Rony Gregory - u1011077
Rubin Geo Varghese - u0922815
Sarathkrishna Swaminathan- u0941334

Process Book - Gender Wage Gap in USA and other OECD countries

Visualization
CS-6630

THE UNIVERSITY OF UTAH

Team Details

Project Repository: https://github.com/sarathkrishna/gender-wage-gap

Project Website : http://sarathkrishna.github.io/gender-wage-gap/united-states.html

Rony Gregory
rony.gregory@utah.edu
u1011077
https://github.com/ronygregory/dataviscourse15-hw-gregory-rony

Rubin Geo Varghese
rubin.varghese@utah.edu
u0922815
https://github.com/rubinutah/dataviscourse15-hw-varghese-rubin

Sarathkrishna Swaminathan
sarath@cs.utah.edu
u0941334
https://github.com/sarathkrishna/dataviscourse15-hw-swaminathan-sarathkrishna

Background and Motivation

It's a fact that even in 2014, women working full time in a capitalist country like the United States typically were paid just 79 percent of what men were paid, a gap of whopping 21 percent! The gap has narrowed since the 1970s, due largely to women's progress in education and workforce participation but quite slower when compared to men's wages. The progress has diminished lately and looks like it's going to stay for a while.

This wage gap affects women from all backgrounds, at all ages, ethnicities and of all levels of educational achievement, although earnings and the gap vary depending on a woman's individual situation. Recently this issue came to the limelight, thanks to the comments by Oscar winner Jennifer Lawrence about on low pay received by her for her acting in American Hustle. Selecting this topic and the aforesaid news outbreak was a coincidence, which further fueled our passion to work on this.

Rony Gregory - u1011077
Rubin Geo Varghese - u0922815
Sarathkrishna Swaminathan- u0941334

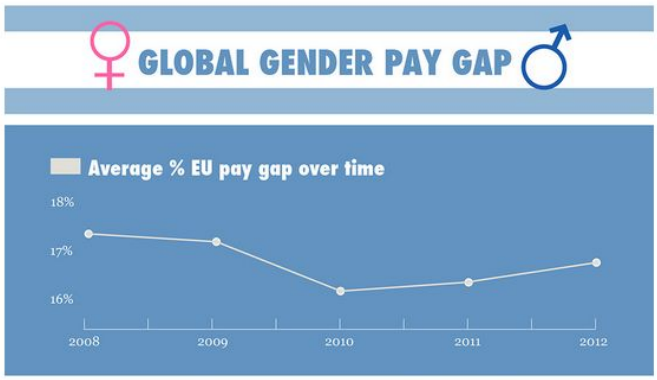
In this project, we plan to visualize the available data on the wage gap over the past many years and show the current state of US. In addition to this, we will be comparing the state of US with that of other OECD countries to provide an idea on where US stands with respect to other nations.

Related Work

Even though the gender wage gap gained much attention recently, there are few visualizations and data available to give a big picture. We considered following works to get motivation and to get data.

1) http://www.payscale.com/data-packages/gender-pay-gap : In this article, the data of United States for a certain year is visualized to understand the gender wage gap. They were trying to study the data for entire United States for different age groups and job level. This study is limited as they are not considering state-wise data and for a range of years. In this work we are trying to overcome those limitations.

2) http://www.movehub.com/blog/global-gender-pay-gap-map :



In this article, they are trying to view the global trends in gender wage gap for years from 2008 to 2012. But they are focusing more on the european countries and doesn't provide any tools to analyze the trends of each countries individually. In our work we are trying to solve the limitation of this data by providing tools to analyze the gender wage gap of different countries and to compare them against United States.

Rony Gregory - u1011077
Rubin Geo Varghese - u0922815
Sarathkrishna Swaminathan- u0941334

Questions

Through visualizing the data on gender wage gap, we are trying to answer these questions:

- How is the gender wage gap varied in United States for past several years?
- Which states in United States have high and low gender wage gaps?
- Are there any changes in gender wage gap across states for past few years?
- Which sectors in United States have high and low gender wage gaps?
- Are there any changes in gender wage gap in different sectors for past few years?
- How does United States perform against other OECD countries on gender wage gap?
- Which OECD countries have high and low gender wage gaps?
- How is the gender wage gap varied in different OECD for past several years?

By answering these questions, we will be able to get a big picture of trends in gender wage gap. Whether it's improving, getting worse or is saturating to a stable value. And by learning about the sectors, states or OECD countries with lower and higher gender wage gaps, we might be able to get some intuitions about the reasons behind the gender wage gap and will be able to propose some directions to reduce the wage gap.

Data & Data Processing

Data for different states, sectors, ethnicities, United States and OECD countries are available online. These are the data sources:

- Wage gap in OECD countries:
  - http://www.oecd.org/els/emp/Gender-wage-gaps-time-series.xlsx
- Median usual weekly earnings in US (1979-2015)
  - http://www.bls.gov/webapps/legacy/cpswktab1.htm
- Sector-wise (Median weekly earnings of full-time wage and salary workers by detailed occupation and sex)
  - 2014 - http://www.bls.gov/cps/cpsaat39.htm
  - 2013 - http://www.bls.gov/opub/reports/cps/highlights-of-womens-earnings-in-20

    - 13.pdf (page 8)
    - 2012 - http://www.bls.gov/cps/cpswom2012.pdf (pages 8 and 9)
    - 2011 - http://www.bls.gov/cps/cpswom2011.pdf (pages 12 and 13)
- State-wise overall data for 2011-2014
    - 2011 - 2013: http://www.aauw.org/resource/archive-data-gender-wage-gap-by-state-and-congressional-district/
    - 2014: http://www.aauw.org/resource/gender-pay-gap-by-state-and-congressional-district/
- Women's Earnings (includes earnings for men and earnings by education, age, race, and Hispanic ethnicity) for 1998 to 2013
    - http://www.bls.gov/cps/earnings.htm

In the OECD data, for some early years, the data for most countries are missing. We are planning to skip some early years which doesn't have data for most countries.

We are not expecting to do any other substantial data cleanup.

All these data sources provide data in terms of men's wage and women's wage. We will convert all such data to earnings ratio.

Earnings ratio = Women's earning * 100 / Men's earning

The comparison among different visualization will be easy by using similar metric across all data.

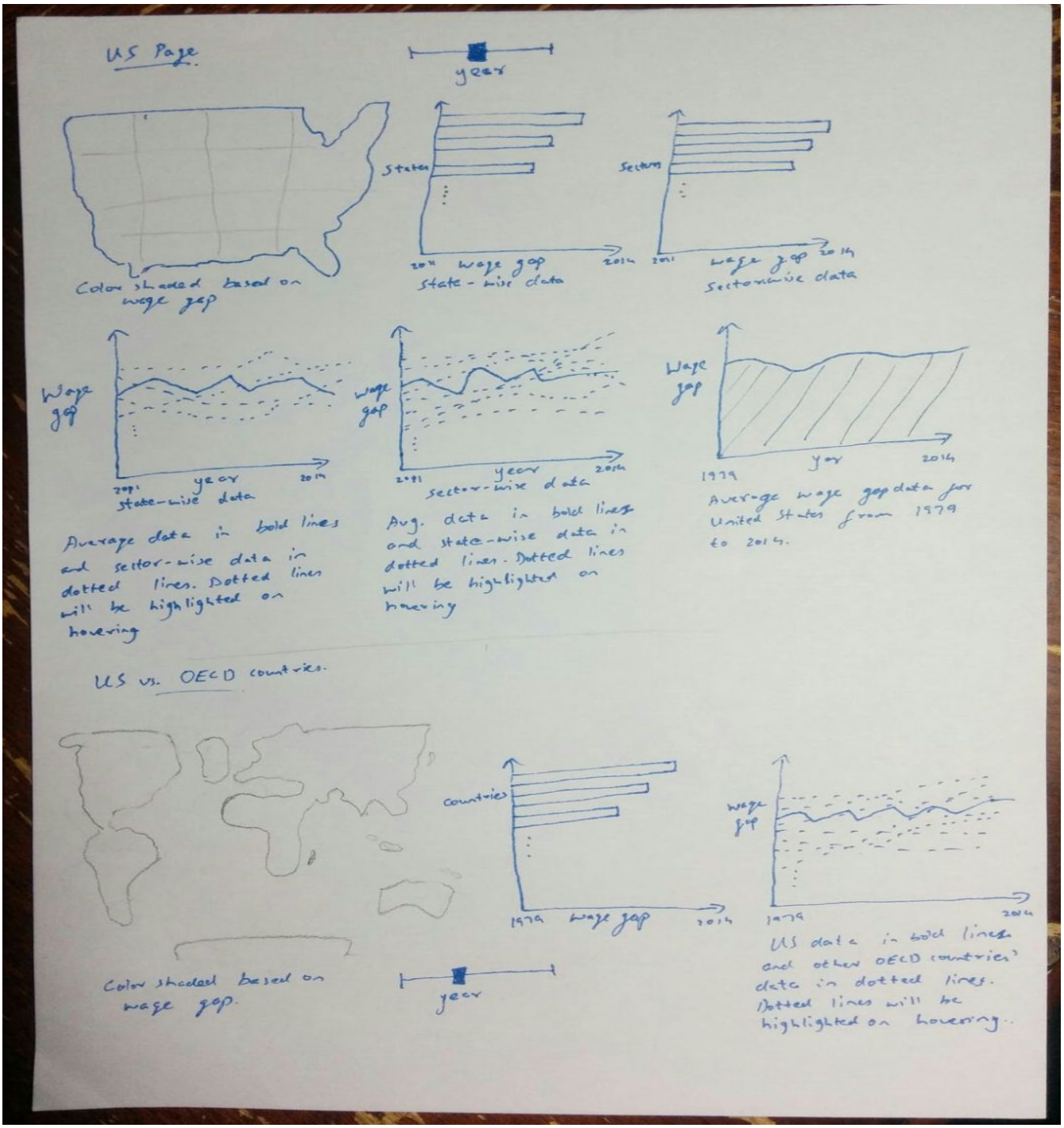
Exploratory Data Analysis

We didn’t have to perform much exploratory data analysis, as the dimension of data we have is much less than and divided into different groups like states, sector and OECD countries. Though we faced one issue, while using the OECD countries as most of the countries doesn’t have data for all years. And we are confused about how it affect our visualization. We did some basic exploration using numpy to identify whether any OECD country have no data for all the years. During this exploration, we noticed that Turkey and Slovenia have data for only one year, and hence there was no line chart corresponding to these countries.

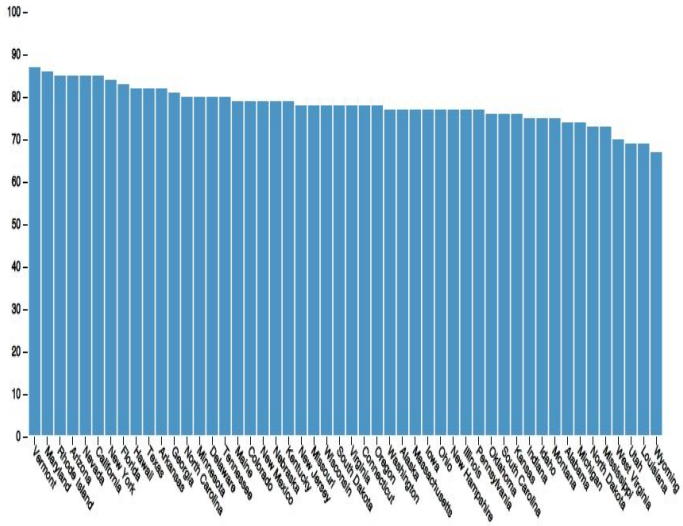
Rony Gregory - u1011077
Rubin Geo Varghese - u0922815
Sarathkrishna Swaminathan- u0941334

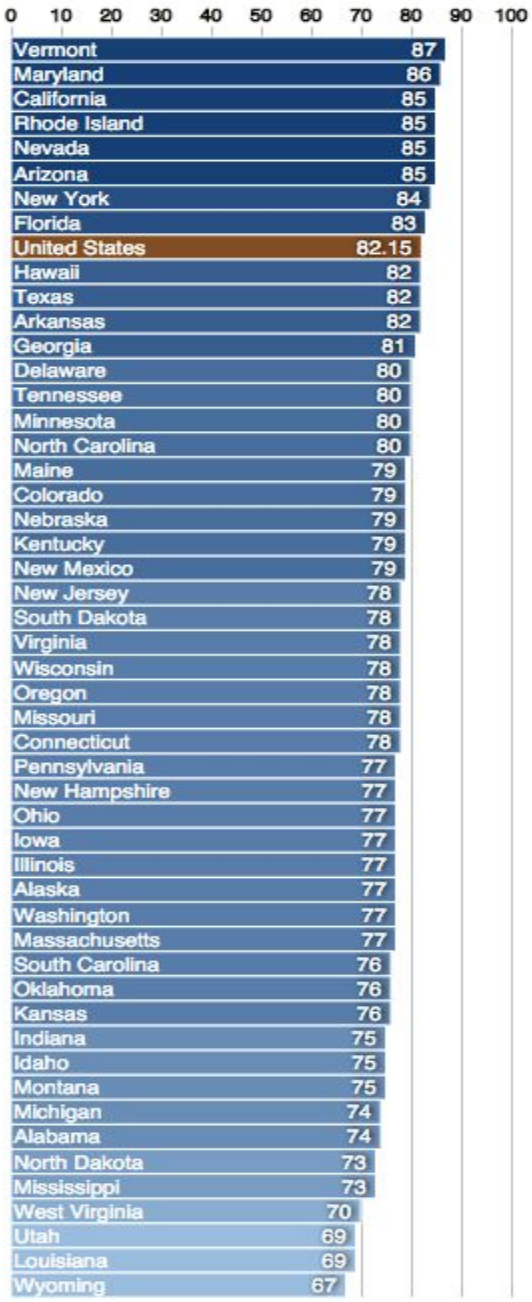
Design Evolution

We mostly stick to our initial design during the course of the project. We made some changes to avoid clutter and provide better readability. Also we have included some extra features to improve our visualization. This is our initial design :



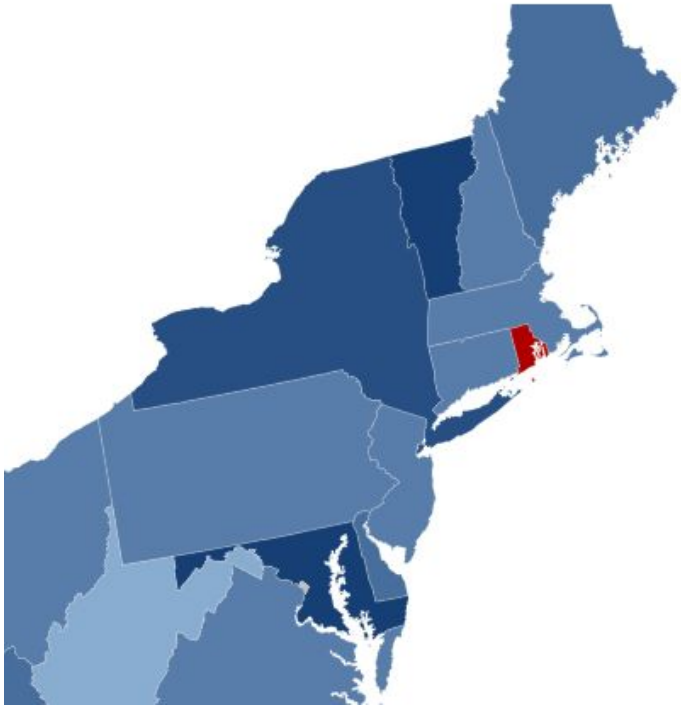
Initially we used a vertical bar chart for representing state-wise, sector-wise and country-wise data. We changed it back to a sorted horizontal bar chart as it provides a better understanding of the data and which states/sectors/countries have highest and lowest wage gap.

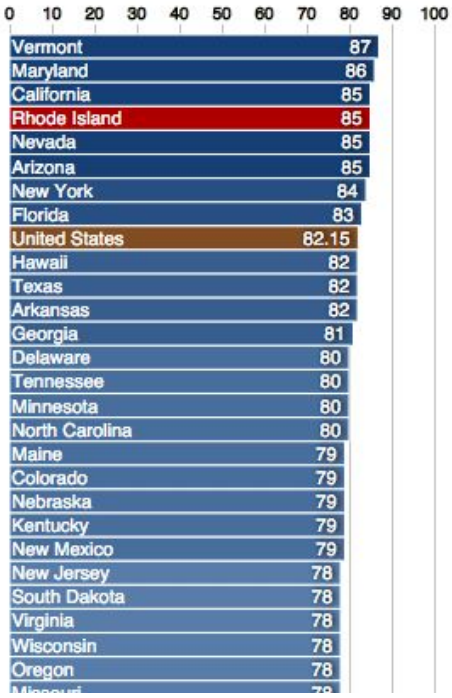




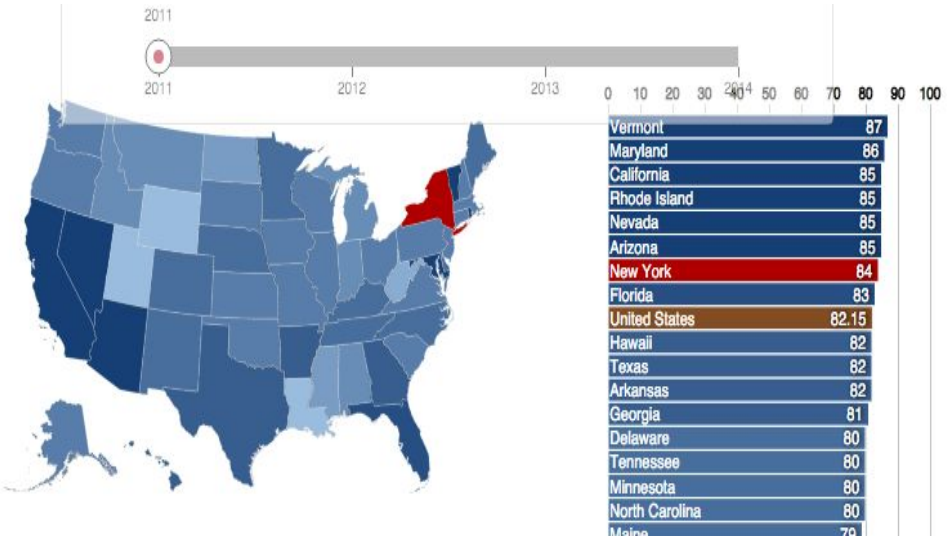
Also we have implemented a zoomable US map, to provide a clear picture for certain states in the east coast as it becomes difficult to identify those states due to their

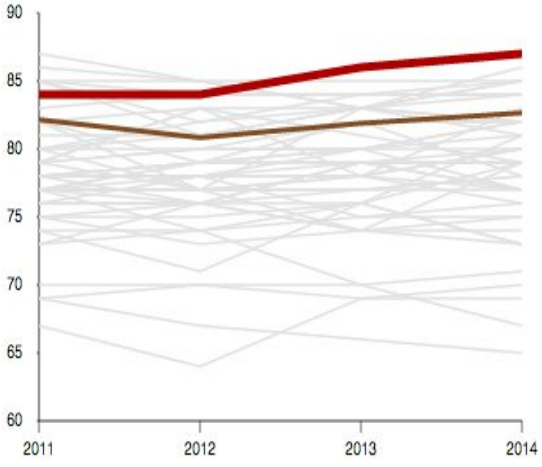
geographical size. But we have to revert that back as it slows down the use of map.





In our initial design, we were planning to plot, the united states map, horizontal bar charts for state-wise data and sector-wise data in one row, as they are dependent of the year selected in the slider. But since it become too cluttered, we re-organised them to different rows and provided a floating transparent year-slider.





Also we provided a way to connect the United States map, state-wise bar chart and state-wise line chart, by selecting the same state in all 3 charts, if a state get selected in any of these three visualization. We implemented a similar connection for 1) United States sector-wise bar chart and sector-wise line chart and 2) OCED countries map, OECD countries line chart and OECD countries bar chart.

Implementation

1. US_Data<place>

This page contains the data only for US. It makes use of the following files.

a) US_Data.html <place>

An HTML page with the required placeholders. Other than the UI part, does nothing fancy or complicated.

b) usPageMain.js

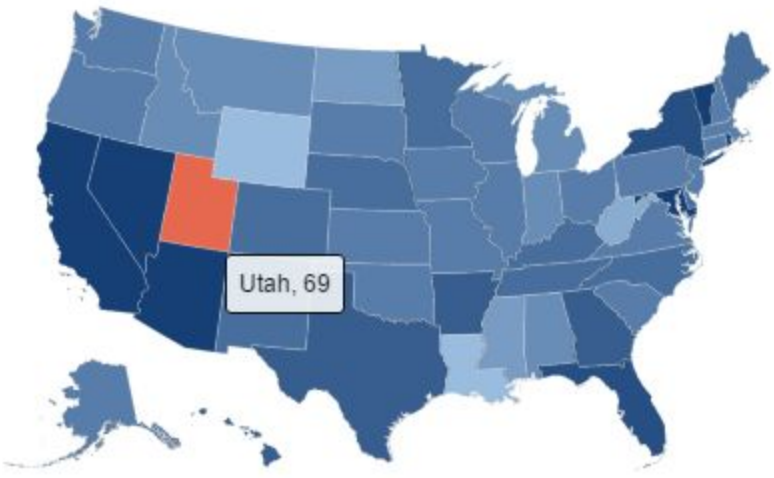
The JavaScript file that acts as the controller for the US page. It

i. creates objects for different visualizations,
ii. loads data required for them and sends them in corresponding objects,
iii. creates the slider for changing years
iv. and handles deselecting of selected state and sector.

c) usMapVis.js

Responsible for rendering the US map. It has the following features:

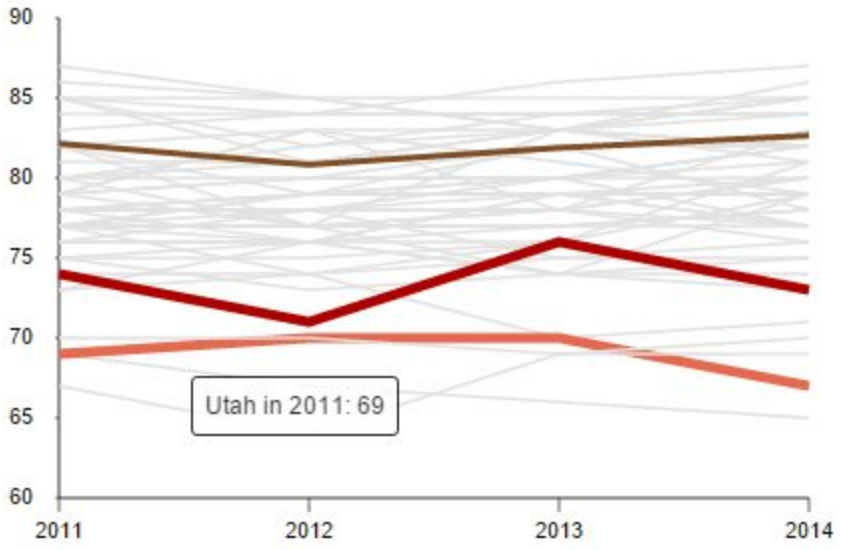
i. Color-shading with darker color implying higher female income.
ii. Hover changes color
iii. Tooltip showing state name and gender gap value.
iv. Clicking a state selects that state and updates state line chart as well as the state bar chart. Clicking elsewhere deselects the state.

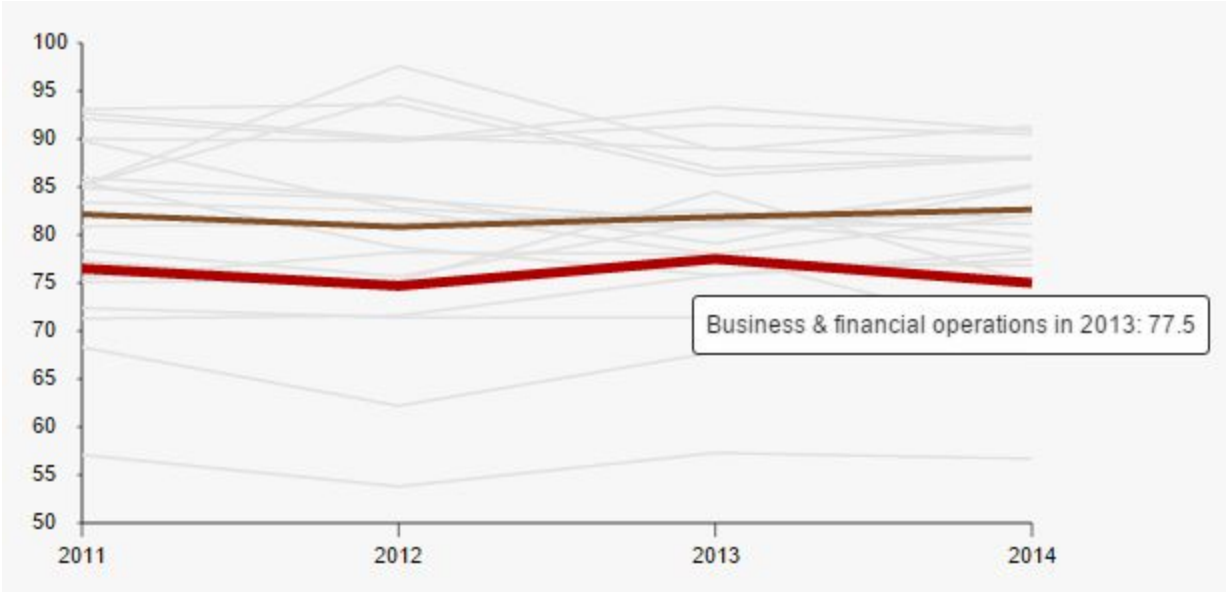


d) lineChartVis.js

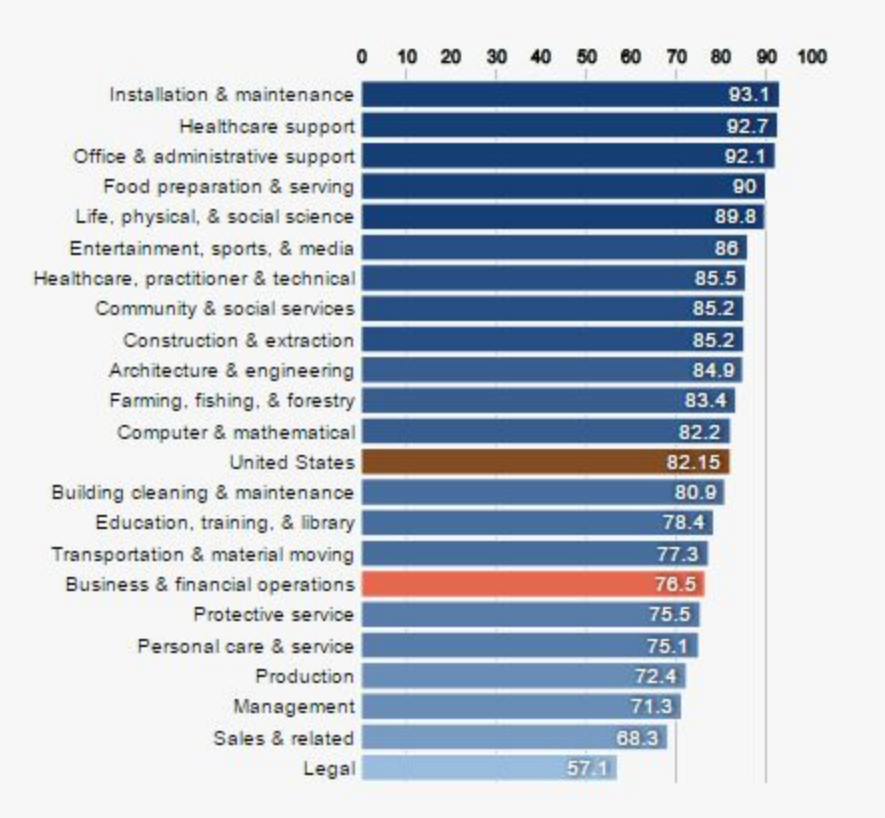
Responsible for 2 line charts – the one for states as well as sectors. This very same file is used for world line chart as well. It has the following features:

i. Hover highlights the line.
ii. Tooltip showing state/sector name and gender gap value.
iii. Clicking a state selects that line and updates the state on the map (if it's on line chart) as well as the state/sector bar chart
v. Clicking on a state brings that particular line to the front, thereby avoiding other lines appearing on top of the selected line. Clicking elsewhere deselects the line.

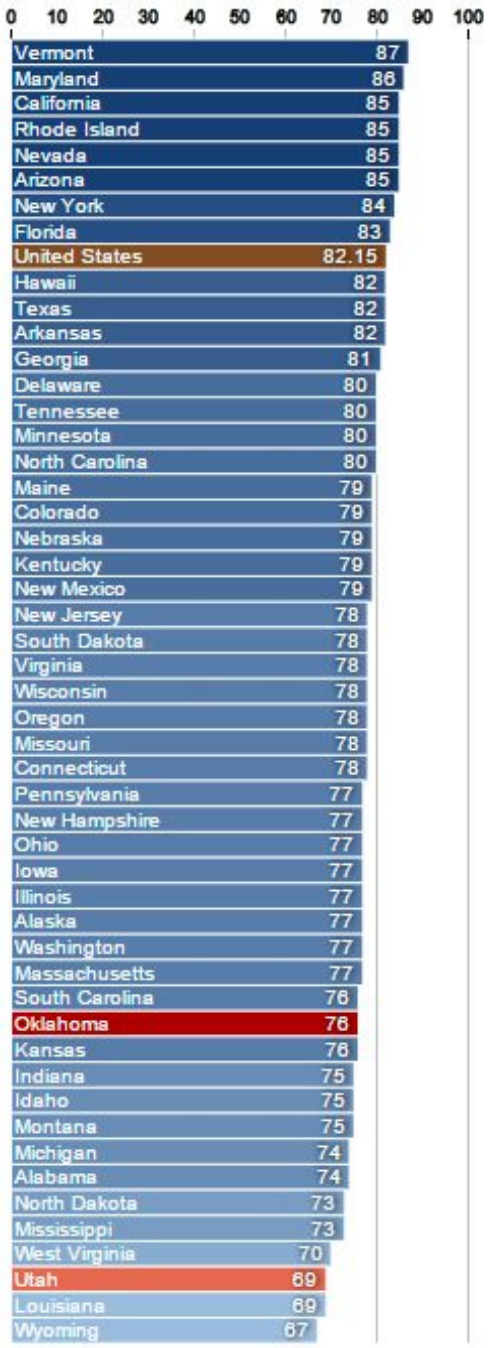




e) usSectorBarChartVis.js
Responsible for the sector bar chart. It has the following features:
   i. Color-shading with darker color implying higher female income.
   ii. The bars are sorted based on the gender gap percentage.
   iii. Hover changes color
   iv. Clicking a sector selects that bar and updates the sector line chart. Clicking elsewhere deselects the bar.



f) usStateBarChartVis.js
Responsible for the state bar chart. It has the following features:
   i. Color-shading with darker color implying higher female income.
   ii. The bars are sorted based on the gender gap percentage.
   iii. Hover changes color.
   iv. Clicking a state selects that bar and updates the state line chart. Clicking elsewhere deselects the bar.



2. Wold_Data<place>

This page contains the data for OECD counties. It makes use of the following files.

a) World_Data.html <place>
   Similar to US_page this too is an HTML page with the required placeholders.
b) worldPageMain.js

The JavaScript file that acts as the controller for the OECD page. It

i. creates objects for different visualizations,
ii. loads data required for them and sends them in corresponding objects,
iii. creates the slider for changing years
iv. and handles deselecting of selected country.

c) worldMapVis.js

Responsible for rendering the world map. It has the following features:

i. Color-shading with darker color implying higher female income.
ii. Hover changes color.
iii. Tooltip shows state name and gender gap value.
iv. Clicking a country selects that country and updates state line chart as well as the bar chart. Clicking elsewhere deselects the country.
v. Pan and zoom on the map.

d) lineChart.js

The same file used in the US data page is reused here.

e) worldBarChart.js

Responsible for the world bar chart. It has the following features:

i. Color-shading with darker color implying higher female income.
ii. The bars are sorted based on the gender gap percentage.

iii. Hover changes color.
iv. Clicking a country selects that bar and updates the country line chart. Clicking elsewhere deselects the bar.

Evaluation

In this project, we are trying to analyse the trend in gender wage gap in United States, different states of United States, different sectors of United States and OECD countries over the past few years. Also we were expecting to compare the position of United States among other OECD countries. These are some of the answers we observed from visualizing the data :

- During the last 36 years, gender wage gap reduced from 37.63 % to 18.58%.
- Only New York, Vermont, Maryland, California, Nevada, Arizona, Florida states always have less than 18% gender wage gap for the last 4 years
- Utah, Wyoming, Louisiana and West virginia always had higher wage gap among different states in United States for the last 4 years and is always greater than 30%.
- While few states like North Carolina and Oregon reduced the gender wage gap over last four years, states like Louisiana steadily increased gender wage gap
- In Office & administrative support, Food preparation & serving, Healthcare support sectors have a remarkably low gender wage gap of less than 10% for the last 4 years.
- Legal sector has a considerably high gender wage gap among all other sectors. It always had greater than 40% wage gap in the last 4 years.
- None of the sectors had a steady increase or decrease in wage gap for the last 4 years.
- Over the last 43 years most of the world countries improved the gender wage gap.
- Several countries like Spain, France, New Zealand, Australia, Hungary, Luxembourg has low gender wage gap among other countries in OECD and has significantly lower gender wage gap than United States in the last 40 years.
- Republic of Korea has the highest gender wage gap among OECD countries but it is slowly reducing the gender wage gap in the last 40 years from 47.4 to 36.6.

So with this visualization, we were able to answer most of the questions we had when we started this project.

References

- D3 world map - http://techslides.com/demos/d3/worldmap-template.html
- topojson - https://github.com/mbostock/topojson/wiki/Gallery
  - https://www.npmjs.com/package/topojson
- coloring states - https://groups.google.com/forum/#!topic/d3-js/kxMRr-CzZkY
- chloropleth - http://bl.ocks.org/mbostock/4060606
- US geojson - http://eric.clst.org/Stuff/USGeoJSON
- geojson to topojson conversion
  - http://blog.webkid.io/maps-with-leaflet-and-topojson/
- topojson installation issue
  - http://stackoverflow.com/a/24592328/812004
- Making maps with D3
  - http://volumelabs.net/making-maps-with-d3/
- Countries data
  -https://raw.githubusercontent.com/datasets/geo-boundaries-world-110m/master/countries.geojson
- Country Codes - https://raw.githubusercontent.com/datasets/country-codes/master/data/country-codes.csv
- d3 slider - http://thematicmapping.org/playground/d3/d3.slider/
- US Map States - Choropleth Plus Bar (An example)
  - https://vida.io/documents/4vZ9mRGyepoyQxFcK
- Line chart - http://projects.flowingdata.com/life-expectancy/
- Tooltips - http://bl.ocks.org/lhoworko/7753a11efc189a936371
- zoom on maps - http://bl.ocks.org/mbostock/eec4a6cda2f573574a11
- Bringing an html element to front
  - https://gist.github.com/trtg/3922684
-
- US state-wise Data: https://docs.google.com/spreadsheets/d/1WgPxsDur8OTUkF8K0KMfKvl8U_3iiiPoF0fPGRzFiGw/edit?usp=sharing
- Horizontal Bar Chart - https://vida.io/documents/4vZ9mRGyepoyQxFcK
- Commands used - topojson states.json -o out.json -p
  topojson countries.json -o out.json -p