

## OVERVIEW

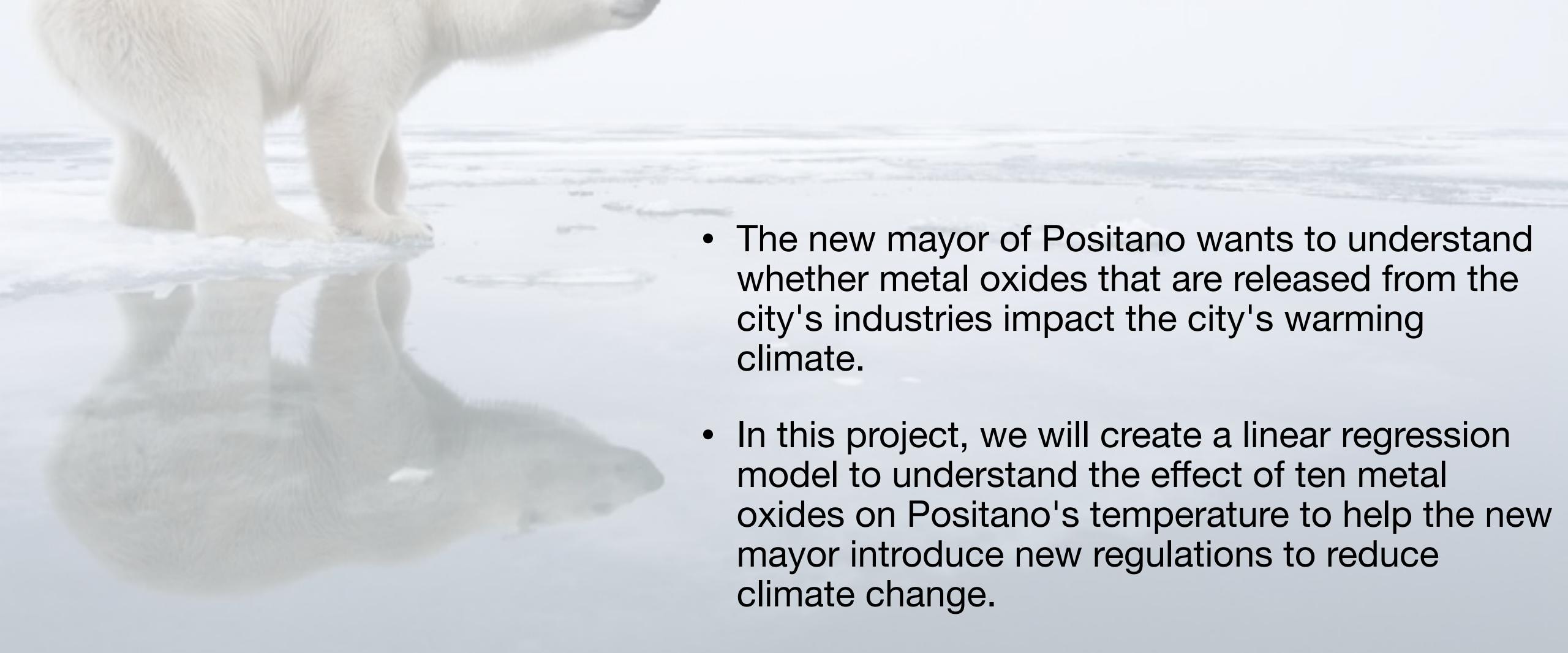


# OUTLINE



- Business problem
- Data
- Methods
- Results
- Conclusions

## BUSINESS PROBLEM



### BUSINESS PROBLEM



#### **Air Quality Data Set**

The dataset contains 9358 instances of hourly averaged responses from an array of 5 metal oxide chemical sensors embedded in an Air Quality Chemical Multisensor Device.

Data were recorded from March 2004 to February 2005.

Metal oxides studied include:

- CO
- Tin oxide
- Non Metanic HydroCarbons
- Benzene
- Titania
- NOx
- Tungsten oxide (nominally NO2 & NOx targeted)
- NO2
- Indium oxide

## METHODS

- 1. Data exploration:
  - How do variables change over time?
  - How did the average of each variable change between 2014 and 2015?
- 2. Linear regression model:
  - Iteration process to improve the model by studying multicollinearity and interactions.
  - Model validation using train and testing subset.

## RESULTS

### Change of variables over time

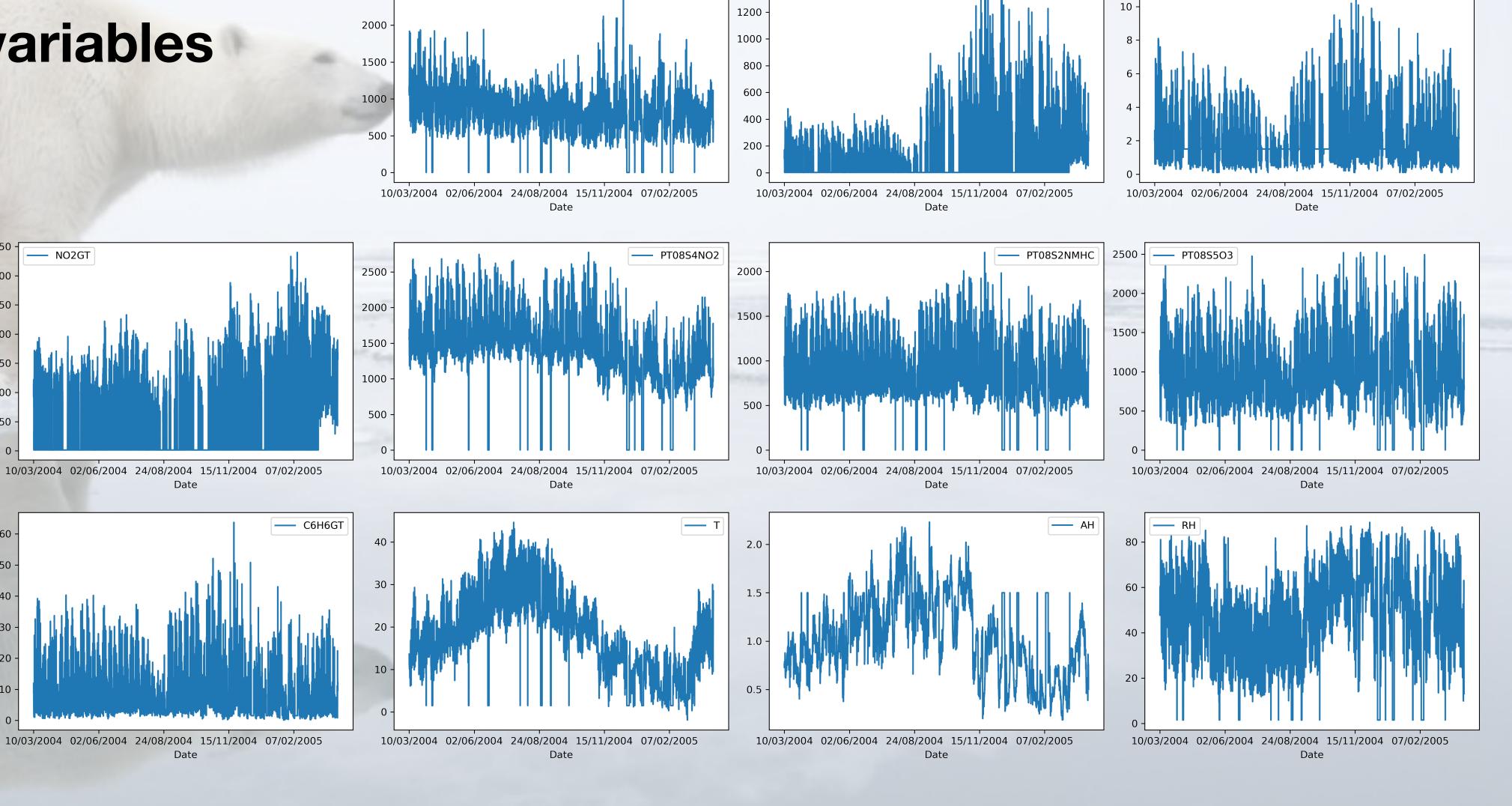
— NMHCGT

10/03/2004 02/06/2004 24/08/2004 15/11/2004 07/02/2005

10/03/2004 02/06/2004 24/08/2004 15/11/2004 07/02/2005

NO2GT

Date



— PT08S3NOx

2500

# RESULTS

#### Average differences between 2014 and 2015

| Variable                            | Average 2014 | Average 2015 |
|-------------------------------------|--------------|--------------|
| CO(GT)                              | 2.036723     | 2.030975     |
| Tin oxide - PT08.S1(CO)             | 1061.597046  | 1041.919893  |
| Non Metanic HydroCarbons - NMHC(GT) | 29.435724    | 1.500000     |
| Benzene - C6H6(GT)                  | 10.338383    | 7.877303     |
| Titania - PT08.S2(NMHC)             | 930.068214   | 815.172230   |
| NOx - NOx(GT)                       | 170.898312   | 308.375834   |
| tungsten oxide - PT08.S3(NOx)       | 828.916174   | 720.461949   |
| NO2 - NO2(GT)                       | 79.611744    | 137.482198   |
| Tungsten oxide - PT08.S4(NO2)       | 1502.120956  | 1074.207388  |
| indium oxide - PT08.S5(O3)          | 980.447117   | 990.885180   |
| Temperature - (T)                   | 20.240605    | 9.494393     |
| Relative humidty (RH) - PT08S5O3    | 46.753657    | 49.308055    |
| Absolute humidty (AH)               | 1.151260     | 0.704977     |

# RESULTS

#### Linear Regression Model

Final R-squared = 0.668 (Initial R-squared = 0.653)

Variables that positively impact Temperature

| Variable     | Coefficient |  |
|--------------|-------------|--|
| C6H6(GT)     | 11.09684    |  |
| NO2(GT)      | 1.925704    |  |
| PT08.S4(NO2) | 6.673217    |  |

# Variables that negatively impact Temperature

| Variable       | Coefficient |  |
|----------------|-------------|--|
| CO(GT)         | -0.561600   |  |
| PT08.S1(CO)    | -4.501010   |  |
| PT08.S2(NMHC)) | -5.678621   |  |
| NOx(GT)        | -2.7477100  |  |
| PT08.S3(NOx)   | -3.156382   |  |
| PT08.S5(O3)    | -5.449668   |  |

## CONCLUSIONS



All the metal oxides studied in this model significantly affect T in Positano.

Benzene, NO2 and tungsten oxide (nominally NO2 targeted) positively contribute to T raise

Benzene is the one with the highest positive impact.

CO, tin oxide, Titania (PT08.S2), NOx(GT), indium oxide, and tungsten oxide (nominally 03 targeted) negatively contribute to T

Titania is the one with the highest negative impact.

### BUSINESS RECOMMENDATIONS



- 1. Check the Temperature Sensor and repeat the study with newly collected data.
- 2. Investigate why the levels of NOx almost double in 2015.
- 3. In terms of regulation, since
  Benzene is the one that contributes
  the most to Temperature increase,
  we recommend that the new
  regulations aim to reduce the
  emissions of Benzene.

