**HW 1 – Business analytics.**

To solve part b) below, you will need to do the following:

* 1. Use the Generalized Analytics Procedure (GAP) to set up your problem as follows:
     1. Define your model in words
        1. Identify the objective function in words
        2. Identify the random variables in words
        3. Identify the decision variables in words
        4. Identify the constraints in words
     2. Formulate your model mathematically
        1. Define the decision variables
        2. Define the random variables
        3. Define the objective function in terms of decision variables and random variables
        4. Define the constraints in terms of the decision variables. Please include any non-negativity constraints in your formulation
  2. Solve the problem in Excel
  3. Answer the questions stated in the problem (in words).

Please submit only one file in PDF format with your write-up. Do not submit your Excel file. Your writeup must include the screenshots from your Excel Spreadsheets. If you make any additional assumptions, state them clearly.

**Australian Tabaco Production**

1. Tab “training data” of HW1 spreadsheet.xlsx spreadsheet contains 1990-2001 quarterly Australian Tobacco Production (in metric tons). Use the following three forecasting methods (see Lecture 1 slides for details) to forecast the production in the 10 quarters starting with Q1 2002, using the 1990-2001 data as your training data:
   * Simple Average
   * Naïve Seasonal
   * Drift

Create a line graph in Excel showing the predictions from the three forecasts. In addition to the forecasts, include the actual data (from 1990 to 2004) in your graph. Which method do you expect to perform best/worst?

Using the three forecasting methods, we can get the predictions:

A table with numbers and a few simple average

Description automatically generated

The line graph for actual data and predictions:

A graph with blue lines and orange lines

Description automatically generated

We expect Seasonal Naïve to have the best performance, as it includes seasonal effects while not putting too much emphasis on previous observations. Simple Average is expected to have the worst performance.

1. b) Complete part a) using R instead of Excel. Include screenshots of your R code, and the graph produced by R. Do not attach the file with the R code

A screenshot of a computer program

Description automatically generated

A graph showing the growth of tobacco production

Description automatically generated

1. You have been tasked to select the best of the three methods to forecast the production in the next 10 quarters (Q1 2002 – Q2 2004). Which method do you prefer? Answer this question using actual (realized) production quantities found in Q1 2002 – Q2 2004 displayed in tab “full data”. Follow the procedure on the previous page (in red), and an evaluation method of your choice (MAD or MSE).

**Define model in words:**

**Objective:** Predict tobacco production volume from Q1 2002 to Q2 2004

**Random Variables:** Actual tobacco production volume from Q1 2002 to Q2 2004

**Decision Variables:** Predictions for tobacco production volume from Q1 2002 to Q2 2004

**Constraints:** Non-negative

**Formulate model mathematically:**

**Decision Variables:** Predictions , , , etc… for tobacco. Each quarter

**Random Variables:** Actual volume , , , etc… for tobacco. each quarter

**Objective Function:** MAD: -

MSE: - -

**Constraints:**

We’ll calculate MAD and MSE using actual data. MAD is the average of absolute deviations for all predictions, and MSE is the average of squared deviations for all predictions.

|  |  |  |
| --- | --- | --- |
| **Absolute Deviations (MAD)** | | |
| **Simple Average** | **Seasonal Naïve** | **Drift** |
| 1749 | 374 | 636 |
| 786 | 153 | 371 |
| 312 | 84 | 888 |
| 883 | 187 | 361 |
| 1230 | 145 | 57 |
| 382 | 251 | 948 |
| 334 | 62 | 1040 |
| 1066 | 4 | 351 |
| 1618 | 243 | 157 |
| 565 | 68 | 939 |
| 892.65 | 157.10 | 574.81 |

|  |  |  |
| --- | --- | --- |
| **Squared Deviations (MSE)** | | |
| **Simple Average** | **Seasonal Naïve** | **Drift** |
| 3059511 | 139876 | 403982 |
| 618025 | 23409 | 137499 |
| 97435 | 7056 | 788922 |
| 779947 | 34969 | 130045 |
| 1513259 | 21025 | 3251 |
| 146035 | 63001 | 899511 |
| 111653 | 3844 | 1081246 |
| 1136667 | 16 | 123365 |
| 2618396 | 59049 | 24763 |
| 319390 | 4624 | 881801 |
| 1040031.83 | 35686.90 | 447438 |

Seasonal Naïve method has the smallest MAD and MSE. We’ll select Seasonal Naïve method.

1. (Bonus) Come up with a method that outperforms the methods used in a)/b).