

Daniel Finley

10/27/2020

U1325465

IS6420

Air Quality Data

To the political leaders of Utah,

I have run an analysis on the air quality index (AQI) data published by the EPA. The data for this analysis uses AQI data captured in 1999, 2009, and 2019. Because of the range of years involved I am able to look at trends over time. Let's first take a look at the average AQI per season for each of those three years.

| | 123 year T↑ | 123 average_aqi T↑ | season T↑ |
|----|-------------|--------------------|-----------|
| 1 | 1,999 | 42.5147582196 | fall |
| 2 | 1,999 | 54.8510300687 | spring |
| 3 | 1,999 | 62.6037358742 | summer |
| 4 | 1,999 | 38.8244949286 | winter |
| 5 | 2,009 | 34.0803997808 | fall |
| 6 | 2,009 | 42.9762318708 | spring |
| 7 | 2,009 | 43.8959388117 | summer |
| 8 | 2,009 | 39.2156758867 | winter |
| 9 | 2,019 | 34.1227011494 | fall |
| 10 | 2,019 | 39.7717294952 | spring |
| 11 | 2,019 | 40.5238351403 | summer |
| 12 | 2,019 | 35.9388522092 | winter |

As you can see in the table, the same trend holds true for each year for the national averages. The average AQI is highest in the spring and summer, peaking in the summer months, but as we approach the fall the AQI falls and then in winter it hits the lowest average point. Living in Salt Lake County, you may not have expected this result.

In addition to this query I also looked at the average for the whole year and the data shows that from 1999 to 2009 the average dropped by 10 and then from 2009 to 2019 it dropped again by another 2 points. Bringing us to the best national average AQI yet.

For now let us focus on Utah. I want to illustrate how many days each of these years the air qualified as unhealthy or worse. In 1999 we had 92 days, in 2009 that went up to 139 days, and in 2019 it dropped back down to 61. In 1999 and 2009 Salt Lake County led the counties in the number of days with unhealthy air. In 2009 they dropped to number two behind Uintah County.

| | 123 year T↑ | ABC county Name T↑ | 123 days_unhealthy T↑ |
|----|-------------|--------------------|-----------------------|
| 1 | 1,999 | Davis | 13 |
| 2 | 1,999 | Salt Lake | 36 |
| 3 | 1,999 | San Juan | 5 |
| 4 | 1,999 | Utah | 26 |
| 5 | 1,999 | Weber | 12 |
| 6 | 2,009 | Box Elder | 5 |
| 7 | 2,009 | Cache | 28 |
| 8 | 2,009 | Davis | 12 |
| 9 | 2,009 | Salt Lake | 39 |
| 10 | 2,009 | San Juan | 2 |
| 11 | 2,009 | Tooele | 10 |
| 12 | 2,009 | Uintah | 2 |
| 13 | 2,009 | Utah | 22 |
| 14 | 2,009 | Washington | 3 |
| 15 | 2,009 | Weber | 16 |
| 16 | 2,019 | Box Elder | 2 |
| 17 | 2,019 | Cache | 7 |
| 18 | 2,019 | Carbon | 1 |
| 19 | 2,019 | Davis | 5 |
| 20 | 2,019 | Duchesne | 11 |
| 21 | 2,019 | Salt Lake | 15 |
| 22 | 2,019 | Tooele | 1 |
| 23 | 2,019 | Uintah | 17 |
| 24 | 2,019 | Weber | 2 |

Let's look at just Salt Lake county now:

| | 123 year | 123 month | 123 days_unhealthy | ABC county Name |
|----|----------|-----------|--------------------|-----------------|
| 1 | 1,999 | 1 | 2 | Salt Lake |
| 2 | 1,999 | 2 | 2 | Salt Lake |
| 3 | 1,999 | 5 | 1 | Salt Lake |
| 4 | 1,999 | 6 | 6 | Salt Lake |
| 5 | 1,999 | 7 | 10 | Salt Lake |
| 6 | 1,999 | 8 | 5 | Salt Lake |
| 7 | 1,999 | 12 | 10 | Salt Lake |
| 8 | 2,009 | 1 | 9 | Salt Lake |
| 9 | 2,009 | 5 | 1 | Salt Lake |
| 10 | 2,009 | 6 | 3 | Salt Lake |
| 11 | 2,009 | 7 | 9 | Salt Lake |
| 12 | 2,009 | 8 | 4 | Salt Lake |
| 13 | 2,009 | 9 | 2 | Salt Lake |
| 14 | 2,009 | 11 | 1 | Salt Lake |
| 15 | 2,009 | 12 | 10 | Salt Lake |
| 16 | 2,019 | 1 | 3 | Salt Lake |
| 17 | 2,019 | 6 | 3 | Salt Lake |
| 18 | 2,019 | 7 | 4 | Salt Lake |
| 19 | 2,019 | 8 | 5 | Salt Lake |

* This data set does not include data for Salt Lake county for Nov or Dec of 2019


It does appear that historically December has been the worst month for Salt Lake County in terms of air quality. The unhealthy days also appear to peak in July, matching the rest of the country and falling in with the normal trends. We can account the poor air quality of December to the inversion effect we see here. It is difficult to tell if the air quality is getting better over time as we are missing the crucial November and December months for 2019 but we can see that January of 2019 does show the most unhealthy days of any of the years. This may be a predictor or worsening air quality in the winters. On the other hand, the air quality in the summers appears to be improving.


Data







Q1.

```
-- Questions
-- 1. What is the average AQI (air quality index) by year by season (winter, spring, summer, fall)?

select
  extract (year from "Date") as year, avg(aqi) as average_AQI
  ,(case when to_char("Date", 'MMDD') between '0321' and '0620' then 'spring'
        when to_char("Date", 'MMDD') between '0621' and '0922' then 'summer'
        when to_char("Date", 'MMDD') between '0923' and '1220' then 'fall'
        else 'winter'
      end) as season
FROM aqi
group by year, season
;
```

Results 




`select extract (year from "Date") as year, avg(aqi) as average_AQI`  *Enter a SQL expression to filter results (use Ctrl+Space)*

| |  year  |  average_aqi  |  season  |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | 1,999 | 42.5147582196 | fall |
| 2 | 1,999 | 54.8510300687 | spring |
| 3 | 1,999 | 62.6037358742 | summer |
| 4 | 1,999 | 38.8244949286 | winter |
| 5 | 2,009 | 34.0803997808 | fall |
| 6 | 2,009 | 42.9762318708 | spring |
| 7 | 2,009 | 43.8959388117 | summer |
| 8 | 2,009 | 39.2156758867 | winter |
| 9 | 2,019 | 34.1227011494 | fall |
| 10 | 2,019 | 39.7717294952 | spring |
| 11 | 2,019 | 40.5238351403 | summer |
| 12 | 2,019 | 35.9388522092 | winter |
| | | | |
| | | | |
| | | | |

days_unhealthy:

-- 2. What were the top 10 locations with worst AQI in each year?

```
(select
  extract (year from "Date") as year, aqi."county Name"
  ,avg(aqi) over (order by aqi desc) avg_aqi
from aqi
where extract (year from "Date") = 2019
group by year, aqi."county Name", aqi.aqi
order by avg_aqi desc, year
limit 10)
union
(select
  extract (year from "Date") as year, aqi."county Name"
  ,avg(aqi) over (order by aqi desc) avg_aqi
from aqi
where extract (year from "Date") = 2009
group by year, aqi."county Name", aqi.aqi
order by avg_aqi desc, year
limit 10)
union
(select
  extract (year from "Date") as year, aqi."county Name"
  ,avg(aqi) over (order by aqi desc) avg_aqi
from aqi
where extract (year from "Date") = 1999
group by year, aqi."county Name", aqi.aqi
order by avg_aqi desc, year
limit 10)
;
```

| | 123 year  | ABC county Name  | 123 avg_aqi  | |
|----|--------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|--|
| 1 | 1,999 | Carbon | 1,409.1111111111 | |
| 2 | 1,999 | Inyo | 2,017.4 | |
| 3 | 1,999 | Inyo | 1,829.8333333333 | |
| 4 | 1,999 | Inyo | 2,356 | |
| 5 | 1,999 | Inyo | 2,797 | |
| 6 | 1,999 | Inyo | 1,409.1111111111 | |
| 7 | 1,999 | Inyo | 1,320.3 | |
| 8 | 1,999 | Inyo | 1,657.4285714286 | |
| 9 | 1,999 | Inyo | 2,631 | |
| 10 | 1,999 | CHIHUAHUA STATE | 2,200.25 | |
| 11 | 2,009 | Mono | 5,600.6666666667 | |
| 12 | 2,009 | Inyo | 7,722.5 | |
| 13 | 2,009 | Mono | 3,106.2857142857 | |
| 14 | 2,009 | Pinal | 4,535.75 | |
| 15 | 2,009 | Mono | 14,043 | |
| 16 | 2,009 | Inyo | 3,889 | |
| 17 | 2,009 | BAJA CALIFORNIA NOR | 2,816.75 | |
| 18 | 2,009 | Pinal | 2,401.1 | |
| 19 | 2,009 | Mono | 3,453 | |
| 20 | 2,009 | Pinal | 2,590.7777777778 | |
| 21 | 2,019 | Luna | 1,738.5 | |
| 22 | 2,019 | Dona Ana | 1,898.7142857143 | |
| 23 | 2,019 | Mono | 3,089.6666666667 | |
| 24 | 2,019 | Mono | 2,720 | |
| 25 | 2,019 | Mono | 3,852 | |
| 26 | 2,019 | Mono | 2,355 | |
| 27 | 2,019 | Pinal | 2,110.1666666667 | |
| 28 | 2,019 | Dona Ana | 1,610.5555555556 | |
| 29 | 2,019 | Mono | 1,501 | |
| 30 | 2,019 | Mono | 3,587.5 | |

-- 4. In Utah counties, how many days of "Unhealthy" air did we have in each year? Is it improving?
 --Ordered by year and county. Showing the breakdown of unhealthy or worse air for each Utah county for each year.

```
select
  extract (year from aqi."Date") as year
  ,aqi."county Name"
  ,count("Date") as days_unhealthy
from aqi
where "State Name" = 'Utah'
  and ( category = 'Unhealthy'
    or category = 'Very Unhealthy'
    or category = 'Unhealthy for Sensitive Groups')
group by
  extract (year from aqi."Date")
  ,aqi."county Name"
;
```

| aqi | | | |
|-------------------------------------------------------------|-------|-------------|----------------|
| select extract (year from aqi."Date") as year,aqi."county N | | | |
| | year | county Name | days_unhealthy |
| 1 | 1,999 | Davis | 13 |
| 2 | 1,999 | Salt Lake | 36 |
| 3 | 1,999 | San Juan | 5 |
| 4 | 1,999 | Utah | 26 |
| 5 | 1,999 | Weber | 12 |
| 6 | 2,009 | Box Elder | 5 |
| 7 | 2,009 | Cache | 28 |
| 8 | 2,009 | Davis | 12 |
| 9 | 2,009 | Salt Lake | 39 |
| 10 | 2,009 | San Juan | 2 |
| 11 | 2,009 | Tooele | 10 |
| 12 | 2,009 | Uintah | 2 |
| 13 | 2,009 | Utah | 22 |
| 14 | 2,009 | Washington | 3 |
| 15 | 2,009 | Weber | 16 |
| 16 | 2,019 | Box Elder | 2 |
| 17 | 2,019 | Cache | 7 |
| 18 | 2,019 | Carbon | 1 |
| 19 | 2,019 | Davis | 5 |
| 20 | 2,019 | Duchesne | 11 |
| 21 | 2,019 | Salt Lake | 15 |
| 22 | 2,019 | Tooele | 1 |
| 23 | 2,019 | Uintah | 17 |
| 24 | 2,019 | Weber | 2 |

-- 5. In Salt Lake County, which months have the most "Unhealthy" days? Has that changed since 1999?

```
select
  extract (year from aqi."Date") as year
,extract (month from aqi."Date") as month
,count("Date") as days_unhealthy
, aqi."county Name"
from aqi
where aqi."county Name" = 'Salt Lake'
and (category = 'Unhealthy'
or category = 'Very Unhealthy'
or category = 'Unhealthy for Sensitive Groups')
group by
  extract (year from aqi."Date")
,extract (month from aqi."Date")
, aqi."county Name"
```

aqi

select extract (year from aqi."Date") as year,extract (mont | Enter a SQL expression to filter results (use Ctrl+Space)

| | year | month | days_unhealthy | county Name |
|----|-------|-------|----------------|-------------|
| 1 | 1,999 | 1 | 2 | Salt Lake |
| 2 | 1,999 | 2 | 2 | Salt Lake |
| 3 | 1,999 | 5 | 1 | Salt Lake |
| 4 | 1,999 | 6 | 6 | Salt Lake |
| 5 | 1,999 | 7 | 10 | Salt Lake |
| 6 | 1,999 | 8 | 5 | Salt Lake |
| 7 | 1,999 | 12 | 10 | Salt Lake |
| 8 | 2,009 | 1 | 9 | Salt Lake |
| 9 | 2,009 | 5 | 1 | Salt Lake |
| 10 | 2,009 | 6 | 3 | Salt Lake |
| 11 | 2,009 | 7 | 9 | Salt Lake |
| 12 | 2,009 | 8 | 4 | Salt Lake |
| 13 | 2,009 | 9 | 2 | Salt Lake |
| 14 | 2,009 | 11 | 1 | Salt Lake |
| 15 | 2,009 | 12 | 10 | Salt Lake |
| 16 | 2,019 | 1 | 3 | Salt Lake |
| 17 | 2,019 | 6 | 3 | Salt Lake |
| 18 | 2,019 | 7 | 4 | Salt Lake |
| 19 | 2,019 | 8 | 5 | Salt Lake |