Visualizing Data

What is Data Visualization?

Data Visualization is the process of using graphs, charts, or images to visualize complex data.

As humans, we are much better at interpreting and extracting meaning from information when it's presented to us in a visual form rather than just raw numbers!

Data vs. Visualization

Average Weather Conditions in San Jose

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Average high in °F	60	64	67	70	75	81	84	84	82	76	67	61
Average low in °F	42	45	47	49	53	56	58	58	57	53	46	42
Av. precipitation in inch	2.99	3.32	2.04	1.06	0.39	0.09	0.00	0.00	0.23	0.78	1.88	2.12

Data source: https://www.usclimatedata.com/climate/san-jose/california/united-states/usca0993

Data vs. Visualization

Average Weather Conditions in San Jose

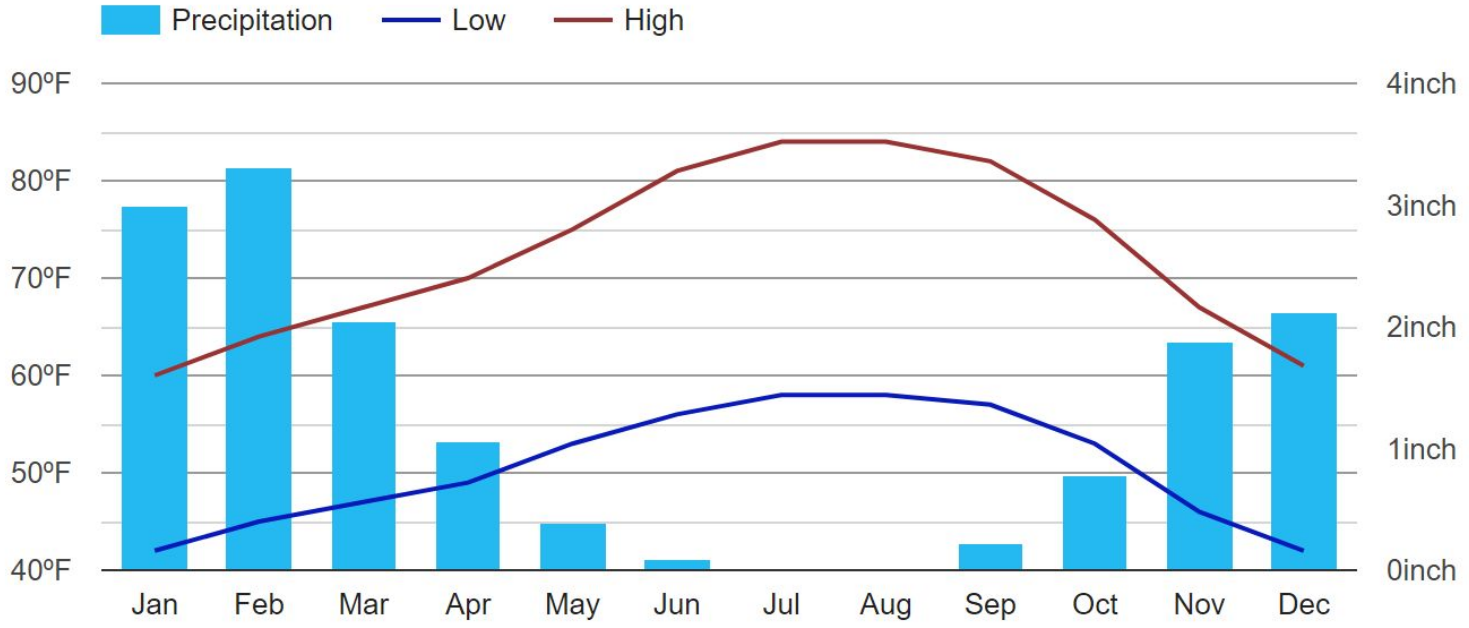
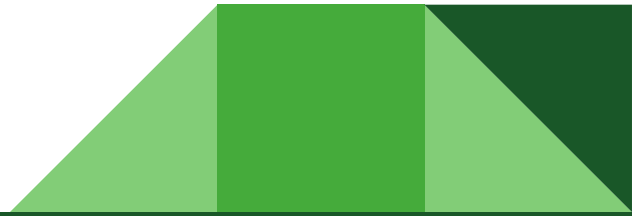


Chart source: https://www.usclimatedata.com/climate/san-jose/california/united-states/usca0993

Data Visualization is used EVERYWHERE

Since data comes from all kinds of different sources, many different fields make use of data visualization to better understand their information!

- Science
- Government
- Sports
- Agriculture
- Music
- Economics
- Medicine and disease
- So many more!

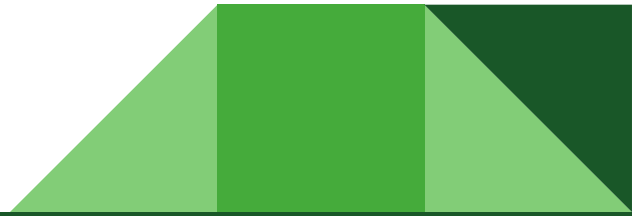


Historical Data Visualization

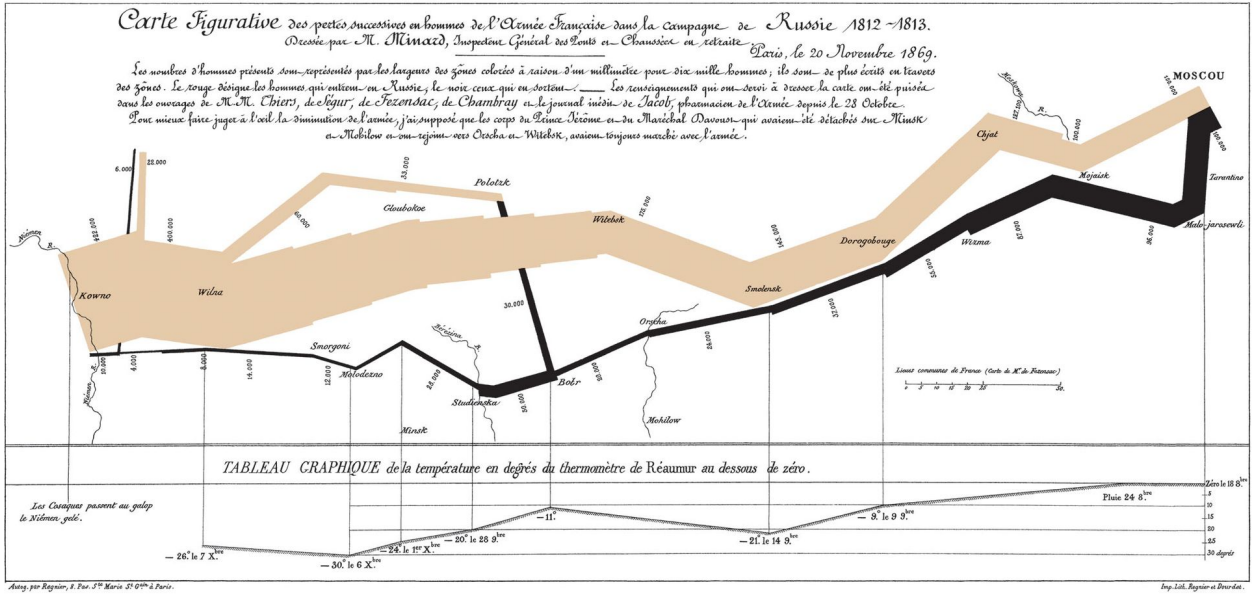
People have been using pictures to understand data for centuries - we don't NEED computers to visualize our data!

Maps are a prime example of data visualization that have existed since at least the 17th century! We are representing information about how the world as a whole looks, to make it easier for us to understand!

The Pie chart was invented in the 19th century, which makes it a good deal older than computers!

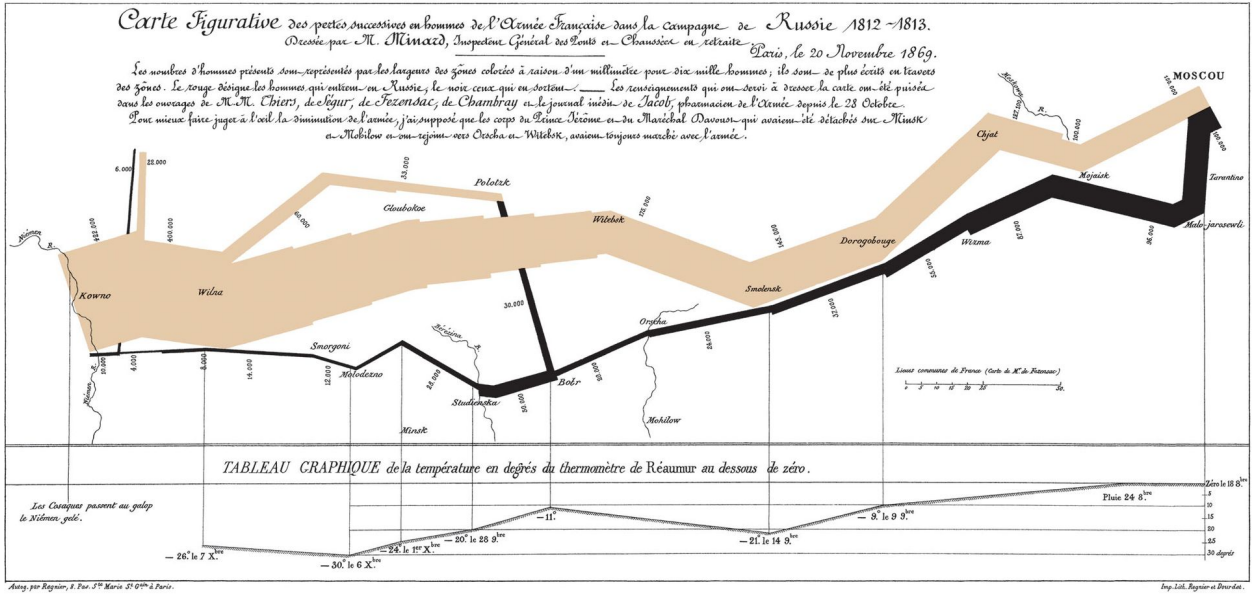


Historic Data Visualization



Charles Joseph Minard was one of the first to develop informational graphics. He made a famous visualization of Napoleon's losses during the Russian campaign of 1812!

Historic Data Visualization



He likely had a table of values, then had to carefully plot each point himself.

He probably needed a good deal of artistic talent to accomplish this feat!

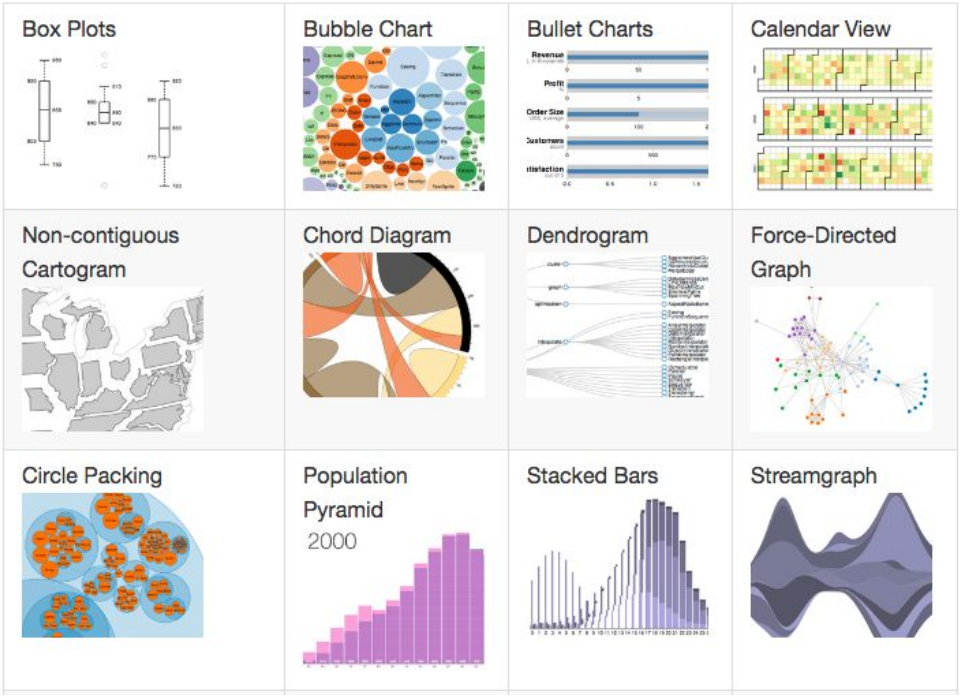
Today's Data Visualization

Fortunately, technology has come a long way since Charles Joseph Minard made his visualization. Computers have made incredible contributions to the field of data visualization!

They allow us to process large amounts of data quickly, update our visualizations to stay accurate if the data changes, and even add interactivity to our visualizations!

Computers also let us be super creative with *how* we choose to visualize our data, allowing us to blend science and art! There are entire communities dedicated to this idea, like r/dataisbeautiful!

Lots of Different Kinds



Check out more types of visualization here:

https://github.com/d3/d3/wiki/Gallery

Example: Table

Tables are great for showing **precise** data.

We can even color-code different cells to represent different things!

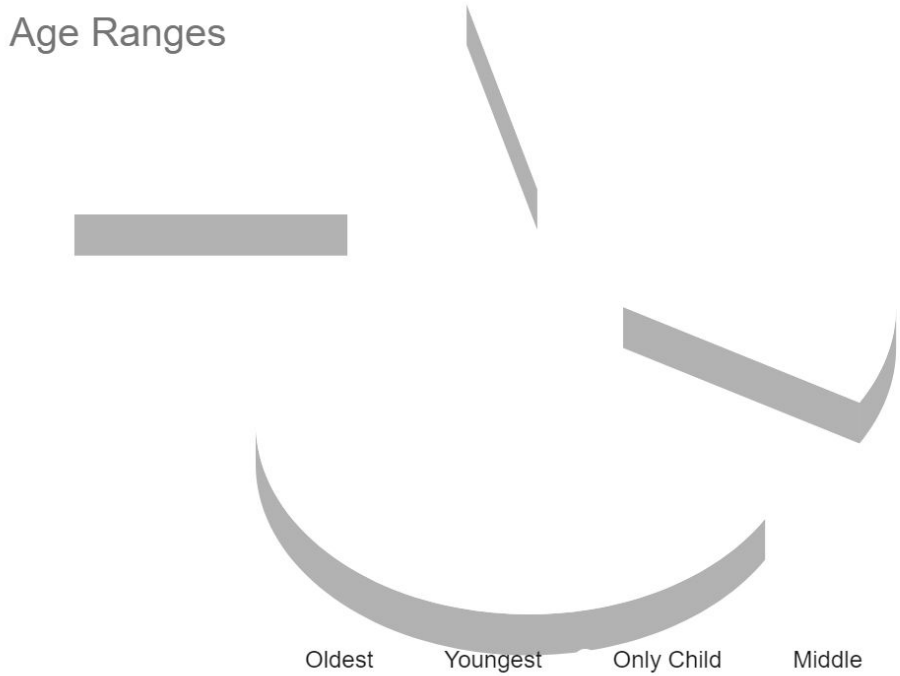
Southbound Service

Zone	Southbound Train No.	102	104	206	208	210	312	314	216	218	220	322	324
1	San Francisco	4:55	5:25	6:06	6:24	6:44	6:56	7:12	7:19	7:24	7:44	7:56	8:12
1	22nd Street	5:00	5:30	6:11	6:29	6:50	7:02	7:18	7:25	7:29	7:50	8:02	8:18
1	Bayshore	5:05	5:35	—	6:35	—	—	—	—	7:35	—	—	—
1	So. San Francisco	5:11	5:41	—	6:41	—	—	—	—	7:41	—	—	—
1	San Bruno	5:15	5:45	—	6:44	—	—	—	7:37	7:44	—	—	—
2	Millbrae	5:19	5:49	6:24	6:49	7:02	7:17	7:32	—	7:49	8:02	8:17	8:32
2	Burlingame	5:23	5:53	6:28	6:53		—	—	7:44	7:53	—	—	—
2	San Mateo	5:28	5:58	6:32	6:56	7:09	—	—	7:48	7:56	8:09	—	—
2	Hayward Park	5:31	6:01	—	7:00	—	—	—	—	8:00	—	—	—
2	Hillsdale	5:34	6:04	6:36	7:04	—	—	7:42	7:52	8:04	—	—	8:42
2	Belmont	5:37	6:07		7:07	—	—	—	—	8:07	—	—	—
2	San Carlos	5:40	6:10	6:40	7:11	7:16	—	—	7:58	8:11	8:16	—	—
2	Redwood City	5:45	6:15	6:45	7:15	7:22	7:32	—	—	8:15	8:22	8:32	—

Source: Caltrain http://www.caltrain.com/schedules/weekdaytimetable.html

Example: Pie Chart

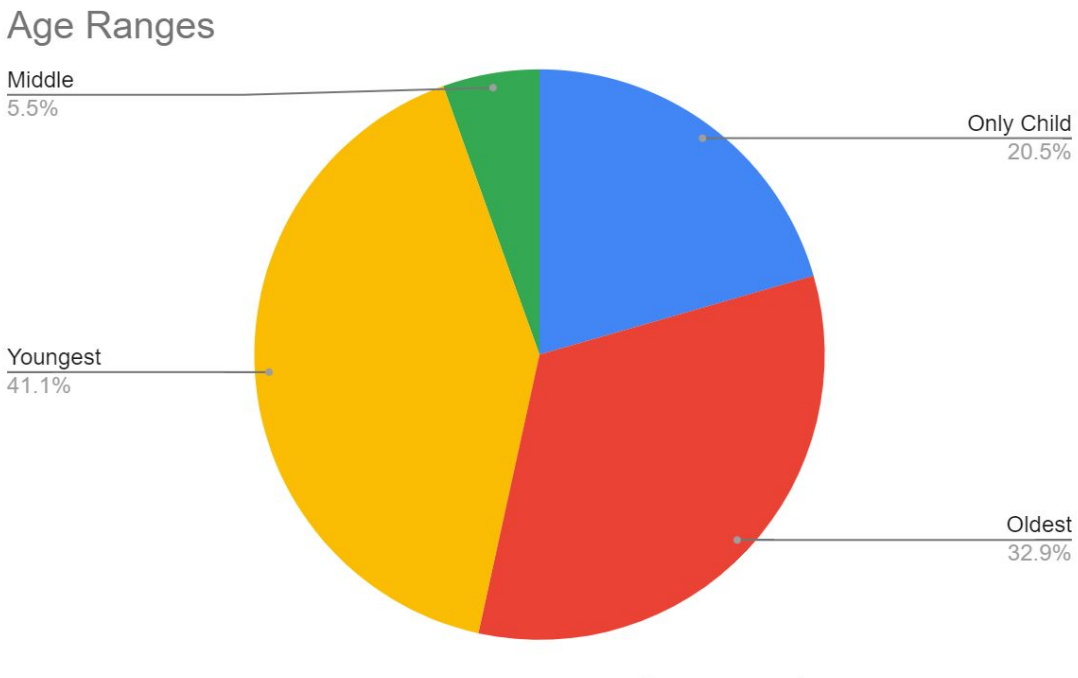
Pie Charts are great for showing **percentages** of the whole!



Example: Pie Chart

Pie Charts are great for showing **percentages** of the whole!

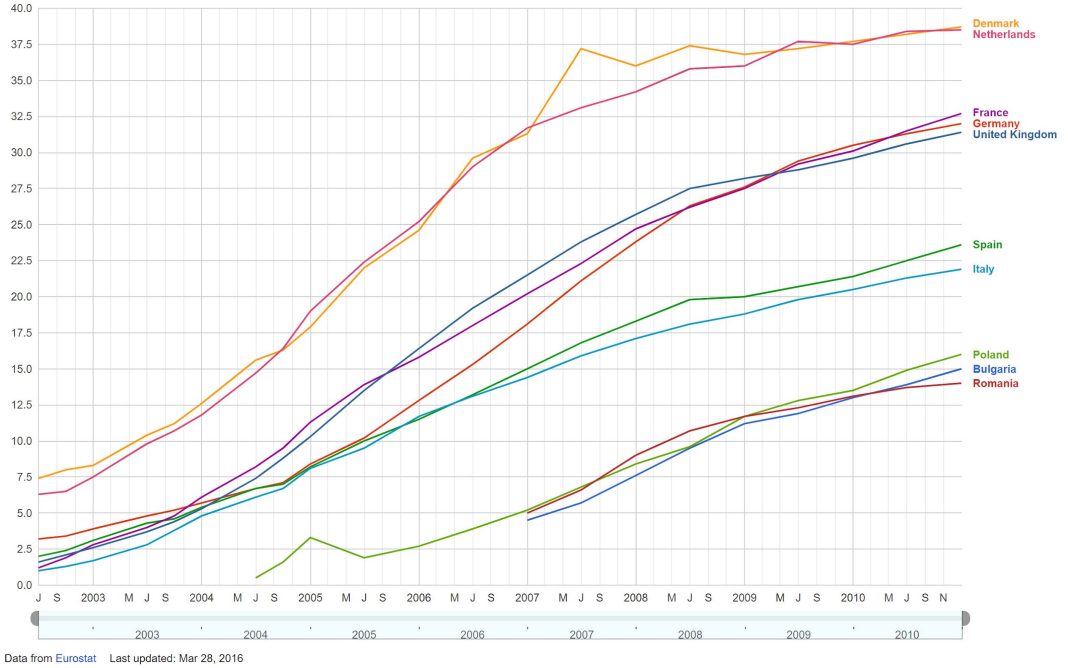
What percentage of people in our class are the **youngest** in their family?



Example: Line Chart

Line charts are good at displaying changes in a system over time.

Has access to high speed internet gotten better or worse for these countries over the last 10 years?

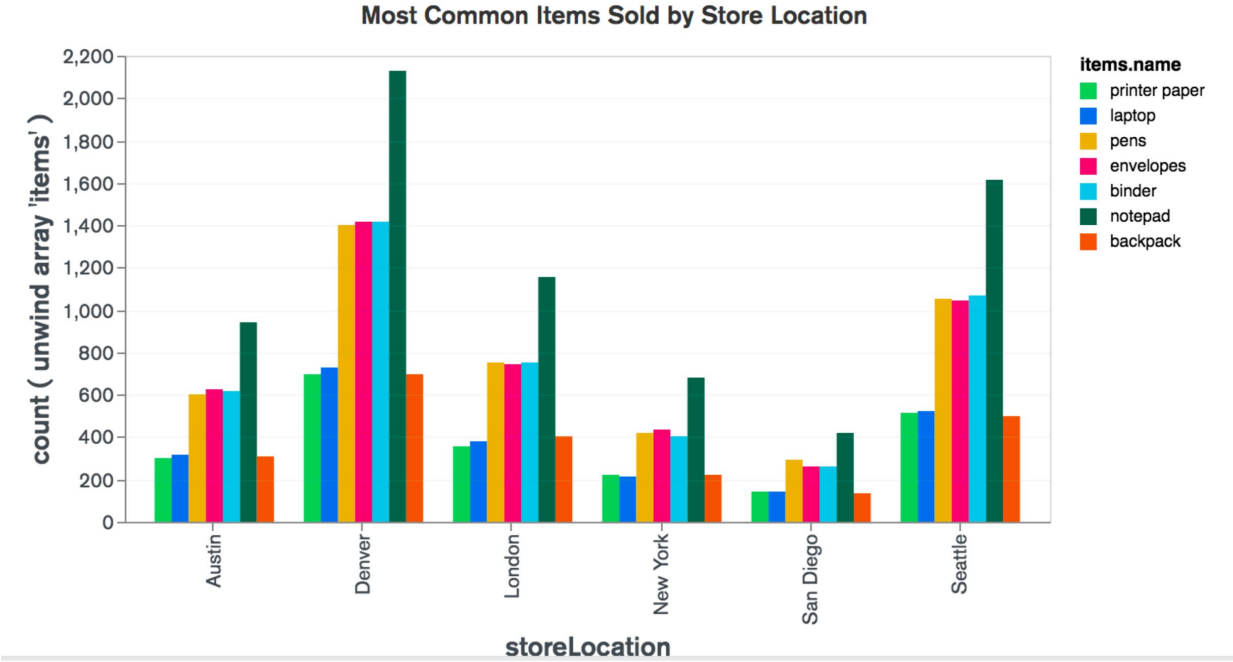


Source: Google Public Data explorer
https://www.google.com/publicdata/explore?ds=l6t8doc0rch3u_&ctype=l&strail=false&nselm=h&met_y=i132&scale_y=lin&ind_y=false&rdim=country_group&idim=country:dk:nl:bg:ro:fr:de:uk:it:es:pl&tstart=1025481600000&tunit=M&tlen=90&iconSize=0.5&uniSize=0.035

Example: Bar Chart

Bar charts are useful when you want to compare different categories against one another.

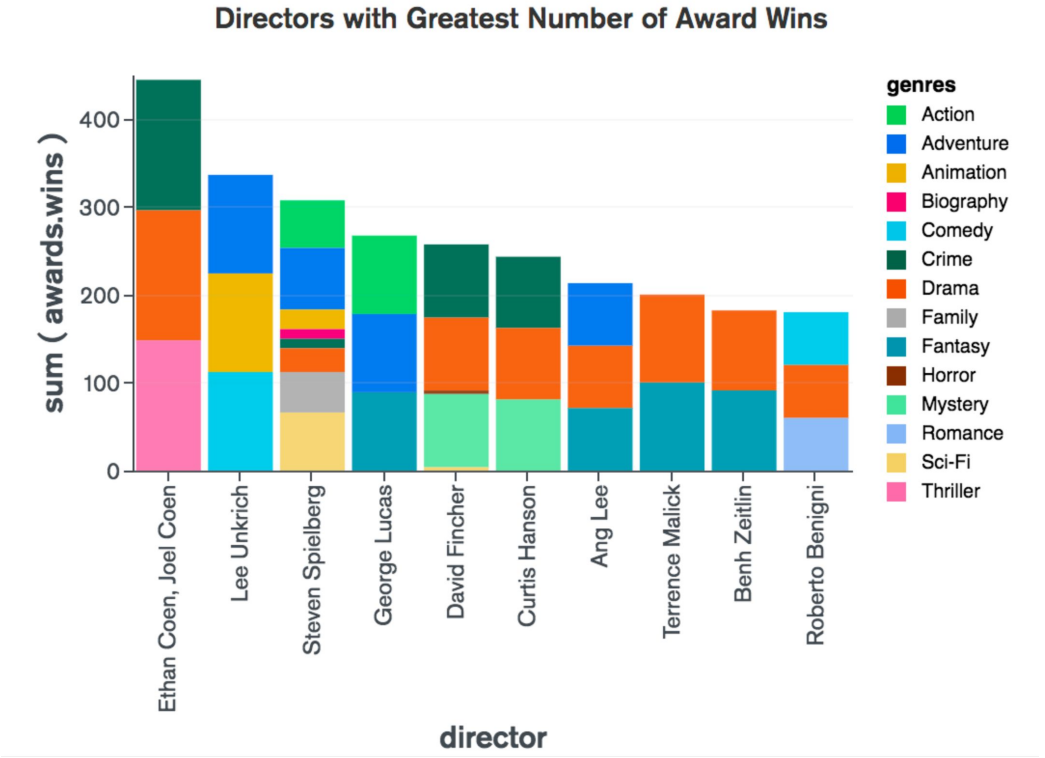
What's the best selling item in Denver? What about in London?



Source: https://docs.mongodb.com/charts/master/chart-type-reference/column-bar-chart

Example: Stacked Bar Chart

We can even extend our bar charts by stacking the bars on top of one another!

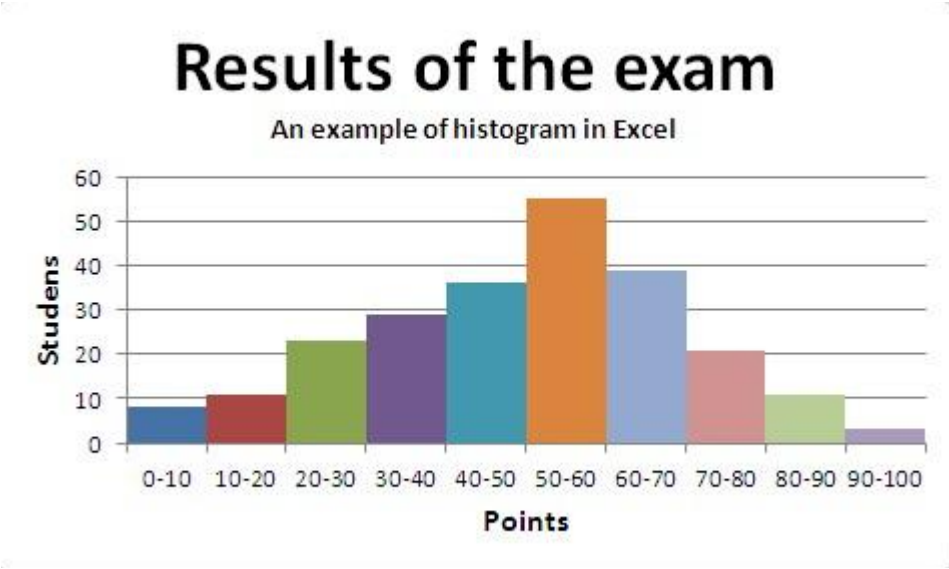


Source: https://docs.mongodb.com/charts/master/chart-type-reference/column-bar-chart

Example: Histogram

Histograms, while similar to Bar Charts, are distinct from them!

Histograms are useful for showing the **frequency** of events.



Source: http://best-excel-tutorial.com/56-charts/76-histogram

Example: Interactive Data

One of the best ways to engage someone in the information you're sharing is by making it **interactive**!

Here's two examples of interactive data - let's check them out!

https://public.tableau.com/en-us/gallery/cities-around-world?gallery=votd

https://public.tableau.com/en-us/gallery/fifa-19-top-ranked-players?gallery=votd

You can check out more interactive visualizations here:
https://public.tableau.com/en-us/gallery/

Creativity in Visualization

Often, we can be creative in how we display our data!

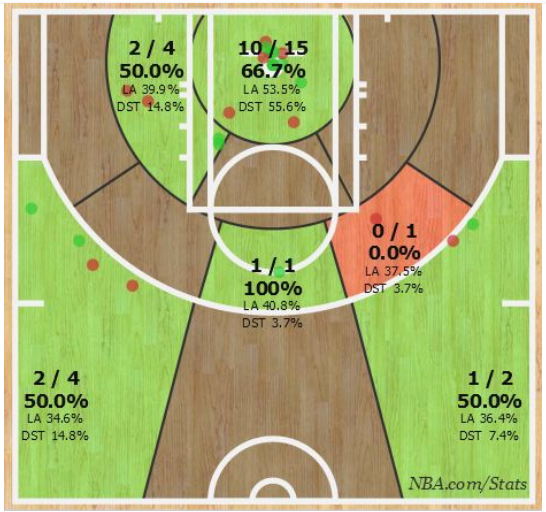
Player	MIN	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	OREB	DREB	REB	AST	TOV	STL	BLK	PF	PTS	+/-
LeBron James - F	42:35	16	27	59.3	3	6	50.0	6	8	75.0	2	6	8	11	1	4	3	3	41	26

By itself, this row of data doesn't show a whole lot of information.

Creativity in Visualization

Often, we can be creative in how we display our data!

Player	MIN	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	OREB	DREB	REB	AST	TOV	STL	BLK	PF	PTS	+/-
LeBron James - F	42:35	16	27	59.3	3	6	50.0	6	8	75.0	2	6	8	11	1	4	3	3	41	26



If we visualize the data on a basketball court, though, we can get a much better idea of the information it's trying to convey!

How are these visualizations used?

Public data sets and data visualization tools allow widespread access to knowledge! Websites like Gapminder, Tableau, and the Google Public Data Explorer surface all kinds of interesting information that's publically available!

These tools enable anyone who wants to draw conclusions about the world around them to do so!