



# Data



# What is Data?

- **Data** = the collected observations we have about something.
- Data can be **continuous**:  
*"What is the stock price?"*
- or **categorical**:  
*"What car has the best repair history?"*



# Why Data Matters

- Helps us **understand things as they are:**

*"What relationships if any exist between two events?"*

*"Do people who eat an apple a day enjoy fewer doctor's visits than those who don't?"*



# Why Data Matters

- Helps us **predict future behavior** to guide business decisions:

*"Based on a user's click history which ad is more likely to bring them to our site?"*



# Visualizing Data

- Compare a **table**:

## Flights

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	year	month	passengers	year	month	passengers	year	month	passengers	year	month	passengers			
2	1950	January	115	1952	July	230	1955	January	242	1957	July	465			
3	1950	February	126	1952	August	242	1955	February	233	1957	August	467			
4	1950	March	141	1952	September	209	1955	March	267	1957	September	404			
5	1950	April	135	1952	October	191	1955	April	269	1957	October	347			
6	1950	May	125	1952	November	172	1955	May	270	1957	November	305			
7	1950	June	149	1952	December	194	1955	June	315	1957	December	336			
8	1950	July	170	1953	January	196	1955	July	364	1958	January	340			
9	1950	August	170	1953	February	196	1955	August	347	1958	February	318			
10	1950	September	158	1953	March	236	1955	September	312	1958	March	362			
11	1950	October	133	1953	April	235	1955	October	274	1958	April	348			
12	1950	November	114	1953	May	229	1955	November	237	1958	May	363			
13	1950	December	140	1953	June	243	1955	December	278	1958	June	435			
14	1951	January	145	1953	July	264	1956	January	284	1958	July	491			
15	1951	February	150	1953	August	272	1956	February	277	1958	August	505			
16	1951	March	178	1953	September	237	1956	March	317	1958	September	404			
17	1951	April	163	1953	October	211	1956	April	313	1958	October	359			
18	1951	May	177	1953	November	180	1956	May	219	1958	November	210			

Not much  
can be  
gained by  
reading it.

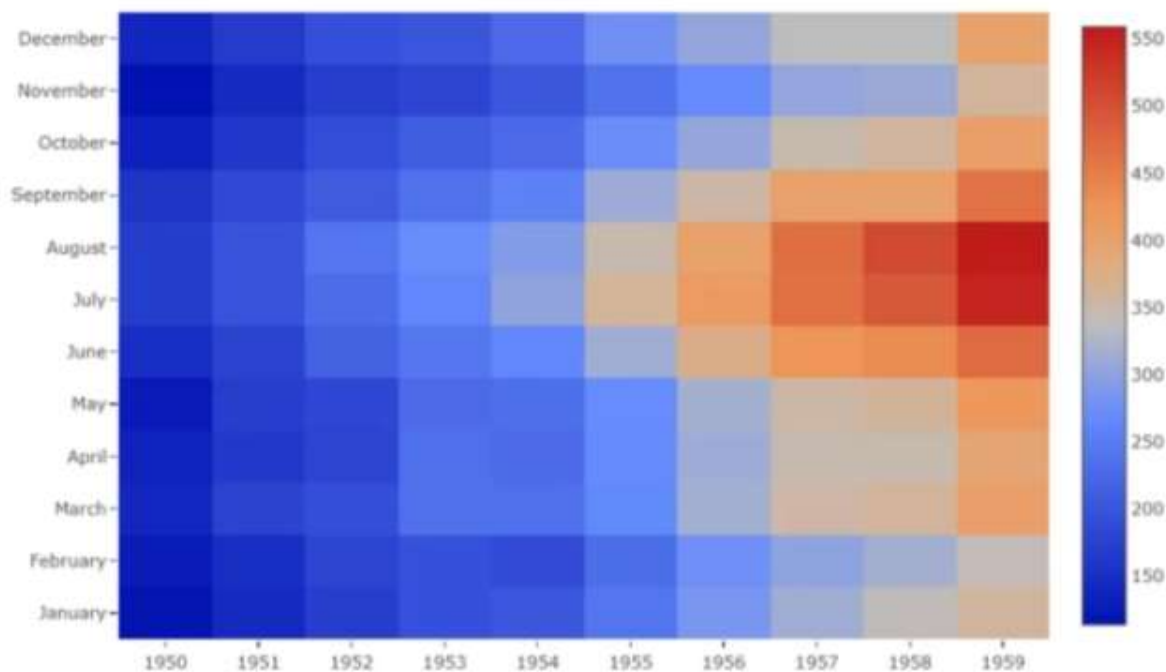




# Visualizing Data

- to a **graph**:

Flights



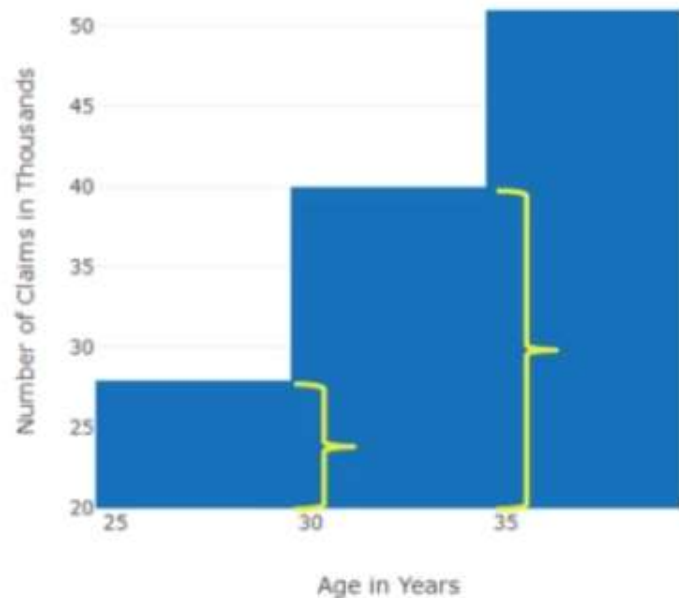
The graph uncovers two distinct trends - an increase in passengers flying over the years and a greater number of passengers flying in the summer months.



# Analyze Visualizations Critically!

- Graphs can be misleading:

Insurance Claims by Age



Insurance Claims by Age

