

# Air Quality Analysis

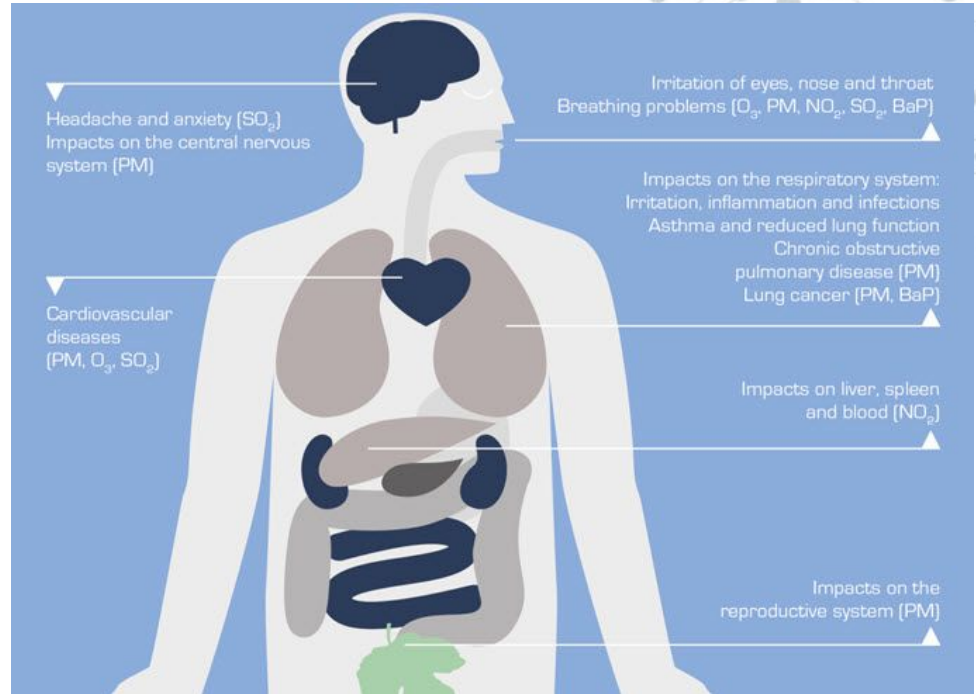


Haocheng Zhu  
Isabel Souza Shiratsubaki  
Anh Hoang

Group 20

# Overview

- Air quality is fundamental to our well-being
- On average, a person inhales about 14,000 litres of air every day, and the presence of contaminants in this air can adversely affect people's health
- Global warming is serious concern



# Motivation

Pollution levels from 2014 to 2016



## Top 10 most polluted metro areas in the U.S.

Ozone pollution

<b>1</b>	Los Angeles-Long Beach	<b>6</b>	San Diego-Carlsbad
<b>2</b>	Bakersfield, Calif.	<b>7</b>	Modesto-Merced, Calif.
<b>3</b>	Visalia-Porterville-Hanford, Calif.	<b>8</b>	Phoenix-Mesa-Scottsdale
<b>4</b>	Fresno-Madera, Calif.	<b>9</b>	Redding-Red Bluff, Calif.
<b>5</b>	Sacramento-Roseville, Calif.	<b>10</b>	New York-Newark-N.Y., N.J., Conn., Pa.

SOURCE lung.org


- California has eight out of ten most polluted cities in the nation
- The Los Angeles/Long Beach and San Diego area took the dubious distinction of being the nation's most ozone-polluted city as it has for nearly the entire 19-year history of the report.
- Outdoor air pollution continues to threaten the lives and health of millions of people in the U.S.

Source: <https://www.usatoday.com/story/news/nation/2018/04/18/california-has-eight-10-most-polluted-u-s-cities/524815002/>



# Objective

Creating a system to analyze the air quality  
based on locations, its content and time



## Datasets

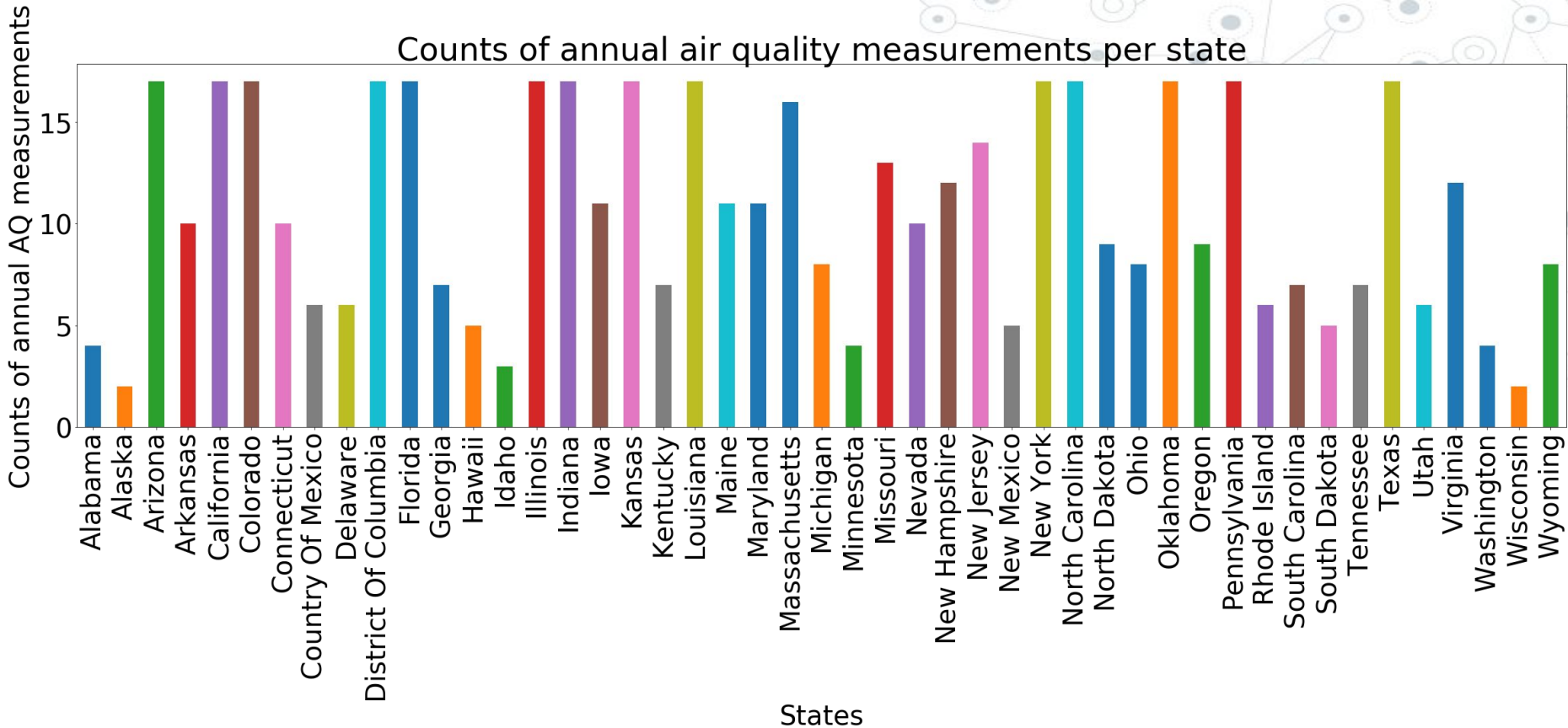
- © Pollution in US open dataset:  
<https://www.kaggle.com/sogun3/uspollution>
- © The dataset presents information about the state, the county, the city, the date of the measurements (from 2000 to 2016), and measurements of the pollutants level in the air
- © The dataset format is csv (size: 382 MB)

# Methodology: EPA Category

Using the data from **Air Quality Index (AQI) Basics** to rate the air quality of each state and difference time of the year

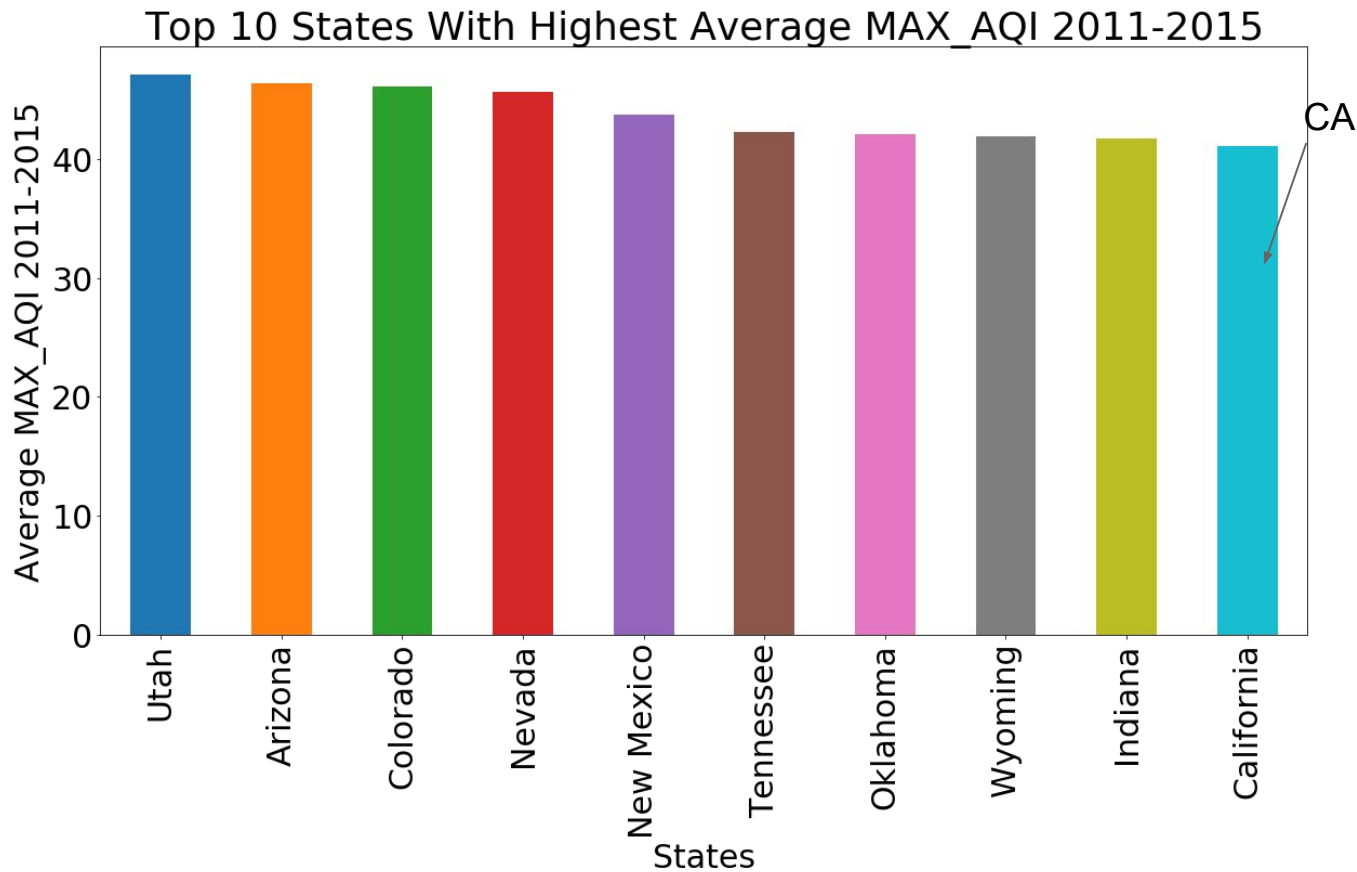
Air Quality Index (AQI) Values	Levels of Health Concern	Colors
<i>When the AQI is in this range:</i>	<i>..air quality conditions are:</i>	<i>...as symbolized by this color:</i>
0 to 50	Good	Green
51 to 100	Moderate	Yellow
101 to 150	Unhealthy for Sensitive Groups	Orange
151 to 200	Unhealthy	Red
201 to 300	Very Unhealthy	Purple
301 to 500	Hazardous	Maroon

Note: Values above 500 are considered Beyond the AQI. Follow recommendations for the Hazardous category. Additional information on reducing exposure to extremely high levels of particle pollution is available [here](#).



Air quality was measured across 50 states. However, the number of data values were vary from state to state.

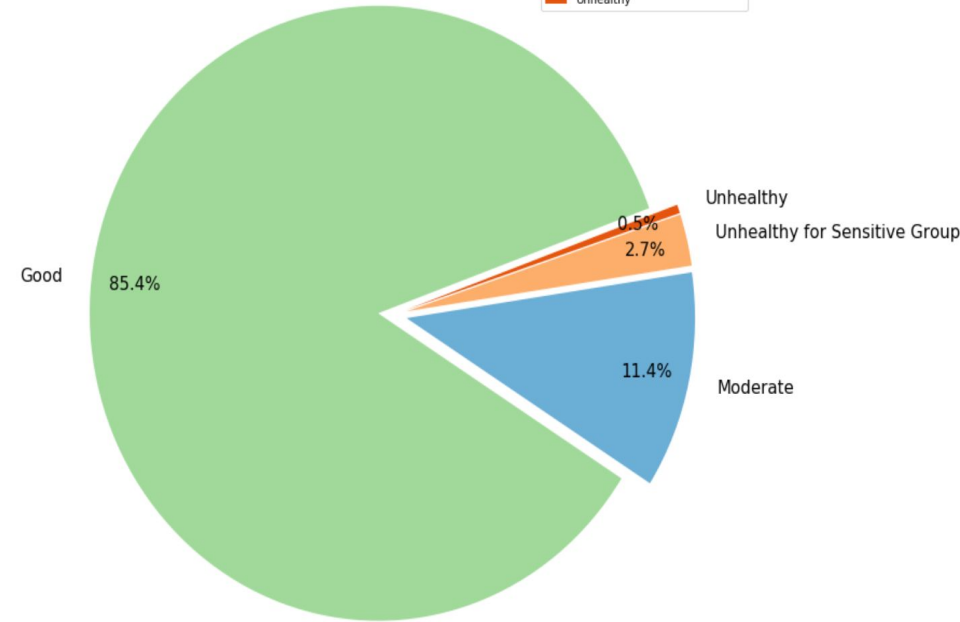
## Top 10 States with Highest Average of AQI Max from 2011-2015



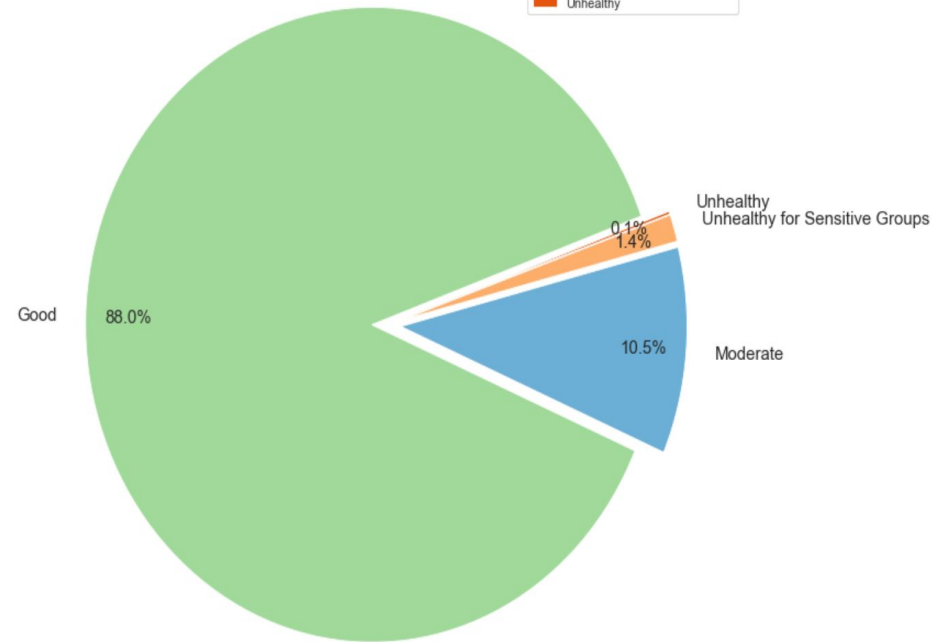


# EPA Classification CA vs U.S.A

CA EPA Percentage



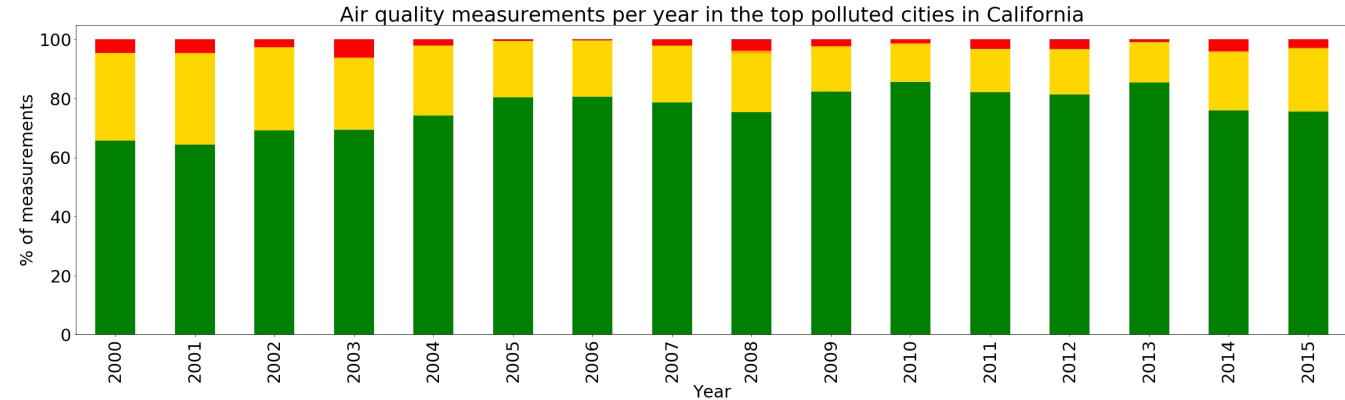
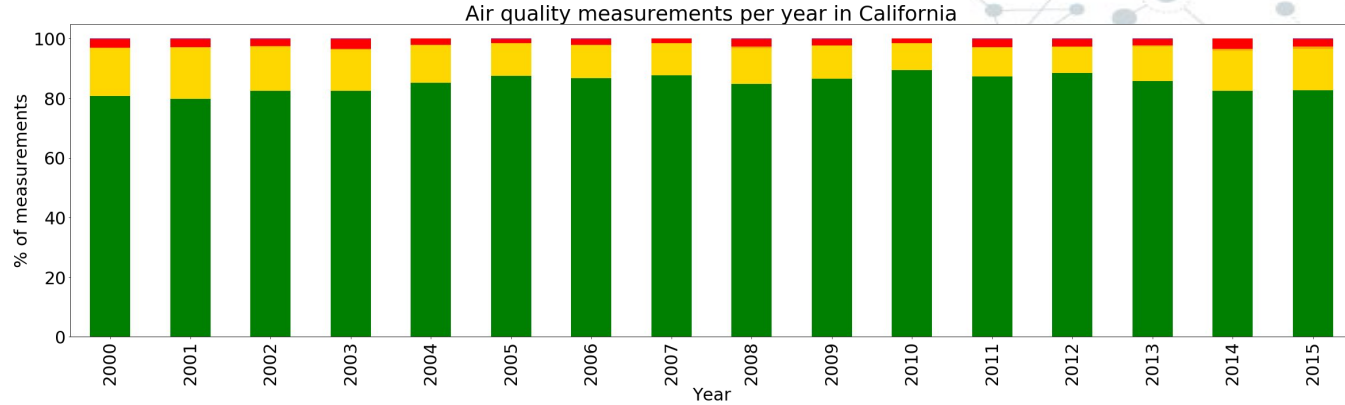
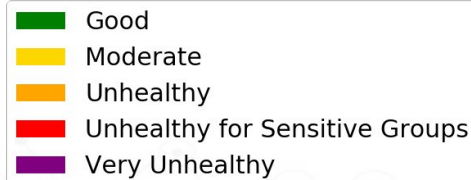
U.S EPA Percentage



- California has much more percentage of unhealthy days than U.S average

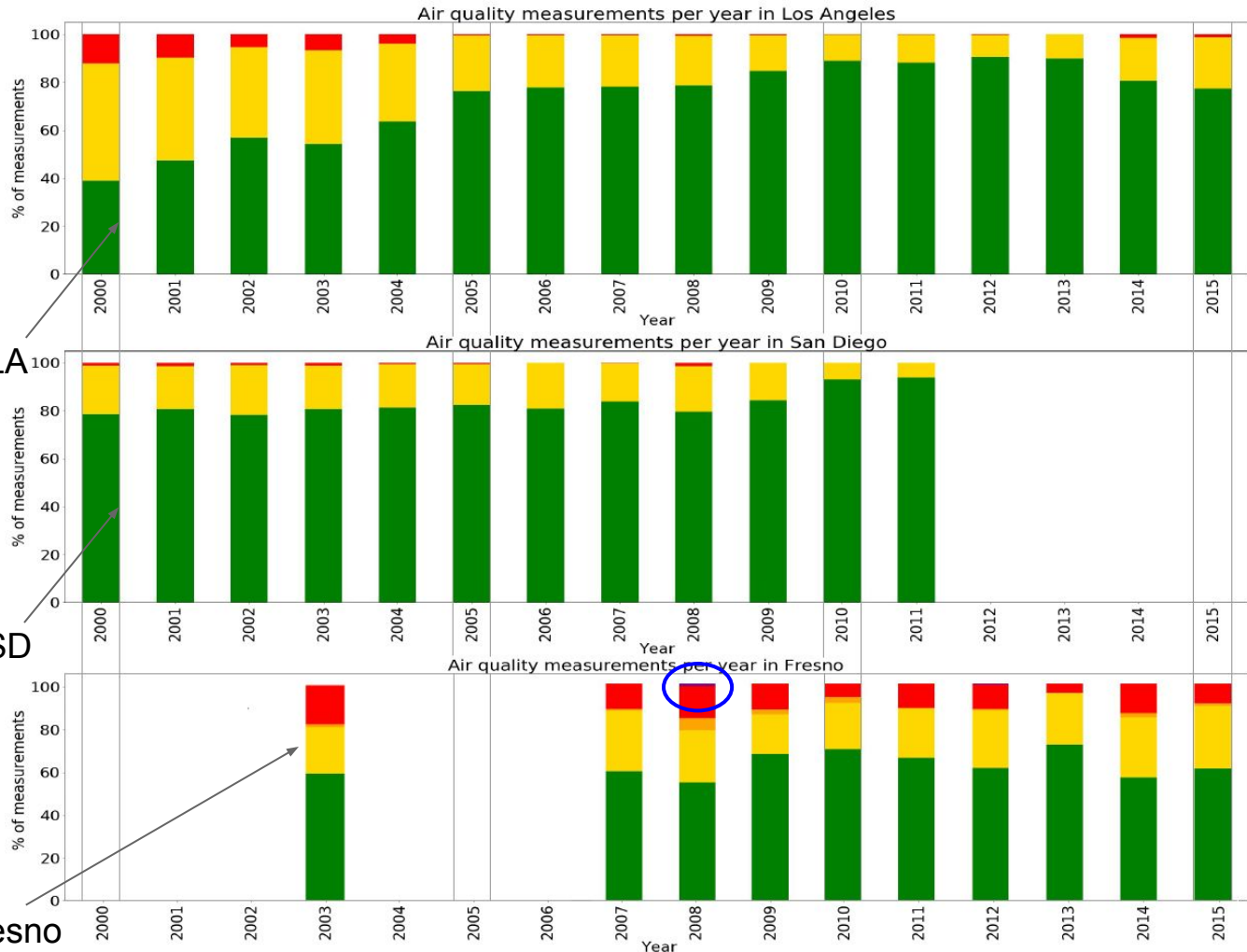
# Air Quality CA versus Top Polluted Cities in CA

- ◎ The number of Moderate air quality measurements increase when we filtered only top polluted cities in CA
- ◎ Unhealthier air quality measurements were verified between 2000 and 2004

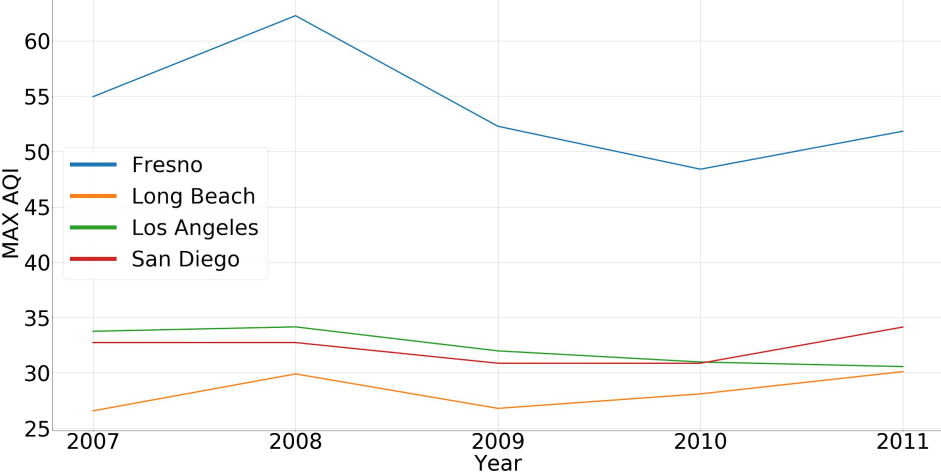


© Fresno was the city that presented the worst air quality among other polluted cities, especially in 2008

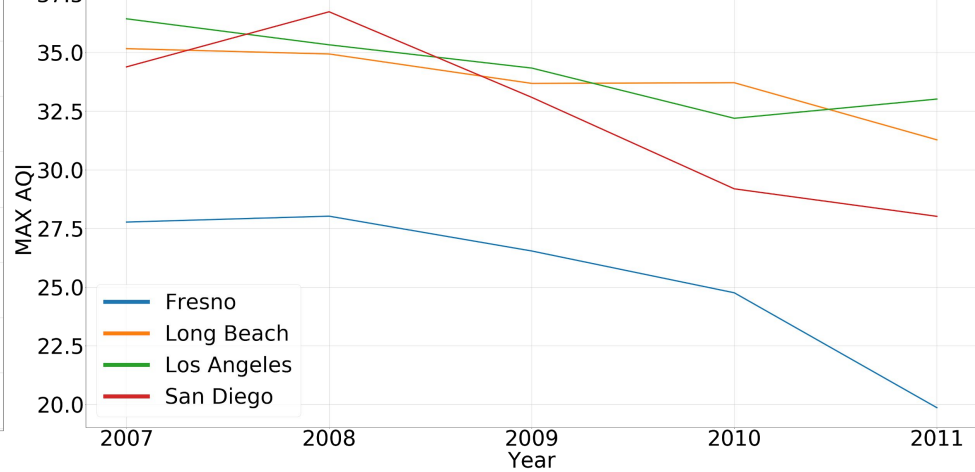
© Overall, LA presents worse air quality than SD over time



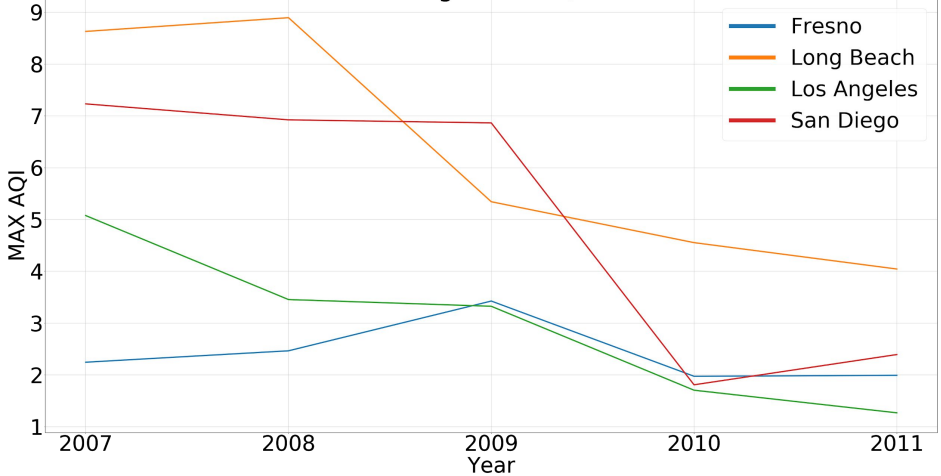
Annual average O3 AQI over time



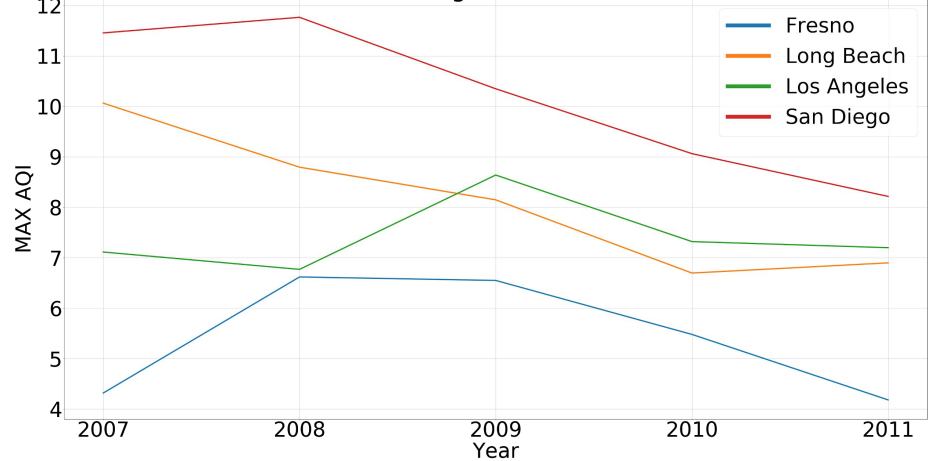
Annual average NO2 AQI over time

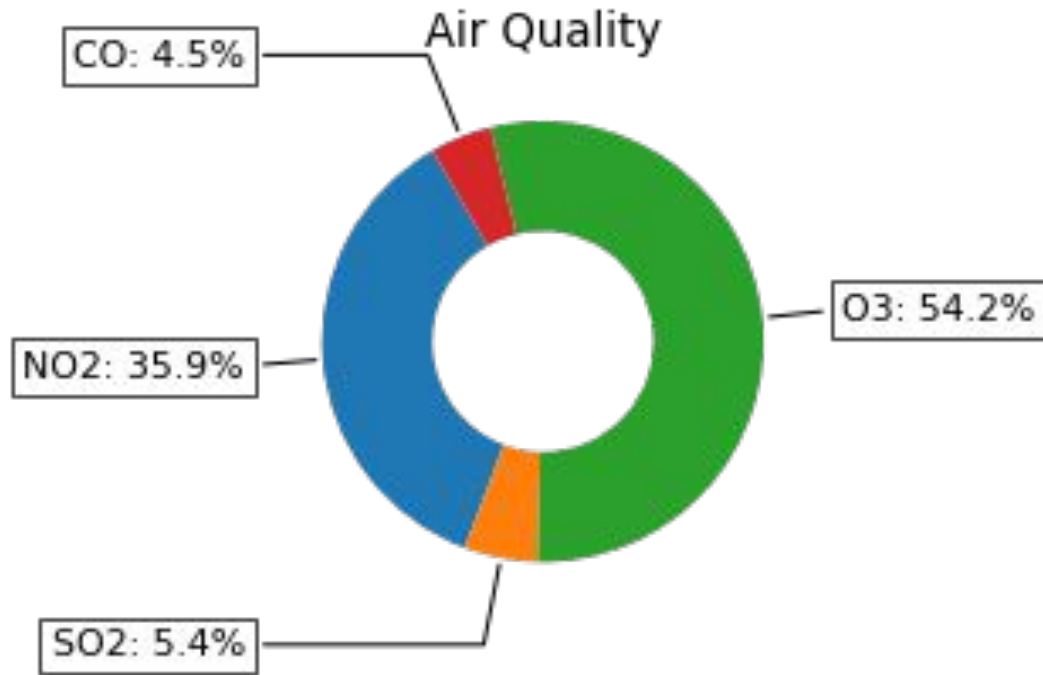


Annual average SO2 AQI over time



Annual average CO AQI over time





- O3 was the most prevalent pollutant gas followed by NO2, SO2 and CO over the period of 2000 to 2016

## Conclusion

- ◎ California air quality is one of the worst in US and it presents bad performance compared to the average of US
- ◎ Top most polluted cities in CA presented higher AQI than the average in CA
- ◎ Ozone is the most prevalent pollutant gas in the US, with the highest recorded O3 AQI levels recorded in Fresno
- ◎ Our results confirm that California is a leader in air pollution among other states



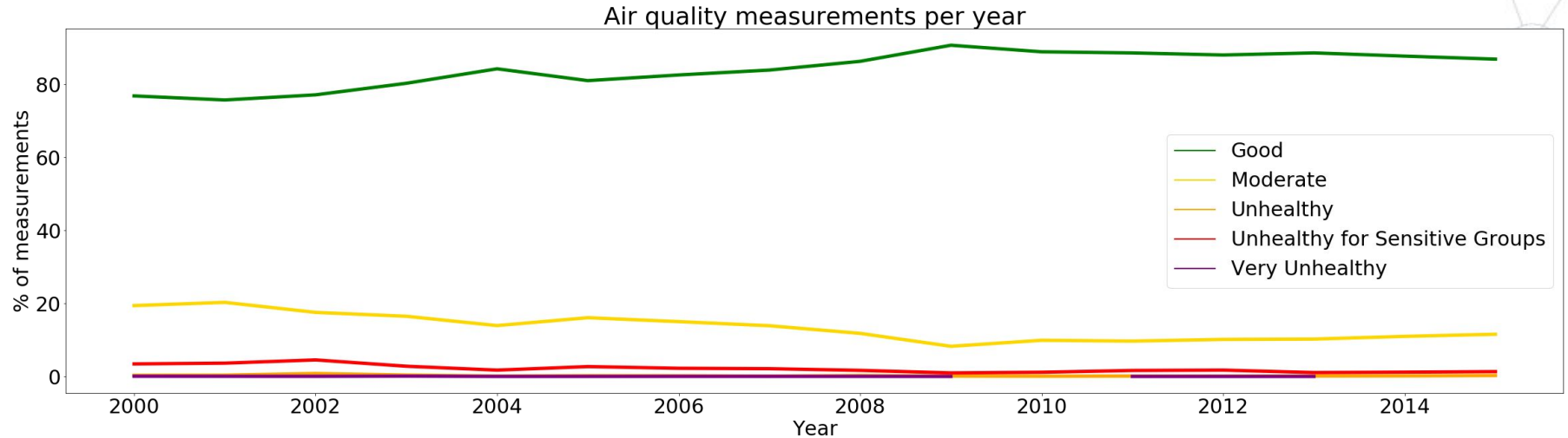
THANK YOU



# Appendix



# Air quality measurements over time per category



## Air Quality Index (AQI) - EPA

$$I = \frac{I_{high} - I_{low}}{C_{high} - C_{low}}(C - C_{low}) + I_{low}$$

where:

$I$  = the (Air Quality) index,

$C$  = the pollutant concentration,

$C_{low}$  = the concentration breakpoint that is  $\leq C$ ,

$C_{high}$  = the concentration breakpoint that is  $\geq C$ ,

$I_{low}$  = the index breakpoint corresponding to  $C_{low}$ ,

$I_{high}$  = the index breakpoint corresponding to  $C_{high}$ .