Time Series Analysis of Northern Italy Restaurants: A Data Science application in a post-pandemic business context

MARCO SALLUSTIO 906149 , MARCO GUARISCO 789244 , AND SALVATORE RASTELLI 903949

Dipartimento di Informatica Sistemistica e Comunicazione, DISCO, Università degli Studi di Milano Bicocca

Piazza dell'Ateneo Nuovo, 1, 20126 Milano MI

Abstract
This study focuses on the ongoing process of recovery that restaurants are experiencing after the suspension of activities between March and June 2020 caused by Covid-19. In such a rapidly changing context, in fact, the importance of utilizing time series analysis to forecast future revenues and effectively model restaurant management becomes evident. This approach can provide valuable tools for resource planning, staffing, and marketing strategies, considering seasonal fluctuations and potential shifts in customer preferences. The insights drawn from this study can aid restaurant owners, policymakers, and industry stakeholders in making more informed and targeted decisions, thus contributing to the revitalization and sustenance of the restaurant sector in the post-pandemic era.
Keywords
Sarima – Prophet – Restaurant – Time Series

1. INTRODUCTION

Without a doubt, one of the economic sectors most affected by the Covid-19 pandemic has been the restaurant sector. The consequences of the health emergency have manifested themselves unequivocally through the forced closure of places of socialization, such as pubs, bars and restaurants. This scenario, which has prevented any form of social contact, has given rise to a series of unprecedented challenges for operators in the sector, putting a strain on their ability to adapt.

The unstoppable rise of the virus has imposed the need to adopt very strict measures, including the suspension of restaurant activities and the adoption of regional restrictions differentiated according to the level of risk of each region.

The response to these challenges has been characterized by a series of courageous efforts and continuous adaptations. The vaccination campaign and the widespread use of personal protective equipment (PPE) have provided a ray of hope, gradually allowing businesses to reopen. However, this new phase has not been without changes and compromises. The need to maintain safety distances has led, for example, to a significant reduction in the number of tables and, consequently, in the total capacity of the restaurants. The introduction of protective measures, such as the mandatory use of masks and the frequent sanitization of environments, has led to additional costs for restaurateurs, helping to redefine the very face of the restaurant experience.

The main goal of this study is to use time series to forecast future revenues and effectively model restaurant management. In more detail, we analyzed the data on the gross receipts of six restaurants located in Northern Italy; moreover, to obtain a more accurate model, we have integrated our data with the meteorological data of the cities where the restaurants are located.

2. DATASET

2.1. Restaurant Data

The data used in the following analysis refer to six restaurants, three of which are located in Piacenza (restaurants 1, 2 and 3) while the remaining three located in provinces of Northern Italy (restaurant 0 in the province of Rimini and the restaurants 4 and 5 restaurant in the province of Pavia). The available variables are:

- **Restaurant ID**: identification number of the restaurant, from 0 to 5;

- **Date**: date to which the information refers (format: *yyyy/mm/dd*);
- **Total gross**: total gross collected on the reference day;
- **Receipts**: number of receipts made during the day of reference;

Therefore, for each of the six restaurants, two daily time series are available.

2.2. Meteorological Data

To obtain a more accurate model, the set of regressors have been integrated with the meteorological data (obtained from https://www.meteo.it) of the areas where the restaurants are located. The initial data have been enriched with the variables:

- **Heat Index**: measure used to express the perceived temperature, calculated with the average temperature and humidity[1] :

$$\begin{aligned} HI = & -42.379 + 2.049\,T + 10.143\,R - 0.225\,TR \\ & - (6.837 \times 10^{-3})\,T^2 - (5.481 \times 10^{-3})\,R^2 \\ & + (1.229 \times 10^{-3})\,T^2R + (8.528 \times 10^{-4})\,TR^2 \\ & - (1.99 \times 10^{-6})\,T^2R^2 \end{aligned}$$

  Where:
  HI = Heat Index (in Fahrenheit),
  T = ambient dry-bulb temperature (in Fahrenheit) and
  R = relative humidity (percentage value between 0 and 100).

- **Atmospheric phenomena**: Categorical variable equal to one of the following values: *storm*, *snow*, *fog*, *rain*.

We chose to use the Heat Index to enrich the dataset instead of the the raw temperature and humidity data because it is an indicator of the human-perceived temperature.

3. PRE-PROCESSING

We observed few anomalies in the data, namely:

- It seems that the data from 2018 have been collected in a monthly manner, rather than in a daily manner as it is for the other years. Furthermore, the data for Restaurant 2 begin on 1st October of 2019. We then decided to neglect these data;
- Starting from the 3rd of May 2023 the data has stopped being updated, as we can see being 0 the total gross for the rest of the month. We decided to neglect these data as well;
- The value of the total gross is 0 also in the period from March 2020 to June 2020 which is obviously given by the Covid-19 emergency. For this reason, this portion of four months was not used in subsequent analyses. The data were then divided into two parts:
  - **Pre-Lockdown data**, i.e. **436** daily observations from January 2019 to February 2020.
  - **Post-Lockdown data**, i.e. **1092** daily observations from June 2020 to April 2023.

  In the following analyses we only used the second chunk of data.
- A further anomaly was found in all six restaurants: for some days the total gross and the number of receipts is zero: this is probably caused by the closure of restaurants. In calculating metrics of performance, the null values in the test set were not considered, since it is very probable that the restaurateur has not the need to make predictions for a particular day if he intends to leave the restaurant closed.

Figures 1 and 2 shows the graphical representation of the Total Gross and Receipts time series after the pre-processing.

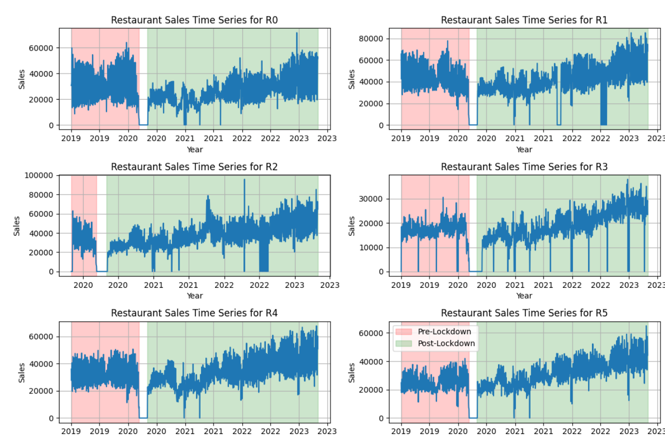

Fig. 1. Total Gross Time Series for Each Restaurant. The Colors indicate Pre (red) and Post-Lockdown(green)

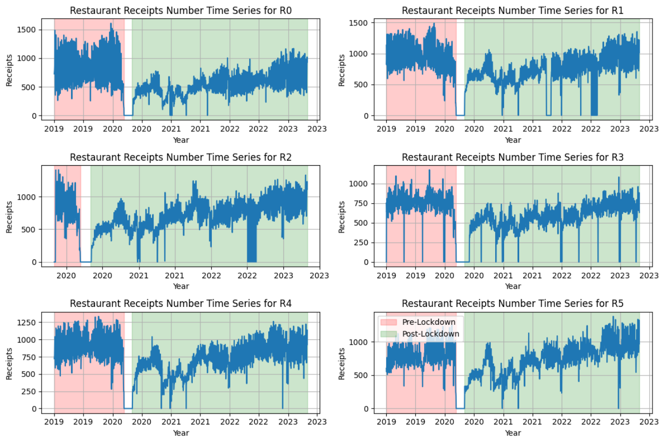

Fig. 2. Receipts Time Series for Each Restaurant. The Colors indicate Pre (red) and Post-Lockdown(green)

4. TREND ANALYSIS

From Fig.1, we can observe an important aspect for the sales: prior to the lockdown it is evident that there is a lack of discernible increasing or decreasing trends in the historical series of all six restaurants. However, after the lockdown we can identify an increasing trend for all six restaurants.

To highlight the trend, we decompose the time series. Time series decomposition is a method used to break down a series into different components, facilitating the identification and understanding of patterns and behaviors within the data. These components represent the various influences that contribute to the overall time series.

For this analysis, an STL decomposition ("Seasonal and Trend decomposition using Loess") was utilized. STL decomposition[2] is an advanced method that divides a time series into three primary components:

- **Trend**: The trend component represents the general direction or trend of the data over time. It can be a curve or line showing how the time series is changing over time. The trend is an indication of how the series evolves over the long term, regardless of seasonal patterns or short-term noise.

- **Seasonality**: The seasonal component represents cyclical or periodic patterns that repeat at regular intervals. It can include variations that follow a regular sequence of periods, such as weeks, months, or years. This component captures seasonal effects, such as sales increasing during the holidays.

- **Noise or Residual**: This component represents the random or unexplained variations that remain after removing trend and seasonality. These are the "irregular" effects that may be due to random factors or unforeseen influences

This method uses the **Loess smoothing method** to estimate the trend and seasonality. Loess is a non-parametric regression method that locally estimates mean values in the vicinity of each point, providing a smooth curve that fits the data well. For seasonality, Loess is adjusted to larger seasonal intervals than the trend. We used it to visualize graphically the restaurants' trend.

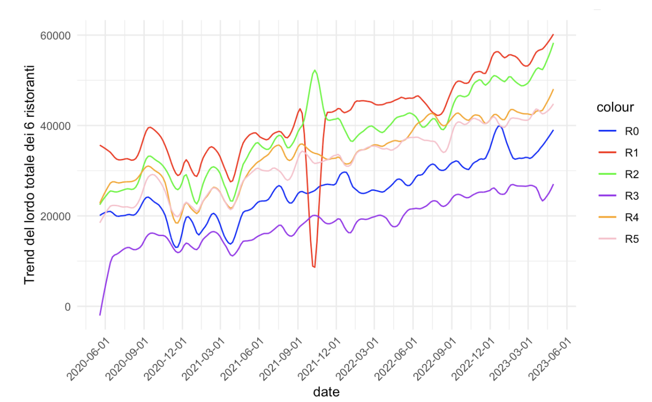

Fig. 3. Graphical Trend visualization using STL

From the results in the figure, several observations can be made:

- It's immediately noticeable that the trends of all restaurants show a slight upward trajectory. For instance, the average gross in December 2022 surpasses that of December 2021 and December 2020.

- In the months following the lockdown there is a clear growth in the average gross, but then immediately after the end of the summer, starting from the month of October 2020 there a sudden drop in the restaurant gross due to a further increase in infections which led to the subsequent introduction of new restrictive measures including the introduction of a curfew and the division of the country into different zones.

- During the Christmas holidays in the period December 2020-January 2021, a slight increase in total gross can be observed, likely due to the holiday season.

- A recovery phase for all restaurants began only after June 2021, probably driven by the easing of restrictions and an increase in the number of vaccinated individuals in Italy.

- In October 2021, precisely from 27/09/2021 to 20/10/2021, Restaurant 1 closed, consequently bringing the takings to zero. During this same period, Restaurant 2 experienced a sharp increase in its takings in the same period. Since these restaurants, together with Restaurant 3, are the only ones located in Piacenza and not in the province, one could put forward the hypothesis that the closure of Restaurant 1 has benefited Restaurant 2 due to the change in customer flow.

To further investigate the increase in average total gross during the post-lockdown period, we illustrated the trend line for the Average Transaction Value (ATV), which is the total gross divided by number of receipts of the day: it represents the average amount spent in a single transaction. This visual representation shows an increase in ATV following the spike right after the lockdown. Considering the fact that the number of receipts issued after the lockdown is lower on average to the period before (as we can see from the Figure 2), the rise in total gross could be attributed to a price increase implemented by restaurateurs to offset the financial impact of the mandatory closure of their establishments. To better understand this phenomenon we would need data about the price lists of the restaurants.

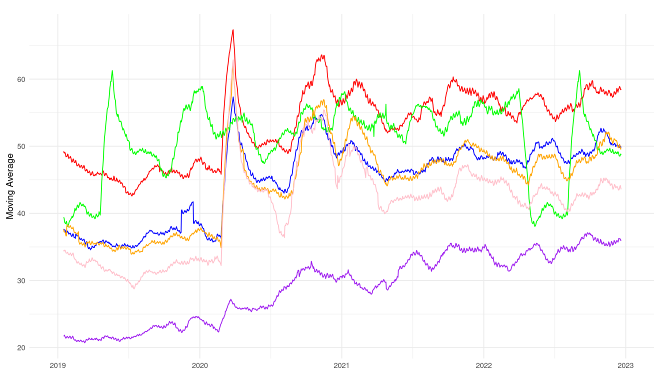

Fig. 4. Average Transaction Value trend-line

5. COVARIATE'S IMPACT ANALYSIS

Lately, we conducted an analysis to understand how different atmospheric phenomena may have influenced the total gross of Restaurants. Additionally, we monitored how various days of the week and also public holidays, could affect the restaurant's total gross. For ease of presentation, we show the results of these analyses only for one of the restaurant, in particular the R3 one.

5.1. Welch *t*-test

First, we employed a *t*-test for each atmospheric phenomenon during the post-lockdown periods. This test aimed to determine if there was a significant difference between the averages of the total gross with and without the occurrence of an atmospheric phenomenon.

The only statistically significant result was observed in the analysis of the storm variable (in figure labeled as Temporale):

```
        Welch Two Sample t-test

data:  r3_post$lordototale[r3_post$temporale == 1] and r3_post$lordototale[r3_post$temporale == 0]
t = -2.8583, df = 89.626, p-value = 0.005296
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3524.0865  -633.9042
sample estimates:
mean of x mean of y 
 16547.29  18626.29 
```

Fig. 5. Results of the Welch Two Sample t-test with storm = 1 and storm=0

Looking at the Fig.4, we can say that:

- Test Statistic (t): The t value is -2.8583. This value represents how much the averages of the two groups (those with storm = 1 and storm = 0) differ from each other, taking into account the variability of the data within each group. The fact that the t value is negative suggests that the mean value of Total Gross of the group with storm = 1 is lower than that of the group with storm = 0.

- P-value: The p-value associated with the test is 0.005296. This is the result of the probability of obtaining such an extreme t-value simply by chance, assuming that there is no real difference between the means of the two groups (null hypothesis). A very low p-value, below the common significance level of 0.05, suggests that there is significant statistical evidence to reject the null hypothesis.

- Sample Estimates: The sample estimates indicate that the mean of the group with storm = 1 is approximately 16547.29, while the mean of the group with storm = 0 is approximately 18626.29.

With a very low p-value, there is substantial statistical evidence to reject the null hypothesis, indicating that the difference between the averages is not equal to zero. Therefore, we can conclude that, based on this test, the restaurant tends to have a smaller total gross when there is a storm compared to periods without a storm.

5.2. Linear regression with phenomena covariates

However, since the *t*-test does not take into account other factors that could affect the total gross of the restaurant and may not even consider the variability of the data and also any combinations of two atmospheric phenomena, we have also fitted a logistic model that helps us to explore in more detail the relationships between the total gross and the atmospheric phenomena, also allowing to control them simultaneously. We have obtained the following results:

```
Call:
lm(formula = lordototale ~ pioggia * nebbia * temporale * neve +
    HI, data = r3_post)

Residuals:
     Min       1Q   Median       3Q      Max
-19112.7  -4248.7   -248.8   4582.2  19025.5

Coefficients: (7 not defined because of singularities)
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                   19308.92     796.20  24.251   <2e-16 *
pioggia                         243.32     656.76   0.370   0.7111
nebbia                          194.07     789.90   0.246   0.8060
temporale                     -2283.84     968.14  -2.359   0.0185 *
neve                          -8324.51    6463.88  -1.288   0.1981
HI                              -28.97      32.82  -0.883   0.3776
pioggia:nebbia                -1593.66    1780.96  -0.895   0.3711
pioggia:temporale                   NA         NA      NA       NA
nebbia:temporale                    NA         NA      NA       NA
pioggia:neve                   4627.88    6868.95   0.674   0.5006
nebbia:neve                    3405.59    9153.14   0.372   0.7099
temporale:neve                      NA         NA      NA       NA
pioggia:nebbia:temporale            NA         NA      NA       NA
pioggia:nebbia:neve           -5477.34   11536.59  -0.475   0.6350
pioggia:temporale:neve              NA         NA      NA       NA
nebbia:temporale:neve               NA         NA      NA       NA
pioggia:nebbia:temporale:neve       NA         NA      NA       NA
---
```

Fig. 6. Linear Model Meteorological Data Post-Lockdown

We can interpret the output as follows :

- the intercept is estimated at 19308.92 and this represents the expected value of Total Gross in a sunny day, when all the phenomena variables are equal to 0;
- The effects of major variables such as rain, fog and snow (labeled in figure as pioggia,nebbia and neve, respectively) do not seem to be significant (p-value > 0.05);
- The coefficient associated with the storm variable has an estimate of -2283.84 with a p-value of 0.0185. This suggests that the presence of a storm is associated with a significant decrease in the dependent variable compared to the reference situation, which is a sunny day.
- The interaction between the variables are all not significant, having a very low p-value.

This confirms the same result we had with the *t*-test, that is that the presence of a storm is statistically significant for the total gross of the restaurant.

5.3. Linear regression with holiday covariate

We also fitted a further linear model considering this time the holiday variable (a variable added by us that indicates whether a day is a holiday or not) as an independent variable. We also considered Sunday as holiday. The output of the model is:

```
Call:
lm(formula = lordototale ~ holiday, data = r3_post)

Residuals:
     Min       1Q   Median       3Q      Max
-18909.7  -4213.9   -423.9   4554.7  19101.8

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  18909.7      211.8  89.275  < 2e-16 ***
holiday      -2774.7      536.3  -5.174 2.74e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6386 on 1075 degrees of freedom
Multiple R-squared:  0.0243,    Adjusted R-squared:  0.02339
F-statistic: 26.77 on 1 and 1075 DF,  p-value: 2.736e-07
```

Fig. 7. Linear Model Holiday Post-Lockdown

The intercept is 18909.7. This represents the estimated value of total gross (labeled as lordototale) when the holiday variable is zero (working day). The coefficient associated with holiday is -2774.7. This indicates the estimated change in the variable total gross associated with a holiday exchange rate unit. Since holiday is binary, this indicates the average difference between public holidays and non-public holidays.

5.4. Linear regression with day variability covariate

Lastly, we fitted a linear model considering, this time, the day variability as an independent variable. With this model we wanted to observe how the value of the total gross changes according to the day of the week. Below is the model output always with the Post-Lockdown period data:

```
Call:
lm(formula = lordototale ~ as.factor(giorno), data = r3_post)

Residuals:
     Min       1Q   Median       3Q      Max
-22672.4  -3701.4     -5.9   4413.1  17406.5

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)         16666.6      495.6  33.627  < 2e-16 ***
as.factor(giorno)2    477.2      699.8   0.682    0.495
as.factor(giorno)3    681.6      699.8   0.974    0.330
as.factor(giorno)4    956.7      699.8   1.367    0.172
as.factor(giorno)5   3938.4      699.8   5.628 2.33e-08 ***
as.factor(giorno)6   6005.8      699.8   8.582  < 2e-16 ***
as.factor(giorno)7    600.6      699.8   0.858    0.391
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6131 on 1070 degrees of freedom
Multiple R-squared:  0.105,     Adjusted R-squared:  0.09996
F-statistic: 20.92 on 6 and 1070 DF,  p-value: < 2.2e-16
```

Fig. 8. Linear Model Days Post-Lockdown

We can then interpret the output as follows:

- The intercept represents the estimate of total gross when day is equal to 1, which corresponds to Monday. The estimated value is 16666.6. Monday is the reference level, this value is more a reference point than a practical interpretation.
- The coefficient associated with day 5 (Friday) is 3938.4, with a very low p-value (2.33e-08). This indicates that, compared to Monday, on Friday has an average estimate of significantly higher total gross of 3938.4 units.
- The coefficient associated with day 6 (Saturday) is 6005.8, with a very low p-value (2e-16). This indicates that, compared to Monday, on Saturday has an average estimate of significantly higher total gross of 6005.8 units.
- Among other coefficients (Day 2,3,4,7) none of these have a significant p-value, indicating that there is no significant statistical difference compared to Monday.

6. TOTAL GROSS FORECASTING

In this section we show the different methods used for historical time series forecasting. We esteemed 2 different models, making predictions and validating the performances firstly with a simple train/test partition of the data, then we exploited a different strategy of sampling known as Time-Series Cross Validation.

6.1. Train/Test Partitioning

The models were trained on a train set composed of 852 daily observation, while the evaluation of the forecast performance of the models was carried out by comparing the forecasts of the total gross over the next 210 days. The train set, then, goes from the 3rd of June 2020 to the 2nd of October 2022, and the last observation of the test set is dated 30th of April 2023.

We evaluated 2 different models: a **Sarimax** model[4] and a **Prophet**[5] model. To evaluate the models' performance, we considered 2 metrics: **RMSE** (Root Mean Square Error) and **MAPE** (Mean Absolute Percent Error).

6.2. Rolling Window Analysis

We employed a different strategy to partition the observations into train and test: the rolling window time-series cross-validation[3]. This type of subdivision is based on the choice of the value of an hyperparameter called prediction window p. The forecasts and estimates of the model are calculated iteratively a number of times equal to the number of observations in the test set minus the prediction window. For each iteration the model is estimated on the observations belonging to the train set and a review is carried out of the first p observations belonging to the test set. Then the train set is updated, adding one observations to the queue. The process is repeated until a review is available for all the observations in the test set. Figure 8 briefly shows how this method works with a prediction window of size p=1.

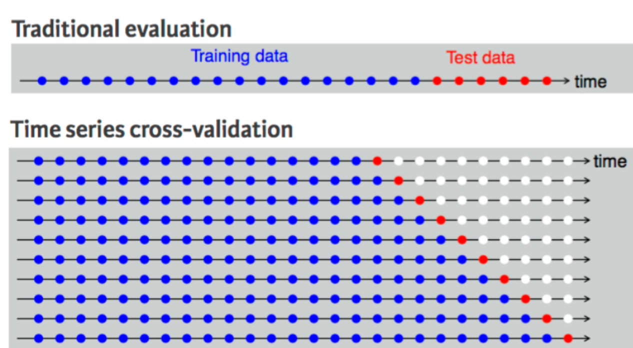

Fig. 9. Difference between the classic partitioning in train and test set and the time series Cross Validation with a prediction window of size p=1

We applied cross-validation to the two precedent models, using a **weekly** (p=7) window. We utilized the same metrics as in classic partitioning for model evaluation. Specifically, we calculated the RMSE and MAPE for each day within the window, spanning from the first day to the seventh day. This approach enabled us to assess the forecasting performance across various time lags, where a single time lag represents one day.

7. RESULTS

In this section we show the results of the forecasting on the restaurants.

7.1. Restaurant 0

7.1.1. Train/Test

Model	RSME	MAPE
Sarimax	15811.5	36.17
Prophet	10109.5	21.75

Table 1. Evaluation measures for the models used on Restaurant 0 with the train/test partitioning

7.1.2. Rolling Window

	RMSE		MAPE	
Day	**Sarimax**	**Prophet**	**Sarimax**	**Prophet**
+1	4654.1	7114.7	15.18	22.03
+2	5632.3	7188.4	18.07	21.91
+3	5561.6	7127.0	17.60	21.54
+4	5565.4	7096.2	17.63	21.50
+5	5556.5	7026.1	17.58	21.26
+6	5547.7	6976.5	17.53	21.12
+7	5532.8	7033.3	17.50	21.33

Table 2. Evaluation measures for the models used on Restaurant 0 applying time series cross validation and considering different time lags (with a single time lag being a day)

For this restaurant, the Prophet model is the best one when considering the traditional partitioning, while Sarimax is better when considering the short term forecasting. In particular, we can see that it has the best performances in terms of RMSE and MAPE when considering a one day forecast and the worst performances in a forecast of two days. The following time lags have similar values, altough it is descending.

7.2. Restaurant 1

7.2.1. Train/Test

Model	RSME	MAPE
Sarimax	19130.6	26.01
Prophet	10760.7	16.00

Table 3. Evaluation measures for the models used on Restaurant 1 with the train/test partitioning

7.2.2. Rolling Window

	RMSE		MAPE	
Day	**Sarimax**	**Prophet**	**Sarimax**	**Prophet**
+1	5995.6	7592.9	11.93	15.46
+2	6469.8	7705.8	13.17	15.62
+3	6314.6	7697.5	12.77	15.53
+4	6118.7	7692.9	12.37	15.55
+5	6127.0	7667.8	12.34	15.50
+6	6056.2	7577.3	12.18	15.37
+7	6143.7	7508.3	12.36	15.29

Table 4. Evaluation measures for the models used on Restaurant 1 applying time series cross validation and considering different time lags (with a single time lag being a day)

Also for this restaurant, the prophet model is the best one considering the traditional partitioning, while Sarimax is better in a short term forecasting. Looking at the RMSE and MAPE values for different time lags, we see that it has the best performances for a day forecast and the worst performances for a forecast of two days ahead; in particular, we can notice how the MAPE decrease from the second day, where the model has a percent error of $\sim$ 13.2%, to the sixth day, where it has a percent error of $\sim$ 12.2%, increasing again on the seventh day. Therefore, the Sarimax model for this restaurant could be useful to predict, for example, the total gross of Monday and Saturday.

7.3. Restaurant 2

7.3.1. Train/Test

Model	RSME	MAPE
Sarimax	19222.5	27.33
Prophet	14950.7	23.74

Table 5. Evaluation measures for the models used on Restaurant 2 with the train/test partitioning

7.3.2. Rolling Window

	RMSE		MAPE	
Day	**Sarimax**	**Prophet**	**Sarimax**	**Prophet**
+1	5554.6	7189.1	12.43	14.93
+2	5697.6	7331.6	12.86	15.13
+3	5716.3	7369.1	12.78	15.11
+4	5699.3	7408.7	12.74	15.20
+5	5655.6	7388.2	12.63	15.14
+6	5642.7	7414.6	12.63	15.19
+7	5626.2	7512.3	12.58	15.37

Table 6. Evaluation measures for the models used on Restaurant 2 applying time series cross validation and considering different time lags (with a single time lag being a day)

This case is similar to the Restaurant 1, being Prophet the best model considering the traditional partitioning, while is Sarimax the best model in short term forecasting. In this case too the performances of the model are the best in a day forecast, with a percent error of $\sim$ 12.4%, and the worst in a two days forecast, with a percent error of $\sim$ 12.9%. The percent error gets better as the time lags increase, but we can see how the values are all within 0.4% of percent error, therefore if the restaurateur would predict the total gross, it could use the model for the entire week with similar performances.

7.4. Restaurant 3

7.4.1. Train/Test

Model	RSME	MAPE
Sarimax	5801.7	15.92
Prophet	2964.4	8.71

Table 7. Evaluation measures for the models used on Restaurant 3 with the train/test partitioning

7.4.2. Rolling Window

	RMSE		MAPE	
Day	**Sarimax**	**Prophet**	**Sarimax**	**Prophet**
+1	2135.0	2053.5	8.10	7.84
+2	2294.1	2094.9	8.64	7.98
+3	2297.0	2105.8	8.65	8.01
+4	2322.7	2100.3	8.75	7.96
+5	2332.4	2100.8	8.79	7.95
+6	2314.2	2123.5	8.72	8.04
+7	2325.6	2146.3	8.76	8.12

Table 8. Evaluation measures for the models used on Restaurant 3 applying time series cross validation and considering different time lags (with a single time lag being a day)

For this restaurant, Prophet is the best model considering both the traditional partitioning and the short term forecasting. It is also the restaurant with the best performance out of all, even reaching a percent error of $\sim$ 8%. In this case the worst performances are for the seventh day, with a percent error of $\sim$ 8.1%, but just like the restaurant 2, the error during a week varies in a little span (the difference between the best day and the worst is just of 0.28%). Also in this case, the restaurateur that would use Prophet for forecasting, could use it to predict the entire week with similar performances

7.5. Restaurant 4

7.5.1. Train/Test

Model	RSME	MAPE
Sarimax	14877.4	29.76
Prophet	8807.9	15.66

Table 9. Evaluation measures for the models used on Restaurant 4 with the train/test partitioning

7.5.2. Rolling Window

	RMSE		MAPE	
Day	**Sarimax**	**Prophet**	**Sarimax**	**Prophet**
+1	4263.9	5415.2	11.47	13.82
+2	4617.1	5486.3	12.63	13.92
+3	4544.2	5458.7	12.34	13.75
+4	4508.4	5399.9	12.23	13.60
+5	4498.9	5357.2	12.20	13.49
+6	4499.3	5293.4	12.21	13.35
+7	4497.5	5322.7	12.21	13.43

Table 10. Evaluation measures for the models used on Restaurant 4 applying time series cross validation and considering different time lags (with a single time lag being a day)

The performances on this restaurant are comparable with the restaurant 2, being Prophet the best model considering traditional partitioning, while Sarimax has better performances for a short term forecast. However, the performance in terms of MAPE differs more than 1% between the various time lags considered, being $\sim$ 11.5% the best percent error (achieved on the first day) and $\sim$ 12.6% the worst percent error (achieved on the second day).

7.6. Restaurant 5

7.6.1. Train/Test

Model	RSME	MAPE
Sarimax	10518.6	21.68
Prophet	8309.4	15.14

Table 11. Evaluation measures for the models used on Restaurant 5 with the train/test partitioning

7.6.2. Rolling Window

	RMSE		MAPE	
Day	**Sarimax**	**Prophet**	**Sarimax**	**Prophet**
+1	3815.6	4657.6	10.50	12.26
+2	4079.2	4724.3	11.39	12.42
+3	3992.1	4743.9	11.13	12.44
+4	3984.0	4742.8	11.09	12.43
+5	4003.5	4750.7	11.12	12.44
+6	3984.0	4736.6	11.08	12.39
+7	3981.4	4758.5	11.08	12.44

Table 12. Evaluation measures for the models used on Restaurant 5 applying time series cross validation and considering different time lags (with a single time lag being a day)

For this last restaurant we have similar performances to restaurant 4.

8. FUTURE DEVELOPMENTS

8.1. CMW

One might ask whether the same data have a propensity to be intrinsically explained by multiple models. In order to delve deeper into this research question, we adopted a Cluster-Weighted Models approach implemented through the R package *flexCWM* [2]. Given that a number of clusters too high would not provide appreciable insights into the evolution of restaurants, it was decided to search for mixtures of optimal models up to a maximum of 5 clusters, all initialized via a k-means algorithm. The results obtained seem to show, after a graphical analysis, the presence of distinct bands of total gross depending on the day of the week. Figure 9 shows what has just been said for the best model according to the AICu criterion, with 4 clusters. Since this approach requires a much more in-depth study, predictions were not performed on a test dataset, but we want to show the results obtained on the training set using the data of Restaurant 3, which is the one on which the models' have the best overall performances.

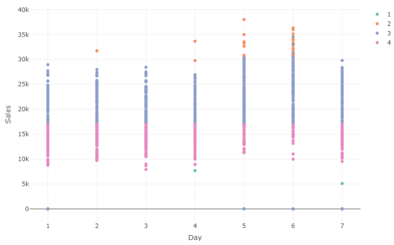

Fig. 10. Clusters identified with the CMW method

It can be observed that in all cases there is a significant and high-value intercept, indicating that a large part of the total gross does not depend on other external factors, but, as already seen in previous analyses, on trends and seasonality already present.

```
Best fitted model according to AICu

Clustering table:
  1   2   3   4
  8  37 607 410

Prior: comp.1 = 0.0071885, comp.2 = 0.0514888, comp.3 = 0.5810758, comp.4 = 0.3602468

Distribution used for GLM: gaussian(identity). Parameters:

Component 1
               Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)  5.0870e+02  2.9021e-01   1752.9 < 2.2e-16 ***
HI           1.6821e+02  8.0903e-03  20791.9 < 2.2e-16 ***
pioggia     -7.5624e+02  1.1765e-01  -6428.0 < 2.2e-16 ***
neve         3.7320e+02  1.2693e-01   2940.1 < 2.2e-16 ***
nebbia      -2.7747e+03  2.0507e-01 -13530.4 < 2.2e-16 ***
temporale   -4.5680e+03  1.7865e-01 -25569.7 < 2.2e-16 ***
holiday     -3.7509e+03  1.5735e-01 -23837.5 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
sigma = 1.0343

Component 2
             Estimate Std. Error  t value  Pr(>|t|)
(Intercept) 38647.673    269.153 143.5897 < 2.2e-16 ***
HI           -404.395     11.468 -35.2632 < 2.2e-16 ***
pioggia     -3912.282    217.235 -18.0094 < 2.2e-16 ***
neve         2402.177    787.658   3.0498  0.002347 **
nebbia       5708.049    297.882  19.1621 < 2.2e-16 ***
temporale    2531.325    330.738   7.6536 4.398e-14 ***
holiday     -7643.688    522.564 -14.6273 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
sigma = 2303.6

Component 3
               Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)  21732.4276    495.6037  43.8504 < 2.2e-16 ***
HI              -8.6164     20.4834  -0.4207    0.6741
pioggia        532.4650    403.6530   1.3191    0.1874
neve          -994.5584   1399.1917  -0.7108    0.4774
nebbia        -390.9395    454.0151  -0.8611    0.3894
temporale    -3215.8510    615.9821  -5.2207 2.146e-07 ***
holiday     -21035.8107    854.8509 -24.6076 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
sigma = 4161.2

Component 4
               Estimate Std. Error t value Pr(>|t|)
(Intercept) 14076.5298   313.9171 44.8416  < 2e-16 ***
HI              7.9134    12.9706  0.6101  0.54193
pioggia      -482.6625   229.2930 -2.1050  0.03553 *
neve        -2569.4124   789.3555 -3.2551  0.00117 **
nebbia        338.0708   266.7997  1.2671  0.20539
temporale    -274.8594   368.5243 -0.7458  0.45593
holiday       144.2086  1070.9727  0.1347  0.89291
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
sigma = 2452.7
```

Fig. 11. Results of the models' mixture

8.2. Price Lists

As has already been mentioned in the analysis carried out in section 4, it would be useful to have data on the price lists of the various restaurants which would allow us to carry out further analysis in order to confirm or reject our hypothesis that there was a price increase following the lockdown.

9. CONCLUSIONS

Based on our analysis of the data, we can conclude that the restaurant sector is experiencing a period of recovery after the lockdown. However, this recovery has not been without changes and compromises, as restaurateurs have had to adapt to new safety measures and reduced capacity. An important finding from our study is the increase in the average transaction value of restaurants during the post-lockdown period. While this increase may be attributed to the surge in prices, it could suggests that customers are willing to spend more on dining experiences.

This study provides a management basis for undertaking some strategic investment decisions of the resources of the entrepreneurs. By forecasting future revenues and modeling restaurant management, these tools can provide valuable insights for resource planning, staffing, and marketing strategies. The best results for long term forecasting, reached with Prophet, are within a range from 15% to 23% in terms of percent error, with the exception of the 8% error for Restaurant 3. However, in this sector it's more useful to have short term forecasting, and these models perform better in those cases, with percent error that span from 8% to 18%, with the exception of the $\sim$ 22% percent error for Restaurant 0.

Furthermore, we saw that there isn't a "one-for-all" formula, as the models performed differently for the different restaurants. The owner of Restaurant 3 should use a Prophet model to manage their venue, being the best choice in every forecasting situation. The owners of the other Restaurants, instead, could benefit more from a Sarimax model in a weekly forecast.

10. APPENDIX

10.1. Methodological Aspects

In order to make the forecasting, we employed different instruments: Sarima, Prophet and CMW.

10.1.1. Arima

In order to understand the SARIMA model we first need to introduce the ARIMA model. The **ARIMA** model(AutoRegressive Integrated Moving Average) is a class of statistical models used to analyze time series. It's composed of three parts:

- **AR**: AutoRegressive term: $X_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t$
- **I**: this component is the one related to differencing order (d) that has to be chosen to make the series stationary.
- **MA**: Moving Average: $X_t = \mu + \theta_1 \epsilon_{t-1} + ... + \theta_q \epsilon t - q$

The ARIMA model derives from the ARMA model to which the d-order differences were applied. Its parameters are:

1. **p**: number of lag in the auto-regressive term;
2. **d**: number of integrations;
3. **q**: number of lag in the moving average term.

The limit of the ARIMA model is that it doesn't support the seasonality in the historical series, which is defined as a cyclic recurrence in the series.

10.1.2. Sarima

Seasonal ARIMA is an extension of the arima models and it introduces 4 new parameters:

1. **P**: number of seasons in the auto-regressive component;
2. **D**: number of integrations in the seasons;
3. **Q**: number of seasons in the moving average term;
4. **m**: number of points composing the season.

10.1.3. Sarimax

There is the possibility of contemplate exogenous variables, which are external data in the forecast. We used as exogenous variables:

- Heat Index of the day;
- Holiday, a dichotomous variable that is 1 if the day considered is an Holiday, 0 otherwise;
- 4 other dichotomous variables linked to the meteorological phenomenon of the day (Fog, Snow, Rain and Storm).

10.1.4. Prophet

Prophet is open source software released by Facebook's Core Data Science team, implemented in R and Python. It uses a time series model that can be decomposed (Harvey & Peters 1990) into three main components: trend, seasonality and holidays. They are combined into the following equation:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

Where:

- g(t): trend function that models non-periodic changes in the value of the time series;
- s(t): periodic changes (weekly, monthly and/or annual seasonality);
- h(t): recurring holidays that do not always fall on the same day or period (e.g. Easter);
- ϵ_t: error that the model cannot explain.

10.1.5. CMW

The purpose of CWMs is to identify data clusters on which perform different estimates of the parameters of a statistical model and obtain an optimal fitting (a mixture of linear regression models). The data distribution (X,Y) can be written as:

$$p(x, y; \theta) = \sum_{j=1}^{k} \pi_j p(y|x; \beta_j, \gamma_j) p(x; \alpha_j)$$

where, for the j-th component, π_j is the mixing proportion, with $\pi_j > 0$ and $\sum_{j=1}^{k} \pi_j = 1$. $p(y|x; \beta_j, \gamma_j)$ is the parametric distribution of Y given X = x, with respect to β_j and γ_j, and $p(x; \alpha_j)$ is the parametric distribution of X with respect to α_j.

The fitting of the parameters is done using the expectation-maximization (EM) algorithm.

REFERENCES

1. Lans P. Rothfusz. "The Heat Index 'Equation' (or, More Than You Ever Wanted to Know About Heat Index)", Scientific Services Division (NWS Southern Region Headquarters), 1st July 1990.https://www.weather.gov/media/ffc/ta_htindx.PDF
2. Robert B. Cleveland et al. "STL: A Seasonal-Trend Decomposition Procedure Based on Loess (with Discussion)". In: Journal of Official Statistics 6 (1990). https://api.semanticscholar.org/CorpusID:64570714
3. Berta, P., Ingrassia, S., Punzo, A. et al. Multilevel cluster-weighted models for the evaluation of hospitals. METRON 74, 275–292 (2016). https://doi.org/10.1007/s40300-016-0098-3
4. Hyndman, R.J., & Athanasopoulos, G. (2021) Forecasting: principles and practice, 3rd edition, OTexts: Melbourne, Australia. OTexts.com/fpp3

5. Ahmed M. Elshewey et al. "A Novel WD-SARIMAX Model for Temperature Forecasting Using Daily Delhi Climate Dataset" https://doi.org/10.3390/su15010757
6. Taylor SJ, Letham B. 2017. Forecasting at scale. PeerJ Preprints 5:e3190v2 https://doi.org/10.7287/peerj.preprints.3190v2