Superstore Data Time-Series Analysis

Tableau Visualization

Problem Statement

Business leaders of a global superstore have noted that seasonal trends in the sales data are hard to detect, making it difficult for the business to plan resources and marketing campaigns effectively. To address this, they have provided a dataset covering four years (January 2015 to December 2018) for analysis.

Objective

Analyzed the dataset using time-series in Excel to highlight seasonal peaks and troughs. Create visual forecasts to help the business prepare for demand fluctuations, ensuring better resource allocation.

Questions we can answer with our analysis:

1. What seasonal trends are evident in the sales data, and how do these patterns inform marketing and inventory strategies?
2. How have the low-sales periods evolved over time, and what does this suggest about the business's growth or operational changes?
3. How does the six-month sales forecast support decision-making for resource allocation in the upcoming period?
4. What was the highest-performing month in the dataset, and what factors might have contributed to its exceptional sales performance?
5. How can the identification of sales peaks and troughs be leveraged to drive targeted promotions and enhance customer engagement?

Key Terms Used in This Project

- **Time-Series Analysis**: This is a method of analyzing data points collected or recorded at successive points in time. It focuses on understanding patterns, such as trends, seasonality, and cycles, to gain insights into the data's behavior over time.

- **Trend**: the trend represents the long-term movement or direction of data over time. It can show whether values are generally increasing, decreasing, or remaining steady.
- **Moving Average**: used to smooth out fluctuations in data, helping to identify trends more clearly. By averaging data points over a specific window, moving averages reduce noise and highlight the underlying pattern.
- **Forecasting**: forecasting predicts future values based on historical patterns. It combines time-series techniques like trend identification and moving averages to make accurate, data-driven projections.

Dataset Used

https://www.kaggle.com/datasets/rohitsahoo/sales-forecasting

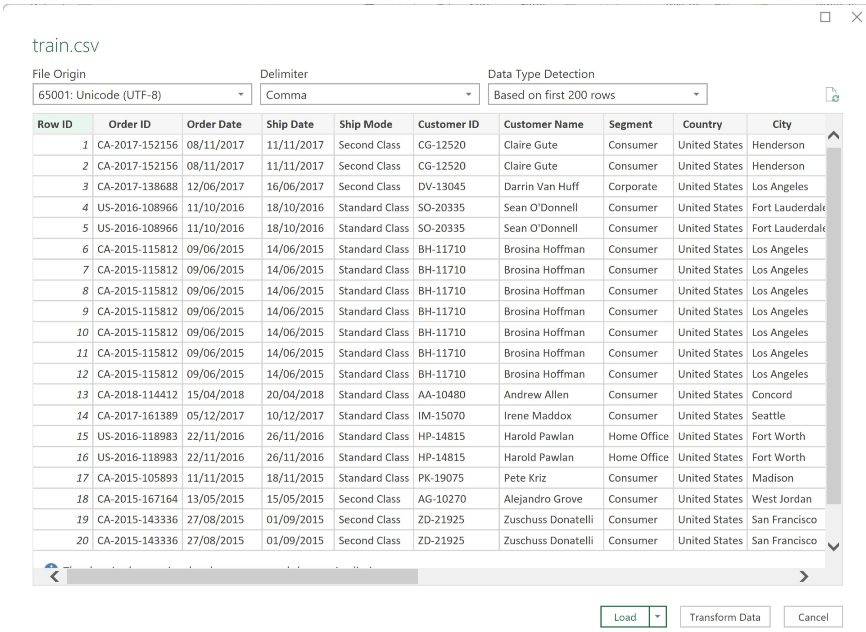
Method

Import the Data

For this project, I opted to use the Power Query Editor for more control over the data import and transformation process.

Open Excel, go to Data > Get Data, and choose From Text/CSV to load the file.

- After loading the file in Power Query, I reviewed the dataset for missing or incorrect data then clicked Close & Load to bring the transformed data into Excel for further work.


train.csv

File Origin: 65001: Unicode (UTF-8)

Delimiter: Comma

Data Type Detection: Based on first 200 rows

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City
1	CA-2017-152156	08/11/2017	11/11/2017	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson
2	CA-2017-152156	08/11/2017	11/11/2017	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson
3	CA-2017-138688	12/06/2017	16/06/2017	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles
4	US-2016-108966	11/10/2016	18/10/2016	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale
5	US-2016-108966	11/10/2016	18/10/2016	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale
6	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles
7	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles
8	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles
9	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles
10	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles
11	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles
12	CA-2015-115812	09/06/2015	14/06/2015	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles
13	CA-2018-114412	15/04/2018	20/04/2018	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord
14	CA-2017-161389	05/12/2017	10/12/2017	Standard Class	IM-15070	Irene Maddox	Consumer	United States	Seattle
15	US-2016-118983	22/11/2016	26/11/2016	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth
16	US-2016-118983	22/11/2016	26/11/2016	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth
17	CA-2015-105893	11/11/2015	18/11/2015	Standard Class	PK-19075	Pete Kriz	Consumer	United States	Madison
18	CA-2015-167164	13/05/2015	15/05/2015	Second Class	AG-10270	Alejandro Grove	Consumer	United States	West Jordan
19	CA-2015-143336	27/08/2015	01/09/2015	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco
20	CA-2015-143336	27/08/2015	01/09/2015	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco

Load | Transform Data | Cancel

Data Cleaning and Transformation

Properly cleaning and preparing the dataset was crucial for analysis. The following steps were taken:

- Verified that the Order Date and Ship Date columns were properly formatted as valid dates.
- Addressed issues such as blank cells, non-date entries, and dates stored as text.

For entries stored in the format dd/mm/yyyy, the following steps were taken:

- Inserted a new column next to the Order Date column then used the date formula to transform text dates into proper date format:

```
=DATE(VALUE(RIGHT(C2,4)), VALUE(MID(C2,4,2)), VALUE(LEFT(C2,2)))
```

In the formula above, RIGHT(C2,4) extracts the year, MID(C2,4,2) extracts the month and the (C2,2) extracts the day.

- Copied the formula down the column and replaced the original date column with the corrected values.

Add New Columns

The next step was to add Month and Year columns for more granular analysis using the Order Date column:

- Inserted a new column named Year
- Used the following formula in the first cell of the Year column to extract the year

```
=YEAR(C2)
```

- Dragged the formula down to apply it to all rows

Added another new column named Month:

- Entered the following formula in the first cell of the Month column to extract the full name of the month

```
=TEXT(C2, "mmmm")
```

- Applied the formula to all rows to create a complete month column

C	D	E
Order Date	Year	Month
1/2/2015	2015	January
1/3/2015	2015	January
1/3/2015	2015	January
1/3/2015	2015	January
1/3/2015	2015	January
1/3/2015	2015	January
1/3/2015	2015	January
1/3/2015	2015	January

The dataset now includes 19 columns and 9,801.

Row ID	Order ID	Order Date	Year	Month	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State	Postal Code	Region	Product ID	Category
541	CA-2015-140795	1/2/2015	2015	January	3/2/2015	First Class	BD-11500	Bradley Drucker	Consumer	United States	Green Bay	Wisconsin	54302	Central	TEC-AC-10001432	Technology
5714	US-2015-143707	1/3/2015	2015	January	5/3/2015	Standard Class	HR-14770	Hallie Redmond	Home Office	United States	New York City	New York	10035	East	TEC-PH-10003655	Technology
6548	CA-2015-113880	1/3/2015	2015	January	5/3/2015	Standard Class	VF-21715	Vicky Freymann	Home Office	United States	Elmhurst	Illinois	60126	Central	FUR-CH-10000863	Furniture
6549	CA-2015-113880	1/3/2015	2015	January	5/3/2015	Standard Class	VF-21715	Vicky Freymann	Home Office	United States	Elmhurst	Illinois	60126	Central	OFF-PA-10003036	Office Supplie
7948	CA-2015-131009	1/3/2015	2015	January	5/3/2015	Standard Class	SC-20380	Shahid Collister	Consumer	United States	El Paso	Texas	79907	Central	OFF-FA-10004395	Office Supplie
7949	CA-2015-131009	1/3/2015	2015	January	5/3/2015	Standard Class	SC-20380	Shahid Collister	Consumer	United States	El Paso	Texas	79907	Central	FUR-CH-10001270	Furniture
7950	CA-2015-131009	1/3/2015	2015	January	5/3/2015	Standard Class	SC-20380	Shahid Collister	Consumer	United States	El Paso	Texas	79907	Central	FUR-FU-10001095	Furniture
7951	CA-2015-131009	1/3/2015	2015	January	5/3/2015	Standard Class	SC-20380	Shahid Collister	Consumer	United States	El Paso	Texas	79907	Central	OFF-ST-10001469	Office Supplie
158	CA-2015-104269	1/3/2015	2015	January	6/3/2015	Second Class	DB-13060	Dave Brooks	Consumer	United States	Seattle	Washington	98115	West	FUR-CH-10004063	Furniture
8310	CA-2015-168312	1/3/2015	2015	January	7/3/2015	Standard Class	GW-14605	Giulietta Weimer	Consumer	United States	Houston	Texas	77036	Central	OFF-ST-10003692	Office Supplie
8311	CA-2015-168312	1/3/2015	2015	January	7/3/2015	Standard Class	GW-14605	Giulietta Weimer	Consumer	United States	Houston	Texas	77036	Central	FUR-TA-10001866	Furniture
1373	US-2015-157021	1/4/2015	2015	January	6/4/2015	Second Class	KM-16720	Kunst Miller	Consumer	United States	Vallejo	California	94591	West	OFF-LA-10002312	Office Supplie
1374	US-2015-157021	1/4/2015	2015	January	6/4/2015	Second Class	KM-16720	Kunst Miller	Consumer	United States	Vallejo	California	94591	West	OFF-BI-10000042	Office Supplie
7954	CA-2015-138359	1/4/2015	2015	January	6/4/2015	Standard Class	KH-16330	Katharine Harms	Corporate	United States	Revere	Massachusetts	2151	East	OFF-ST-10000636	Office Supplie
7955	CA-2015-138359	1/4/2015	2015	January	6/4/2015	Standard Class	KH-16330	Katharine Harms	Corporate	United States	Revere	Massachusetts	2151	East	OFF-BI-10000145	Office Supplie
245	CA-2015-131926	1/6/2015	2015	January	6/6/2015	Second Class	DW-13480	Dianna Wilson	Home Office	United States	Lakeville	Minnesota	55044	Central	FUR-CH-10004063	Furniture
246	CA-2015-131926	1/6/2015	2015	January	6/6/2015	Second Class	DW-13480	Dianna Wilson	Home Office	United States	Lakeville	Minnesota	55044	Central	OFF-ST-10002276	Office Supplie
247	CA-2015-131926	1/6/2015	2015	January	6/6/2015	Second Class	DW-13480	Dianna Wilson	Home Office	United States	Lakeville	Minnesota	55044	Central	OFF-PA-10004082	Office Supplie
248	CA-2015-131926	1/6/2015	2015	January	6/6/2015	Second Class	DW-13480	Dianna Wilson	Home Office	United States	Lakeville	Minnesota	55044	Central	OFF-AP-10002945	Office Supplie
249	CA-2015-131926	1/6/2015	2015	January	6/6/2015	Second Class	DW-13480	Dianna Wilson	Home Office	United States	Lakeville	Minnesota	55044	Central	OFF-PA-10000061	Office Supplie
307	CA-2015-111003	1/6/2015	2015	January	6/6/2015	Standard Class	CR-12625	Corey Roper	Home Office	United States	Lakewood	New Jersey	8701	East	OFF-BI-10001072	Office Supplie
308	CA-2015-111003	1/6/2015	2015	January	6/6/2015	Standard Class	CR-12625	Corey Roper	Home Office	United States	Lakewood	New Jersey	8701	East	OFF-AR-10002135	Office Supplie
2201	US-2015-165659	1/6/2015	2015	January	6/6/2015	Standard Class	LT-17110	Liz Thompson	Consumer	United States	Little Rock	Arkansas	72209	South	FUR-FU-10001935	Furniture
2202	US-2015-165659	1/6/2015	2015	January	6/6/2015	Standard Class	LT-17110	Liz Thompson	Consumer	United States	Little Rock	Arkansas	72209	South	TEC-PH-10002563	Technology
9150	US-2015-157070	1/6/2015	2015	January	6/6/2015	Standard Class	QJ-19255	Quincy Jones	Corporate	United States	Detroit	Michigan	48234	Central	OFF-BI-10001765	Office Supplie
9151	US-2015-157070	1/6/2015	2015	January	6/6/2015	Standard Class	QJ-19255	Quincy Jones	Corporate	United States	Detroit	Michigan	48234	Central	OFF-AP-10004859	Office Supplie
1319	CA-2015-160773	1/7/2015	2015	January	5/7/2015	Standard Class	LW-16825	Laurel Workman	Corporate	United States	Deltona	Florida	32725	South	TEC-PH-10004586	Technology
1320	CA-2015-160773	1/7/2015	2015	January	5/7/2015	Standard Class	LW-16825	Laurel Workman	Corporate	United States	Deltona	Florida	32725	South	OFF-BI-10000546	Office Supplie
716	CA-2015-153150	1/7/2015	2015	January	6/7/2015	Second Class	DI-13600	Dorris liebe	Corporate	United States	Seattle	Washington	98105	West	OFF-BI-10003355	Office Supplie
2173	CA-2015-152296	1/8/2015	2015	January	3/8/2015	First Class	IL-15100	Ivan Liston	Consumer	United States	San Francisco	California	94122	West	OFF-BI-10004506	Office Supplie
1838	CA-2015-117345	1/8/2015	2015	January	5/8/2015	Standard Class	BF-10975	Barbara Fisher	Corporate	United States	Charlotte	North Carolina	28205	South	OFF-LA-10000240	Office Supplie

Data Aggregation and Trend Identification

To analyze the sales trends effectively, I aggregated the data by year and month using a PivotTable.

Inserted a Pivot Table:

- Highlighted the data range (including headers) then on the Insert tab selected PivotTable and placed the PivotTable in a new worksheet.
- Dragged Year into the Rows area to group data by year
- Dragged Month into the Rows area under Year to further group data by month within each year
- Dragged Sales into the Values area

	A	B
1	**Month**	**Sales by Month**
2	**2015**	
3	Jan	28828.25
4	Feb	12588.48
5	Mar	54027.69
6	Apr	24710.02
7	May	29520.49
8	Jun	29181.33
9	Jul	35194.56
10	Aug	37349.27
11	Sep	65956.40
12	Oct	34561.95
13	Nov	64369.46
14	Dec	63568.31
15	**2016**	**459436.01**
16	**2017**	**600192.55**
17	**2018**	**722052.02**
18	**Grand Total**	**2261536.78**

Preparing Data for Forecasting

Copied the data from the PivotTable into a new worksheet to simplify analysis.

Added a Helper Column to assign each month a sequential numeric identifier for use in forecasting calculations:

- In a new column, I started the sequence with **1** for the first month in the dataset (January 2015).
- Dragged the sequence down incrementally to number all months up to the final historical month (December 2018 = 48).
- Extended the sequence beyond the historical data to account for future months up to June 2019 = 54, for a six-month forecast).

	A	B	C	D
1	Year	Month	Month Sequence	Sales by Month
2	2015	Jan	1	28828.25
3	2015	Feb	2	12588.48
4	2015	Mar	3	54027.69
5	2015	Apr	4	24710.02
6	2015	May	5	29520.49
7	2015	Jun	6	29181.33
8	2015	Jul	7	35194.56
9	2015	Aug	8	37349.27
10	2015	Sep	9	65956.40
11	2015	Oct	10	34561.95
12	2015	Nov	11	64369.46
13	2015	Dec	12	63568.31
14	2016	Jan	13	29347.39
15	2016	Feb	14	20728.35
16	2016	Mar	15	34489.68
17	2016	Apr	16	38056.97
18	2016	May	17	30761.56
19	2016	Jun	18	28515.91
20	2016	Jul	19	28573.31

Using the TREND Formula for Forecasting

Added a new column titled Forecast to store the forecasted values.

Entered the following formula in the first cell of the Forecast column for the future months:

```
=TREND(Sales_Range, Date_Range, New_Date_Range)

--for this analysis, the formula is:
=TREND(D2:D49, C2:C49, C50:C55)
```

Replaced the placeholders with the appropriate ranges from the data:

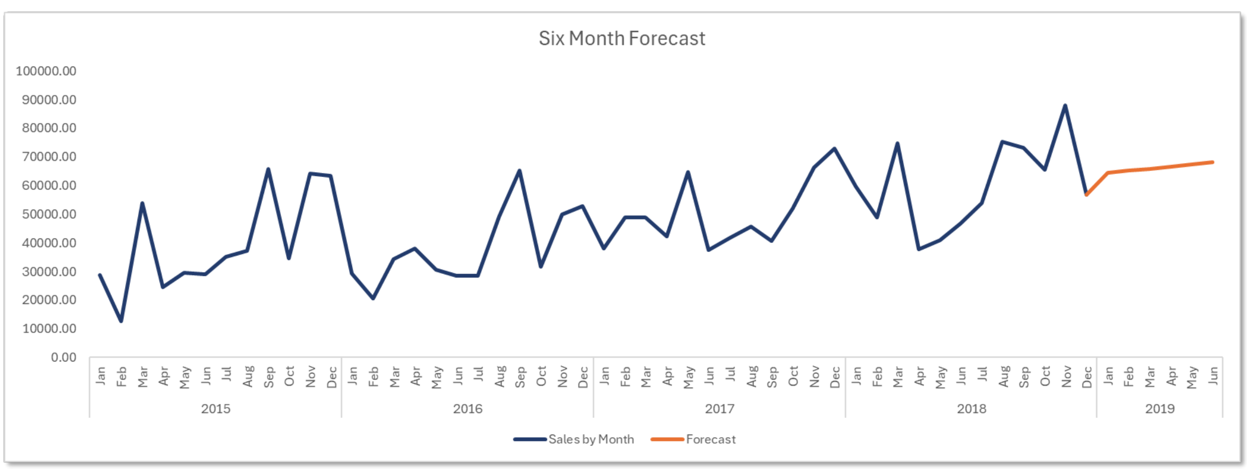
- Sales_Range: The cells containing the Sales by Month column data
- Date_Range: The cells containing the Helper Column data
- New_Date_Range: The numeric sequence for the future months representing January 2019 to June 2019

SUM | =TREND(D2:D49, C2:C49, C50:C55)

	A	B	C	D	E
31	2017	Jun	30	37424.68	
32	2017	Jul	31	41761.94	
33	2017	Aug	32	45766.81	
34	2017	Sep	33	40692.31	
35	2017	Oct	34	52156.96	
36	2017	Nov	35	66392.55	
37	2017	Dec	36	72847.09	
38	2018	Jan	37	59767.09	
39	2018	Feb	38	48928.83	
40	2018	Mar	39	74748.62	
41	2018	Apr	40	37849.22	
42	2018	May	41	40882.45	
43	2018	Jun	42	46912.85	
44	2018	Jul	43	53942.78	
45	2018	Aug	44	75408.78	
46	2018	Sep	45	73153.36	
47	2018	Oct	46	65501.16	
48	2018	Nov	47	87997.64	
49	2018	Dec	48	56959.24	
50	2019	Jan	49		=TREND(D2:D49, C2:C49, C50:C55)
51	2019	Feb	50		TREND(known_ys, [known_xs], [new_xs], [const])
52	2019	Mar	51		65949
53	2019	Apr	52		66660
54	2019	May	53		67370
55	2019	Jun	54		68081

Forecast Output and Visualization:

The TREND formula generates forecasted values based on historical data. It assumes a linear relationship between the independent and dependent variables. In my analysis, the TREND formula was used to project sales for the next six months, providing an initial look at future performance.



However, as demonstrated in the chart, the TREND formula generates a straight-line forecast. While this is useful for identifying general growth or decline, it did not accurately capture the seasonal or cyclical variations present in the data. For instance, the historical analysis shows recurring peaks in sales during September and November, suggesting a pattern that is not purely linear.

To account for these nuances and provide a more realistic projection, i used the FORECAST.ETS function as well. Unlike the TREND function, FORECAST.ETS is designed to handle time-series data with seasonality, making it better suited for projecting future sales in datasets like this.

Steps to Apply FORECAST.ETS

Created a new column labeled FORECAST.ETS in the dataset for storing the forecasted values.

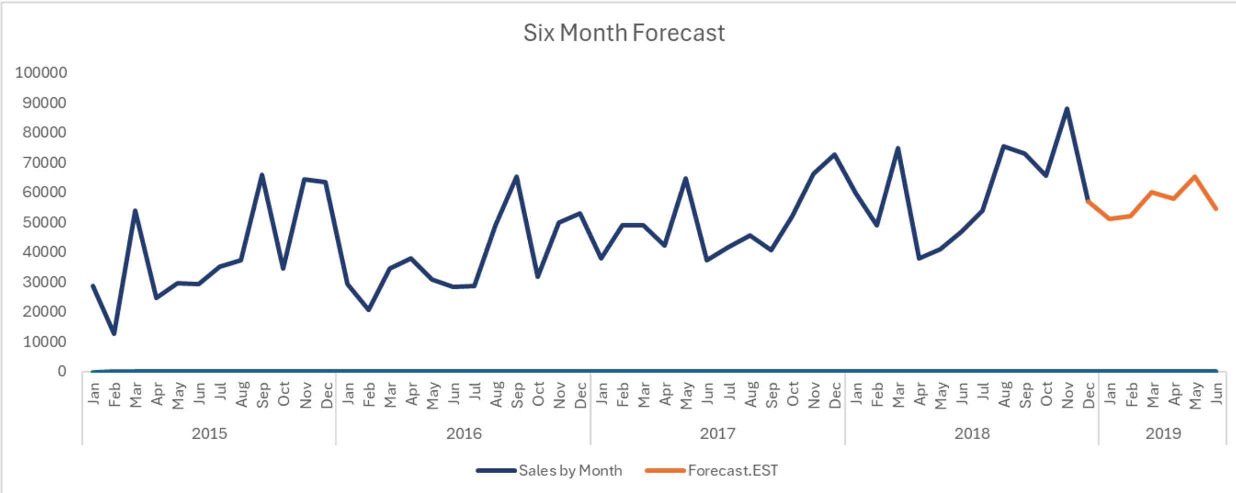
In the first cell of the FORECAST.ETS column I entered the formula:

```
=FORECAST.ETS(target_date, values, timeline)

--for this analysis, the formula is:
=FORECAST.ETS(C50:C55, D2:D49, C2:C49)
```

TREND vs. FORECAST.ETS Analysis

In contrast to the TREND chart above, the FORECAST.ETS formula demonstrates its ability to capture seasonal fluctuations more effectively. By factoring in both seasonality and historical patterns, FORECAST.ETS provided a realistic outlook for the next six months. This method aligns more closely with observed sales dynamics, offering actionable insights for planning inventory and marketing strategies.



Forecasting Using Trendlines

Forecasting using trendlines helps identify and project patterns in data to predict future values. By analyzing historical trends, trendlines provide a visual and mathematical representation of data behavior over time.

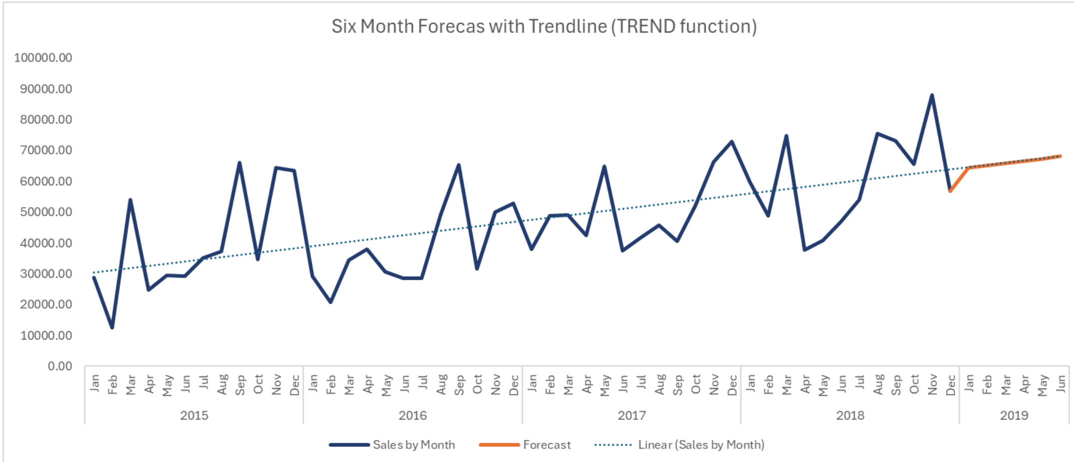
I added a Trendline by clicking on the chart and then clicking on the Chart Elements Icon and checking the box next to Trendline.

Then I extend the Trendline for Forecasting:

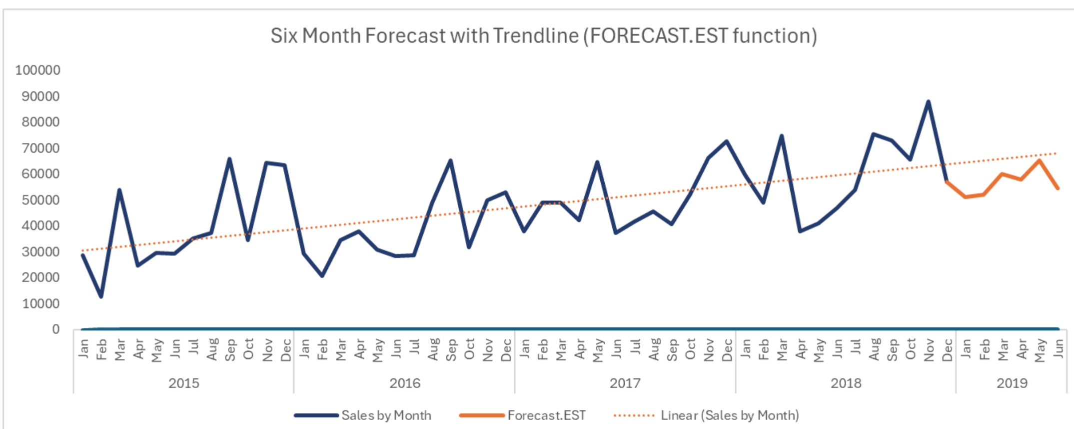
- Right-clicked on the trendline in the chart and select Format Trendline.
- Scrolled down to the Forecast section and input the desired number of months (6).

Analysis of Trendlines:

- A trendline was added to both the TREND and FORECAST.ETS charts for comparison.
    - On the TREND chart, the trendline aligns perfectly with the projected values, confirming the linear nature of the TREND function's predictions but lacks the ability to model seasonal or cyclical variations.



- In the FORECAST.ETS chart, the forecasted values generally trend upward, reflecting seasonality. However, they fall slightly below the extended trendline, suggesting that FORECAST.ETS accounts for seasonal adjustments, moderating the linear growth projection to better match real-world fluctuations.



Highlighting High and Low Values Using Conditional Formatting

To identify months with consistently high or low sales, I used the Mean and Standard Deviation Approach. This method allowed me to determine thresholds for "high" and "low" sales without predefined limits.

- Copied the sales data to a new sheet.
- Created a Pivot Table:
  - Placed Year in the column.
  - Placed Month in the rows.

- Placed Sales in the values.

Calculated the Mean:

```
=AVERAGE(Sales_Range)
=AVERAGE(F3:I14)
```

Result: $47,115.35

Calculated the Standard Deviation to measure how spread out the sales are:

```
=STDEV.S(Sales_Range)
=STDEV.S(D2:D49)
```

Result: $16,750.96

Set High and Low Thresholds:

- High Sales Threshold: Values greater than one standard deviation above the mean:

  ```
  High Sales Threshold = 47,115.35 + 16,750.96 = $63,866.3
  1
  ```

  Sales above $63,866.31 are considered "high."

- Low Sales Threshold: Values less than one standard deviation below the mean:

  ```
  Low Sales Threshold = 47,115.35 - 16,750.96 = $30,364.39
  ```

  Sales below $30,364.39 are considered "low."

Applying the Thresholds with a Conditional Formatting Rule

Highlighting high Sales:

- Selected the Range > Home tab → Conditional Formatting → New Rule
- Selected “Use a formula to determine which cells to format.”

- Entered the formula:

```
=F3 > 63866.31
```

- Clicked Format → Fill tab → Selected a green color → Clicked OK.

Highlighting Low Sales:

- Repeated steps above.
- Entered the formula:

```
=F3 < 30364.39
```

- Click Format → Fill tab → Select a red color → Click OK.

This approach visually emphasizes months with significantly high or low sales, revealing seasonal patterns or anomalies.

E	F	G	H	I	J
Sum of Sales by Month	**Column Labels**				
Row Labels	**2015**	**2016**	**2017**	**2018**	**Grand Total**
Jan	$ 28,828.25	$ 29,347.39	$ 38,048.18	$ 59,767.09	$ 155,990.92
Feb	$ 12,588.48	$ 20,728.35	$ 48,907.59	$ 48,928.83	$ 131,153.26
Mar	$ 54,027.69	$ 34,489.68	$ 48,990.14	$ 74,748.62	$ 212,256.13
Apr	$ 24,710.02	$ 38,056.97	$ 42,368.05	$ 37,849.22	$ 142,984.25
May	$ 29,520.49	$ 30,761.56	$ 64,836.25	$ 40,882.45	$ 166,000.75
Jun	$ 29,181.33	$ 28,515.91	$ 37,424.68	$ 46,912.85	$ 142,034.77
Jul	$ 35,194.56	$ 28,573.31	$ 41,761.94	$ 53,942.78	$ 159,472.59
Aug	$ 37,349.27	$ 49,076.93	$ 45,766.81	$ 75,408.78	$ 207,601.79
Sep	$ 65,956.40	$ 65,353.00	$ 40,692.31	$ 73,153.36	$ 245,155.07
Oct	$ 34,561.95	$ 31,631.89	$ 52,156.96	$ 65,501.16	$ 183,851.96
Nov	$ 64,369.46	$ 50,009.15	$ 66,392.55	$ 87,997.64	$ 268,768.79
Dec	$ 63,568.31	$ 52,891.88	$ 72,847.09	$ 56,959.24	$ 246,266.52
Grand Total	**$ 479,856.21**	**$459,436.01**	**$600,192.55**	**$722,052.02**	**$2,261,536.78**
	Average Sales	$ 47,115.35		**High Sales**	$ 63,866.31
	Standard Dev.	$ 16,750.96		**Low Sales**	$ 30,364.39

Calculating Sales Increase Over a 6-Month Forecast Period

To analyze the expected percentage increase in forecasted sales over the next six months (January to June 2019), I compared the forecasted values at the beginning and end of this period. By determining the percentage change between the starting value and the ending value, I can quantify the anticipated growth and evaluate the trend's significance.

Steps to Calculate the Percentage Increase:

Using the Starting Value (January): $64,527.54 and Ending Value (June): $68,081.04 I calculated the Total Increase by subtracting the starting value from the ending value:

$68,081.04 - $64,527.54 = $3,553.51

Then, divided the total increase by the starting value and multiply by 100:

($3,553.51 / $64,527.54) × 100 = 5.51%

This analysis shows a projected 5.51% increase in sales over the six-month period, indicating a steady upward trend.

Exponential Moving Average (EMA)

I chose the Exponential Moving Average (EMA) because it adapts to cyclical trends, smooths fluctuations, and retains sensitivity to older data. It is particularly suited for capturing seasonal patterns and peaks while accounting for recent changes effectively.

Choosing the Smoothing Factor (α)

I used the alpha formula method to calculate the smoothing factor, as it ties α to a specific period, ensuring consistency and replicability. This approach can be more reliable and transparent compared to subjective manual selection.

For 48 months of data (2015–2018) and a 12-month EMA span, the formula is:

```
α  =  2 / (n + 1)
```

Where n=12. Using this formula, the smoothing factor is: 0.1538

Process for Calculating EMA:

- Copied the sales data into a new sheet.

- Set months for the EMA span in cell D1 to 12.
- Calculated the Smoothing Factor (α):
    - In cell D2:

```
=2/(D1+1)
```

Result: α=0.1538

Calculated the Initial EMA (SMA of the First n Periods):

- Added a new column called EMA.
- Calculated the Simple Moving Average (SMA) for the first 12 months (January to December 2015):
    - In cell C2:

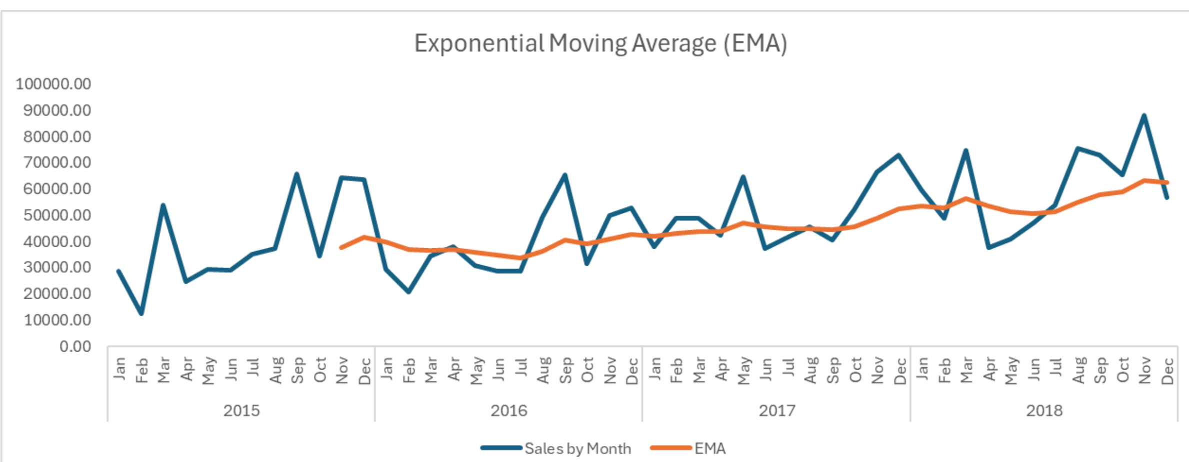
```
=AVERAGE(B1:B12)
```

Calculate EMA for Subsequent Periods using the EMA formula:

- Starting from row 13 (corresponding to the 13th month):

```
=(D$2*B13) + ((1-D$2)*C12)
```

- Dragged this formula down for all rows in Column C to compute the EMA for the remaining months.



The EMA effectively smoothed out short-term fluctuations while preserving the overall trend and captured cyclical behavior, highlighting trends and seasonal sales patterns.

Sales Trend Analysis Summary

With the analysis completed, I was able to answer the initial questions:

1. What seasonal trends are evident in the sales data, and how do these patterns inform marketing and inventory strategies?
    - September and November consistently show strong sales, indicating key months for focused marketing efforts and increased inventory planning.
2. How have the low-sales periods evolved over time, and what does this suggest about the business's growth or operational changes?
    - Low-sales months in 2015 and 2016 are no longer below the threshold in subsequent years, pointing to consistent growth and fewer dips in demand.
3. How does the six-month sales forecast support decision-making for resource allocation in the upcoming period?
    - The 5.51% projected increase suggests strong upcoming performance, aiding in budgeting and workforce planning.
4. What was the highest-performing month in the dataset, and what factors might have contributed to its exceptional sales performance?
    - November 2018 had the best sales at $87,997.64, potentially linked to holiday shopping or promotions.
5. How can the identification of sales peaks and troughs be leveraged to drive targeted promotions and enhance customer engagement?
    - Seasonal sales insights can guide promotional timing, ensuring campaigns align with high-demand periods like September and November.

This analysis highlights a positive trend with seasonal sales peaks in September and November and a projected upward trajectory in the coming months. This insight provides a foundation for optimizing inventory, staffing, and marketing strategies during high-demand periods and positioning for continued growth.

Recommendations for the Business

- Capitalize on high-sales periods and focus marketing efforts during high-sales months by allocating more budget to targeted marketing campaigns to maximize revenue.
- Optimize Inventory Management to ensure adequate stock levels for high-demand products during high sales months to avoid low stock.
- It might be beneficial to add offers or discounts to further drive sales during peak periods.
- Address Low-Sales Periods by implement strategies to stimulate demand, such as discounts, flash sales, or loyalty rewards during low-performing months.
- Investigate factors causing low performance (e.g., market trends, competition, or seasonality) to mitigate future slowdowns.

Visualization

When calculating the Exponential Moving Average I used the alpha formula method to calculate the smoothing factor. Based on my research, for monthly data, an α between 0.1 and 0.3 is often suitable. To make this process more efficient, I’ve created an interactive visualization using Tableau. This tool allows business users to test different alpha values and N-periods, and compare them with the alpha formula-based approach. The interactive nature of this visualization reduces the need for manual calculations and enhances accessibility for non-technical stakeholders.

Validation:

- By calculating EMA with various values, business users can visually confirm how well the selected α aligns with the actual data.

Improved Decision-Making:

- EMA testing provides qualitative validation (trend visualization). This empowers business users to make informed decisions when selecting an α, ensuring the model aligns with both accuracy metrics and business needs.

By integrating these visualizations, I've created a solution that enhances the decision-making process, simplifies forecasting analysis, and ensures greater confidence in the selected smoothing factors.

Link: Tableau Visualization