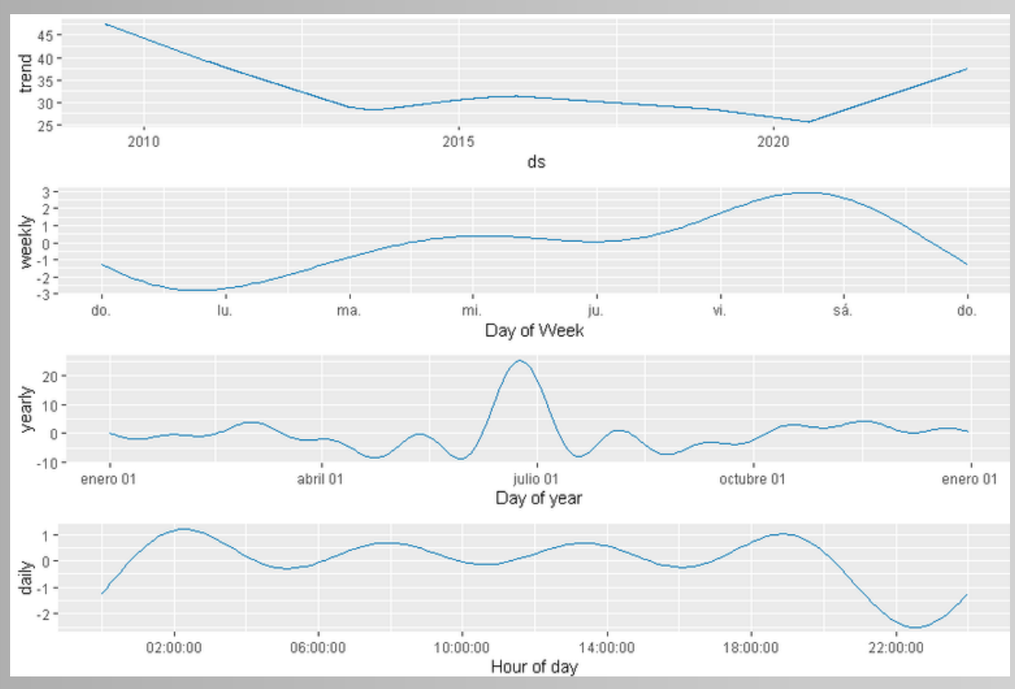
FORECASTING THE WEATHER IN EIXAMPLE(BARCELONA)

Fabián Birch Torrejón, student of the Pompeu Fabra high school, in the B1 F class done by the teacher Francesc Perez

This project aims to compare two forecasting algorithms in R, namely Prophet and AutoARIMA. The objective is to predict PM10 data for the month of January 2023 based on historical data spanning from 1991 to 2022. In order to determine which algorithm performed better, we compared them using the following evaluation metrics: Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Scaled Error (MASE). By analyzing these metrics, we were able to assess and compare the performance of the Prophet and AutoARIMA algorithms.

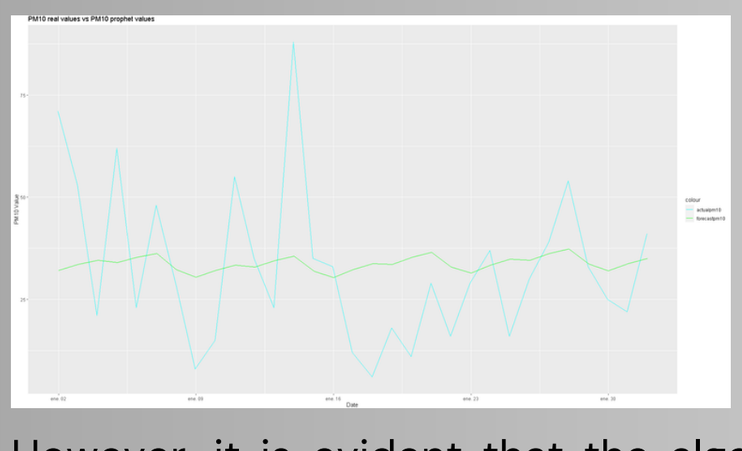


In this article, we begin by examining a graph generated by the Prophet algorithm, which predicts the behavior of PM10 levels on an hourly, daily, monthly, and yearly basis. In the annual analysis, we observe a trend of decline in previous years followed by a period of relative stability in the more recent years. However, according to the algorithm, the PM10 values are projected to significantly increase once again.

Moving on to the weekly analysis, we notice a gradual increase in PM10 levels throughout the week, reaching a peak and then experiencing a significant decrease on Sundays. The lowest value is observed on Mondays, after which the levels start to rise once more.

Zooming in on the monthly analysis, we observe generally stable PM10 levels with occasional fluctuations, except for the month of July, where a clear disruption and increase in PM10 levels can be observed.

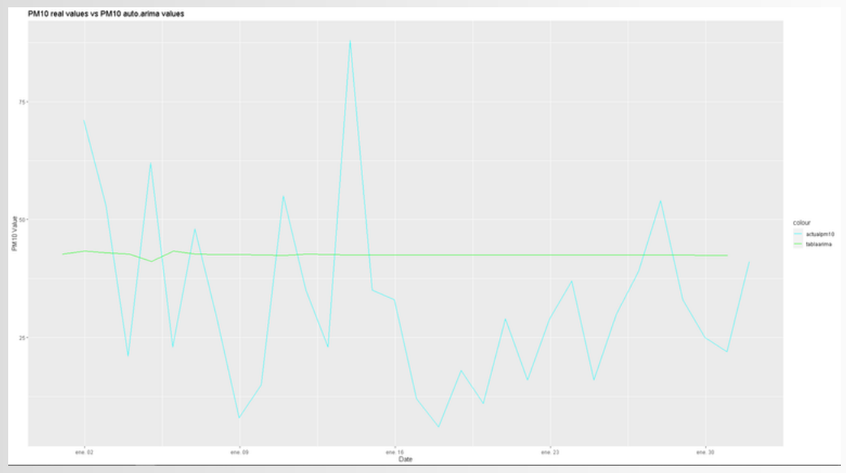
Finally, in the hourly analysis, we see a gradual increase in PM10 levels throughout the day, which remains relatively stable. However, as the evening hours approach, the levels start to decline.



This graph compares the actual data from January 2023 with the predictions generated by the Prophet algorithm. It shows that the predicted values are generally stable when compared to the actual data.

However, it is evident that the algorithm was able to accurately capture most of the upward and downward trends in the PM10 values.

After working with Prophet, we decided to also explore AutoARIMA, and here are the results it provided:



The results of the auto.arima analysis clearly indicate that the predictions align with the average of the actual values. However, it falls short in accurately forecasting the peaks and dips observed in the real PM10 data.

To determine which algorithm is better, we can use the forecast library and the accuracy function to assess how closely each algorithm's predictions align with the actual data. This comparison allows us to identify the algorithm that has achieved a closer approximation to the real values.

PROPHET	ME	RMSE	MAE	MPE	MAPE	MASE
	-0.925073	18.4728	14.40695	-3.20789	0.4261506	8.855227
AUTO.ARIMA						
	-9.736052	21.25436	18.45188	-2.2801	0.43387	84.99589

Values closer to zero in the evaluation metrics indicate greater accuracy in the predictions. Based on the evaluation metrics provided, the Prophet algorithm outperforms auto.arima in terms of accuracy. Prophet consistently achieves lower values in metrics such as RMSE, MAE, MAPE, and MASE, indicating more accurate predictions. The Mean Error (ME) metric also suggests that Prophet exhibits less bias in its predictions compared to auto.arima. Overall, Prophet demonstrates superior performance across multiple evaluation metrics, making it the preferred choice for forecasting.

In conclusion, both algorithms exhibit some level of similarity to the real data, but they are not entirely accurate. Prophet, however, has shown the ability to detect and differentiate the peaks and dips to a greater extent, whereas auto.arima has remained closer to the average values without distinct variations.