MSBX 5405 Team Database Project: Team 15

Project Scenario:

3-5 paragraphs related to the theme of your research project or a description of your 'simulated' organization. Explain why your team is building a database related to this research project or 'simulated' organization.(Samantha Fildish)

The main theme of our project is Air Quality in the US and how that may be impacted by a few different factors, including different types of energy consumption and production. We decided to focus on three types of energy being consumed and produced in the US. Those are coal, hydro energy, natural gas, and liquefied petroleum gas. One of the reasons we decided to pursue this subject is over the past year we have seen numerous Air Quality alerts issued across Colorado. While we know that a lot of those issues were a result of wildfires further West, we thought learning a little more about other factors that could affect Air Quality would be an interesting topic to explore.

We have three main goals with this research project. The first goal is to complete some analysis on each type of energy selected, the consumption and production of that energy, and see if there is any correlation between poor or good AQI where certain energies are being produced and consumed. We will look at each year from 2010 through 2014 and compare the level of energy consumption and production to the AQI for that State. Since the AQI locations are at a county level and the energy data is at the state level, we will take the average AQI across all counties in the State and use that to compare.

The second goal is to compare the Census data with the Air Quality data to see how a higher or lower population estimate may affect the Air Quality Index. This comparison could also provide some insight into whether or not a poor AQI results in lower population or vice versa. As with the Energy data, we would need to use the average AQI across all Counties in a State to compare the census data more accurately with the AQI data. We may also compare other census related data at this point, such as Death Rate to see if there is any correlation between the two.

Our third goal is to just review the data and see what sort of connections we can make between types of energies being produced and consumed, as well as seeing if we can find any interesting correlations between energy,census, and overall AQI ratings.

3-5 paragraphs related to the tables, relationships, and/or 'themes' that appear in your database and some of the 'business logic' that would flow through your database.(Samantha Fildish)

Our database consists of ten tables. Seven tables will reflect the data related to Census and Energy data, and the remaining three will represent the Air Quality data. We have two identifying tables; States and AQI Locations. Each individual State in the Energy and Census tables can map to multiple AQI locations, or counties within the States. Each State/County will have AQI data from years 2010-2014. Each State will also have energy and census data from years 2010-2014.

The States table represents all of the United States, including Washington DC, and contains some geographical information. A few examples of this information are whether or not the state borders the Great Lakes, which region the state is in, and whether or not the state is a coastal state. The AQI Locations table will hold states, counties, latitude and longitude of those counties, and a location id that will map the State/County combination back to a state on the States table. There would be a one-to-many relationship between a State and an AQI Location, and a many-to-one relationship between an AQI Location to a State.

The next theme, set of tables, covers the different types of energy and the state's consumption, production, price, expenditure of that energy type. Each table houses values from the Years 2010 through 2014. The states on the State Codes table will be mapped to a state code on each of these tables. We will then be able to map the different types of energy consumed and produced most in those states and see how that may or may not affect the Air Quality for the counties within those states. Along similar lines we have a table that tracks GDP per state per year between 2010 and 2014.

The other theme we have within our database is the census data. The census table houses census related data from the years 2010 through 2014. This data includes the Population Estimate, Birth Rate, and Death of each state. Each state listed in the States data table is mapped to a state code on the Census data table. This can also be mapped to the AQI Locations table and AQI information. This could be used to see how population estimates may or may not affect the AQI. We could also consider comparing AQI to death rates and see if at all how the AQI may affect it.

3-5 paragraphs describing some of the relevant, important columns/fields/data used in your team's database. Briefly describe any 'transformations' that were made to data from its original form. (Danielle Allen)

Two of the most important fields for the structure and utilization of our database are the locationID and stateCode fields. One of these 2 fields is present in each table in our database. These fields allowed us to bring together the AQI data set and the US energy and census data set from different sources that did not share common fields. These fields also facilitate the necessary joins between tables in our database for analysis and visualization creation in both SQl and Tableau.

Two additional fields that are also important for our data set and analysis are the year and state fields. Throughout our business questions you can see that these two fields are frequently included. We used them in many of our select statements as well as for creating our visualizations that represented change over location and time. Being able to segment our data by both state and year made it possible to analyze trends over time as well as geographical trends. Specifically for looking at the different relationships between the consumption and cost of each energy type, AQI, and birth rates the year and state fields were essential to dig deeper than just generalized relationships over the four years.

The transformations that we made to our data were the addition of state codes and the removal of data that was not relevant to our scenario. For any tables that did not already include either

the state code or location id we added those fields prior to inserting into our database so that we could connect each of our tables. We also removed unnecessary data from our set that was not relevant to our project scenario. Specifically the energy and census data set came with a vast amount of data that was useful for our scenario.

3-5 paragraphs describing how you normalized your database (i.e., normalized each table) (Rachel Raifsnider)

In order to create the best database for our project, we normalized the data and tables themselves which allowed better focus on the subject at hand, air quality of the United States. We also incorporated more data from other sources to create a more complete picture regarding what does and does not affect overall air quality.

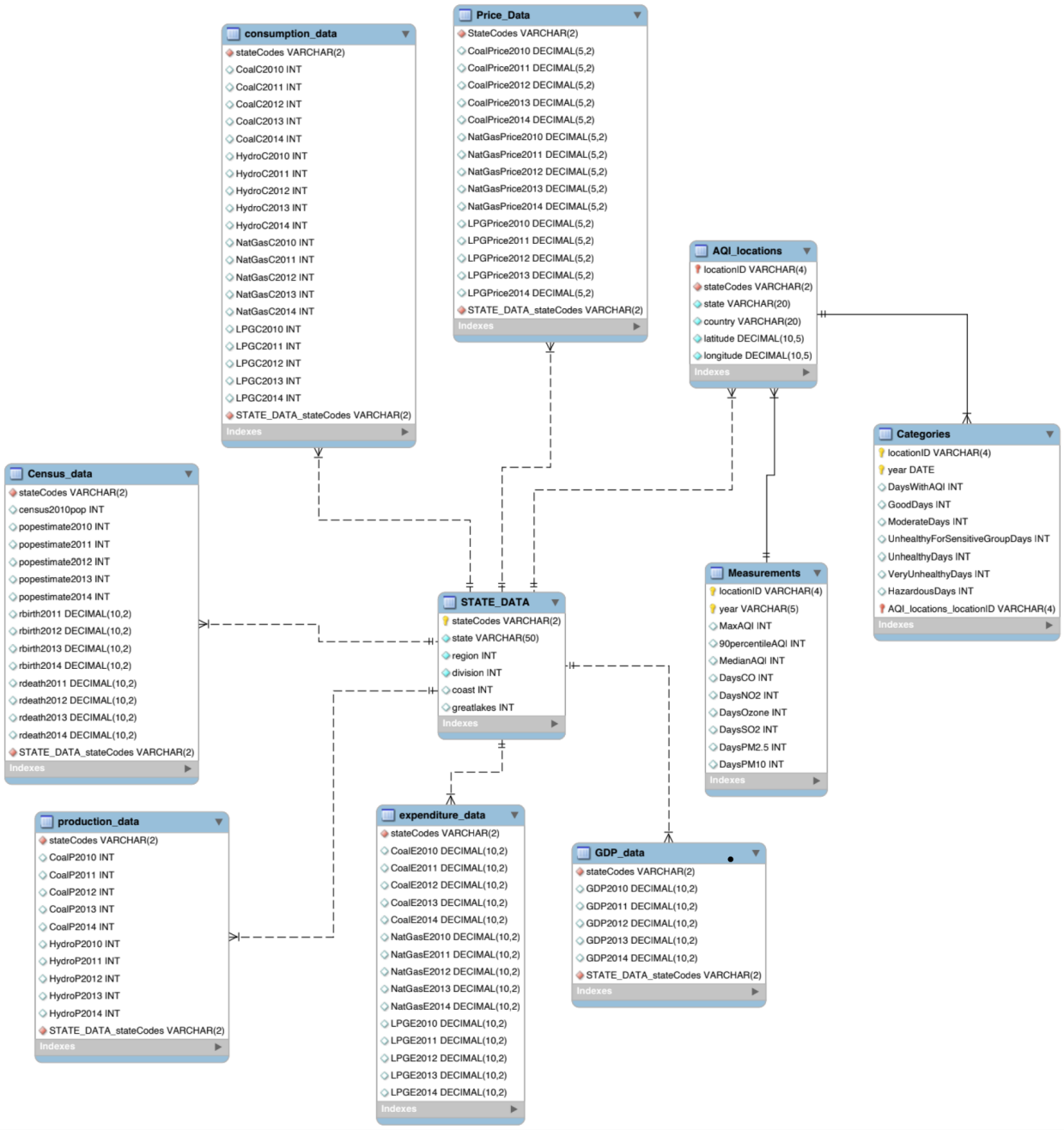
To begin, since we chose to use the AQI and the Census/ Energy dataset, we had to make changes to the years of each dataset so that they were cohesive and then connected the data between the two datasets. This required only using the 2010-2014 data from the AQI dataset as the Census and Energy dataset only included those four years. After that, we chose to focus on the yearly AQI data rather than the daily since that appeared to fit better with our project's required tasks and would likely give us more meaningful and interesting results in regards to the project. We also did not feel that it would have a negative impact on our analysis given that most of our other variables were also yearly values. Other modifications that were made to the data pertained to removing state totals so we would be able to create our own calculations. In addition, we also removed certain types of energy data completely such as geothermal and electric. This was specifically for the project in order to have more focus, in a different setting it would be good to include said data. GDP's quarterly data was removed to maintain on trend with our yearly information trend, as well as the death rate data after 2014.

Once we chose the relevant datasets and customized what data to include, we then organized the data into related tables. The data was broken into 10 separate tables each with data that pertained to similar topics which removed redundancy and made it easier to analyze each subject. This resulted in a table that only contained: state information, energy production, energy consumption, energy price, energy expenditure, state GDP, state census information , county location info regarding AQI, air quality levels and air quality measurements. Then, we established relationships between the tables with the state codes, state names and location ids. The energy, census and economic tables were connected with the state codes while relating the AQI census, locations, categories and measurements required a little further construction. The result was matching the state names to state codes on the state data table and creating a location id for each individual AQI location. After that, we connected the state name column to the location id column in the AQI location table.

The normalizations detailed above pertaining to the datasets, data and tables helped ensure the quality of our database and focus on our project theme and effectively answer our business questions. It also reduced redundancy and allowed us to establish relationships between these datasets to better maneuver through the database and access the information we needed.

MSBX 5405 Team Database Project: Team 15

ER Model:

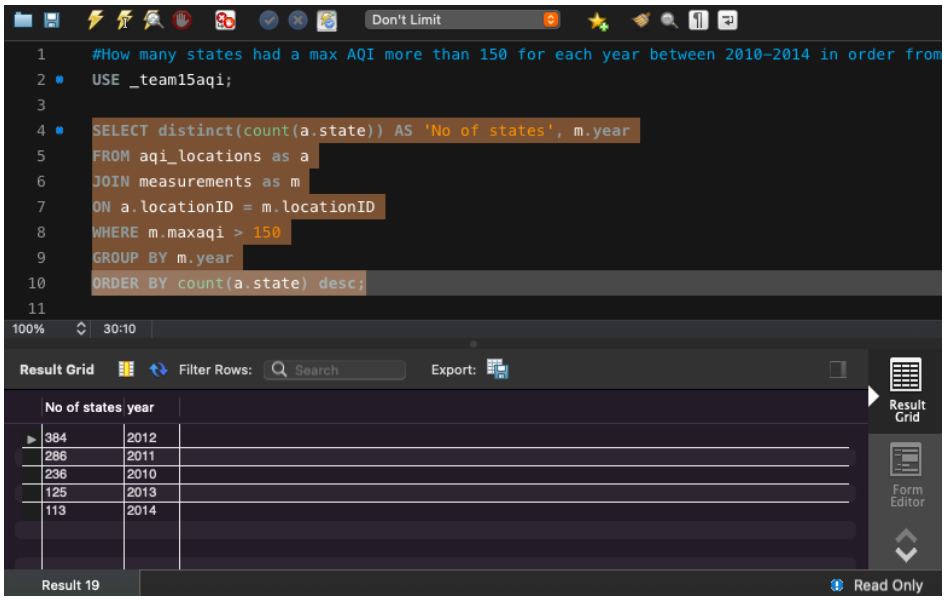


Business Questions:

MSBX 5405 Team Database Project: Team 15

Bella Franz-

Business Question 1: How many states had a max AQI more than 150 for each year between 2010 and 2014 in order from highest to lowest number of states?



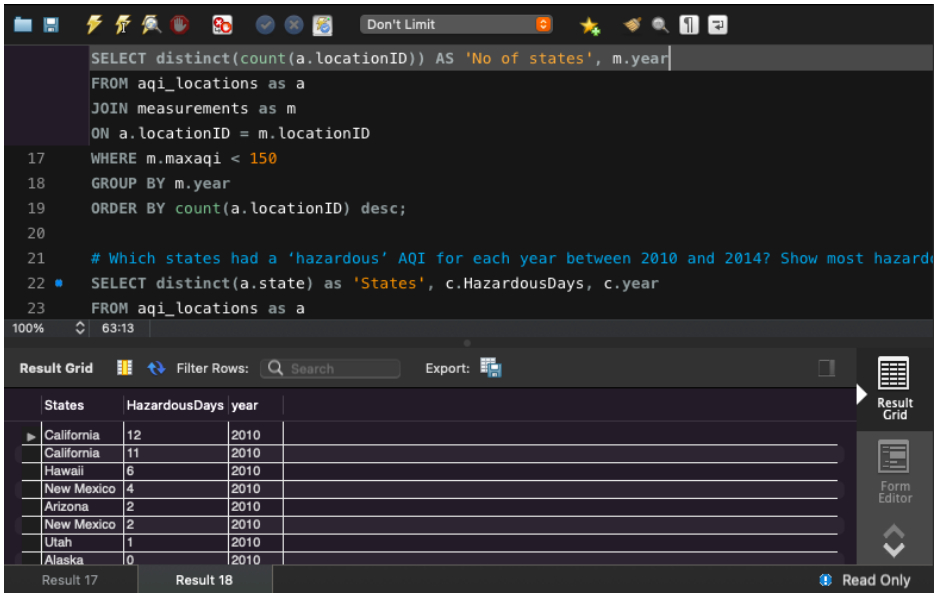
```
#How many states had a max AQI more than 150 for each year between 2010-2014 in order from
USE _team15aqi;

SELECT distinct(count(a.state)) AS 'No of states', m.year
FROM aqi_locations as a
JOIN measurements as m
ON a.locationID = m.locationID
WHERE m.maxaqi > 150
GROUP BY m.year
ORDER BY count(a.state) desc;
```

Result Grid

No of states	year
384	2012
286	2011
236	2010
125	2013
113	2014

Result 19 — Read Only

Business Question 2: How many states had a max AQI less than 150 for each year between 2010 and 2014 in order from highest to lowest number of states?



```
SELECT distinct(count(a.locationID)) AS 'No of states', m.year
FROM aqi_locations as a
JOIN measurements as m
ON a.locationID = m.locationID
WHERE m.maxaqi < 150
GROUP BY m.year
ORDER BY count(a.locationID) desc;

# Which states had a 'hazardous' AQI for each year between 2010 and 2014? Show most hazard
SELECT distinct(a.state) as 'States', c.HazardousDays, c.year
FROM aqi_locations as a
```

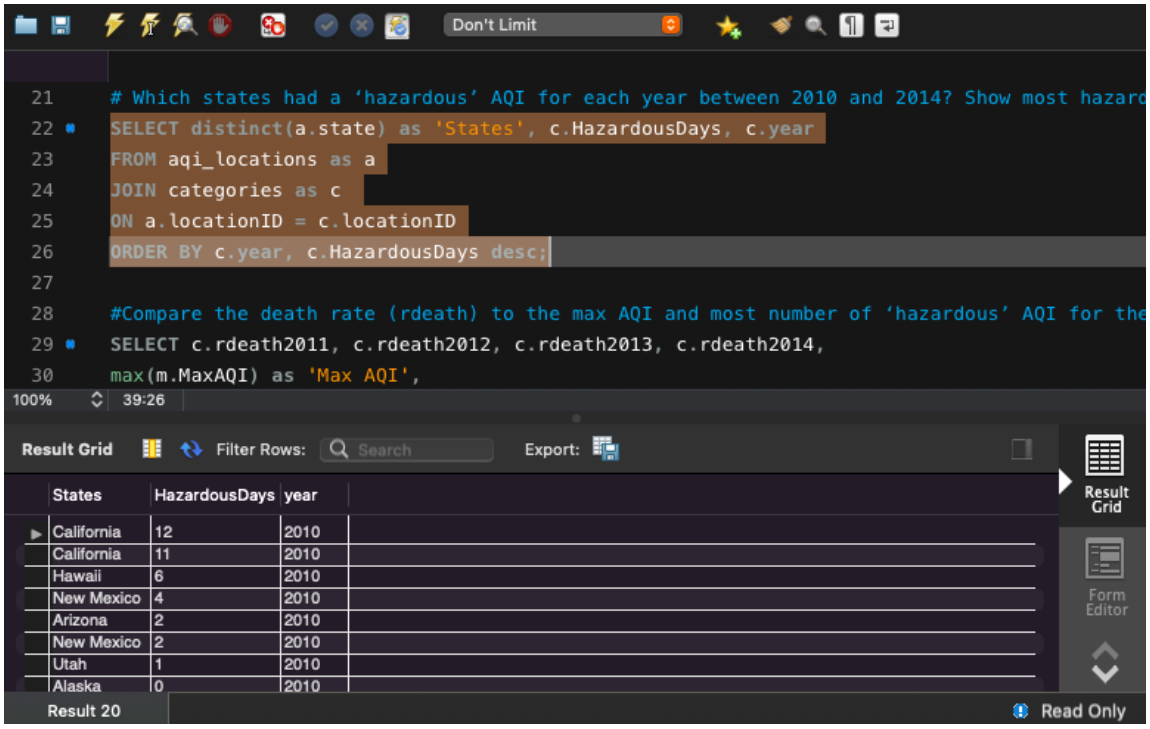
Result Grid

States	HazardousDays	year
California	12	2010
California	11	2010
Hawaii	6	2010
New Mexico	4	2010
Arizona	2	2010
New Mexico	2	2010
Utah	1	2010
Alaska	0	2010

Result 17 | Result 18 — Read Only

Business Question 3: Which states had a 'hazardous' AQI for each year between 2010 and 2014? Show the most hazardous by year.

MSBX 5405 Team Database Project: Team 15



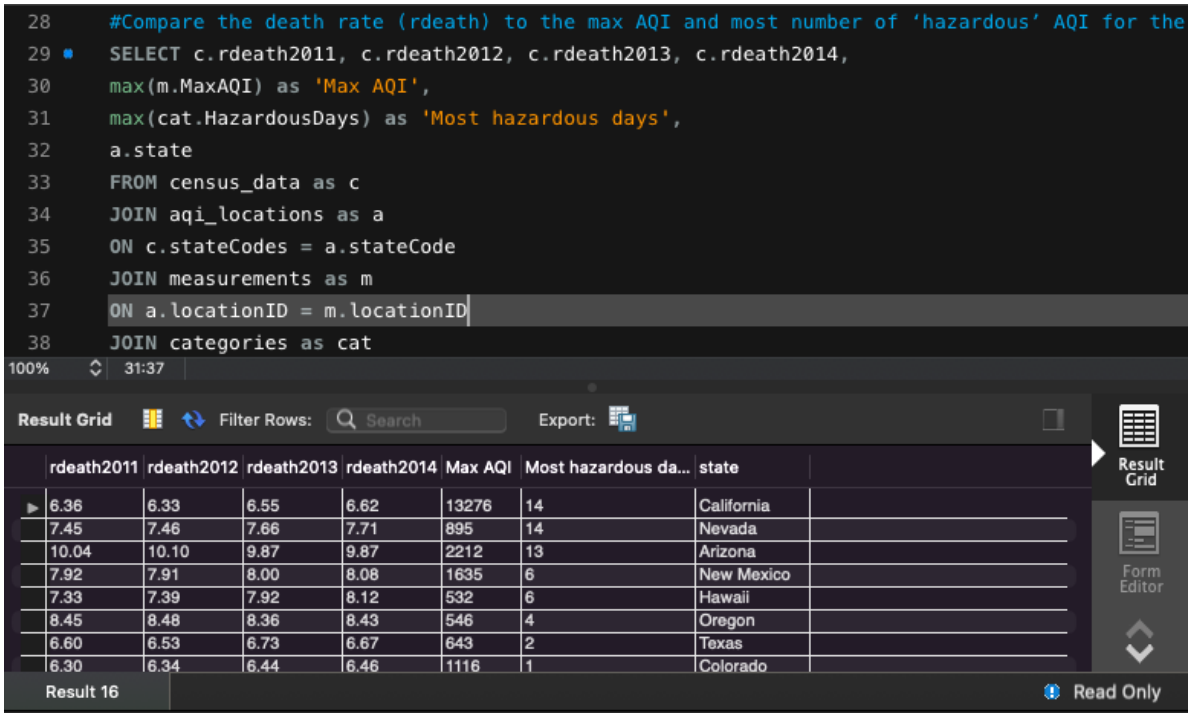
```
21  # Which states had a 'hazardous' AQI for each year between 2010 and 2014? Show most hazard
22  SELECT distinct(a.state) as 'States', c.HazardousDays, c.year
23  FROM aqi_locations as a
24  JOIN categories as c
25  ON a.locationID = c.locationID
26  ORDER BY c.year, c.HazardousDays desc;
27
28  #Compare the death rate (rdeath) to the max AQI and most number of 'hazardous' AQI for the
29  SELECT c.rdeath2011, c.rdeath2012, c.rdeath2013, c.rdeath2014,
30  max(m.MaxAQI) as 'Max AQI',
```

Result Grid | Filter Rows: | Export:

States	HazardousDays	year
California	12	2010
California	11	2010
Hawaii	6	2010
New Mexico	4	2010
Arizona	2	2010
New Mexico	2	2010
Utah	1	2010
Alaska	0	2010

Result 20 — Read Only

Business Question 4 : Compare the death rate (rdeath) to the max AQI and most number of 'hazardous' AQI for the years 2010-2014 by state and death rate.



```
28  #Compare the death rate (rdeath) to the max AQI and most number of 'hazardous' AQI for the
29  SELECT c.rdeath2011, c.rdeath2012, c.rdeath2013, c.rdeath2014,
30  max(m.MaxAQI) as 'Max AQI',
31  max(cat.HazardousDays) as 'Most hazardous days',
32  a.state
33  FROM census_data as c
34  JOIN aqi_locations as a
35  ON c.stateCodes = a.stateCode
36  JOIN measurements as m
37  ON a.locationID = m.locationID
38  JOIN categories as cat
```

Result Grid | Filter Rows: | Export:

rdeath2011	rdeath2012	rdeath2013	rdeath2014	Max AQI	Most hazardous da...	state
6.36	6.33	6.55	6.62	13276	14	California
7.45	7.46	7.66	7.71	895	14	Nevada
10.04	10.10	9.87	9.87	2212	13	Arizona
7.92	7.91	8.00	8.08	1635	6	New Mexico
7.33	7.39	7.92	8.12	532	6	Hawaii
8.45	8.48	8.36	8.43	546	4	Oregon
6.60	6.53	6.73	6.67	643	2	Texas
6.30	6.34	6.44	6.46	1116	1	Colorado

Result 16 — Read Only

Business Question 5: Which states were the least energy efficient between 2010 and 2014? Look at the production and consumption tables and show max/ min for each state for coal, hydro, natgas, and LPG.

```
43    # Which states were the least energy efficient between 2010 and 2014? Look at the producti
44    SELECT s.state,
45    min(c.CoalC2010), min(c.CoalC2011), min(c.CoalC2012), min(c.CoalC2013), min(c.CoalC2014),
46    min(c.HydroC2010), min(c.HydroC2011), min(c.HydroC2012), min(c.HydroC2013), min(c.HydroC20
47    min(c.NatGasC2010), min(c.NatGasC2011), min(c.NatGasC2012), min(c.NatGasC2013), min(c.NatGa
48    min(c.LPGC2010), min(c.LPGC2011), min(c.LPGC2012), min(c.LPGC2013), min(c.LPGC2014),
49    max(p.CoalP2010), max(p.CoalP2011), max(p.CoalP2012), max(p.CoalP2013), max(p.CoalP2014),
50    max(p.HydroP2010), max(p.HydroP2011), max(p.HydroP2012), max(p.HydroP2013), max(p.HydroP20
51    FROM consumption_data as c
52    JOIN production_data as p
53    ON c.StateCodes = p.stateCodes
```

100% 90:45

Result Grid | Filter Rows: Search | Export:

state	min(c.CoalC2010)	min(c.CoalC2011)	min(c.CoalC2012)	min(c.CoalC2013)	min(c.CoalC2014)	min(c.HydroC201...	mi
Texas	1568060	1695239	1498818	1597357	1585961	12310	54
Indiana	1449371	1333442	1193465	1198611	1221492	4426	39
Ohio	1355137	1222606	1019140	1104486	1057377	4186	37
Pennsylvania	1310670	1212984	1093178	1126118	1039161	22753	31
Illinois	1069037	1052224	969294	1026925	1017855	1157	13
Kentucky	1009823	1010623	909705	914774	913452	25173	28
West Virginia	848129	822645	756716	771188	816460	13340	14
Missouri	801625	825650	768297	806549	780707	15018	11

Result Grid | Form Editor

Result 15 | Read Only

Danielle Allen-

Business Question 1: Which state had the most expensive coal prices in 2012?

In this query I used the state field alongside the CoalPrice2012 field grouped by state and ordered by coal price to determine which state had the highest prices for coal of all states in 2012.

state	MAX(CoalPrice2012)
Alaska	4.69

Business Question 2: How does coal consumption relate to birth rates year over year by state?
In this query I used each of the coal consumption fields for 2011- 2014 alongside each of the birth rate fields for 2011-2014 to assess if the rates of consumption affected the rates of borth by state for each of the years.

state	CoalC2011	CoalC2012	CoalC2013	CoalC2014	rbirth2011	rbirth2012	rbirth2013	rbirth2014
Alabama	651032	547004	565051	575912	12.45	12.28	12.03	11.99
Alaska	15481	15521	14819	18225	16.30	15.30	15.24	15.23
Arizona	459909	420570	454865	447849	13.12	13.11	12.93	12.84
Arkansas	306119	296732	327099	339214	13.36	13.14	13.18	13.00
California	55264	43832	38151	39486	13.59	13.13	13.22	13.10
Colorado	368871	370085	363532	350526	12.97	12.56	12.53	12.47
Connecticut	6081	9290	7679	9097	10.50	10.26	10.11	10.14
Delaware	17881	17384	18254	10238	12.50	12.23	11.96	11.97
District of Columbia	48	77	4	48	15.01	14.71	14.94	14.75
Florida	552730	482984	505155	557882	11.23	11.13	10.95	10.87
Georgia	634756	435546	426184	482657	13.58	13.37	13.00	12.88
Hawaii	16080	16572	15306	17241	13.97	13.64	13.52	13.52
Idaho	7846	5186	7965	7472	14.55	14.07	14.27	14.13
Illinois	1052224	969294	1026925	1017855	12.64	12.45	12.26	12.21
Indiana	1333442	1193465	1198611	1221492	12.83	12.73	12.71	12.64
Iowa	463138	422595	402402	401151	12.38	12.53	12.57	12.48
Kansas	346521	307585	326817	316572	14.00	13.75	13.96	13.79
Kentucky	1010623	909705	914774	913452	12.69	12.63	12.65	12.53
Louisiana	269993	238811	228055	209973	13.66	13.56	13.50	13.44
Maine	1542	1299	1657	2142	9.51	9.63	9.59	9.52
Maryland	241190	192334	183191	201220	12.65	12.39	12.32	12.26

Business Question 3: What is the price of coal in the state that had the highest GDP for 2014?
This query uses the state field, coal price for 2014 field, gdp for 2014 field, a select statement in the WHERE clause, and groups by state to determine the price of coal in the state that had the highest GDP for 2014.

state	CoalPrice2014	GDP2014
California	3.43	2324995.50

Business Question 4: How does GDP relate to birth rate year over year per state?

MSBX 5405 Team Database Project: Team 15

In this query I used the yearly GDP columns as well as the yearly birth rate columns grouped and ordered by state to analyze if GDP has an impact on birth rates by state per year.

state	GDP2011	GDP2012	GDP2013	GDP2014	rbirth2011	rbirth2012	rbirth2013	rbirth2014
Alabama	181923.00	187283.25	191605.25	197534.50	12.45	12.28	12.03	11.99
Alaska	59318.25	61613.50	59890.75	58066.75	16.30	15.30	15.24	15.23
Arizona	255620.50	266130.75	271072.50	281558.75	13.12	13.11	12.93	12.84
Arkansas	109378.25	111541.00	116651.50	121064.75	13.36	13.14	13.14	13.00
California	2031347.75	2121602.25	2215231.50	2324995.50	13.59	13.13	13.22	13.10
Colorado	266796.25	277646.75	288809.00	305366.75	12.97	12.56	12.53	12.47
Connecticut	232270.75	238322.25	242417.00	250763.75	10.50	10.26	10.11	10.14
Delaware	59763.25	60774.25	61423.75	65484.75	12.50	12.23	11.96	11.97
District of Columbia	108029.00	109760.00	111816.50	116539.25	15.01	14.71	14.94	14.75
Florida	735244.25	764645.50	797343.50	835578.50	11.23	11.13	10.95	10.87
Georgia	418916.50	434977.75	450933.50	471879.50	13.58	13.37	13.00	12.88
Hawaii	69689.25	72100.25	74011.75	76424.75	13.97	13.64	13.52	13.52
Idaho	57081.00	58544.75	61110.75	63364.00	14.55	14.07	14.27	14.13
Illinois	679269.00	712604.25	720494.50	742027.75	12.64	12.45	12.26	12.21
Indiana	291889.75	299906.00	310669.25	324289.00	12.83	12.73	12.71	12.64
Iowa	149402.00	159095.75	163495.25	170715.00	12.38	12.53	12.57	12.48
Kansas	137654.25	141600.25	142773.50	146562.00	14.00	13.75	13.96	13.79
Kentucky	172255.25	178209.75	182696.25	188518.00	12.69	12.63	12.65	12.53
Louisiana	236958.25	243308.75	238308.25	245790.75	13.66	13.56	13.50	13.44
Maine	51480.75	52606.00	53299.75	55029.00	9.51	9.63	9.59	9.52
Maryland	323942.50	331423.75	337422.00	350262.50	12.65	12.39	12.32	12.26

Business Question 5: In which state and in which year did we see the most Good AQi Days This query uses the state, year, and count of good AQI days grouped by year then state and ordered by the count of good AQI days to identify which state had the highest count of good days and in which year.

state	year	COUNT(GoodDays)
California	2010	53
California	2011	53
California	2012	53
California	2013	53
California	2014	53
Texas	2010	41
Texas	2011	41
Texas	2012	41
Ohio	2010	41
Texas	2013	41
Ohio	2011	41
Texas	2014	41
Ohio	2012	40
Ohio	2013	40
Ohio	2014	40
Florida	2014	39
Florida	2011	39
Florida	2012	39
Pennsyl...	2013	38
Pennsyl...	2014	38

Morgan Likens-

MSBX 5405 Team Database Project: Team 15

Business Question 1: Did the top 5 states that consumed the most Nat Gas in 2013 also pay more than the average price for Nat Gas in 2013?

I looked at the consumption and price by state for 2013, limiting my results to the top 5 consumers of Natural Gas. I thought it would be interesting to look at these values compared to the average pricing because most of these states actually pay less even though their consumption is the highest.

Result Grid | Filter Rows: Search | Export:

STATECODES	NATGASC2013	NATGASPRICE2013	AVG 2013 NAT GAS PRICE
TX	4143328	4.46	7.681961
CA	2483208	6.51	7.681961
LA	1499625	4.19	7.681961
NY	1315282	8.22	7.681961
FL	1245287	5.59	7.681961

Business Question 2: Did the state with the most coal production in 2011 also have the highest days with SO2?

I reviewed the sum of days with SO2 by state and pulled in the coal production for 2011. I was curious to see if the state with the most coal production also showed the leading emissions from sulfur dioxide. Surprisingly, the 7th highest producer has the most days with SO2, which would make for an interesting look into state regulations.

Result Grid | Filter Rows: Search | Export:

STATECODES	2011 DAYS SO2	COALP2011
WY	400	7591675
WV	803	3321102
KY	457	2623804
PA	576	1511491
IL	1184	864228
IN	844	840986
MT	177	746650
OH	1688	679187
TX	75	605290
CO	7	586847

Business Question 3: What years had the highest Max AQI values overall regardless of location?

MSBX 5405 Team Database Project: Team 15

I was curious to see if we are creating more pollution as time went on, or if we have been decreasing the impacts on the environment. I did a simple sum of the MAXAQI and grouped it by year to review the trend.

Result Grid Filter Rows: Search

YEAR	YEARLY AQI
2011	145174
2012	143426
2010	131844
2013	116128
2014	111770

Business Question 4: Which division produced the most Coal, and which division produced the most Hydro in 2012? (with the averages for reference)

I was interested in looking at the ways the US was broken up into divisions which also corresponds to the census data, and looked at division first which covers 1-9. Because these are smaller groups I was surprised to see that the main coal producers from above don't necessarily dictate their division as one of the highest producers. The query groups by division and I have manually filtered the table depending on which product I want to review.

Result Grid Filter Rows: Search Export:

DIVISION	SUM(COALP...	AVG COAL	SUM(HYDROP2012)	AVG HYDRO
8	9221073	197452.3333	34741	2125.0000
5	3606576	197452.3333	11667	2125.0000
6	2740269	197452.3333	18093	2125.0000
3	2563195	197452.3333	3696	2125.0000
2	1390644	197452.3333	26905	2125.0000
7	655819	197452.3333	4608	2125.0000
4	376411	197452.3333	11766	2125.0000
9	31332	197452.3333	157401	2125.0000
1	0	197452.3333	7359	2125.0000

Result Grid Filter Rows: Search Export:

DIVISION	SUM(COALP2012)	AVG COAL	SUM(HYDROP...	AVG HYDRO
9	31332	197452.3333	157401	2125.0000
8	9221073	197452.3333	34741	2125.0000
2	1390644	197452.3333	26905	2125.0000
6	2740269	197452.3333	18093	2125.0000
4	376411	197452.3333	11766	2125.0000
5	3606576	197452.3333	11667	2125.0000
1	0	197452.3333	7359	2125.0000
7	655819	197452.3333	4608	2125.0000
3	2563195	197452.3333	3696	2125.0000

Business Question 5: Which region had the highest prices by product in 2014?
Now looking at the larger regions (1-4) to see where prices were highest and how geography affects the pricing of products. I used MAX to pull in the 2014 regions highest prices.

Result Grid Filter Rows: Search Export:

REGION	MAX(COALPRICE2014)	MAX(NATGASPRICE2014)	MAX(LPGPRICE2014)
1	4.89	10.78	38.40
3	3.65	12.11	35.47
4	4.87	41.71	33.38
2	2.88	8.92	31.78

Rachel Raifsnider-

MSBX 5405 Team Database Project: Team 15

Business Question 1: Which states produce hydro and coal?

Discovering which states produced both hydropower and coal is an important question in relation to the air quality. We may use this question to than further analyze the hypothesis that the more energy and types of energy produced will have an impact on issues such as air quality and death rates, as well as, GDP.

```
4 •  SELECT StateCodes
5    FROM production_data
6    WHERE (CoalP2010 + CoalP2011 + CoalP2012 + CoalP2013 + CoalP2014 ) > 0 AND
7    (HydroP2010 + HydroP2011 + HydroP2012 + HydroP2013 + HydroP2014) > 0;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Cor

StateCodes
AK
AL
AR
AZ
CO
IL
IN
KS
KY
LA
MD
MO
MT
ND
NM
OH
OK

Business Question 2: Which states produce only hydropower?

This question allows us to gain a better picture into these specific states and maybe conclude connections that are a result of only producing hydropower.

```
10 •  SELECT StateCodes, (CoalP2010 + CoalP2011  + CoalP2012 + CoalP2013 + CoalP2014 ) AS 'TotalCoalP',
11    HydroP2010, HydroP2011, HydroP2012, HydroP2013, HydroP2014
12    FROM production_data
13    WHERE (CoalP2010 + CoalP2011  + CoalP2012 + CoalP2013 + CoalP2014 ) = 0 AND
14    (HydroP2010 + HydroP2011 + HydroP2012 + HydroP2013 + HydroP2014) > 0;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

StateCodes	TotalCoalP	HydroP2010	HydroP2011	HydroP2012	HydroP2013	HydroP2014
CA	0	33431	42557	26837	23755	16531
CT	0	391	567	312	402	434
FL	0	177	182	151	254	211
GA	0	3322	2705	2236	3714	3064
HI	0	70	93	115	78	94
IA	0	948	925	766	749	879
ID	0	9154	13405	10940	8473	9002
MA	0	996	1149	912	992	902
ME	0	3810	3979	3733	3560	3623
MI	0	1251	1357	1215	1419	1600
MN	0	840	746	561	511	548
NC	0	4757	3893	3728	6901	4756
NE	0	1314	1617	1257	1124	1158
NH	0	1478	1605	1289	1427	1381
NJ	0	18	24	11	18	17
NV	0	2157	2191	2440	2682	2389
NY	0	25472	27997	24652	24973	26087

Business Question 3: Which state has the most amount of unhealthy AQI days overall ?

Knowing the total amount of unhealthy days per state gives the most extreme cases to focus on which in turn, aids in discovering any trends/relations with the other data. This may then point to what categories have the most influence on poor AQI.

```
17 •  SELECT stateCode, SUM(UnhealthyDays) AS TotUnhealthy4yrs
18    FROM categories
19    JOIN aqi_locations
20    USING(locationID)
21    GROUP BY stateCode
22    ORDER BY SUM(UnhealthyDays) DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

stateCode	TotUnhealthy4yrs
CA	1586
HI	890
TX	191
AZ	185
UT	174
MD	129
OH	127
PA	127
AK	117
MO	95
NJ	93
WI	90
LA	89
OR	88
IL	87
MI	84

Business Question 4: Which states produced more coal than they consumed in 2014?

This question may be used to show whether production or consumption has a greater effect on other categories. I thought it would be interesting to analyze the relationship between the production or consumption of coal with the air quality of states.

```
25 •  SELECT StateCodes, CoalP2014, CoalC2014, CoalP2014 - CoalC2014 AS DiffCoalPandC
26    FROM consumption_data
27    JOIN production_data
28    USING(StateCodes)
29    WHERE CoalP2014  > CoalC2014
30    GROUP BY StateCodes;
31
```

esult Grid | Filter Rows: | Export: | Wrap Cell Content:

StateCodes	CoalP2014	CoalC2014	DiffCoalPandC
AK	22944	18225	4719
CO	528242	350526	177716
IL	1293541	1017855	275686
KY	1869335	913452	955883
MT	790673	175357	615316
NM	400158	215348	184810
PA	1566449	1039161	527288
UT	410997	344054	66943
VA	393225	278183	115042
WV	2858022	816460	2041562
WY	6880205	489300	6390905

Business Question 5: Which states produced more than the average amount of coal production and hydro production?

It is always important to analyze the average because it offers great insights and sets a specific trend to follow. I analyzed the states that were above average in coal and hydropower energy because having a more extreme value would hopefully point to a greater amount of relationships.

```
32    #5 Which states produce more than the average amount of coal production and hydro production?
33 •  SELECT StateCodes, CoalP2010, (SELECT ROUND(AVG(CoalP2010),2) FROM production_data) AS avgCP2010,
34    HydroP2010,(SELECT ROUND(AVG(HydroP2010),2) FROM production_data) AS avgHP2010
35    FROM production_data
36    WHERE (CoalP2010) > (SELECT AVG(CoalP2010) FROM production_data) AND
37    (HydroP2010) > (SELECT AVG(HydroP2010) FROM production_data);
38
```

esult Grid | Filter Rows: | Export: | Wrap Cell Content:

StateCodes	CoalP2010	avgCP2010	HydroP2010	avgHP2010
AL	493094	428064.75	8704	5102.00
MT	797035	428064.75	9415	5102.00

Samantha Fildish-

MSBX 5405 Team Database Project: Team 15

Business Question 1: Do states with a higher average population have a worse average AQI rating over the years 2010-2014?

Answer: It does not necessarily seem that having a higher average population means you will have a higher AQI rating. California has one of the highest average populations, but a lower AQI rating than DC.

stateCodes	Avg Population	Avg AQI	AVG Status
HI	6963515	59.26	Moderate
DC	3168681	51.20	Moderate
CA	190334585	48.58	Good
AZ	14735161	47.22	Good
MD	29438897	44.41	Good
DE	4585295	44.13	Good
PA	63793620	43.34	Good
OK	19094236	43.21	Good

Business Question 2: Does having a higher number of Unhealthy Days result in higher death rates?

Answer: I am not seeing any strong correlation between a higher Death rate and number of Unhealthy Days.

stateCodes	Avg Death Rate	Avg Unhealthy Days
WV	47	0.20
AL	41	0.19
AZ	40	3.08
PA	40	0.51
OK	39	0.64
KY	39	0.43
MS	39	0.08
MF	39	0.08

Business Question 3: Do coastal states consume or produce more hydro energy?

Answer: All coastal states have consumed more energy than they have produced over all years between 2010-2014.

State	2010 Status	2011 Status	2012 Status	2013 Status	2014 Status
Alaska	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Alabama	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
California	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Connecticut	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Delaware	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Florida	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Georgia	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Hawaii	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Louisiana	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Massachusetts	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Maryland	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Maine	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Mississippi	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
North Carolina	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
New Hampshire	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
New Jersey	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
New York	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More
Oregon	Consumed More	Consumed More	Consumed More	Consumed More	Consumed More

Business Question 4: Which States produced the most Hydro Energy overall, and are those states coastal or on the Great Lakes?

Answer: The top five states that produced the most Hydro Energy are Washington, Oregon, California, New York, and Montana. Each of these are either coastal or located on the Great Lakes, except Montana.

State	Hydro Production	Coast	Great_Lakes
Washington	407188	1	0
Oregon	180627	1	0
California	143111	1	0
New York	129181	1	1
Montana	54415	0	0

Business Question 5: Which States consume the most Hydro Energy, overall?

Answer: It does appear that the top five hydro energy producing states, are also the top for consumers.

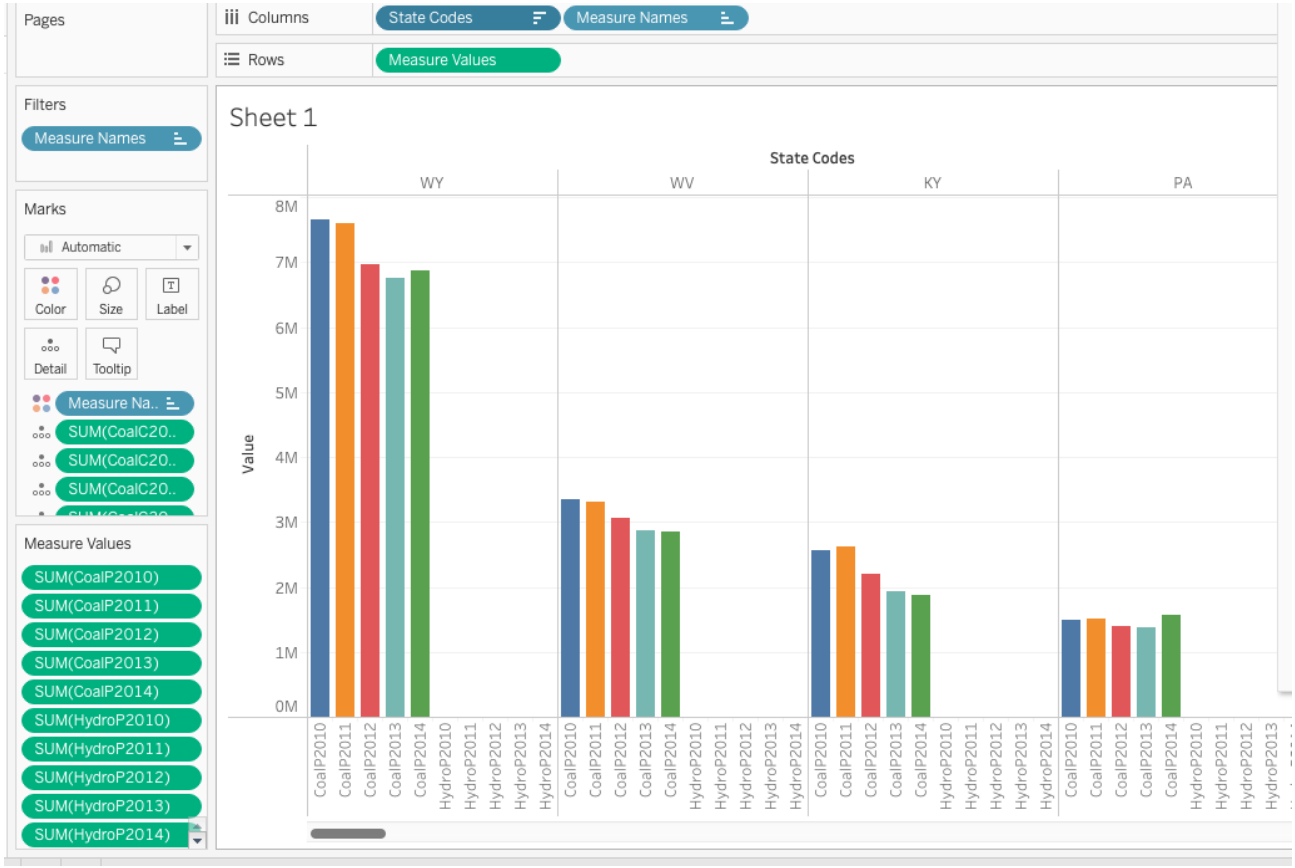
State	Hydro Consumption	Coast	Great_Lakes
Washington	3911038	1	0
Oregon	1735254	1	0
California	1378880	1	0
New York	1241470	1	1
Montana	522764	0	0

Data Visualizations:

MSBX 5405 Team Database Project: Team 15

Bella Franz-

Visualization 1: Show which states produce the most coal in comparison to how much energy they consume between 2010 and 2014. Show visualization by comparing consumption and production of coal, hydro, natgas and LPG.

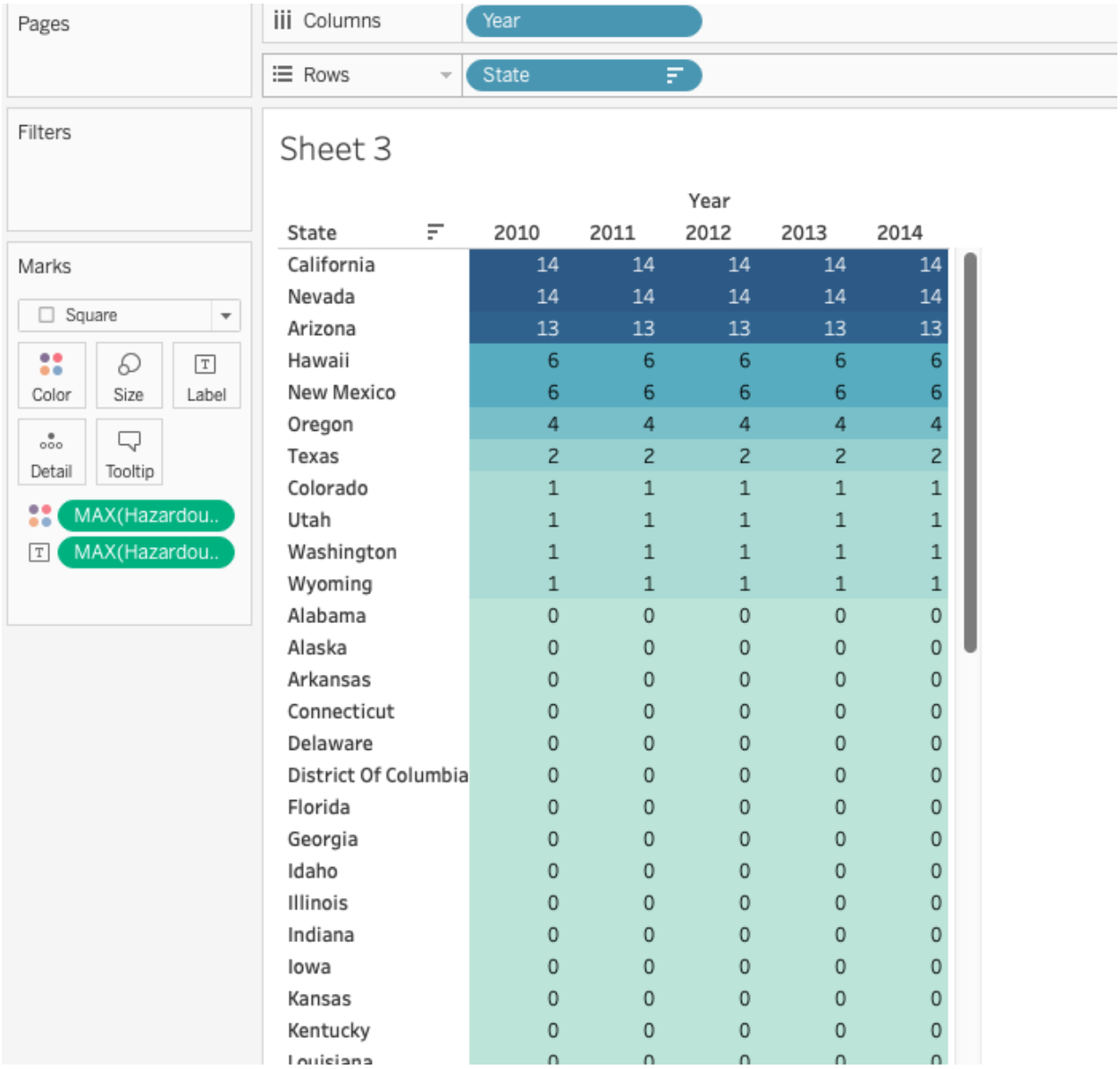


MSBX 5405 Team Database Project: Team 15

Visualization 2: With respect to the death rate, compare which states have the most hazardous days, and the most max aqi days for 2010, 2011, 2013, and 2014.

MSBX 5405 Team Database Project: Team 15

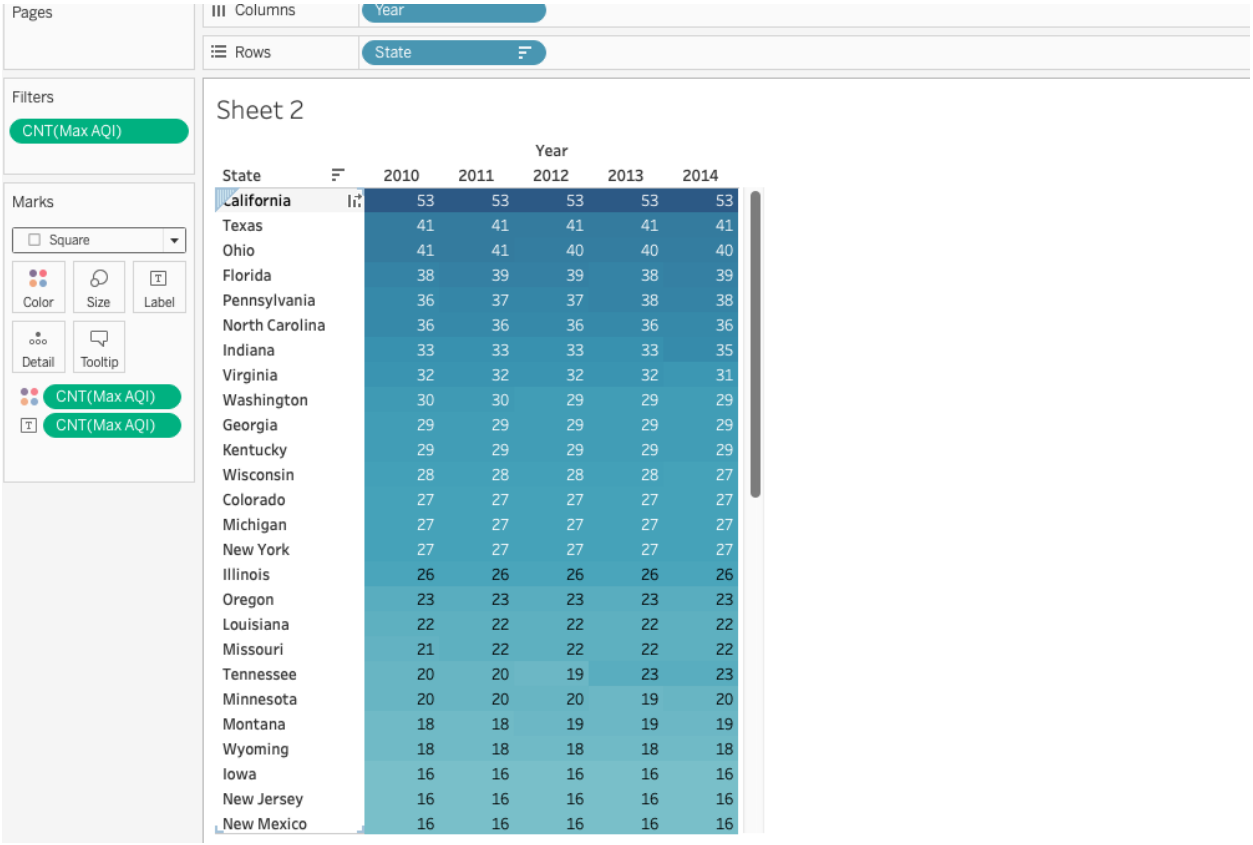
Visualization 3: Show which states had the most number of hazardous days by year.



State	Year 2010	2011	2012	2013	2014
California	14	14	14	14	14
Nevada	14	14	14	14	14
Arizona	13	13	13	13	13
Hawaii	6	6	6	6	6
New Mexico	6	6	6	6	6
Oregon	4	4	4	4	4
Texas	2	2	2	2	2
Colorado	1	1	1	1	1
Utah	1	1	1	1	1
Washington	1	1	1	1	1
Wyoming	1	1	1	1	1
Alabama	0	0	0	0	0
Alaska	0	0	0	0	0
Arkansas	0	0	0	0	0
Connecticut	0	0	0	0	0
Delaware	0	0	0	0	0
District Of Columbia	0	0	0	0	0
Florida	0	0	0	0	0
Georgia	0	0	0	0	0
Idaho	0	0	0	0	0
Illinois	0	0	0	0	0
Indiana	0	0	0	0	0
Iowa	0	0	0	0	0
Kansas	0	0	0	0	0
Kentucky	0	0	0	0	0
Louisiana	0	0	0	0	0

MSBX 5405 Team Database Project: Team 15

Visualization 4: Show which states had a max AQI below 150 for 2011, 2012, 2013, and 2014.



State	Year 2010	2011	2012	2013	2014
California	53	53	53	53	53
Texas	41	41	41	41	41
Ohio	41	41	40	40	40
Florida	38	39	39	38	39
Pennsylvania	36	37	37	38	38
North Carolina	36	36	36	36	36
Indiana	33	33	33	33	35
Virginia	32	32	32	32	31
Washington	30	30	29	29	29
Georgia	29	29	29	29	29
Kentucky	29	29	29	29	29
Wisconsin	28	28	28	28	27
Colorado	27	27	27	27	27
Michigan	27	27	27	27	27
New York	27	27	27	27	27
Illinois	26	26	26	26	26
Oregon	23	23	23	23	23
Louisiana	22	22	22	22	22
Missouri	21	22	22	22	22
Tennessee	20	20	19	23	23
Minnesota	20	20	20	19	20
Montana	18	18	19	19	19
Wyoming	18	18	18	18	18
Iowa	16	16	16	16	16
New Jersey	16	16	16	16	16
New Mexico	16	16	16	16	16

MSBX 5405 Team Database Project: Team 15

Visualization 5: Show which states had a max AQI above 150 for 2011, 2012, 2013, and 2014.

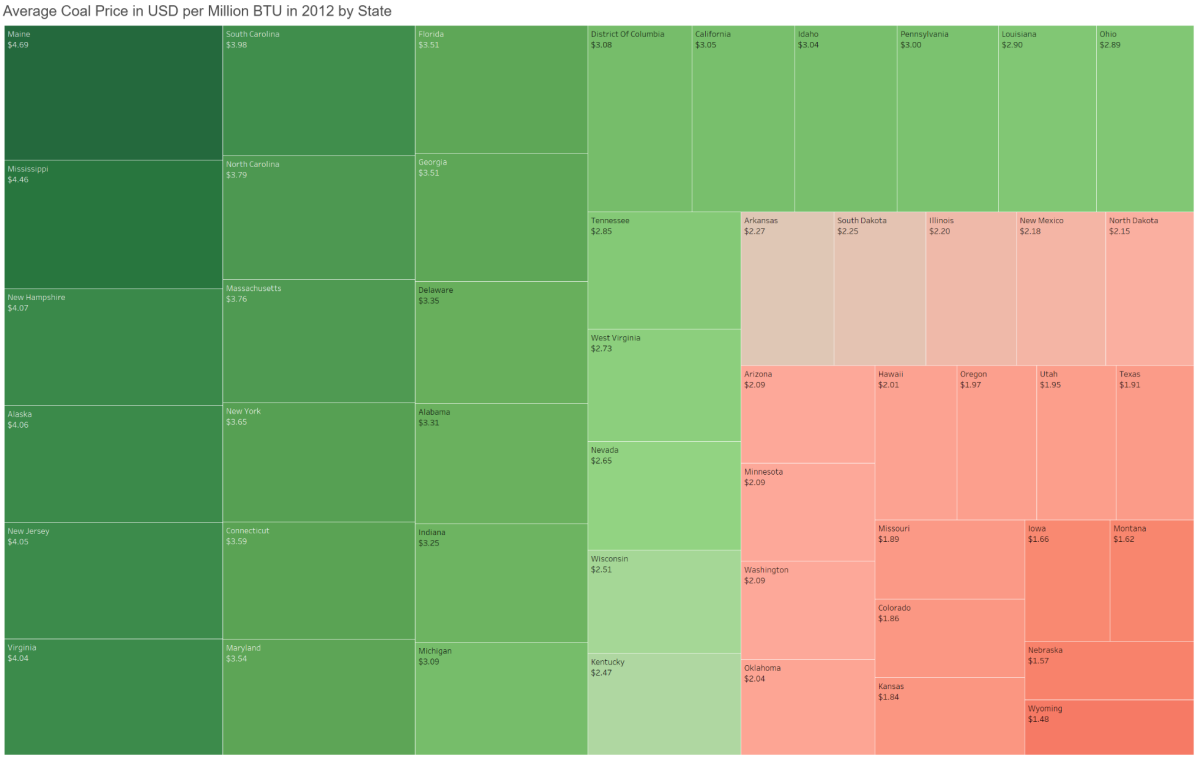
	Year				
State	2010	2011	2012	2013	2014
California	29	26	33	22	29
Ohio	17	14	32	4	4
Texas	12	21	21	11	6
Pennsylvania	21	12	21	4	5
Colorado	8	10	11	11	3
Maryland	14	13	15		1
Indiana	4	13	18	3	4
Michigan	3	14	19	1	1
New Jersey	13	11	8	3	3
Wisconsin	8	9	18	1	1
Georgia	12	10	9	1	2
New York	10	11	9	2	2
Missouri	4	8	16	4	1
Illinois	5	12	12	2	1
Utah	7	7	5	7	4
Idaho	3	2	10	7	5
Kentucky	2	6	18		1
North Carolina	6	6	15		
Oregon	3	5	4	8	7
Connecticut	5	7	6	6	2
Louisiana	6	10	5	3	2
Washington	1	6	8	4	6
Virginia	10	6	6	1	1
Tennessee	3	7	12		1
Arizona	3	3	4	6	5
Massachusetts	6	7	7	1	

MSBX 5405 Team Database Project: Team 15

Danielle Allen-

Visualization 1: Coal prices by state for 2012

This treemap visualization uses the SUM of coal prices as the title, size, and color marks with state acting as another title measure and a filter for 2012 as the year. Green indicates the states with the highest coal prices for 2012 with red indicating the lowest priced coal in 2012.



MSBX 5405 Team Database Project: Team 15

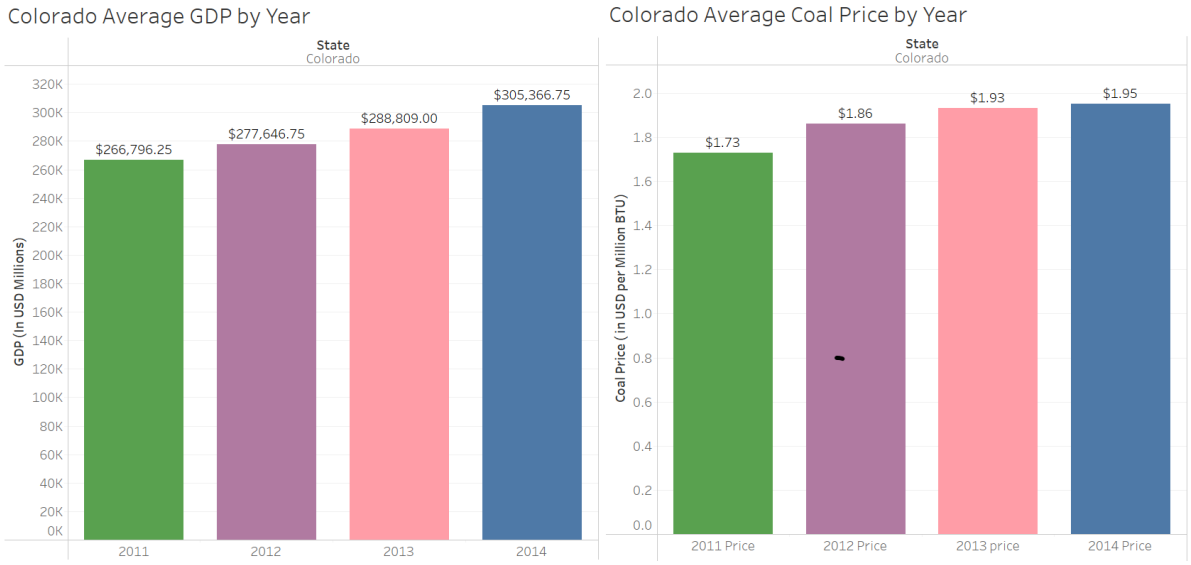
Visualization 2: Coal consumption rates and birth rates year over year by state
This text table visualization has the state as the dimension in rows and the SUM of birth rates and coal consumption for each year between 2011-2014 as the columns and measure values. It is then filtered by the state to remove any states without data.

Coal Consumption Rates and Birth Rates Year Over Year by State

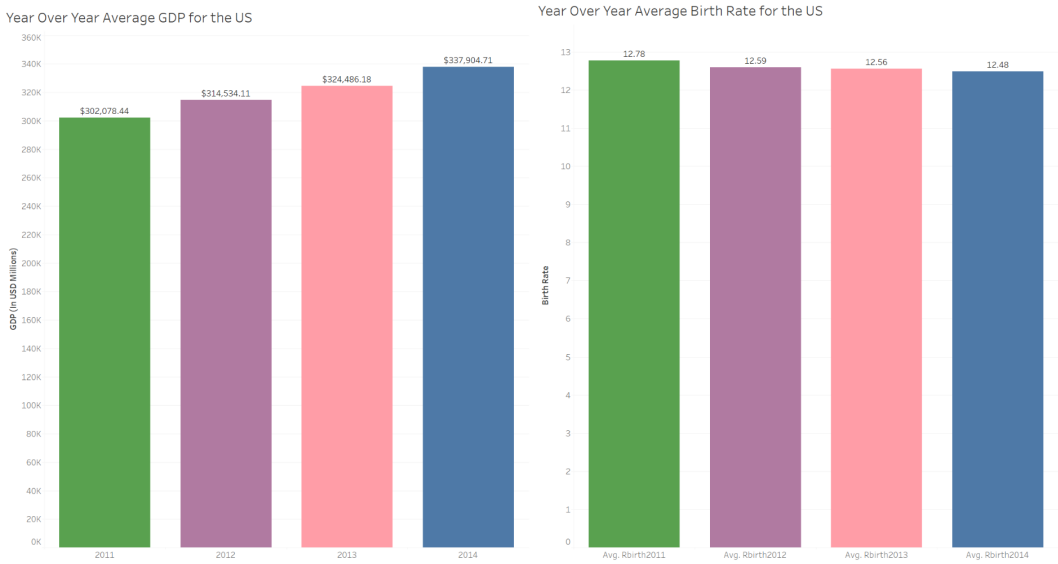
State	2011 Consuption	2011 Birth Rate	2012 Consuption	2012 Birth Rate	2013 Consuption	2013 Birth Rate	2014 Consuption	2014 Birth Rate
Alabama	651,032	12.45181625	547,004	12.28088177	565,051	12.02634207	575,912	11.99148272
Alaska	15,481	16.29737098	15,521	15.29663544	14,819	15.24170151	18,225	15.22533041
Arizona	459,909	13.11919153	420,570	13.11303338	454,865	12.93113735	447,849	12.84325384
Arkansas	306,119	13.36451617	296,732	13.13613071	327,099	13.17647865	339,214	12.99788628
California	55,264	13.58689192	43,832	13.13081487	38,151	13.22215746	39,486	13.10054383
Colorado	368,871	12.96547405	370,085	12.55681835	363,532	12.53426697	350,526	12.47352265
Connecticut	6,081	10.49752283	9,290	10.25790342	7,679	10.11023113	9,097	10.14422143
Delaware	17,881	12.50304278	17,384	12.22879252	18,254	11.95686928	10,238	11.96654869
District Of Colu..	48	15.00607439	77	14.71006406	4	14.94372547	48	14.75071942
Florida	552,730	11.23331539	482,984	11.13002763	505,155	10.95124579	557,882	10.86591025
Georgia	634,756	13.58401017	435,546	13.36596967	426,184	13.00236686	482,657	12.87958821
Hawaii	16,080	13.9719882	16,572	13.63831402	15,306	13.51725152	17,241	13.52283928
Idaho	7,846	14.55228364	5,186	14.06567968	7,965	14.26989437	7,472	14.13170975
Illinois	1,052,224	12.64104635	969,294	12.44582335	1,026,925	12.25912663	1,017,855	12.20978574
Indiana	1,333,442	12.83445023	1,193,465	12.72740588	1,198,611	12.71037648	1,221,492	12.63802093
Iowa	463,138	12.37931914	422,595	12.52923257	402,402	12.56558559	401,151	12.47623384
Kansas	346,521	14.00370123	307,585	13.75450818	326,817	13.95559524	316,572	13.79387161
Kentucky	1,010,623	12.68710702	909,705	12.63402777	914,774	12.6516444	913,452	12.53165764
Louisiana	269,993	13.66017388	238,811	13.56343013	228,055	13.49876782	209,973	13.43879055
Maine	1,542	9.51383483	1,299	9.633648809	1,657	9.594723053	2,142	9.524629804
Maryland	241,190	12.64587643	192,334	12.39266468	183,191	12.31624279	201,220	12.26120305
Massachusetts	42,963	11.10793792	24,036	10.88656333	42,208	10.90050411	29,871	10.88842942
Michigan	691,150	11.54168182	621,633	11.44403256	658,172	11.41092335	618,479	11.38404701
Minnesota	315,604	12.8300033	257,875	12.68565824	267,695	12.7797976	313,073	12.6814087
Mississippi	107,546	13.35893869	82,456	13.22469013	97,767	12.84503081	116,537	12.76451088
Missouri	825,650	12.70015089	768,297	12.53408055	806,549	12.47585168	780,707	12.38600369
Montana	165,687	12.0689898	157,308	12.00704605	166,104	12.06221501	175,357	12.01210924
Nebraska	285,384	14.01814821	272,587	13.93345465	292,956	13.92955105	276,544	13.84412415
Nevada	62,701	13.22112791	52,850	12.6806984	64,838	12.64707065	79,230	12.48642905
New Hampshire	24,489	9.862500408	14,191	9.371805626	16,778	9.396678333	14,867	9.313704953
New Jersey	49,592	12.0914459	25,589	11.79719813	25,910	11.69281668	30,731	11.59012569
New Mexico	284,737	13.41810517	263,439	12.98678525	256,377	12.896594	215,348	12.84851384
New York	125,181	12.49181324	72,914	12.26311458	68,707	12.23830758	64,655	12.20579928
North Carolina	624,783	12.60132872	534,713	12.32638698	493,771	12.25691682	501,592	12.13557541
North Dakota	394,771	13.53499261	406,344	14.25865588	393,193	14.49112701	399,195	14.73342814
Ohio	1,222,606	11.95027366	1,019,140	11.92449247	1,104,486	11.96545106	1,057,377	11.91383918
Oklahoma	378,320	13.93186967	327,089	13.77034468	335,925	13.82627806	336,080	13.73427486
Oregon	35,135	11.78133891	28,259	11.56531117	38,900	11.50157818	34,238	11.44776976
Pennsylvania	1,212,984	11.23288907	1,093,178	11.13990659	1,126,118	11.15284017	1,039,161	11.10991824
Rhode Island	0	10.45272002	0	10.45871132	0	10.29918931	0	10.35604477
South Carolina	366,160	12.34737915	298,572	12.24478284	257,319	12.01531405	305,656	11.95304949
South Dakota	32,112	14.37364778	35,574	14.35121407	34,246	14.47368891	33,062	14.38406768
Tennessee	481,127	12.32597338	423,086	12.4956763	399,768	12.38376134	427,509	12.29007879
Texas	1,695,239	14.99890164	1,498,818	14.61229471	1,597,357	14.64625734	1,585,961	14.46940613
Utah	346,181	18.57784091	322,055	17.79237805	355,230	17.99450189	344,054	17.60168904
Vermont	0	9.664266172	0	9.727061093	0	9.541952748	0	9.58340281
Virginia	288,348	12.68564272	222,270	12.60775988	290,508	12.53564874	278,183	12.48409768
Washington	56,951	12.78293321	42,656	12.64269392	74,955	12.67073908	76,547	12.53912286
West Virginia	822,645	11.03646704	756,716	11.14274128	771,188	11.15175902	816,460	11.05099164
Wisconsin	447,370	11.89308385	373,281	11.83206831	454,554	11.65555051	417,096	11.60556521
Wyoming	467,671	13.19447451	490,128	12.99754308	520,699	13.08662237	489,300	13.06519922

MSBX 5405 Team Database Project: Team 15

Visualization 3: Price of coal and state GDP year over year for Colorado
This visualization uses 2 bar graphs combined into a dashboard to represent the GDP by year and the Average coal price by year for Colorado. The average GDP chart uses the dimension state and the measures of average GDP for each year from 2011-2014 and is then filtered by state to only include Colorado. The average coal price chart uses the dimension state and SUM of coal price for each year as the measures and is then filtered by state to include only colorado.

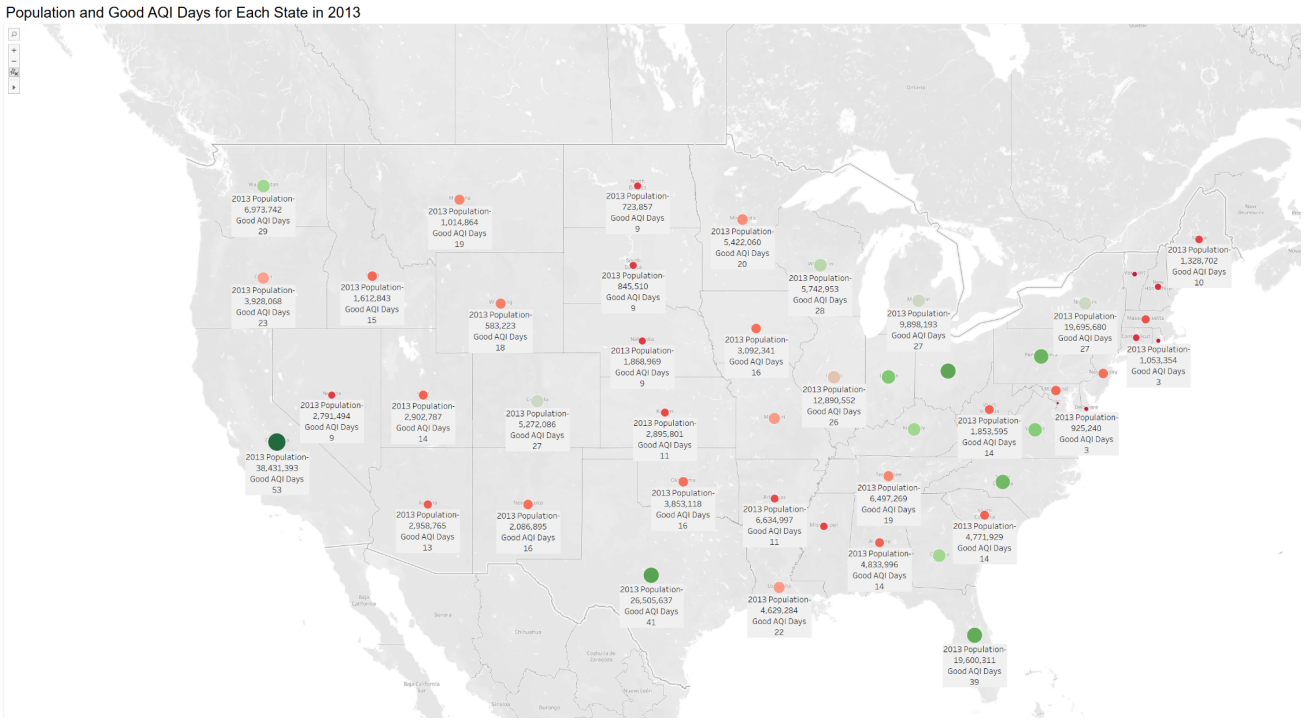


Visualization 4: GDP and birth rate year over year
This visualization uses two bar graphs combined into a dashboard to show GDP year over year and birth rate year over year for the US. The GDP chart uses the average yearly GDP fields as the rows and measures. The Birth Rate chart uses the Average birth rate field for each year as the rows and measures.

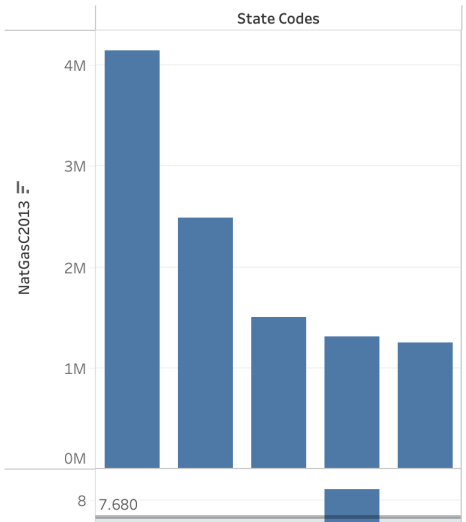


MSBX 5405 Team Database Project: Team 15

Visualization 5: Population and good AQI days for each state in 2013
This symbol map visualization uses longitude as the columns and latitude as the rows with the size, color, and text of the symbols set by the count of good AQI days and uses the sum of the population estimate field as a text label. It is then filtered by year to only represent 2012.

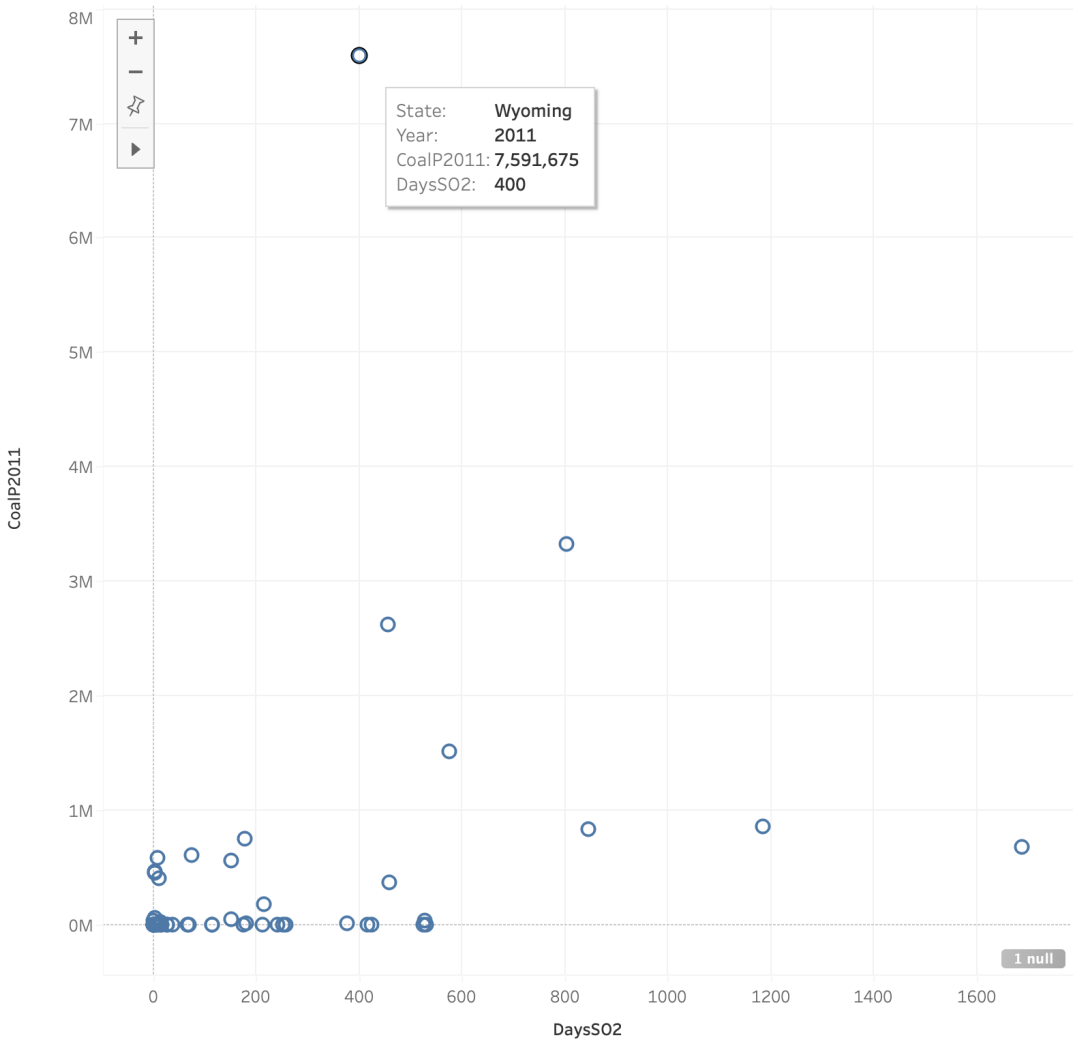


Morgan Likens-Visualization 1: The top 5 states that consumed the most Nat Gas in 2013 with their Nat Gas prices in 2013 compared to a static average line. Joined the Price and Consumption tables, it is a nice view to see how they compare stacked against one another. You can see that while Texas consumed the most gas, they are on the lower end of what the price could have been for the year. The average line is much higher than I would have expected.



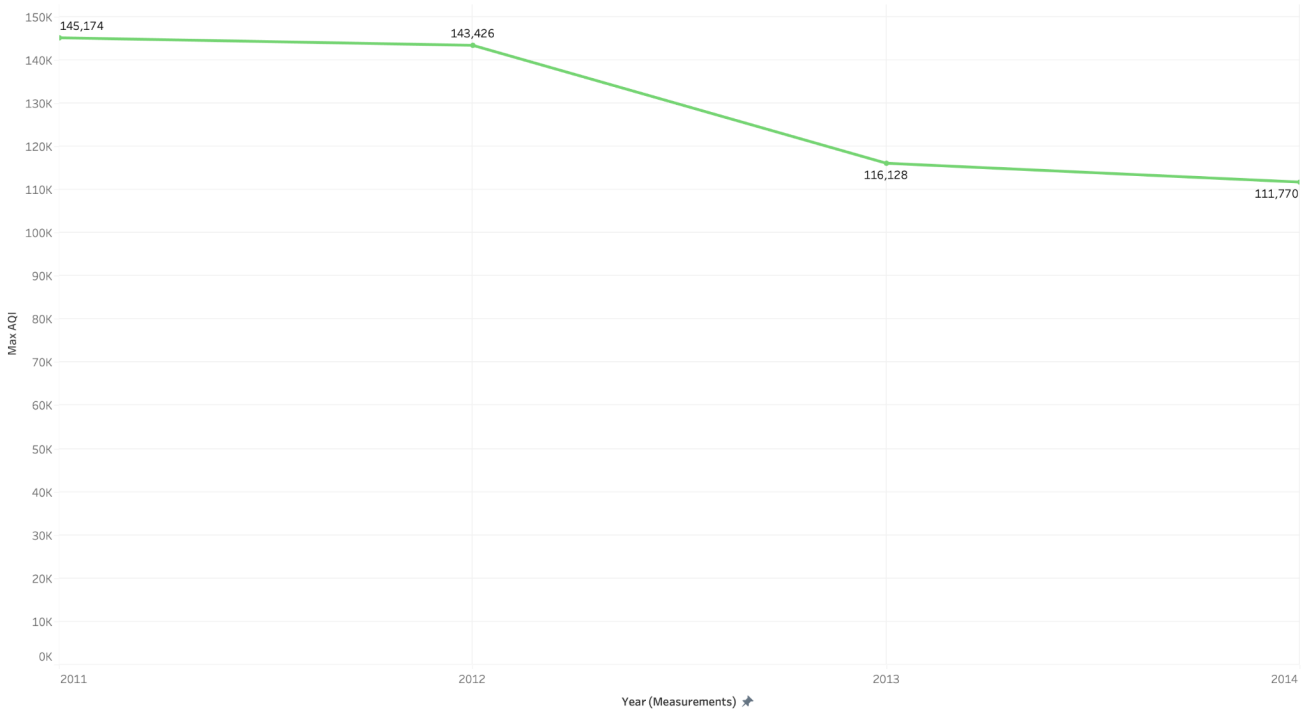
MSBX 5405 Team Database Project: Team 15

Visualization 2: 2011's coal production and SO2 days. Joined the tables on Production and Measurement to visualize the difference in coal production and the days with SO2. It is interesting that WY is so far to the upper left hand corner on this plot given they are the highest producers by a decent margin.



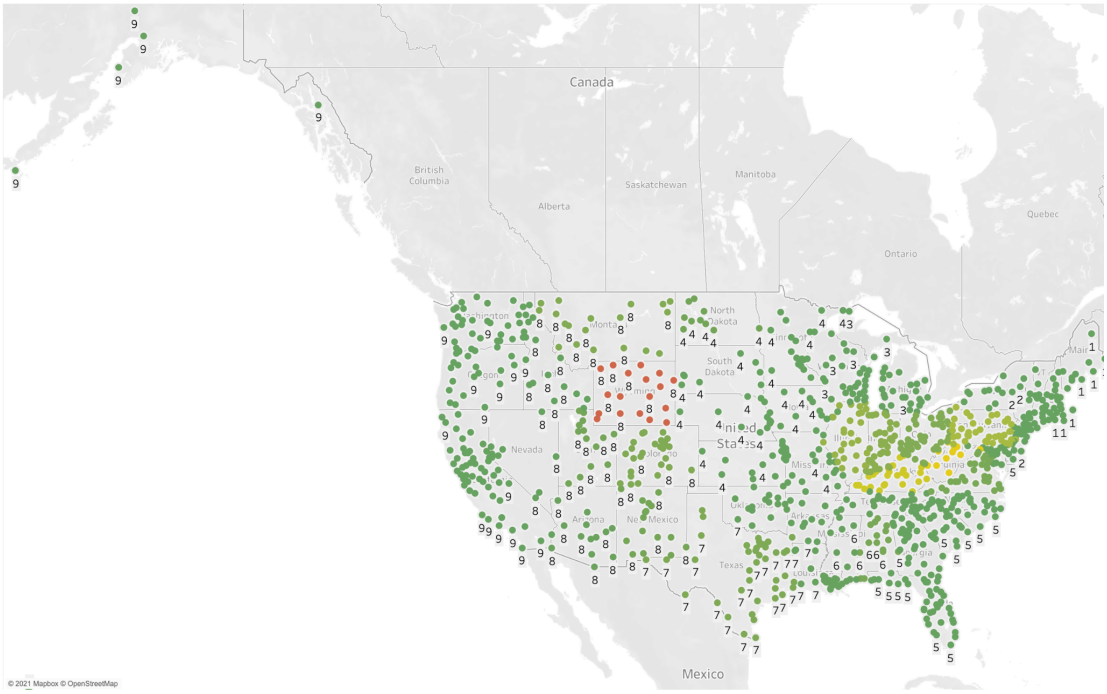
MSBX 5405 Team Database Project: Team 15

Visualization 3: Max AQI volumes by year. No tables were joined to look at this visualization. I was just curious to see the trend down from year to year.

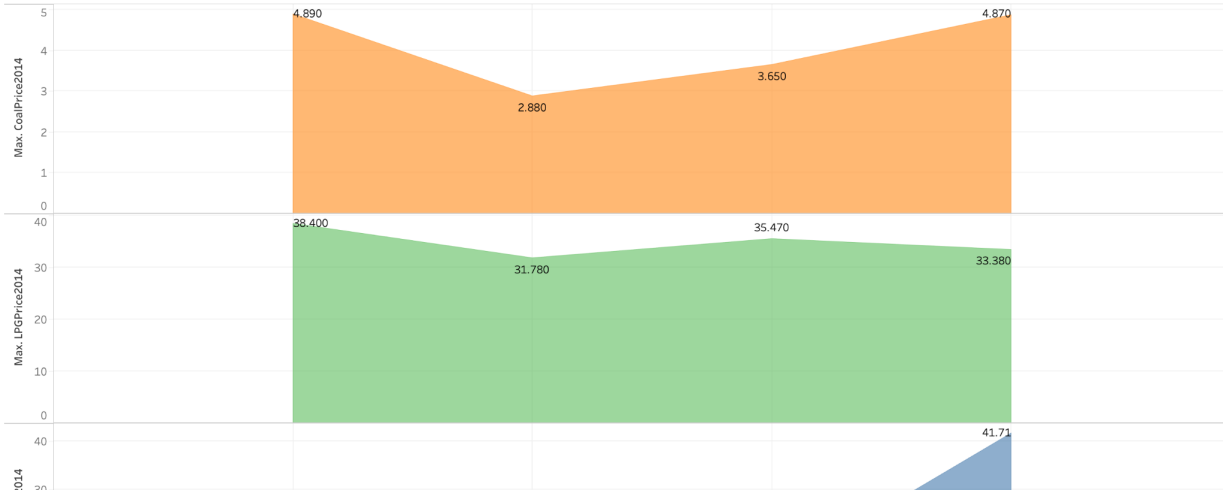


MSBX 5405 Team Database Project: Team 15

Visualization 4: Coal and Hydro production in 2012. I was really interested in doing a map visualization for this since we have latitude and longitude data. I joined State and Production data and included labels by division so you could see how the states were broken up. The red and yellow dots show the highest concentration of coal production, while the greens were a look at hydro locations.



Visualization 5: 2014's product prices. I had these by region just as another way to break up the amount of location data we had. This also used the Price and State tables showing the max price per region per product. Region 4 definitely feels the pricing the worst.



MSBX 5405 Team Database Project: Team 15

Rachel Raifsnider-

Visualization 1: States that produce both hydropower and coal

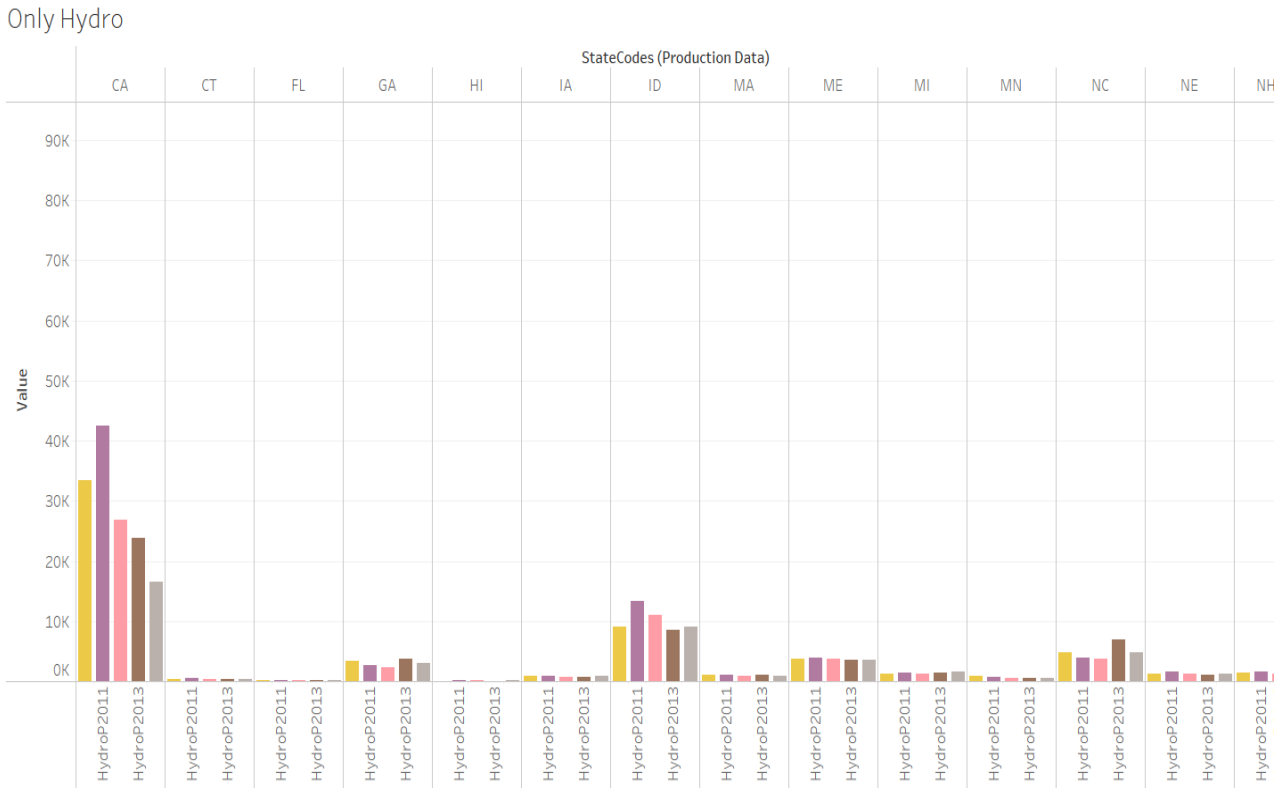
This information was used to gain detail about the states production of energy. In order to create the visual, I incorporated the state codes as dimensions and the coal and hydro production columns for the four years as measures. I then filtered the coal and hydro production columns to be a value greater than 0. I chose the text table because it did the best at displaying the large and small values, since the other graphs made the small values look as if they were valueless.

Hydro and Coal

StateCod..	CoalP20..	CoalP20..	CoalP20..	CoalP20..	CoalP20..	HydroP2..	HydroP2..	HydroP2..	HydroP2..	HydroP2..
AK	33,556	33,524	31,332	24,917	22,944	1,433	1,345	1,575	1,435	1,539
AL	493,094	468,671	488,084	469,162	414,366	8,704	8,884	7,435	12,899	9,467
AR	718	2,985	2,077	1,433	1,864	3,659	2,958	2,198	2,655	2,640
AZ	167,930	174,841	161,374	163,691	173,337	6,622	9,174	6,717	5,915	6,118
CO	551,825	586,847	629,645	529,096	528,242	1,578	2,083	1,497	1,213	1,770
IL	767,419	864,228	1,094,204	1,149,607	1,293,541	119	140	111	120	132
IN	790,889	840,986	826,842	883,260	886,370	454	409	434	387	371
KS	3,069	840	388	541	1,474	13	15	10	15	16
KY	2,556,093	2,623,804	2,193,276	1,940,060	1,869,335	2,580	2,969	2,362	3,275	3,144
LA	54,346	52,440	53,041	38,079	35,541	1,109	1,044	680	1,045	1,090
MD	58,798	65,871	54,088	45,290	46,194	1,667	2,547	1,657	1,727	1,616
MO	9,816	10,075	9,197	9,101	8,206	1,539	1,185	714	1,136	697
MT	797,035	746,650	660,148	753,246	790,673	9,415	12,596	11,283	9,638	11,483
ND	377,669	367,624	366,826	369,479	389,673	2,042	2,580	2,477	1,852	2,531
NM	381,448	405,952	409,050	400,228	400,158	217	195	223	92	98
OH	644,947	679,187	642,149	612,270	541,830	429	384	414	549	478
OK	17,580	19,094	22,004	25,054	20,607	2,809	1,507	1,146	2,178	1,428
PA	1,485,775	1,511,491	1,390,644	1,379,262	1,566,449	2,332	3,217	2,242	2,525	2,641
TN	45,030	38,576	28,372	28,376	21,711	8,138	9,576	8,296	12,443	8,901
TX	538,258	605,290	578,697	564,850	576,803	1,262	563	584	480	386
UT	445,692	453,869	387,129	385,736	410,997	696	1,230	748	505	633
VA	564,326	562,805	493,398	456,766	393,225	1,500	1,210	1,044	1,254	955
WV	3,346,117	3,321,102	3,059,090	2,874,653	2,858,022	1,367	1,453	1,431	1,739	1,242
WY	7,658,286	7,591,675	6,973,727	6,760,447	6,880,205	1,024	1,224	893	711	869

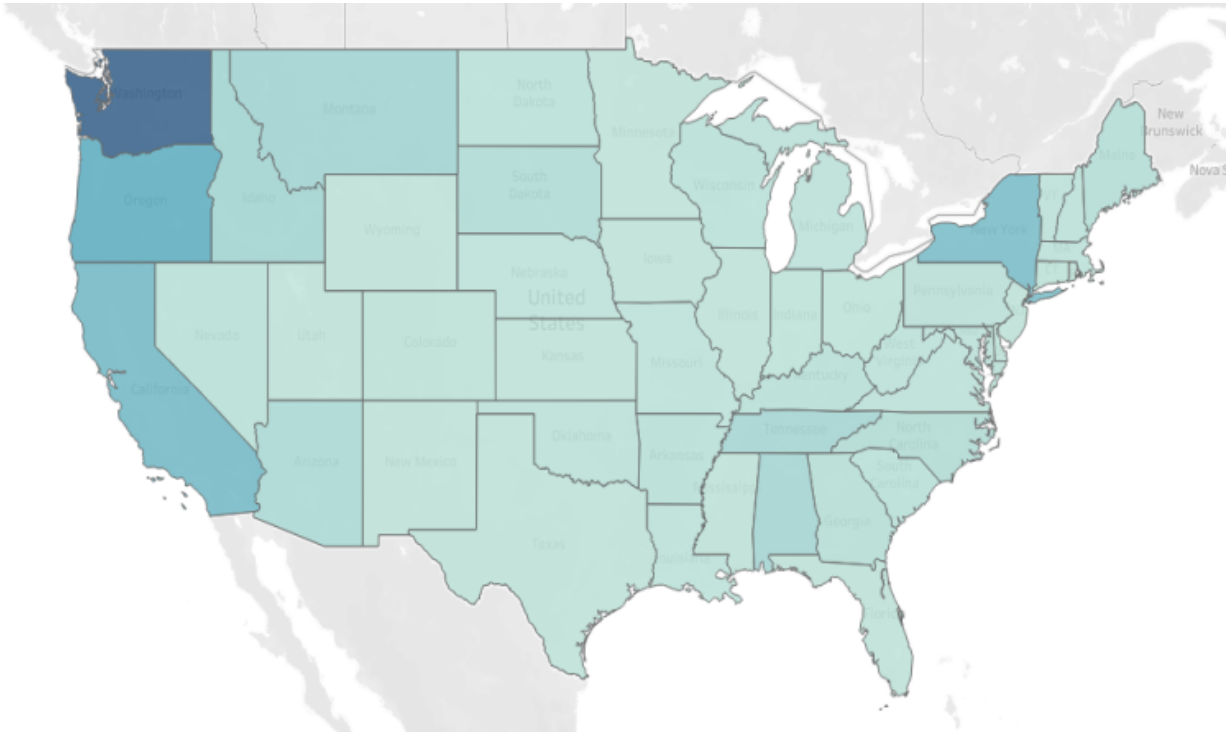
Visualization 2: States that produce only hydropower

Which states produced only hydropower helps to gain further understanding of its effects. The state codes were used as dimensions and the hydropower columns for 2011-2014 were used as measures. I filtered coal production for the four years to have 0 and hydropower to have greater than 0 for the four years. The side-by-side bars were effective in seeing the progression of hydropower production over the years based on each state.



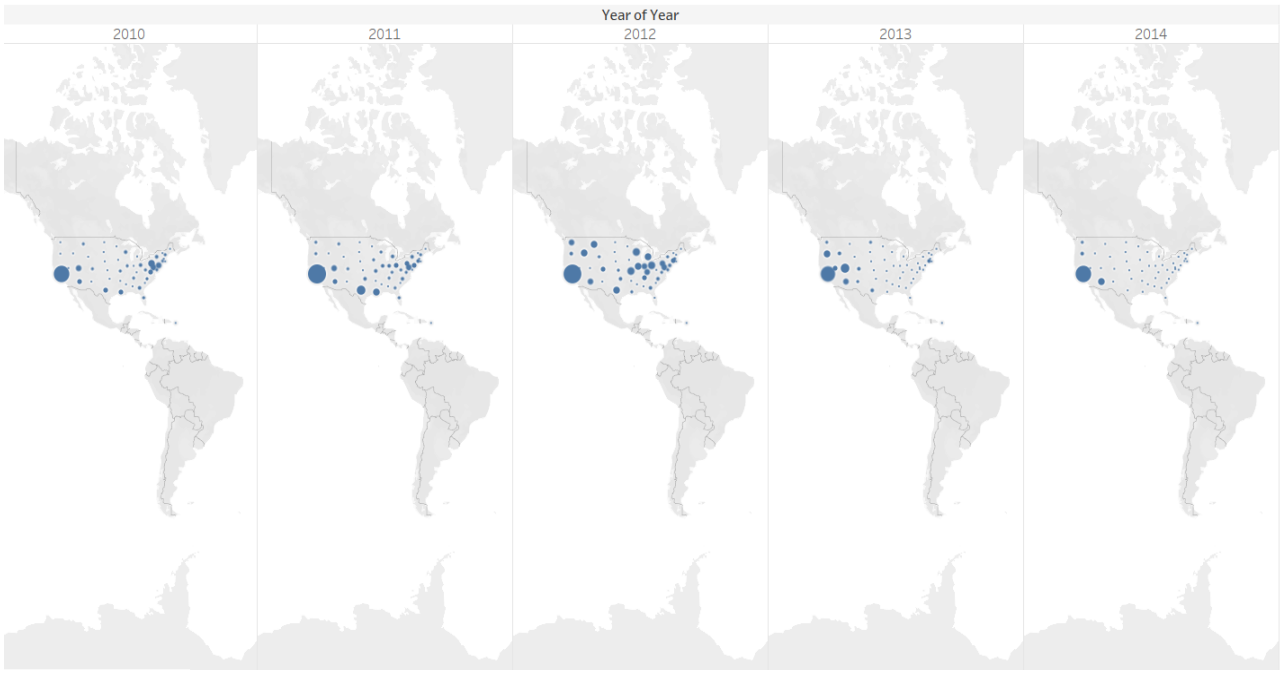
Visualization 3: Hydropower production for each state over the entire four year period

A map was used to better visualize the difference in hydropower production between states. This showed who the main producers were. A join was used between AQI locations with the production data in order to use states as a dimension and the 2011-2014 hydropower columns as measures.



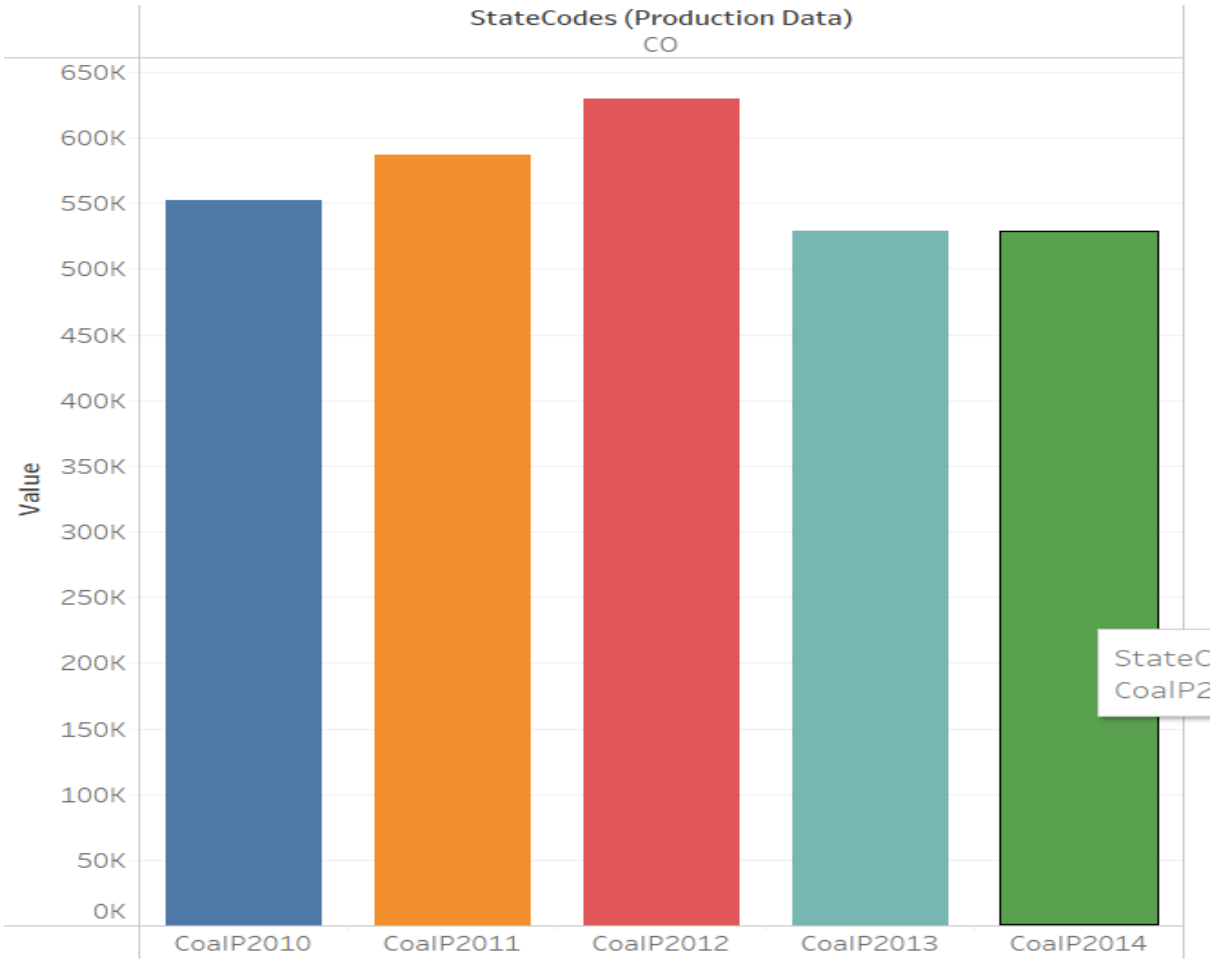
Visualization 4: Number of unhealthy AQI days for each year

The symbol map was the most effective at capturing the yearly changes that occurred to the unhealthy AQI days for each state. I used the categories and production tables to employ the year column and state code as the dimensions. I then used the sum of the unhealthy days as measures.



Visualization 5: Colorado's coal production each year

This visualization was used because of personal curiosity as to how much coal we produce each year. A simple bar chart was used to show the change in our coal production. The state codes were used as a dimension, while the coal production for each year was used as measures. I filtered the state codes to only use Colorado data.

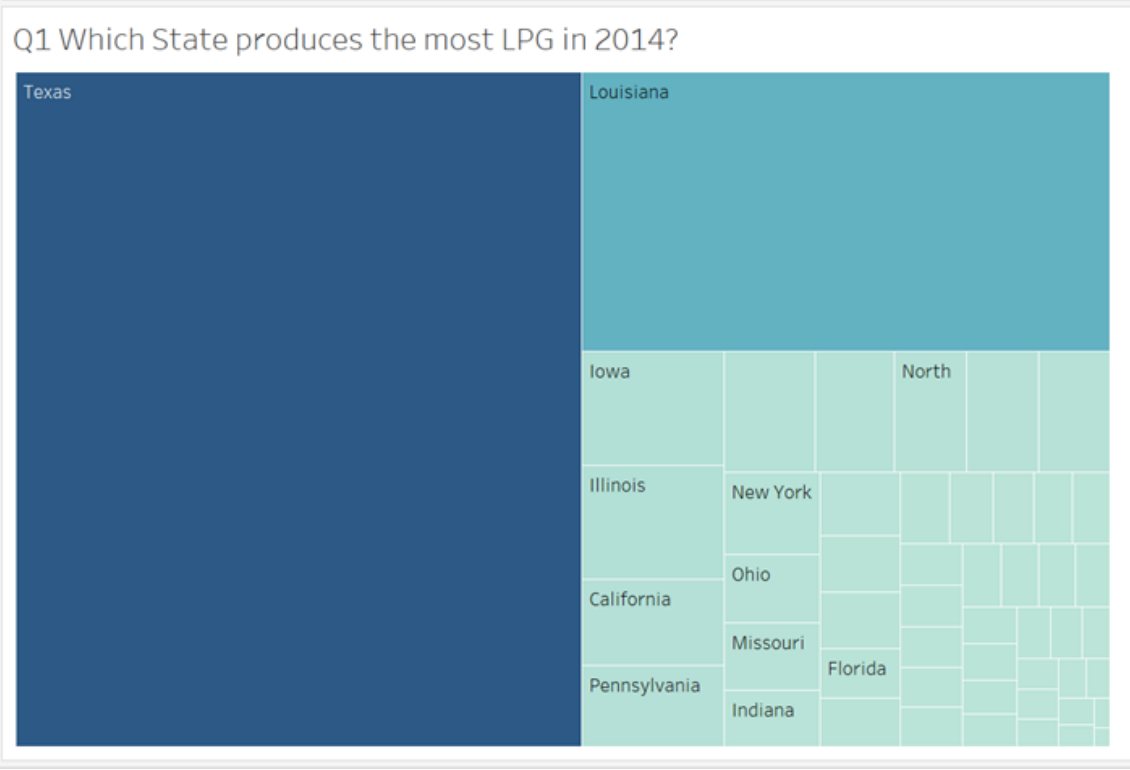


MSBX 5405 Team Database Project: Team 15

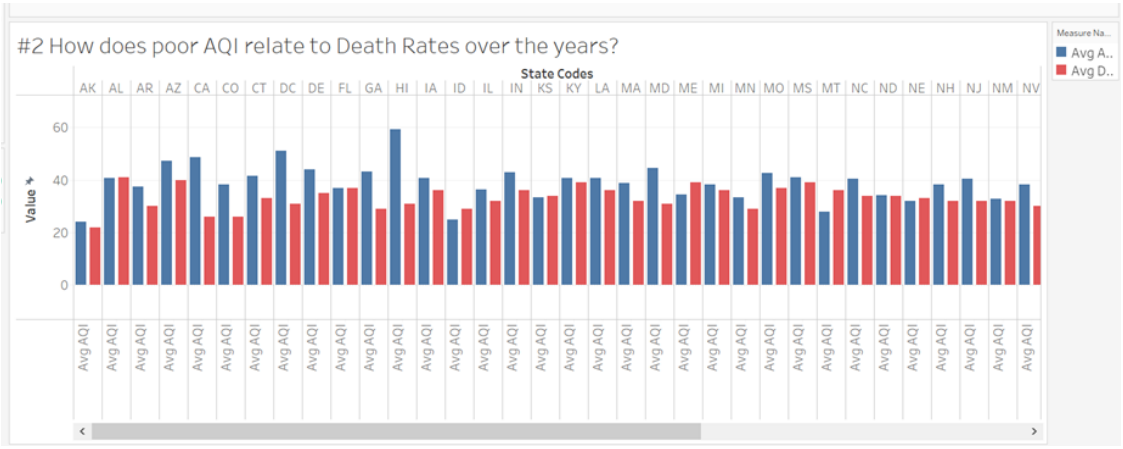
Samantha Fildish-

Visualization 1: Which State(s) produced the most LPG in 2014?

This tree map visualization uses the COAL Production data from the year 2014. The darker blue represents the state(s) with the higher production levels. Those states with lighter blue(almost green) have the lowest production levels of LPG.

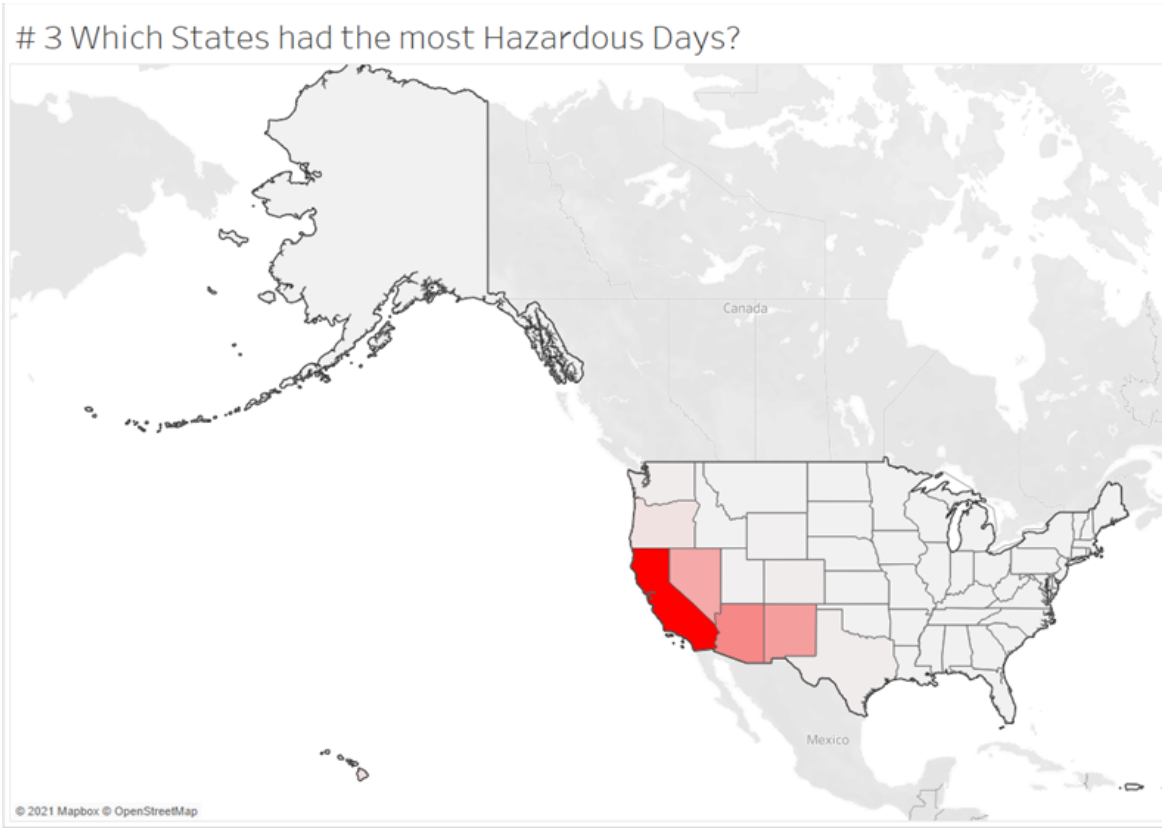


Visualization 2: How does poor AQI relate to Death Rates over the years?
For each state, I tried to compare the Average AQI with the Average Death rate to see if there seemed to be any correlations. Based on my results, there does not seem to be a strong correlation between higher AQI and higher death rates.



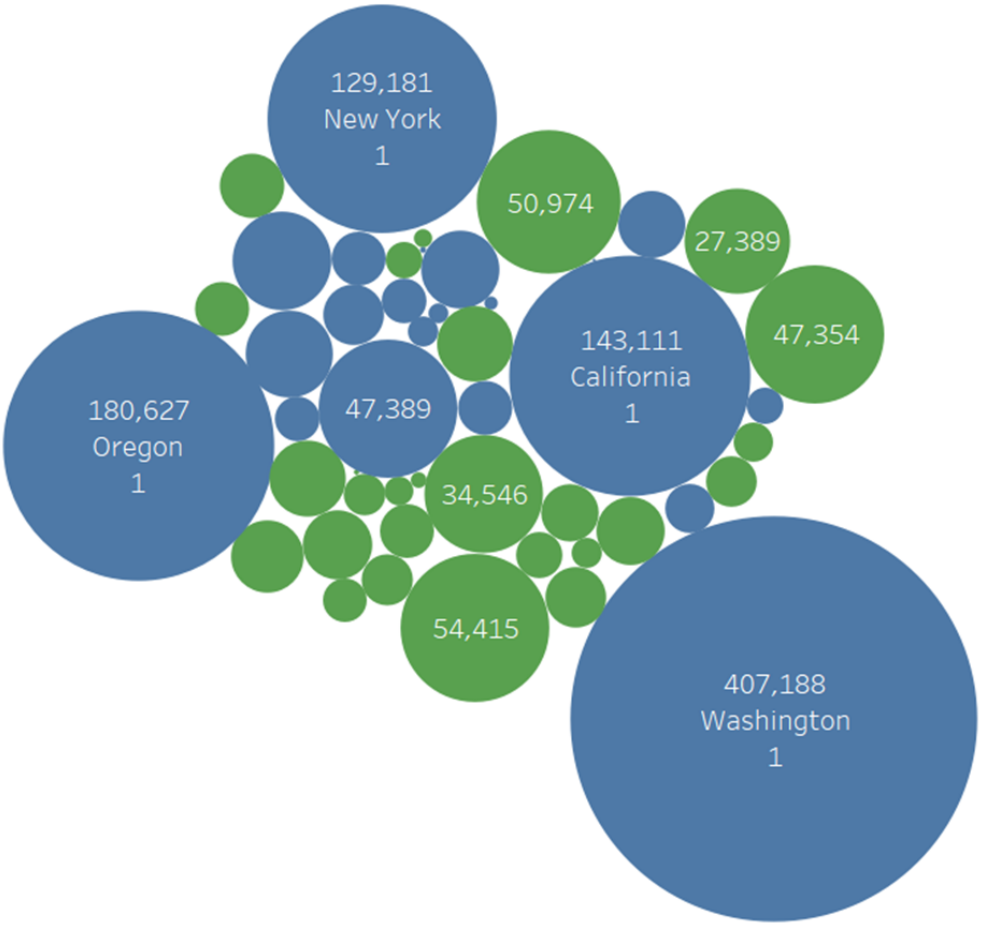
MSBX 5405 Team Database Project: Team 15

Visualization 3: Which States had the most Hazardous Days?
From this visualization, we see that over the 5 years of this study California had the highest number of Hazardous days overall, likely due to wildfires if I had to guess. Those in lighter red had some Hazardous days, and those in grey, almost with no color, had the least or 0 hazardous days overall.



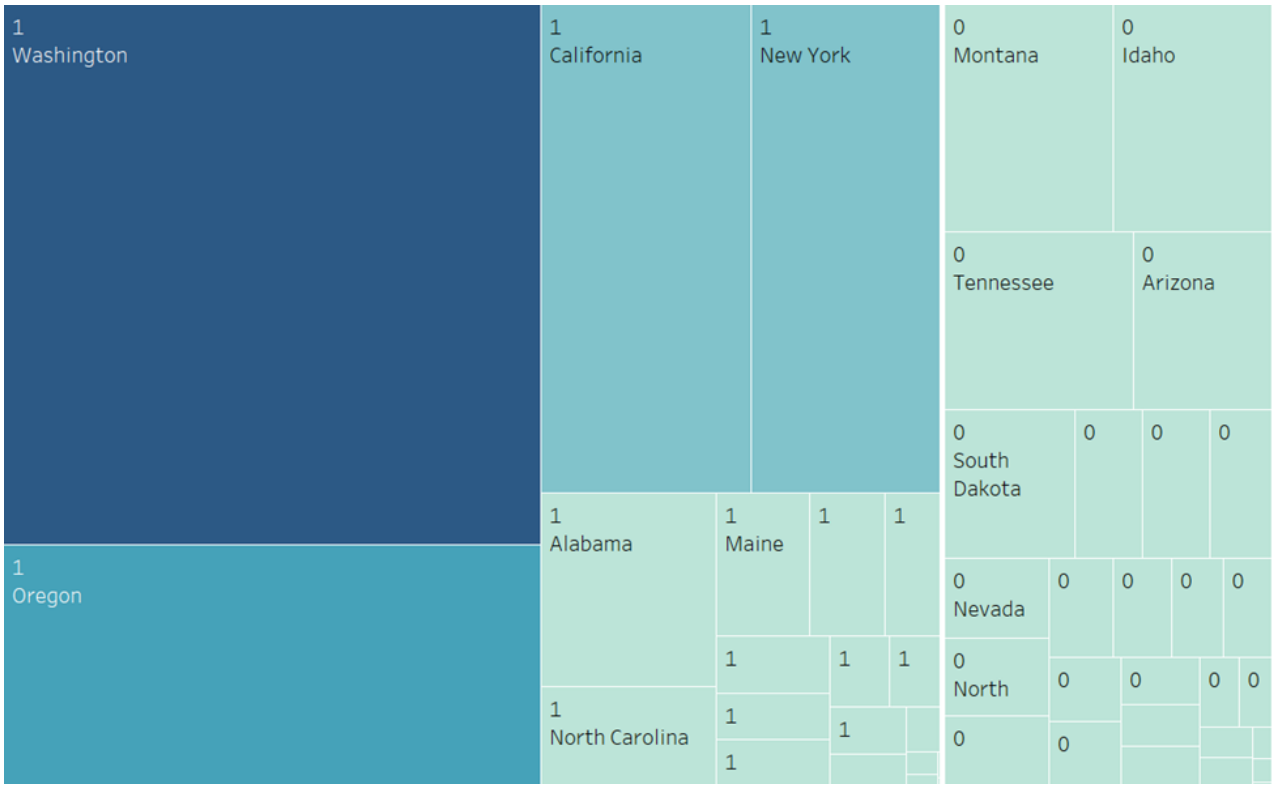
MSBX 5405 Team Database Project: Team 15

Visualization 4: Do coastal states produce more hydroelectric energy than non-coastal states? For this question I wanted to create a visualization that showed a comparison of the Hydro Energy production of Coastal States compared to Non-Coastal States. The Coastal States are represented in blue, while the landlocked states are represented in green. As one might expect, the Coastal states produce a significant more amount of hydro energy than landlocked states.



MSBX 5405 Team Database Project: Team 15

Visualization 5: Which States consume the most Hydro Energy?
For this question I wanted to see if those states that produced more hydro energy are the same states that consume the most hydro energy. It does appear that the top four hydro energy producing states are also the top for consumers.



MSBX 5405 Team Database Project: Team 15

References:

https://www.kaggle.com/threnjen/40-years-of-air-quality-index-from-the-epa-daily

https://www.kaggle.com/lislejoem/us_energy_census_gdp_10-14