

## Data Cleaning, Exploration, Analysis & Forecast in MS Excel

**Dataset:** The dataset is from an e-commerce company having daily sales data from two of their warehouses, EW1 & EW2. It is important to note that for EW1, we have 4 months of data from Aug-2021 to Nov-2021 and for EW2, we have 3 months of data from Sep-2021 to Nov-2021.

### Data Exploration

By using Pivot tables in excel, the daily orders are converted into monthly to have a general idea of total monthly orders, as shown in *Figure 1*.

Row Labels	Sum of total_orders
<b>Aug</b>	<b>12430</b>
EW1	12430
<b>Sep</b>	<b>21125</b>
EW1	15016
EW2	6109
<b>Oct</b>	<b>47652</b>
EW1	13138
EW2	34514
<b>Nov</b>	<b>15325</b>
EW1	3379
EW2	11946
<b>Grand Total</b>	<b>96532</b>

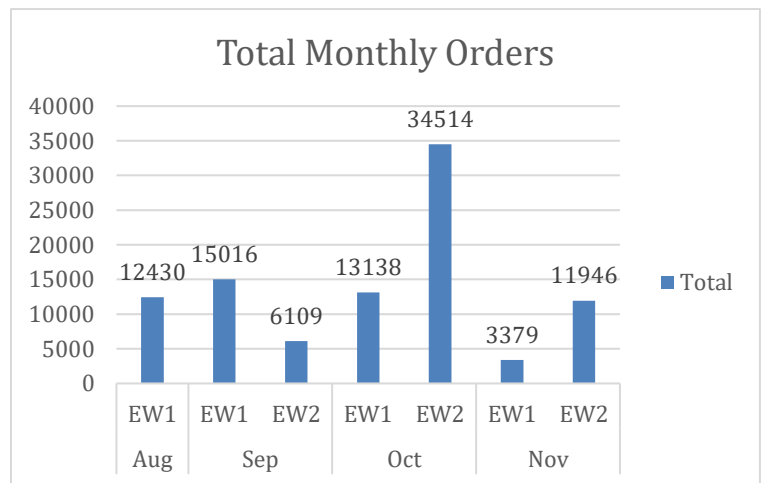


Figure 1

### Daily Analysis on EW1

Daily timeframe shows the high volatility of the number of orders and there is no seasonality either in EW1. There is a sudden deviation on 05/09/2021 and then dropped back to its previous state. A trendline of moving average with period value of 7 is used to smooth out the deviation and see weekly performance, as shown in the *Figure 2*.

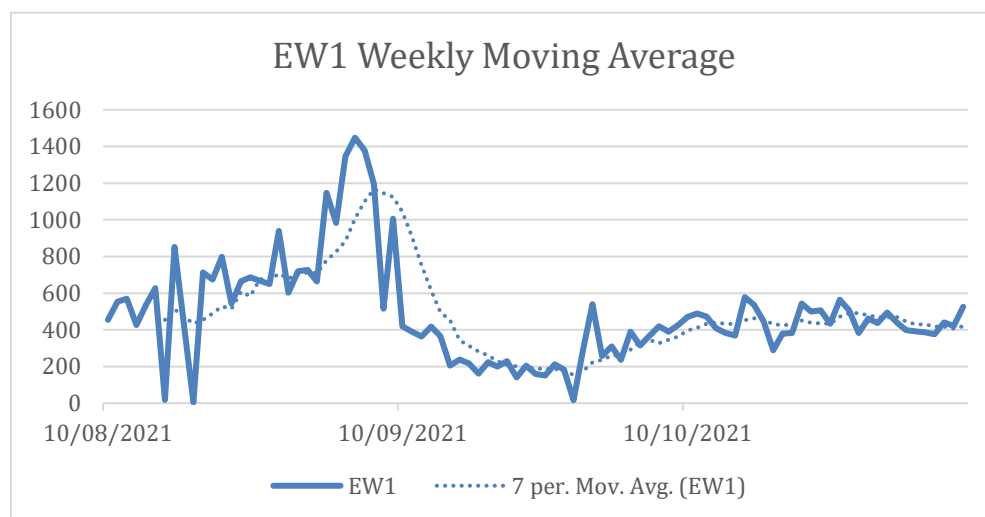


Figure 2

---

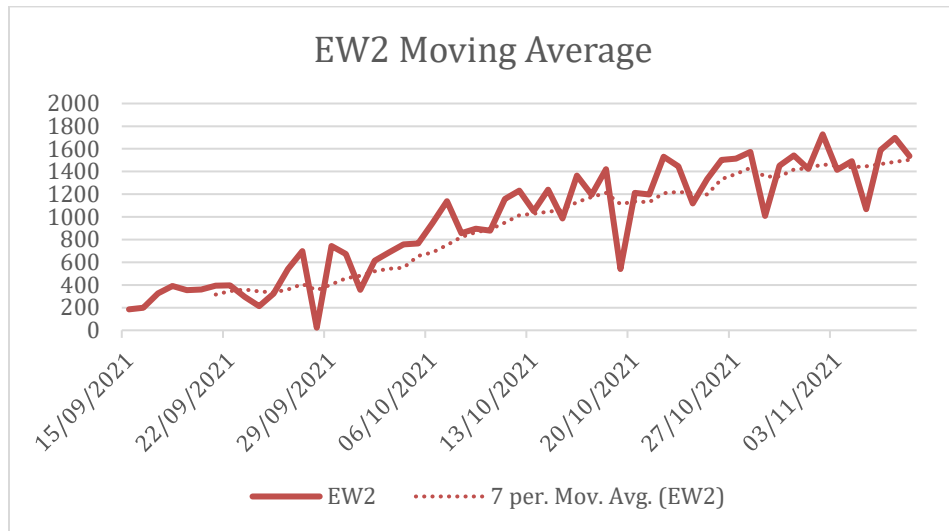
## Data Cleaning, Exploration, Analysis & Forecast in MS Excel

---

### Daily Analysis on EW2

There is not much difference between EW1 and EW2 in terms of high volatility. However, there is a clear upward trend in EW2, as shown in

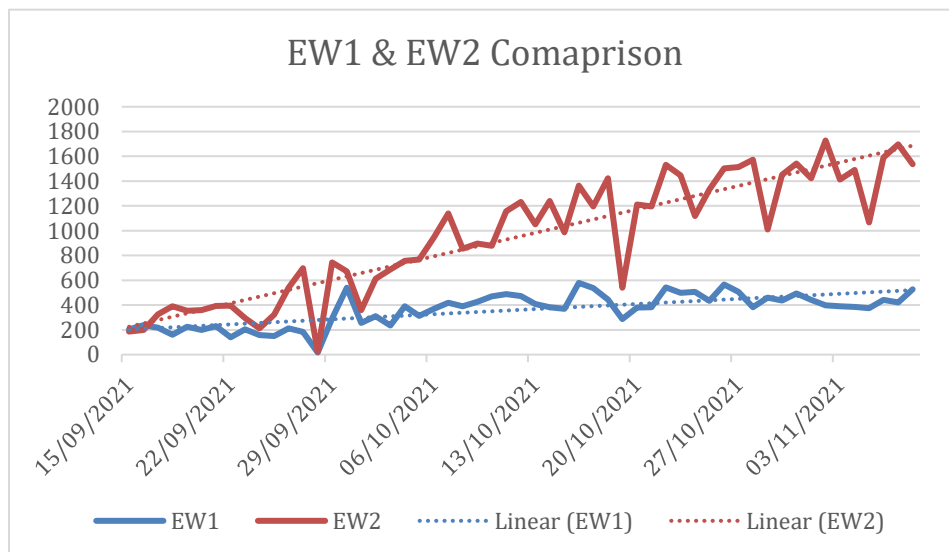
*Figure 3*. But just like EW1, there is no seasonality either in EW2. The trendline of moving average with period value of 7 is used for EW2 as well to smooth out the deviation.



*Figure 3*

### Comparing EW1 & EW2

Comparing both EW1 & EW2 with linear trendline, clearly shows better performance from EW2 and increasing number of orders over the period of 3 months, as shown in *Figure 4*.



*Figure 4*

## Data Cleaning, Exploration, Analysis & Forecast in MS Excel

It is important to note that only data available for both EW1 and EW2 is used for the above chart, so it results in showing a slightly upward trend in EW1, however, that is not the case. The **Error! Reference source not found.** shows a clear downtrend in EW1.

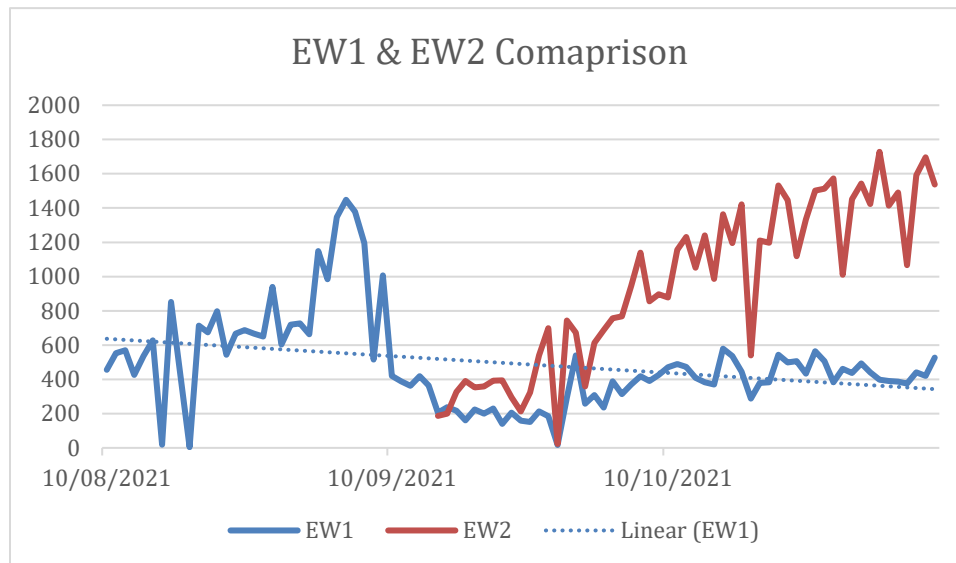


Figure 5

The comparison clearly shows better performance from EW2, and it is gradually increasing each week. Whereas EW1 is almost flat after the big drop in the first week September. Also, there is no seasonality in both EW1 and EW2 and are very volatile, with EW2 being slightly less volatile.

### Forecasting Orders for EW1

Linear Regression model is used to forecast next week orders for EW1. It is important to note here that R-squared value is very low as shown in upper right corner in Figure 6. Although, a low value is not always necessarily bad, but it leaves room for error. The forecast of daily orders for the next week are shown in Table 1 which continues the downtrend for EW1.

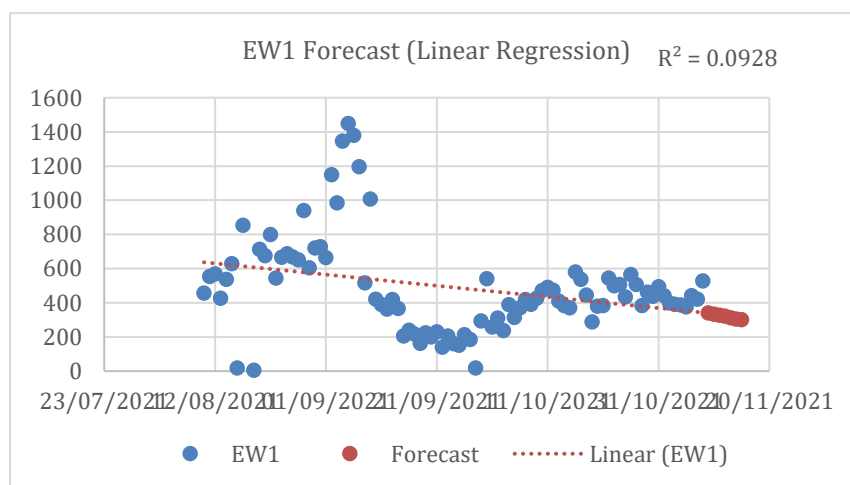


Figure 6

Table 1

09/11/2021	339.5671
10/11/2021	331.8583
11/11/2021	326.1777
12/11/2021	320.7292
13/11/2021	311.1246
14/11/2021	304.081
15/11/2021	299.5526

# Data Cleaning, Exploration, Analysis & Forecast in MS Excel

## Forecasting EW2

Using same Linear regression model for EW2; the R-squared value is high enough to be considered as fit for the linear regression model, as shown in

Figure 7. The forecast of daily orders for the next week are shown in *Table 2*, which continues the upward trend in EW2.

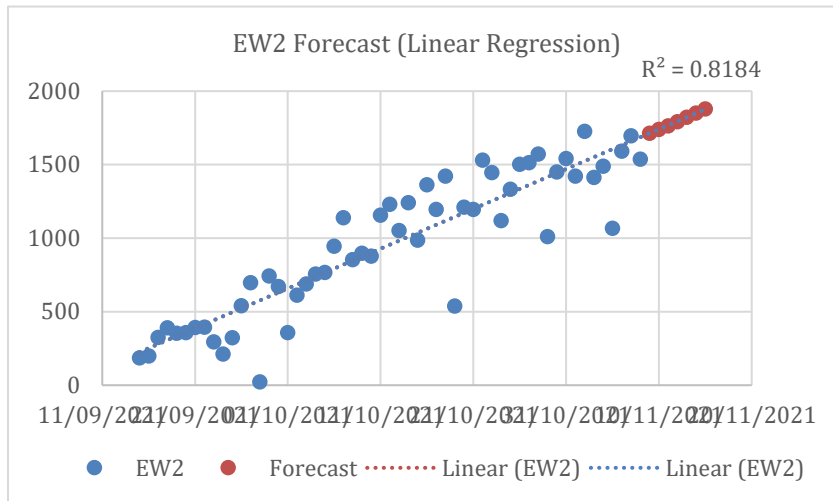


Table 2

09/11/2021	1713.772
10/11/2021	1739.238
11/11/2021	1763.755
12/11/2021	1792.52
13/11/2021	1823.564
14/11/2021	1851.94
15/11/2021	1879.325

Figure 7

## EW1 Forecast with Exponential Smoothing (ETS)

Using exponential triple smoothing (ETS) to forecast next week orders, with confidence interval of 95%, as shown in Figure 8.

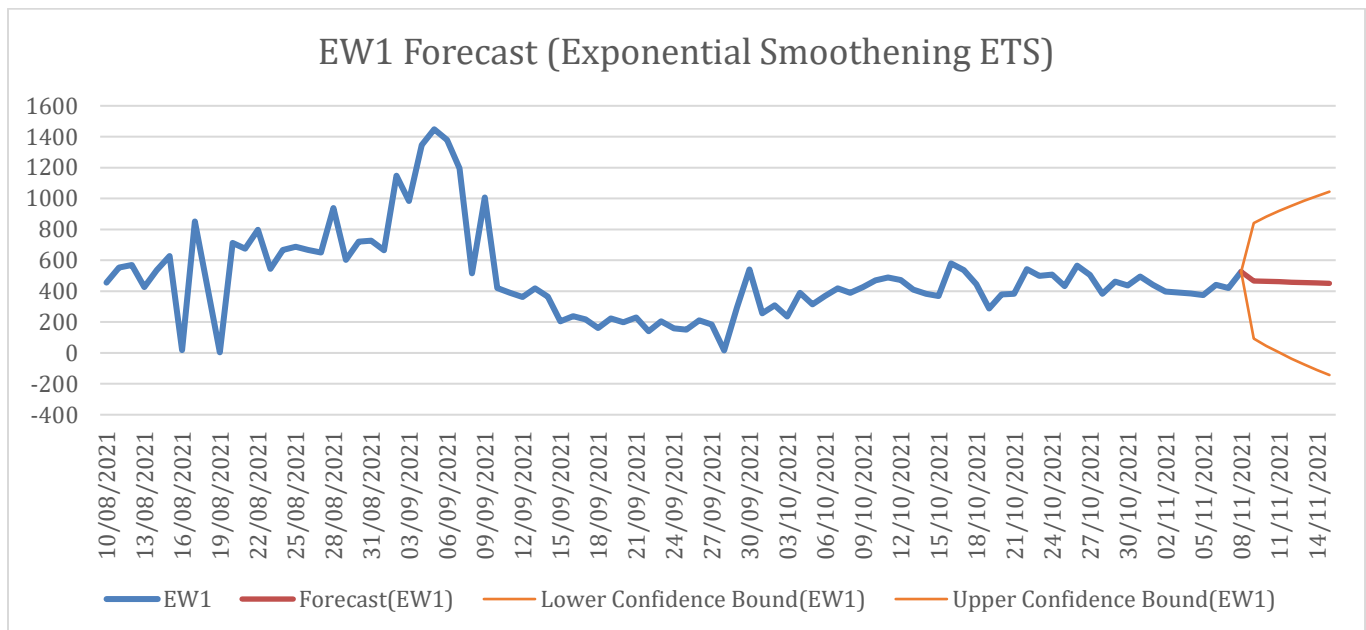


Figure 8

## Data Cleaning, Exploration, Analysis & Forecast in MS Excel

Table 3

Date	Forecast(EW1)	Lower Confidence Bound(EW1)	Upper Confidence Bound(EW1)
09/11/2021	466.2829304	92.53	840.03
10/11/2021	463.6643346	45.63	881.70
11/11/2021	461.0457389	2.84	919.25
12/11/2021	458.4271431	-36.84	953.70
13/11/2021	455.8085474	-74.08	985.69
14/11/2021	453.1899517	-109.31	1015.69
15/11/2021	450.5713559	-142.86	1044.01

### Conclusion for EW1

The next week orders forecast using ETS for EW1 are shown in *Table 3*. Looking at the forecast numbers for EW1 from linear regression model in *Table 1*, there is a huge difference between the results. It is probably due to the low r-squared value that shows linear regression model is not fit for the dataset. ETS in *Table 3* shows more accurate results with almost a flat line, slightly downwards.

### EW2 Forecast with Exponential Smoothing (ETS)

Using the same exponential triple smoothing (ETS) model to forecast next week's orders for EW2, with confidence interval of 95%, as shown in *Figure 9*.

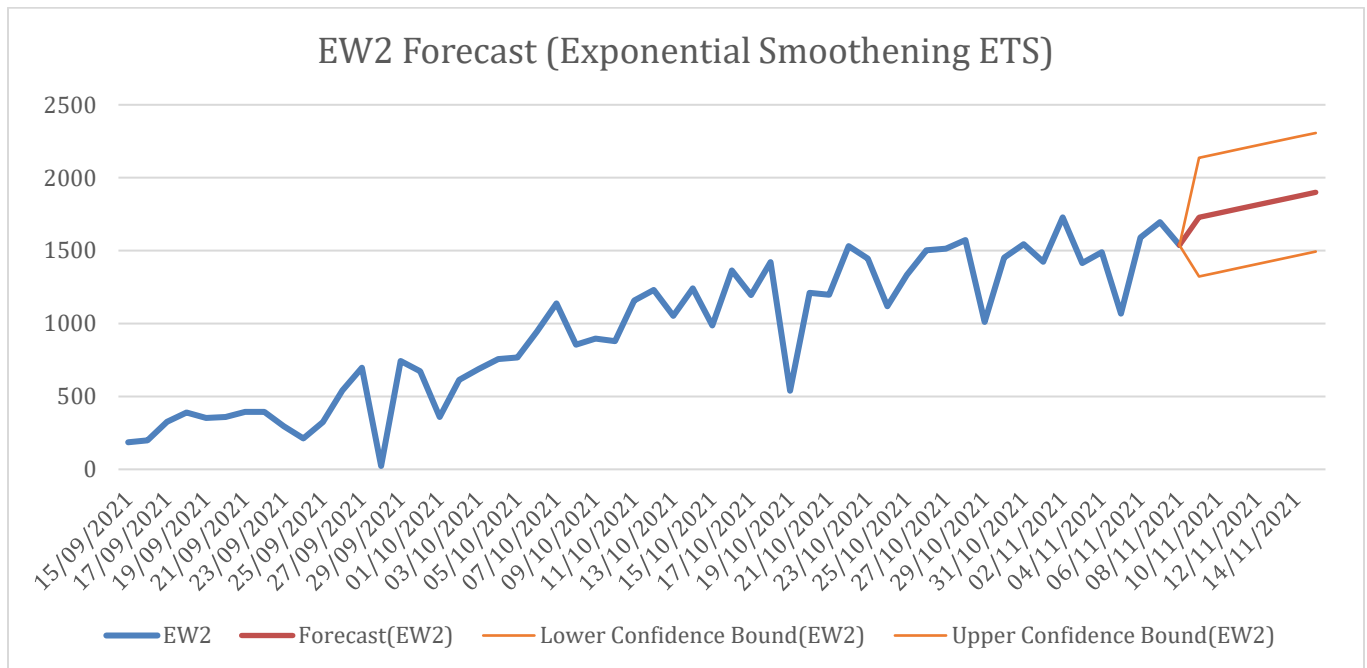


Figure 9

---

## Data Cleaning, Exploration, Analysis & Forecast in MS Excel

---

Table 4

Date	Forecast(EW2)	Lower Confidence Bound(EW2)	Upper Confidence Bound(EW2)
09/11/2021	1729.369	1322.59	2136.15
10/11/2021	1757.794	1351.01	2164.58
11/11/2021	1786.22	1379.43	2193.00
12/11/2021	1814.645	1407.85	2221.44
13/11/2021	1843.071	1436.27	2249.87
14/11/2021	1871.496	1464.69	2278.30
15/11/2021	1899.922	1493.10	2306.74

### Conclusion for EW2

The next week orders forecast using ETS for EW2 are shown in *Table 4*. Interestingly this time, there is not much difference when comparing forecast numbers for EW2 with EW1 using linear regression model in *Table 2*. It is because the R-squared value shown in *Figure 7* was high enough to be fit for the EW2 dataset. In this case, both ETS and linear regression model shows almost equal results, continuing the upwards trend in EW2.