Inter and Intra Regional Analysis of Air Quality Index

Final Project Presentation
Car Barn - 203, 5100 - Section 4
06:30 PM : November 29, 2023

Sheeba Moghal, Tiana Le
Pranav Patil, Dheeraj Oruganty

Contents

Introduction

Introduction

What:

Air Quality Index is used to provide warnings to citizens about the intensities of air pollutions

Why:

1. Causes of health ailments and premature death in the world as WHO estimates that more than 3 million people die every year due to air pollutants
2. EPA categorizes NO_2 as one of the air pollutants
3. NO_2 is emitted by industrial facilities and cars, contributing majorly to lung-related issues such as asthma, and other respiratory symptoms like coughing and wheezing.

[8] Who Health Organization. (2016). Ambient air pollution - world health organization. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/250141/9789241511353-eng.pdf?sequence=1
[3] Environmental Protection Agency. (n.d). Criteria Air Pollutants. U.S. Environmental Protection Agency. https://www.epa.gov/criteria-air-pollutants

Objective

Objective

- To assess the differences in air quality index between the European Region and Southeast Asia Region in 2019.
- To evaluate the change in air quality within France and India separately.
- To assess if global compliance with NO_2 is dependent on year.

Data Exploration

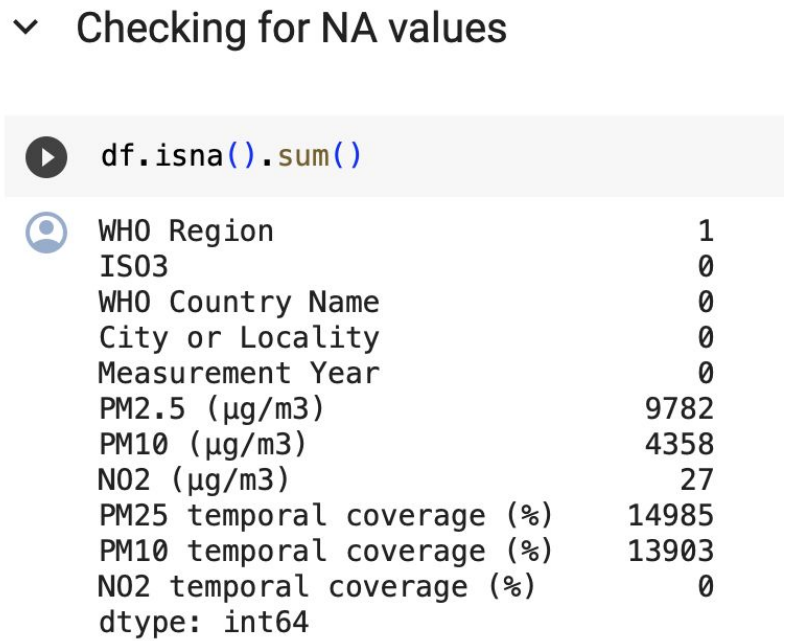
Description of the Collected Data:

```
[ ] df.describe()
```

	Measurement Year	PM2.5 (µg/m3)	PM10 (µg/m3)	NO2 (µg/m3)	PM25 temporal coverage (%)	PM10 temporal coverage (%)	NO2 temporal coverage (%)
count	16857.000000	7075.000000	12499.000000	16830.000000	1872.000000	2954.000000	16857.000000
mean	2015.602717	14.485423	26.477121	20.415367	93.344188	96.279787	96.794899
std	2.720185	9.302019	22.313548	11.648195	12.746337	4.736524	2.649329
min	2010.000000	0.010000	2.320000	0.000000	8.653846	33.333333	90.011000
25%	2014.000000	9.150000	16.780000	11.940000	93.694580	95.000000	94.927667
50%	2016.000000	12.530000	21.010000	18.730000	97.539851	97.808000	97.328767
75%	2018.000000	17.095000	27.400000	27.067500	99.000000	99.417000	99.132000
max	2021.000000	121.000000	322.000000	107.000000	100.000000	100.000000	100.000000

[1] Air quality database. (2023, May 15). Retrieved from https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database.

Checking for Missing Values:



[1] Air quality database. (2023, May 15). Retrieved from https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database.

Final Merged Data:

Merging the data

```
[ ] merged_df
```

	WHO Region	ISO3	WHO Country Name	City or Locality	Measurement Year	NO2 (µg/m3)	NO2 temporal coverage (%)	Population
0	European Region	ALB	Albania	Elbasan	2015	23.96	97.853881	2880703.0
1	European Region	ALB	Albania	Elbasan	2016	26.26	96.049636	2876101.0
2	European Region	AND	Andorra	Escaldes-Engordany	2012	31.64	100.000000	71013.0
3	European Region	AND	Andorra	Escaldes-Engordany	2014	27.62	100.000000	71621.0
4	European Region	AND	Andorra	Escaldes-Engordany	2015	26.65	94.554795	71746.0
...	...	...	...	...	...	...	...	...
16387	South East Asia Region	THA	Thailand	Rayong	2019	21.00	90.400000	71307763.0
16388	South East Asia Region	THA	Thailand	Samut Sakhon	2019	35.50	93.000000	71307763.0
16389	South East Asia Region	THA	Thailand	Saraburi	2019	32.00	92.000000	71307763.0
16390	South East Asia Region	THA	Thailand	Satun	2019	6.00	95.000000	71307763.0
16391	South East Asia Region	THA	Thailand	Tak	2019	11.00	93.000000	71307763.0

16392 rows × 8 columns

[1]. Air quality database. (2023, May 15). Retrieved from https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database.

Addition of Average No2

Final Dataframe

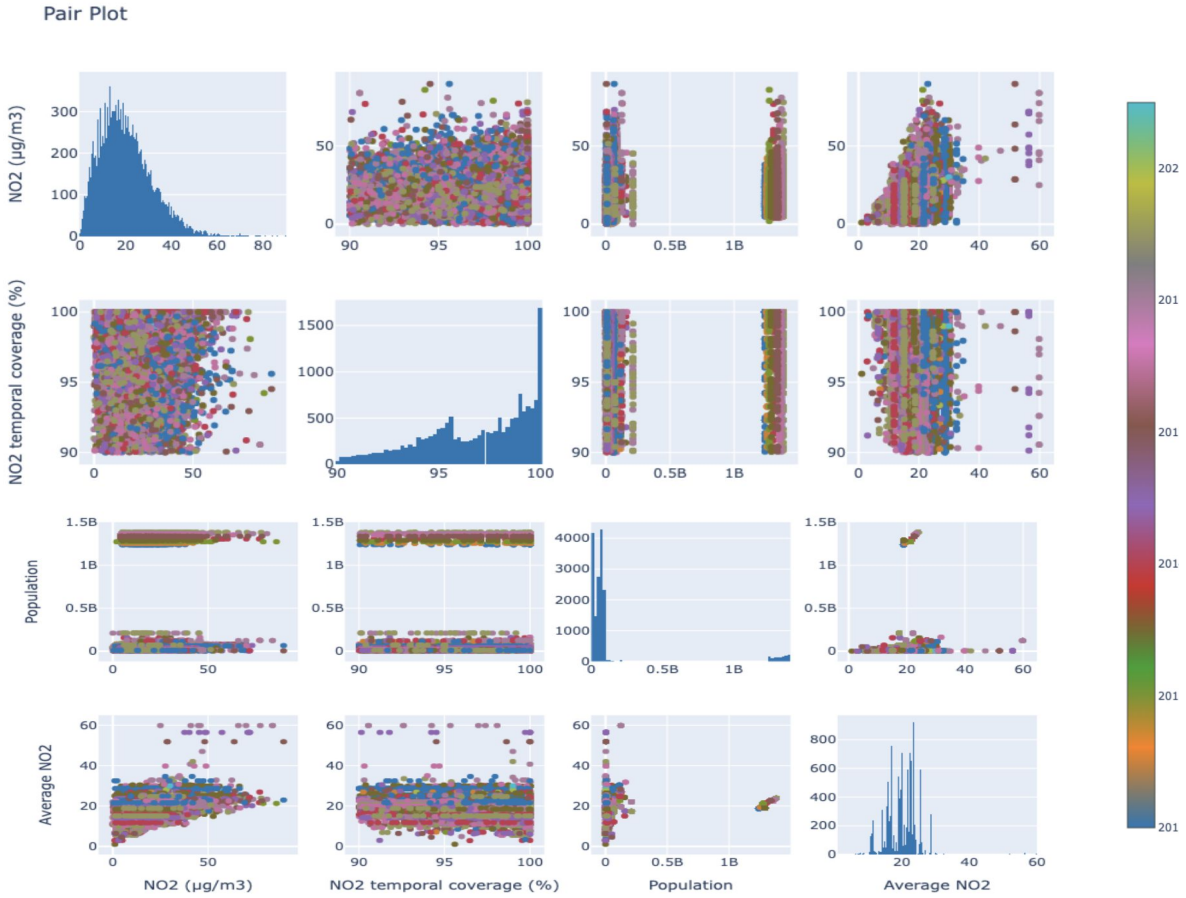
```
[ ] merged_df
```

	WHO Region	ISO3	WHO Country Name	City or Locality	Measurement Year	NO2 (µg/m3)	NO2 temporal coverage (%)	Population	Average NO2
0	European Region	ALB	Albania	Elbasan	2015	23.96	97.853881	2880703.0	23.960
1	European Region	ALB	Albania	Elbasan	2016	26.26	96.049636	2876101.0	26.260
2	European Region	AND	Andorra	Escaldes-Engordany	2012	31.64	100.000000	71013.0	31.640
3	European Region	AND	Andorra	Escaldes-Engordany	2014	27.62	100.000000	71621.0	27.620
4	European Region	AND	Andorra	Escaldes-Engordany	2015	26.65	94.554795	71746.0	26.650
...	...	...	...	...	...	...	...	...	...
16359	South East Asia Region	THA	Thailand	Rayong	2019	21.00	90.400000	71307763.0	18.828
16360	South East Asia Region	THA	Thailand	Samut Sakhon	2019	35.50	93.000000	71307763.0	18.828
16361	South East Asia Region	THA	Thailand	Saraburi	2019	32.00	92.000000	71307763.0	18.828
16362	South East Asia Region	THA	Thailand	Satun	2019	6.00	95.000000	71307763.0	18.828
16363	South East Asia Region	THA	Thailand	Tak	2019	11.00	93.000000	71307763.0	18.828

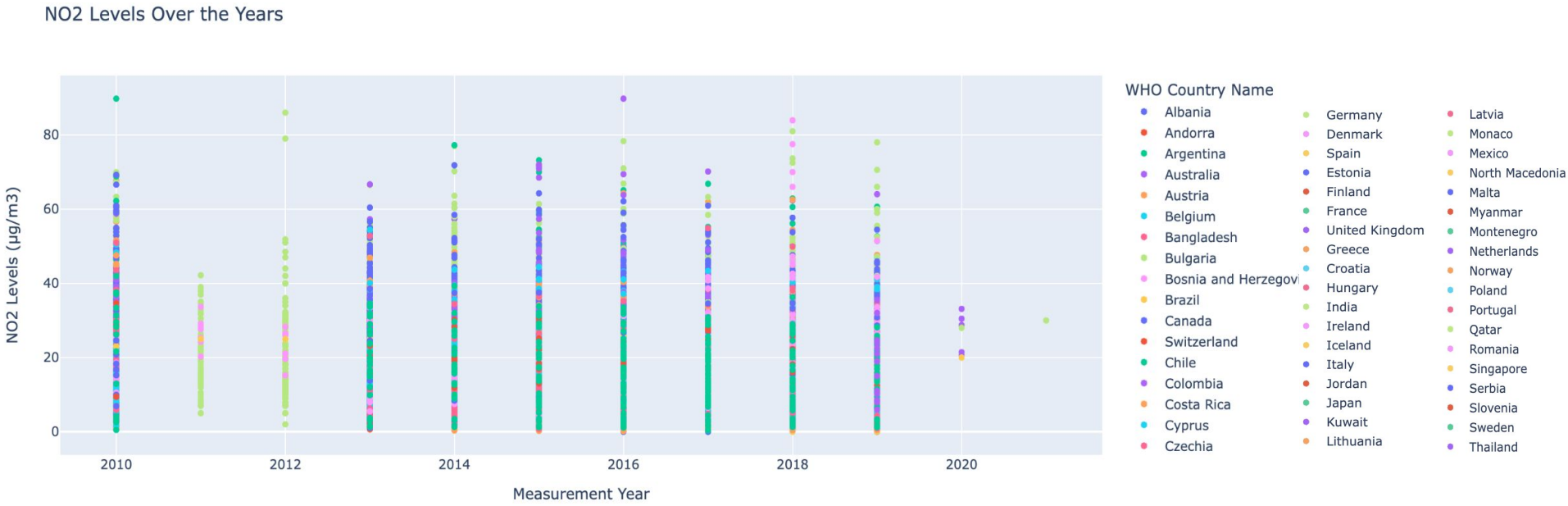
16364 rows × 9 columns

Exploratory Analysis

Exploratory Data Analysis: Pairplot of the Variables

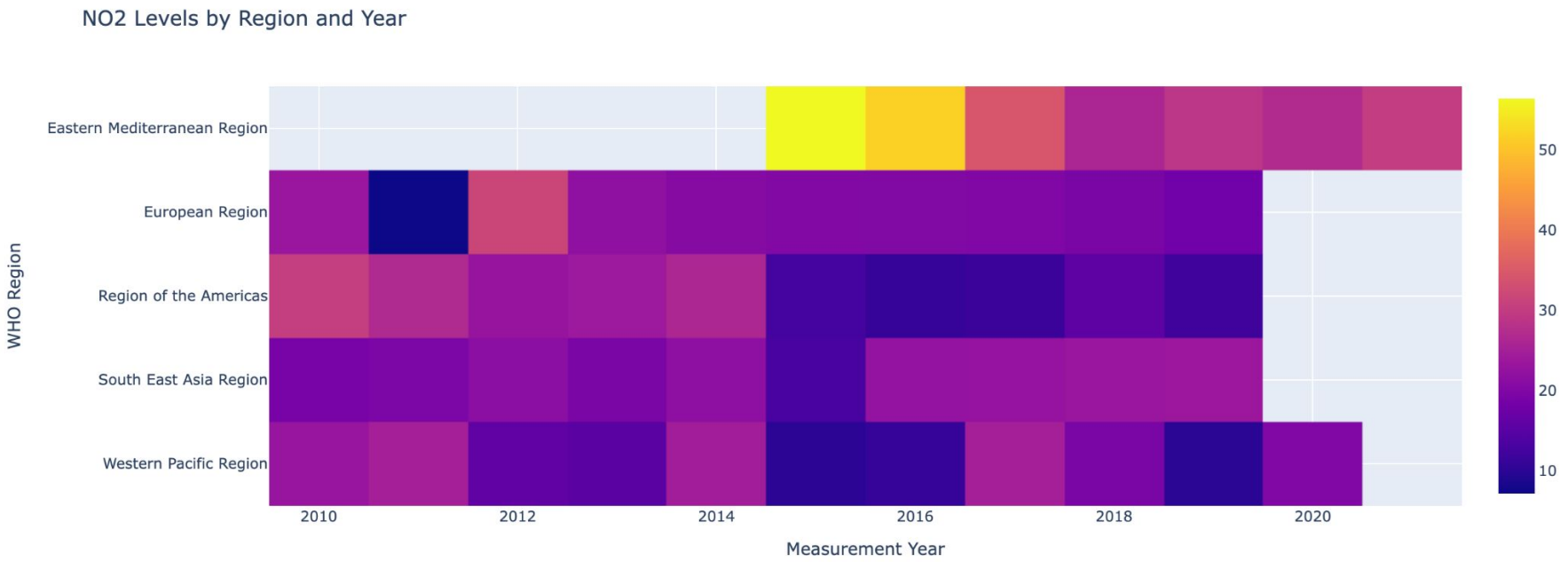


Exploratory Data Analysis: NO2 Levels Over the Years



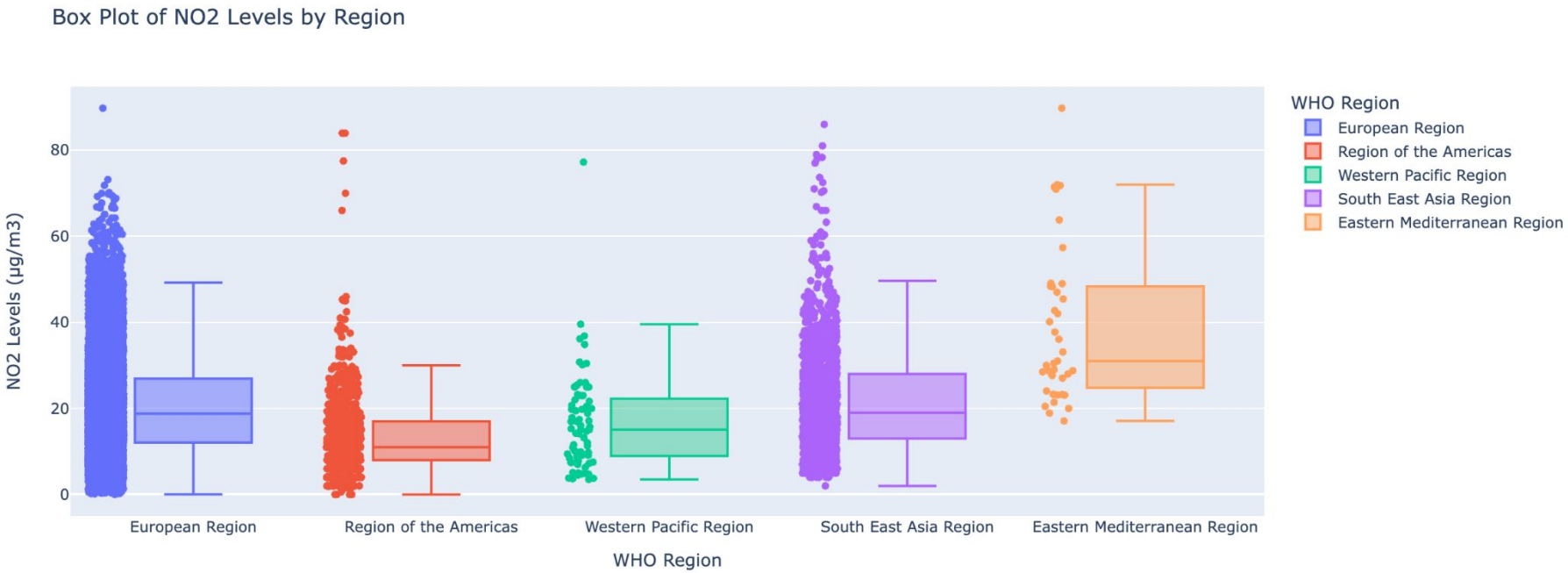
[4] Plotly Technologies Inc. (2015). Collaborative data science. Montréal, QC: Plotly Technologies Inc. Retrieved from https://plot.ly

Exploratory Data Analysis: NO2 Levels by Region and Year



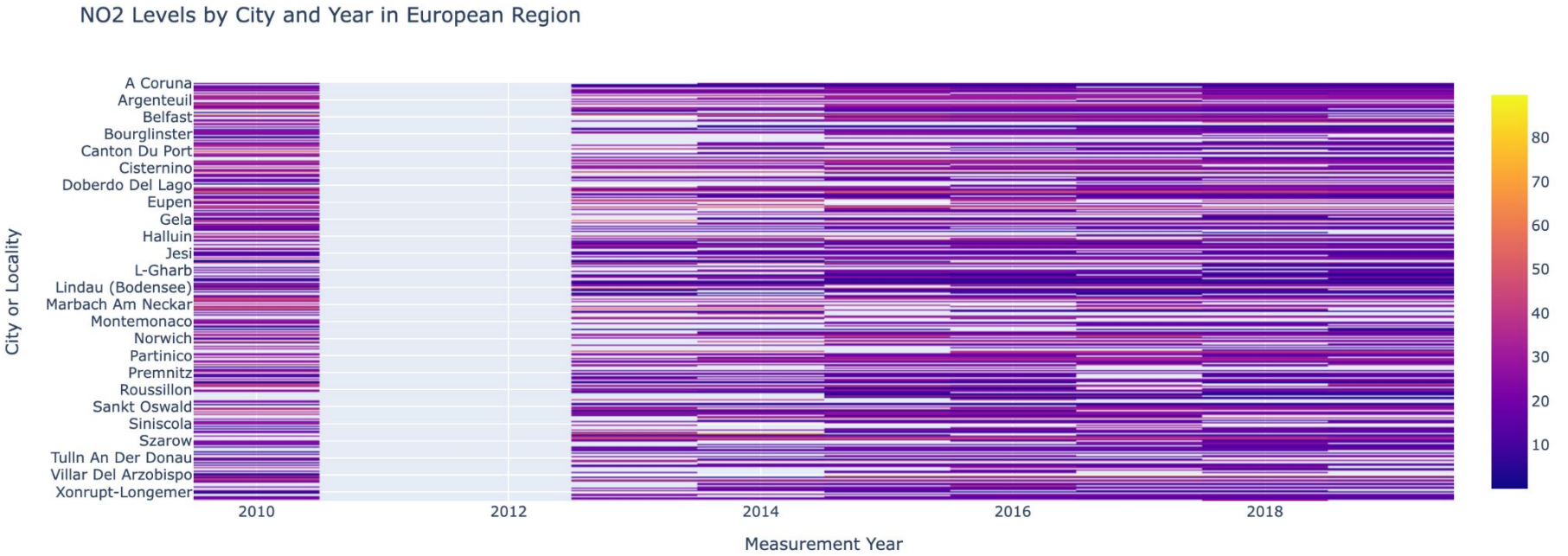
[4] Plotly Technologies Inc. (2015). Collaborative data science. Montréal, QC: Plotly Technologies Inc. Retrieved from https://plot.ly

Exploratory Data Analysis: Boxplots of NO2 Levels by Region



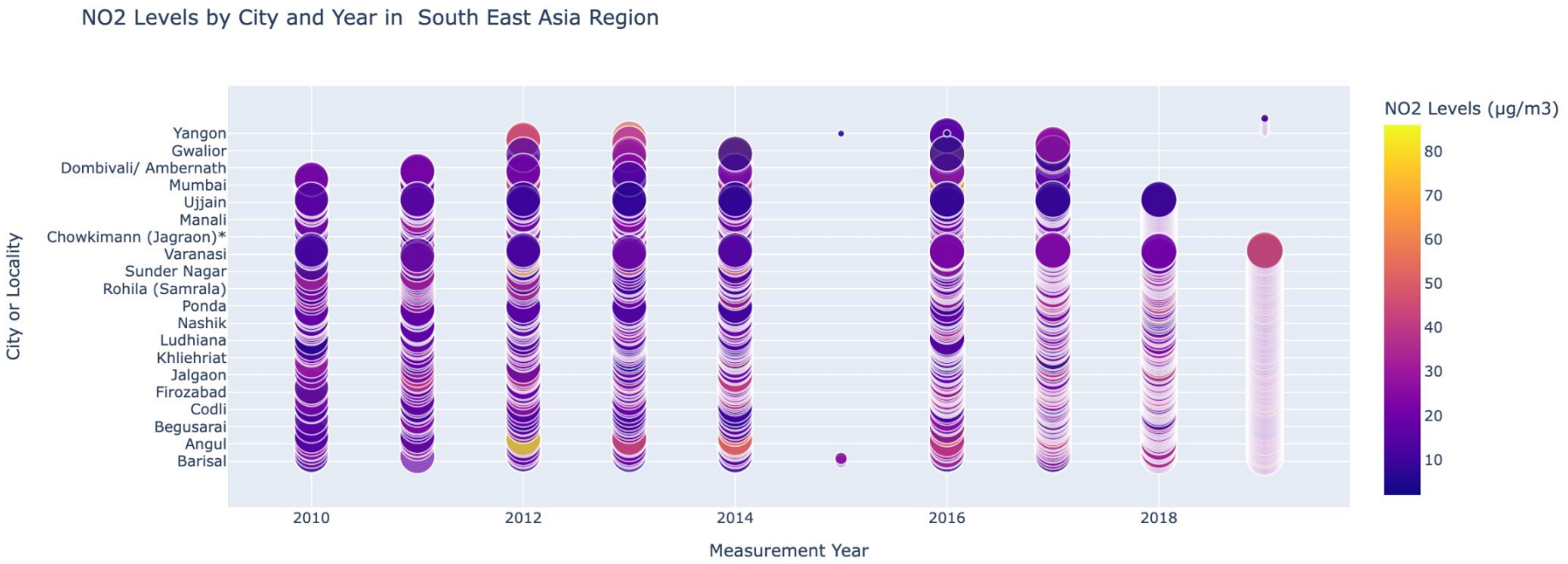
[4] Plotly Technologies Inc. (2015). Collaborative data science. Montréal, QC: Plotly Technologies Inc. Retrieved from https://plot.ly

Exploratory Data Analysis: NO2 Levels by City and Year in ER



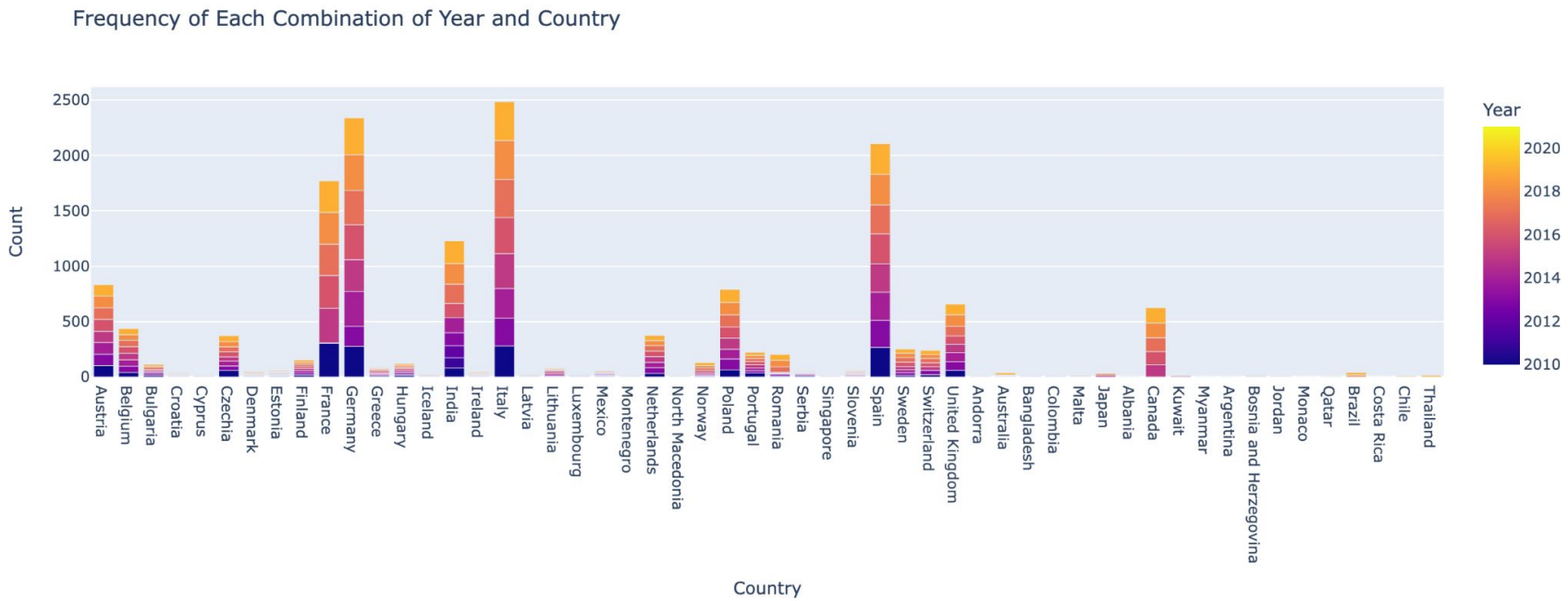
[4] Plotly Technologies Inc. (2015). Collaborative data science. Montréal, QC: Plotly Technologies Inc. Retrieved from https://plot.ly

Exploratory Data Analysis: NO2 Levels by City and Year in SEAR



[4] Plotly Technologies Inc. (2015). Collaborative data science. Montréal, QC: Plotly Technologies Inc. Retrieved from https://plot.ly

Exploratory Data Analysis: Frequency of Each Combination of Year and Country



[4] Plotly Technologies Inc. (2015). Collaborative data science. Montréal, QC: Plotly Technologies Inc. Retrieved from https://plot.ly

Statistical Testing

Does South East Asia Region have more NO2 emissions than the European Region in 2019?

Method used:

Welch Two Sample t-test

Hypothesis:

Ho: The NO2 emissions in the South East Asia Region and European Region are similar.

H1: The NO2 emissions in the South East Asia Region is greater than in the European Region

```
    Welch Two Sample t-test

data:  sea2 and eur2
t = 6.1356, df = 246.22, p-value = 1.682e-09
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 4.023139      Inf
sample estimates:
mean of x mean of y
 23.53045  18.02611
```

[7] T.test: Student's t-test. RDocumentation. (n.d.). https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/t.test

Does France have less NO_2 emissions in 2016 than in 2010?

Method used:

Two-sample bootstrap test for the ratio of means of NO_2

Hypothesis:

Ho: France did not decrease NO_2 emissions in 2016 compared to 2010.

H1: France did decrease NO_2 emissions in 2016 compared to 2010.

```
    Welch Two Sample t-test

data:  france2016 and france2010
t = -4.3778, df = 597.82, p-value = 7.075e-06
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
      -Inf -2.275993
sample estimates:
mean of x mean of y
 19.30578  22.95500
```

2.5%: 0.787455957026454 **97.5%:** 1.83886516406054

```
[1] "The estimated bias is 0.033"
```

[7] T.test: Student's t-test. RDocumentation. (n.d.). https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/t.test

Does India have less NO_2 emissions in 2018 than in 2014?

Method used:

T-test

Hypothesis:

Ho: India did not decrease NO_2 emissions in 2018 compared to 2014.

H1: India did decrease NO_2 emissions in 2018 compared to 2014.

```
    Welch Two Sample t-test

data:  india2018 and india2014
t = 1.0464, df = 268.02, p-value = 0.8518
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf 4.302835
sample estimates:
mean of x mean of y
 23.35289  21.68338
```

[7] T.test: Student's t-test. RDocumentation. (n.d.). https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/t.test

Is Global No2 dependant on the year?

Method used:

Chi-Square Test of Independence

Hypothesis:

Ho: The compliance with recommended NO2 are independent from years. The compliance with recommended NO2 does not differ between years.

H1: The compliance with recommended NO2 are dependent on the year. The compliance with recommended NO2 differs between the years.

	above	within
2013	119	1347
2014	111	1520
2015	103	1869
2016	118	2056
2017	104	2149
2018	91	2317
2019	53	2409

```
    Pearson's Chi-squared test

data:  table_year
X-squared = 94.367, df = 6, p-value < 2.2e-16
```

[3] Jarvis DJ, Adamkiewicz G, Heroux ME, et al. Nitrogen dioxide. In: WHO Guidelines for Indoor Air Quality: Selected Pollutants. Geneva: World Health Organization; 2010. 5. Available from: https://www.ncbi.nlm.nih.gov/books/NBK138707/

[6] RDocumentation. (n.d.-a). Chisq.test: Pearson's chi-squared test for count data. Retrieved from https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/chisq.test

Conclusion

Conclusion

- **The NO_2 in South East Region is greater than in the European Region.**
- **Using NO_2 as a parameter to assess air quality,**
  - Air quality in France improved in 2016 when compared to 2010.
  - Air quality in India did not improve in 2018 compared to 2014.
- **Yes, global compliance with NO_2 is dependent on year.**

References

1. Air quality database. (2023). Retrieved from https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database
2. Education, U. C. for S. (n.d.). Center for Science Education. Nitrogen Dioxide molecule | Center for Science Education. Retrieved from https://scied.ucar.edu/image/nitrogen-dioxide-molecule
3. Environmental Protection Agency. (n.d). Criteria Air Pollutants. U.S. Environmental Protection Agency. https://www.epa.gov/criteria-air-pollutants
4. Jarvis DJ, Adamkiewicz G, Heroux ME, et al. (2010). Nitrogen dioxide. In: WHO Guidelines for Indoor Air Quality: Selected Pollutants. Geneva: World Health Organization. Available from: https://www.ncbi.nlm.nih.gov/books/NBK138707/
5. Plotly Technologies Inc. (2015). Collaborative data science. Montréal, QC: Plotly Technologies Inc. Retrieved from https://plot.ly
6. R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/
7. RDocumentation. (n.d.-a). Chisq.test: Pearson's chi-squared test for count data. Retrieved from https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/chisq.test
8. RDocumentation. (n.d.). T.test: Student's t-test. Retrieved from https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/t.test
9. Who Health Organization. (2016). Ambient air pollution - world health organization. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/250141/9789241511353-eng.pdf?sequence=1

Questions & Answers