Visualization of Unemployment and Crime Rate across USA

Mingxuan Luo u1266548 u1266548@utah.edu
Yulan Wang u1264235 youlan9115@gmail.com
Qing Ye u6007114 q.ye@utah.edu

Basic Info.

Project repository:
https://github.com/blackispower/Unemployment-in-USA

Overview and Motivation.

Unemployment is describing the state that a person above some specified age is not in paid employment or self-employment and are currently available for work during the reference period. It is measured by the unemployment rate. Factors influencing unemployment rate may be caused by the recession of economy, new technologies, policies of the government and so forth.

Crime rate is defined by the ratio of crimes in an area to the population of the area, which is a count of crimes complied to assess the effectiveness of a crime control policy, and the impact of the policy on the risk of crime victimization.

In our project, we decided to visualize the unemployment and crime rates across USA between 2007 to 2018. We would like to reveal the trend of unemployment of specific regions (country- level, state- level) over time. Meantime, we analyzed the relationship between unemployment rate and Crime rate, income, population and visualize their change trend over time. We also found there is a strong influence of Financial crisis (2007–2008) on the unemployment rate, income and so forth.

We are interested in this topic because this is very relevant to our daily life. For individuals, these information could help them find an ideal place to settle down. For government, knowing the relationship between unemployment rate and crime rate could help goverment to make more effective crime control policy.

Related Work.

Local Area Unemployment Statistics Map

https://data.bls.gov/lausmap/showMap.jsp;jsessionid=79A96FEAED81F05DC79399CE15D59304._t3_07v

Questions.

1. Reveal the trend of unemployment, crime rate of specific regions over time.
2. Explore and compare the country- level and state- level unemployment rate, crime rate.
3. Find the regions with extremely high or low unemployment or crime rate.
4. Visualize the trend of unemployment, crime rate, income and population for specific regions. Check the correlation between crime rate and unemployment rate.

Data

Data sources

Unemployment rate data:

Kaggle: https://www.kaggle.com/jayrav13/unemployment-by-county-us

Bureau of Labor Statistics https://data.bls.gov/timeseries/LNS14000000

Crime rate and population data:

Crime in the U.S. https://ucr.fbi.gov/crime-in-the-u.s

Income data :

https://www.census.gov/topics/income-poverty/data/tables.All.html (Following TA's advice to get more date relevant to the unemployment rate. Poverty rate in state-level were not found, so we moved forward with household income data)

Data Clean

The dataset provides us with the unemployment rate date in different counties across USA from 2007 to 2018.

Unemployment rate data:

First, we acquire the unemployment rate data for each county of each state from the website of United States Department of Labor (https://data.bls.gov/lausmap/showMap.jsp;jsessionid=79A96FEAED81F05DC79399CE15D59304._t3_07v). Since the raw data is county-level, we need to preprocess the data to get aggregation of the data from county to state. Finally, we have the county level- unemployment rate data shown as Figure 1, and state level unemployment data shown as Figure 2. Since we only could get unemployment rate in county level, after meeting with our TA Ilkin, we decided not to show the county-level data.

A	B	C	D	E	F	G	H
County	State	Labor Force	Employed	Unemployed	Unemployment Rate	Year	
Autauga County	AL	24,383	23,577	806	3.3	2007	
Baldwin County	AL	82,659	80,099	2,560	3.1	2007	
Barbour County	AL	10,334	9,684	650	6.3	2007	
Bibb County	AL	8,791	8,432	359	4.1	2007	
Blount County	AL	26,629	25,780	849	3.2	2007	
Bullock County	AL	3,653	3,308	345	9.4	2007	
Butler County	AL	9,099	8,539	560	6.2	2007	
Calhoun County	AL	54,861	52,709	2,152	3.9	2007	
Chambers County	AL	15,474	14,469	1,005	6.5	2007	
Cherokee County	AL	11,984	11,484	500	4.2	2007	
Chilton County	AL	19,737	19,067	670	3.4	2007	
Choctaw County	AL	5,183	4,875	308	5.9	2007	
Clarke County	AL	10,358	9,730	628	6.1	2007	
Clay County	AL	6,113	5,775	338	5.5	2007	
Cleburne County	AL	6,483	6,246	237	3.7	2007	

Figure 1. Unemployment data for each county of the US from 2007 to 2018.

	A	B	C
1	State	Unempolyment-rate	Year
2	Alabama	4	2007
3	Alabama	5.7	2008
4	Alabama	11	2009
5	Alabama	10.5	2010
6	Alabama	9.6	2011
7	Alabama	8	2012
8	Alabama	7.2	2013
9	Alabama	6.8	2014
10	Alabama	6.1	2015
11	Alabama	5.8	2016
12	Alabama	4.4	2017
13	Alabama	3.9	2018
14	Alaska	6.3	2007
15	Alaska	6.7	2008
16	Alaska	7.7	2009
17	Alaska	7.9	2010
18	Alaska	7.6	2011
19	Alaska	7.1	2012
20	Alaska	7	2013
21	Alaska	6.9	2014
22	Alaska	6.5	2015
23	Alaska	6.9	2016
24	Alaska	7	2017
25	Alaska	6.6	2018
26	Arizona	3.9	2007

Figure 2. Unemployment data for each state of the US from 2007 to 2018.

Crime rate data:

For the crime rate data, we collected the data from the website https://ucr.fbi.gov/crime-in-the-u.s. After simple data clean process, the crime data is shown as Figure 3.

	A	B	C	D	E
1	Area	Year	Population	number	rate
2	UnitedStates	2007	301,621,157	1,408,337	466.9
3	UnitedStates	2008	304,059,724	1,382,012	454.5
4	UnitedStates	2009	307,006,550	1,325,896	431.9
5	UnitedStates	2010	308,745,538	1,246,248	403.6
6	UnitedStates	2011	311,587,816	1,206,031	387.1
7	UnitedStates	2012	313,914,040	1,214,464	386.9
8	UnitedStates	2013	316,497,531	1,199,684	379.1
9	UnitedStates	2014	318,857,056	1,197,987	375.7
10	UnitedStates	2015	320,896,618	1,234,183	384.6
11	UnitedStates	2016	323,127,513	1,283,058	397.1
12	UnitedStates	2017	325,147,121	1,283,875	394.9
13	UnitedStates	2018	327,167,434	1,245,065	380.6
14	Connecticut	2007	3,502,309	8,965	256
15	Connecticut	2008	3,501,252	10,427	297.8
16	Connecticut	2009	3,518,288	10,588	300.9
17	Connecticut	2010	3,574,097	10,057	281.4
18	Connecticut	2011	3,586,717	9,889	275.7
19	Connecticut	2012	3,590,347	10,160	283
20	Connecticut	2013	3,599,341	9,439	262.2
21	Connecticut	2014	3,596,677	8,522	236.9
22	Connecticut	2015	3,584,730	7,938	221.4
23	Connecticut	2016	3,576,452	8,123	227.1
24	Connecticut	2017	3,573,880	8,190	229.2
25	Connecticut	2018	3,572,665	7,411	207.4
26	Maine	2007	1,317,207	1,554	118
27	Maine	2008	1,316,456	1,547	117.5
28	Maine	2009	1,318,301	1,580	119.9
29	Maine	2010	1,328,361	1,621	122
30	Maine	2011	1,328,544	1,638	123.3
31	Maine	2012	1,329,192	1,631	122.7
32	Maine	2013	1,328,702	1,761	132.5
33	Maine	2014	1,330,089	1,700	127.8
34	Maine	2015	1,329,453	1,726	129.8
35	Maine	2016	1,331,479	1,648	123.8
36	Maine	2017	1,335,063	1,610	120.6

Figure 3. the crime data for each state in the US from 2007 to 2018.

Income data :

Original income data are difficult to read and use in this project (Figure 4), we used a python script to automate the cleaning and get each state's information including unemployment rate, population, unemployment rate by year (Figure 5).

State	2018		2017 (40)		2017		2016		2015		2014		2013 (39)	
	Median income	Standard error	Median income	Standard error	Median income	Standard error	Median income	Standard error	Median income	Standard error	Median income	Standard error	Median income	Standard error
United States	63,179	420	61,136	322	61,372	335	59,039	436	56,516	321	53,657	392	53,585	654
Alabama	49,936	2,423	50,865	1,094	51,113	845	47,221	2,301	44,509	3,419	42,278	1,529	47,320	5,511
Alaska	68,734	3,390	77,987	3,718	72,231	2,719	75,723	4,086	75,112	3,485	67,629	3,153	72,472	5,446
Arizona	62,283	2,291	59,700	2,689	61,125	2,642	57,100	1,971	52,248	2,008	49,254	2,304	52,611	4,365
Arkansas	49,781	2,108	49,751	2,491	48,829	2,642	45,907	2,165	42,798	1,572	44,922	2,546	39,376	2,402
California	70,489	1,233	70,038	1,576	69,759	1,452	66,637	1,075	63,636	1,711	60,487	894	60,794	1,811
Colorado	73,034	3,562	74,984	2,661	74,172	2,982	70,566	4,125	66,596	3,682	60,940	2,414	67,912	4,129
Connecticut	72,812	5,119	74,304	3,453	72,780	3,509	75,923	3,404	72,889	4,449	70,161	2,665	69,291	3,238
Delaware	65,012	3,599	64,961	5,230	62,318	3,300	58,046	2,992	57,756	3,392	57,522	2,970	54,091	4,350
D.C.	85,750	2,659	81,282	1,903	83,382	3,319	70,982	2,734	70,071	2,861	68,277	4,402	60,057	5,529
Florida	54,644	1,887	53,086	1,438	53,681	1,816	51,176	835	48,825	1,395	46,140	950	48,532	2,815
Georgia	55,821	2,032	57,985	2,288	57,016	2,434	53,527	1,679	50,768	1,401	49,555	2,009	46,992	4,236
Hawaii	80,108	3,482	73,599	2,352	73,575	2,227	72,133	2,510	64,514	2,840	71,223	2,676	64,235	3,918
Idaho	58,728	2,405	59,497	2,259	60,208	2,400	56,564	1,261	51,624	1,260	53,438	2,572	48,467	5,484
Illinois	70,145	2,253	65,969	1,667	64,609	1,999	61,386	1,761	60,413	1,999	54,916	1,625	53,937	3,287
Indiana	59,892	1,997	58,767	1,655	58,873	2,026	56,094	2,877	51,983	1,682	48,060	1,943	49,455	4,508
Iowa	68,718	2,910	63,467	4,760	63,481	5,512	59,094	2,484	60,855	3,275	57,810	3,425	60,156	6,134
Kansas	63,938	2,027	56,900	1,937	57,872	2,898	56,810	3,268	54,865	3,731	53,444	2,863	47,820	4,496
Kentucky	54,555	3,668	49,672	1,885	51,348	1,364	45,369	2,005	42,387	2,269	42,786	1,580	44,879	3,924
Louisiana	49,973	1,754	43,565	2,088	43,903	1,453	42,196	1,197	45,922	2,140	42,406	1,862	46,425	2,543
Maine	58,663	5,683	53,316	3,121	51,664	2,146	50,856	3,346	50,756	1,929	51,710	2,171	54,957	3,863
Maryland	86,223	3,331	82,093	3,643	81,084	3,652	73,760	2,591	73,594	2,692	76,165	4,272	69,353	2,970
Massachusetts	86,345	2,785	76,243	2,394	73,227	3,328	72,266	2,911	67,861	2,790	63,151	2,662	62,529	5,303

Figure 4. Original data of household income

State	Unemployment-rate	Year	Income	Population
Alabama	4	2007	42,212	4,627,851
Alabama	5.7	2008	44,476	4,661,900
Alabama	11	2009	39,980	4,708,708
Alabama	10.5	2010	40,933	4,779,736
Alabama	9.6	2011	42,590	4,803,689
Alabama	8	2012	43,464	4,822,023
Alabama	7.2	2013	47,320	4,833,996
Alabama	6.8	2014	42,278	4,849,377
Alabama	6.1	2015	44,509	4,853,875

Figure 5. Python-cleaned data combining unemployment rate, income, population.

Exploratory Data Analysis:

We have used 4 kinds of visulizations for our data, including map chart, line chart. bar chart and bubble chart.

First, we used map chart to shown the unemployment or crime data. The map makes it easy to understand the distribution of the organization's presence across the city, state. And it is convenient to compare the activity across several locations at a glance, and contextualizing data in the real world.

Second, to visualize the data change trend, we used line chart. Line chart can show data variables and trends clearly and directly, which makes it easy to find some sharp increase and decrease, helping to let us think what might causes this abnormality.

Third, bar chart was used to visualize the data for each state or county. Compare to map chart, bar chart was marginated by length. The benefits of bar chart includes displaying relative differences of multiple categories, summarize the extremely value of the data.

Lastly, we are using bubble chart and related selection button to shown the correlation between the factors we are interested. Readers could use this function to explore what they are interested.

In addition, we also used several interaction in our project that including hover, zoom, sorting, linking among different graph.

Design Evolution

The original design:

Overview

The overview of the visualization consists of three parts: map chart, line chart, bar chart and bubble chart, and can toggle between the unemployment rate, the crime rate and the combined view.

Map chart is coded by color, suggesting the rate of unemployment/crime rate in each state. And on the combined view, the color indicates the unemployment rate while the crime rate is represented by the size of a circle. A year slider on the bottom can switch between different years. A tool tip should appear on hover. If a state is clicked, the map would zoom in and show a more detailed unemployment rate information in different counties. (Figure 6).

Line chart for the first two view shows the change over the years and it would highlight the corresponding line for a state when hovering on a specific one. It changes into a scatter plot with X-axis and Y-axis showing the stastitics. When clicking on a state, because we only have data for unemployment rate for each county, the line chart only show changes for the unemployment rate for each county. A description box shows the crime rate and overall unemployment rate in the state. (Figure 6).

Bar chart shows the comparison between states/counties for a specific year. Hovering on map or a bar would highlight each other. The whole chart could sort by alphabetic order or the statistics (Figure 6).

Must-Have Features.

Three views for the map chart and the corresponding functions in line and bar charts

Optional Features.

Zooming function for each counties' data. However, after discussing and with the permission from TA, we decided to drop this freature cause we could only obtain the unemployment rate in county level.

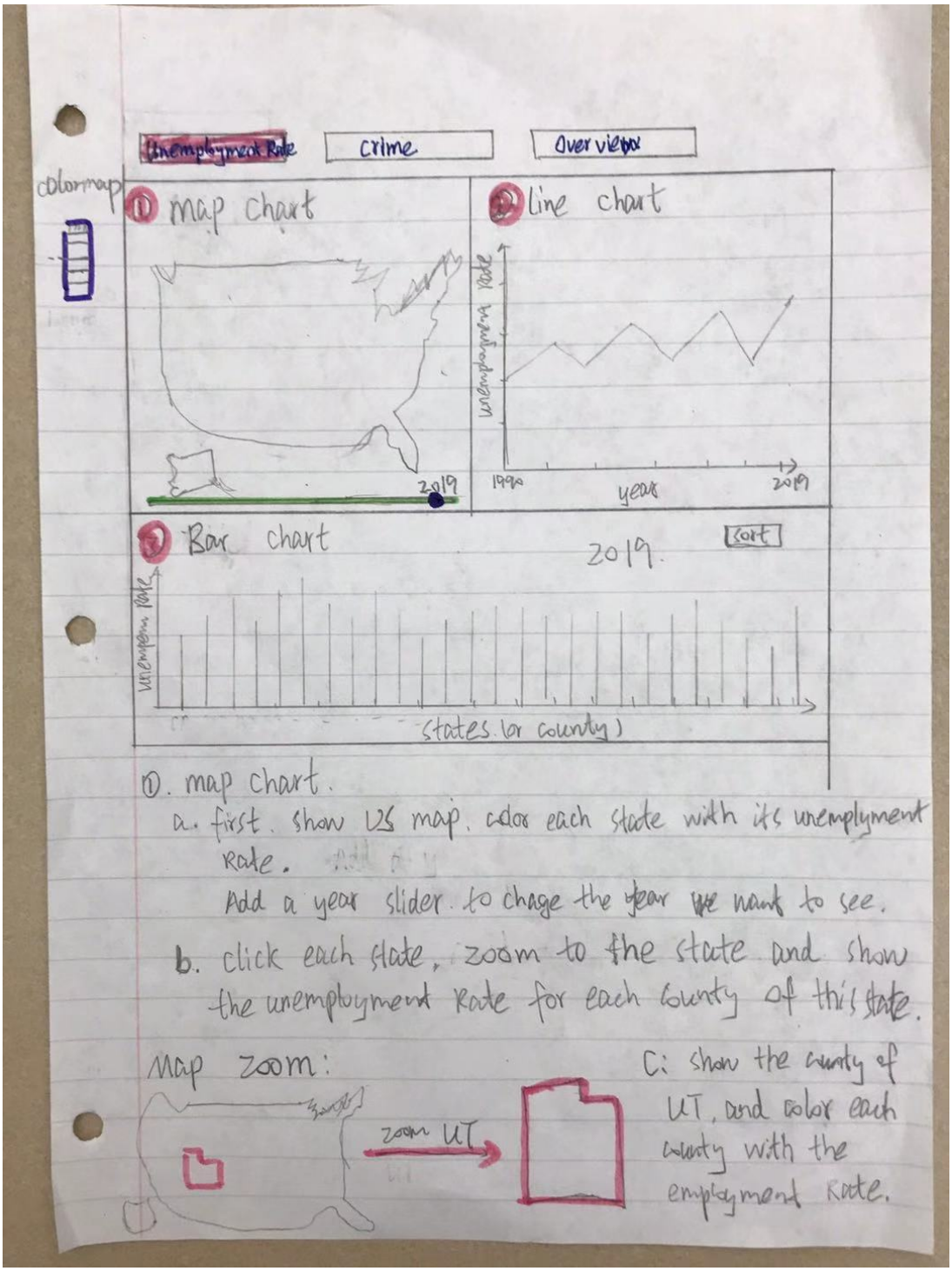


Figure 6. Original design- Part 1

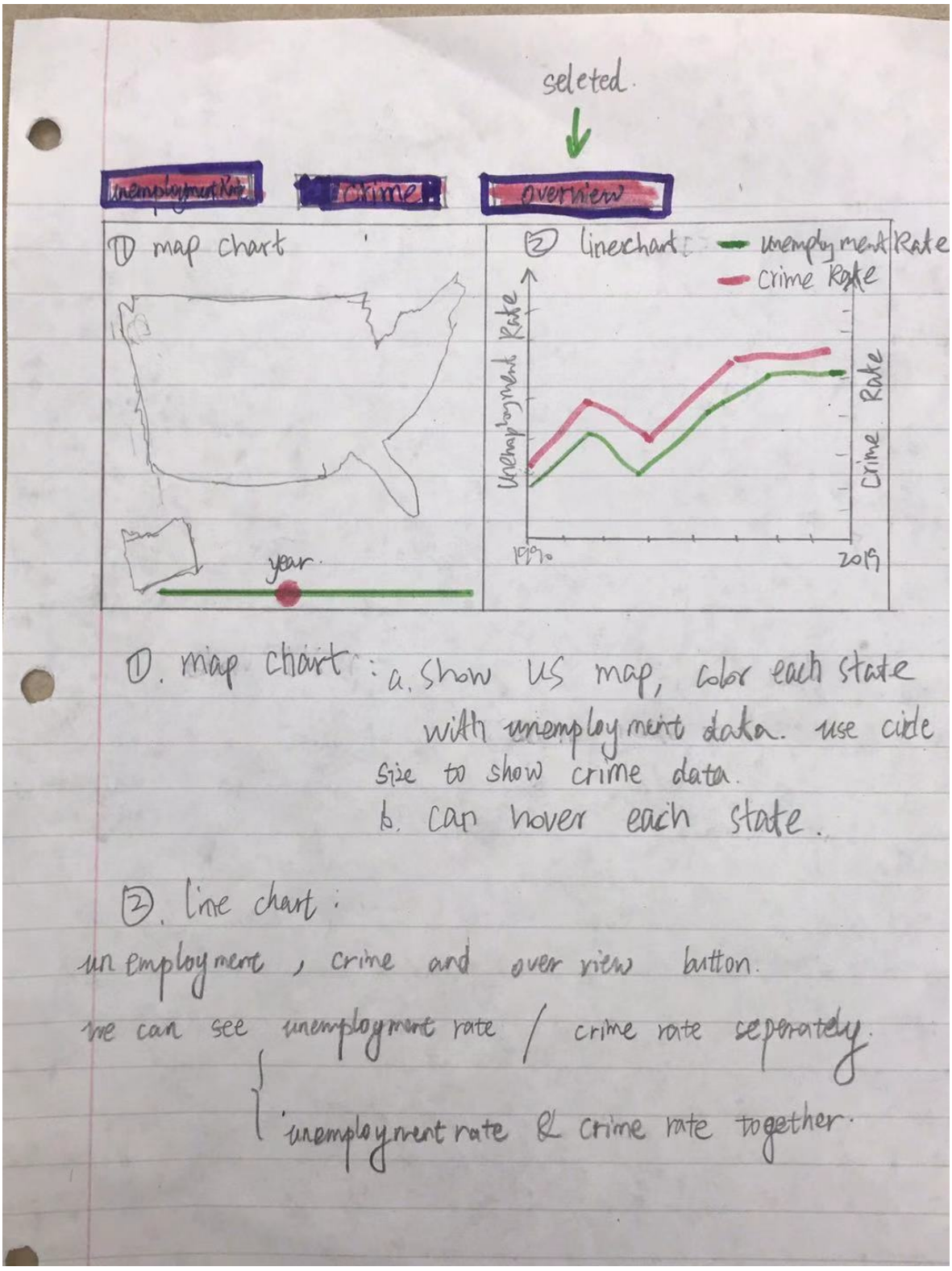


Figure 6. Original design- Part 2

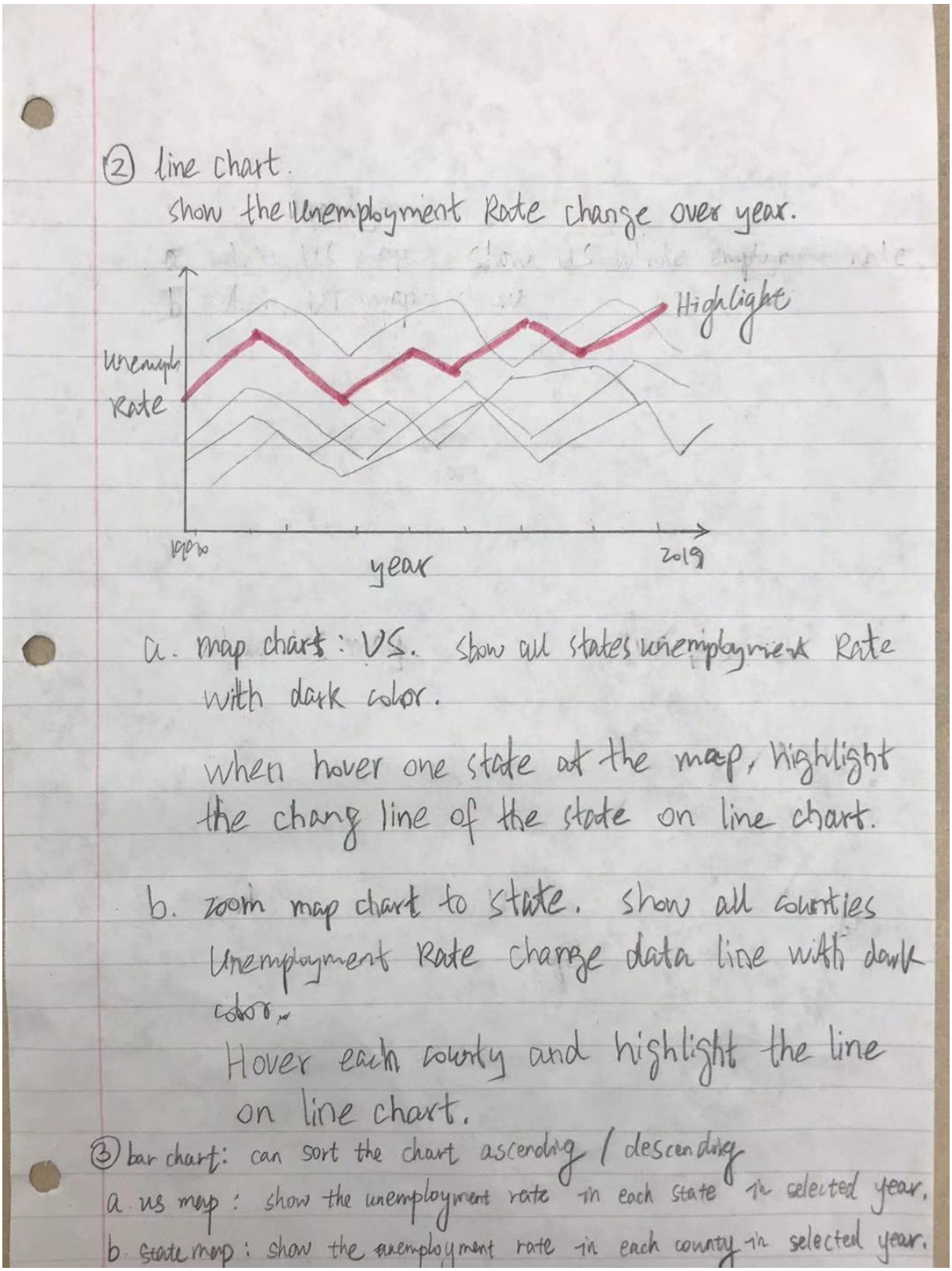


Figure 6. Original design- Part 3

We changed design over time, with the discussion ourselves and together with TA. For example, we abandoned the county-level data because we could only find the county-level data of unemployment rate and we think we should be better consistent with data.

Design pattern we tried to follow here: Balance the design, propotion, emphasize the key areas, variety and so forth.

Implementation.

1. There are 4 main views in our design and we are using four buttons to allow users choose from the 4 views.

We used dark black as the background because we want to highlight the chart we show here.

a) Unemployment rate among different states (2007-2018). Shown with map chart, line chart and bar chart.

We used the color blue here to represent the degree of unemployment because the semantic meaning of blue comes with things like economic recession.

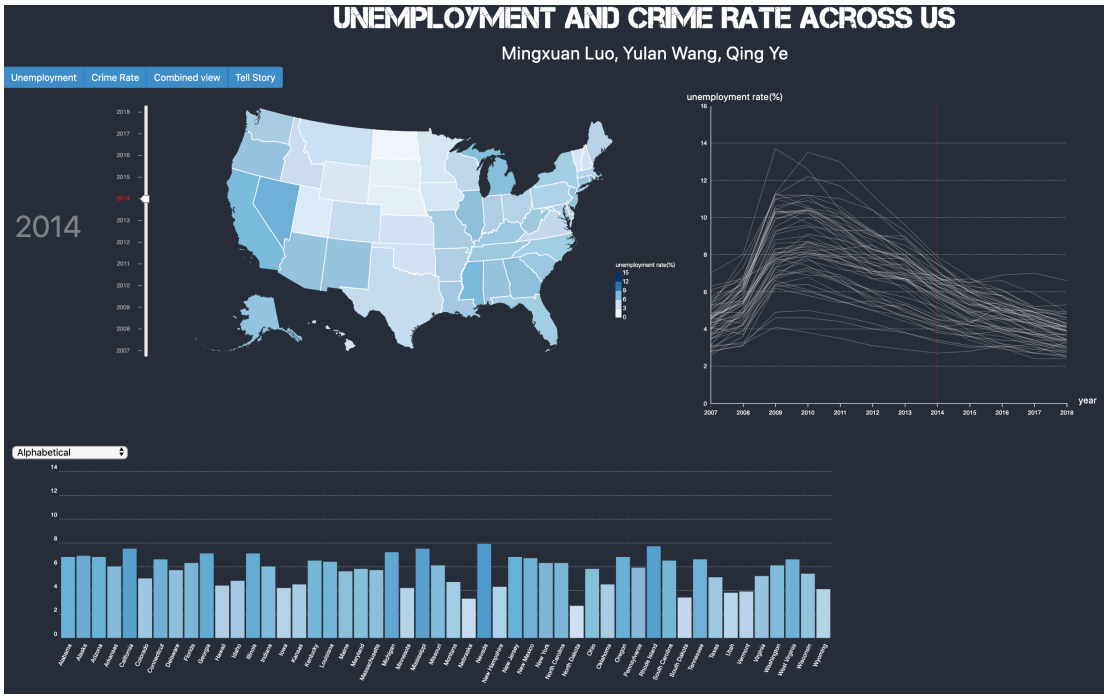


Figure 7. overview of unemployment rate page

b) Crime rate among different states (2007-2018) Shown with map chart, line chart and bar chart.

c) We used the color red here to represent the crime rate because the semantic meaning of red comes with things like violence, crime, bleeding and so forth.

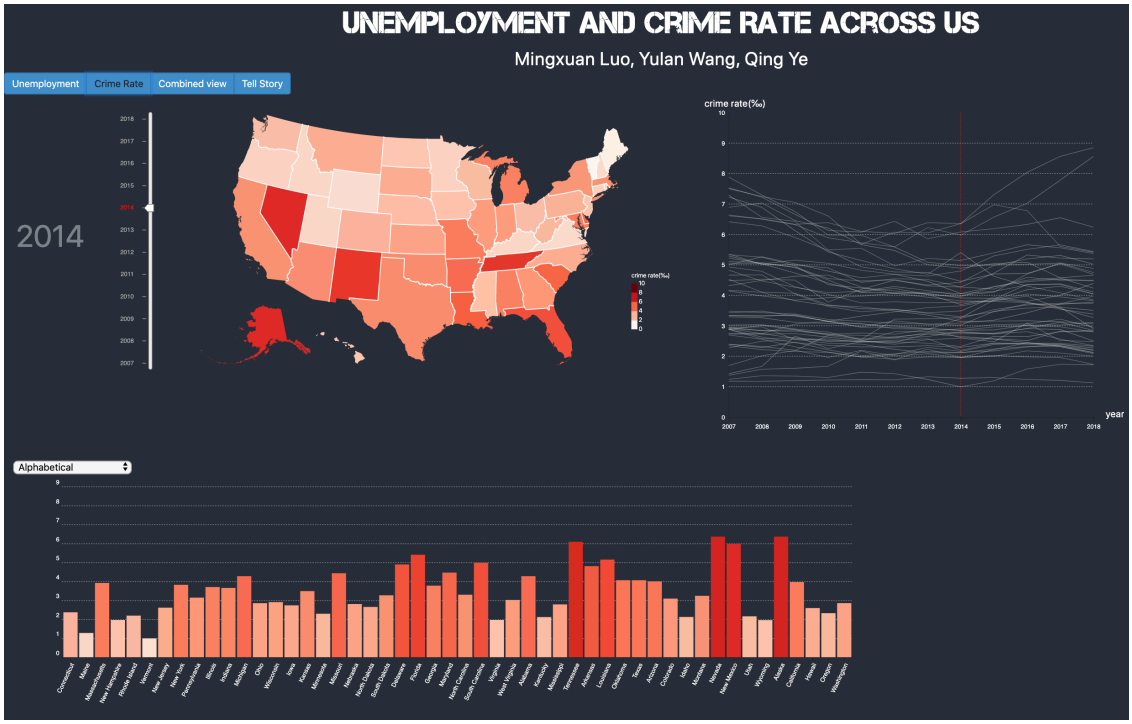


Figure 8. overview of crime rate page

d) Combined view to show the correlation between population, household income, crime rate and unemployment rate (among different states (2007-2018). Shown with line chart and bubble chart. In the combined view, the default data in map chart and line chart is based on the unemployment rate.

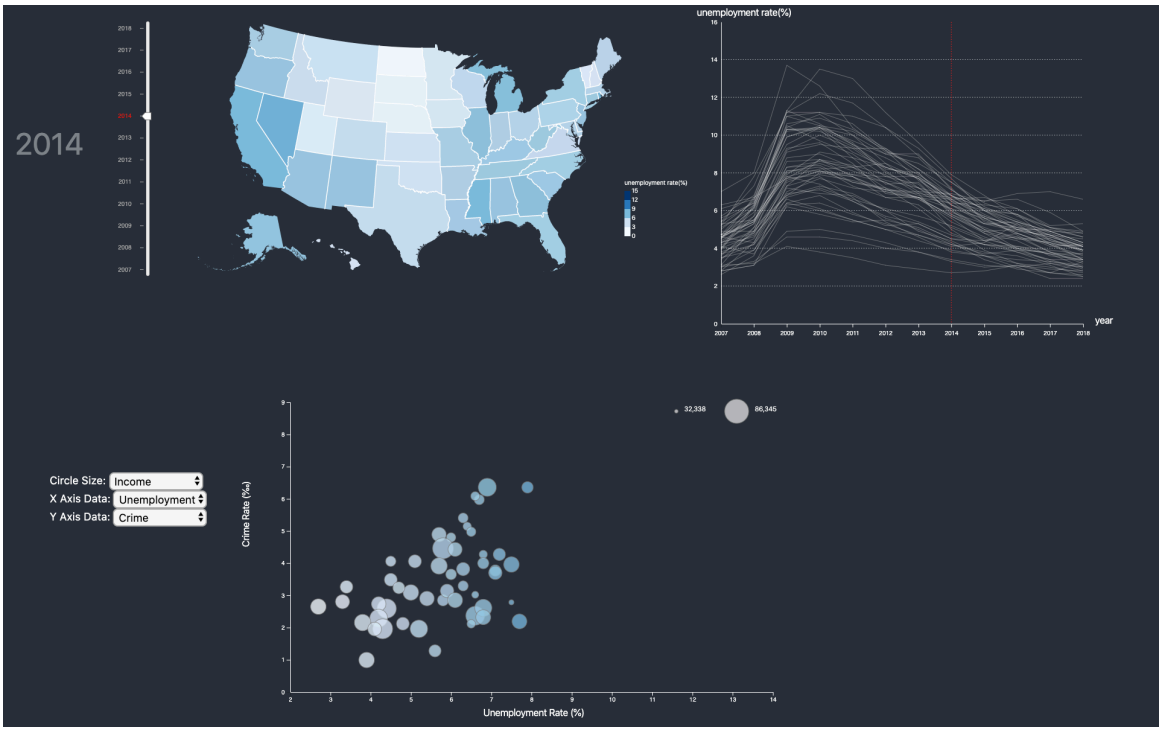


Figure 9. overview of combined page

The map chart

In the map chart, we colored each state based on the crime rate or unemployment rate based on the view the reader choose.

The year slider is added to allow users to choose the data based on the year. When hovering on a state, the detailed information about the state name, crime rate, unemployment rate, population and household income were shown. And the related states were also highlighted in orange in the line chart and bar chart.

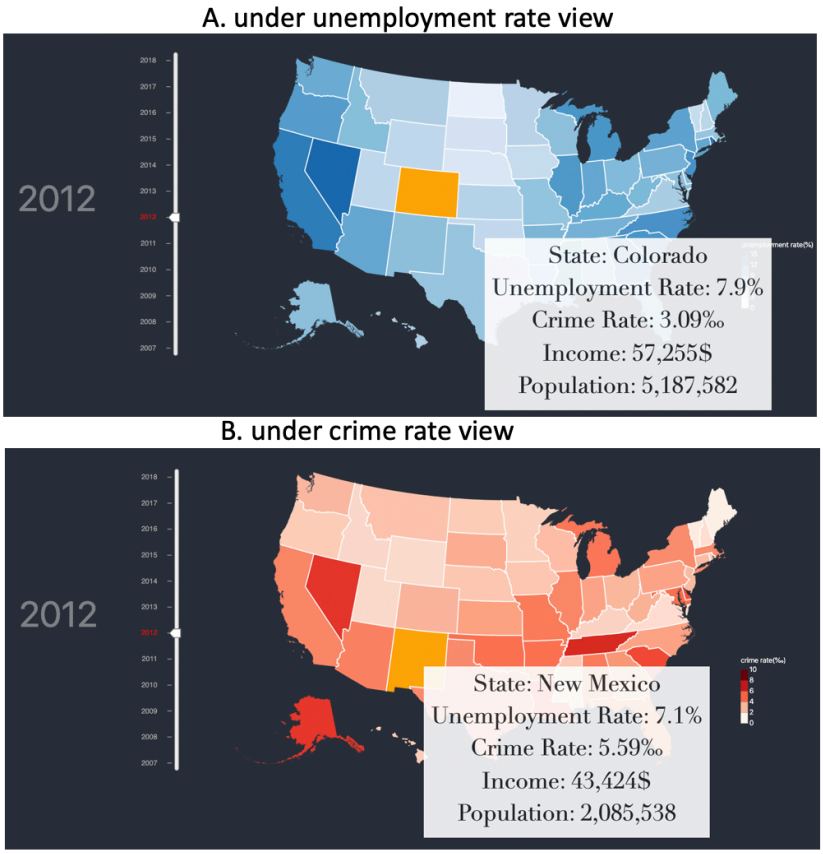


Figure 10. map chart of unemployment rate and and crime rate of the US.

The line chart

The change trend of unemployment or crime rate for specific area. Each line represents a state trend of the related rate over the 12 years. The line corresponding to seleted state would be hightlighted in red. Selected year would be shown in dotted line.

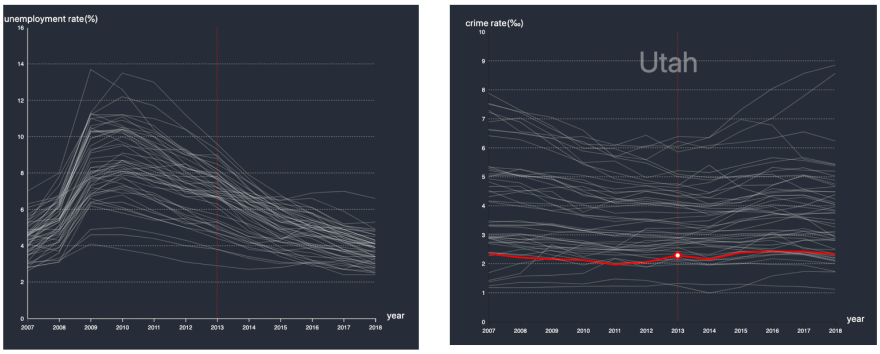
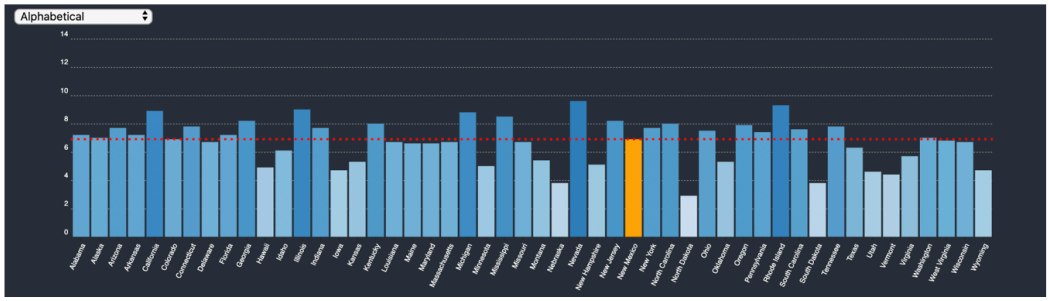


Figure 9. trend of unemployment and crime rate at the US from 2007 to 2018 (selected year was shown in dotted line, selected state was shown in red line).

The bar chart

For the bar chart, in unemployment rate view or crime view, when we select specific year with year slider, the bar chart shows the unemployment rate or crime rate for all states. The bars can be sorted by “Alphabetical”, “Frequency, ascending” or “Frequency, descending”. Each bar can be hovered. Transition was added during the sorting process.

A. under unemployment rate view (alphabetical)



B. under crime rate view (frequency, ascending)

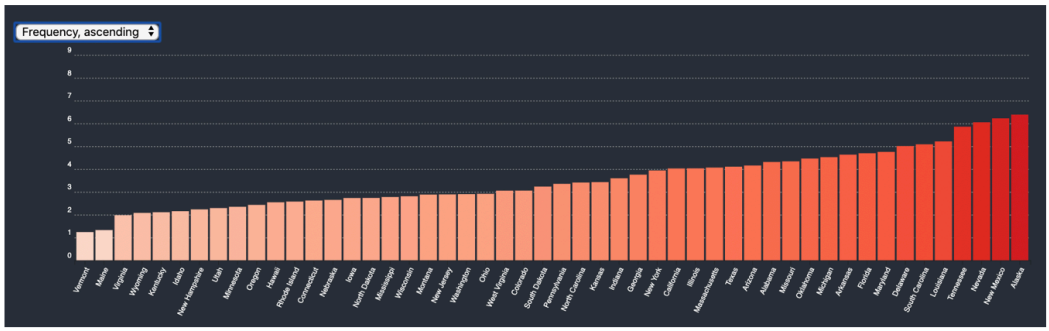


Figure 10. unemployment and crime rate of each state in the US in selected year (selected state is highlighted in orange)

The Bubble Chart:

In the combined view, the default data in map chart and line chart is based on the unemployment rate. The bubble chart is shown in combined view. Each state is represented by the bubble. The color of bubble is based on the unemployment rate. With selection button, the readers could use the date among the four factors to plot the bubble chart with x-axis data, y-axis data and bubble size. When a state is hovered, the related information of the state is shown and the bubble representing the state is highlighted.

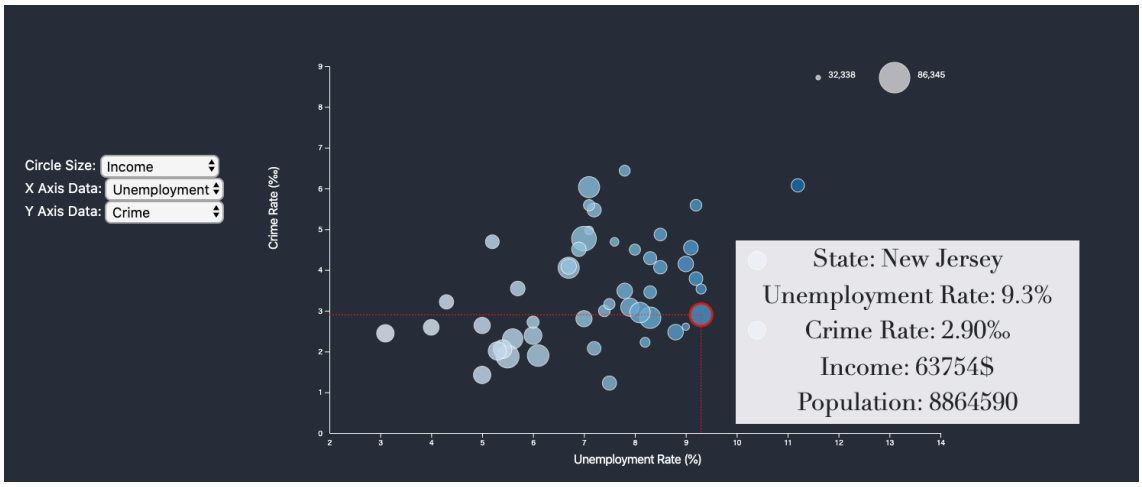


Figure 11. bubble chart shown the correlation between the four factors

Story Telling:

In the storytelling, the data is default to be the unemployment rate. We find there is a sharp increase in unemployment rate among 2008- 2010, which might be caused by the famous 2007-2008 financial crisis. The unemployment rate takes about 5 years to recover the level of 2007 (until 2014). We highlighted it with text together with the red rectangle. The page could be closed with the 'X' button in the top left of the page.

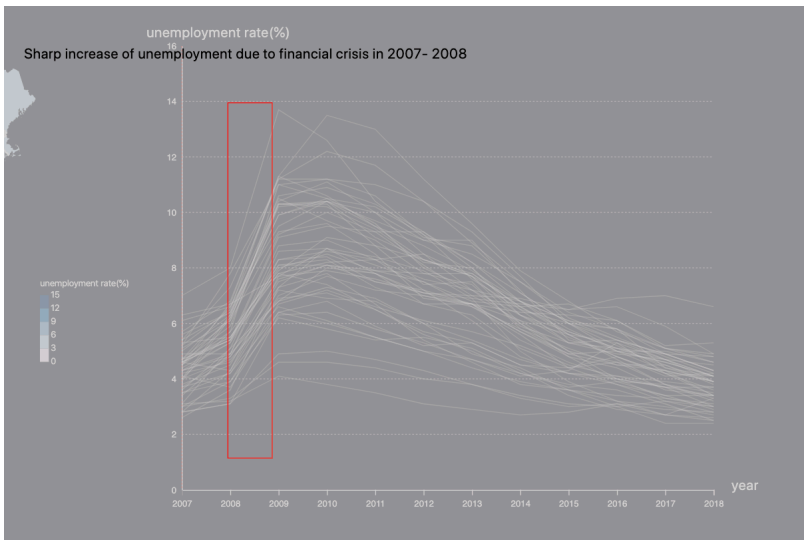


Figure 12. StoryTelling highlights the increase of unemployment rate

Evaluation.

What we learn:

1. Based on the map chart, we could know the unemployment or crime rate difference around the distributions of the US. For example, The crime rate/ unemployment rate of Nevada is very high.
2. Based on the bar chart, the unemployment or crime rate order of specific state could be identified. With sorting, the extremely value (the highest and the lowest) of data are easy to distinguished.
3. Based on the line chart, the change trend of data (unemployment or crime rate) is easy to understand. For example, we find there is a sharp increase in unemployment rate among 2008- 2010, which might be caused by the financial crisis (2007–2008).
4. Based on the bubble chart, correlations between those factors (crime rate, unemployment rate, population and household income) are easily to be seen. However, we could not draw a clear conclusion about the correlation between those factors based on this chart.
5. Storytelling: we find there is a sharp increase in unemployment rate among 2008- 2010, which might be caused by the financial crisis (2007–2008). Accordingly, we highlight it in the 'Tell Story' button.

How did you answer your questions

1. We could easily find the relative unemployment rate by checking the mapchart and bar chart. We could also get each states trendancy of the unemployment rate in the line chart. There is a sharp increase in unemployment rate among 2008- 2010.
2. We could easily find the relative crime rate by checking the mapchart and bar chart. We could also get each states trendancy of the crime rate in the line chart. The crime rate seems to be steady in a level compared with unemployment rate.
3. We didn't find the correlation between the four factors we explored, which is a pity.

How well does your visualization work

1. We are using black as background color to allow readers more focused on our data and charts. When choosing the color for unemployment rate and crime rate, we connect them with sementic meaning, using blue and red respectively.
2. We didn't find the correlation between the four factors we explored, which is a pity.

how could you further improve it?

Our team members meet every week to make sure we are on the same page. After being consistent with design, we would be assigned with different tasks in each meeting and communicate with each other later if one of us get stuck. To improve the visualization, we think we should use much more stuff like transition here to catch reader's eye. In addition, we should connect the data trend with more social events and provide much more for the readers, allowing this visualization project to be more interactive.