**Time Series Forecasting Playbook – Excel Models**

**by Awale Abdi**

**General Process**

1. Create an “Input Sheet” where you’ve cleaned the date and sales (or whatever equivalent of sales is available in your dataset like tonnes of chocolate) columns into two separate and neat columns where you will also be adding future columns as instructed below like the “Period” column. The Professor prefers to create individual input sheets for each regression model, but I prefer one master input sheet to save time.
2. Create a period (numerical index) column (1, 2, 3, 4 etc) that counts the number of instances all the way down to the bottom for the dates and sales columns.
3. Create a lined scatter plot using the period and sales columns to observe the trend.
4. Choose an appropriate forecasting model based on the trend. If it looks linear then a linear model, if it looks exponential then an exponential model and so on. Or just go off the R-squared for each model to see which one returns the best number and go with that one.

**Linear Model**

* **Setup**:
  + Make sure you have columns for sales and period (numerical index).
* **Regression**:
  + Use Excel's Data Analysis tool and choose “Regression”.
  + Y = Sales, X = Period.
  + Make sure to press “Labels” and “Residuals” in the output for the Regression and keep in the titles of the columns in your selections of them.
  + Name the output sheet “Linear Regression”.
* **Error Calculation**:
  + A Residuals column should be available to you in the Regression’s output if you follow the last set of steps but if it isn’t present, you can make one with the formula “**=Actual Sales - Predicted Sales**”.
  + Add an “Actual Sales” column and just copy-paste the actual sales values from your Input Sheet.
  + Create an “Error” column next to the Actual Sales column that is calculated through “**=Actual Sales – Predicted Sales**”.
  + Create an “Error²” column next to the Error column through “**=Error^2**”.
  + Calculate the MSE through “**=AVERAGE(Error^2)”**.
  + Calculate the RMSE through “**=SQRT(MSE)”**.
* **Forecast**:
  + Create columns for future date and future period in the Regression sheet.
  + Create a “Forecasted Sales” column next to these and populate it with the Regression equation “**=Intercept + (Period \* Coefficient of Period)**”. Keep intercept and all coefficients constant (ctrl+f4 or direct dollar sign addition).
  + Have a “Total Sales” at the bottom that is simply calculated via “**=SUM(Forecasted Sales)**”.
  + Plot forecasted sales against the future dates using a LINE GRAPH with points.
* **Note**: Not suitable for data with seasonality.

**Exponential Model**

* **Setup**:
  + Add a “Log of Sales” column to the “Input Sheet” through “**=LN(Sales)**”. Be careful NOT to use =LOG(Sales) instead.
* **Regression**:
  + Use Excel's Data Analysis tool and choose “Regression”.
  + Y = Log of Sales, X = Period.
  + Ensure to tick "Labels" and "Residuals" and include the titles of your columns.
  + Name the output Sheet “Exponential Regression”.
* **Error Calculation**:
  + A Residuals column should be available to you in the Regression’s output if you follow the last set of steps but if it isn’t present, you can make one with the formula “**=Actual Sales - Predicted Sales**”.
  + Create a “Forecast” column next to Residuals with the formula **“=EXP(Predicted Log Sales)**”.
  + Create an “Actual Sales” column next to the Forecast one that is just the copy-pasted actual sales.
  + Create an “Error” column that is basically **“=Actual Sales-Forecast**”.
  + Create an “Error^2” via “**=Error^2**”.
  + Calculate the MSE via “**=AVERAGE(Error^2)”**.
  + Calculate the RMSE via “**=SQRT(MSE)”**.
* **Forecast**:
  + Create columns for future date and future period in the Regression sheet.
  + Create a “Forecasted Log” column next to these with the formula “**=Intercept + (Period \* Coefficient of Period)**”. Keep all coefficients and intercept constant (ctrl+f4 or directly dollar sign addition).
  + Create a “Forecasted Sales” column next to that converts “Forecasted Log” to actual forecasted sales via the formula “**=EXP(Forecasted Log)”**.
  + Have a “Total Sales” at the bottom that is simply calculated via “**=SUM(Forecasted Sales)**”.
  + Plot forecasted sales against the future period using a LINE GRAPH with points.
* **Note**: Useful for exponential growth trends.

**Quadratic Model**

* **Setup**:
  + Add a column for “Period²” next to the Period column through “**=Period^2”** in the “Input Sheet”.
* **Regression**:
  + Use Excel's Data Analysis tool and choose “Regression”.
  + Y = Sales, X = Period, Period² (make sure the columns are next to one another or you can’t select them both together like this).
  + Ensure to tick "Labels" and "Residuals" and include the titles of your columns.
  + Name the output sheet “Quadratic Regression”.
* **Error Calculation**:
  + A Residuals column should already exist in the Regression’s output but if it somehow doesn’t you can calculate it via “**Actual Sales - Predicted Sales**”.
  + Create an “Error²” column that is calculated through “**=Residual^2**”.
  + Calculate the MSE via “**=AVERAGE(Error^2)”**.
  + Calculate the RMSE via “**=SQRT(MSE)”**.
* **Forecast**:
  + Create columns for future date and future period in the Regression sheet.
  + Create a “Forecasted Sales” column next to these and use the regression equation “**=Intercept + (Period \* Coefficient of Period) + (Period^2 \* Coefficient of Period^2)**.” Keep all coefficients and intercept constant (ctrl+f4 or directly dollar sign addition).
  + Sum forecasted sales at the bottom via “**=SUM(Forecasted Sales)**.”
  + Plot forecasted sales against the future period using a LINE GRAPH with points.
* **Note**: Better if the relationship is not linear.

**Linear with Dummy Variables**

* **Setup**:
  + Create a “Year” column that extracts the Year value from the date through methods like “=YEAR()” or whatever method you come up with.
  + Add dummy variable columns based on what you have (one for each month, one for each quarter etc) and populate them with 1s and 0s depending on the month, quarter or whatever else they correspond with (but always keep one (like a month or quarter or what have you) out and populate its row with 0s).
* **Regression**:
  + Use Excel's Data Analysis tool and choose “Regression”.
  + Y = Sales, X = Year + Dummy Variables. (make sure the Year column is now pasted in as values if you used a formula to extract the year and that its format is numeric now).
  + Ensure to tick "Labels" and "Residuals" and include the titles of your columns.
  + Label the output sheet “Linear Regression w Dumm”.
* **Error Calculation**:
  + A Residuals column should be present in your output but in case it’s somehow not, it is calculated via “**=Actual Sales - Predicted Sales”**.
  + Create an “Error²” column next to the Residuals column and populate it with the formula **“=Residual^2”**.
  + Calculate MSE through **“=AVERAGE(Error^2)”**.
  + Calculate RMSE through “**=SQRT(MSE)**”.
* **Forecast**:
  + Create columns for future date, future period, and future year in the Regression sheet.
  + Create columns for the corresponding dummy variables next to these.
  + Create a “Forecasted Sales” column after the dummy variable ones and use this regression formula to calculate the forecasted sales “**=Intercept + (Year \* Year Coefficient) + SUMPRODUCT(Dummy Variables, Dummy Coefficients)”**. Keep all coefficients and intercept constant (ctrl+f4 or directly dollar sign addition).
    - Copy-paste the dummy variable coefficients in a separate column or row when you use them for the SUMPRODUCT part of the formula. You will experience issues otherwise. (see doc @ end of cheat sheet for example)
  + Sum the forecasted sales to get “Total Sales” via “**=SUM(Forecasted Sales)**”
  + Plot forecasted sales against the future dates using a LINE GRAPH with points.
* **Note**: Helps account for seasonality.

**Exponential with Dummy Variables**

* **Setup**:
  + Create a “Log of Sales” column through the formula **“=LN(Sales)**”. Ignore this step if you already have it made from doing Exponential Regression previously.
* **Regression**:
  + Use Excel's Data Analysis tool and choose “Regression”.
  + Y = Log of Sales, X = Year, Dummy Variables.
  + Follow the steps in the Linear Regression with Dummy Variables example above to know how to create the “Year” and “Dummy Variables” columns.
  + Ensure to tick "Labels" and "Residuals" and include the titles of your columns.
  + Name your new sheet “Exponential Regression w Dumm”.
* **Error Calculation**:
  + Create a “Forecast” column next to Residuals with the formula **“=EXP(Predicted Log Sales)**”.
  + Create an “Actual Sales” column next to the Forecast one that is just the copy-pasted actual sales from your Input Sheet.
  + Create an “Error” column next to this that is basically “**=Actual Sales-Forecast**”
  + Create an “Error^2” column next to this with the formula “**=Error^2**”.
  + Calculate MSE through **“=AVERAGE(Error^2)”**.
  + Calculate RMSE through “**=SQRT(MSE)**”.
* **Forecast**:
  + Create columns for future date, future period, and future year in the Regression sheet.
  + Create columns for the corresponding dummy variables next to these.
  + Create a “Forecast Log” column after these with the formula “**=Intercept + (Year \* Year Coefficient) + SUMPRODUCT(Dummy Variables, Dummy Coefficients)”**. Keep all coefficients and intercept constant (ctrl+f4 or directly dollar sign addition).
    - Copy-paste the dummy variable coefficients in a separate column or row when you use them for the SUMPRODUCT part of the formula. You will experience issues otherwise.
  + Create a “Forecasted Sales” column next to Forecast Log with the formula ”**=EXP(Forecast Log)**”.
  + Get the total forecasted sales through “**=SUM(Forecasted Sales)**”
  + Plot forecasted sales against the future Date using a LINE GRAPH with points.
* **Note**: Combines exponential growth with seasonality.

**Quadratic with Dummy Variables**

* **Setup**:
  + Create a Year² column next to the Year column you would have made for the Linear method above through the formula “**=Year^2”**.
* **Regression**:
  + Use Excel's Data Analysis tool and choose “Regression”.
  + Y = Sales, X = Year, Year², Dummy Variables.
  + Follow the steps in the Linear Regression with Dummy Variables example above to know how to create the “Year” and “Dummy Variables” columns.
  + Ensure to tick "Labels" and "Residuals" and include the titles of your columns.
  + Name your new sheet “Quadratic Regression w Dumm”.
* **Error Calculation**:
  + A Residuals column should be present in your output but in case it’s somehow not it is calculated via “**=Actual Sales - Predicted Sales”**.
  + Create an Error² column next to the Residuals column and populate it with the formula **“=Residual^2”**.
  + Calculate MSE through **“=AVERAGE(Error^2)”**.
  + Calculate RMSE through “**=SQRT(MSE)**”.
* **Forecast**:
  + Create columns for future date, future period and future year in the Regression sheet.
  + Create columns for the dummy variables next to these.
  + Create a “F-Year^2” column next to all of these that is calculated via **“=futureyear^2**”
  + Create a “Forecasted Sales” columns and populate via the regression equation: **=Intercept + (Year \* Year Coefficient) + (Year² \* Quadratic Coefficient) + SUMPRODUCT(Dummy Variables, Dummy Coefficients)**. Keep all coefficients and intercept constant (ctrl+f4 or direct dollar sign addition).
    - Copy-paste the dummy variable coefficients in a separate column or row when you use them for the SUMPRODUCT part of the formula. You will experience issues otherwise. (see doc @ end of cheat sheet for example)
  + Sum the forecasted sales for a “Total Sales” through “**=SUM(Forecasted Sales)”**.
  + Plot forecasted sales against the future dates using a LINE GRAPH with points.
* **Note**: Suitable for quadratic trends with seasonality.

**Seasonal Factors Initial Setup**

1. Create a new sheet named "Seasonal Factors Sheet".
2. **Input Data**: Copy the Date, Sales, and Period columns from your Input Sheet to the Seasonal Factors Sheet.
3. **Calculate Moving Averages**:
   * Create a column for “Moving Averages (MA)” by averaging sales over each period of the cycle (e.g., 4 periods for quarterly data or 12 for Monthly data over a year).
   * Formula is “**=AVERAGE(previous periods)**”.
   * Start at the Period you’re beginning the averages so if you have 4 quarters you start the calculating at Quarter 4’s row or 12 months you start at Month 12’s row.
   * See my attached Excel doc at the end of this word doc to see an example.
4. **Calculate Ratios to Moving Averages**:
   * Create a column called “Ratio-to-Moving Average”.
   * Formula: “**=Sales/Moving Average**”.
5. **Calculate Seasonal Factors**:
   * Create a column for Seasonal Factors by averaging the ratios for the same periods in different cycles.
   * Formula “=AVERAGE(all ratios for the same period across cycles)”.
   * Once you’ve established the Seasonal Factor for each period (like all four Quarters or all 12 Months) highlight and save them in your Seasonal Factors Sheet.
6. Create a “Seasonal Input Sheet” with the Date, Sales, and period columns.
7. Add a Seasonal Factors column to this sheet. Then input the corresponding seasonal factors per period from your “Seasonal Factors Sheet”. Do it on repeat so every time it’s the same Seasonal Factor for the same Period (like Quarter 3) going all the way down the list (see my attached Excel file at the very bottom of this Word doc).
8. You are now ready for the individual Regressions with Seasonal Factors that will be covered below. Use the “Seasonal Input Sheet” for all of their inputs. Again, the professor himself prefers separate input sheets per model but I prefer one master sheet to save time.

**Deseasonalized Linear Model**

1. Add a “Deseasonalized Sales” column to the “Seasonal Input Sheet” and populate it with the formula “**=Seasonal Factors/Sales”**
2. **Regression**:
   * Use Excel's Data Analysis tool and choose "Regression".
   * Y = Deseasonalized Sales, X = Period.
   * Ensure to tick "Labels" and "Residuals" and include the titles of your columns.
   * Name the output sheet "Deseasonalized Linear Regression".
3. **Error Calculation**:
   * Use the “Residuals” column provided in the regression output or if it’s somehow not present you can calculate it via **“=Deseasonalized Sales - Predicted Sales**”.
   * Create a “Seasonal Factors” column next to it that is basically a copy-paste of your Seasonal Factors column from your “Seasonal Input Sheet”.
   * Create a “Forecast Seasonalized” column next to this that is populated via **“=Predicted Deseasonalized Sales\*Seasonal Factors**”.
   * Create an “Actual Sales” column next to this one that is basically a copy-paste of your Actual Sales from your “Seasonal Input Sheet”.
   * Create an “Error” column next to this that is calculated via “**=Actual Sales-Forecast Seasonalized**”.
   * Create an “Error²” column next to this via “**=Error^2**”.
   * Calculate MSE via “**=AVERAGE(Error^2)**”.
   * Calculate RMSE via “**=SQRT(MSE)**”.
4. **Forecast**:
   * Create columns for future date and future period in the Regression sheet.
   * Create a “Deseasonalized Forecast” column next to these and populate it with **“=Intercept + (Period \* Coefficient of Period)**”. Make sure to keep the intercept and coefficient of period constants (ctrl+f4 or directly added dollar signs).
   * Create a column for “Seasonal Factors” next to this and fill in the corresponding seasonal factor for each period in the future date column.
   * Create a “Forecasted Sales” column and population it with **“=Deseasonalized Forecast \* Seasonal Factors**”.
   * Sum the forecasted sales via **“=SUM(Forecasted Sales)**”.
   * Plot forecasted sales against the future date using a LINE GRAPH with points.

**Deseasonalized Exponential Model**

1. Create a “Deseasonalized Log Sales” column in your “Seasonal Input Sheet” through the formula **“=LN(Deseasonalized Sales)**”.
2. **Regression**:
   * Use Excel's Data Analysis tool and choose "Regression".
   * Y = Deseasonalized Log Sales, X = Period.
   * Ensure to tick "Labels" and "Residuals" and include the titles of your columns.
   * Name the output sheet "Deseasonalized Expo Regression".
3. **Error Calculation**:
   * Use the residuals provided in the regression output or calculate them via “**=Deseasonalized Log Sales - Predicted Log Sales**”.
   * Create a “Deseasonalized Forecast” column with the formula **“=EXP(Predicted Deasonalized Log Sales)**”.
   * Create a column next to it titled “Seasonal Factors” and copy-paste in the seasonal factors from top to bottom from your Seasonal Input Sheet.
   * Create a column named “Forecast” next to this and populate it via the formula **“=Deseasonalized Forecast\*Seasonal Factors**”.
   * Create a column next to this one named “Actual Sales” and populate it with the copy-pasted sales data from top to bottom from your Seasonal Input Sheet.
   * Create a column named “Error” next to this one and populate it with the formula **“=Actual Sales-Forecast**”.
   * Create an “Error²” column next to this one and populate it with the formula “**=Error^2**”.
   * Calculate MSE via “**=AVERAGE(Error^2)**”.
   * Calculate RMSE via “**=SQRT(MSE)**”.
4. **Forecast**:
   * Create columns for future date and future period.
   * Create a “Forecast Log” column using the regression equation “**=Intercept + (Period \* Coefficient of Period)”.** And make sure to make the Intercept and Coefficient of Period constants by pressing “ctrl+F4” or adding dollar signs to them.
   * Create a “Deasonalized Forecast” column via the formula “**=EXP(Forecast Log)**”**.**
   * Create a “Seasonal Factors” column and fill it with the appropriate corresponding Seasonal Factor per period (i.e. Quarter 4).
   * Create a “Forecasted Sales” column and populate it via the formula **“=Deseasonalized Forecast\*Seasonal Factors**”.
   * Sum the forecasted sales via “**=SUM(Forecasted Sales**” at the bottom.
   * Plot forecasted sales against the future period using a LINE GRAPH with points.

**Deseasonalized Quadratic Model**

1. Create a new column in your “Seasonal Input Sheet” titled “Period^2” and populate it with “=Period^2”. Make sure it is right next to the Period column (before or after).
2. **Regression**:
   * Use Excel's Data Analysis tool and choose "Regression".
   * Y = Deseasonalized Sales, X = Period, Period².
   * Ensure to tick "Labels" and "Residuals" and include the titles of your columns.
   * Name the output sheet "Deseasonalized Quad Regression".
3. **Error Calculation**:
   * There should be a Residuals column present but if somehow it isn’t, calculate it via “**=Deseasonalized Sales - Predicted Sales**”.
   * Next to it create a “Seasonal Factors” column and simply fill it with the Seasonal Factors data from your Seasonal Input Sheet.
   * Next to it create a “Forecast Seasonalized” column via “**=Predicted Deseasonalized Sales\*Seasonal Factors**”.
   * Next to it create an “Actual Sales” column and populate it with the copy-pasted Actual sales values from your Seasonal Input Sheet.
   * Now create an “Error” column that is “**=Actual Chocolate-Forecast Seasonalized**”.
   * Then create an “Error²” column calculated via “**=Error^2**”.
   * Calculate MSE: “**=AVERAGE(Error^2)**”.
   * Calculate RMSE: “**=SQRT(MSE)**”.
4. **Forecast**:
   * Create columns for future date, future period and future period^2 (=futureperiod^2).
   * Create a “Deseasonalized Forecast” column with the formula “**=Intercept + (Period \* Coefficient of Period) + (Period² \* Coefficient of Period²)**” and make sure to keep the elements other than Period and Period² constant (ctrl+f4 or directly added dollar signs).
   * Add a “Seasonal Factors” column and make sure to add the correct Seasonal Factors for the corresponding Future Periods.
   * Create a “Forecasted Sales” column and populate it with “**=Deseasonalized Forecast \* Seasonal Factors**”.
   * Sum forecasted sales with a total at the bottom using “**=SUM(Forecasted Sales)**”.
   * Plot forecasted sales against the future dates using a LINE GRAPH with points.

**Notes**

Here is a Google Drive link to my Excel sheet using the Chocolate Dataset provided in class along with the original dataset:

[Original Dataset](https://docs.google.com/spreadsheets/d/1zat3xUk3u54GkoiDWY7tueFODyrwNjAf/edit?usp=sharing&ouid=108775598614327692311&rtpof=true&sd=true)

[Awale Abdi’s File](https://docs.google.com/spreadsheets/d/1e8oRgHbEV53-nuSm7PMSrLqpl8_M0cIx/edit?usp=sharing&ouid=108775598614327692311&rtpof=true&sd=true)

You may notice at times that some of your numbers like in the Regression sheets are slightly different from mine or the professor’s sheets. Don’t immediately panic. Both myself and the professor just have a habit of reducing the number of decimal points to 1-2 or 0 and introducing commas to larger numbers in the thousands, hundreds of thousands, or millions. Via this part of the home ribbon:

A screenshot of a computer

Description automatically generated

Whilst also turning the Adjusted R-square into a 1 decimal pointed percentage value. So, keep that in mind but if your numbers are still different from mine and the professor’s then you probably messed up somewhere and need to recheck your steps.

**Good luck with your exam and all your future endeavours!**