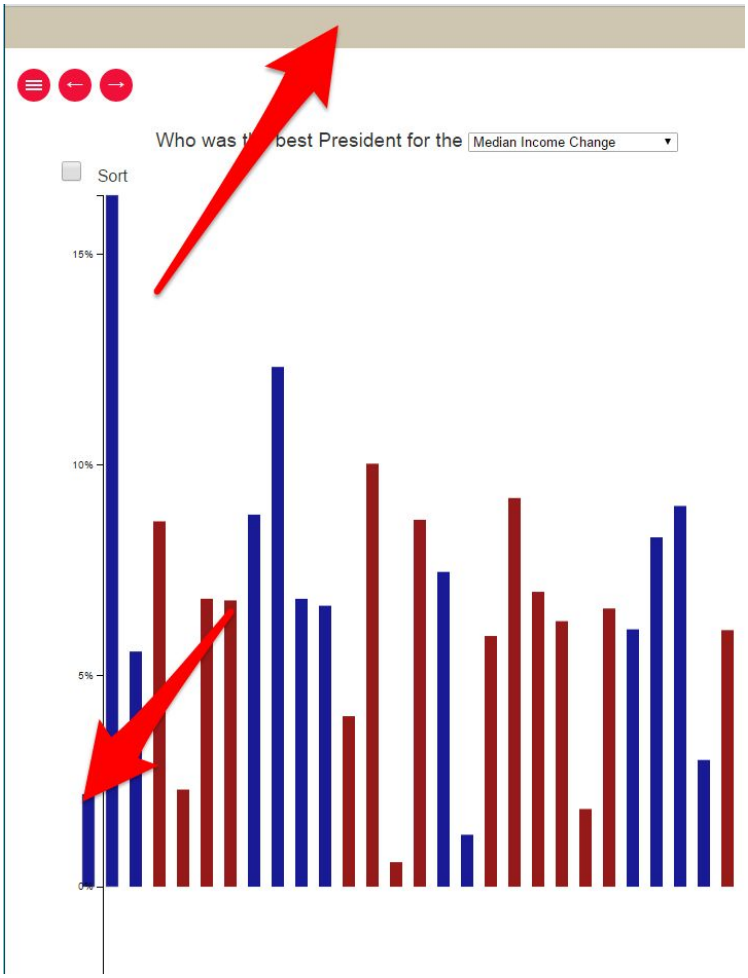
2.3 Map Viz - World GDP



- Real data
- Circle size domain
- Indicators changes

3. Bars viz
    3.1. Implement some labels for each bar with Icons -- -- can somebody implement?
        3.1.1. Heads of all presidents -- png; Ghaith
        3.1.2. Some icon/pngs for US House/US Senate/color_coded for Republicans and Democrats -- Ghaith
    3.2. Extra drop-down for president/house/senate -- vg will try

3.3. Bug when integrated all the external css and js



3.4. review

4. HTML page -- some incompatibilities/conflicts when the js scripts (for the page turn jquery effect) are activated -- bug: can't show tooltips of the bars chart -- can somebody investigate?

4.1. review

5. Html page -- content walktru; VG

5.1. review

6. Submission of Monday's submission and to vielgi.github.io; VG

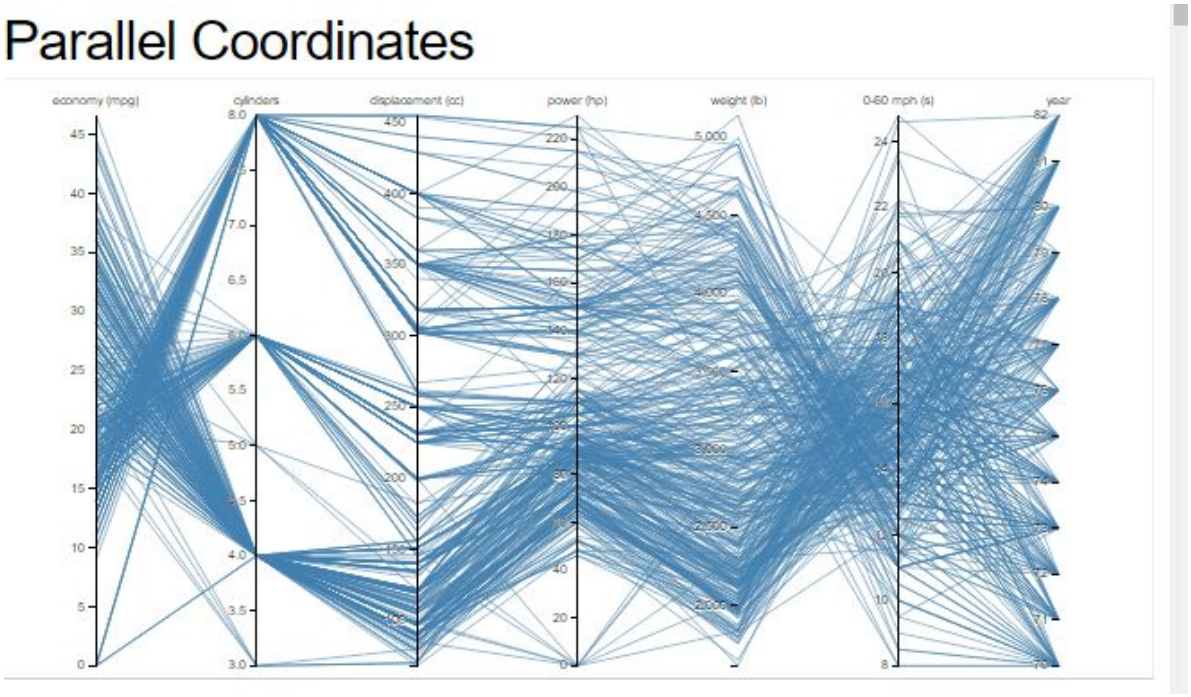
Next Week:

7. responsive?

8. OOP code
9. Coordinated view
10. Complex viz -- maybe scatterplot?

Search for an Innovative Viz

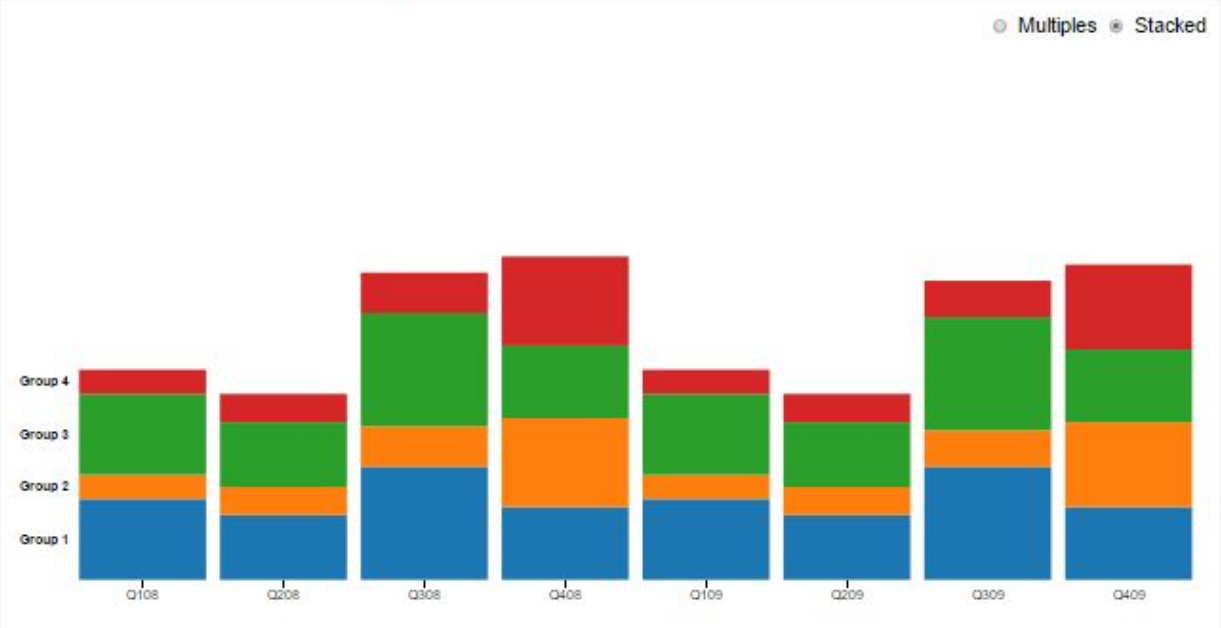
Ghaith: I have collected these visualizations from http://bl.ocks.org/mbostock the code is available



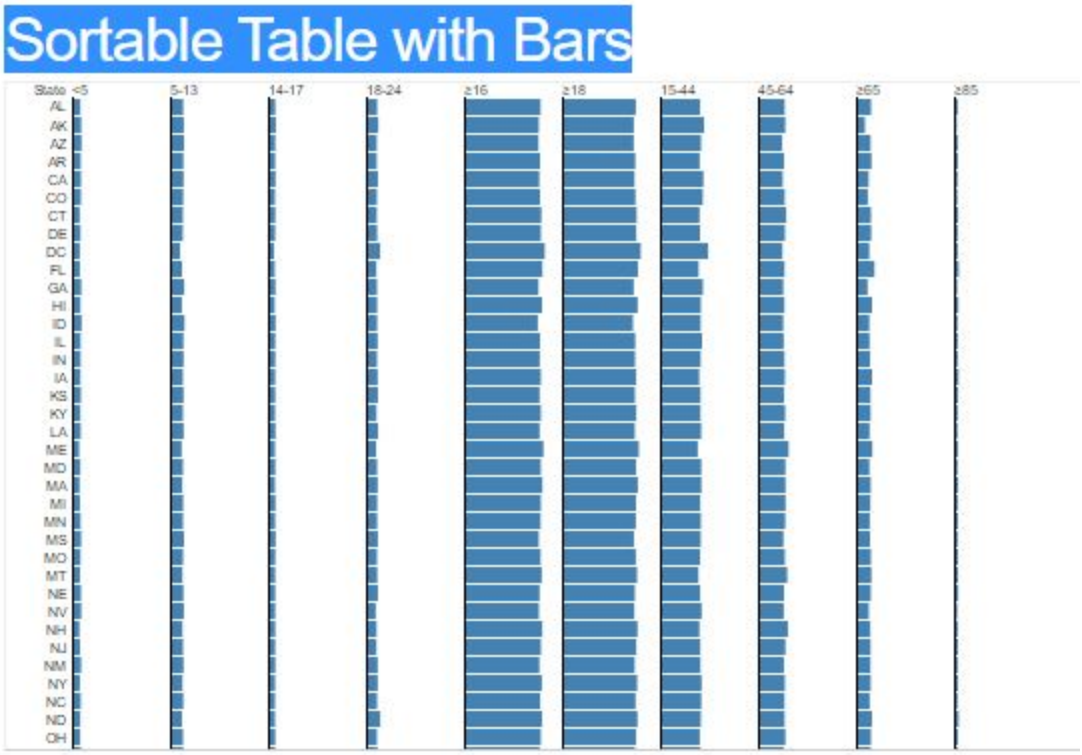
http://bl.ocks.org/mbostock/7586334

http://bl.ocks.org/mbostock/1558011

Stacked-to-Multiples

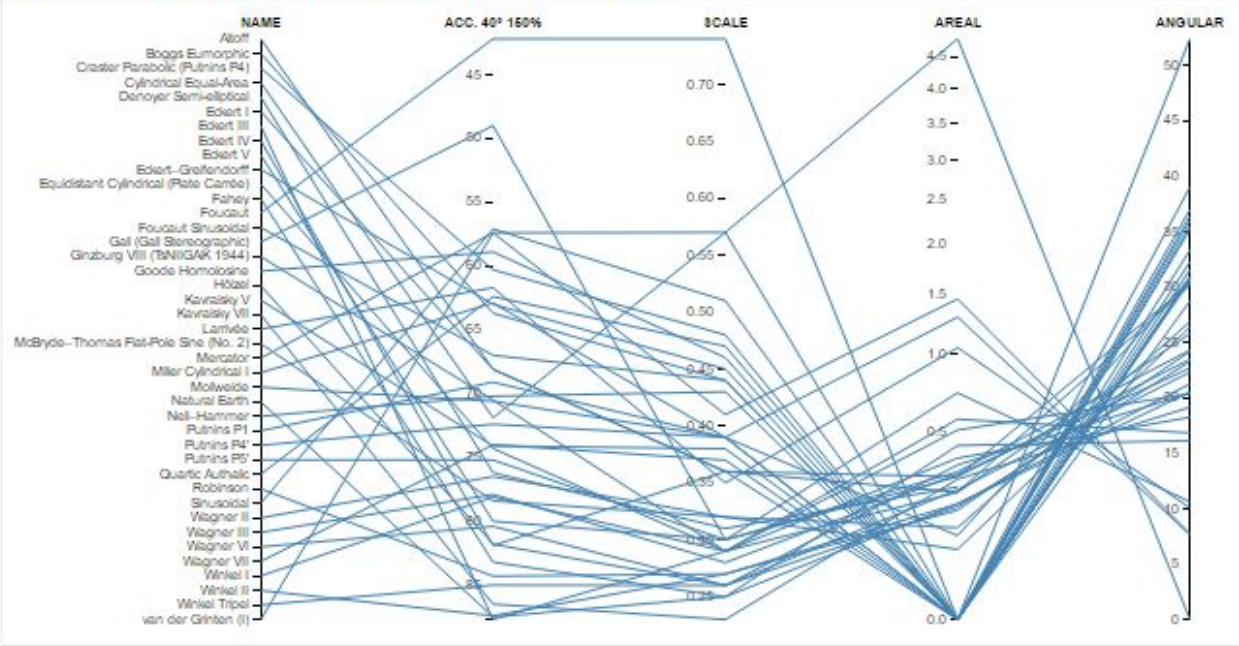


http://bl.ocks.org/mbostock/4679202

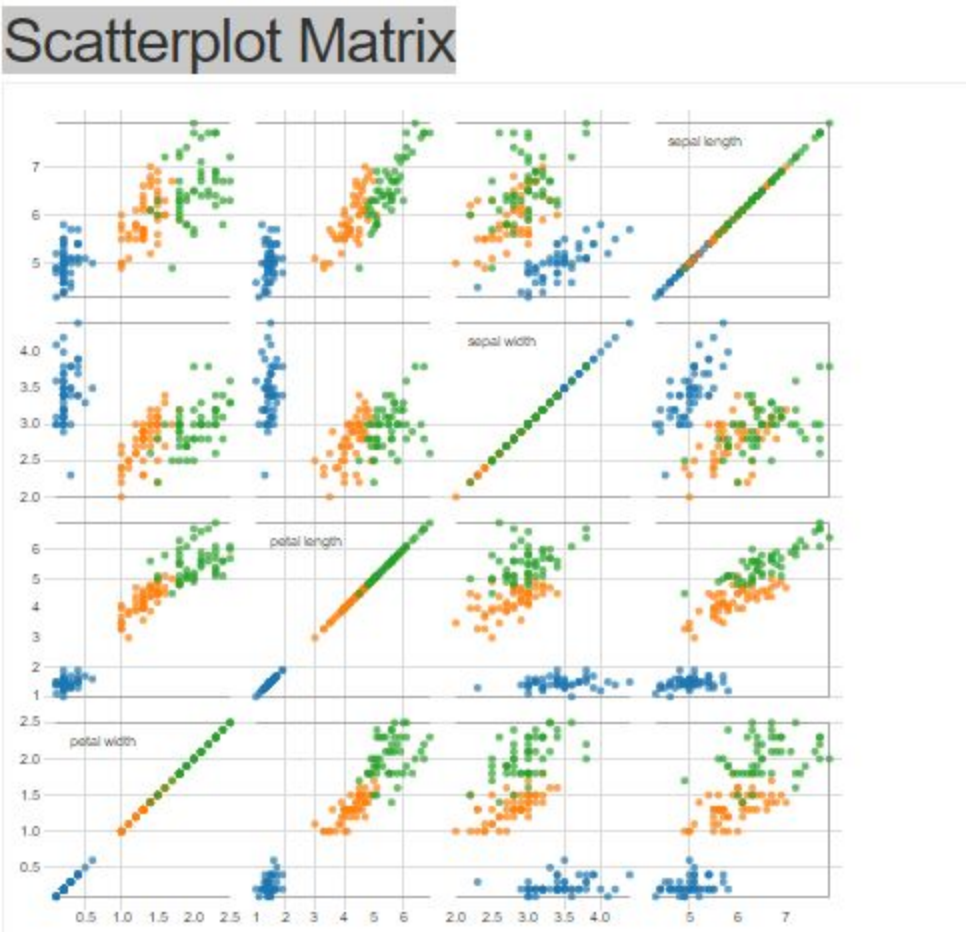


http://bl.ocks.org/mbostock/3719724

Map Projection Distortions



http://bl.ocks.org/mbostock/3709000



http://bl.ocks.org/mbostock/3213173

http://bl.ocks.org/mbostock/4063663

Prototype v2 Working

Separated all into individual simple html files with working visualizations for:

1. intro/index.html
2. Line -- viz
3. bars -- viz
4. matrix -- viz
5. mapprojection -- viz

6. map -- viz
7. methodology
8. process

- text
- optimize html/css/js structure
- bootstrap style
- check working