Analysis for PM 2.5 pollutant

First, I need to specify that the allowed safe norm in EU for PM2.5 is . You can see the table for allowed values below.

A yellow and blue rectangular sign with black text

Description automatically generated

Source of table : <https://www.atmo-france.org/article/lindice-atmo>

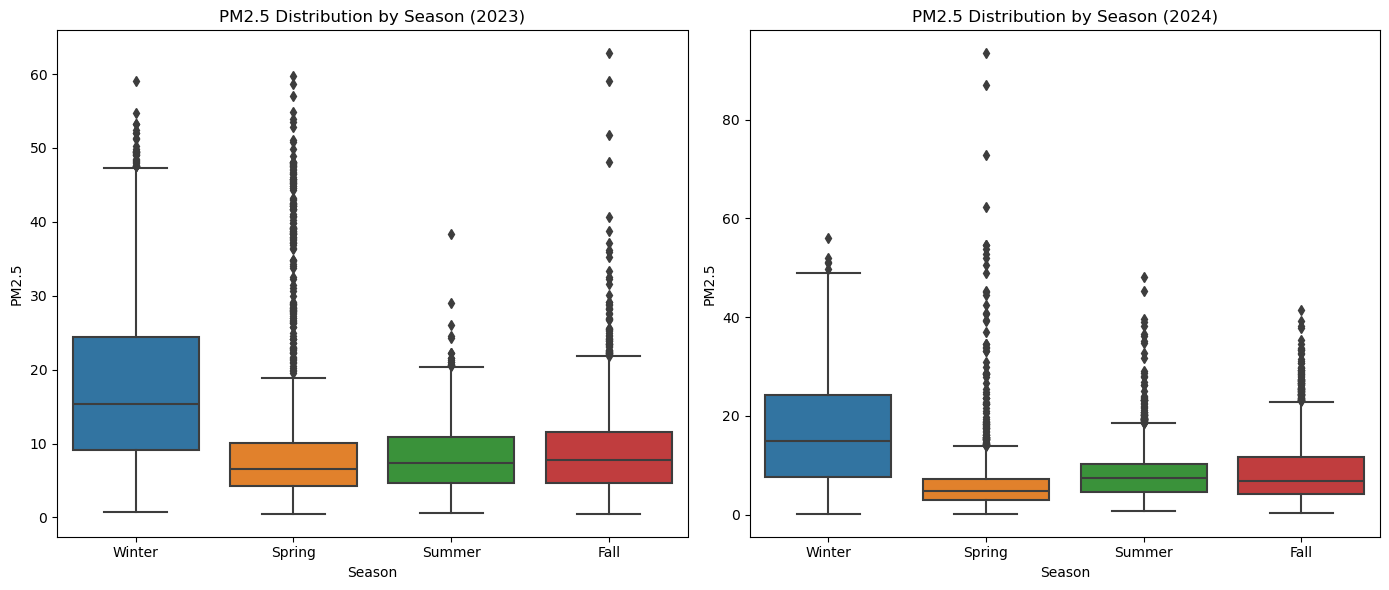
***Saint Martin d'Hères reference data:***

Using Python I could find out that 1227 out of 17354 values exceeding. Percentage of time exceeding safe limit: 7.07% (in 2 years interval)

A graph of a graph

Description automatically generated with medium confidence

Now let’s analyze data distribution. I will provide analysis seasonally (in winter, summer etc), so it can be easier to see the effect of the season on air pollution.



A group of pink and white bars

Description automatically generated

(Distribution for each season in 2024)

|  |  |
| --- | --- |
| Information |  |
| mean | 10.334591 |
| std | 8.540692 |
| min | 0.200000 |
| 25% | 4.500000 |
| 50% | 8.000000 |
| 75% | 13.000000 |
| max | 93.400000 |
| mean | 8.0 |

(Table represents the information about 2 year period)

From the provided figures, we can make the following conclusions:

* The distribution of the data is positively skewed (right skewed) since the mean is greater than the median. Also, we can observe it from box plots where top whiskers are longer than bottom ones.
* As we can observe from box plots, winter has noticeably higher median PM2.5 levels compared to other seasons in both years. Spring, Summer, and Fall generally show lower and more stable PM2.5 levels
* The median PM2.5 levels in Winter are above 20 µg/m³, approaching the safe limit of 25 µg/m³. For other seasons (Spring, Summer, and Fall), the medians are significantly below 25 µg/m³, suggesting cleaner air.
* Both years show several extreme outliers in all seasons. These represent short periods of very high pollution levels. Spring appears to have more frequent and higher outliers during 2 years. Winter has noticeably the lowest frequency of outliers in 2024.
* 2023 and 2024 have similar trends in terms of seasonal distribution and median PM2.5 levels.
* PM2.5 levels are highest in winter, with a wider interquartile range (IQR), indicating more variability in air quality. The higher levels could be due to increased heating activities, industrial emissions, or temperature inversions trapping pollutants.
* PM2.5 levels drop significantly in Spring, Summer and Fall. However, in spring we still can observe many outliers.

***Saint Martin d'Hères microsensor data:***

Let’s identify what percentage of exceeding values for microsensor data. 1331 out of 8766 values exceed safe value which is 15.18%.

A graph of a graph

Description automatically generated with medium confidence

Analysis of distribution of microsensor data:

A graph of a bar chart

Description automatically generated with medium confidence

The data provided by microsensor:

Since we have data for 2023 only, we will analyze 2023.

* In general, the situation here is roughly the same as in reference data. Again, we can observe that in winter we have significantly higher levels of PM2.5 than other seasons as well as in spring we have the greatest number of outliers.
* However, microsensor, compared to reference data, shows us that in 2023 the level of PM2.5 in winter was much higher.
* In the reference plot, Fall shows a larger spread, and more extreme outliers compared to the plot of microsensor, where the spread is slightly narrower.
* The data also has positive skewed distribution.

|  |  |
| --- | --- |
|  | Information |
| mean | 15.430489 |
| Std | 12.571453 |
| Min | 0.000000 |
| 25% | 6.722819 |
| 50% | 12.108637 |
| 75% | 19.485943 |
| Max | 79.072069 |
| Median | 12.1086365 |

(Table represents the information for 2023)

***Les Frenes reference data:***

1117 out of 17210 are higher than accepted norm. Percentage of time exceeding safe limit: 6.49% (in 2 years interval)

A graph of blue lines

Description automatically generated

Analysis of distribution of reference data:

A comparison of a graph

Description automatically generated with medium confidence

|  |  |
| --- | --- |
|  | Information |
| mean | 9.824596 |
| Std | 8.431114 |
| Min | 0.200000 |
| 25% | 4.200000 |
| 50% | 7.400000 |
| 75% | 12.300000 |
| Max | 91.500000 |
| Median | 7.4 |

* Winter looks again like having the highest level of pollution of the air.
* It can be observed that spring has the greatest number of extreme values for both years.
* During spring, fall and summer we have cleaner air compared to winter.
* Data is positively skewed.

***Les Frenes microsensor data:***

1892 out of 24852 exceeding the norm. Percentage of time exceeding safe limit: 7.61%.

***A graph with blue lines

Description automatically generated***

Analysis of distribution of microsensor data:

A graph of a graph of a graph

Description automatically generated with medium confidence

|  |  |
| --- | --- |
|  | Information |
| mean | 10.727481 |
| Std | 8.974209 |
| Min | 0.000000 |
| 25% | 4.811747 |
| 50% | 8.305294 |
| 75% | 13.149408 |
| Max | 65.195516 |
| Median | 8.305140 |