KEY PRINCIPLES OF DATA VISUALIZATION

Strive for **CLARITY** & **SIMPLICITY**

- *Maximize **impact**, minimize **noise***
- *If it doesn't **add value** or **serve a purpose**, get rid of it*

Focus on creating a **NARRATIVE**

- *Don't just show data, **tell a story***
- *Communicate key insights **clearly, quickly** and **powerfully***

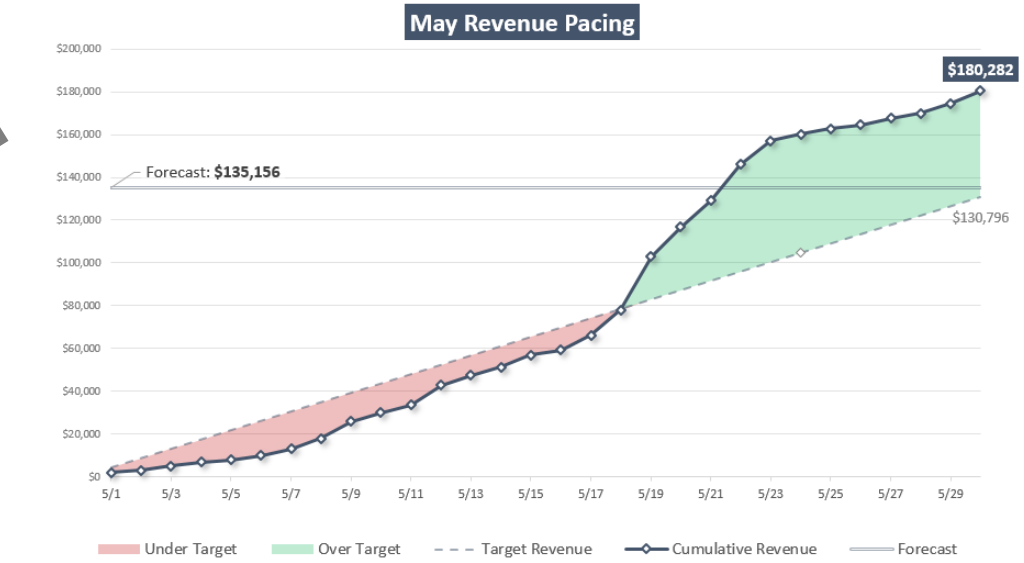
Strike a balance between **DESIGN** & **FUNCTION**

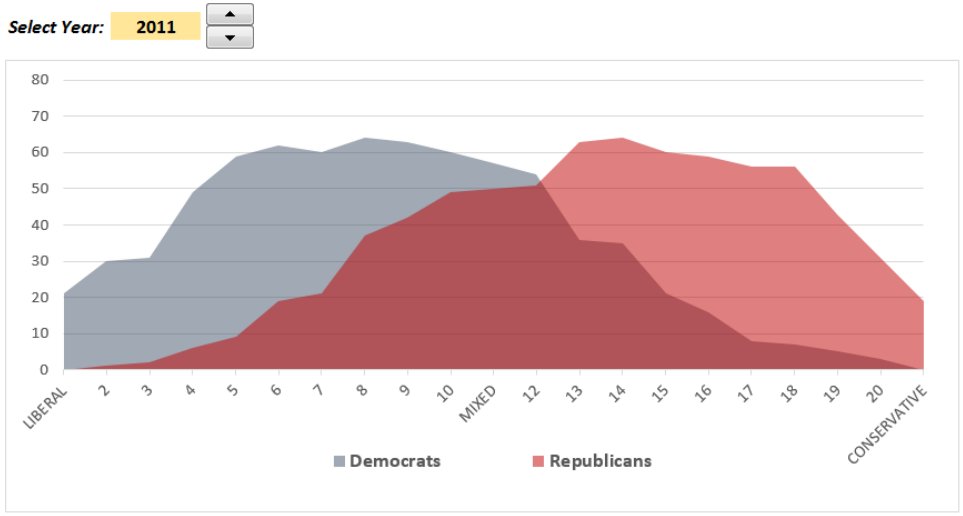
- *Selecting the right type of chart is **critical***
- ***Beautiful** is good, **functional** is better, **BOTH** is ideal*

THE GOOD, THE BAD, AND THE UGLY

THE GOOD

Dynamic formatting helps to strengthen the story





Clean, simple visualization with animation over time

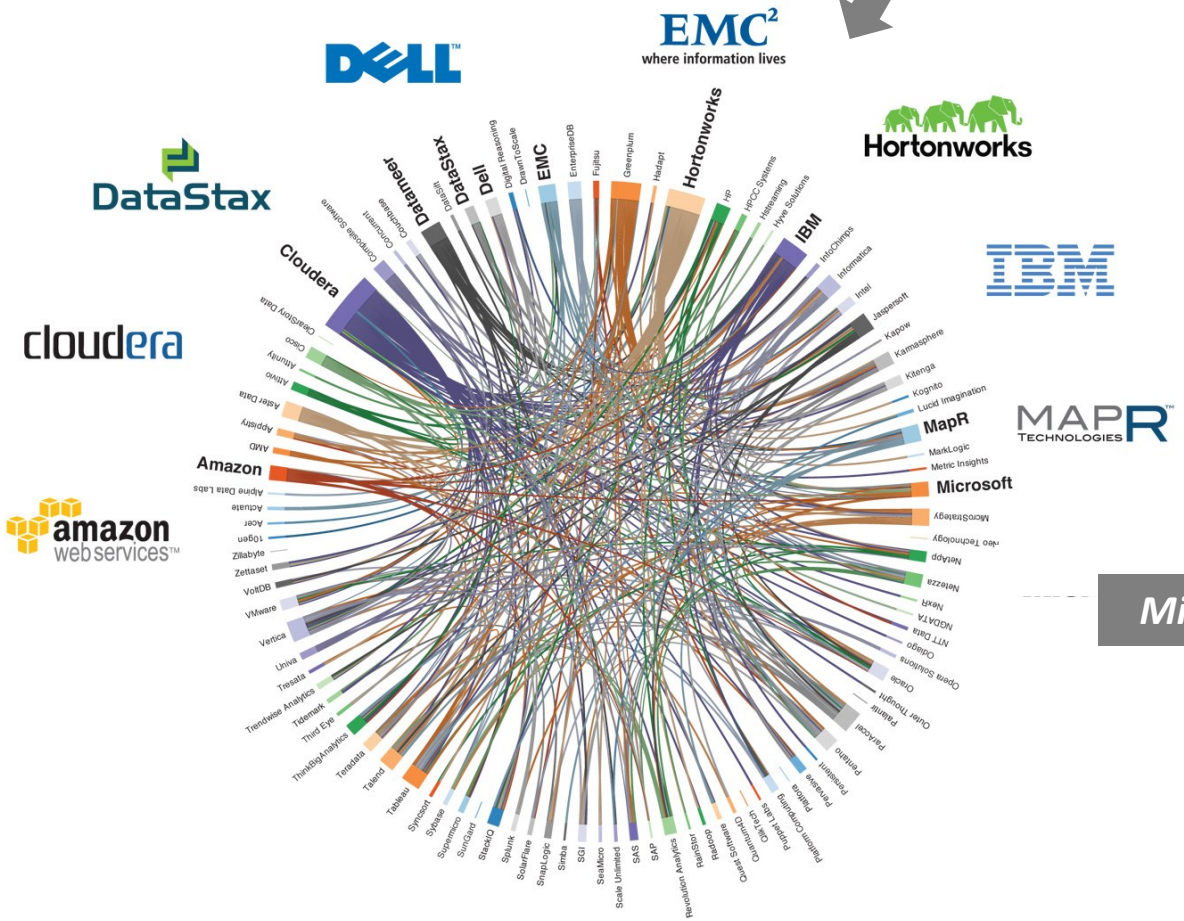
Simple, intuitive custom chart design

EXCEL MAVEN

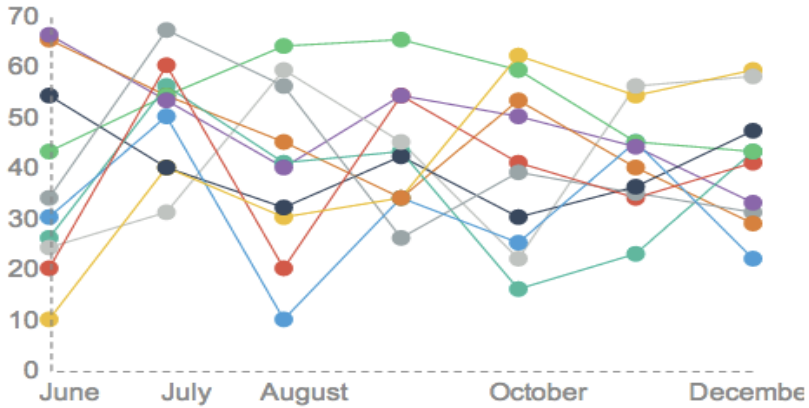
THE GOOD, THE BAD, AND THE UGLY

THE BAD

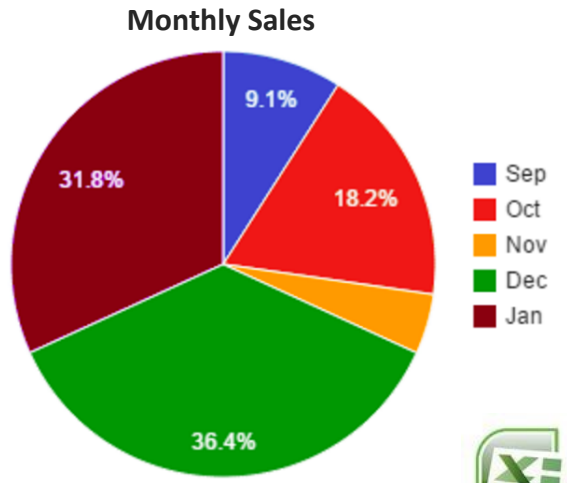
All design, no function



Busy, no clear narrative



Misleading chart type

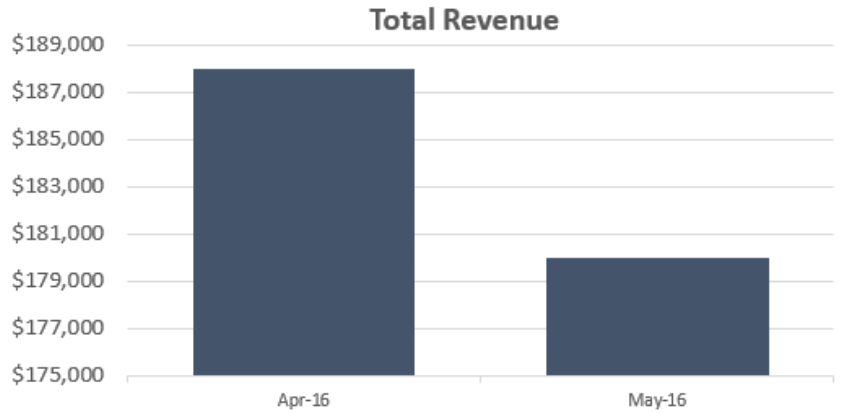


EXCEL MAVEN

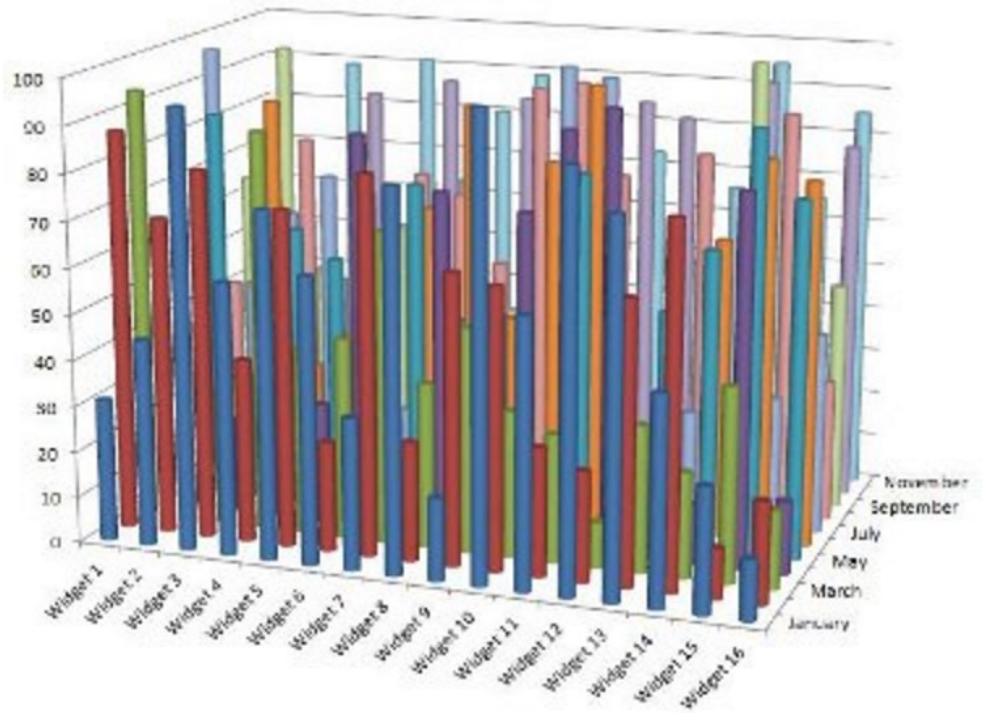
THE GOOD, THE BAD, AND THE UGLY

THE UGLY

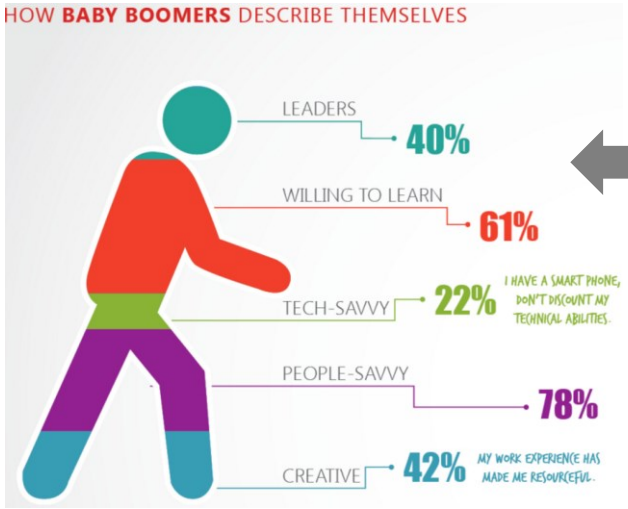
Misleading y-axis scale



Too many elements, distracting 3D design



Improper use of percentages & inconsistent scaling



EXCEL MAVEN

THE 3 KEY QUESTIONS

1. What **type of data** are you working with?
   - *Integer, real, categorical, time-series, geo-spatial, etc.*

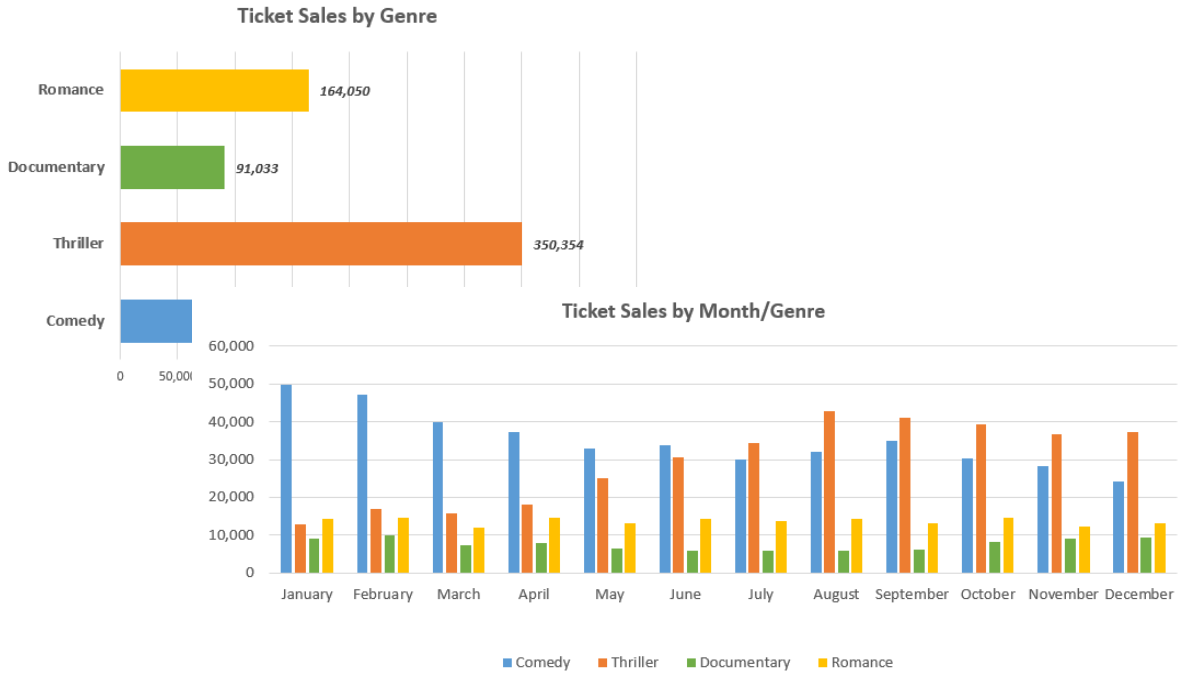
2. What are you trying to **communicate**?
   - *Relationship, comparison, composition, distribution, trending, etc.*

3. Who is the **end user** consuming this information?
   - *Analyst, CEO, client, intern, etc.*

BAR & COLUMN CHARTS

COMMONLY USED FOR:

- **Comparing numerical data across categories**



EXAMPLES:

- *Total sales by product type*
- *Population by country*
- *Revenue by department, by quarter*

PRO TIPS:

Use **stacked** or **clustered** bars/columns to group by subcategory or compare multiple metrics

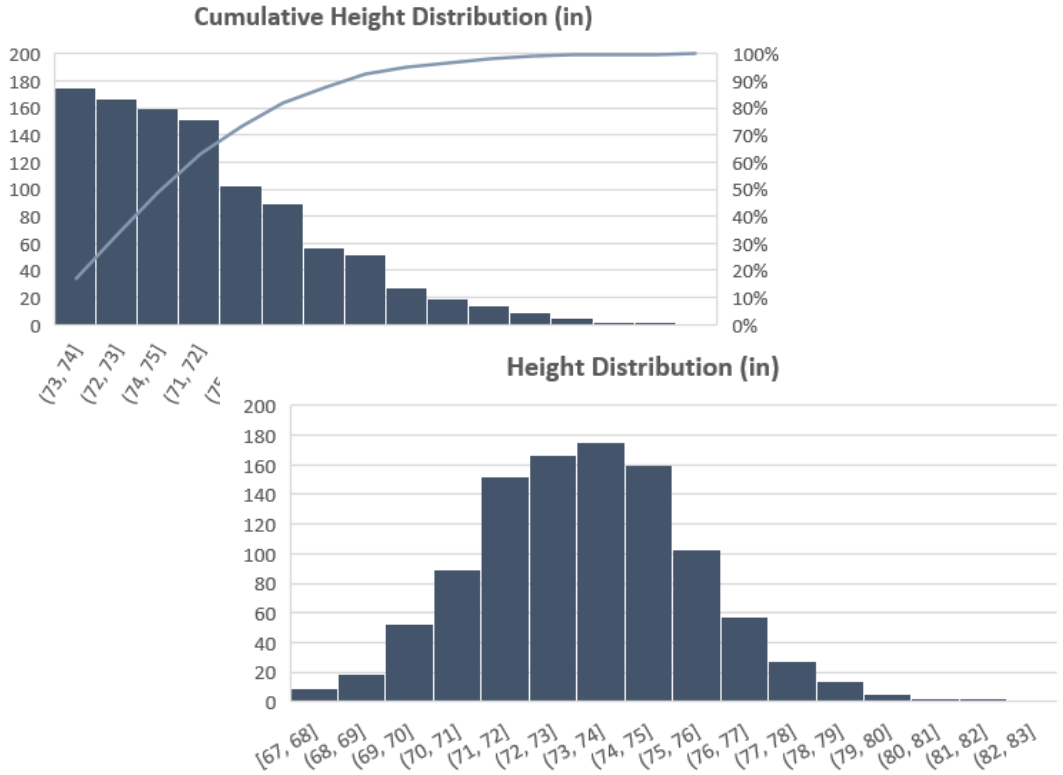
Create **custom formatting rules** to color-code bars/columns based on their values

EXCEL MAVEN

HISTOGRAMS & PARETO CHARTS

COMMONLY USED FOR:

- **Showing the distribution of a continuous data set**



EXAMPLES:

- *Frequency of test scores among students*
- *Distribution of population by age group*
- *Distribution of heights or weights*

PRO TIPS:

- Adjust the bin size to customize the grouping of values
- Use Pareto Charts to show the cumulative impact of each bin, ordered by significance

EXCEL MAVEN

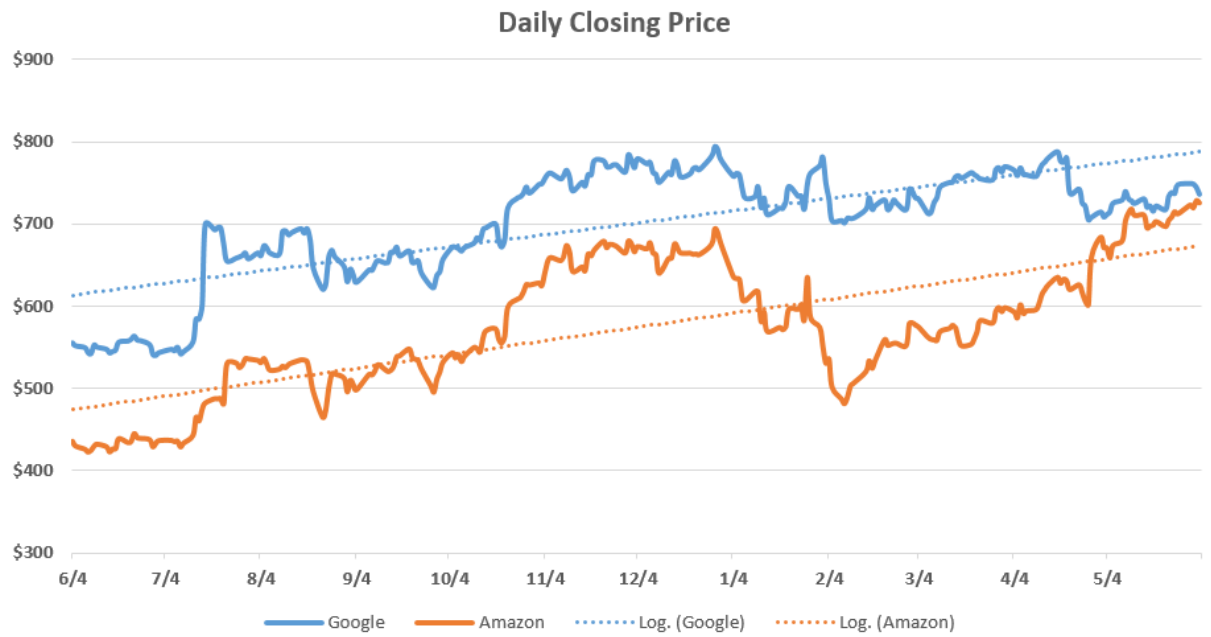
LINE CHARTS

COMMONLY USED FOR:

- **Visualizing trends over time**

EXAMPLES:

- *Stock price by hour*
- *Average temperature by month*
- *Profit by quarter*



PRO TIPS:

Use **linear** or **polynomial trendlines** to visualize patterns or forecast future periods

Combine **line** & **column** charts to trend two variables on different scales

EXCEL MAVEN

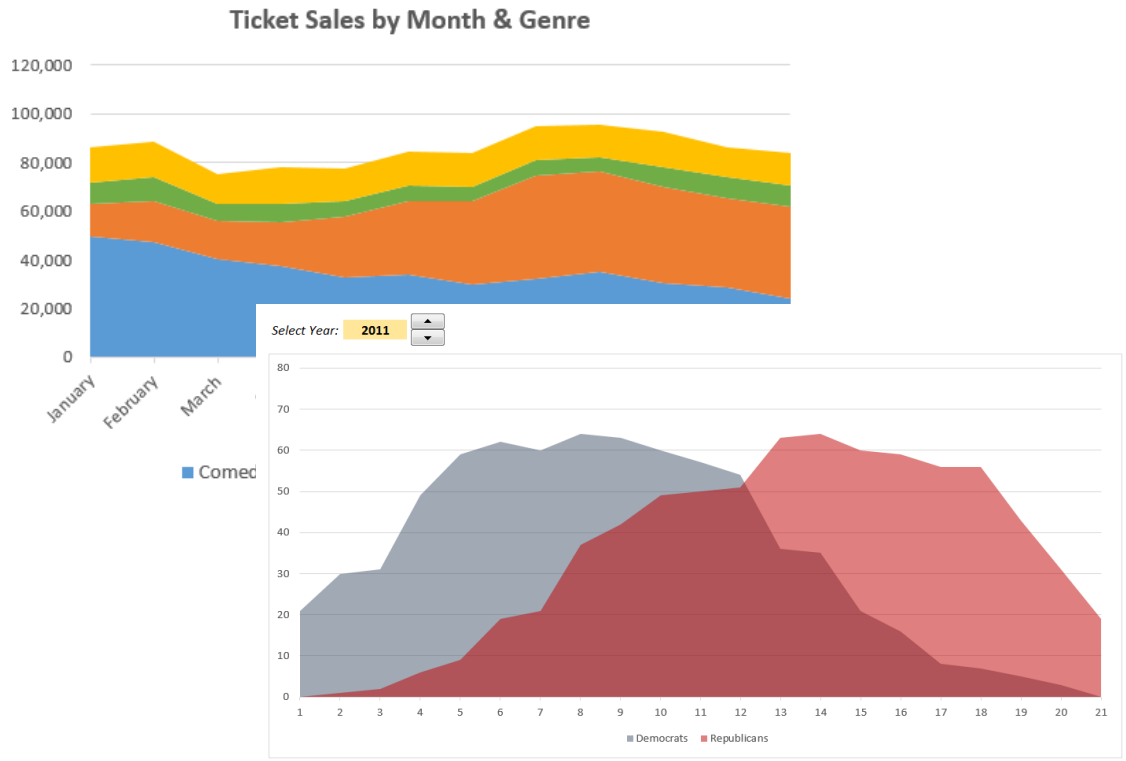
AREA CHARTS

COMMONLY USED FOR:

- **Showing changes in data composition over time**

EXAMPLES:

- *Sales by department, by month*
- *% of total downloads by browser, by week*
- *Population by continent, by decade*



PRO TIPS:

Keep the number of unique categories relatively low **(<6)** to maintain clarity

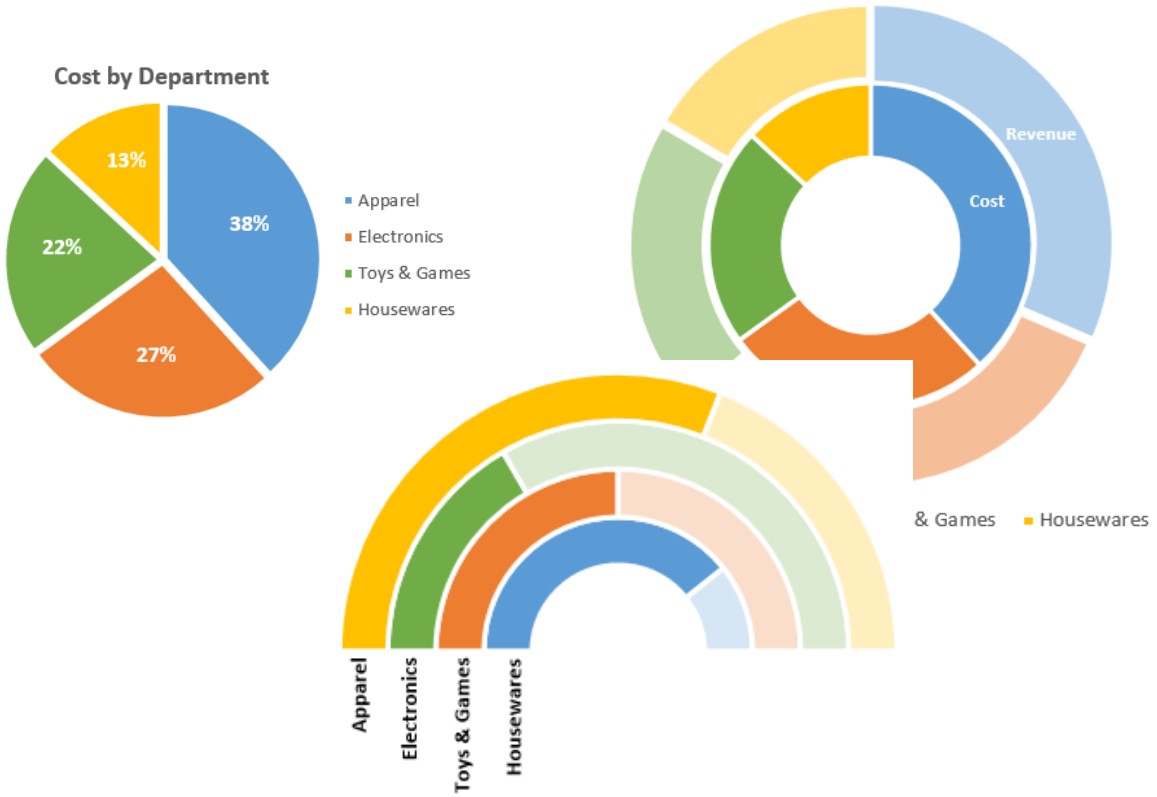
Use **data validation** and **custom formatting** to dynamically highlight specific data series

EXCEL MAVEN

PIE & DONUT CHARTS

COMMONLY USED FOR:

- **Comparing proportions totaling 100%**



EXAMPLES:

- *Percentage of budget spent by department*
- *Proportion of internet users by age range*
- *Breakdown of site traffic by source*

PRO TIPS:

Keep the **number of slices small (<6)** to maximize readability

Use a **donut chart** to visualize more than one series at once, or use transparent segments to create a custom “race track” visualization

EXCEL MAVEN

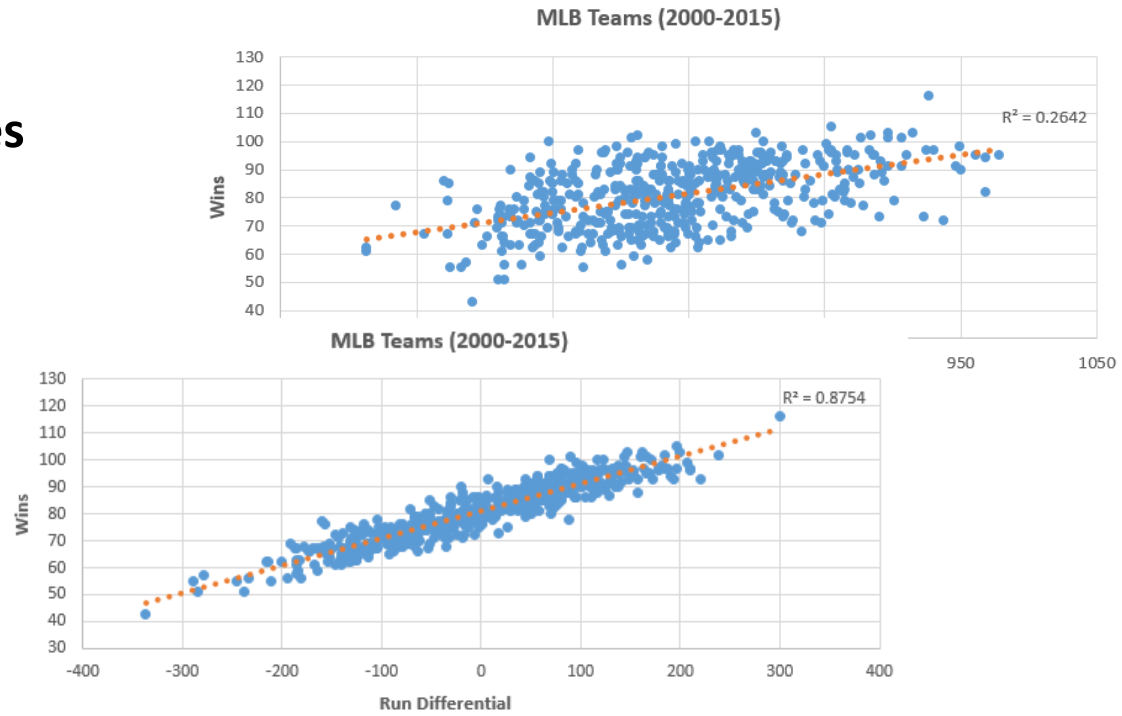
SCATTER PLOTS

COMMONLY USED FOR:

- **Exploring correlations or relationships between series**

EXAMPLES:

- *Number of home runs and salary by player*
- *Ice cream sales and average temperature by day*
- *Hours of television watched by age*



PRO TIPS:

Add a **trendline** or **line of best fit** to quantify the correlation between variables

Remember that **correlation *does not* imply causation**

EXCEL MAVEN

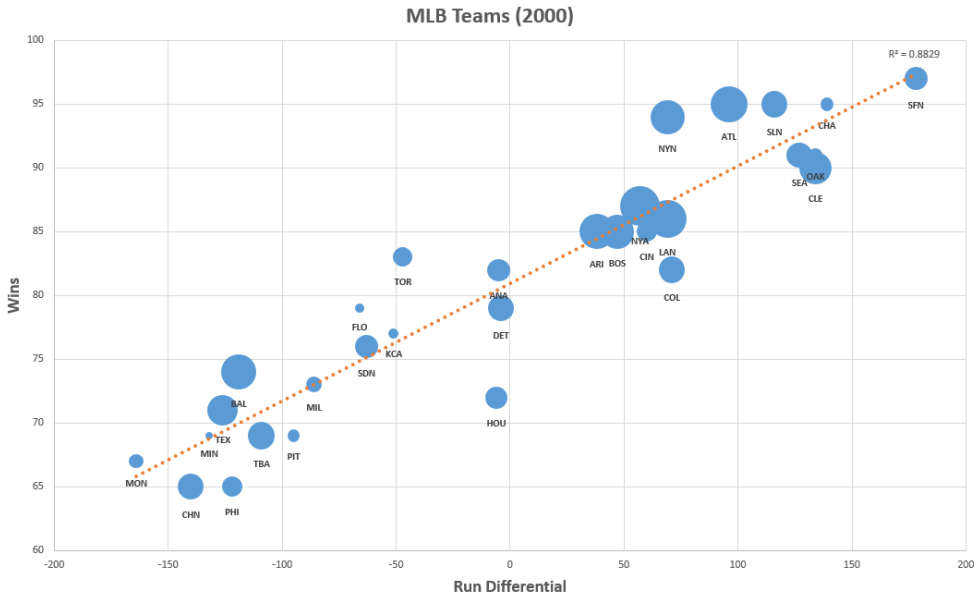
BUBBLE CHARTS

COMMONLY USED FOR:

- **Adding a third dimension (size) to a scatter plot format**

EXAMPLES:

- *Product sales (**X**), Revenue (**Y**), and Market Share (**size**) by Company*
- *Income per Capita (**X**), Life Expectancy (**Y**) and Population (**size**) by Country*



PRO TIPS:

Use **color** as a fourth dimension to differentiate between categories

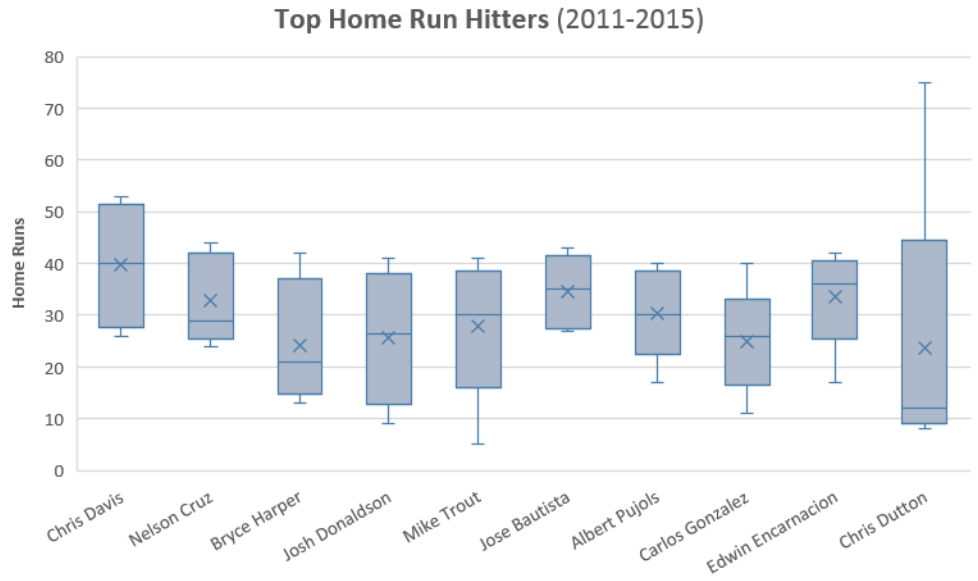
Use **cell formulas** and **form controls** to create a dynamic, animated bubble chart

EXCEL MAVEN

BOX & WHISKER CHARTS

COMMONLY USED FOR:

- **Visualizing statistical characteristics across data series**



EXAMPLES:

- *Comparing historical annual rainfall across cities*
- *Analyzing distributions of values and identifying outliers*
- *Comparing mean and median height/weight by country*

PRO TIPS:

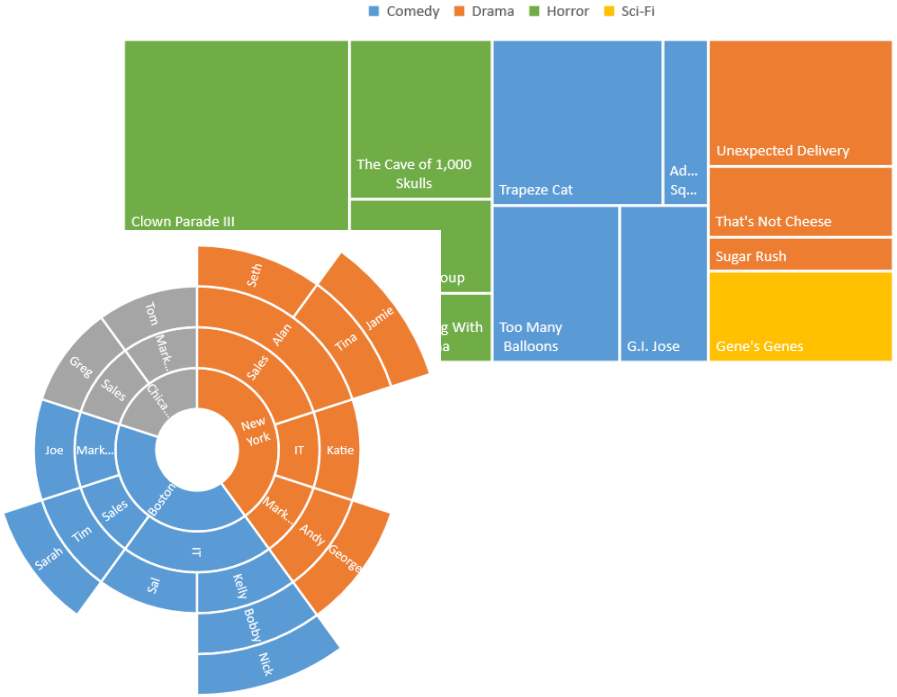
By default, quartiles are calculated by **excluding the median**; this calculation can be adjusted to **include** the median, but may significantly change the result (particularly for smaller data samples)

EXCEL MAVEN

TREE MAPS & SUNBURST CHARTS

COMMONLY USED FOR:

- **Visualizing hierarchical data with natural groups/sub-groups**



EXAMPLES:

- *Revenue by Book Title, Sub-Genre, and Genre*
- *Number of Employees by Department and Office*
- *Population by City, State, and Region*

PRO TIPS:

Use **Tree Maps** when you are only visualizing 1 or 2 hierarchical levels (i.e. topic & sub-topic) or when relative sizes are important, and **Sunburst charts** to visualize the depth of multiple hierarchical levels

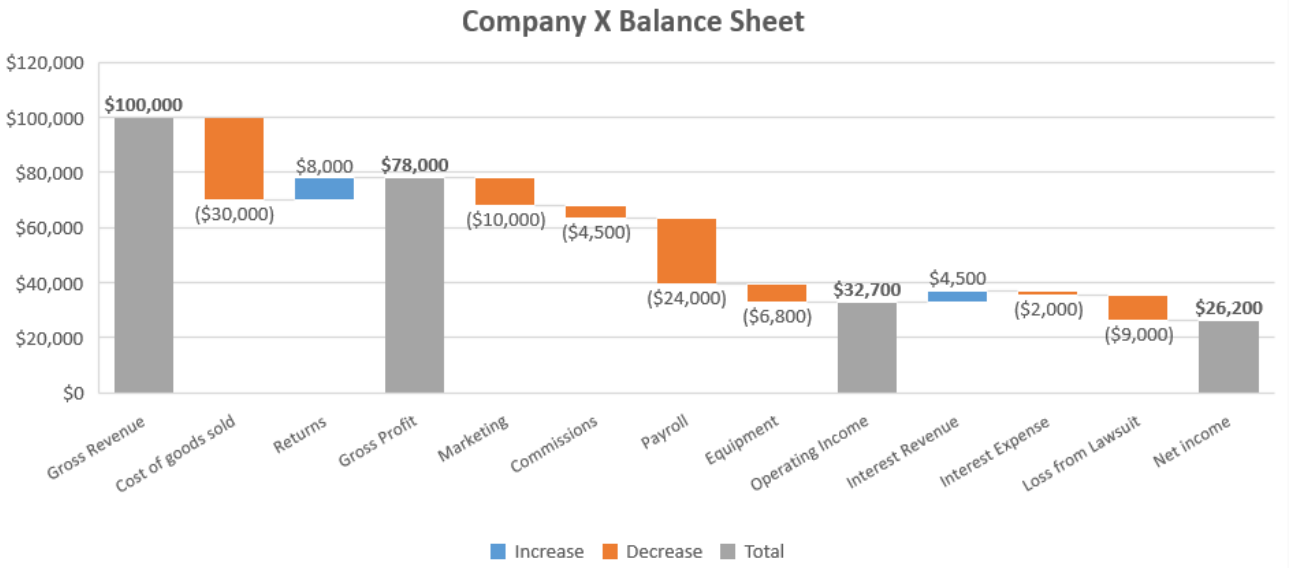
Make sure your raw source data is **grouped** and **sorted** before creating hierarchical charts

EXCEL MAVEN

WATERFALL CHARTS

COMMONLY USED FOR:

- **Showing the net value after a series of positive and negative contributions**



EXAMPLES:

- *Corporate balance sheet analysis*
- *Personal income and spending*

PRO TIPS:

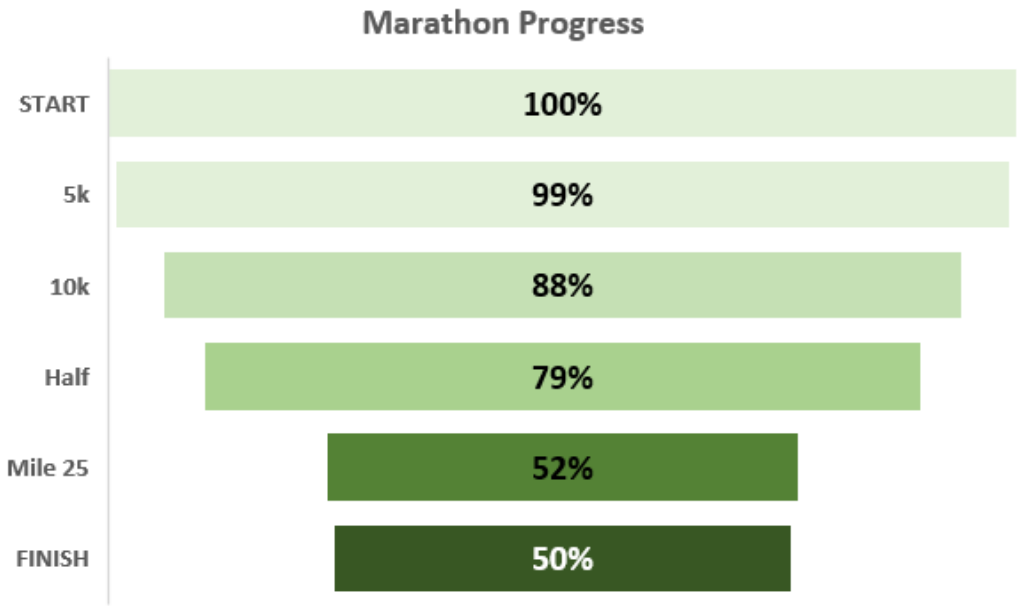
Use **sub-totals** to create "checkpoints" and split up certain types of gains/losses (i.e. **Gross Revenue** - Cost of Goods Sold = **Gross Profit**, Gross Profit - Operating Expenses = **Operating Income**, etc.)

EXCEL MAVEN

FUNNEL CHARTS

COMMONLY USED FOR:

- **Showing progress through the stages of a funnel**



EXAMPLES:

- *Volume of views, clicks, and sales on an ecomm site*
- *Number of runners who reach each checkpoint in a marathon (5k, 10k, half, etc.)*

PRO TIPS:

Use "**percent of total**" calculations to show the % of users (rather than #) at each funnel stage

Customize colors to emphasize progression towards an end goal

EXCEL MAVEN

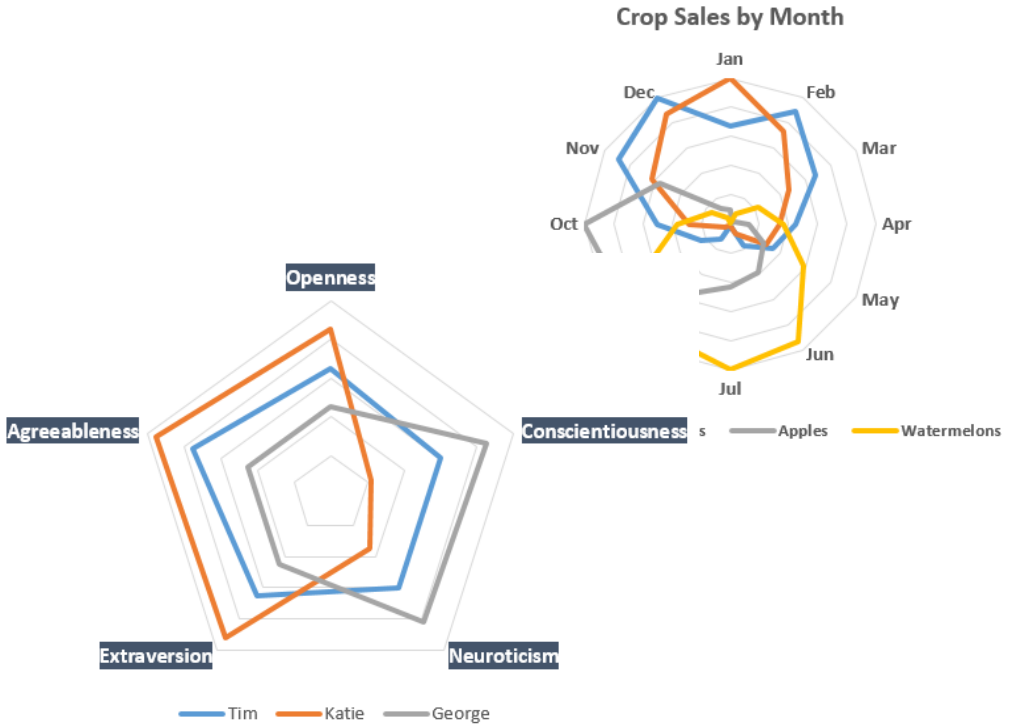
RADAR CHARTS

COMMONLY USED FOR:

- **Plotting three or more quantitative variables on a two-dimensional chart, relative to a central point**

EXAMPLES:

- *Comparing test scores across multiple subjects*
- *Sales of different types of vegetables, by month*
- *Visualizing personality test results across subjects*



PRO TIPS:

Normalize each metric to the same scale (i.e. *0-1*, *1-10*, *1-100*) to improve readability and create more intuitive comparisons across data series

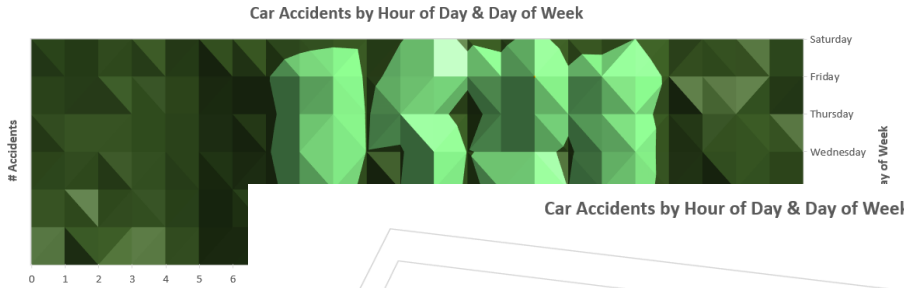
Limit the number of categories or data series to minimize noise and maximize impact

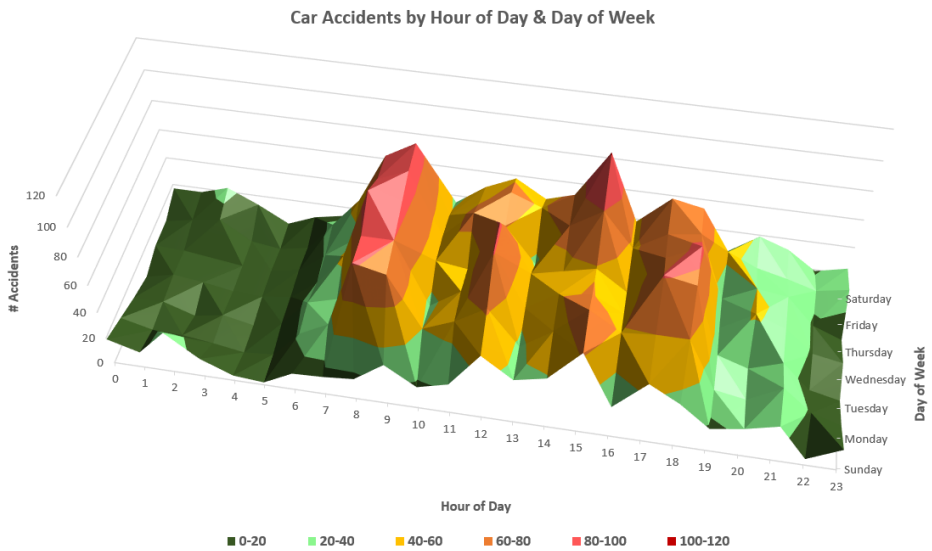
EXCEL MAVEN

SURFACE & CONTOUR CHARTS

COMMONLY USED FOR:

- **Plotting data in three dimensions to find optimum combinations of values**





EXAMPLES:

- *Accident rates by hour of day and day of week*
- *Elevation by latitude and longitude*
- *Cookie deliciousness by oven temp and baking time*

PRO TIPS:

Don’t use surface charts if a simple **heat map** will tell the same story

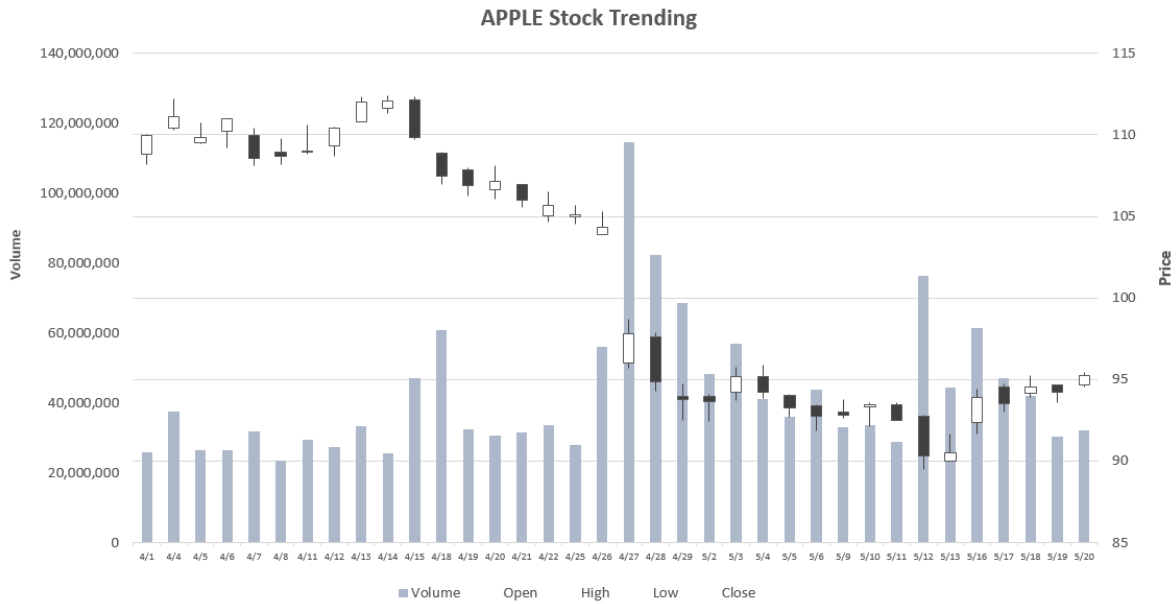
Avoid using **wireframe** chart types when possible, as they can be difficult to interpret

EXCEL MAVEN

STOCK CHARTS

COMMONLY USED FOR:

- **Visualizing stock market data, including volume, high, low, open, and closing prices**



EXAMPLES:

- *Facebook's daily stock performance in 2015*
- *High, low, and closing prices for Google in Q1*
- *Relative performance across multiple stocks*

PRO TIPS:

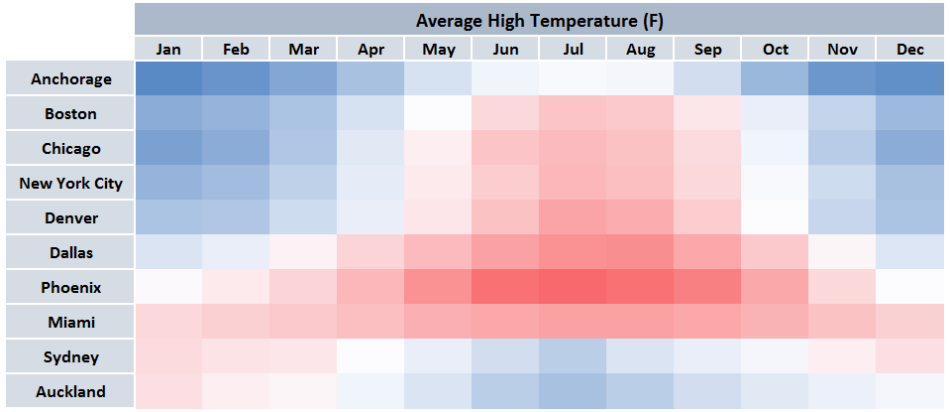
- Manually set **axis minimum/maximum values** to enhance readability
- Switch from a **date** to a **text** axis to eliminate gaps when markets are closed

EXCEL MAVEN

HEAT MAPS

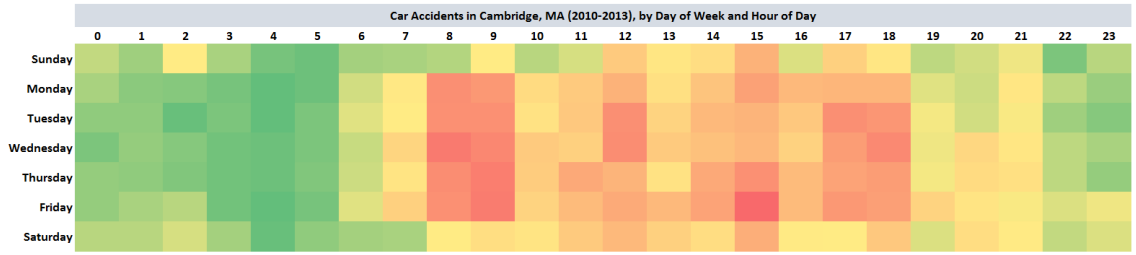
COMMONLY USED FOR:

- **Visualizing trends or relationships using color scales**



EXAMPLES:

- *Accident rates by time of day and day of week*
- *Average temperature by city, by month*
- *Average sentiment by hashtag*



PRO TIPS:

Use intuitive color scales (i.e. **red** to **green**) and apply custom formatting to hide cell values (**;;;**)

Use **data validation** and **cell formulas** to create dynamic heat maps based on user-entered values

EXCEL MAVEN

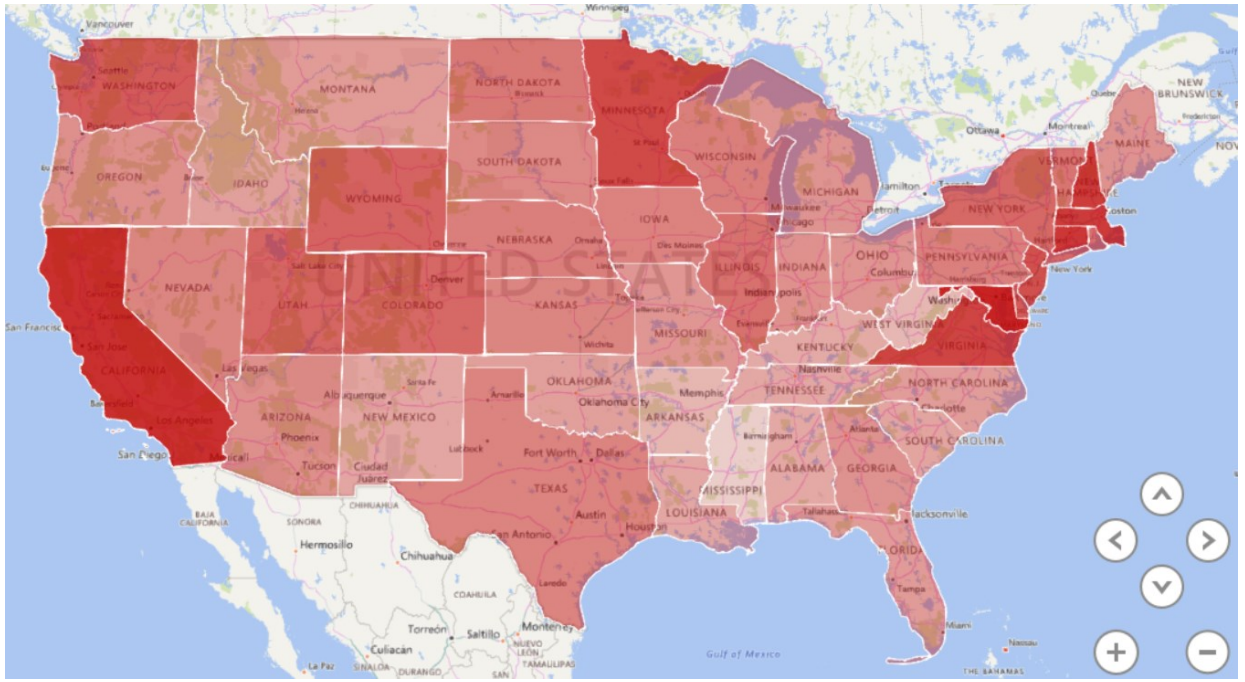
GEOSPATIAL/CHOROPLETH MAP

COMMONLY USED FOR:

- **Visualizing location-based data**

EXAMPLES:

- *Frequency of accidents by street address*
- *Unemployment rate by country*
- *Average rainfall by state*



PRO TIPS:

Use Excel's **Power Map** plug-in to create geo-spatial visualizations and animate changes over time

Utilize attributes like **color** and **size** to visualize multiple attributes at once

EXCEL MAVEN

RESOURCES & NEXT STEPS

- Check out **Excel Analytics – Advanced Formulas & Functions** to master advanced Excel formulas and analytics tools
  - *Stats functions, logical operators, conditional statements, text functions, array formulas, lookup/reference functions, formula-based formatting, and more*
- Head to the following blogs/sites for additional support:
  - ***support.office.com*** *for help with the basic (also check out Office 365)*
  - ***stackoverflow.com*** *for advanced forum support*
  - ***https://sites.google.com/site/e90e50charts/*** *for crazy advanced stuff*
- Rating and reviews are what keeps courses like this alive, so **please** share feedback (for better or for worse!)

EXCEL MAVEN