1. **Plot all the series (an advanced data visualization tool is recommended) - what type of components are visible? Are the series similar or different? Check for problems such as missing values and possible errors.**

Sol:

A graph with blue and orange lines

Description automatically generated

A graph of a number of data

Description automatically generated with medium confidence

A graph of different types of data

Description automatically generated with medium confidence

There’s no seasonality in ‘Y20’ but there’s trend in ‘Y20’. There’s trend in ‘Y152’ and slight seasonality with periodicity 5.

1. **Partition the series into training and validation, so that the last 4 years are in the validation period for each series. What is the logic of such a partitioning? What is the disadvantage?**

**Sol: We train model on train data and we test it on validating data to see how good our model works. But there’s a disadvantage which is, when we use historical data to predict future, its not always perfect predictor of future because if there are significant changes or shifts in the patterns that were not in our historical data then our model might not able to predict future correctly or accurately.**

1. **Generate naive forecasts for all series for the validation period. For each series, create forecasts with horizons of 1,2,3, and 4 years ahead (Ft+1, Ft+2, Ft+3, and Ft+4).**

Sol: A graph of different types of graphs

Description automatically generated with medium confidence

**4&5. For each series, compute MAPE of the naive forecasts once for the training period and once for the validation period. The performance measure used in the competition is Mean Absolute Scaled Error (MASE). Explain the advantage of MASE and compute the training and validation MASE for the naïve forecasts.**

Sol:

|  |  |  |
| --- | --- | --- |
|  | Y20 | Y152 |
| Train MAPE | 11.552918908727575 | 15.966113372749092 |
| Validation MAPE | 14.757780132072615 | 2.604441944525717 |
| Train MASE | 1 | 1 |
| Validation MASE | 4.307816799416224 | 0.38571926589513184 |

MASE is normalized to the mean absolute error of a naive forecast, which is the simplest possible forecast that can be made for a time series. This makes MASE easy to interpret and compare across different time series with different patterns.

**6. Create a scatter plot of the MAPE pairs, with the training MAPE on the x-axis and the validation MAPE on the y-axis. Create a similar scatter plot for the MASE pairs. Now examine both plots. What do we learn? How does performance differ between the training and validation periods? How does performance range across series?**

Sol:

A graph of a number of lines

Description automatically generated with medium confidence

Whenever there is low train mape there’s high validation mape and vice-versa. Train mape has increased and validation mape has decreased as we come from Y20 to Y152

Train MASE is constant but validation MASE has decreased as we come from Y20 to Y152

**8. If you are to consider exponential smoothing, what particular type(s) of exponential smoothing are reasonable candidates? Discuss the results of ES model that you considered.**

Sol: Two models were considered:

1. AAN:

A screenshot of a computer screen

Description automatically generatedA screenshot of a computer screen

Description automatically generated

A graph of a train and ets forecast

Description automatically generated

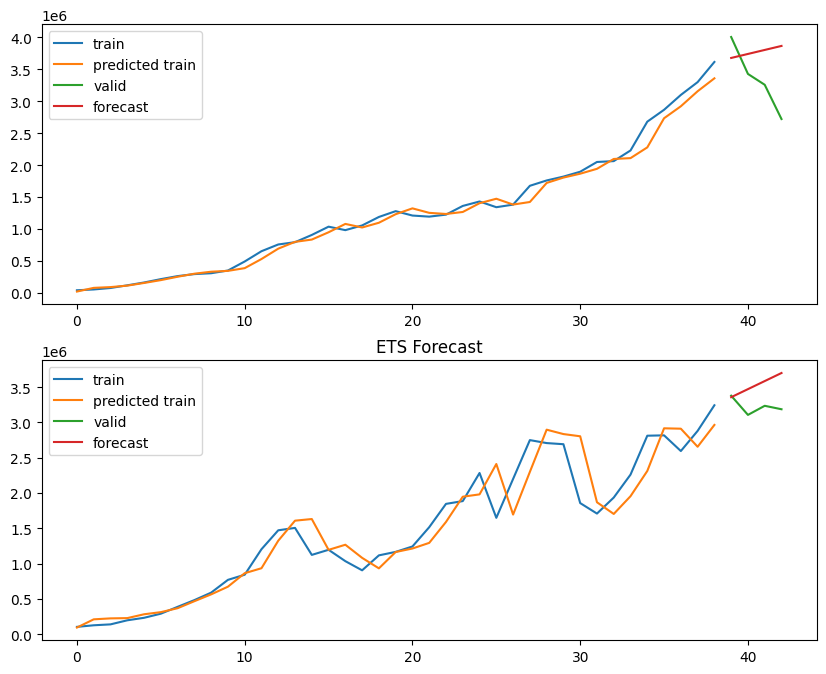
|  |  |  |
| --- | --- | --- |
| ANN Model | Y20 | Y152 |
| Train MAPE | 8.36 | 19.62 |
| Validation MAPE | 30.44 | 3.93 |

1. A screenshot of a computer screen

   Description automatically generatedMAN

A screenshot of a computer screen

Description automatically generated



MAPE values:

|  |  |  |
| --- | --- | --- |
| MAN Model | Y20 | Y152 |
| Train MAPE | 8.69 | 16.24 |
| Validation MAE | 19.02 | 9.82 |

From both model tables you can conclude that whenever there is low train mape there’s high validation mape and vice-versa. Train mape has increased and validation mape has decreased as we come from Y20 to Y152.

By comparing naïve model and Ets models, naïve models are better than ets models.

**7. For forecasting, first compare the three methods and then use an ensemble of the three methods:**

**• Naive forecasts multiplied by a constant trend (global/local trend: "globally tourism has grown "at a rate**

**of 6% annually.")**

**• Linear regression**

**• Polynomial regression**

**• Exponentially-weighted linear regression**

**(a) Write the exact formula used for generating the first method, in the form Ft+k = . . . (k = 1, 2, 3, 4)**

**(b) What is the rational behind multiplying the naive forecasts by a constant? (Hint: think empirical and**

**domain knowledge)**

**(c) What should be the dependent variable and the predictors in a linear and polynomial regression model**

**for this data? Explain.**

**(d) Fit the regression models to both the series and compute forecast errors for the validation period.**

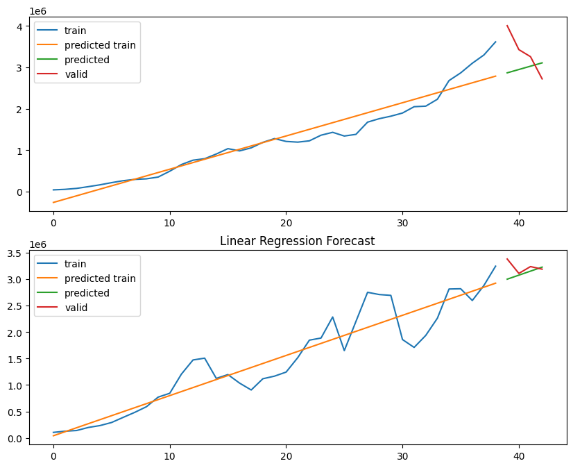
Sol: a): F(t+1) = F(t) \* (1.06)

F(t+k) = F(t) \* (1.06)^(k)

b) By multiplying constant we take trend into consideration(increased 6%) which we didn’t consider in naïve based approach(trend and seasonality is not considered as any value over horizon h is always constant)

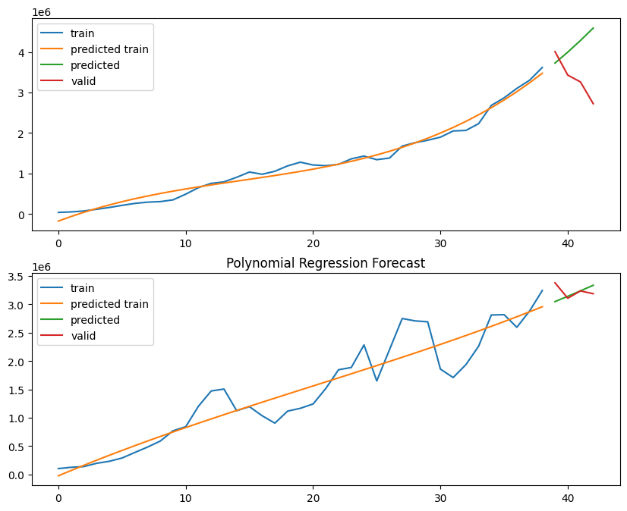
d) Linear Model:

|  |  |  |
| --- | --- | --- |
| Linear Model | Y20 | Y152 |
| Train MAPE | 53.1 | 20.6 |
| Validation  MAPE | 15.9 | 4.05 |



|  |  |  |
| --- | --- | --- |
| Poly degree 3 | Y20 | Y152 |
| Validation MAPE | 30.88 | 3.89 |
| Train MAPE | 33.45 | 22.37 |

Polynomial model(degree 3):

  
Still naïve model is better than linear and polynomial model.

c) Y20 and Y152 are dependentvariables.

Trend is independent(guess).

**9. Can you suggest methods or an approach that would lead to easier automation of the ensemble step?**

Sol: We can use XGBoost or Prophet models.

**Plagiarism Check:**

A screenshot of a social media post

Description automatically generated

These 9 are the QUESTIONS itself which are copied from question paper.