Data Analytics Tools Final Project

Marc Esteller, Oscar Grau, Clara Jimenez & Luke Perreault

2023-01-25

Abstract

One of the main concerns for the citizens of Barcelona is the pollution of the city. For that main reason, this paper analyzes the different types of pollution in Barcelona and how the pollution levels of them change during the year 2022.

In addition to that, we have analyzed if the quantity of trees in the city is making any difference on the pollution levels of the different districts and neighborhoods.

Finally, we will analyze the pollution level of the different neighborhoods of the city and see which of them are less and more polluted on the last year.

In order to develop this analysis we performed a series of different joints/unions/merges between 6 different data sets to be able to work with the data in an efficient way. With that we were able to communicate and perform better Exploratory Data Analysis (EDA).

Analysis

Question 1

How are the pollution indexes evolving during the year? Is there any particular peak in any month, day or hour of the day that you consider concerning?

These questions will be answered in sections broken up by the type of contaminant. For each contaminant, a graph of the year, month, and days of the week will be examined to find trends and peaks or lows.

SO2

Sulfur Dioxide (SO2) is a toxic gas that comes from the burning of fossil fuels and can cause severe health issues. The European Environment Agency (EEA) which is a trusted group that deals with the environment reports that SO2 emissions should not exceed 20 $\mu g/m^3$ as an annual mean, hourly not exceed 350 $\mu g/m^3$ more then 24 times a year, and daily shouldn't exceed 125 $\mu g/m^3$ more than 3 times a year.

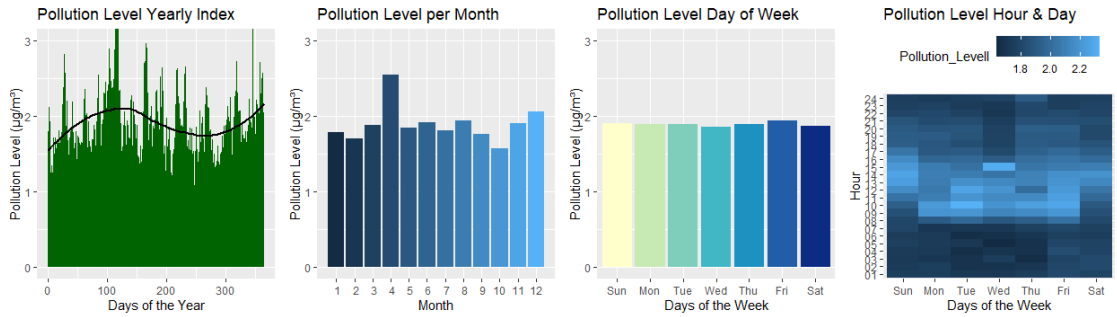

Figure 1. 2. 3. 4.

Trends

The index for SO2 has risen 15% from the beginning to end of the year. From Jan the pollutant level rose and peaked at April but then faulted to another low around October before having a significant increase up to December.

The peak month was **April** and the low was **October** which is interesting how they are opposite of each other in the year. There appears to be pretty steady trend within the days of the week, but **Friday** is the peak day for the pollutant.

Hourly it is observable in graph 4 that during the day is when the pollutant is mostly active especially during rush hours in the morning and at night. There is the highest percentage of contamination around **10am** on Tuesday and oddly at **3pm** on Wednesday.

It seems there is a relationship between daily traffic and this contaminant that is affecting by the burning of fossil fuels. More about this will be discussed at the end of question 1's analysis.

Stability

Throughout the year SO2 is pretty volatile with a daily standard deviation of 0.24 Which is a lot considering that the change from the beginning of the year to the end was ~0.5 $\mu g/m^3$.

Outliers

Yearly the average and any day in the year was no where near the EEA limit of 20 $\mu g/m^3$, so we wouldn't consider anything dangerous according to our sources. The peak month was by far April. The peak hour of the day was 10 am but the hours around it were very close as well. The hourly limit was 350 $\mu g/m^3$, and there was no hour out of every hour in the year close to it. The closest was 75 $\mu g/m^3$ on the **20th of July** at 3 pm.

CO

Carbon monoxide (CO) is an odorless pollutant that is emitted into the air by fossil fuels. It can be lethal at high levels and that level according to Cove says would need to exceed 157 $\mu g/m^3$ at any time.

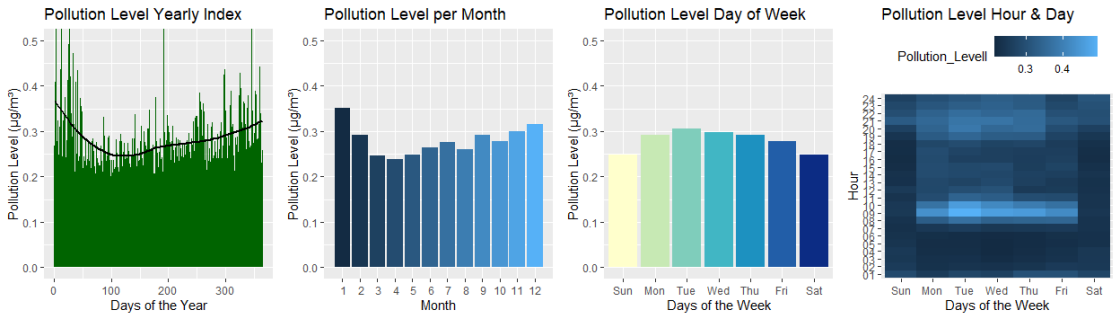

Figure 5. 6. 7. 8.

Trends

The index for CO has decreased -10% from the beginning to end of the year. From January the pollutant level decreased and bottomed out in April but then steadily increased till December.

The peak month was **January** and the lowest was **April** which is the opposite of S02.

There is a trend that begins low on Sunday and slightly rises to a peak on Tuesdays and slowly decreases to Saturday. The days of the week are higher then the weekend which is interesting. **Tuesday** is the peak day for the pollutant.

Hourly it is observable in graph 4 that during the days Monday through Friday is when the pollutant is mostly active especially during rush hours in the morning and at night. There is the highest percentage of contamination around **10am** on Tuesday similarly to SO2 but does seem as high during the middle of the day. Possibly the contaminant isn't affected as much by traffic or the lifespan of each particle is not as long as SO2.

Stability

Throughout the year CO is even more volatile with a daily standard deviation of 0.03 considering that the change from the beginning of the year to the end was ~0.05 µg/m^3.

Outliers

Because we couldn't find info on when this pollutant is considered dangerous or hazardous we cannot assume there is an issue with the levels despite several hours of the day, days of the year, and months having a higher level then others.

NO

Nitrogen Oxide (NO) is the first of three variations of NO we will be looking at. Essentially, they are all related and NOX is the "combination" of NO and NO2. NO is not harmful to humans in the same way the other contaminants we have discussed are. The worst it can do is make it harder to breath and cause fatigue. Two major the sources of the three nitrogens are combustion and lightning.

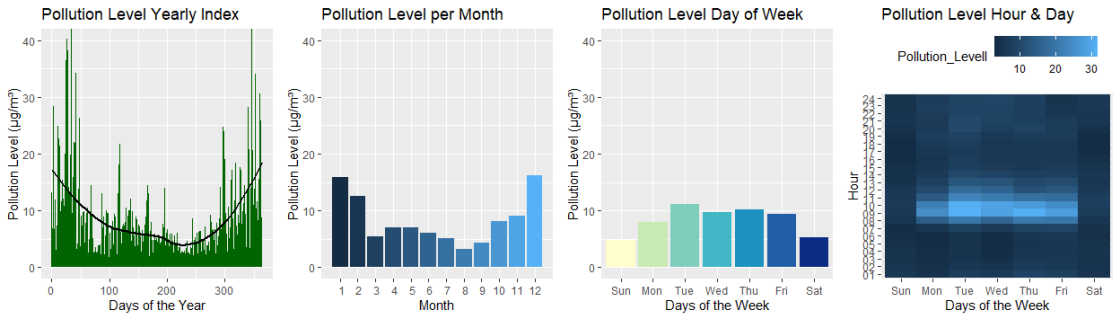

Figure 9. 10. 11. 12.

Trends

The index for NO has increased 2% from the beginning to end of the year. From Jan the pollutant level quickly decreased and began bottoming out in the summer and finally started rising again in August till peaking in December.

The peak month was again **December** with January very close and the lowest was **August**. The trend shows a large dip throughout the summer and the highest in the winter months.

There is the same trend for the days of the week as in the last contaminant which was starting low on Sunday and slightly rising to a peak on Tuesday and slowly decreasing to Saturday. The days of the week are higher then the weekend by far. **Tuesday** is the peak day for the pollutant. Hourly it is observable in graph 4 that during the days Monday through Friday is when the pollutant is mostly active during morning rush hours, but we see differently here that there is higher activity in the night from 7 pm to midnight. There is the highest percentage of contamination around **10am** on Tuesday similarly to SO2 and CO. Saturday and Sunday are very dry until the late hours of night which is very strange compared to the previous trends.

Stability

Throughout the year NO is extremely volatile with a daily standard deviation of 4.34 considering that the change from the beginning of the year to the end was ~1 µg/m^3.

Outliers

Because we couldn't find info on when this pollutant is considered dangerous or hazardous we cannot assume there is an issue with the levels despite several hours of the day, days of the year, and months having a higher level then others.

NO2

Nitrogen Dioxide (NO2) comes from the burning of fossil fuels, and according to the EEA the annual mean level should not surpass 40 µg/m^3. The hourly level should not surpass 200 µg/m^3 in any day.

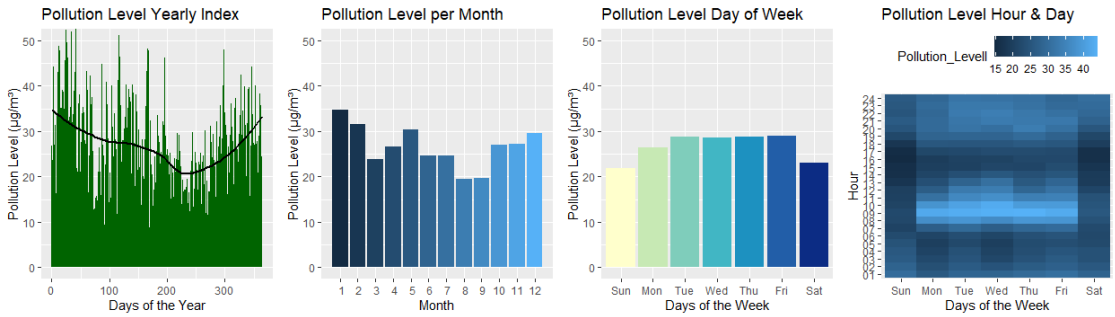

Figure 13. 14. 15. 16.

Trends

The index for NO2 decreased -15% from the beginning to end of the year. From Jan the pollutant level decreased and but took while to get near the bottom in the summer unlike NO. Eventually it hit bottom around August and increased till in December but wasn't able to break even with the beginning of the year.

The peak month was in **January** but February was close and the lowest by a lot were **August** and **September**. The trend shows a large dip throughout the end of the summer and the highest in the winter months again.

The days of the week are higher then the weekend by far. There isn't a very obvious peak but **Friday** and **Tuesday** peak pretty equally at the top. Hourly it is observable in graph 4 that during the days Monday through Friday is when the pollutant is very active during the morning rush hour at **9am**, and slightly active again in the night time. Everyday from around 3pm to 8pm there is a gap of the lowest part of the day then. Even the night has a higher level then the afternoon.

Stability

Throughout the year NO2 is extremely volatile with a daily standard deviation of 4.55 considering that the change from the beginning of the year to the end was ~5 µg/m^3.

Outliers

To stay under dangerous levels it is advised that the annual mean does not surpass 40 µg/m^3; other then a few days the average is not near the limit. The daily mean shouldn't pass 200 µg/m^3 and in Figure 13 we set any day equal to red to see if it would pass 200 but none did.

NOX

Aeroqual says NOX is the collective term used to refer to the two NO and NO2. So, it is affected by the same sources as Nitrogen.

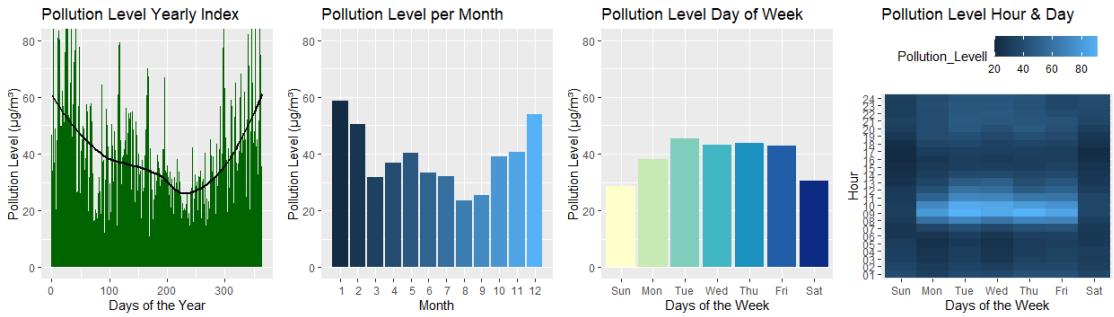

Figure 17. 18. 19. 20.

Trends

The index for NOX descreased -8% from the beginning to end of the year. Similarly to NO2 from January the pollutant level decreased and but took while to get near the bottom in the summer unlike NO. Eventually it hit bottom around August and increased till in December but wasn't able to break even with the beginning of the year.

The peak month was in **January** but February and December were close and the lowest again by a lot were **August** and **September**. The trend shows a large dip throughout the end of the summer and the highest in the winter months again.

The days of the week are higher then the weekend by even greater a difference than NO2. There is a peak on **Tuesday** which is the day that seems to be the trend for most the pollutants. Hourly it is observable in graph 4 that during the days Monday through Friday is when the pollutant is very active during the morning rush hours at **9am and 10am**, and there is a slight active time frame again in the night. There is noticeable gap again in the afternoon but in contrast to NO2 there is also one in the early morning.

Stability

Throughout the year NO is extremely volatile with a daily standard deviation of 10.87 considering that the change from the beginning of the year to the end was ~5 µg/m^3.

Outliers

We couldn't find any hazard limits for this contaminant specifically, but we will use SO2's limits as a substitute to get an idea. Using the annual limit of 40 µg/m^3, SOX doesn't pass it with its mean of 38.82µg/m^3. There is still no day that surpasses 200 µg/m^3.

PM2.5

Particular Matter 2.5 (PM2.5) comes from mainly combustion and burning of natural gas. By EPA standards PM 2.5 should not exceed 35 µg/m3 daily and 12 µg/m3 annually.

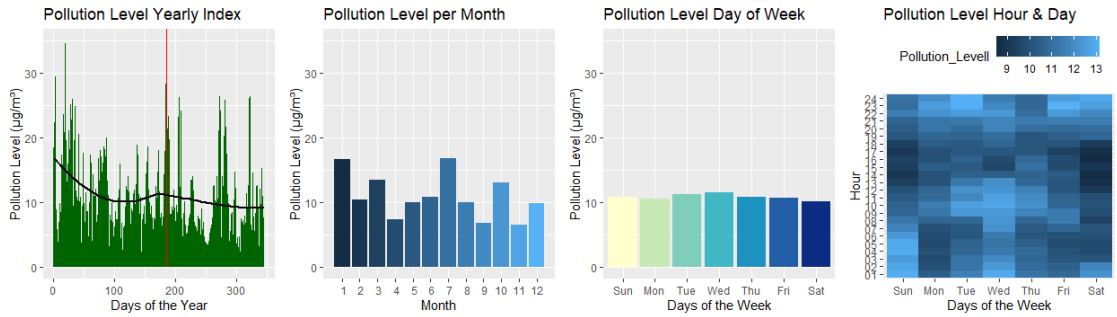

Figure 21. 22. 23. 24.

Trends

The index for PM2.5 significantly decreased -41% from the beginning to end of the year. From Jan the pollutant level decreased a lot initially and never regained a positive trend other then a rise in July. The peak month was in **January** and **July** and the lowest was **September** and **November** with April not too far behind. The days of the week have a very interesting trend. From Sunday to Wednesday there is a steady linear increase and after peaking on Wednesday there is the negative linear decline at the same velocity.

Hourly the graph has no obvious trend other then a chaotic one. There's a large level from **1 am to 7 am** on Sunday and the same type of level in the **middle of the day on Wednesday**. Furthermore, all the nights and early mornings are at a high level. In the afternoon on Saturday is the lowest time of day.

Stability

Throughout the year NO is volatile with a daily standard deviation of 3.44 considering that the change from the beginning of the year to the end was ~7 µg/m^3.

Outliers

The yearly standard set by the EPA for PM2.5 is to not exceed an average of 12 µg/m3; the yearly average we have is barely below it at 10.84µg/m3. PM2.5 is to not exceed a 35 µg/m3 daily average. Our data shows there is one instance of this on the **24th of July** when the daily average is 46 µg/m3. This is a dangerous day according to the EPA; the day is in align with the outlier from SO2 which was on July 20th.

PM10

Particular matter 10 (PM 10) are sourced from any type of stone crushing operation and can get into the air from a variety of sources such as storms, wildfires, sea spray, pollen, human emissions and other natural sources. The main difference from PM 2.5 to PM 10 is the particle size. PM 2.5 is smaller and PM10 can be made up of PM 2.5. PM10 should not exceed a 40 µg/m3 annual mean and should not exceed a 50 µg/m3 daily mean more then 35 times a year.

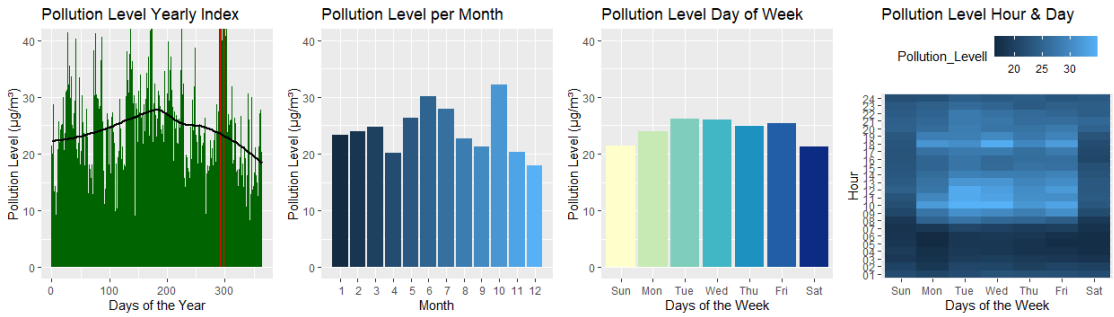

Figure 25. 26. 27. 28.

Trends

The index for PM2.5 significantly decreased -23% from the beginning to end of the year. This pollutant had the opposite trend of others and actually increased right away and peaks around June and then mainly decreasing till December with some super high days in October.

The peak month was **October** with June closely behind. The lowest was **December**. The days of the week have a trend similar to many of the others, however. The contaminant is highest during the week days and peaks on **Tuesday** with Wednesday and Friday close behind.

Hourly the graph has a lot of activity from 9am till 10 at night. The peak hour is around **10 am** most days. Wednesday similar to SO2 has a peak at **6 pm**.

Stability

Throughout the year NO is pretty volatile with a daily standard deviation of 4.28 considering that the change from the beginning of the year to the end was ~6 µg/m^3.

Outliers

According to the EPA PM10 is not supposed to exceed a mean of 40 µg/m3 annually; our data proves that the mean is safely below this standard at 24.23 µg/m^3. Furthermore PM10 should not pass a 50 µg/m^3 daily average more than 35 times a year. Our data shows three days **(the 19th, 20th, 26th) in October** where the highest one hit a max of 64 µg/m^3 (19th). PM10 according to our data proves to be at a safe level in 2022.

O3

Ozone (O3) comes from burning fossil fuels and from paints, cleaners or other solvents. The EEA regulates O3 to not surpass an hourly average of 180 µg/m^3.

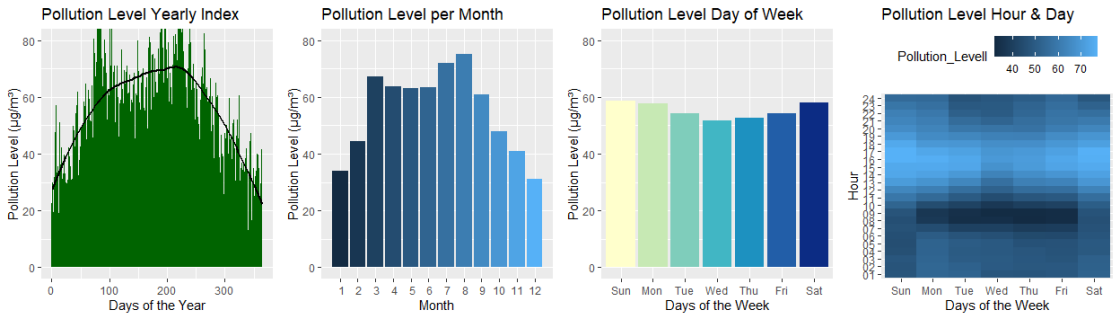

Figure 29. 30. 31. 32.

Trends

The index for O3 decreased -9% from the beginning to end of the year. This pollutant similarly to PM10 had the opposite trend of others and actually increased right away, dipped slightly in the summer months and peaked around August. Then the level decreased rapidly till December.

The peak month was **August**, and again the lowest was **December** with January close behind. The days of the week are concave down the whole time and has the weekends as the highest days and **Wednesday** as the lowest. The day with the highest level is **Sunday** but Saturday is very close too.

Hourly the graph has a trend almost opposite of every other graph with the lowest part of the day being the morning rush hours. The hot times of the day are **2pm to 6pm** especially on Saturday and Sunday. The cause of O3 is combustion (like other pollutants), but it must have a different factors or a strange lifespan compared to the others.

Stability

Throughout the year NO is extremely volatile with a daily standard deviation of 15 considering that the change from the beginning of the year to the end was ~3 µg/m^3.

Outliers

The hourly average is not supposed to surpass 180 µg/m3 and there is not time in the year that an hour is passed that number. The closest appears to be in the summer but are still safely from 180 µg/m3.

Black Carbon

Black Carbon is a life hazard and big global warming influencing pollutant that sources from combustion.

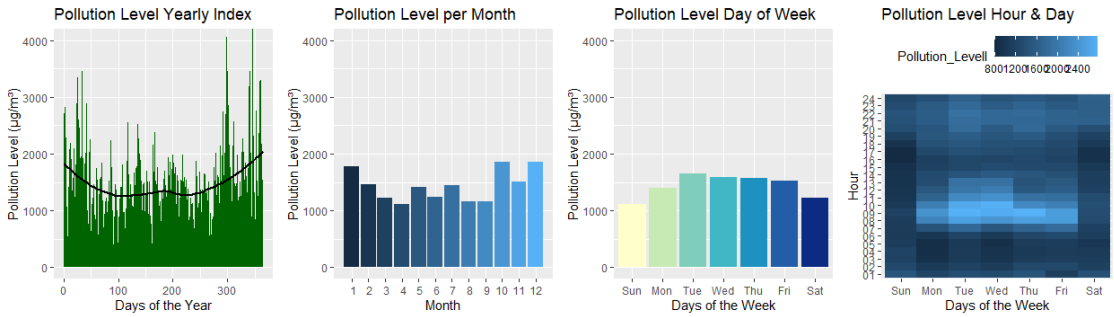

Figure 33. 34. 35. 36.

Trends

The index for Black Carbon increased 5% from the beginning to end of the year. This pollutant started high in January and decreased and stayed at a lower level throughout the summer months but had a big jump in October and rose till December.

The peak months were **October** and **December**, and **April** was the lowest with August and September closely behind. The days of the week are concave up the whole time, and **Tuesday** is the highest with **Sunday** as the lowest.

Hourly the graph has a trend very similar to the first few pollutants with the hot hour being **9-11am** on the weekdays. There is a dip in the middle of the day, but then every night of the week has a less but a good amount of exposure from 6pm to 11pm.

Stability

Throughout the year NO is also extremely volatile with a daily standard deviation of 274.88 considering that the change from the beginning of the year to the end was ~91 µg/m^3.

Outliers

Because we couldn't find info on when this pollutant is considered dangerous or hazardous we cannot assume there is an issue with the levels despite several hours of the day, days of the year, and months having a higher level then others.

Biomass Black Carbon

Biomass Black Carbon (BBC) doesn't have a unit other then %, so it is unclear what the level means, but BBC is the amount of organic Black Carbon there is in the environment at a time and is usually measured in units of energy. We are unclear of any outliers or levels of danger.

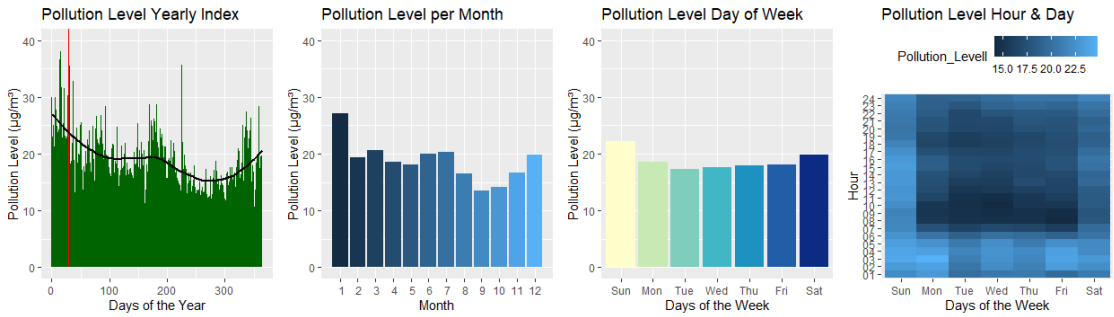

Figure 37. 38. 39. 40.

Trends

The index for Biomass Black Carbon had a substantial decrease of -27% from the beginning to end of the year. This pollutant started high in January and decreased to plateau in the summer then dropped in more in early fall. By November the pollutant bounced back up to finish around average in December.

The peak month was **January**, and September was the lowest with **October** closely behind. The days of the week were affected very differently to usual. **Sunday** was the highest and dipped down to the lowest average on **Tuesday**. After Tuesday the level slowly increased until Saturday.

Hourly the graph is completely flipped from what the past pollutants had shown. Sunday is lit up with the highest rating all throughout the day. Saturday has a high level for most the day as well. Every night from midnight till morning is slightly lit up as well. This pollutant is obviously not affected by the same factors or doesn't affect the environment the same way as the other pollutants. Without further research and understanding of the pollutant it is impossible to determine the reason of these results.

Stability

Throughout the year NO is not as volatile as Black Carbon with a daily standard deviation of 3.55 considering that the change from the beginning of the year to the end was ~6 $\mu g/m^3$.

Review

The month with that peaked the most (4 total) out of the 10 contaminants is January. Overall the winter months tended to have the highest levels so they are the most "dangerous" per say. The lowest months are August and September, so in general the summer are the "safer" months with less pollution.

The day of the week with the highest level is by far Tuesday (6 of the contaminants had it). The lowest were Saturday and Sunday. So, instead of sleeping in on Saturday go shopping in the morning when the air is mostly clear of contaminants.

The hour with the highest level of pollution is 9-10am, morning rush hour. Most of the contaminants had its higher to highest level at this time. The night is generally the time

with the least volume of pollution. July 20th and 24th had two separate contaminants report dangerous levels of exposure to the contaminants on these days. October 19-26 also reported dangerous levels of the PM10.

Question 2

Which is the impact of the "green level" (density of trees) in these indicators? Are the trees helping keep a clean air at all?

To answer this, we related the density of trees with the level of pollution for each district. For that reason you will find a graph for each district, this meaning 10 graphs in total.

SO2

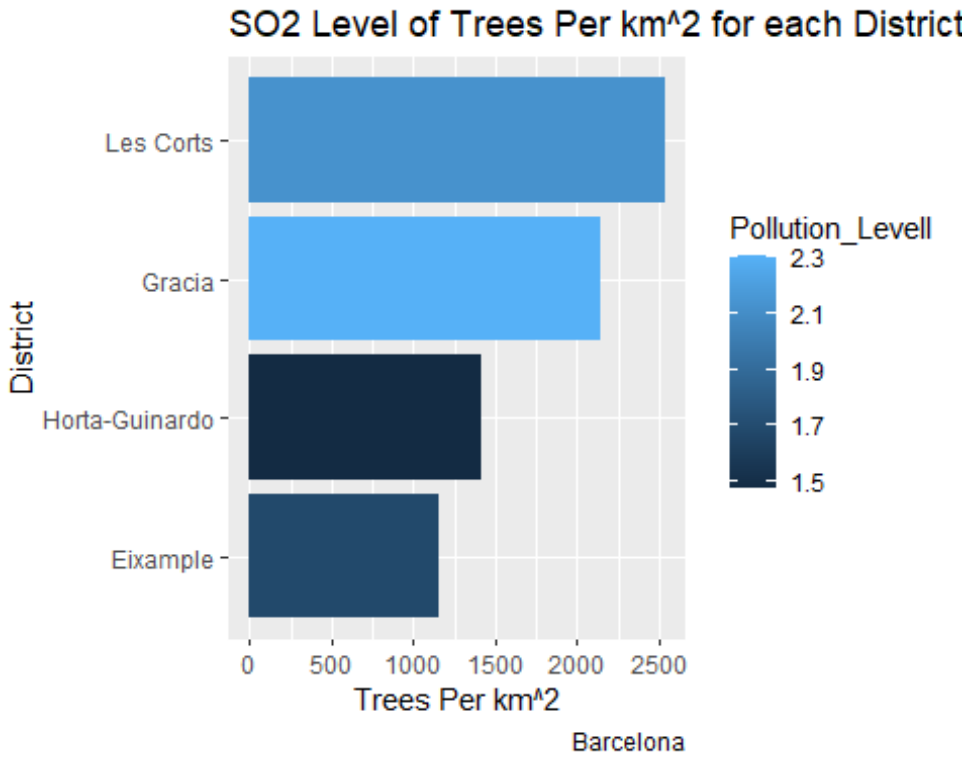

Figure 41

Trends

The trend is completely unexpected, as the intuitive expected result would be that where there is a higher percentage of trees per squared kilometers there would be less pollution. But in reality we can see that it is not like this. Surprisingly, the trend is completely the opposite way, the pollution indexes for SO2 is higher in the districts where there are more trees per km2.

We can see Les Corts and Gracia with a high amount of trees and pollution, while Horta-Guinardo and Eixample have a low amount of both trees and polution.

Plausible Explanations

It is difficult to just say that the number of trees do not really have a positive effect on SO2 emissions. As, there are many things which we did not took into consideration. There might be various external factors which affect SO2 emissions which we are not taking into consideration.

CO

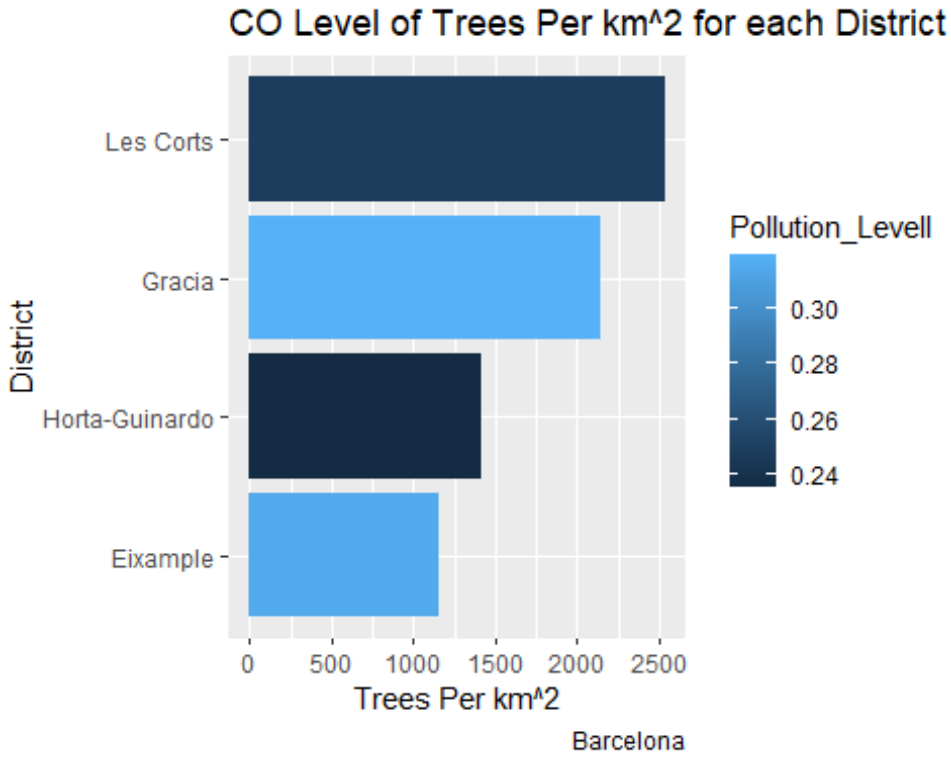

Figure 42

Trends

When trying to follow EDA(exploratory data analysis) methods and looking for a certain relationship between trees per km2 and Pollution levels for each district is rather difficult. As the results contradict themselves.

Firstly, you can see that Les Corts has the peak of number of trees and it is low in pollution levels. So that seems that the tendency will be that for more trees less pollution, but if you then take a look at Gracia is the other way around.

So that basically shows that there is no clear relationship in this graph between trees per km2 and pollution level of CO for each state.

Plausible Explanations

The most logical explanation is that as there is no clear tendency, there might be other factors that affect the Carbon monoxide levels, as the number of factories or something similar.

NO

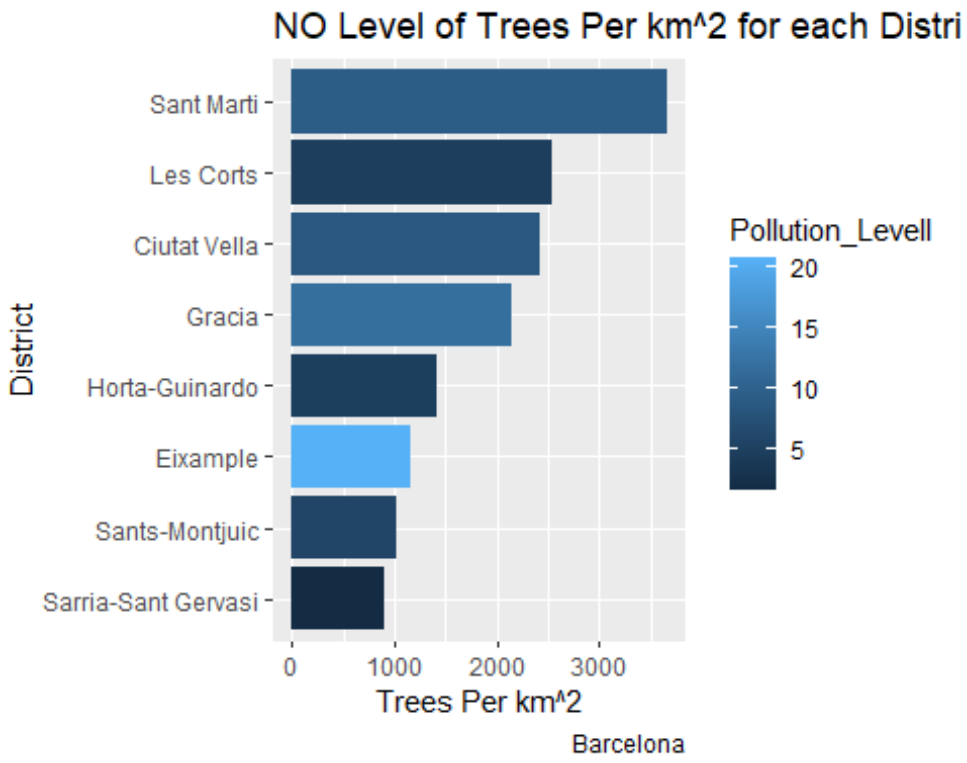

Figure 43

Trends

We can see in this graph that the district of Eixample stands out from the rest due to its elevated levels of NO(nitrogen oxide). But again there is no clear relationship between the trees and the pollution of NO by state.

Plausible Explanations

The reason behind the high number of nitrogen oxide in Eixample might be due to its population. Nitrogen oxide comes from many places one of this places are cars. So, Eixample is the district with the largest population in Barcelona. Therefore with more people comes more cars, and with more cars more NO.

It is important to highlight that at any moment we are affirming this is 100% the reason for Eixample to have the heighest levels of NO. But it is likely that is the reason or one of the reasons.

NO2

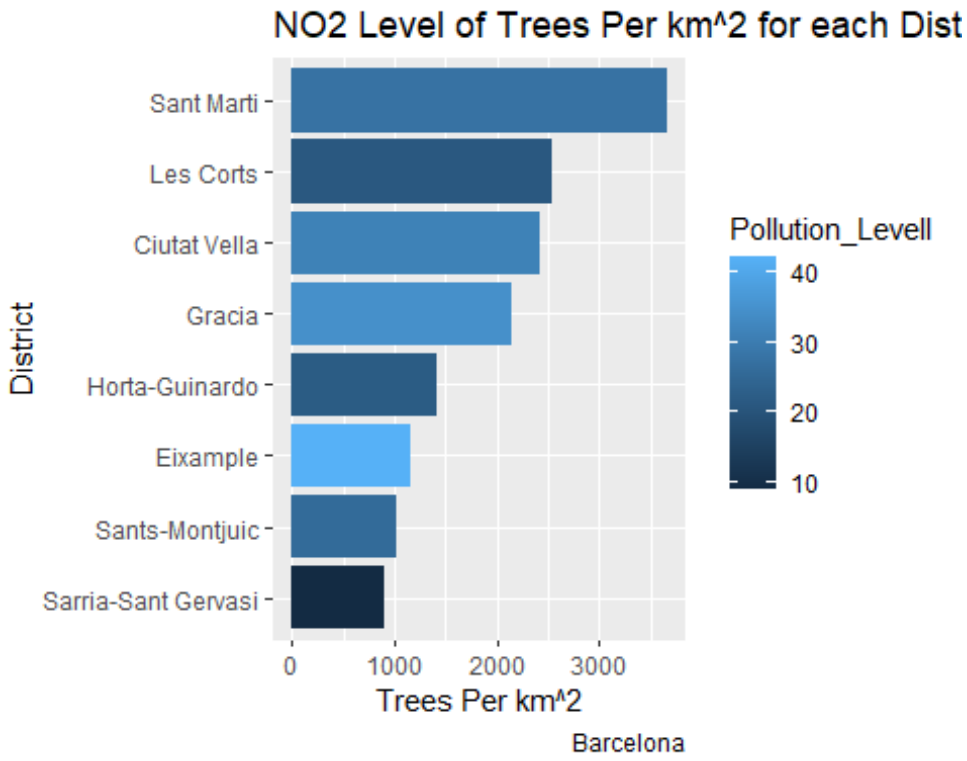

Figure 44

Trends

This trend is pretty similar to the one in NO(figure 43). Simultaneously, if we observe Sants_Montjuic and Sant Marti have a very similar level of pollution. Even though that Sant Marti has much more trees than Sants-Montjuic.

This just reaffirms how there is not an existing relationship

We can not observe any clear relationship or tendency again.

Plausible Explanations

NO and NO2 have similar results as they are similar pollutants.

Also if we take a look at Sarria-Sant Gervasi district we can see that there is a low level of pollution but also not many trees so trees are again not the relationship. But we can related again to the population, as in Sarria-Sant Gervasi district is one of the districts with a lower population rate.

This could be a reason for its low levels of NO2.

PM2.5

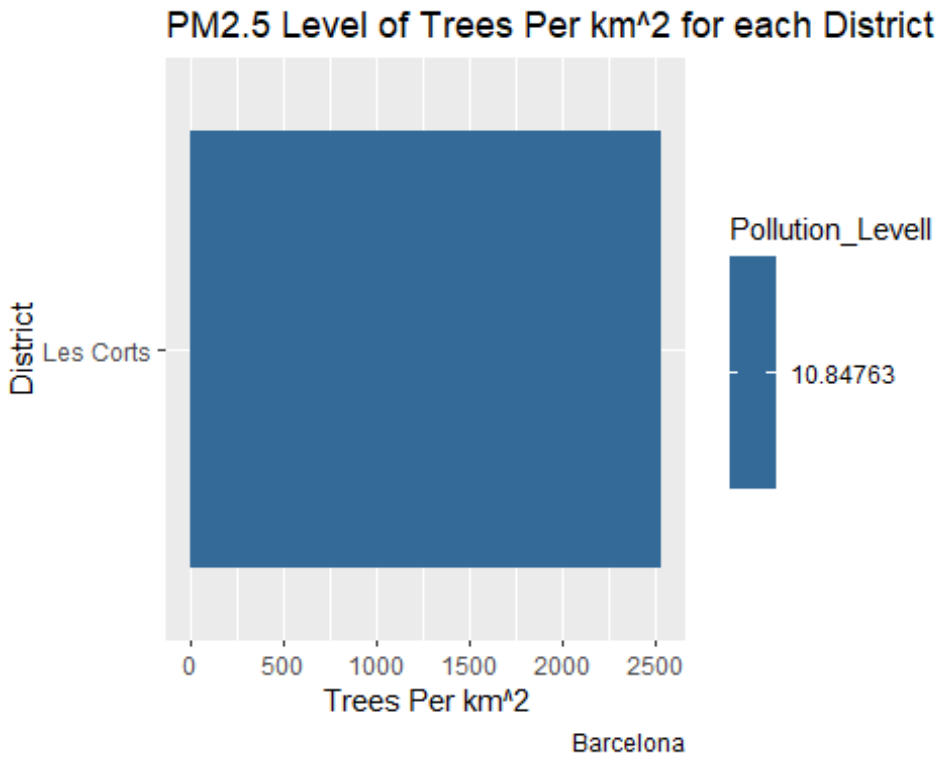

Figure 45

Trend

It is impossible to look for a relationship between trees per square kilometers and the pollution level of PM2.5 with only one kind of observations.

As there are not enough pieces of information to compare and analyze.

PM10

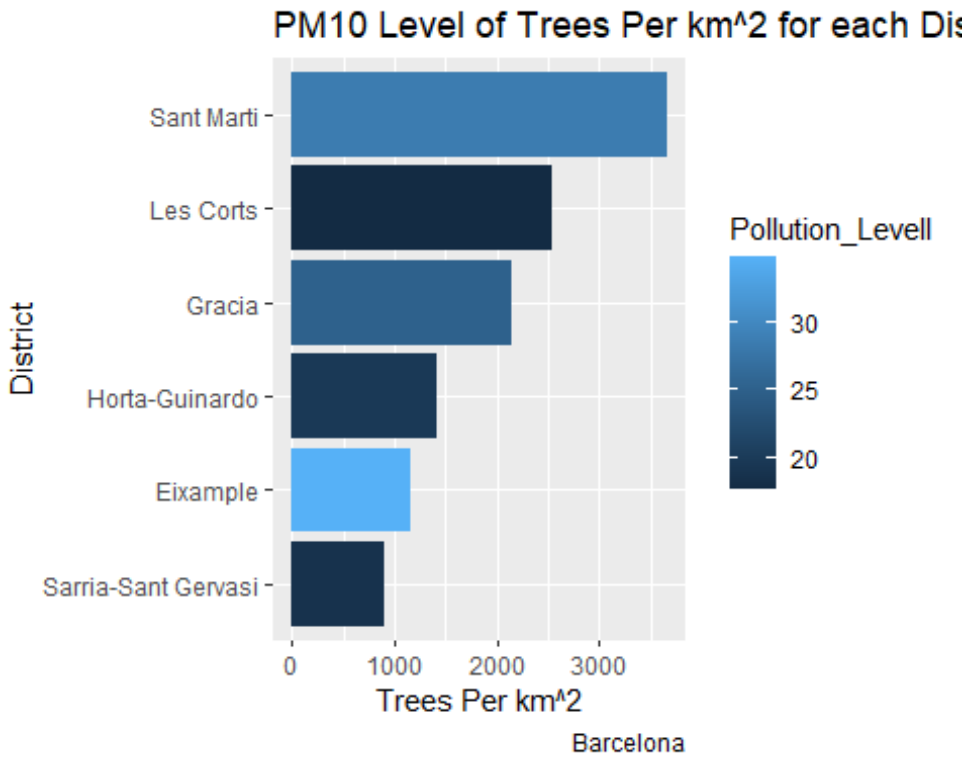

Figure 46

Trends

Here the tendency is unclear as, it happens a little bit as always we can find Sant Marti right on the top with the highest amount of trees per km2 and highest pollution level of PM10, but then once you go down you get to Les Corts. we can see that with less trees there is also less pollution level. Therefore, percentage of trees are not that good indicators for level of pollution.

Plausible Explanations

We could speculate that the possible reason for it to be like this. Is again population and movement, factories involved and new constructions. That is why Sant Marti, Eixample and Gracia have higher levels of pollution.

While Les Corts, Horta-Guinardo and Sarria-Sant Gervasi have a rather low volume of pollution for PM10.

NOx

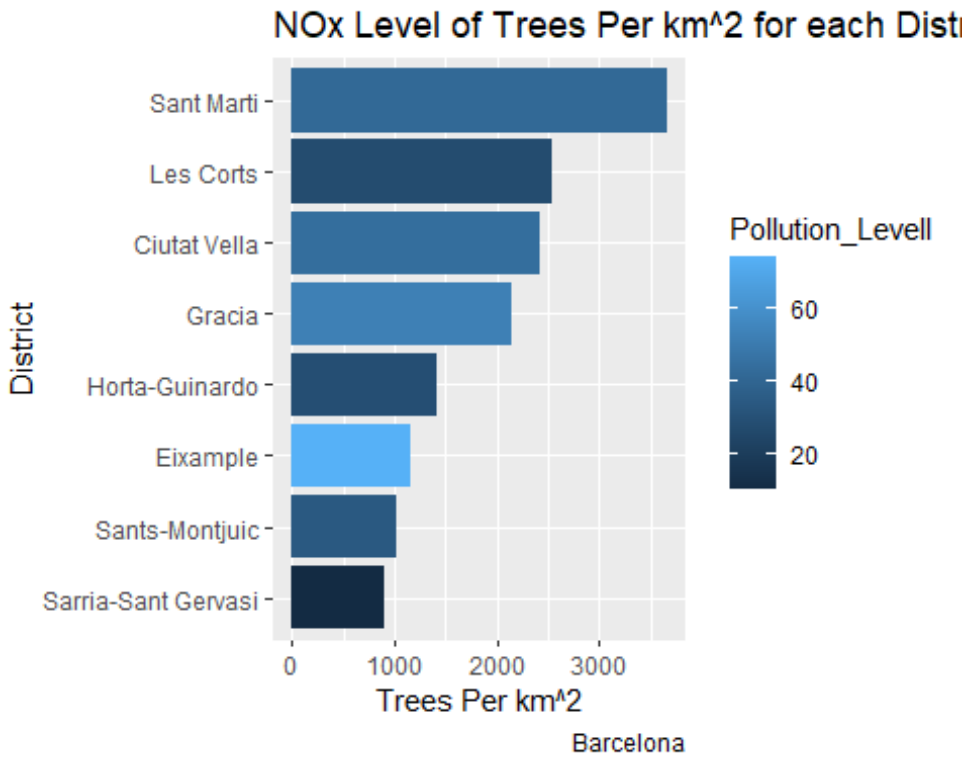

Figure 47

Trends

As with the same graphs for NO and NO2 there is no relationship between trees per square meter and the levels of NOx pollution.

To see this similar relationship you can check both figure 43 and 44.

Plausible Explanations

The explanation for this is obvious. As NOx is a mix of both NO and NO2. Therefore, the tendency they show will be pretty similar to NOx.

O3

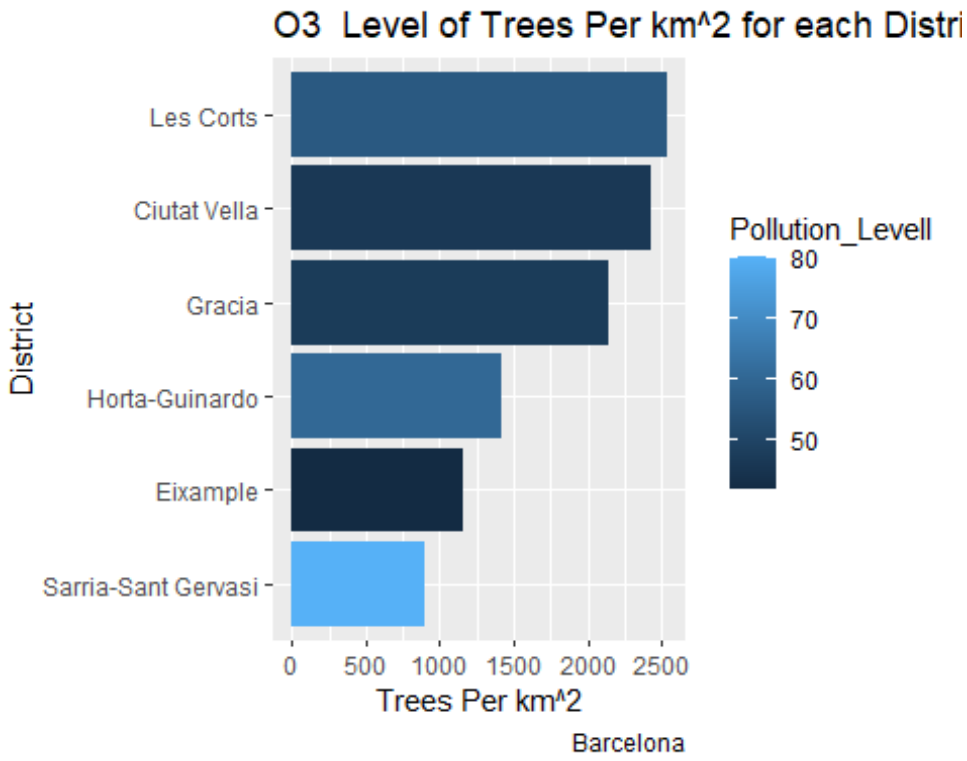

Figure 48

Trends

Here it might look as there might be a relationship between trees per square kilometers and pollution levels of O3 per state. As for the first 3 districts, Les Corts, Ciutat Vella and Gracia we can see that as they have many trees they have less pollution. While in Horta-Guinardo as there are less trees O3 levels increase. So, this might seem as a pattern. But then if you take a look at the district Sarria-Sant Gervasi, you can see that it does not follow the trend.

As a result, it ends up being as always we can not see a clear relationship.

Plausible Explanations

There is no relationship between the trees per km2 as most of O3 comes from natural process. It is also true that there are some human activities which affect it but trees is not one of them or at least is not one of the major things for its level of pollution.

Black Carbon

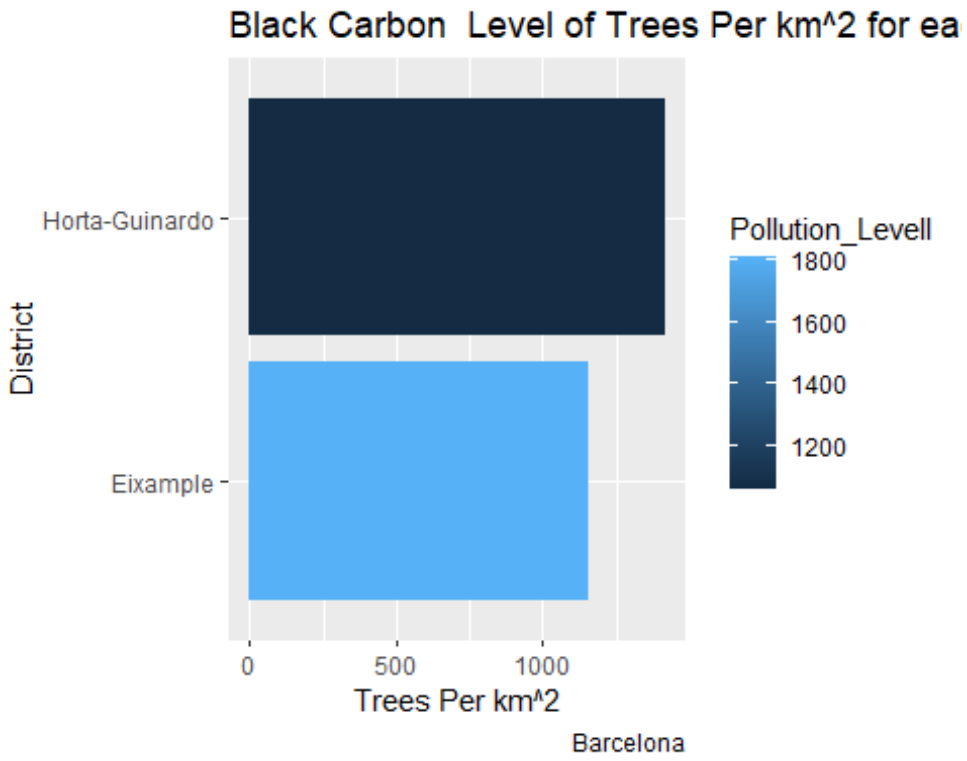

Figure 49

Trends

Trying to figure out if there is a real relationship or trend in this situation is a big challenge as we only have two districts two compare and analyse with. But still with only this two districts it is not useful to show relationship between trees and levels of pollution.

Biomass Black Carbon

Biomass Black Carbon is linked to a station that does not exist, so the website or company that we took the data sets from needs to deal with that info.

Conclusion

To conclude, we`ve seen all around the different graphs that results do not follow a general pattern. Even within the same graphs there were difficulties finding a pattern or tendency. For that reason we can say that there is no relationship between the number of trees per square meter and the level of pollution. But, also it is important to highlight how this conclusion is not fully reliable as to get to this conclusion we just took into consideration some variables. When in reality we can find many more factors that would change the results and have a greater impact on levels of pollution and would help to understand everything in a more clear way.

Question 3

Which neighborhoods are more and which are less polluted? Which is the amount of population involved in the best (and worse) neighborhoods?

We created a Facet wrap to see in a general way the level of pollution in the different neighborhoods of Barcelona.

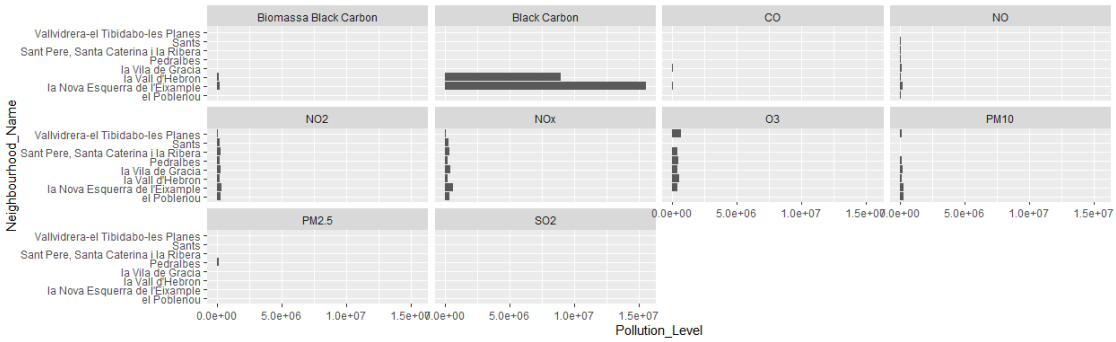

Figure 1

As we can't extract a conclusion from this big graph, we will create a graph per each pollutant code. CO looks non-existent but it is there but very small.

SO2

It is a colorless gas with a characteristic irritating odor. It is a reducing substance that, with time, contact with air and humidity, converts to sulfur trioxide.

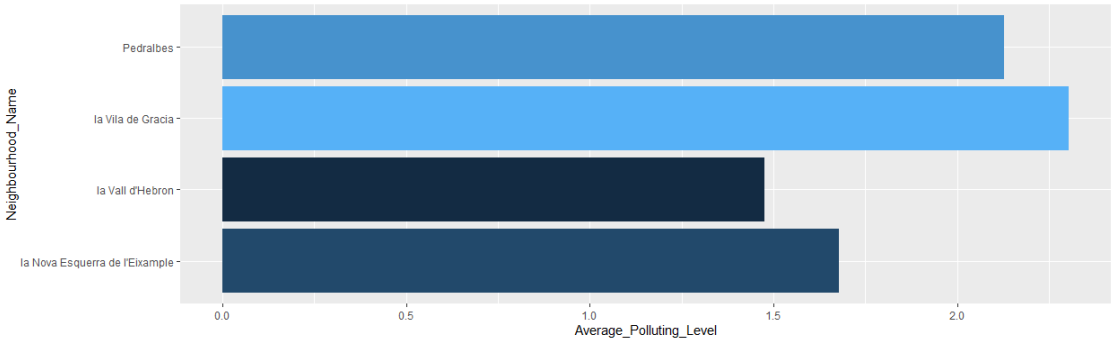

Figure 2

We only have 4 neighborhoods that analyze this pollutant element and we can see how SO2 is not a very pollutant element in our city since the highest average pollutant level is 2 in Vila de Gracia. The rest of the neighborhoods have similar values.

CO

The third pollutant code that we will be analyzing will be CO. CO as we explain in the first exercise, is a compound of carbon and oxygen that exists under standard temperature and pressure conditions. It is closely related to the greenhouse effect

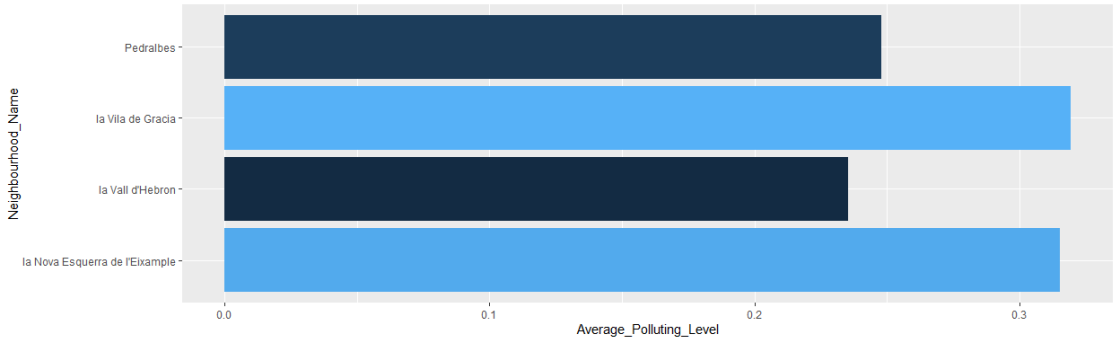

Figure 3

As it is a common pollutant code, we have 4 different stations around Barcelona: la Nova Esquerra de l'Eixample, la Vall d'Hebron, la Vila de Gracia and Pedralbes. The districts with the highest pollution level are la Vila de Gracia and la Nova Esquerra de l'Eixample with 0.3. We can also observe that compared to the black carbons, CO does not have such a high pollutant impact as the others.

NOx

Now it's turn for the NO. Nitrogen oxides (NOx), which includes nitric oxide (NO) and nitrogen dioxide (NO2). Their main source is motor vehicles, although fires and volcanoes also emit nitrogen compounds into the atmosphere.

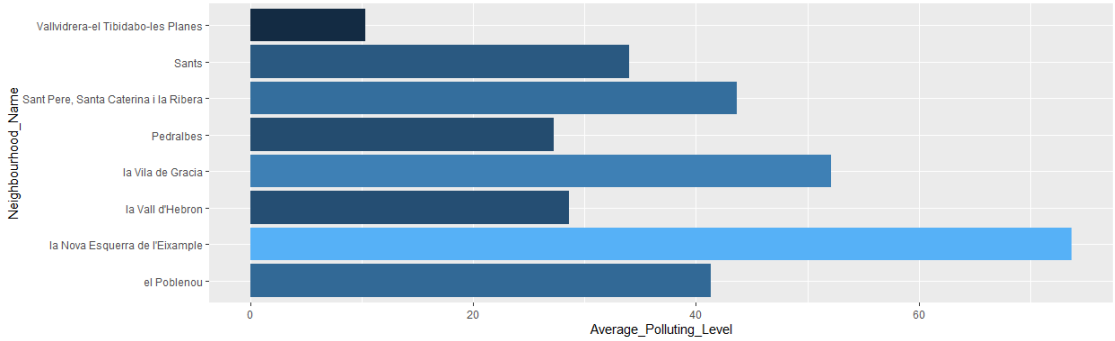

Figure 4

We can see that all Nitric Oxides are tested in the 8 neighborhoods that we have stations in Barcelona: Poblenou, Nova Esquerra de l'Eixample, Vall d'Hebron, Vila de Gracia, Pedralbes, Sant Pere, Sants and Vallvidrera. Regarding Nitrogen Oxides, we can analyze that Nova Esquerra de l'Eixample is the neighborhood with the highest average level with a great difference with respect to the others (75).

NO

And now that we have the context of the nitrics, we will analyze the nitric oxide.

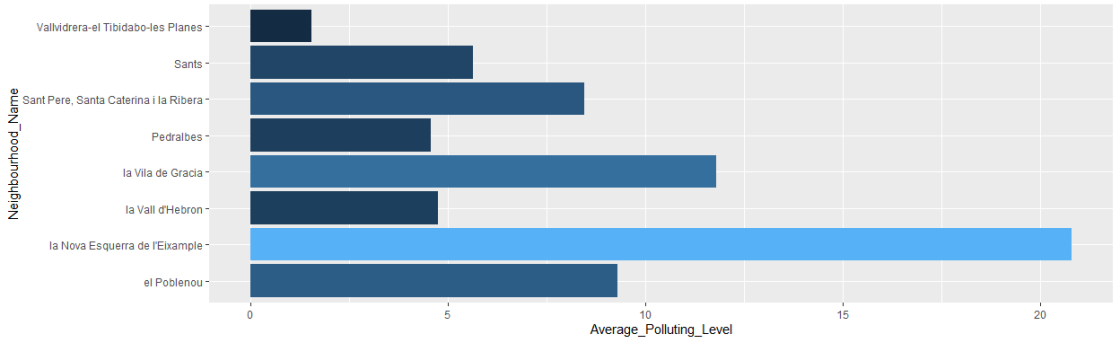

Figure 5

As we have seen in the previous graph, the Nova Esquerra de l'Eixample leads the ranking of nitrogen oxides pollution. The pollutant level of nitric oxides in the Eixample is 20 that compared to the second highest level, Gracia with 12, we can conclude that the Nova Esquerra de l'Eixample leads the ranking with a big difference.

NO2

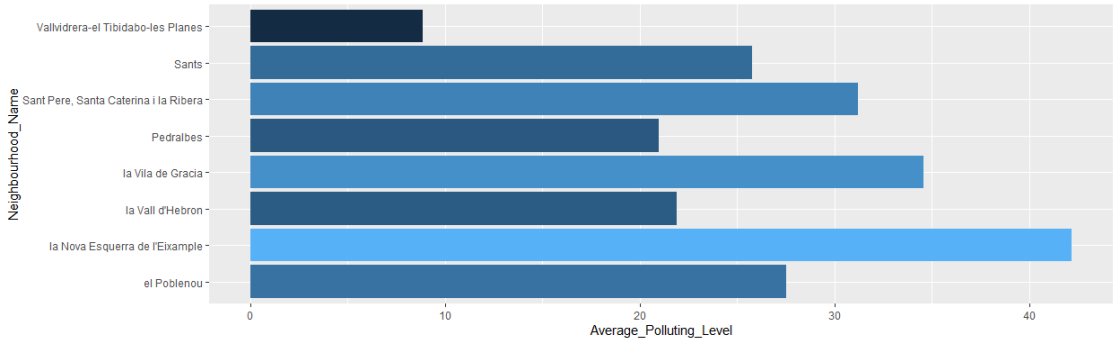

Figure 6

And another time Eixample is the district more contaminated with an average of 40. In this graph we can see how the difference is not as big as the last nitric.

PM

As we say before, PM10 refers to small solid or liquid particles of dust, ash, soot, metal particles, cement or pollen, dispersed in the atmosphere, and whose aerodynamic diameter is less than 10 µm or 2.5 µm.

PM10

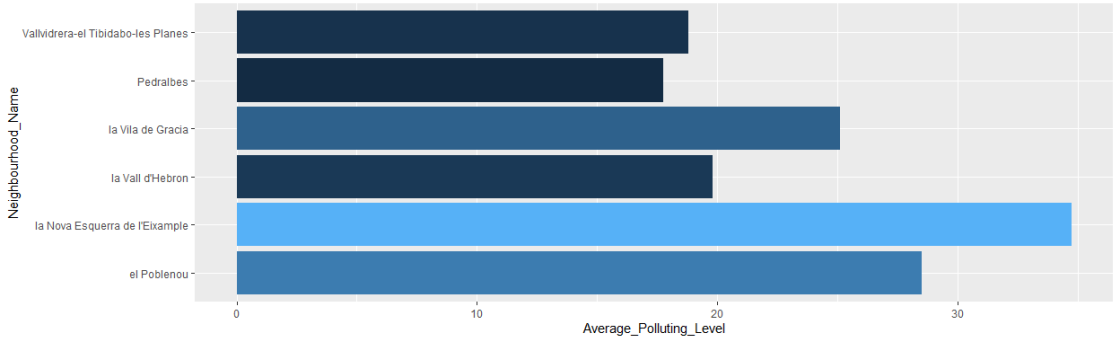

Figure 7

La nova Esquerra de l'eixample is the neighborhood with the highest average has 30 which is a great polluting average. We can also say that the values are very similar, so we can conclude by saying that we can find all these polluting codes around the city of Barcelona.

PM2.5

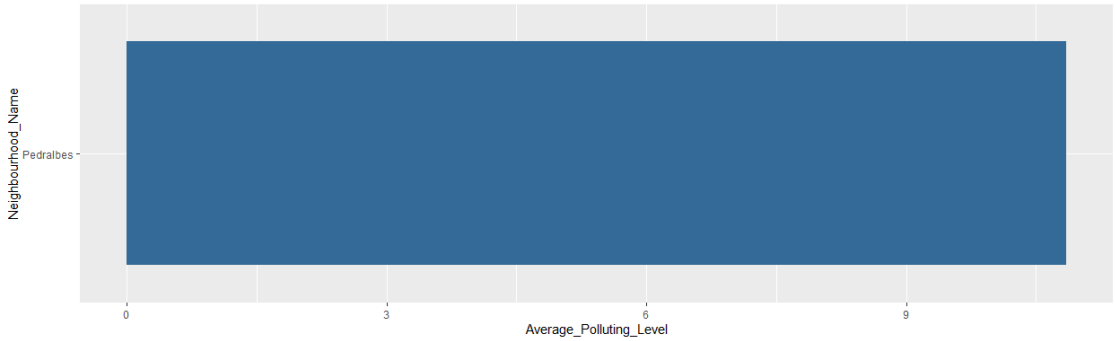

Figure 8

We only have one observation for PM2.5 and its in Pedralbes. This average compared with the PM10 is way smaller so we can conclude the PM's saying that when a PM is bigger, this one, pollutes more.

O3

O3 is one of the more dangerous Pollutant code, as we confirm before, because the ground-level ozone produced by industrial combustion, which is actually ozone mixed with nitrogen dioxide, can aggravate existing respiratory diseases and also cause throat irritation, headache or chest pain.

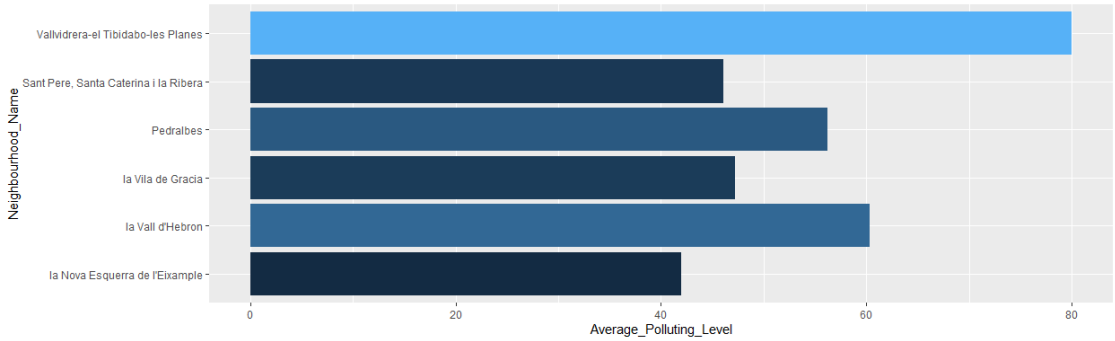

Figure 9

Vallvidrera is the area where we find the highest average pollutant level of O3. Vallvidrera has a big difference with the other neighborhoods as we can see how the second is Vall d'Hebron with 60. I think this may be because Tibidabo has many trees and they produce oxygen particles.

Black carbon

The last Pollutant code analyzed will be the black carbon. Black carbon consists of pure carbon in various bonded forms. It is formed through incomplete combustion of fossil fuels, bio-fuels and biomass.

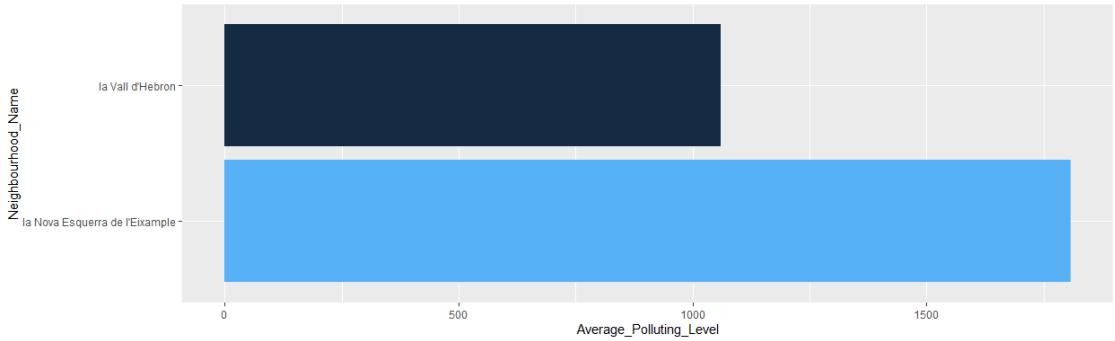

Figure 10

The neighborhood of Vall d'Hebron and Nova Esquerra de l'Eixample are the only ones that have a station that analyzes Black Carbon. Taking this into account, we can see that Vall d'Hebron is very polluted with black carbon and that Nova Esquerra de l'Eixample has an average of 1700. That is the maximum that we will see because within the black carbon we have many pollutants such as Biomass.

Biomassa

Finally we will analyze the Biomassa pollutant code. As we saw in the first exercise analysis, Biomassa black carbon is a specific type of black carbon and is very polluting.

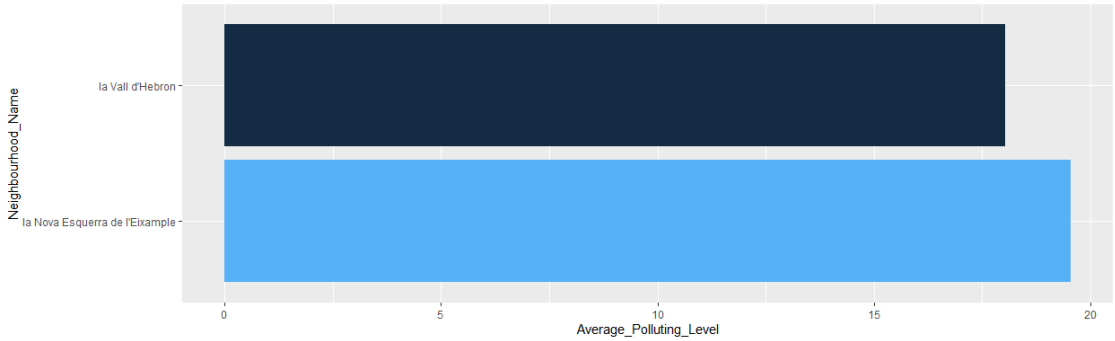

Figure 11

As in the previous analysis, we only have two neighborhoods analyzed: La Vall d'Hebron and L'eixample. We can conclude by saying that it is not a highly polluting black carbon because the average in La Vall d'Hebron is 18 and 20 for Nova Esquerra de L'Eixample.

Which is the amount of population involved in the best (and worse) neighborhoods?

Now we have to use the the average of all pollutant codes because we want to know which are the best and the worst neighbourhoods. Once we analyze this we could see if the pollution depends on the population of the districts.

So lets see the graph:

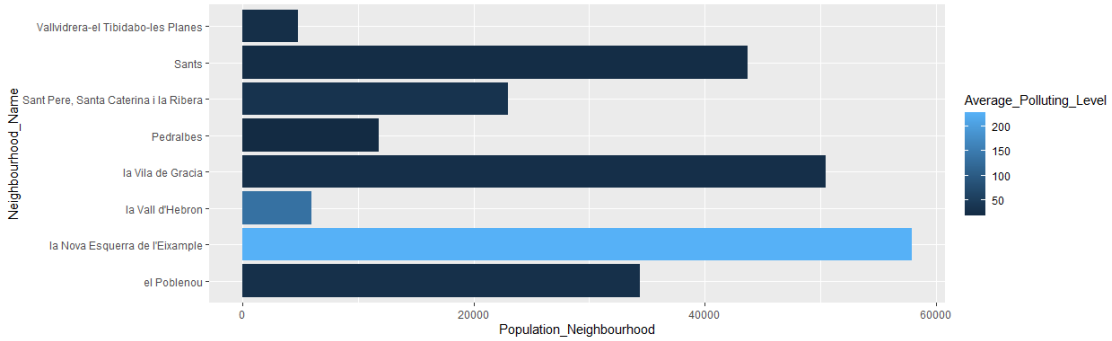

Figure 12

If we look at the result, we can see that the neighborhood with the highest average level of pollution is l'Eixample, which is also the neighborhood with the highest number of inhabitants. If we only took this into account, we could conclude that it is highly correlated, but if we check our final graph, we can see that the second neighborhood with the highest average level of pollution is Vall d'Hebron which only has 5932 inhabitants.

Last but not least, we can state that population and pollution are not correlated.

Data Limitations

In our paper we couldn't extract strong conclusions due to the fact that we have data limitations.

Those limitations range from information about the pollution that industries produce to the pollution that a single person produces.

In addition to that, we do not know what quantity of pollution trees can absorb in order to reduce it.

Biomass Black Carbon is a very sketchy pollutant due to the units and the station (999). There also wasn't information on the internet about it. Nevertheless, we added it in our evaluation best we could.

Not all stations measured each contaminant, so it is difficult to assess the trend of each contaminant throughout the stations.

In summary, there are a lot of external factors that can be affecting this analysis which are not being taking into consideration in this paper. For this reason, the information extracted from this paper is not strong enough to extract many firm and reliable conclusions.

Conclusion

To conclude this analysis, the times in the year to look out for high levels are in January, on Tuesdays, and in the morning rush hour. Areas in Barcelona to be conscious of are Eixample, Gracia for their high air pollution levels. With more focus on capturing and understanding the data related to air pollution, it allows awareness and professionals to make a difference. Our team hopes this analysis was informative enough, easy to follow along, and increased your awareness and knowledge of the different contaminants in the Barcelona air.

References

A Beginner's Guide to NOx, NO and NO2 as Air Pollutants. (n.d.). Aeroqual. https://www.aeroqual.com/blog/meet-the-nitrogen-oxide-family

Air pollution. (n.d.). European Environment Agency. https://www.eea.europa.eu/themes/air/contentbytype?b_start:int=600

Air quality data from the measure stations of the city of Barcelona. (n.d.). Open Data BCN. https://opendata-ajuntament.barcelona.cat/data/en/dataset/qualitat-aire-detall-bcn

Air quality measure stations of the city of Barcelona - Open Data Barcelona. (n.d.). https://opendata-ajuntament.barcelona.cat/data/en/dataset/qualitat-aire-estacions-bcn

Black carbon. (n.d.). Climate & Clean Air Coalition. https://www.ccacoalition.org/en/slcps/black-carbon

Is CO Heavier Than Air? | Cove Security. (2023, January 19). https://www.covesmart.com/blog/is-co-heavier-than-air/

Measured air pollutants by the air quality measurement stations of the city of Barcelona - Open Data Barcelona. (n.d.). https://opendata-ajuntament.barcelona.cat/data/en/dataset/contaminants-estacions-mesura-qualitat-aire

National Emission reduction Commitments Directive reporting status 2020. (n.d.). European Environment Agency. https://www.eea.europa.eu/publications/national-emission-reduction-commitments-directive

Ozono. (n.d.). https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/salud/ozono.aspx

Street trees of the city of Barcelona. (n.d.). Open Data BCN. https://opendata-ajuntament.barcelona.cat/data/en/dataset/arbrat-viari

The Editors of Encyclopaedia Britannica. (1998, July 20). Biomass | Definition, Types, & Facts. Encyclopedia Britannica. https://www.britannica.com/science/biomass

Trees in the parks of the city of Barcelona - Open Data Barcelona. (n.d.). https://opendata-ajuntament.barcelona.cat/data/en/dataset/arbrat-parcs

Wikipedia contributors. (2022, September 26). Districts of Barcelona. Wikipedia. https://en.wikipedia.org/wiki/Districts_of_Barcelona