UNIVERSIDADE FEDERAL FLUMINENSE

Programa de Mestrado e Doutorado em Engenharia de Produção

Forecasting

Lesson: 101

Professor: Valdecy Pereira, D. Sc.

email: valdecy.pereira@gmail.com

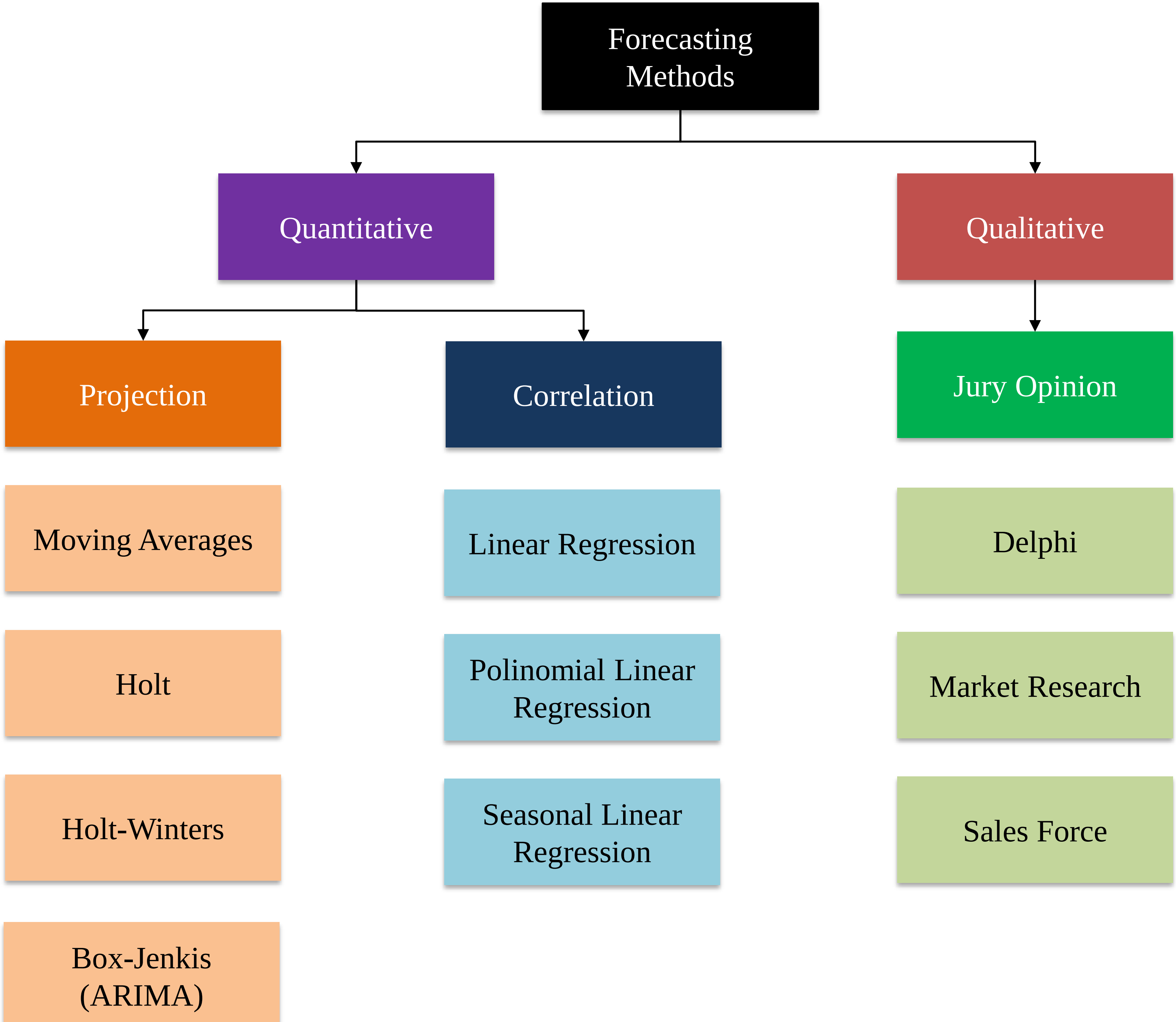
Presentation

- **Graduation:** Production Engineering (UERJ)
- **Specialization:** Construction Engineering and Assembly (PROMINP / UFF)
- **Masters:** Production Engineering (UFF) - Emphasis on Technology and Innovation.
- **PhD:** Production Engineering (UFF) - Emphasis on Operational Research.

Date	Lessons
05/04/2018	Prediction Theory and Errors
12/04/2018	Test
19/04/2018	Decomposition of Time Series
26/04/2018	Test
03/05/2018	Holt Method
10/05/2018	Test
17/05/2018	Holt-Winters Method
24/05/2018	Test
31/05/2018	RECESS
07/06/2018	Linear Regression
14/06/2018	Test
21/06/2018	Seasonal Linear Regression
28/06/2018	Test
05/07/2018	Logistic Regression
12/07/2018	Test

https://github.com/Valdecy/Forecasting

Year	GM - Pontiac Aztek Sales
2000	11,201
2001	27,322
2002	27,793
2003	27,354
2004	20,588
2005	5,020
2006	347
2007	69


Forecasting Methods
Quantitative
Qualitative
Projection
Correlation
Jury Opinion
Moving Averages
Linear Regression
Delphi
Holt
Polinomial Linear Regression
Market Research
Holt-Winters
Seasonal Linear Regression
Sales Force
Box-Jenkis (ARIMA)

Forecasting

Forecasting is a rational process of searching for information about the amount of item sales (or set of items) in the future. However forecasting methods do not lead to perfect results, and the chance of error increases as the planning horizon increases.

A **time series** can be defined as a set of observations of a variable arranged sequentially in time. Usually the time series are analyzed from their main movements described as: trend, seasonality, cycle and random events (prediction errors).

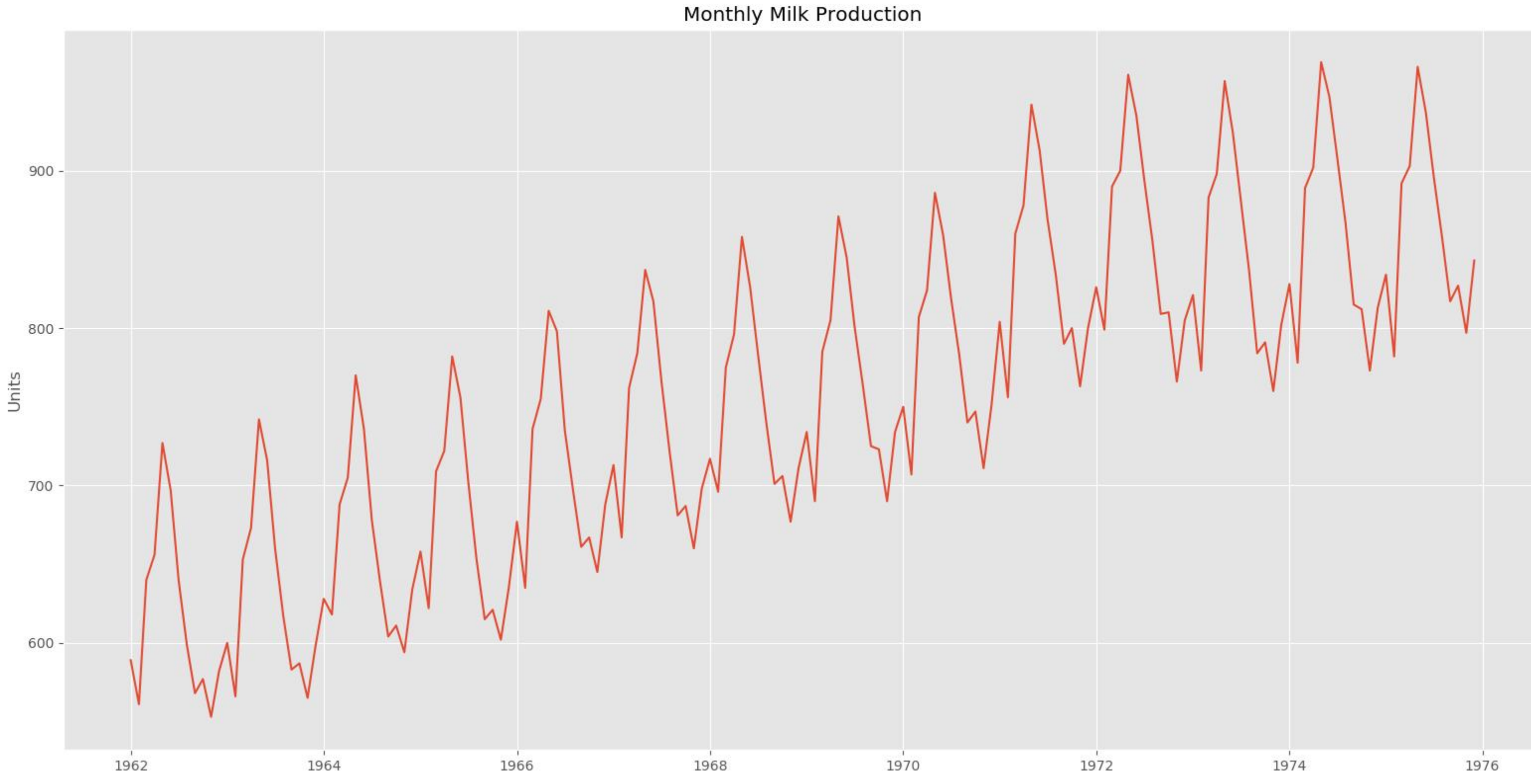
TIME SERIES

Trend

Seasonality

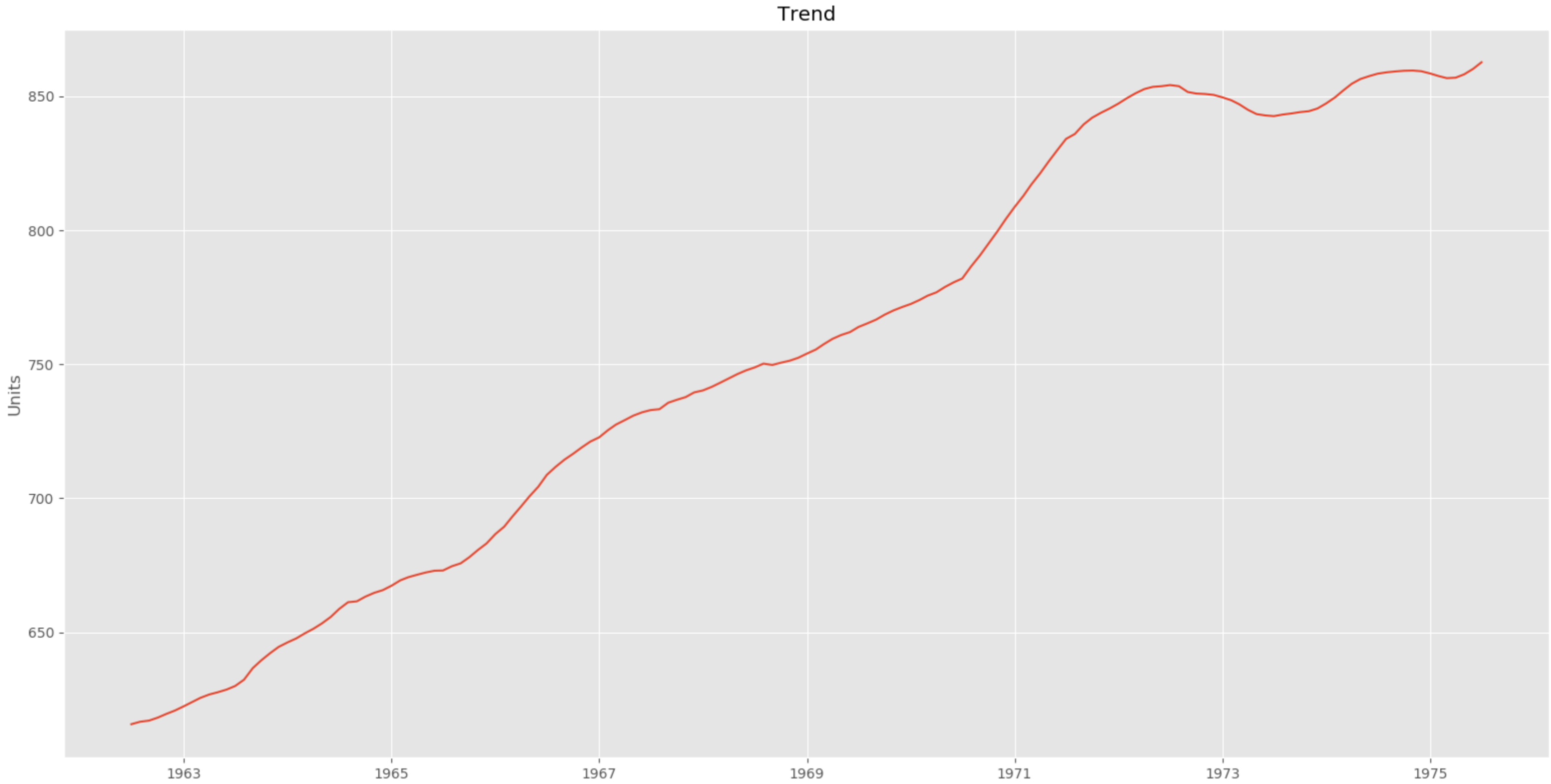
Cycle

Noise


Monthly Milk Production
Units
900
800
700
600
1962
1964
1966
1968
1970
1972
1974
1976

TIME SERIES

Trend	Seasonality
Cycle	Noise



Trend
Units
850
800
750
700
650
1963
1965
1967
1969
1971
1973
1975

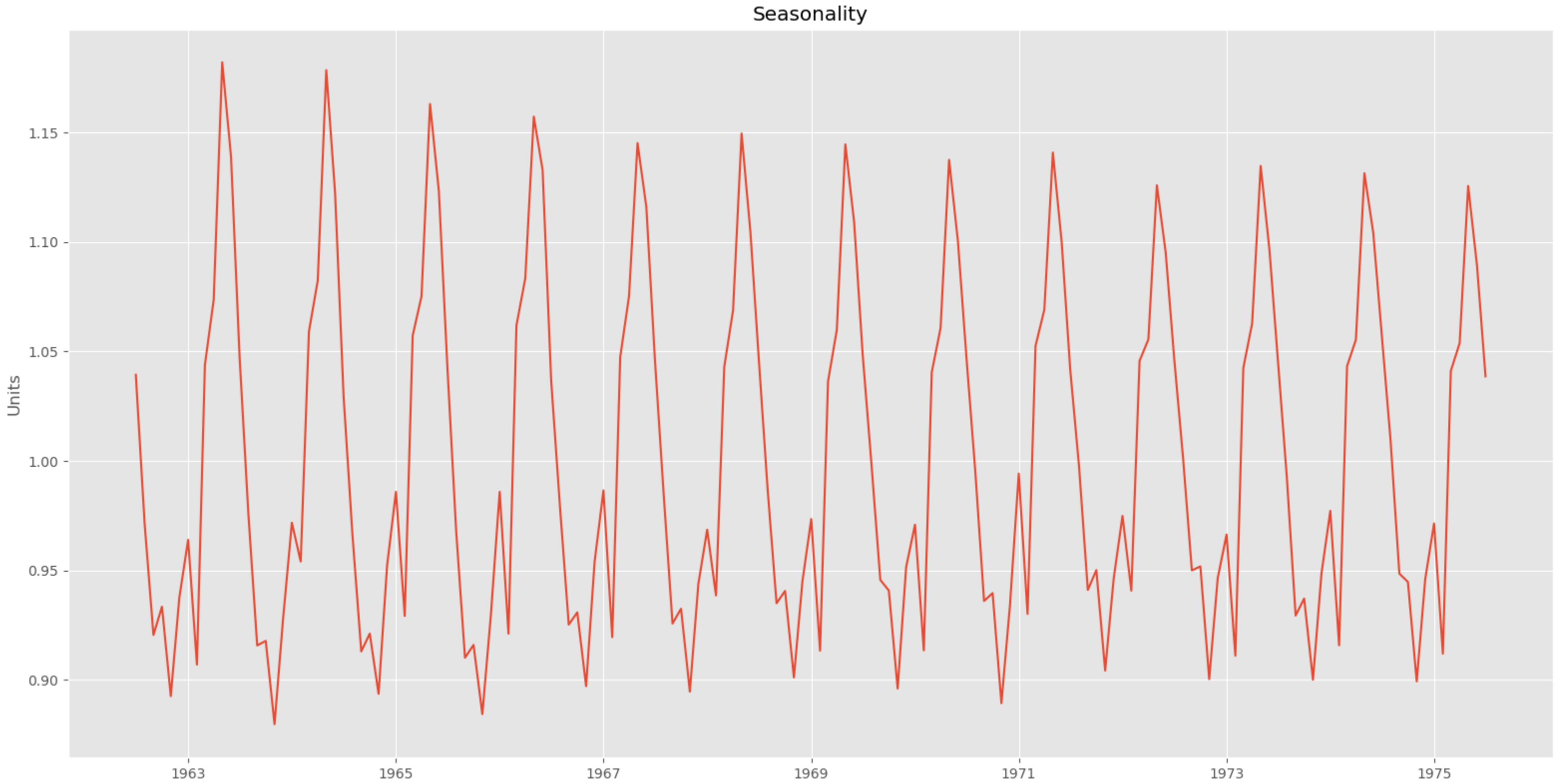
TIME SERIES

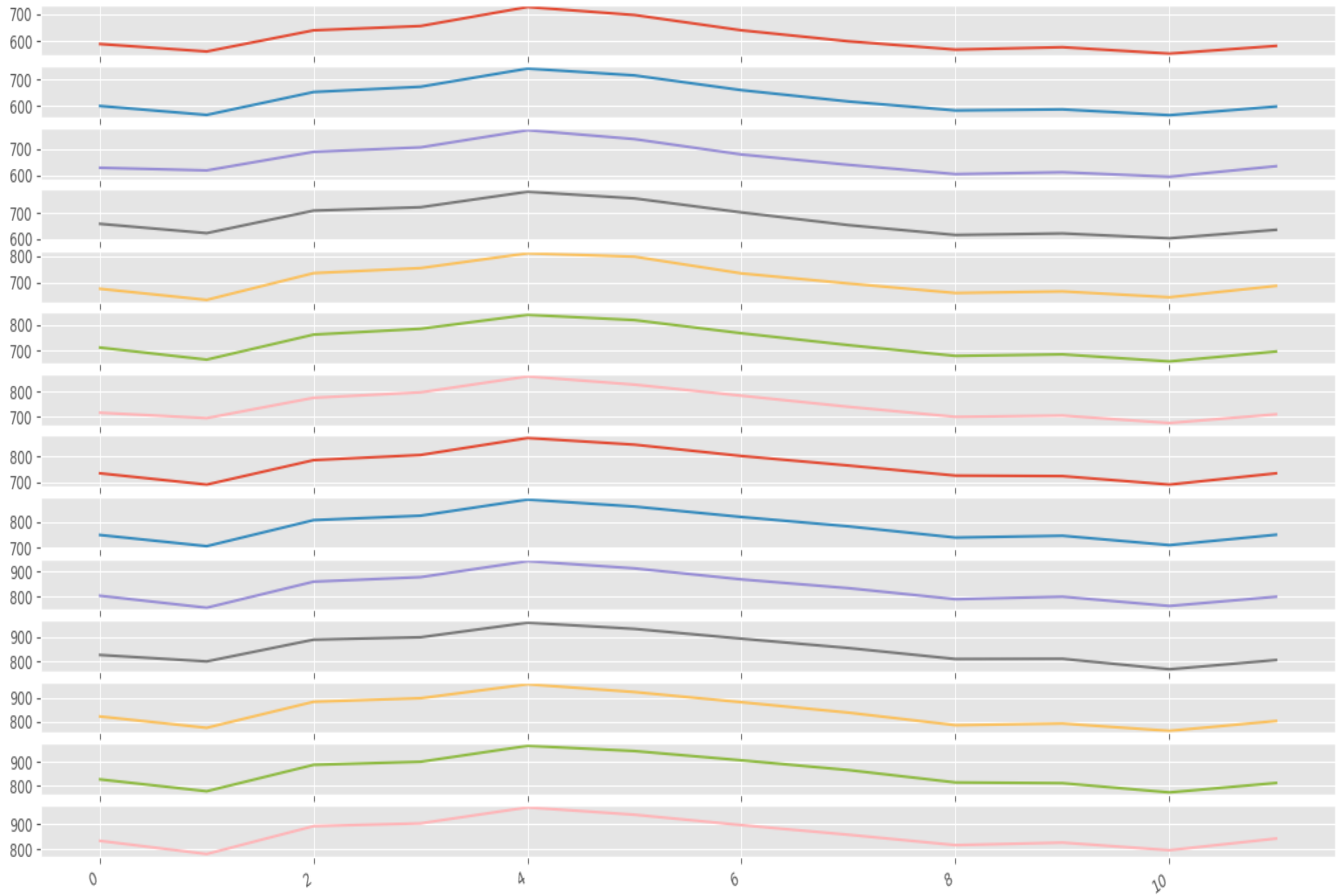
Trend

Seasonality

Cycle

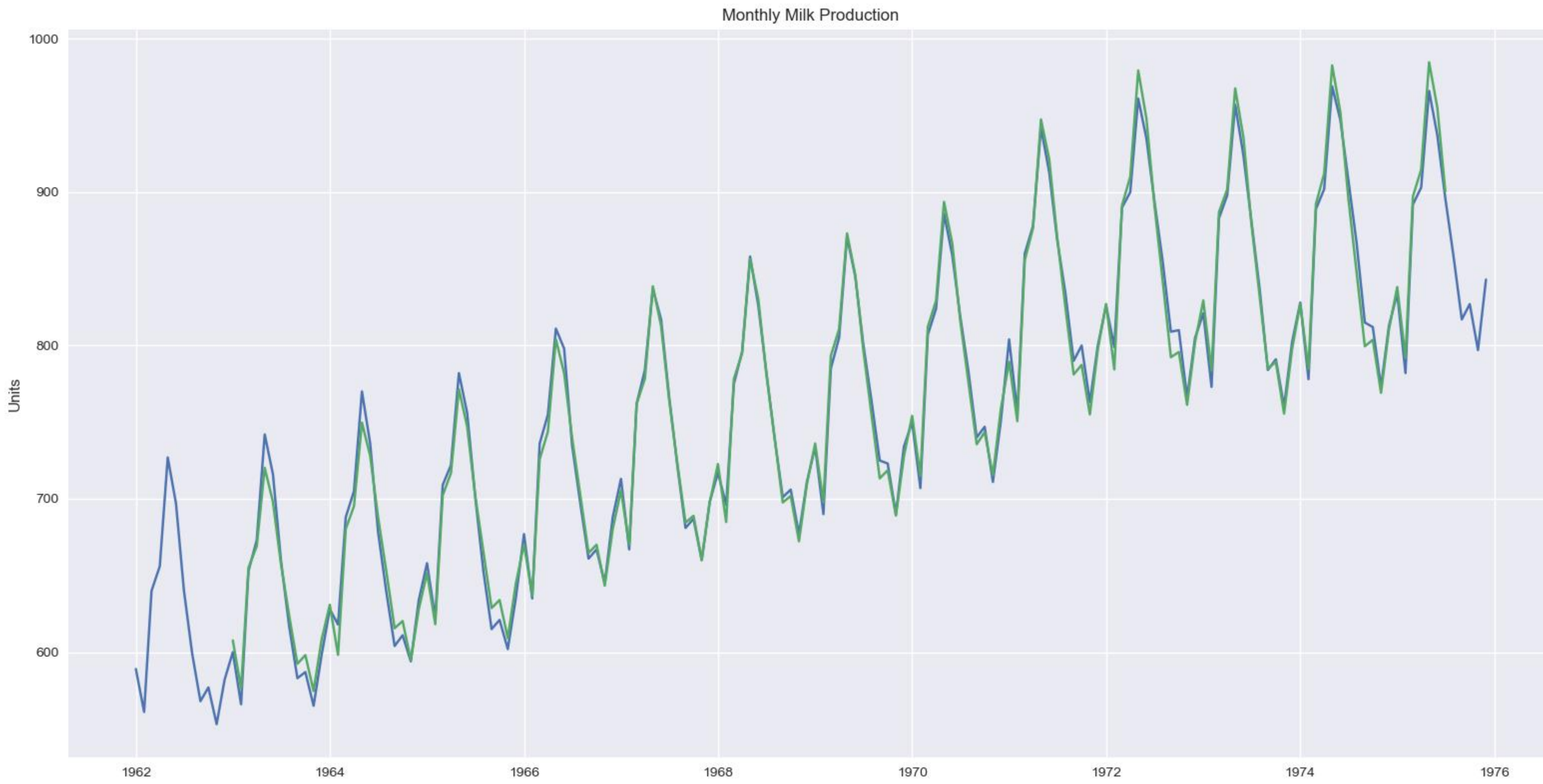
Noise


Seasonality
Units
1.15
1.10
1.05
1.00
0.95
0.90
1963
1965
1967
1969
1971
1973
1975

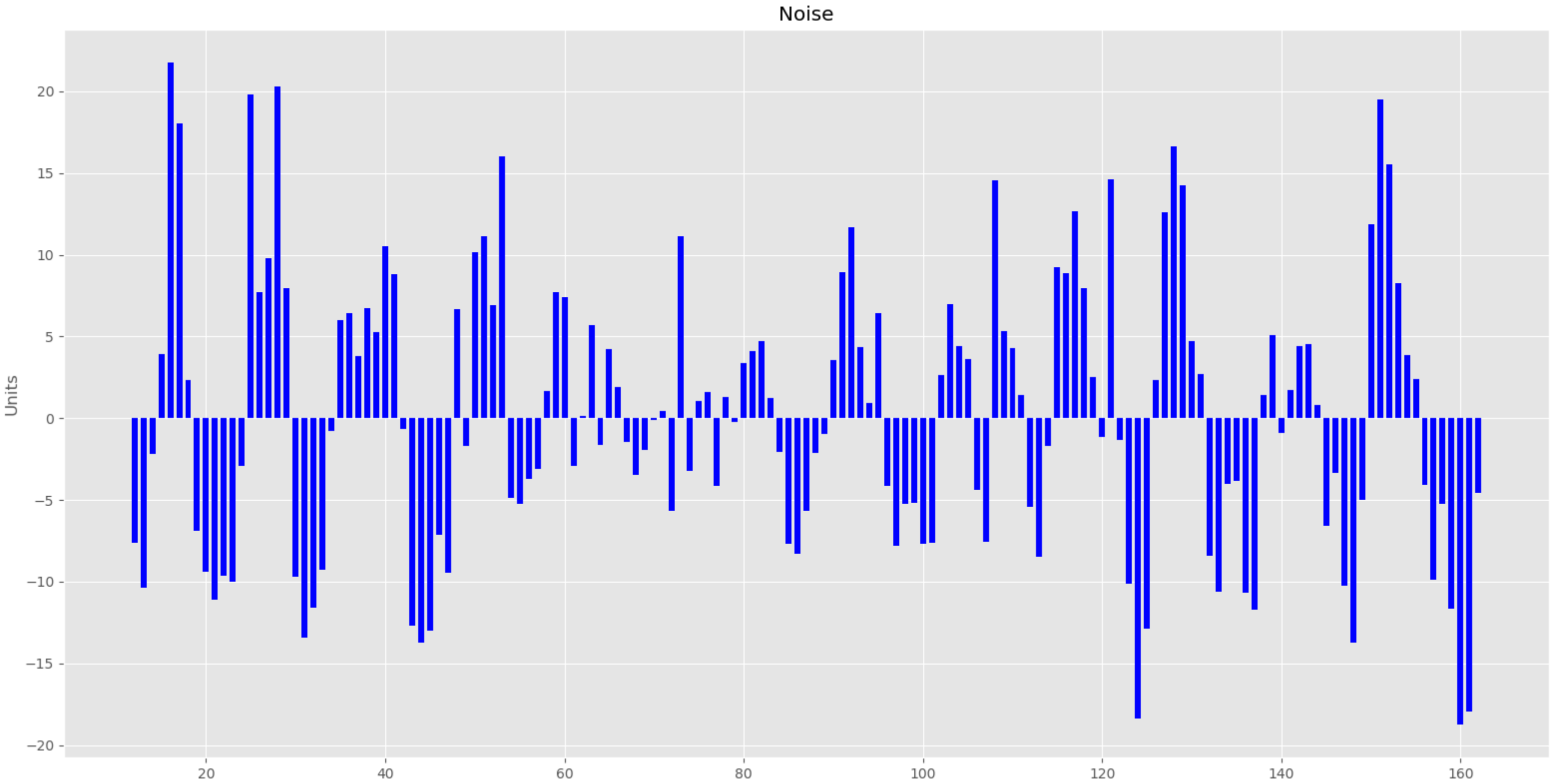

700
600
700
600
700
600
700
600
800
700
800
700
800
700
800
700
800
700
900
800
900
800
900
800
900
800
900
800
0
2
4
6
8
10

TIME SERIES

Trend	Seasonality
Cycle	Noise



Monthly Milk Production
1000
900
800
700
600
Units
1962
1964
1966
1968
1970
1972
1974
1976


Noise
Units
20
15
10
5
0
−5
−10
−15
−20
20
40
60
80
100
120
140
160

Forecasting

The trend (T) represents a growth or fall in the consumption of a certain item during some representative period of time. A time series may have a linear or a nonlinear trend.

Seasonality (S) is defined as the repetitive and predictable movement around the trend line within a year or less. It is detected visually within the time intervals that make up the year, such as: days, weeks, months, or quarters.

The cycle (C) can only be perceived if it is considered a large period of time, for example decades, and can be understood as a repetition of the consumption behavior of a particular item or set of items.

Finally, random events or prediction errors (e) are perturbations that may occur during the actual consumption period and affect it in some way. Prediction errors may also be correlated to variables not contemplated by the forecast model.

Forecasting

The demand forecast (P) can be calculated based on one of the following equations:

$$\text{Aditive} \rightarrow P = (T) + (S) + (C) + (e)$$

$$\text{Multiplicative} \rightarrow P = (T)(S)(C)(e)$$

Forecasting

GODNESS OF FIT

Forecasting

To measure the goodness of fit of a model, the following measurements are used:

a) CFE (Cumulative Forecasting Error)
b) ME (Mean Error)
c) MAE (Mean Absolute Error)
d) MSE (Mean Squared Error)
e) MPE (Mean Percentage Error)
f) MAPE (Mean Absolute Percentage Error)
g) U-Theil

Forecasting

- The error is calculated as:

$$e = D - P$$

- The Cumulative Forecasting Error (CFE) reports the bias of the forecast errors and is given by:

$$CFE = \sum_{i=1}^{n} e_i$$

Forecasting

- The Mean Error (ME) is given by:

$$ME = \frac{\sum_{i=1}^{n} e_i}{n}$$

- The ME will probably generate small values or be equal to zero (0) due to the trade-offs between the positive and negative values of the errors of each period. The ME also reports, on average, the bias of forecast errors.

Forecasting

- The Mean Absolut Error (MAE) is given by:

$$MAE = \frac{\sum_{i=1}^{n} |e_i|}{n}$$

- The MAE turns all errors into positive values and then an average is calculated. Being easy to interpret results and easier to explain to non-specialists.

Forecasting

- The Mean Squared Error (MSE) is given by:

$$MSE = \frac{\sum_{i=1}^{n} e_i^2}{n}$$

- The MSE turns all errors into positive values and then a mean is calculated. The MSE is not so easy to interpret but it manages to punish the model when large prediction errors occur.

Forecasting

- The MSE can also be used to find confidence intervals for the forecast:

$$P_{i+1} \pm z\sqrt{MSE}$$

- The MSE provides an estimate of the variance so the square root of the MSE provides an estimate of the standard deviation of the prediction error. Therefore it is assumed that the errors are normally distributed and have zero mean (0).

Forecasting

- The Mean Percentage Error (MPE) is given by:

$$MPE = \frac{\sum_{i=1}^{n}\left(\frac{D_i - P_i}{D_i}\right)}{n}$$

- The MPE allows to compare time series with different scales of magnitude, but it has the same disadvantages of the ME and can not be computed if the time series allows values equal to zero. Ex: Air temperature.

Forecasting

- The Mean Absolute Percentage Error (MAPE) is given by:

$$MAPE = \frac{\sum_{i=1}^{n} \left| \left(\frac{D_i - P_i}{D_i} \right) \right|}{n}$$

- The MAPE has the same characteristics of the MAE, but it also can not be used in time series that allow the existence of the value zero (0). MAPE transmits more information than the previous error measures.

Forecasting

- The U-Theil Statistic (developed by Theil in 1966) measures whether the model used is as good as the Naïve Forecast.

$$U = \sqrt{\frac{\sum_{i}^{n-1}\left(\frac{P_{i+1} - D_{i+1}}{D_i}\right)^2}{\sum_{i}^{n-1}\left(\frac{D_{i+1} - D_i}{D_i}\right)^2}}$$

Forecasting

- $U = 1 \rightarrow$ The evaluated model is as good as the Naïve Forecast.

- $U < 1 \rightarrow$ The evaluated model is better than the Naïve Forecast. The lower the value, the better the model.

- $U > 1 \rightarrow$ The evaluated model is worse than the Naïve Forecast. The higher the value, the worse the model.

References

- MAKRIDAKIS, S.G., WHEELWRIGHT, S.C, HYNDMAN, R.J. **Forecasting: Methods and Application**. 3rd Edition