

Fundamentals of time series data

TIME SERIES ANALYSIS IN POWER BI



Kevin Barlow

Data Analytics Professional

Course pre-requisites

This course requires intermediate DAX functions throughout the exercises, including:

- `CALCULATE()`
- `IF()`
- `SUMMARIZE()`

Why is it important?

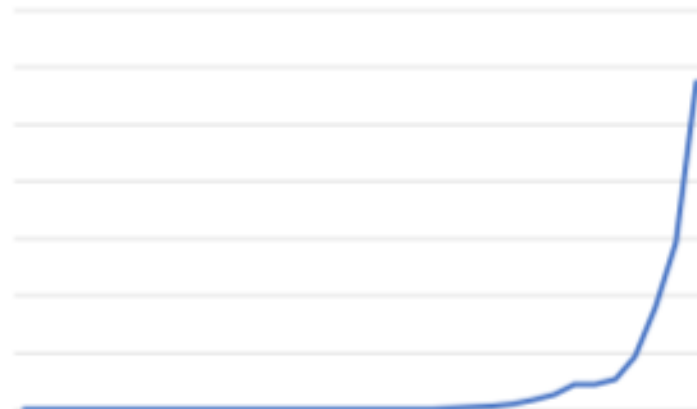
Every dataset has time as a dimension! We can learn some very interesting trends from analyzing data over time, and we need a strong toolkit to do so.

- *How has my data changed over time?*
- *What should my data show based on history?*
- *What will my data show in the future?*

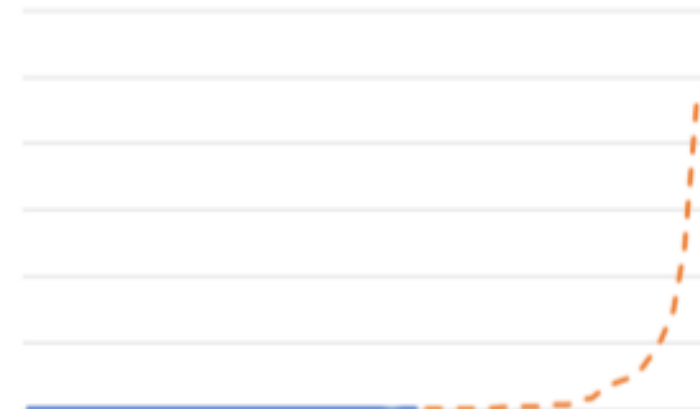
What does history say?



How has my data changed?



How will my data change?

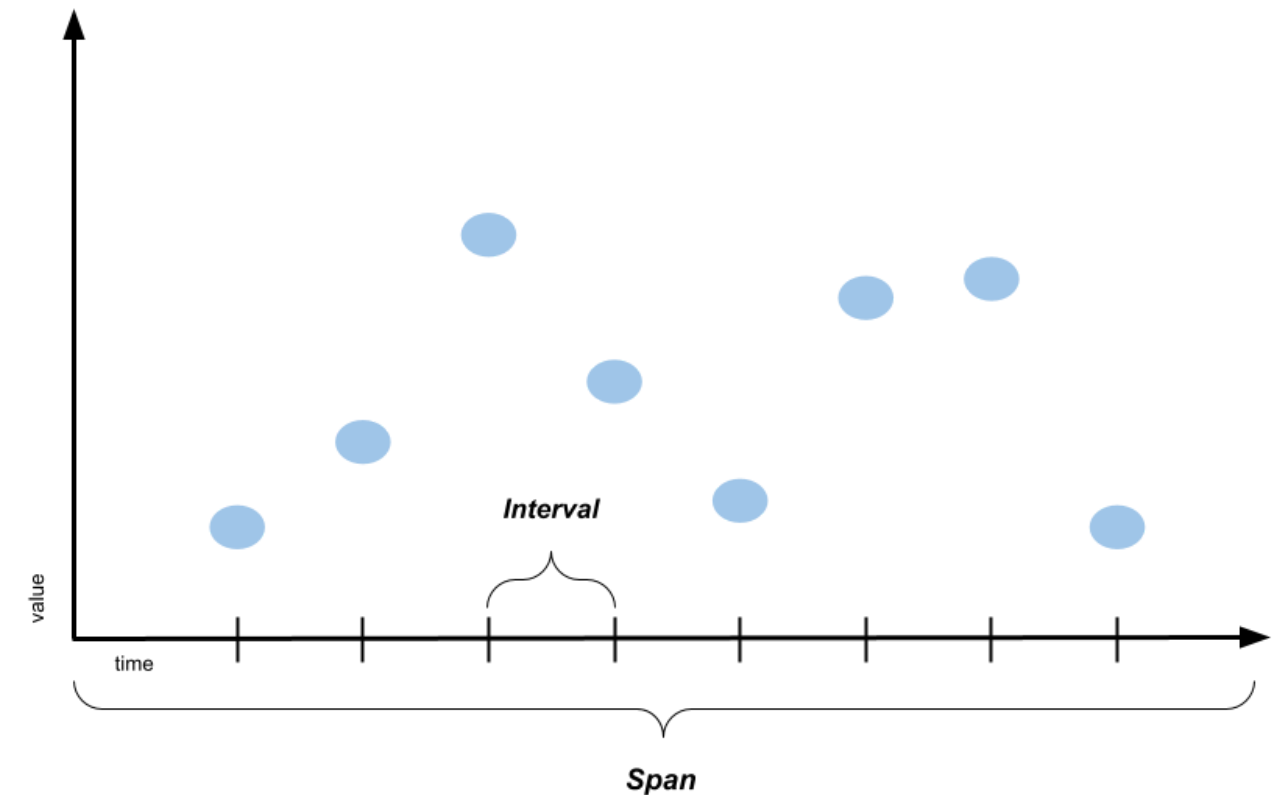


Defining time series data

Time Series Data is a set of observations (i.e. data points) that have been collected about the same subject over a period of time.

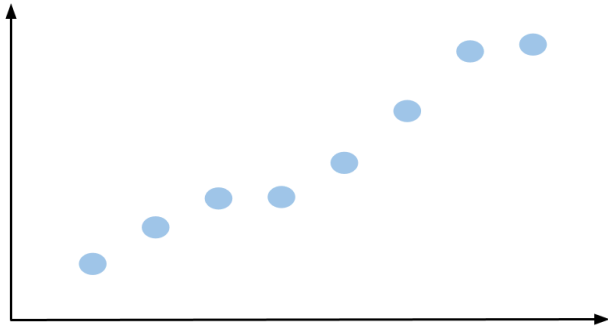
Two key components

- *Span*: the total time period we have data
- *Interval*: the amount of time in between each observation.

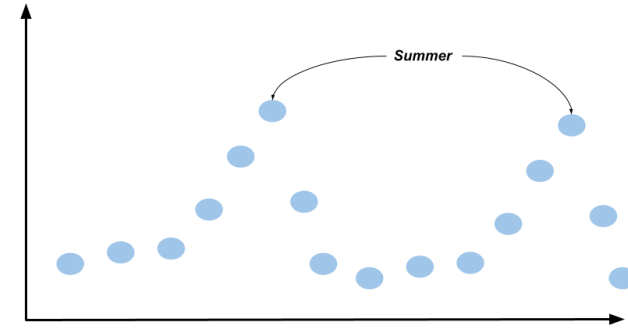


Tracking changes over time

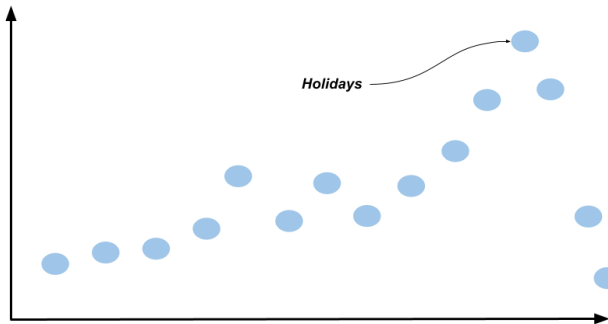
Secular Variation



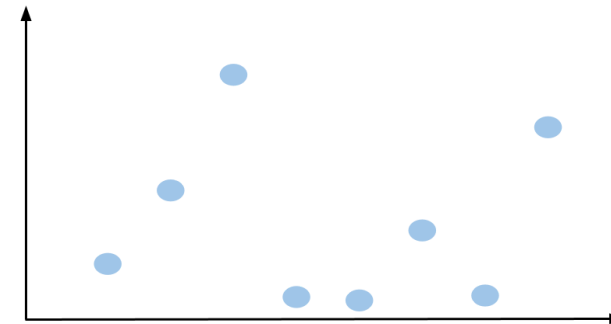
Seasonal Variation



Cyclical Variation



Random (Irregular) Variation



Short term time series analysis

A short-term time span encompasses a timeframe typically less than one year. With these, we want to understand what our data is showing right now.

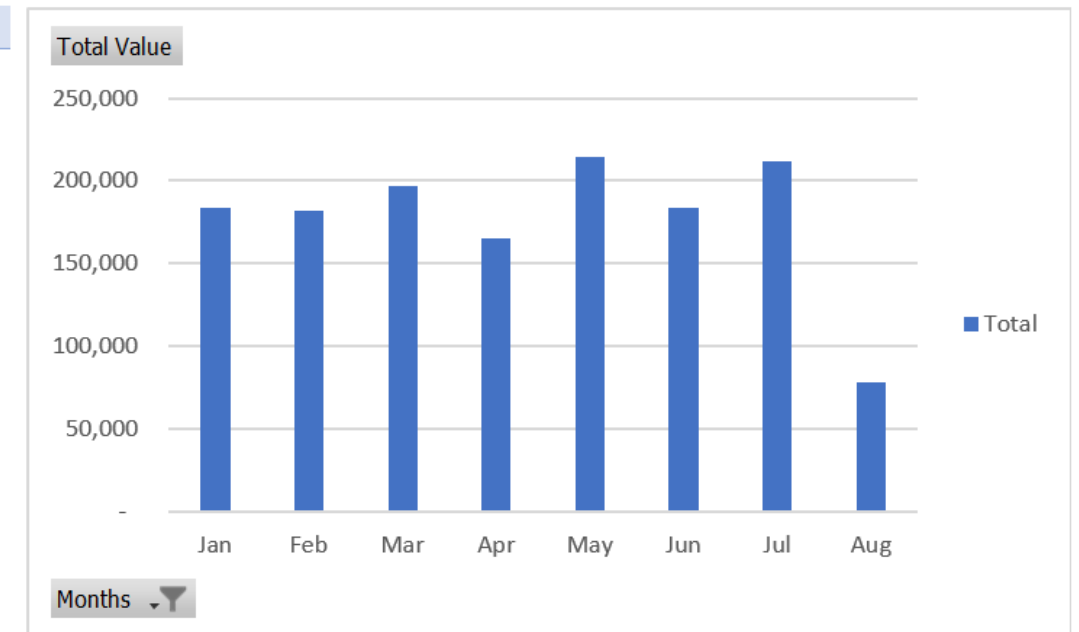
Common Analyses

- *Last x periods* (e.g., Last 90 days, Last 2 quarters)
- *Period to date* (e.g., Year to date, Month to Date)

Example

Month	Total Value
Jan	183,438
Feb	182,341
Mar	196,822
Apr	165,576
May	214,687
Jun	183,320
Jul	212,226
Aug	78,425

Year to Date 1,416,835
Last 2 Months 403,666



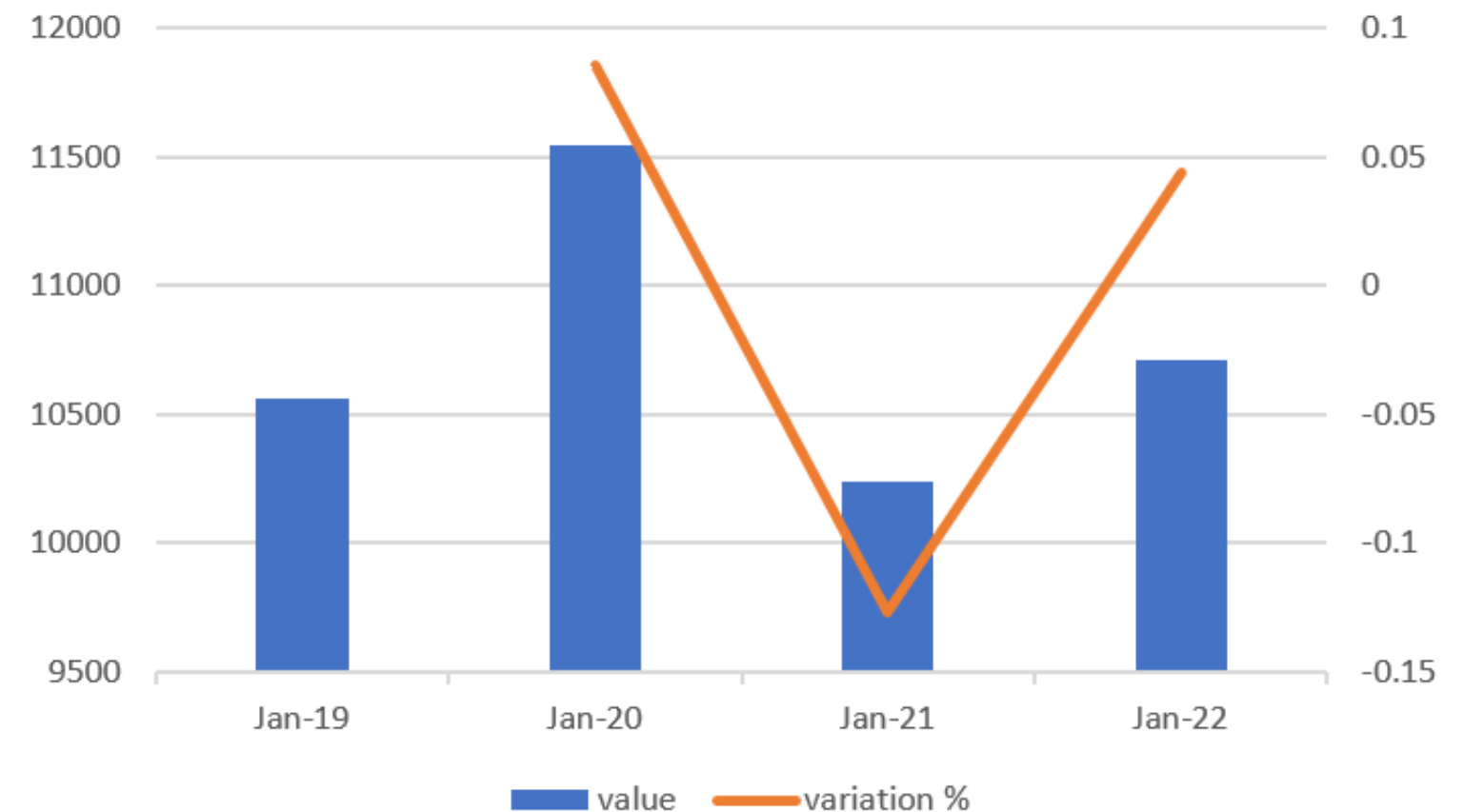
Long term time series analysis

Long term analyses encompass a timeframe longer than one year. We are trying to understand historical data and its relationship to current data.

Common Analyses:

- *Year over Year*
- *Month over Month*
- *Same period last year*

Example Calculating Year over Year (YoY) sales performance for the month of January.



Forecasting the future

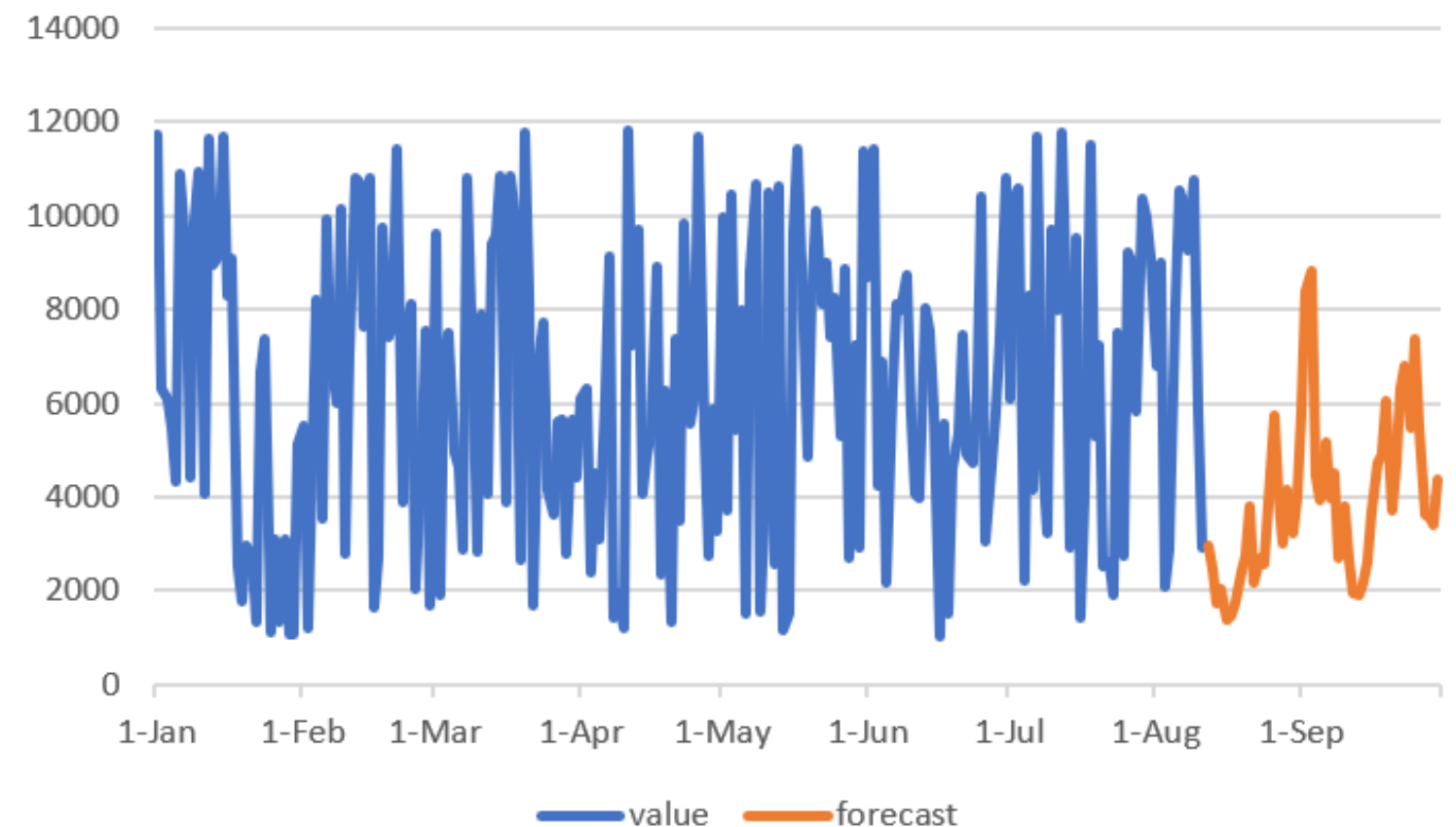
Future analyses, such as *forecasting*, look at historical data and project it into the future. These kinds of analyses allow us to predict where we will be and make decisions based on that.

Common Analyses

- Regressions (e.g. linear)
- Point in time estimations
- Machine Learning

Example

Forecasting stock trade volume several months into the future.



Let's practice!

TIME SERIES ANALYSIS IN POWER BI

Time series variation

TIME SERIES ANALYSIS IN POWER BI



Kevin Barlow

Data Analytics Professional

Let's practice!

TIME SERIES ANALYSIS IN POWER BI

Basics of time series data transformation

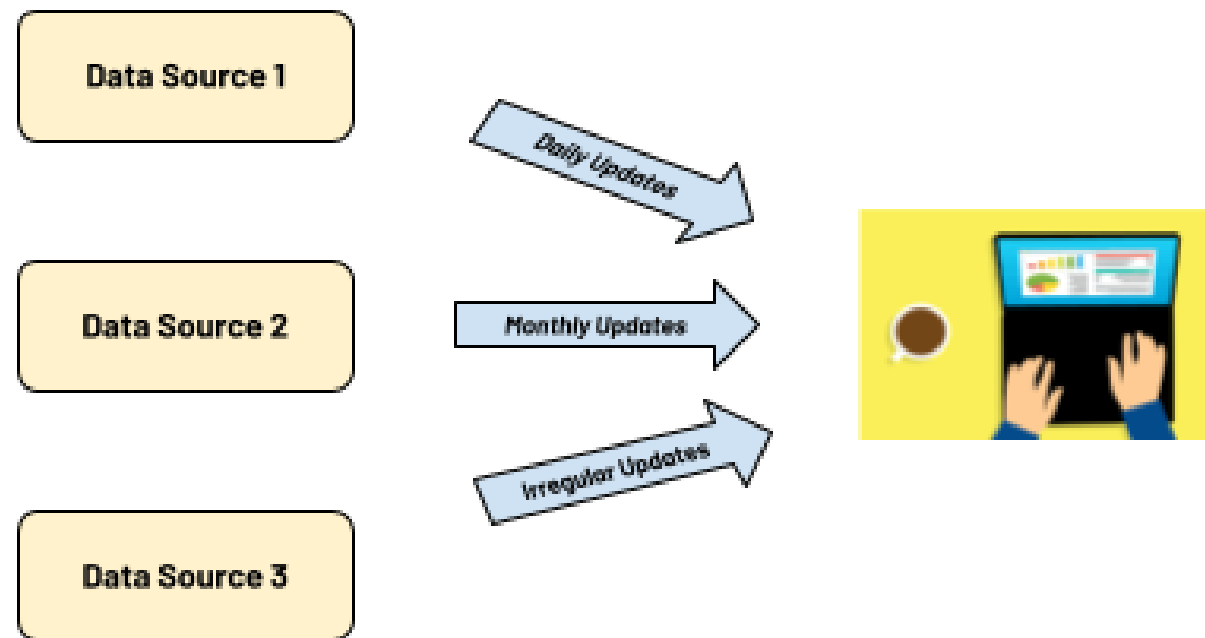
TIME SERIES ANALYSIS IN POWER BI



Kevin Barlow
Data Analytics Professional

Context and importance

- Time series data has become increasingly prevalent in every industry.
- Different systems have different requirements and formats.



ISO as the international standard

ISO 8601

- An international standard for managing date and time data
- Allows for translation of dates and times into different formats
- Provides localization and personalization for individual needs
- Manipulate your date field with `FORMAT()`

Example Date:

```
var date = January 1st 2022
# US Standard (MM DD YYYY)

var date2 = FORMAT(date, "DD MM YYYY")
# UK Standard
# 1 January 2022

var date3 = FORMAT(date, "YYYY MM DD")
# ISO Standard
# 2022 January 01
```

¹ https://en.wikipedia.org/wiki/ISO_8601#Combined_date_and_time_representations

UNIX for high-precision timestamps

UNIX Time

- Also referred to as Epoch time
- Signified the number of seconds that have passed since the Epoch time (January 1st, 1970, at 00:00:00 UTC)
- UNIX time typically applies to computer systems
- Represents a timestamp (date and time)

Example dates and times

January 1st, 2022 06:00:00 UTC

1641016800

January 1st, 2022 18:00:00 UTC

1641060000

¹ https://en.wikipedia.org/wiki/Unix_time

Manipulating dates in Power BI

DATEADD()

Allows you to move a specified number of time intervals from a reference date.

```
DATEADD(<dates>, <number_of_intervals>, <interval>)
```

Examples:

```
DATEADD('date', 30, DAY)
```

```
var timestamp_PST =  
    DATEADD('timestamp_EST', -3, HOUR)
```

DATEDIFF()

Calculates the number of time intervals between two date objects.

```
DATEDIFF(<Date1>, <Date2>, <Interval>)
```

Examples:

```
DATEDIFF(DATE(2019, 2, 1),  
          DATE(2020, 4, 30),  
          MONTH)
```

Result: 14

Summarizing data with DAX

Time Series does not always come in regular intervals, and depends on how the data is collected. This can be problematic:

- Time algorithms and equations assume a consistent interval
- Significant changes can happen in between data points

`SUMMARIZE()` your table to "smooth" out the dataset. By aggregating our dataset, we can still get valuable information on trends, even if data is missing or irregular.

Example:

```
SUMMARIZE('<table>',  
          [group_column(s)],  
          [new aggregated column(s)])
```

```
SUMMARIZE('sales',  
          sales[Quarter],  
          sales[Region],  
          "Total Sales", SUM(sales[revenue]))
```

Handling missing data

Imputation - a useful strategy when we have lots of context about what the missing data value should be; this is often done when in *Power Query*.

date	actual	imputed
1/1/2022	12.3	12.3
1/2/2022	13.4	13.4
1/3/2022	14.5	14.5
1/4/2022		15.6
1/5/2022	16.7	16.7
1/6/2022	17.8	17.8
1/7/2022	18.9	18.9

Dropping data - can be used when a row or column is missing significant amounts of data; should be avoided unless strongly benefits our analysis

date	A	B (can drop)	C
1/1/2022	12.3		red
1/2/2022	13.4		blue
1/3/2022	14.5		green
1/4/2022	15.6		blue
1/5/2022	16.7		green
1/6/2022	17.8	example	red
1/7/2022	18.9		black

Superstore dataset

In the following exercises, you will be looking at the **Superstore** dataset. This dataset represents Point of Sale (PoS) data for Superstore locations across the United States. With it, you will be formatting dates to a correct format for further analysis.

Category	City	Country.Region	Customer.Name	Discount..bin.	Manufacturer	Order.Date	Order.ID	Postal.Code	Product.Name
Furniture	New York City	United States	Parhena Norris	0	Other	09/11/2020	CA-2020-161018	10009	9-3/4 Diameter Round Wall Clock
Furniture	New York City	United States	Cynthia Voltz	0	Other	25/12/2020	CA-2020-146780	10035	Magnifier Swing Arm Lamp
Technology	New York City	United States	Mark Packer	0	Other	12/10/2018	CA-2018-102281	10035	I Need's 3d Hello Kitty Hybrid Silicone Case Cover for HTC One X 4g with 3d Hello Kitty Stylus Pen Gree
Office Supplies	New York City	United States	Mark Packer	0	Other	12/10/2018	CA-2018-102281	10035	4009 Highlighters by Sanford
Office Supplies	New York City	United States	Nathan Mautz	0	Other	04/11/2017	CA-2017-164973	10024	Trav-L-File Heavy-Duty Shuttle II, Black
Technology	New York City	United States	Adam Bellavance	0	Other	01/09/2019	CA-2019-129714	10009	Sabrent 4-Port USB 2.0 Hub
Office Supplies	New York City	United States	Jeremy Lonsdale	0	Other	08/04/2019	CA-2019-138520	10035	#10 Gummed Flap White Envelopes, 100/Box
Office Supplies	New York City	United States	Erin Mull	0	Other	22/07/2017	CA-2017-127691	10024	Premium Writing Pencils, Soft, #2 by Central Association for the Blind
Office Supplies	New York City	United States	Jennifer Ferguson	0	Other	28/12/2020	CA-2020-164826	10024	OIC Bulk Pack Metal Binder Clips
Technology	New York City	United States	Jennifer Ferguson	0	Other	28/12/2020	CA-2020-164826	10024	Cush Cases Heavy Duty Rugged Cover Case for Samsung Galaxy S5 - Purple
Office Supplies	New York City	United States	Tracy Hopkins	0	Other	05/04/2017	CA-2017-113887	10035	Eaton Premium Continuous-Feed Paper, 25% Cotton, Letter Size, White, 1000 Shts/Box
Technology	New York City	United States	Patrick O'Donnell	0	Other	30/12/2020	CA-2020-143259	10009	Gear Head AU3700S Headset
Office Supplies	New York City	United States	Bryan Spruell	0	Other	09/04/2020	CA-2020-135279	10011	Color-Coded Legal Exhibit Labels
Office Supplies	New York City	United States	Bryan Spruell	0	Other	09/04/2020	CA-2020-135279	10011	Computer Printout Paper with Letter-Trim Fine Perforations
Furniture	New York City	United States	Naresj Patel	0	Other	09/08/2018	CA-2018-131338	10024	Artistic Insta-Plaque
Furniture	New York City	United States	Michael Nguyen	0	Other	26/11/2020	CA-2020-100097	10009	DataProducts Ampli Magnifier Task Lamp, Black,
Furniture	New York City	United States	Christina DeMoss	0	Other	05/11/2020	CA-2020-102925	10024	Career Cubicle Clock, 8 1/4", Black
Furniture	New York City	United States	Adam Shillingsburg	0	Other	22/09/2017	CA-2017-156160	10035	Computer Room Manger, 14"
Furniture	New York City	United States	Giulietta Baptist	0	Other	08/03/2019	CA-2019-137337	10011	Coloredge Poster Frame
Office Supplies	New York City	United States	Christy Brittain	0	Other	16/11/2020	US-2020-157896	10009	Tops Green Bar Computer Printout Paper
Office Supplies	New York City	United States	Cindy Chapman	0	Other	02/05/2019	CA-2019-146836	10024	Super Decoflex Portable Personal File
Office Supplies	New York City	United States	Rick Hansen	0	Other	23/08/2017	CA-2017-110639	10009	Personal Creations Ink Jet Cards and Labels
Office Supplies	New York City	United States	Lisa DeCherney	0	Other	29/09/2017	CA-2017-124856	10011	White Computer Printout Paper by Universal
Office Supplies	New York City	United States	Christina DeMoss	0	Other	05/11/2020	CA-2020-102925	10024	Eaton Premium Continuous-Feed Paper, 25% Cotton, Letter Size, White, 1000 Shts/Box
Office Supplies	New York City	United States	John Huston	0	Other	28/03/2018	CA-2018-133536	10009	Security-Tint Envelopes
Office Supplies	New York City	United States	Annie Zypern	0	Other	12/05/2018	CA-2018-130876	10024	Portable Personal File Box
Office Supplies	New York City	United States	James Galang	0	Other	05/12/2019	CA-2019-108434	10024	Dual Level, Single-Width Filing Carts
Office Supplies	New York City	United States	Bart Folk	0	Other	05/12/2020	CA-2020-117128	10024	Telephone Message Books with Fax/Mobile Section, 4 1/4" x 6"

Let's practice!

TIME SERIES ANALYSIS IN POWER BI

Transforming dates

TIME SERIES ANALYSIS IN POWER BI



Kevin Barlow
Data Professional

Let's practice!

TIME SERIES ANALYSIS IN POWER BI