

Madrid Air Quality

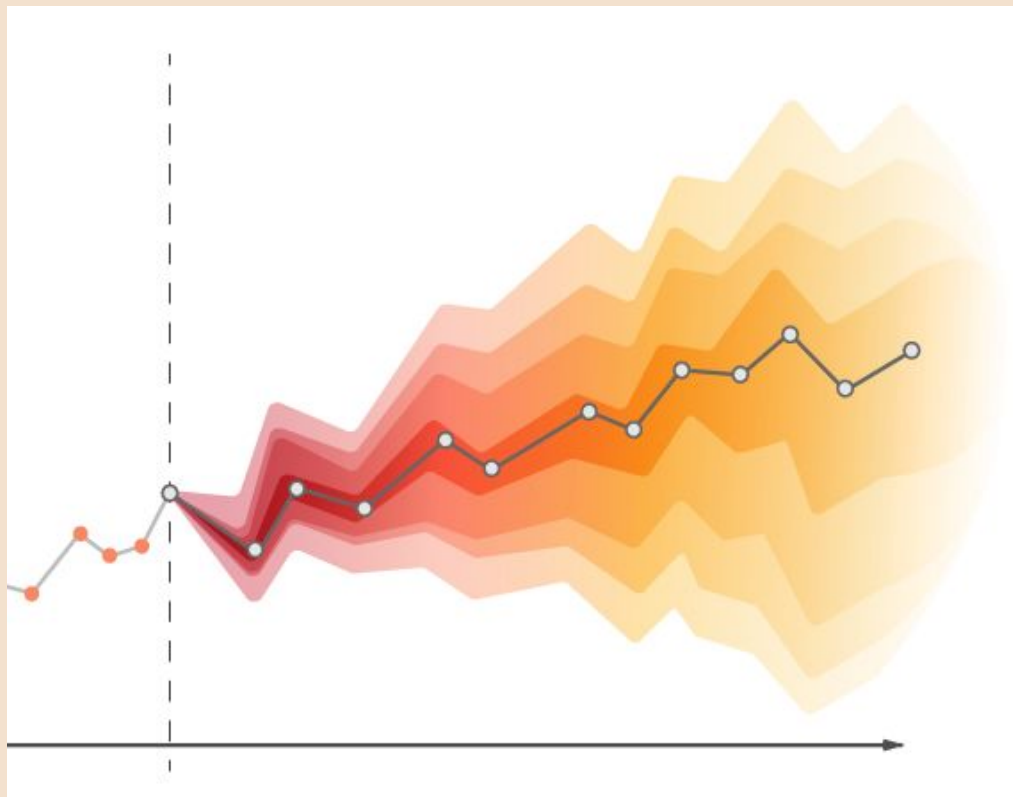
Explanatory Analysis and
Forecasting

Buu NGUYEN

Gwenaëlle DAVRAIN

Aya BEN HRIZ

Anh Thu DOAN



Introduction

Global warming is the long-term gradual heating of Earth's climate system.

→ It is a direct consequence of contamination

The contamination of the air is due to some particles (carbon monoxide)

→ Negative impact on the environment

→ Negative impact on human beings

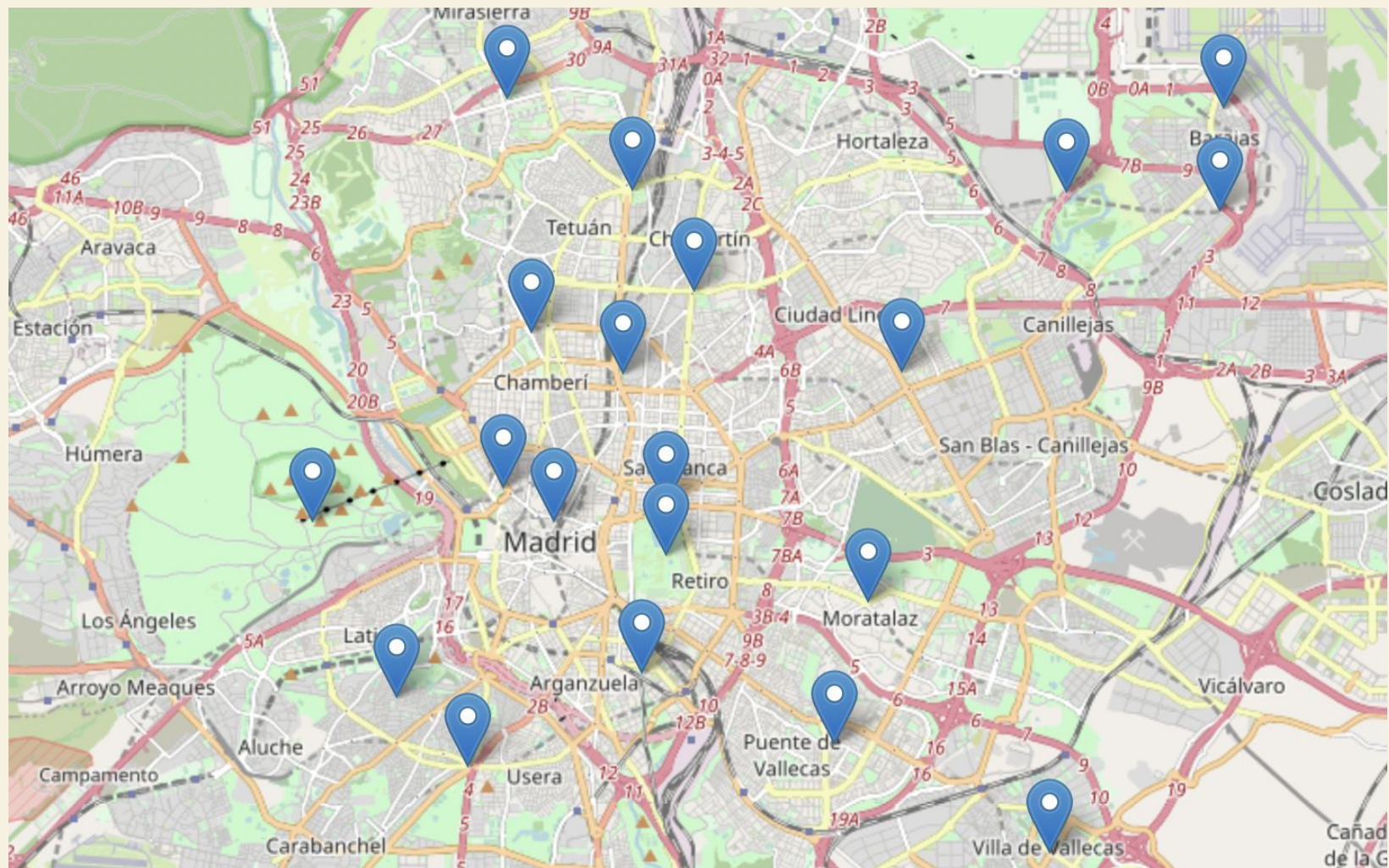


General Information

- This dataset is created in the Madrid's City Council Open Data website. It contains in a practical format 18 years (2001-2018) of daily and hourly data.
- The dataset contains stations at the highest hierarchical level: each station history can be individually extracted from the file for further study
- The complete list of available pollutants: SO₂, CO, NO, NO₂, PM_{2.5}, PM₁₀, NO_x, O₃, TOL, BEN, EBE, MXY, PXY, OXY, TCH, CH₄, NMHC.


Methods

- Data pre-processing
- Forecasting



Data pre-processing



- The mean of all stations was calculated
 - Data was converted into CAQI from $\mu\text{g}/\text{m}^3$
 - The **peak** of each day was taken as the index of the day
 - Monthly index is the mean of days
 - Yearly index is the mean of months
- 

Pollutants and calculation grid for the revised CAQI hourly and daily grid (all changes in *italics*)

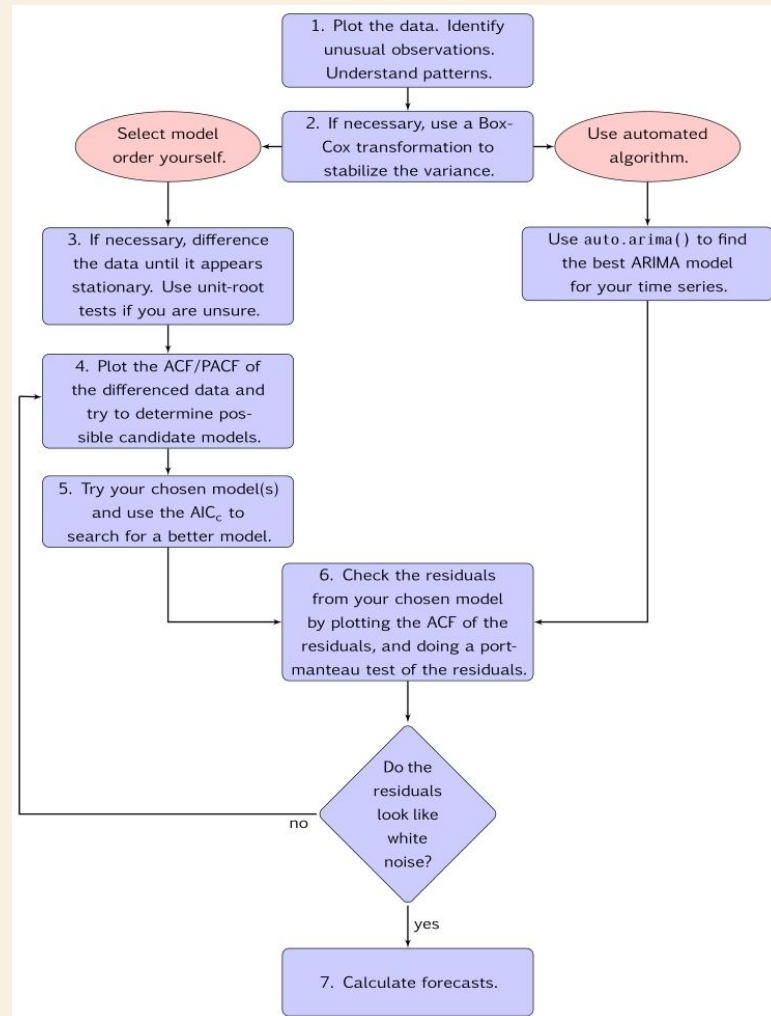
Index class	Grid	Traffic						City Background							
		core pollutants			pollutants			core pollutants				pollutants			
		NO2	PM10		PM2.5		CO	NO2	PM10		O3	PM2.5		CO	SO2
			1-h.	24-h.	1-h.	24-h.			1-h.	24-h.		1-h.	24-h.		
Very low	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	25	50	25	15	15	10	5000	50	25	15	60	15	10	5000	50
Low	25	50	25	15	15	10	5000	50	26	15	60	15	10	5000	50
	50	100	50	30	30	20	7500	100	50	30	120	30	20	7500	100
Medium	50	100	50	30	30	20	7500	100	50	30	120	30	20	7500	100
	75	200	90	50	55	30	10000	200	90	50	180	55	30	10000	350
High	75	200	90	50	55	30	10000	200	90	50	180	55	30	10000	350
	100	400	180	100	110	60	20000	400	180	100	240	110	60	20000	500
Very High*	> 100	> 400	>180	>100	> 110	>60	>20000	> 400	>180	>100	>240	> 110	>60	>20000	>500
NO ₂ , O ₃ , SO ₂ :		hourly value / maximum hourly value in µg/m ³													
CO		8 hours moving average / maximum 8 hours moving average in µg/m ³													
PM ₁₀		hourly value / daily value in µg/m ³													
* An index value above 100 is not calculated but reported as “ > 100”															

Data pre-processing

- Reformatting the structure of the dataset
 - We mainly use forecast, tidyverse, ggplot2, and xts libraries from R to perform our analysis
- Choice of the variables
 - Particles smaller than ten μm (PM10)
 - Ground-level Ozone (O₃)
 - Nitrogen dioxide (NO₂)

Forecasting

- ARIMA model
- Auto Arima in R
- Metric: Akaike Information Criterion (AIC) - Lower is better



ARIMA model

Auto Regression

A model that shows a changing variable that regresses on its own lagged

Integrated

The differencing of raw observations to allow for the time series to become stationary

Moving Average

The dependency between an observation and a residual error from a moving average model applied to lagged observations

ARIMA(p, d, q)(P, D, Q)_m

p

Number of
autoregressive
terms

d

Number of
differences

q

Number of
moving averages

P, D, Q

Seasonal
counterparts

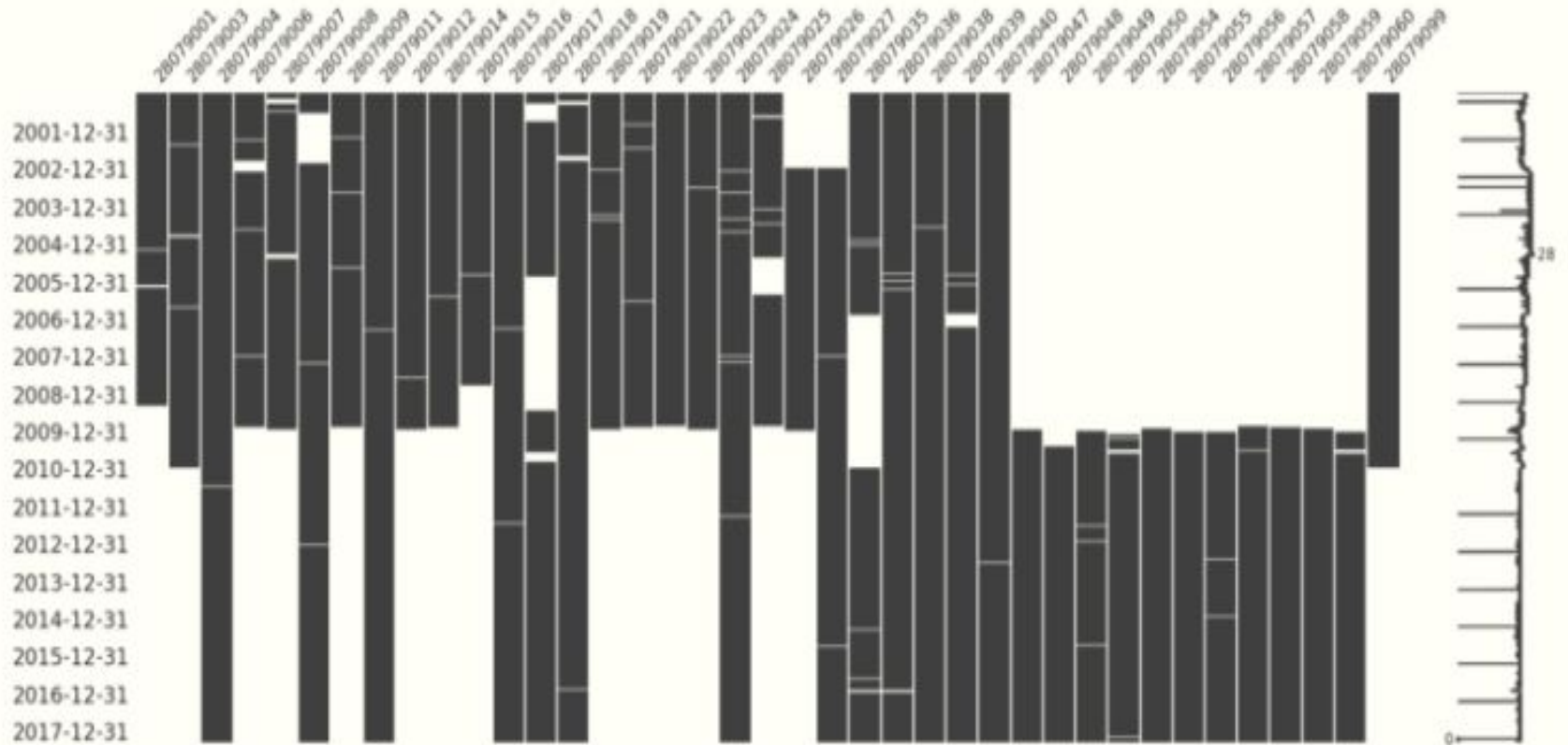
m

Number of
observations per
year

Results

- Missing data
- Explanatory analysis
- Forecasting

Missing data present in each station during the period 2001-2018.

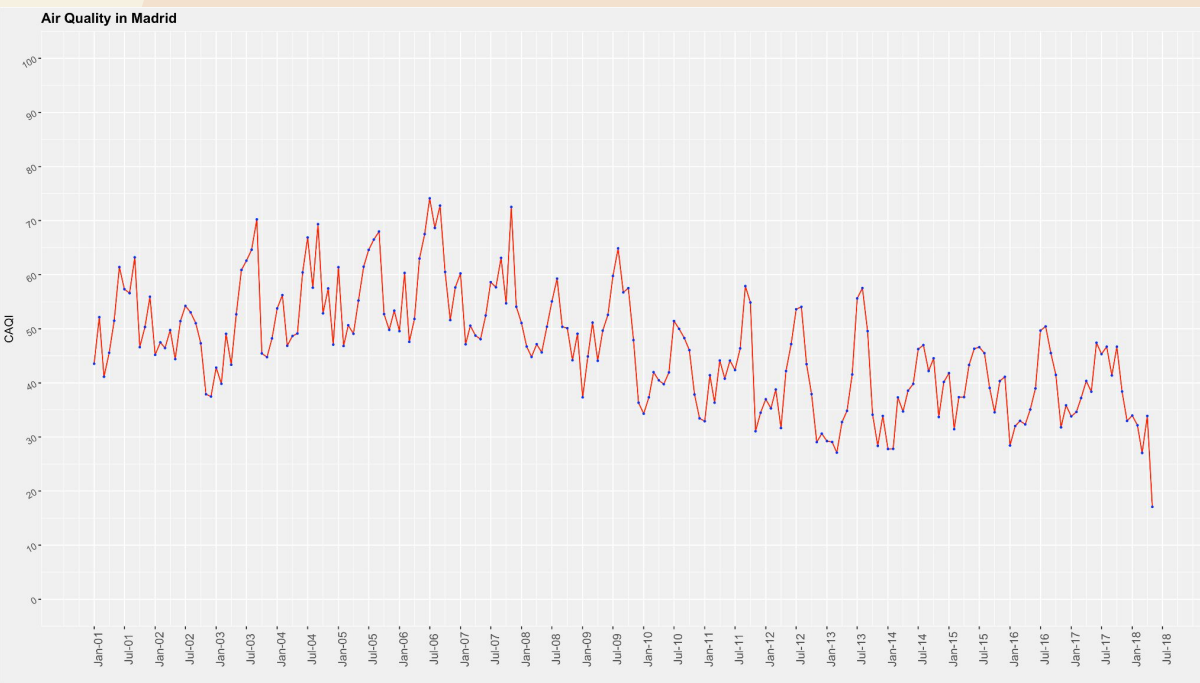


Missing data

- In 2009 almost half of the stations ceased all activity and a good portion of them started
- Only 6 stations span through the whole period
- In 2009 Madrid was breaking all contamination records and endured fierce criticism for its immobility
- At same year, the EU recommendations went from a friendly recommendation to being compulsory, and most of the limits were not being fulfilled

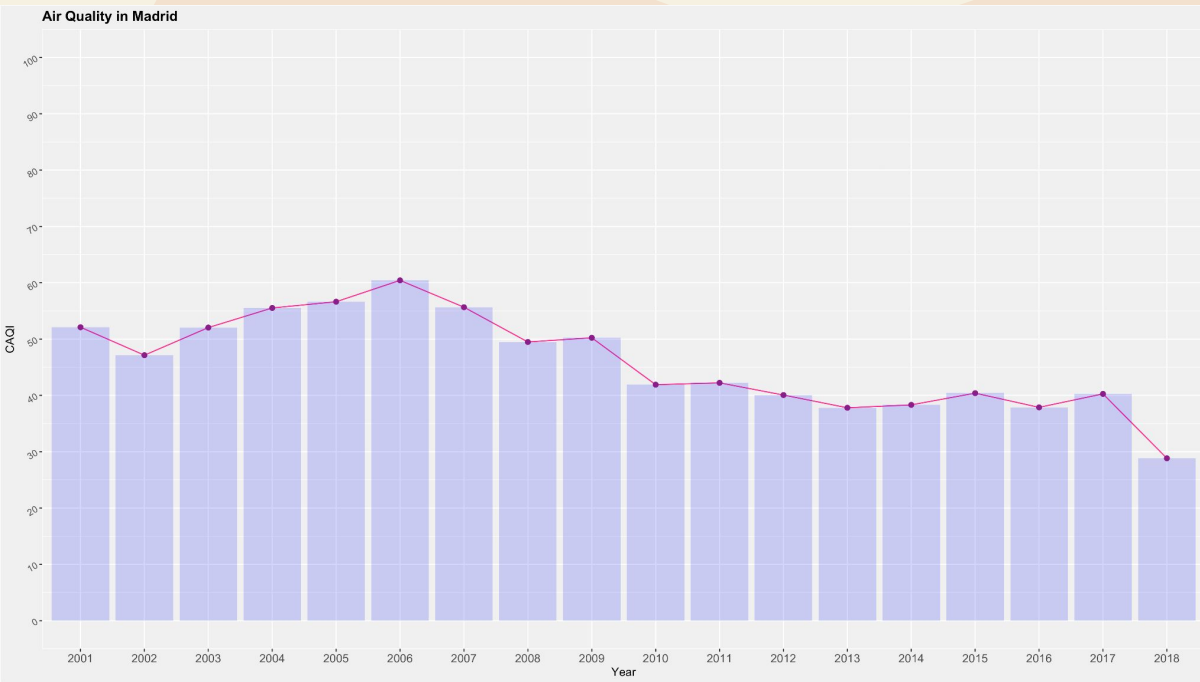
The background features a light beige color with large, flowing, wavy shapes in a slightly darker shade of beige. There are two small, stylized white cloud icons, one in the upper left and one in the lower right. On the right side, there is a circular area with a pattern of thin, parallel diagonal lines.

Explanatory Analysis and Forecasting



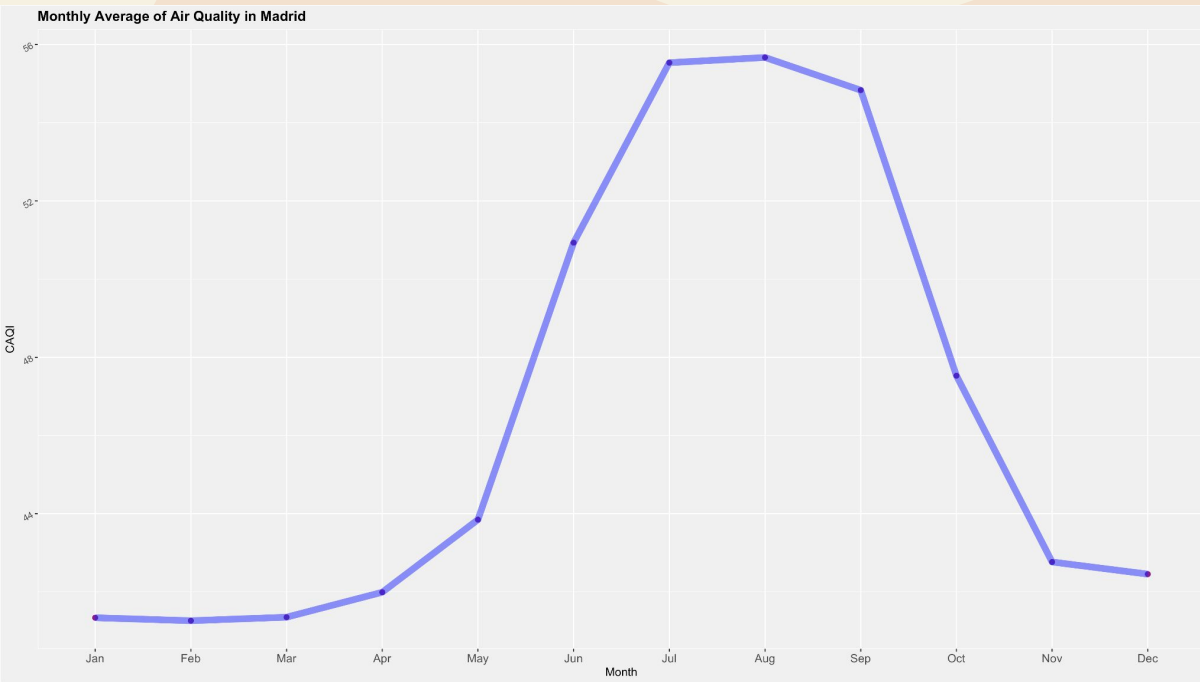
Monthly Moving Average of Madrid's Air Quality

- Increasing trend from 2001 to 2008
- Decreasing trend from 2009 to 2018
- Seasonality: Rise and fall of CAQI that regularly repeats over the same period



Yearly Moving Average of Madrid's Air Quality

- CAQI values are increasing from 52 to 60 on average, between 2001 and 2006
- Then, there had been a gradual decrease during 12 years
- Two stages on the slight drop:
2006-2008 (60 to 49)
and 2009-2017 (50 to 40)

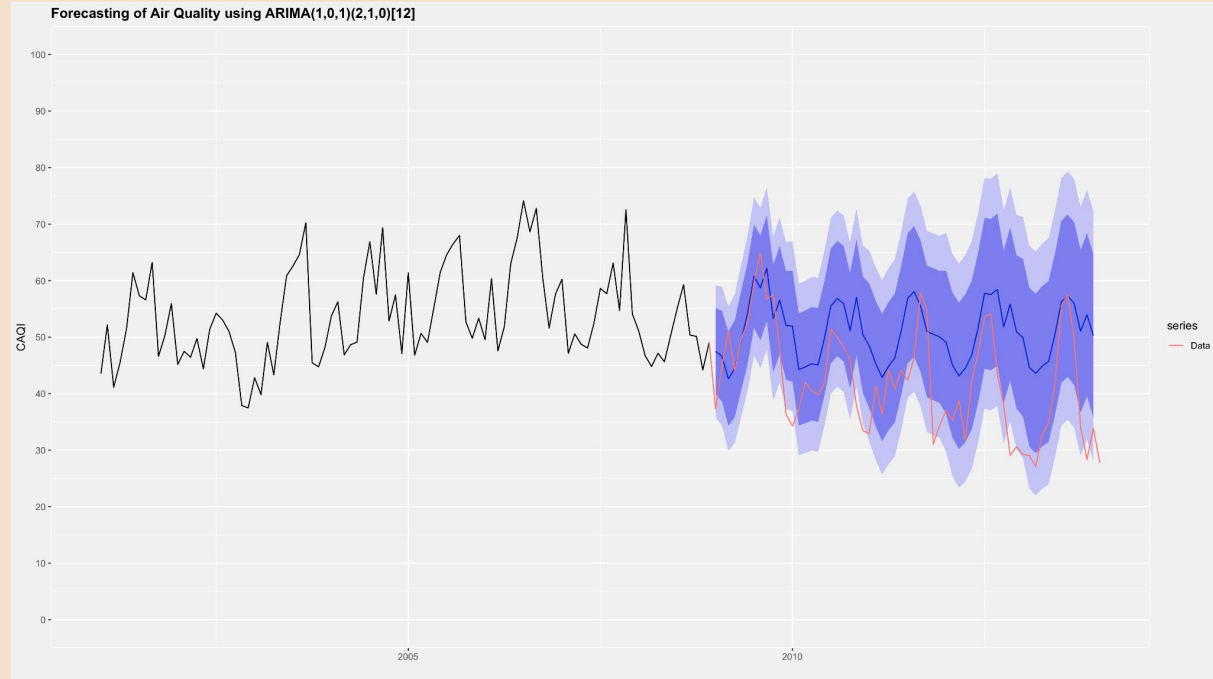


Monthly Average of Madrid's Air Quality

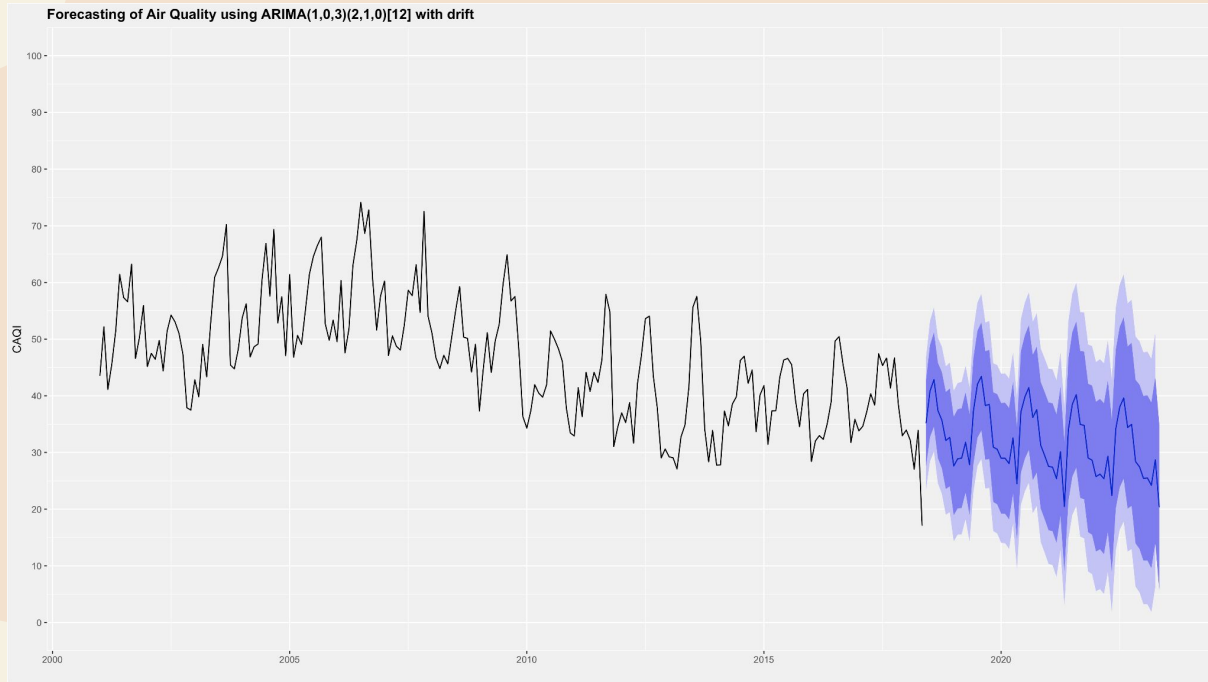
- The pollution increases in summer from the beginning of May and decreases in September and October
- In August, the contamination reaches its peak

Forecast of Madrid's Air Quality from 2009

- Decreasing trend
- The forecast line decreases slower than the actual data



Forecast of Madrid's Air Quality in next 5 years (2023)



- Decreasing trend
- Mean of around 30 CAQI
- Its peak will be around 41 CAQI

Discussion

- Explanatory analysis
- Forecast of Madrid's
Air Quality

Explanatory analysis

- In 2009, Madrid broke all contamination records -> EU air quality recommendations went from friendly to compulsory.
- Since then, Madrid invested in ads to promote electric cars and promulgated laws to reduce the CO emission in the city by limiting the number of cars and the speed limitation.

Explanatory analysis

- Madrid, like many other cities in Spain, is attractive for tourism during the high season. Hence, the increase in CAQI from June to September may be partly due to the population rising during the summer vacation.

Forecast of Madrid's Air Quality




- The developments in EU environment policy during 2009 and the Madrid environment plan have been efficient in improving air quality.
- Data reliability of Madrid Air Quality.

Conclusion

- The air pollution in Madrid was climbing to danger zone until 2009.
- If the EU did not react and forced Madrid's government to change, based on our forecast, the air quality would be terrible.
- From our forecast, it is shown that the pollution in Madrid is heading to the green zone.

Thank You For Your Attention



CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik**