Project - Time Series Forecasting

Submission type	:	File Upload	**Due Date**	:	Oct 10, 11:59 PM
Total Score	:	60	**Available from**	:	Sep 24, 8:00 AM

Description

Dear Participants,

Please find below the Time Series Forecasting Project instructions:

- You have to submit 2 files :

1. **Answer Report**: In this, you need to submit all the answers to all the questions in a sequential manner. **It should include a detailed explanation of the approach used, insights, inferences, all outputs of codes like graphs, tables etc.** Your report should **not** be filled with codes. You will be evaluated based on the business report.

Note: In the business report, there should be a proper interpretation of all the tasks performed along with actionable insights. Only the presence of interpretation of the models is not sufficient to be eligible for full marks in each of the criteria mentioned in the rubric. Marks will be deducted wherever inferences are not clearly mentioned.
THE REPORT HAS TO BE STRICTLY SUBMITTED IN A PDF/DOC FORMAT. ANY OTHER FORMAT WILL NOT BE CONSIDERED FOR GRADING. 6 Marks are allotted for the "Quality of Business Report".

2. **Jupyter Notebook file**: This is a must and will be used for reference while evaluating
- Any assignment found copied/ plagiarized with another person will not be graded and marked as zero.
- Please ensure timely submission as a post-deadline assignment will not be accepted.

Problem:

For this particular assignment, the data of different types of wine sales in the 20th century is to be analysed. Both of these data are from the same company but of different wines. As an analyst in the ABC Estate Wines, you are tasked to analyse and forecast Wine Sales in the 20th century.

Data set for the Problem: Sparkling.csv and Rose.csv

Please do perform the following questions on each of these two data sets separately.

1. Read the data as an appropriate Time Series data and plot the data.
2. Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition.
3. Split the data into training and test. The test data should start in 1991.
4. Build various exponential smoothing models on the training data and evaluate the model using RMSE on the test data.
Other models such as regression,naïve forecast models, simple average models etc. should also be built on the training data and check the performance on the test data using RMSE.
5. Check for the stationarity of the data on which the model is being built on using appropriate statistical tests and also mention the hypothesis for the statistical test. If the data is found to be non-stationary, take appropriate steps to make it stationary. Check the new data for stationarity and comment.
Note: Stationarity should be checked at alpha = 0.05.
6. Build an automated version of the ARIMA/SARIMA model in which the parameters are selected using the lowest Akaike Information Criteria (AIC) on the training data and evaluate this model on the test data using RMSE.
7. Build ARIMA/SARIMA models based on the cut-off points of ACF and PACF on the training data and evaluate this model on the test data using RMSE.
8. Build a table with all the models built along with their corresponding parameters and the respective RMSE values on the test data.
9. Based on the model-building exercise, build the most optimum model(s) on the complete data and predict 12 months into the future with appropriate confidence intervals/bands.
0. Comment on the model thus built and report your findings and suggest the measures that the company should be taking for future sales.

Important Note: Please reflect on all that you have learned while working on this project. This step is critical in cementing all your concepts and closing the loop. Please write down your thoughts here.

All the very best!

Regards,

Program Office

Scoring guide (Rubric) - Time Series Forecasting Project (1)

Criteria	Points
1. Read the data as an appropriate Time Series data and plot the data.	2
2. Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition.	5
3. Split the data into training and test. The test data should start in 1991.	2
4. Build various exponential smoothing models on the training data and evaluate the model using RMSE on the test data. Other models such as regression,naïve forecast models and simple average models. should also be built on the training data and check the performance on the test data using RMSE.	16
5. Check for the stationarity of the data on which the model is being built on using appropriate statistical tests and also mention the hypothesis for the statistical test. If the data is found to be non-stationary, take appropriate steps to make it stationary. Check the new data for stationarity and comment. Note: Stationarity should be checked at alpha = 0.05.	3
6. Build an automated version of the ARIMA/SARIMA model in which the parameters are selected using the lowest Akaike Information Criteria (AIC) on the training data and evaluate this model on the test data using RMSE.	8
7. Build ARIMA/SARIMA models based on the cut-off points of ACF and PACF on the training data and evaluate this model on the test data using RMSE.	8
8. Build a table (create a data frame) with all the models built along with their corresponding parameters and the respective RMSE values on the test data.	2
9. Based on the model-building exercise, build the most optimum model(s) on the complete data and predict 12 months into the future with appropriate confidence intervals/bands.	3
10. Comment on the model thus built and report your findings and suggest the measures that the company should be taking for future sales. Please explain and summarise the various steps performed in this project. There should be proper business interpretation and actionable insights present.	5
Quality of Business Report	6
Please reflect on all that you learnt and fill this reflection report. You have to copy the link and paste it on the URL bar of your respective browser. https://docs.google.com/forms/d/e/1FAIpQLSeBxE1cfP7ugyx8sa1JFGg_Nkv-jlEztsszbc9US911oWo2KQ/viewform https://docs.google.com/forms/d/e/1FAIpQLSeBxE1cfP7ugyx8sa1JFGg_Nkv-jlEztsszbc9US911oWo2KQ/viewform	0
Points	60

Submit Assignment

Drag your file(s) here or browse for a file to upload

Add comments

Submit Assignment

Previous Next