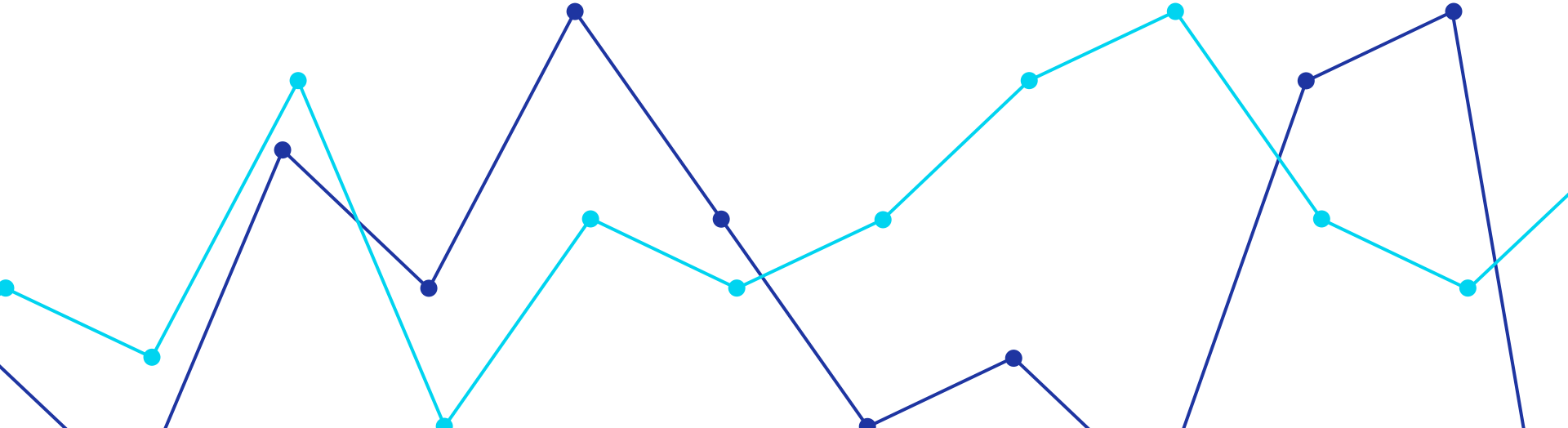
Data Visualization Using Power BI

By: Toaa Gamal Sobeihy



Why do we need data visualization ?

- Summarization gives us the ability to ***Interpret*** large amounts of Data
- Data visualization gives us a clear idea of what the information means by giving it **visual context** through maps or graphs.
- makes it easier to identify **trends**, **patterns, and outliers**

Why do we need data visualization ?

- Summarization gives us the ability to Interpret large amounts of Data
- Data visualization gives us a clear idea of what the information means by giving it ***visual context*** through Graphs.
- makes it easier to identify **trends**, **patterns, and outliers**

Visual Context

Market	Sales
Africa	783,773.21
APAC	3,585,744.13
Canada	66,928.17
EMEA	806,161.31
EU	2,938,089.06
LATAM	2,164,605.17
US	2,297,200.86
Total	**12,642,501.91**

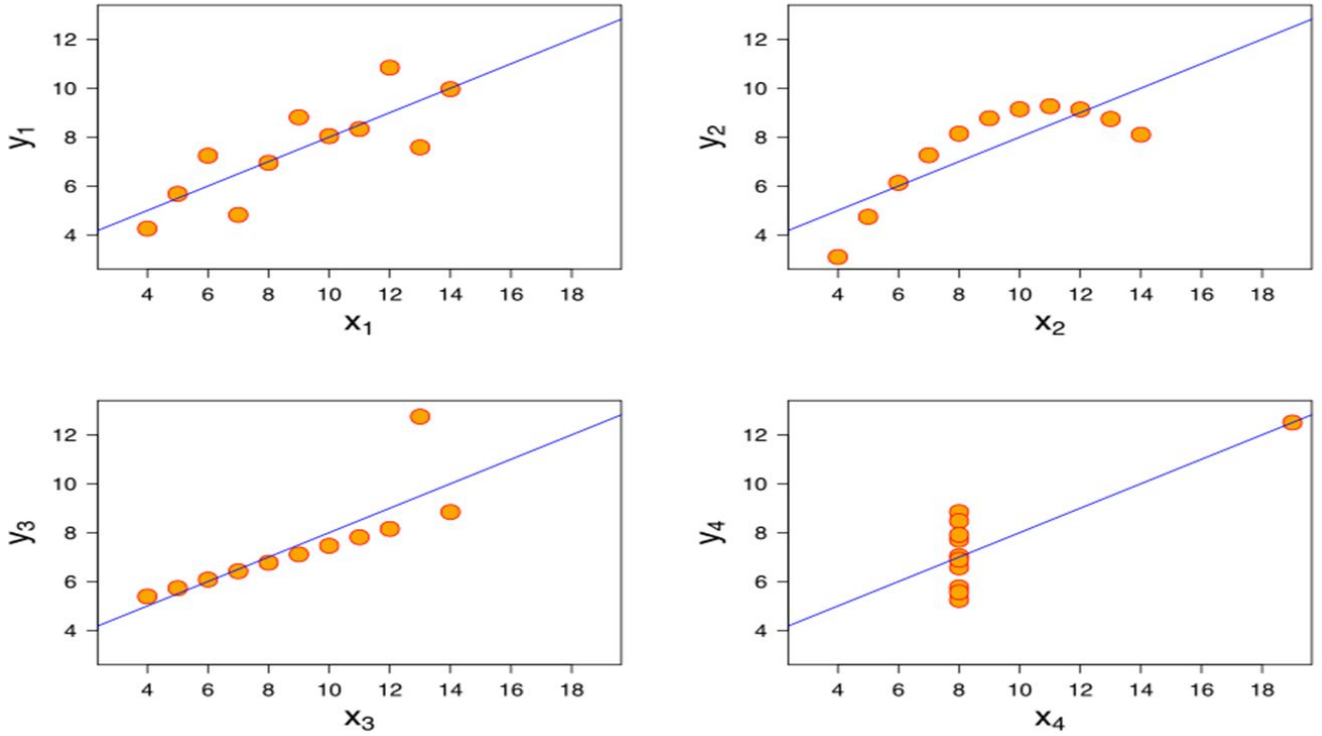
Anscombe's quartet

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Anscombe's quartet

Property	Value
Mean of x	9
Sample variance of x : s_x^2	11
Mean of y	7.50
Sample variance of y : s_y^2	4.125
Correlation between x and y	0.816
Linear regression line	y = 3.00 + 0.500x

Anscombe's quartet



Anscombe's quartet

- Four data sets that have nearly **identical simple descriptive statistics**, yet have very **different distributions** and appear very different when graphed.
- Constructed in 1973 by the statistician **Francis Anscombe**
- To demonstrate both the importance of graphing data before analyzing it and the effect of **outliers** and other influential observations on statistical properties.

Why do we need data visualization ?

- Summarization gives us the ability to Interpret large amounts of Data
- Data visualization gives us a clear idea of what the information means by giving it **visual context** through maps or graphs.
- Makes it easier to identify ***trends, patterns, and outliers***

What is Power BI?

Power BI is a collection of software services, apps, and connectors that work together to turn your unrelated sources of data into coherent, visually immersive, and interactive insights. Your data may be an Excel spreadsheet, or a collection of cloud-based and on-premises hybrid data warehouses. Power BI lets you easily connect to your data sources, visualize and discover what's important, and share that with anyone or everyone you want.

The Main Parts of Power BI

Power BI Desktop

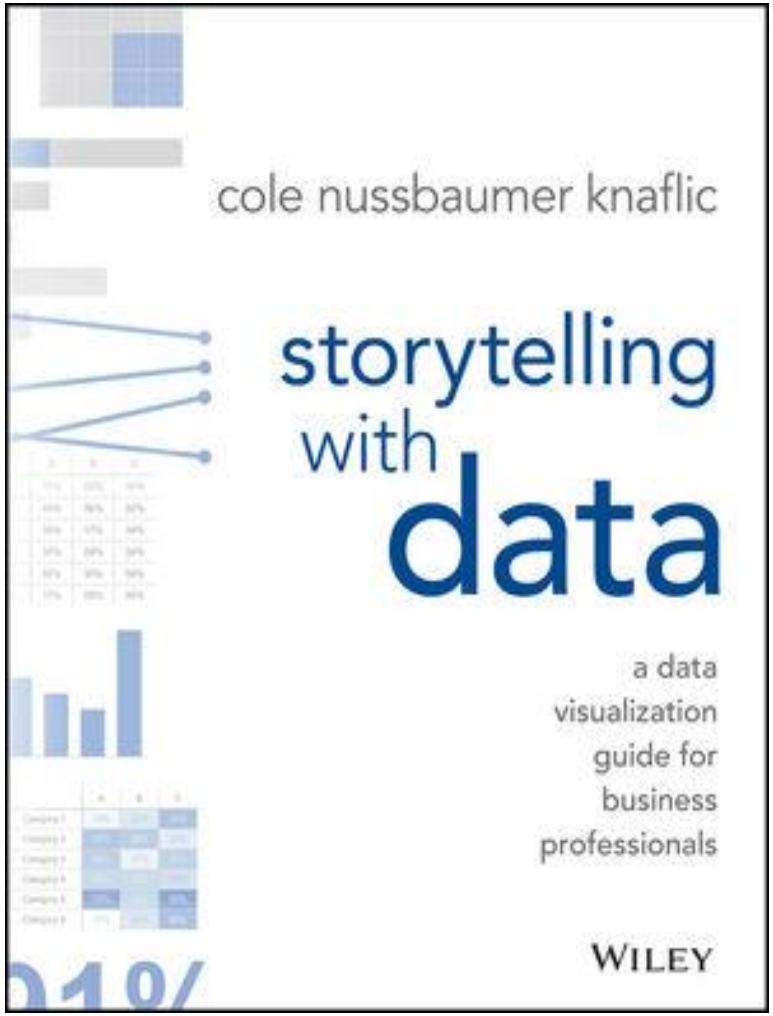
Power BI Desktop is an application that you download and install for free on your local computer. Desktop is a complete data analysis and report creation tool that is used to connect to, transform, visualize, and analyze your data.

Power BI Service

The Power BI service is a cloud-based service, or software as a service (SaaS). It supports report editing and collaboration for teams and organizations.

Power BI Desktop

- Data Preparation (Power Query)
- Creating a Data Model
- Calculations if needed (DAX)
- Visualization


cole nussbaumer knaflic
storytelling with data
a data visualization guide for business professionals
WILEY

Storytelling with data

1) The importance of context
2) Choosing an effective visual
3) Clutter is your enemy!
4) Focus your Audience’s attention
5) Think like a designer

Storytelling with data

1) **The importance of context**
2) Choosing an effective visual
3) Clutter is your enemy!
4) Focus your Audience’s attention
5) Think like a designer

The Importance of context

Exploratory vs. Explanatory Analysis

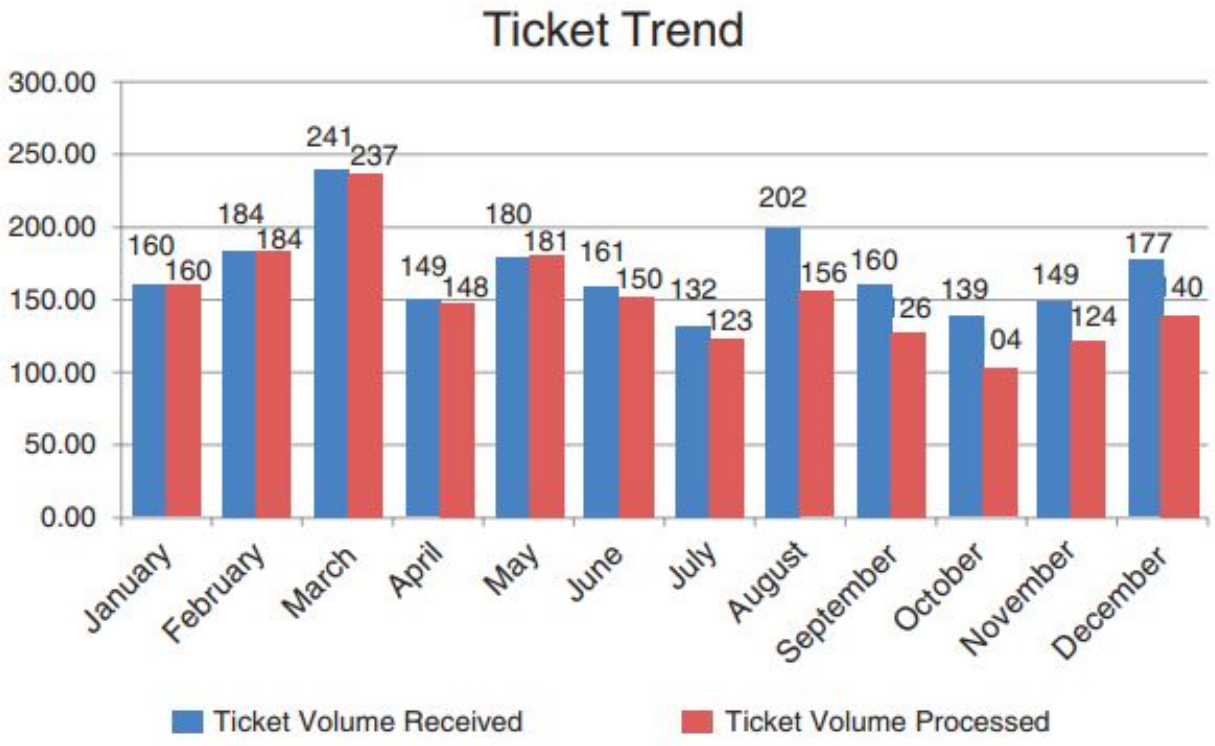
Exploratory Analysis is what you do to understand the data and figure out what might be noteworthy or interesting to highlight to others.

When we're at the point of communicating our analysis to our audience, we really want to be in the EXPLANATORY Space .

Storytelling with data

1) The importance of context
2) **Choosing an effective visual**
3) Clutter is your enemy!
4) Focus your Audience’s attention
5) Think like a designer

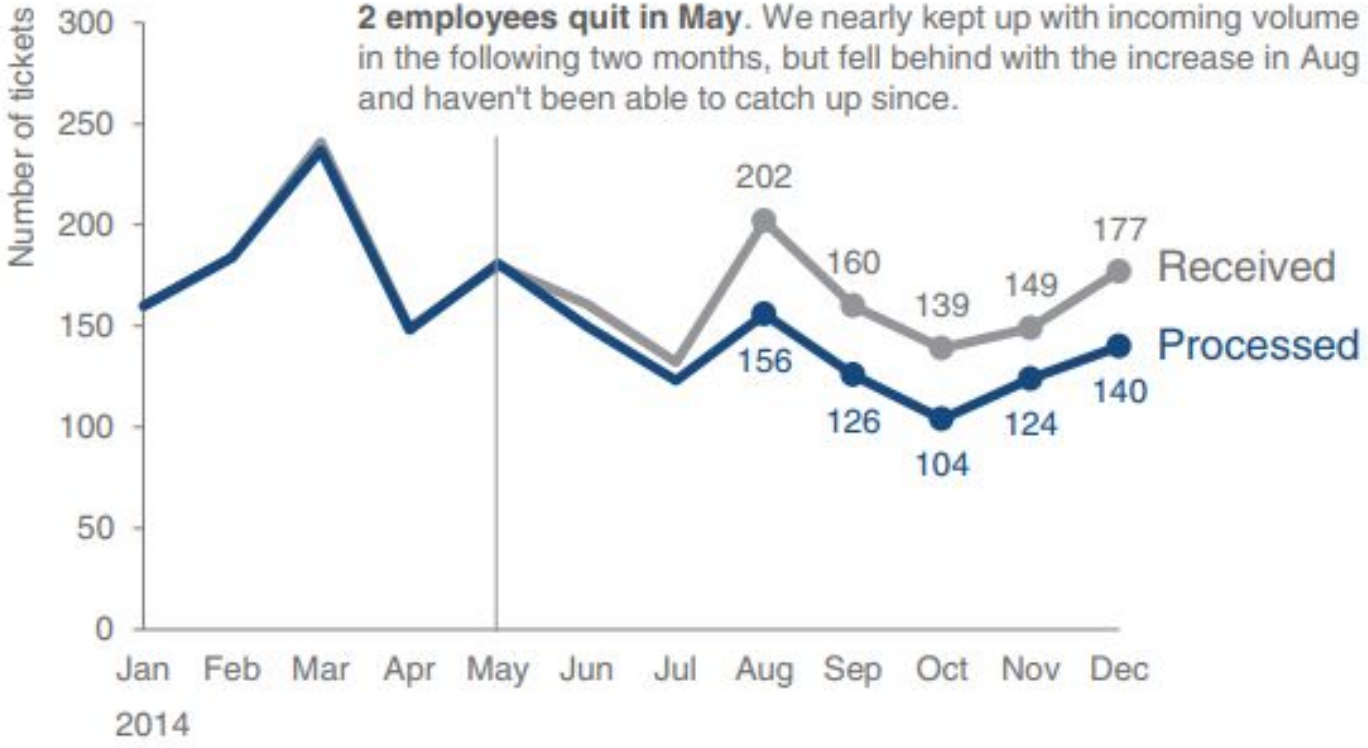
A request for hiring 2 new employees

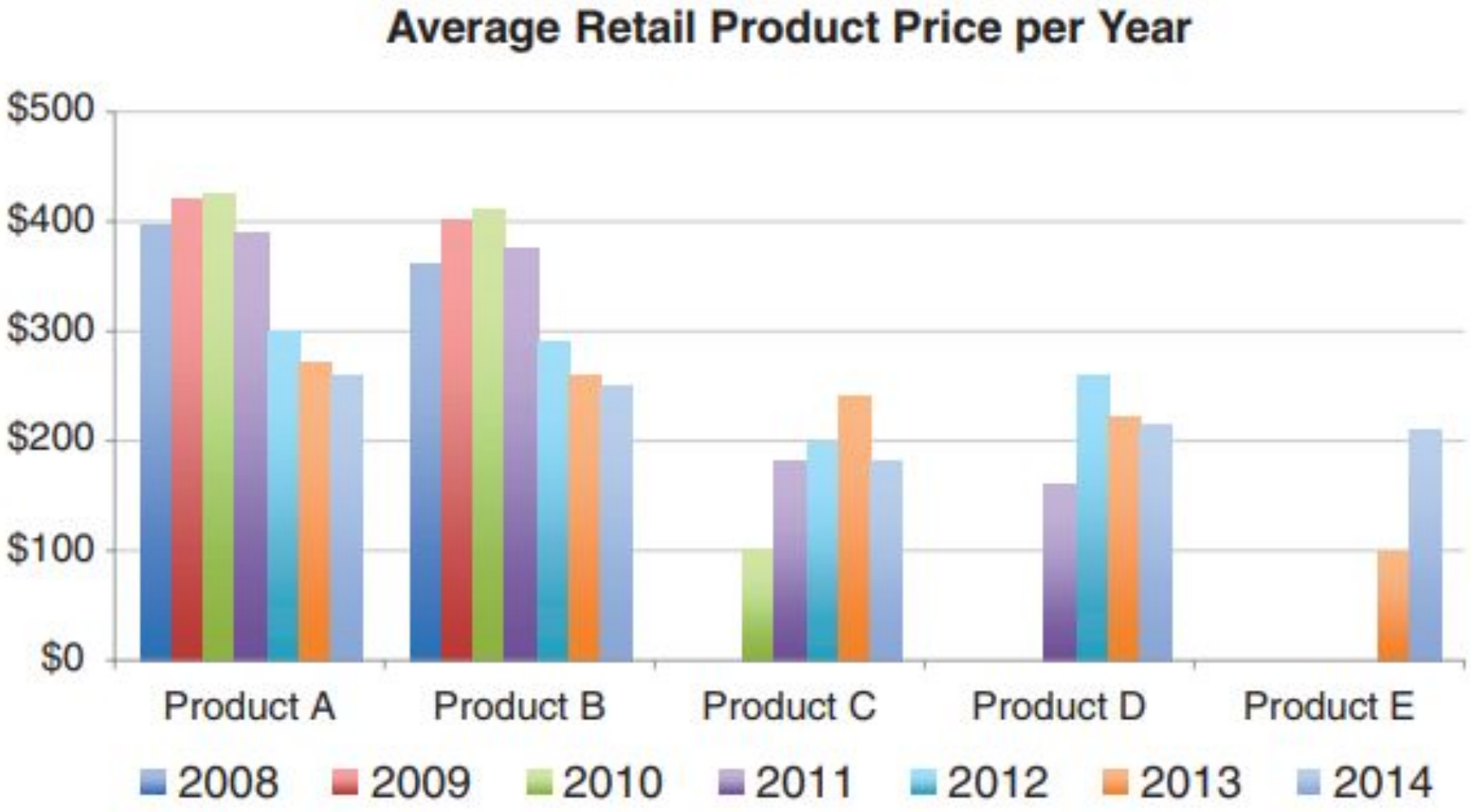


Please approve the hire of 2 FTEs

to backfill those who quit in the past year

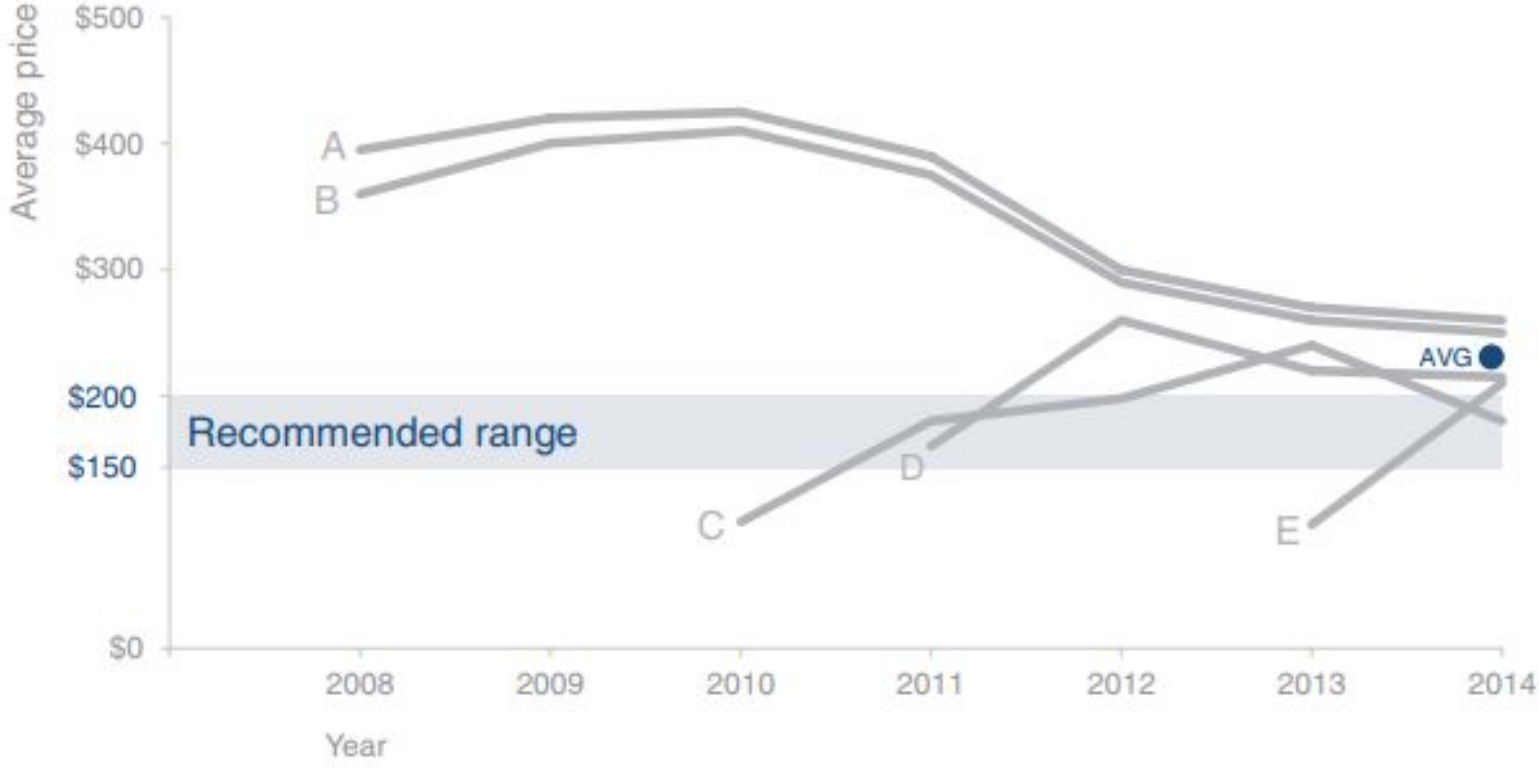
Ticket volume over time




Average Retail Product Price per Year
$500
$400
$300
$200
$100
$0
Product A
Product B
Product C
Product D
Product E
2008
2009
2010
2011
2012
2013
2014

To be competitive, we recommend introducing our product *below the $223 average* price point in the **$150–$200 range**

Retail price over time by product



Types of Data

Qualitative (Categorical)

Sales, HR, PR

Cairo, Giza, Alexandria

Qualitative (Ordinal)

Gold, Silver, Bronze

Excellent, Good, Poor

Quantitative (Discrete or Continuous)

Weight

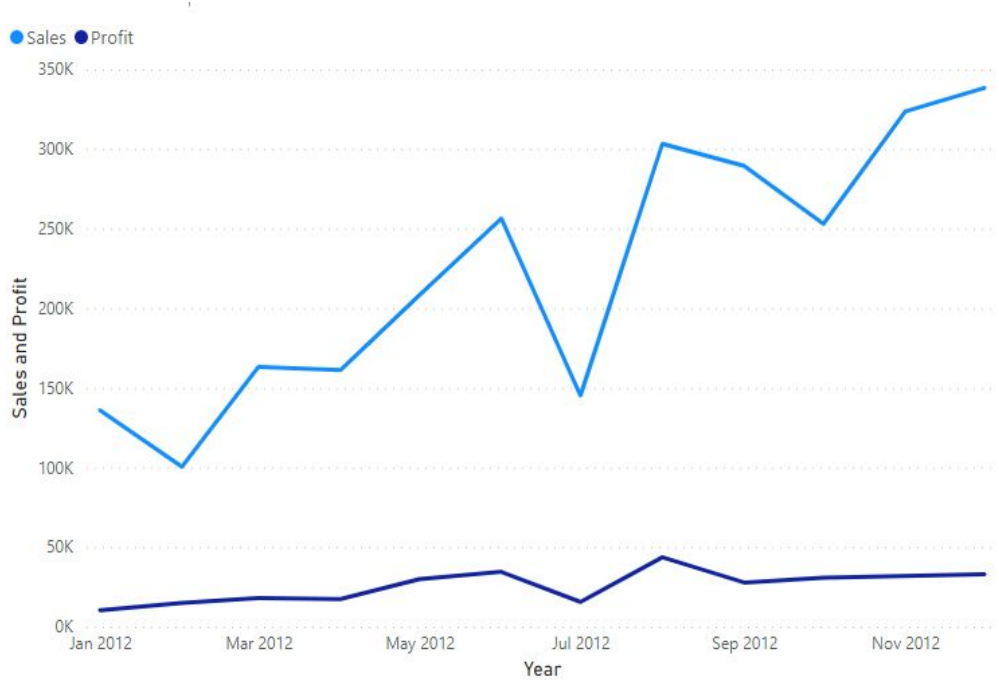
Cost

Card (Simple Text)

Single number cards display a single fact, a single data point. Sometimes a single number is the most important thing you want to track in your dashboard or report, such as total sales, market share year over year, or total opportunities.

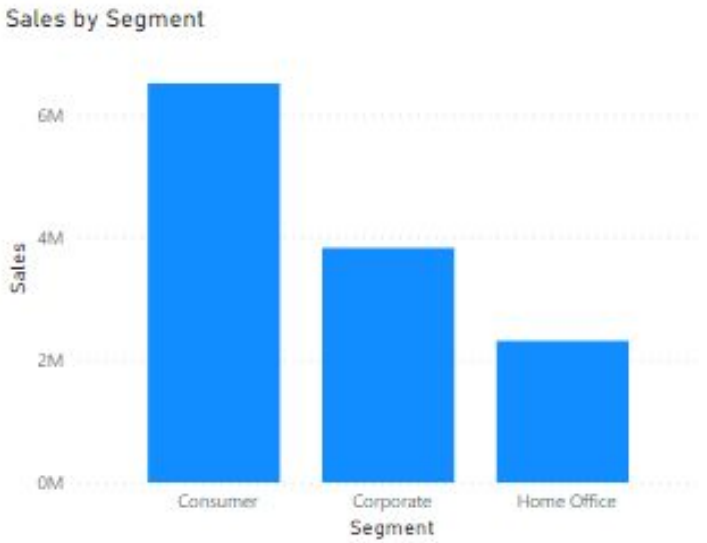
Line Chart

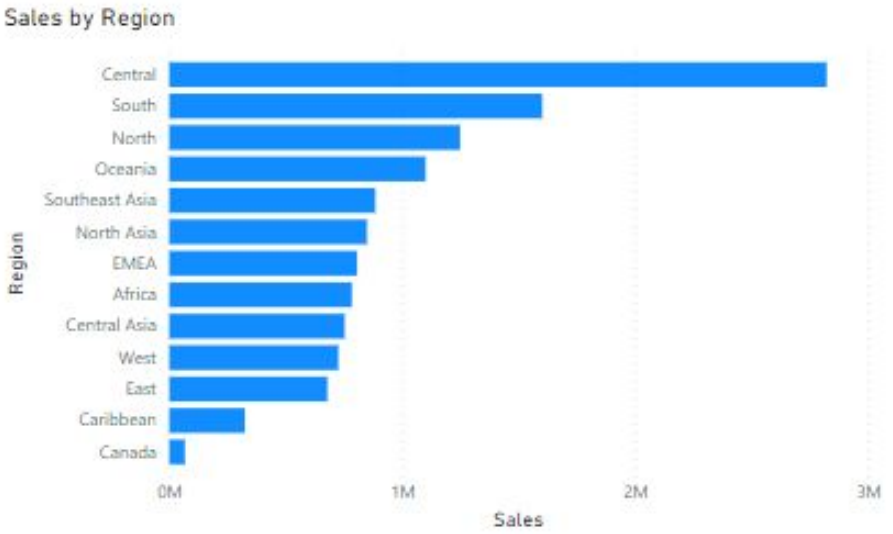
Line Graphs are most commonly used to Plot Continuous data.
Often, our continuous data is in some unit of time: days, months, quarters, or years.



Bar Chart (Vertical and Horizontal)

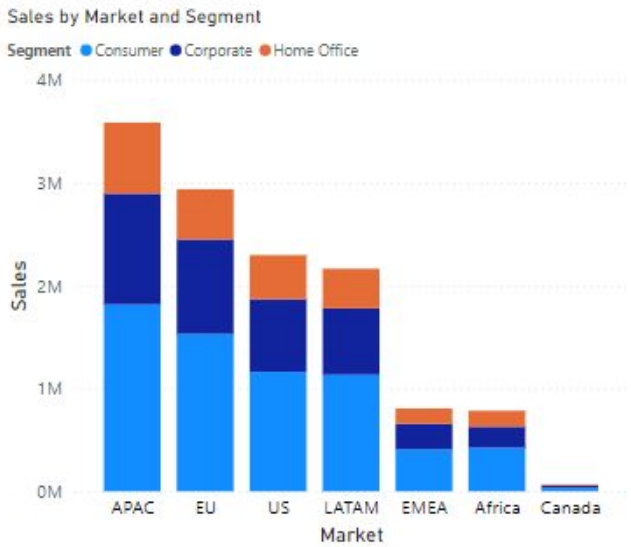
Bar charts are the standard for looking at a specific value across different categories.





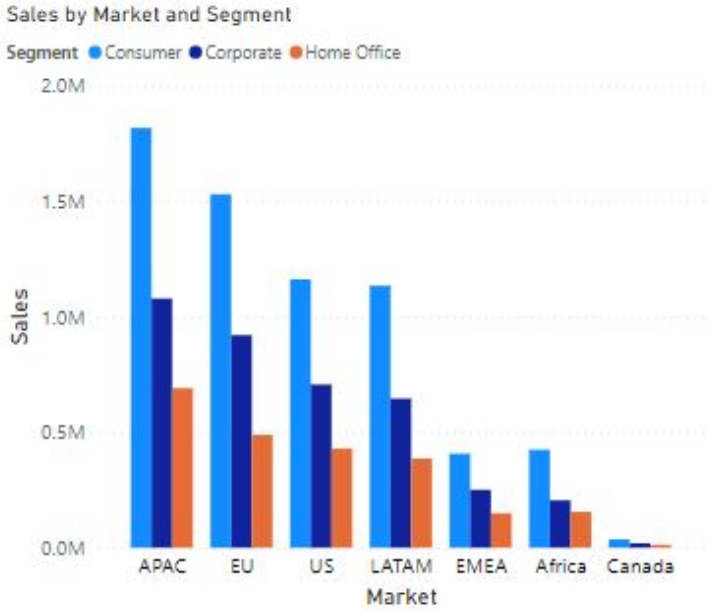
Stacked Bar Chart

Compare values across categories and their subcomponents. But be careful ! It's hard to compare subcomponents across the various categories once you get beyond the bottom series; because you no longer have a consistent baseline to use to compare.



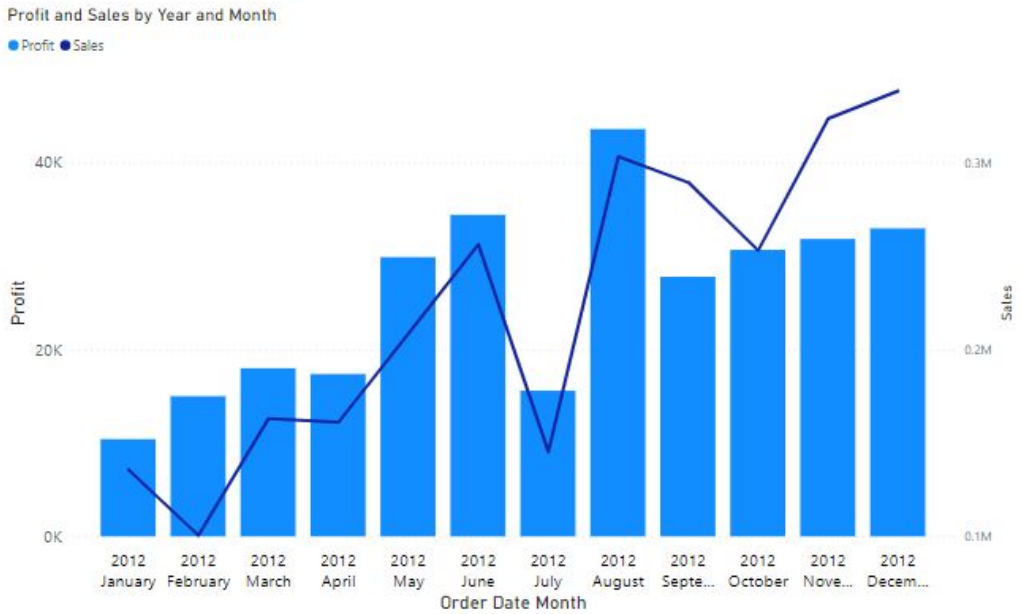
Clustered Bar/Column Chart

Similar in use to stacked bar chart but there is no total for each category, it solves the baseline problem but still use it with caution because it can be difficult to read.



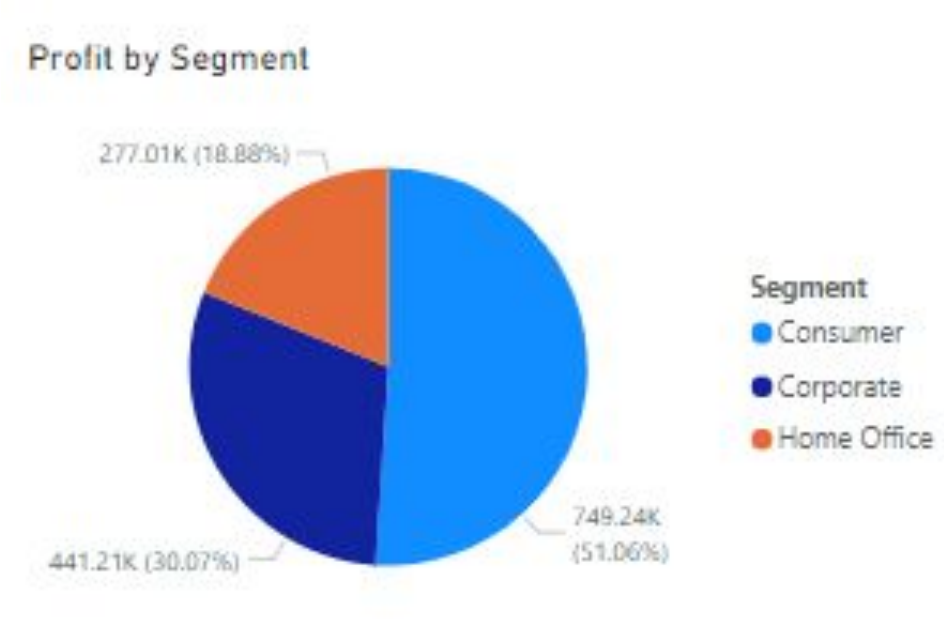
Combo Chart

A combo chart combines a column chart and a line chart. Combining the two charts into one lets you make a quicker comparison of the data. Combo charts can have one or two Y axes, so be sure to look closely.



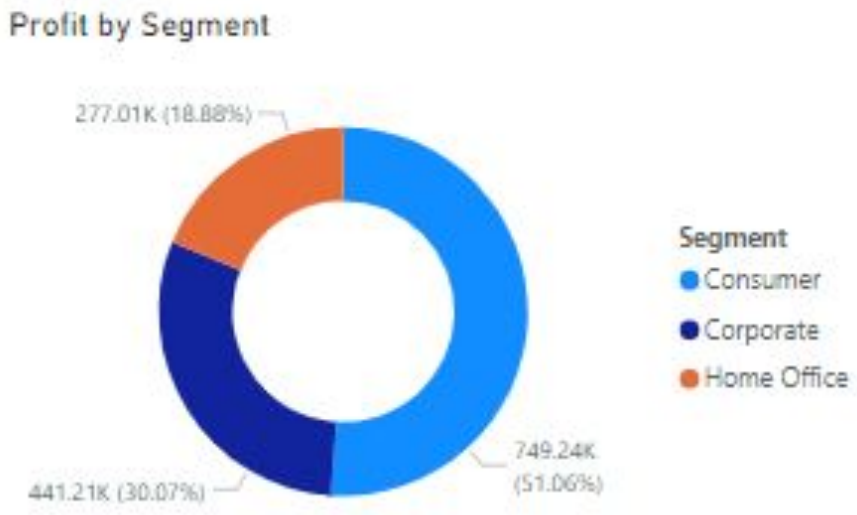
Pie Chart

Pie charts show the relationship of parts to a whole.



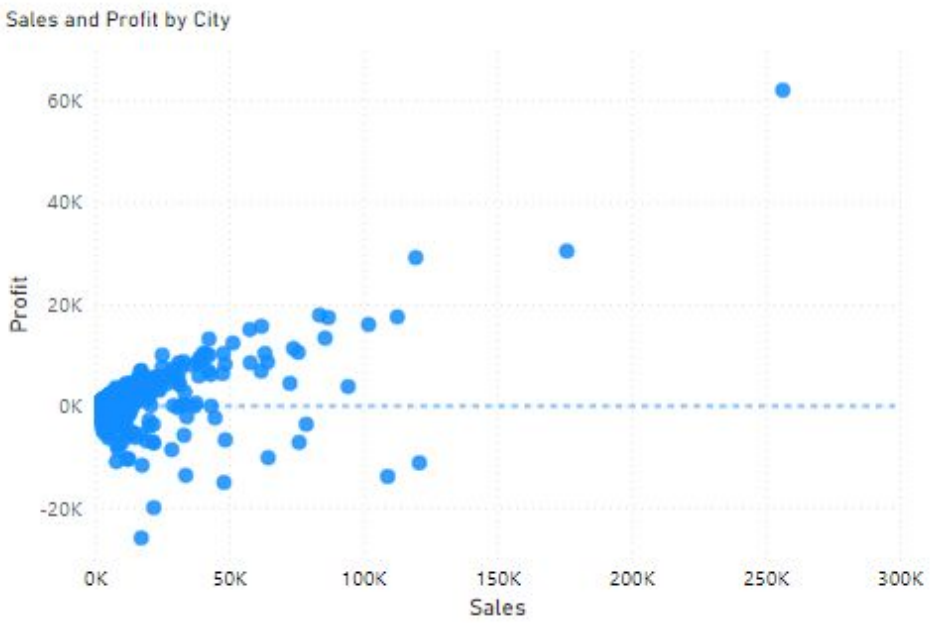
Donut Chart

Doughnut charts are similar to pie charts. They show the relationship of parts to a whole. The only difference is that the center is blank and allows space for a label or icon.



Scatter Plot

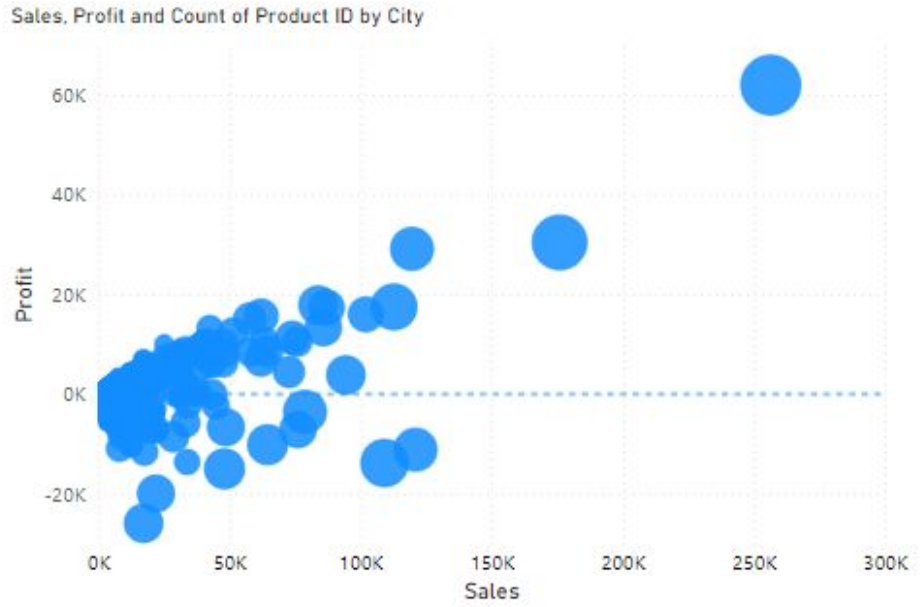
A scatter chart always has two value axes to show one set of numerical data along a horizontal axis and another set of numerical values along a vertical axis.



Bubble Chart

A bubble chart replaces data points with bubbles, with the bubble size representing an additional dimension of the data.

Hanz Rosling: https://www.ted.com/talks/hans_rosling_the_best_stats_you_ve_ever_seen



Gauge Chart

A radial gauge chart has a circular arc and displays a single value that measures progress toward a goal/KPI.

KPI Chart

A Key Performance Indicator (KPI) is a visual cue that communicates the amount of progress made toward a measurable goal.

KPIs are a great choice:

•To measure progress (what am I ahead or behind on?).

•To measure distance to a goal (how far ahead or behind am I?).

Map Chart

Use a basic map to associate both categorical and quantitative information with spatial locations.

Shape Map

Shape maps compare regions on a map using color shading. A shape map can't show precise geographical locations of data points on a map. Instead, its main purpose is to show relative comparisons of regions on a map.

For more info:
https://www.mssqltips.com/sqlservertip/6106/power-bi-bubble-map-shape-map-and-filled-map-examples/

Filled Map

Rather than representing data in points, It fills the countries in the visualizations.It comes with many limitations, but is useful when dealing with clear boundaries data like country, state.

Slicers

A slicer is a standalone chart that can be used to filter the other visuals. It's useful for creating more focused reports.

Table

Tables work well with quantitative comparisons where you are looking at many values for a single category.

Market	Sales	Shipping Cost	Profit
Africa	783,773.21	88,139.47	88,871.63
APAC	3,585,744.13	387,165.11	436,000.05
Canada	66,928.17	7,405.63	17,817.39
EMEA	806,161.31	88,375.73	43,897.97
EU	2,938,089.06	309,422.35	372,829.74
LATAM	2,164,605.17	234,133.61	221,643.49
US	2,297,200.86	238,173.79	286,397.02
Total	**12,642,501.91**	**1,352,815.70**	**1,467,457.29**

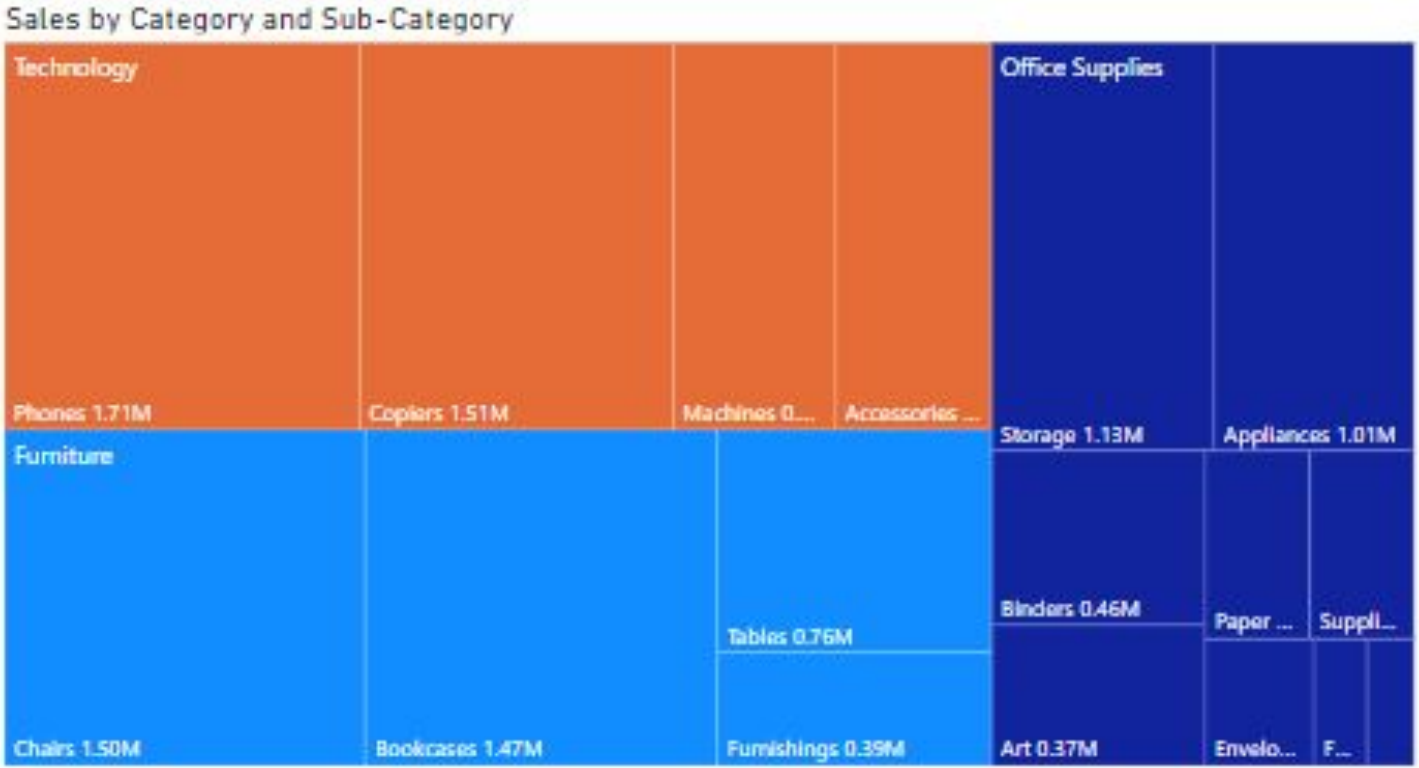
Matrix

Matrix has the same use as tables but can have more dimensions.

Sales Breakdown

Category	2011	2012	2013	2014	Total
Furniture	**756,192.38**	**858,902.56**	**1,117,723.55**	**1,378,055.69**	**4,110,874.19**
Bookcases	259,396.29	317,953.45	376,025.80	513,196.71	**1,466,572.24**
Chairs	285,730.98	295,058.14	427,514.37	493,378.28	**1,501,681.76**
Furnishings	63,933.63	81,804.47	111,819.85	128,020.31	**385,578.26**
Tables	147,131.49	164,086.50	202,363.54	243,460.40	**757,041.92**
Office Supplies	**675,606.45**	**795,094.63**	**1,010,717.64**	**1,305,651.51**	**3,787,070.23**
Technology	**827,652.06**	**1,023,441.51**	**1,277,305.25**	**1,616,158.67**	**4,744,557.50**
Total	**2,259,450.90**	**2,677,438.69**	**3,405,746.45**	**4,299,865.87**	**12,642,501.91**

Treemap


Sales by Category and Sub-Category
Technology
Phones 1.71M
Copiers 1.51M
Machines 0....
Accessories ...
Office Supplies
Storage 1.13M
Appliances 1.01M
Furniture
Chairs 1.50M
Bookcases 1.47M
Tables 0.76M
Furnishings 0.39M
Binders 0.46M
Paper ...
Suppli...
Art 0.37M
Envelo...
F...

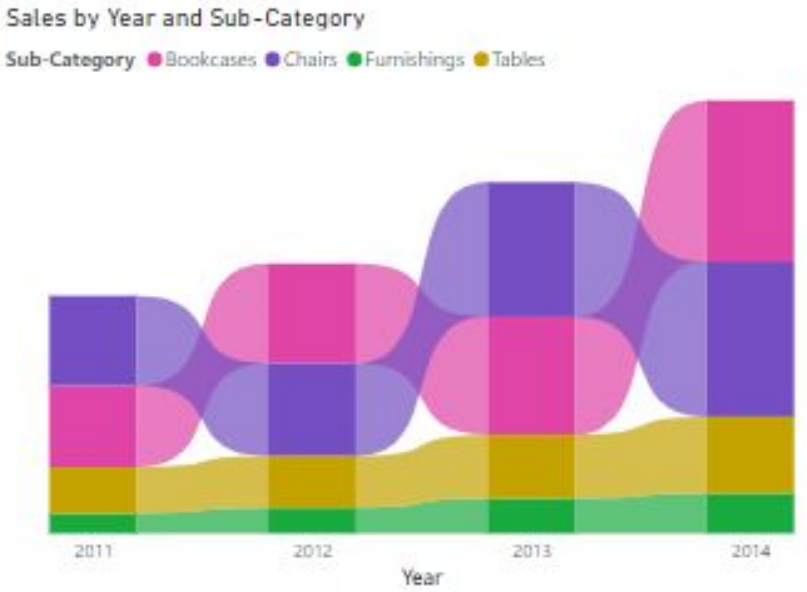
Treemap

- Treemaps are charts of colored rectangles, with size representing value. They can be hierarchical, with rectangles nested within the main rectangles. The space inside each rectangle is allocated based on the value being measured. And the rectangles are arranged in size from top left (largest) to bottom right (smallest).

Check this cool visual: https://finviz.com/map.ashx

Ribbon Chart

Ribbon Char shows ranking of categories according to a certain measure across time.

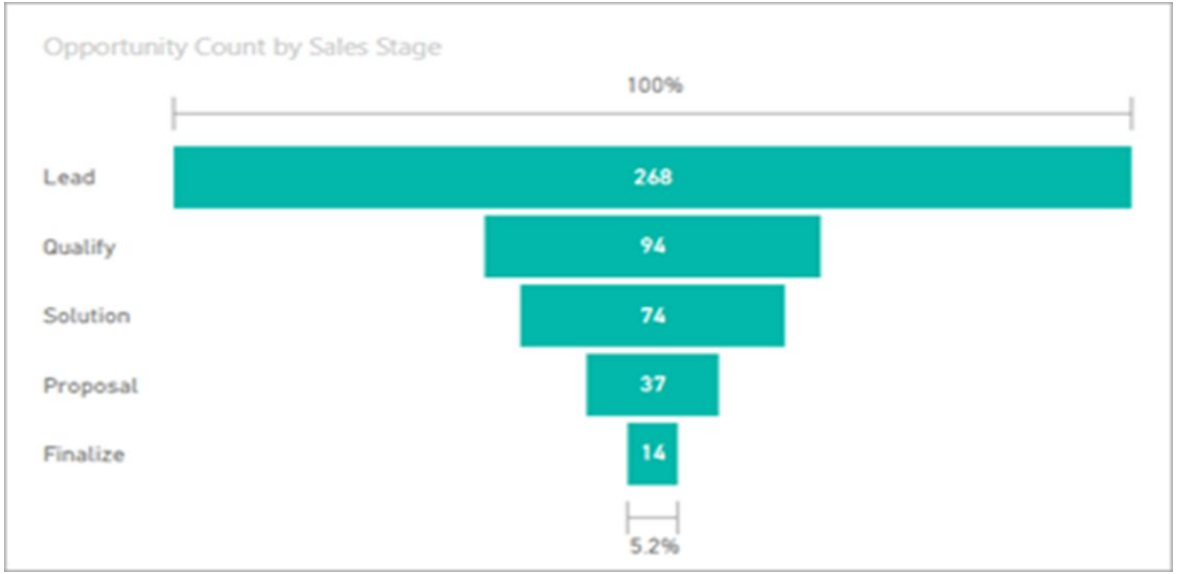


Waterfall Chart

•A waterfall chart shows a running total as values are added or subtracted. It's useful for understanding how an initial value (for example, net income) is affected by a series of positive and negative changes.

Funnel Chart

Funnels help visualize a process that has stages, and items flow sequentially from one stage to the next. One example is a sales process that starts with leads and ends with purchase fulfillment.



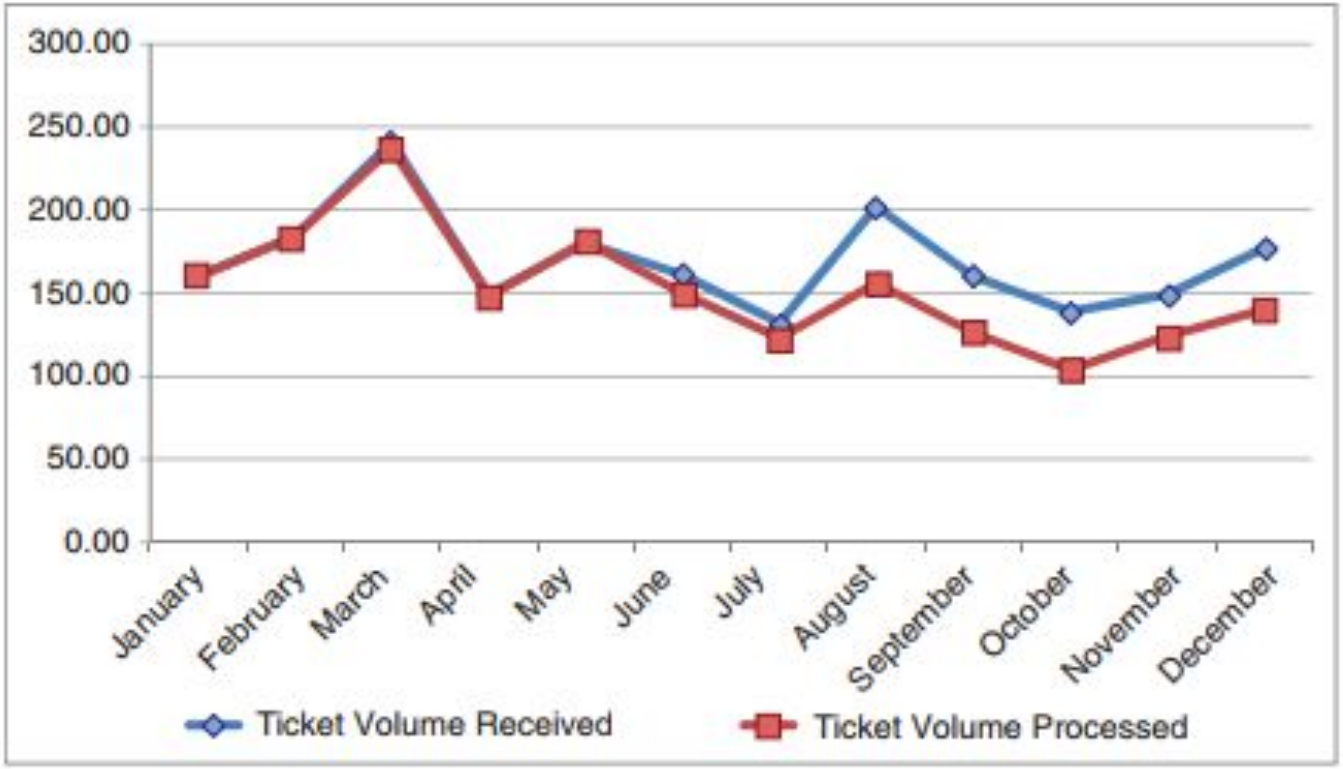
Storytelling with data

1) The importance of context
2) Choosing an effective visual
3) **Clutter is your enemy!**
4) Focus your Audience’s attention
5) Think like a designer

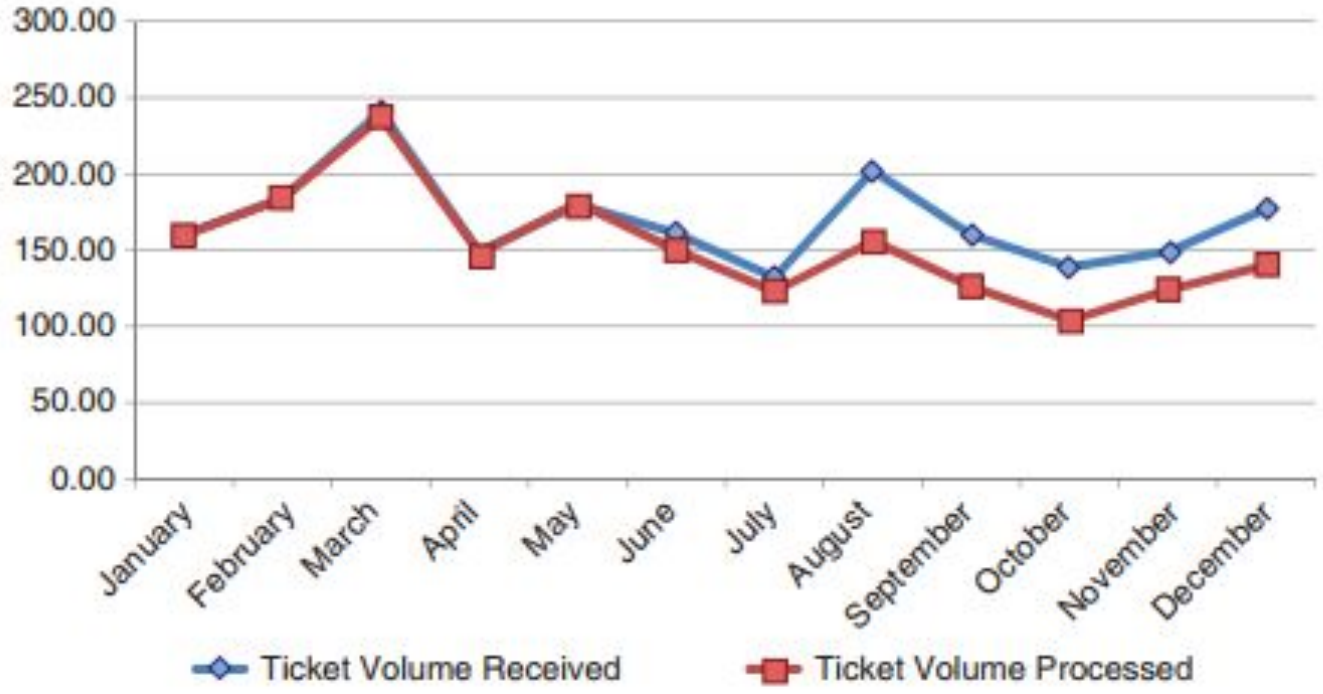
Eliminating Clutter

Maximizing the signal-to-noise ratio (also known as data-ink ratio), where the signal is the information we want to communicate, and the noise are those elements that either don’t add to, or in some cases detract from, the message we are trying to impart to our audience.

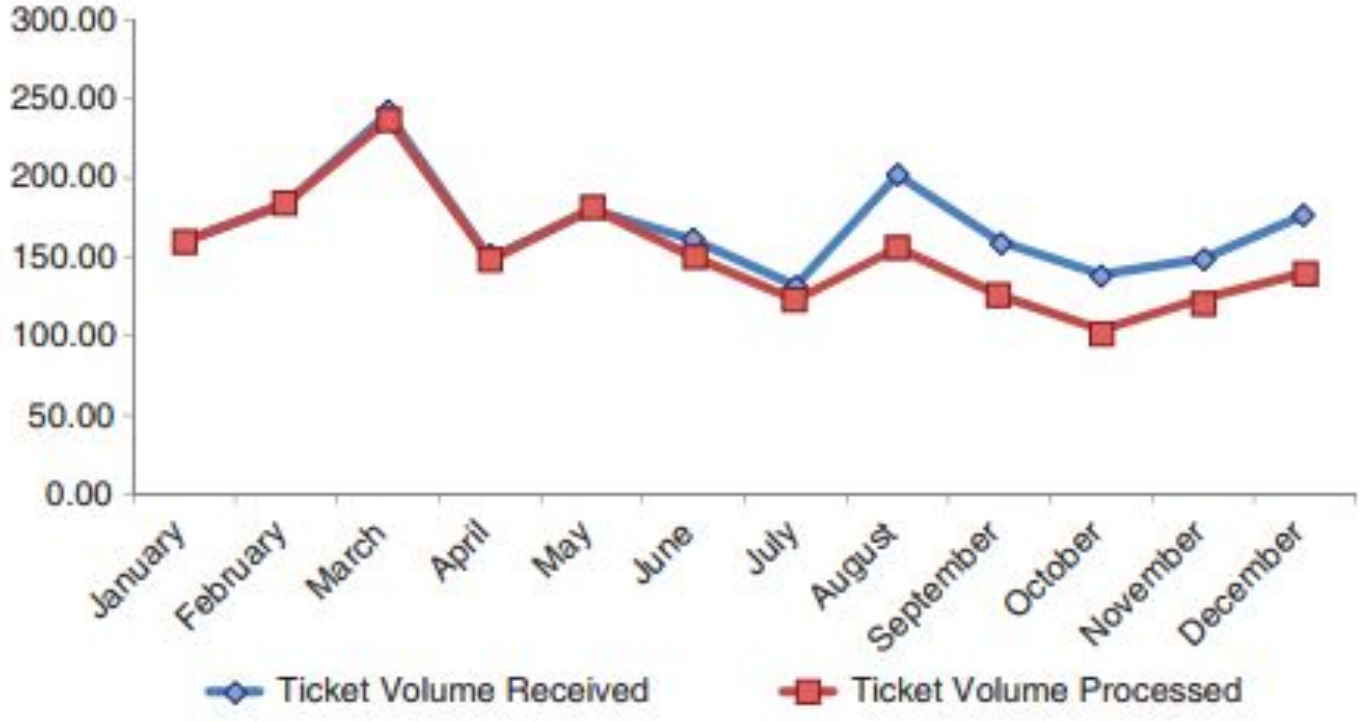
Example


300.00
250.00
200.00
150.00
100.00
50.00
0.00
January
February
March
April
May
June
July
August
September
October
November
December
Ticket Volume Received
Ticket Volume Processed

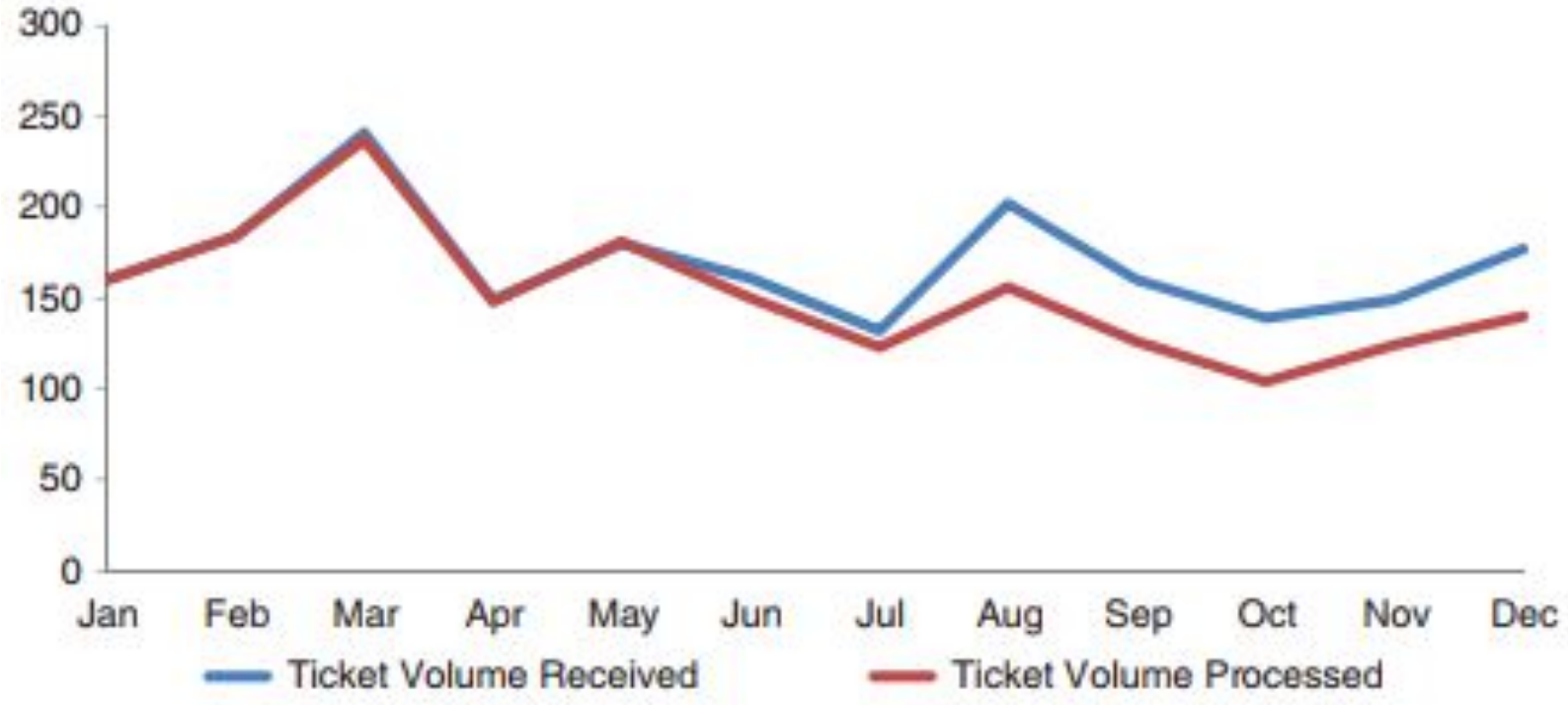
1- Remove chart border



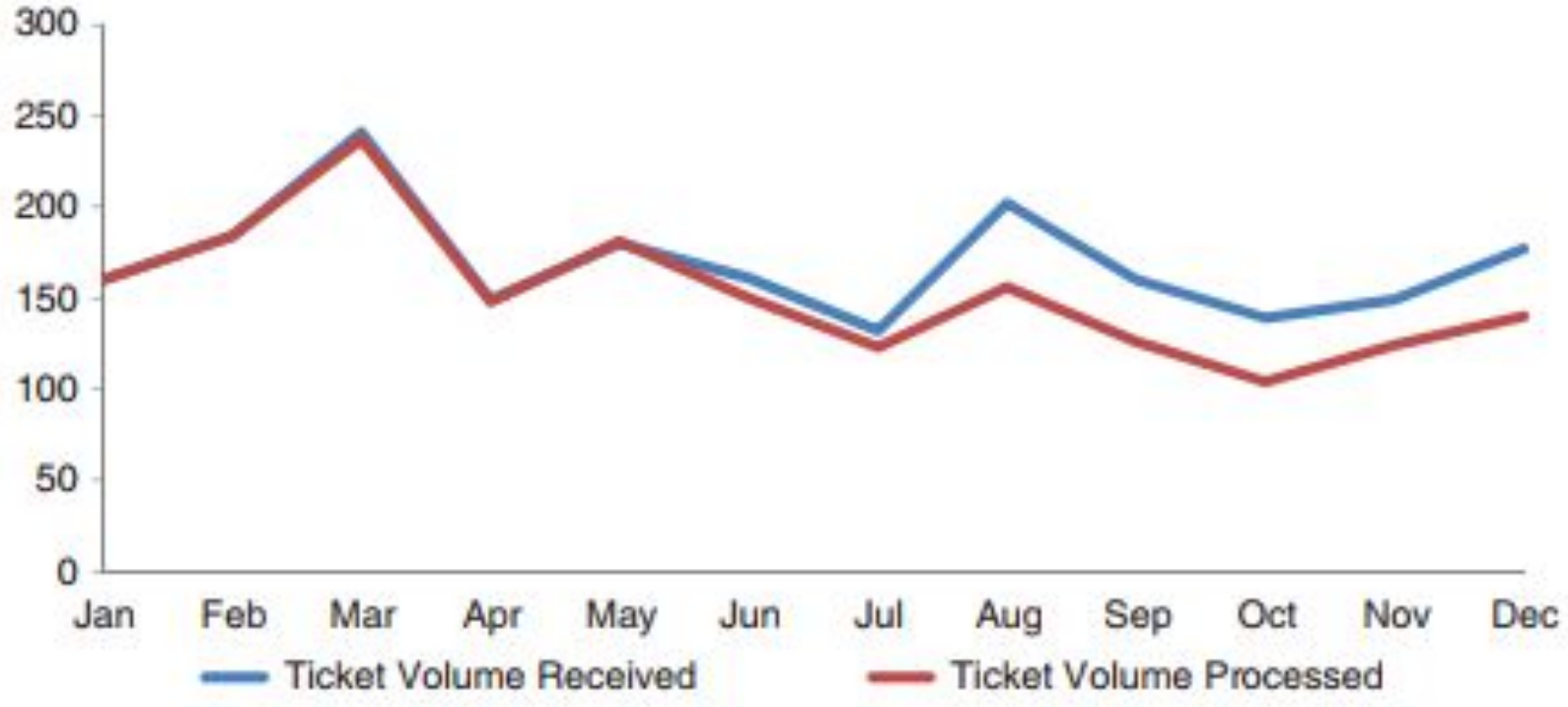
2- Remove gridlines



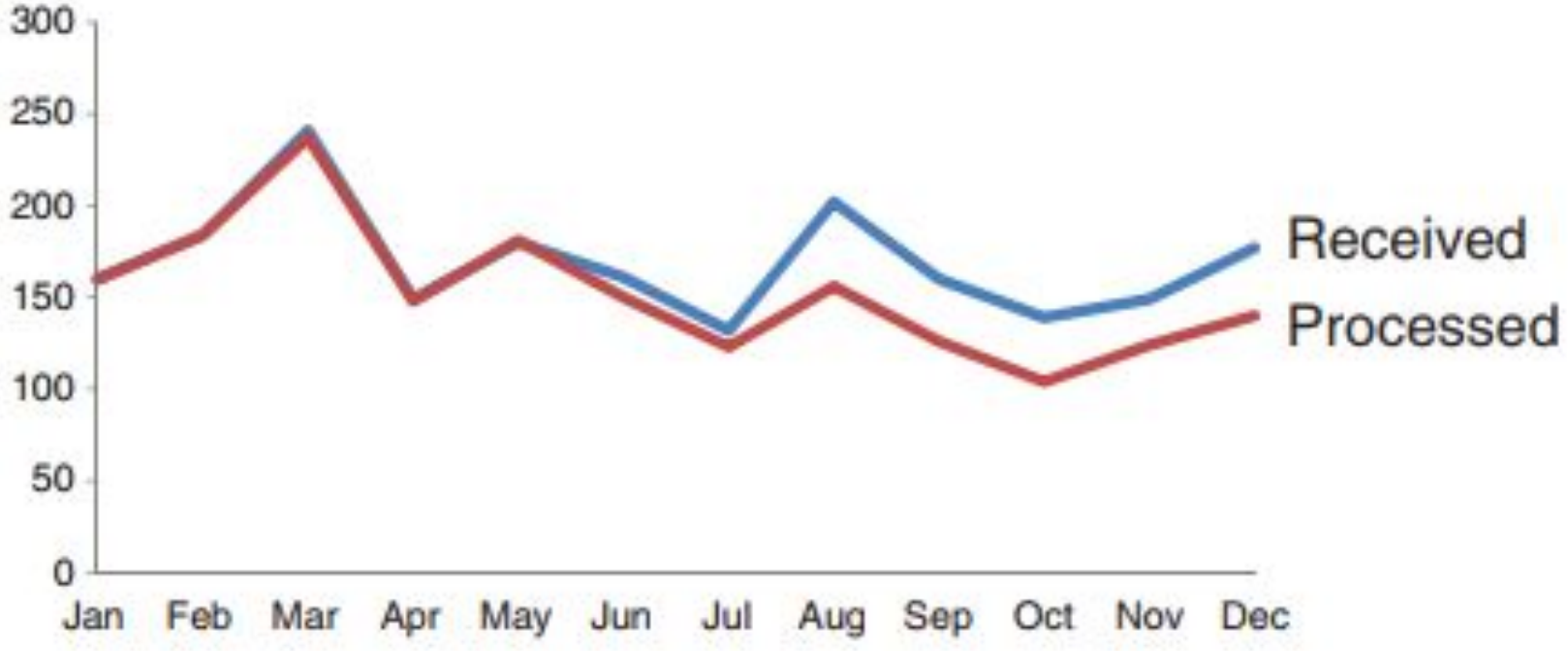
3- Remove data markers



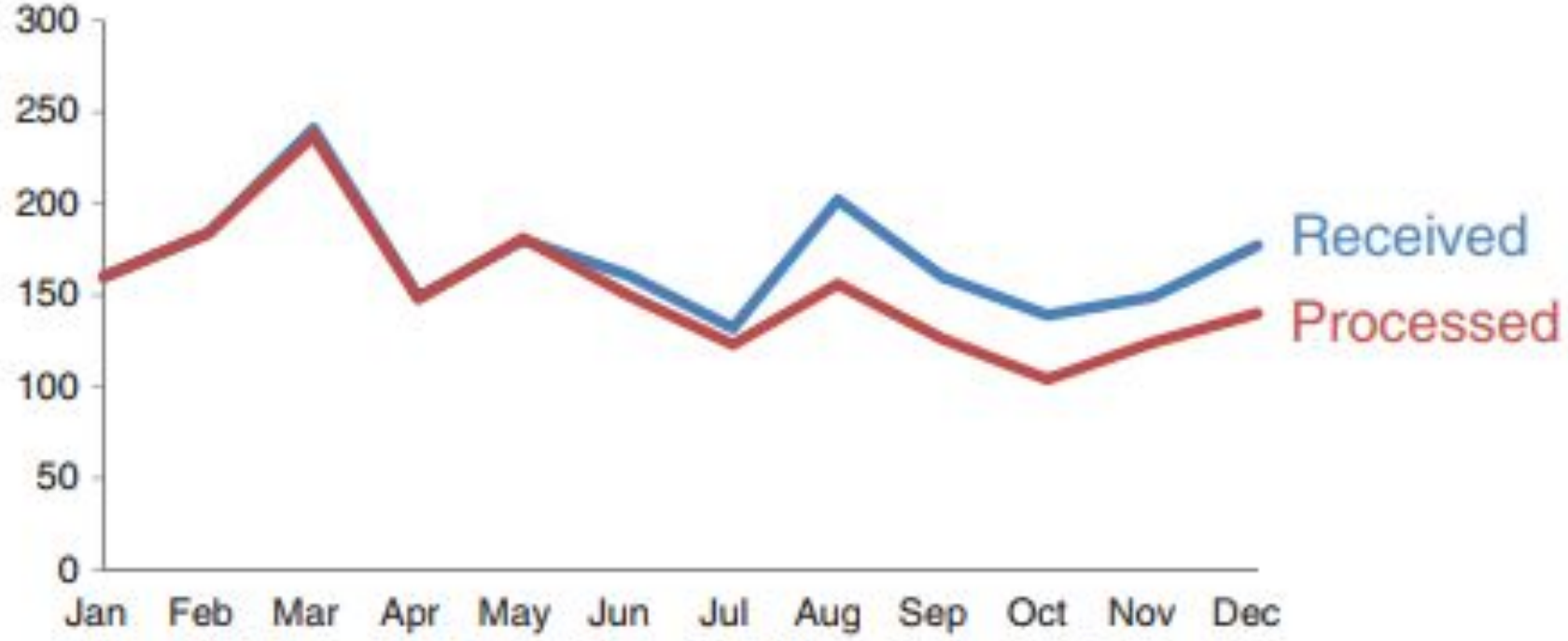
4- Clean up axis labels

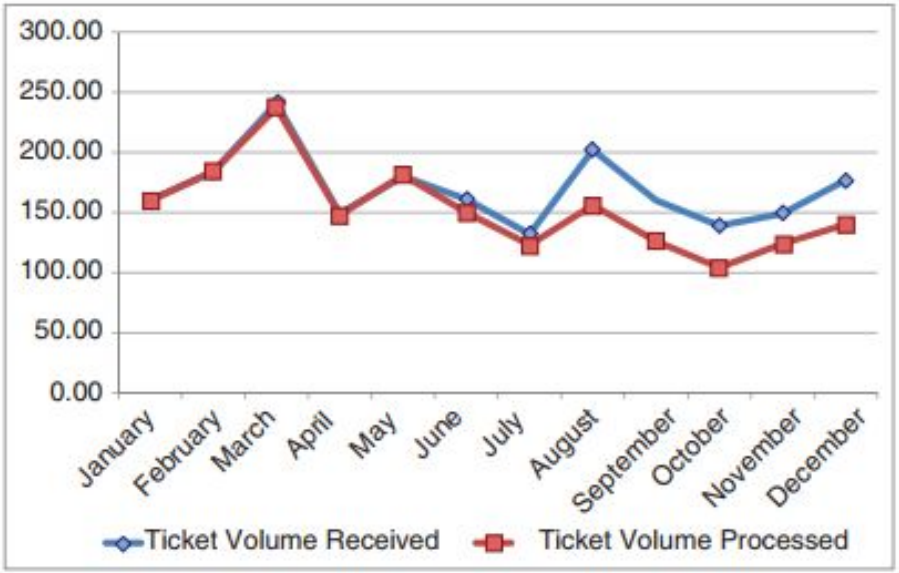


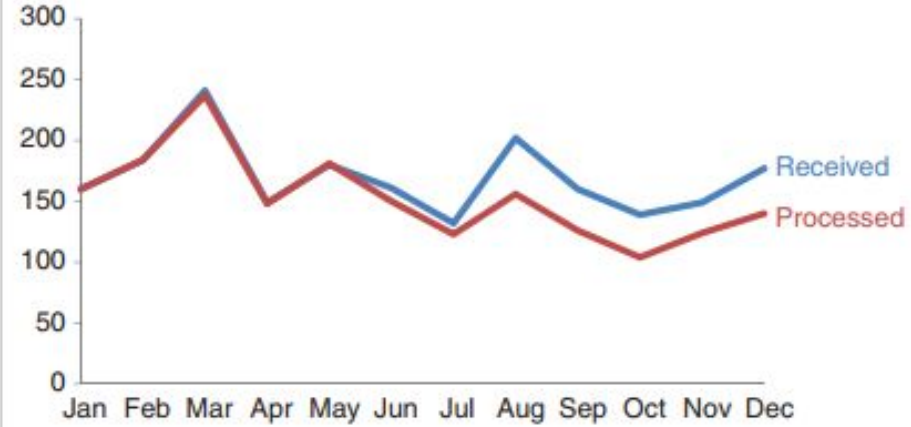
5- Label data directly



6- Leverage Consistent Color




300.00
250.00
200.00
150.00
100.00
50.00
0.00
January
February
March
April
May
June
July
August
September
October
November
December
Ticket Volume Received
Ticket Volume Processed

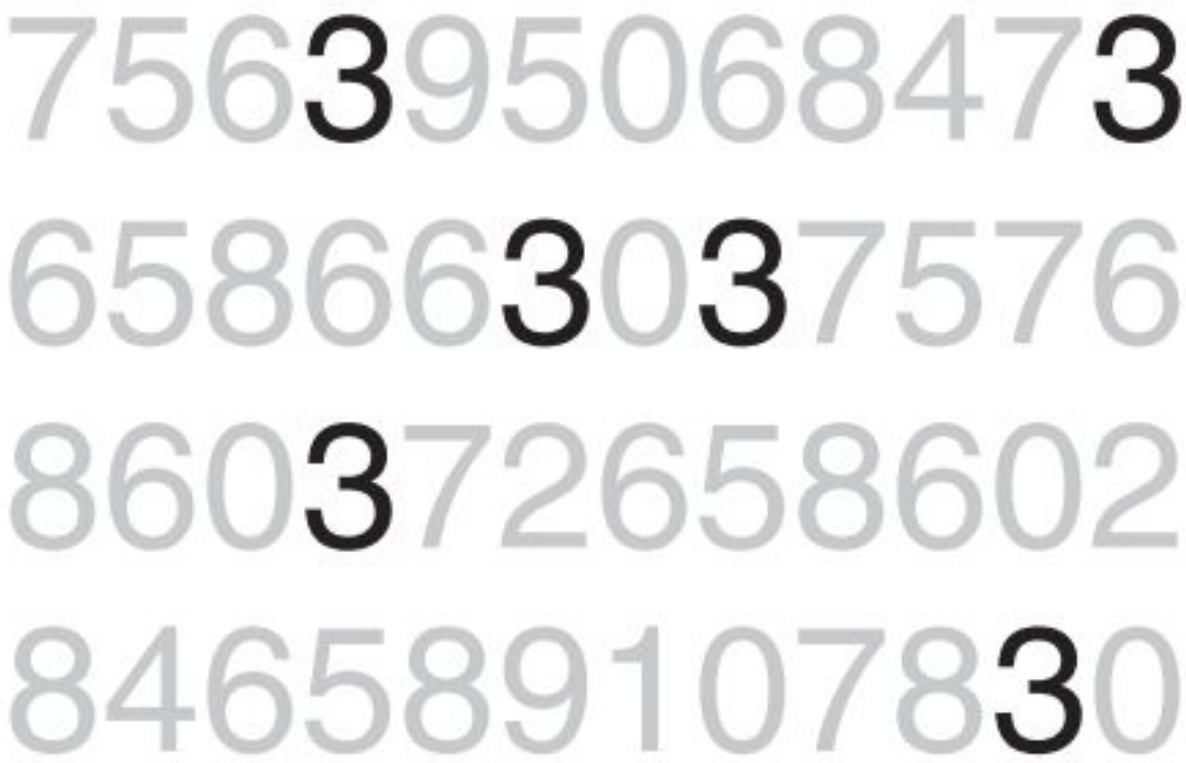

300
250
200
150
100
50
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Received
Processed

Storytelling with data

1) The importance of context
2) Choosing an effective visual
3) Clutter is your enemy!
4) **Focus your Audience's attention**
5) Think like a designer

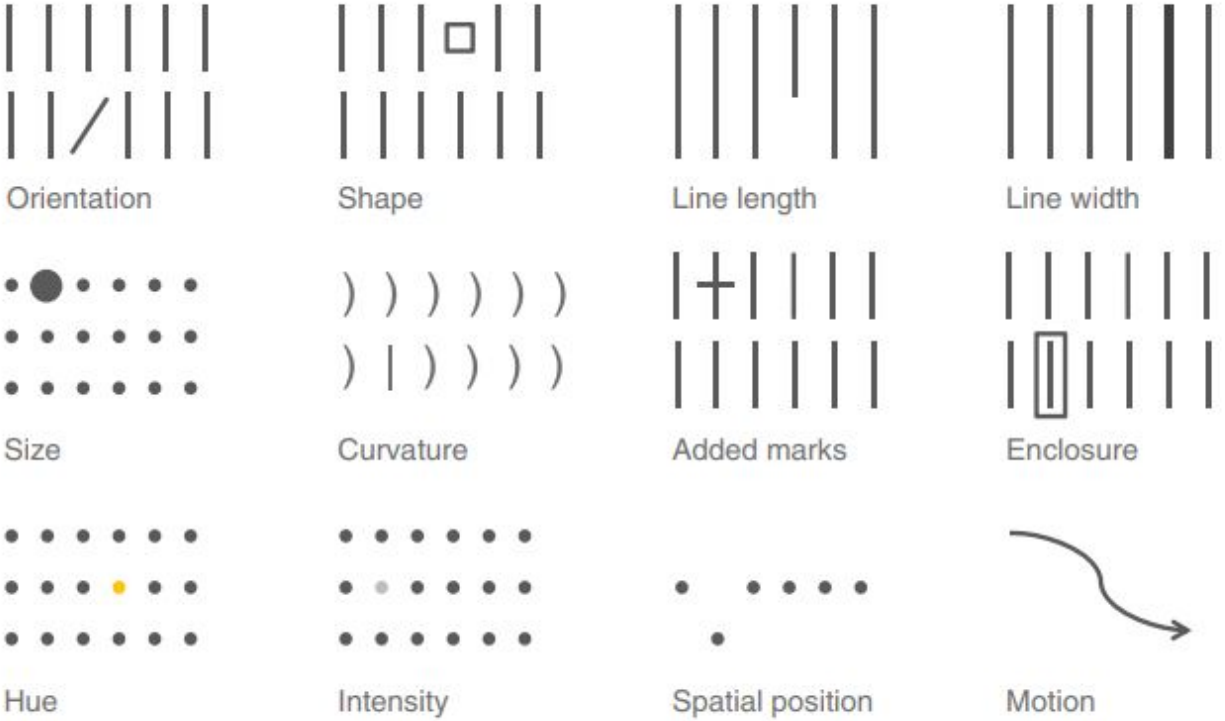
How Many 3s do you see ?

How Many 3s do you see ?



That is the magic of preattentive attributes!

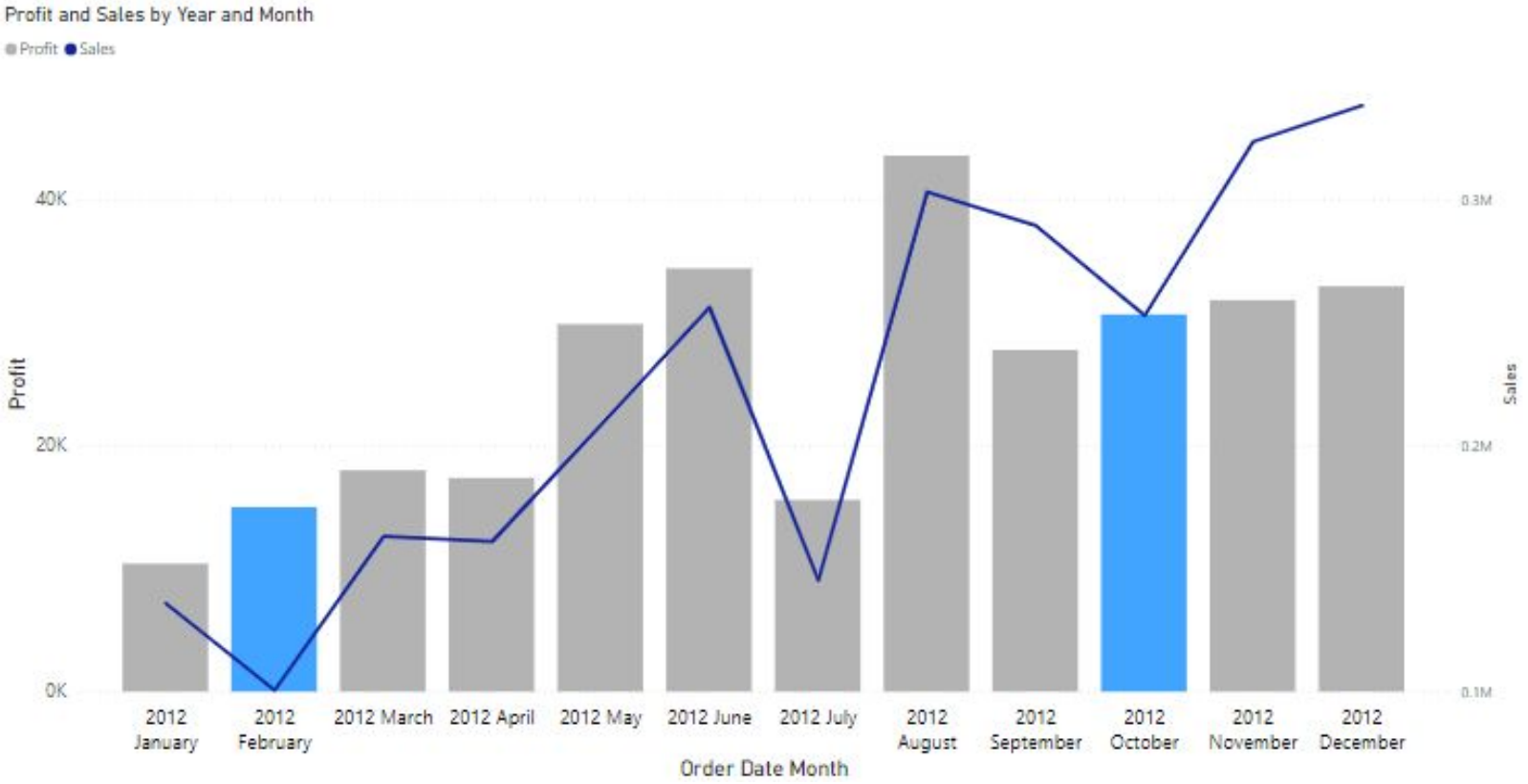
Preattentive attributes



Preattentive attributes in text

- Color
- Bold
- Italic
- Size
- Separate Spatially
- Underline
- Outline

In graphs:


Profit and Sales by Year and Month
Profit
Sales
40K
20K
0K
0.3M
0.2M
0.1M
Profit
Sales
2012 January
2012 February
2012 March
2012 April
2012 May
2012 June
2012 July
2012 August
2012 September
2012 October
2012 November
2012 December
Order Date Month