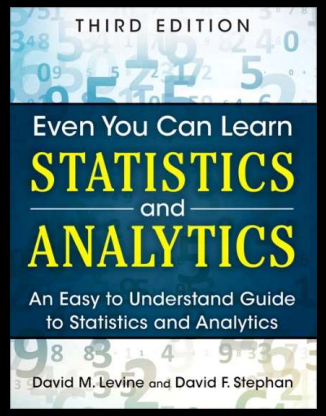
IMGD 2905

Presenting Data

Chapter 2

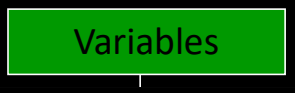


Outline

- Types of Charts (next)
- Game Analytics Examples
- Guidelines for Charts

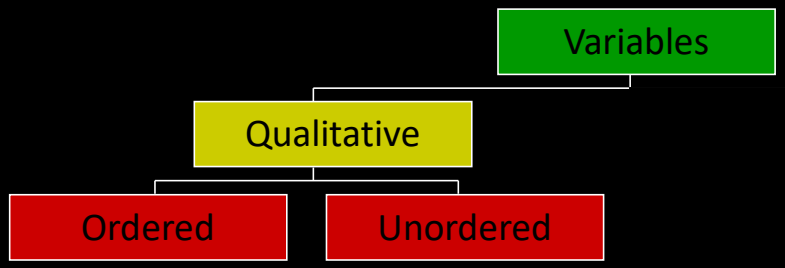
“Right” Chart Depends on Variable Type

- Variable – characteristic of individuals in population analyzing



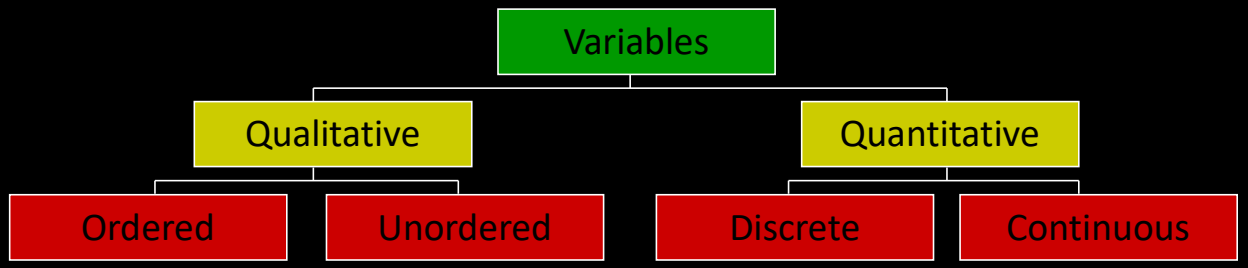
"Right" Chart Depends on Variable Type

- Qualitative (Categorical) variables
  - Can have states or subclasses
    - e.g., position: [striker, goalie, midfield]
  - Can be ordered or unordered
    - e.g., bronze, silver, gold → ordered
    - e.g., support, warrior, specialist → unordered



"Right" Chart Depends on Variable Type

- Quantitative (Numeric) variables
  - Numeric levels
  - Discrete or continuous
    - e.g., goals in season, speed in meters
    - e.g., takedowns, win percentage



Tables

- Generally, independent variable in left column and dependent variables next

		Do you live on campus?	
		Off-campus	On-campus
Class rank	Freshman	37	100
	Sophomore	42	48
	Junior	90	8
	Senior	62	1
Total		231	157

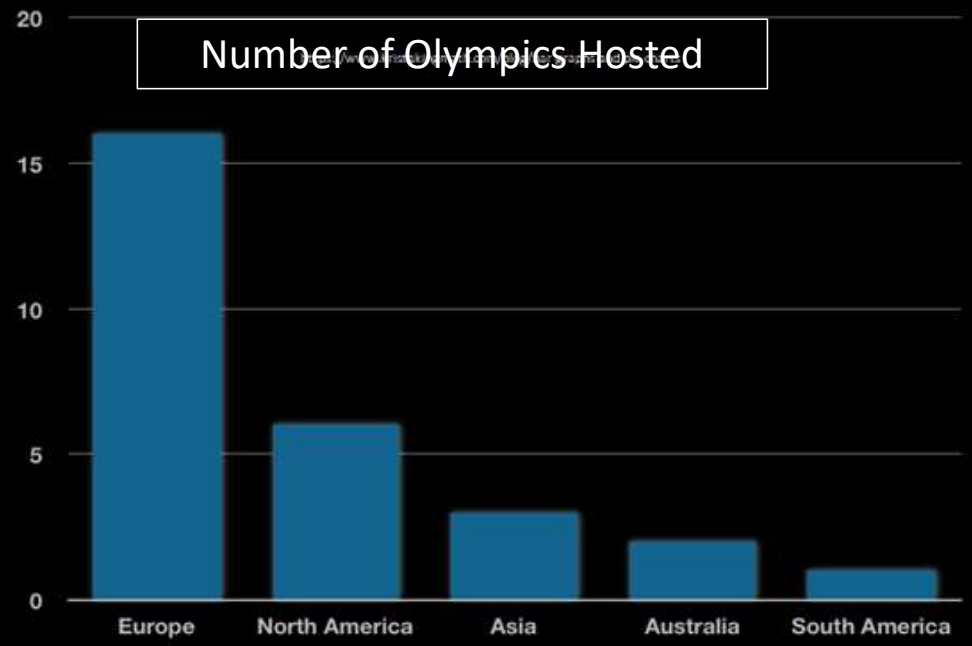
Table 1. Number of student on campus and off by year

- Number & caption
- Units labeled
- Few/no lines
  - Use **bold**

Make sure to consider *message*.
Often much clearer in chart!

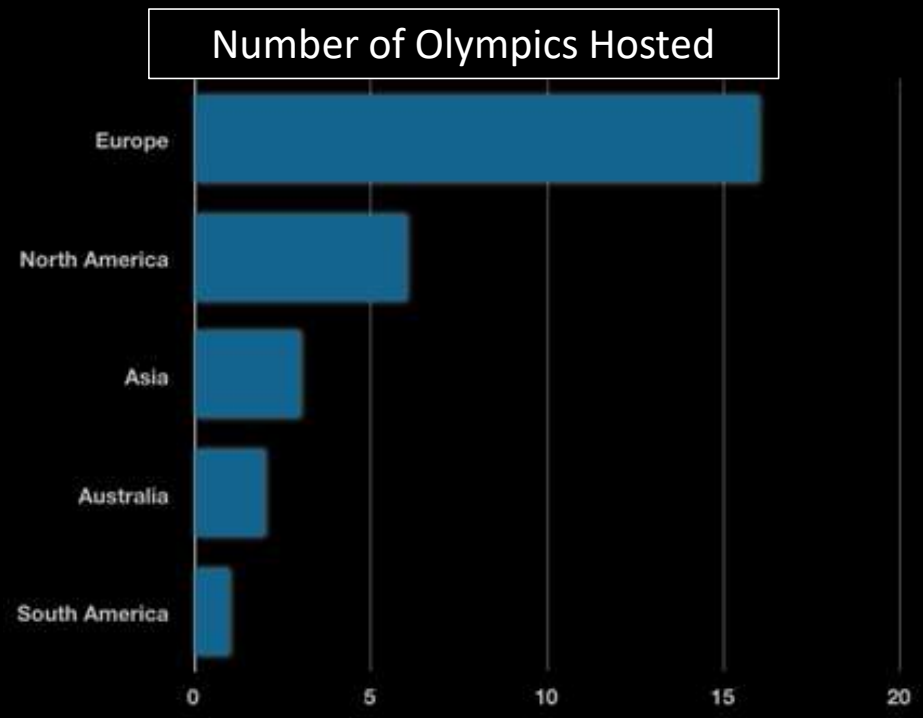
Categorical: Bar Chart (1 of 3)

- Chart containing rectangles (“bars”) where length represents count, amount, or percent (aka “column chart”)
- Better than table for comparing numbers



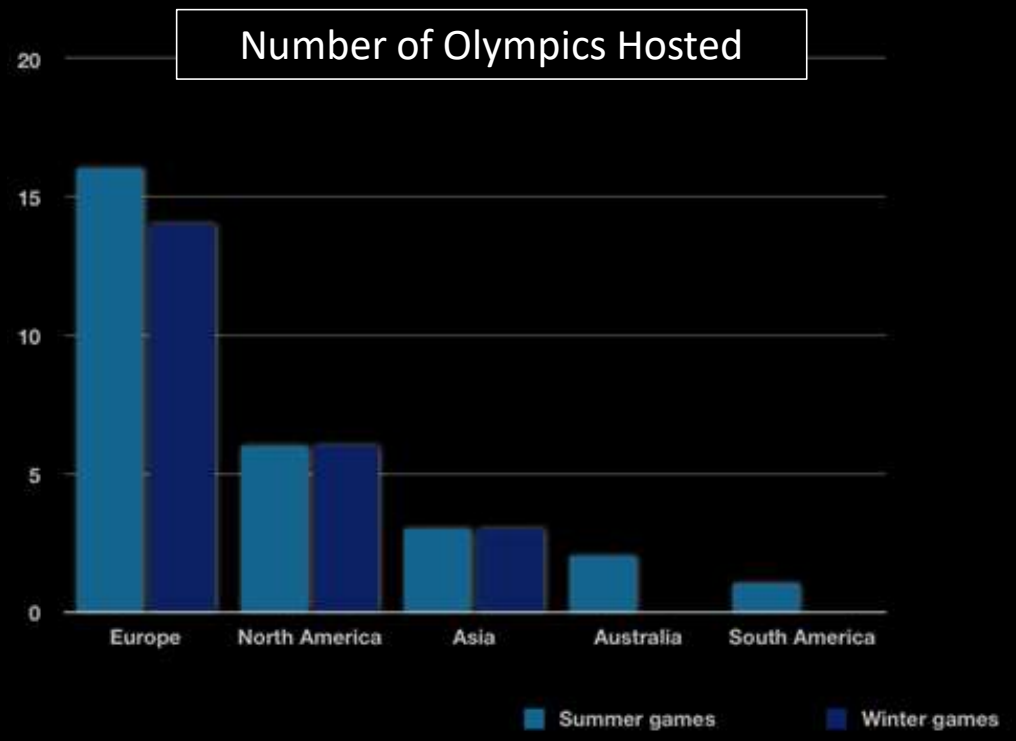
Categorical: Bar Chart (2 of 3)

- Horizontal (good if many observations)



Categorical: Bar Chart (3 of 3)

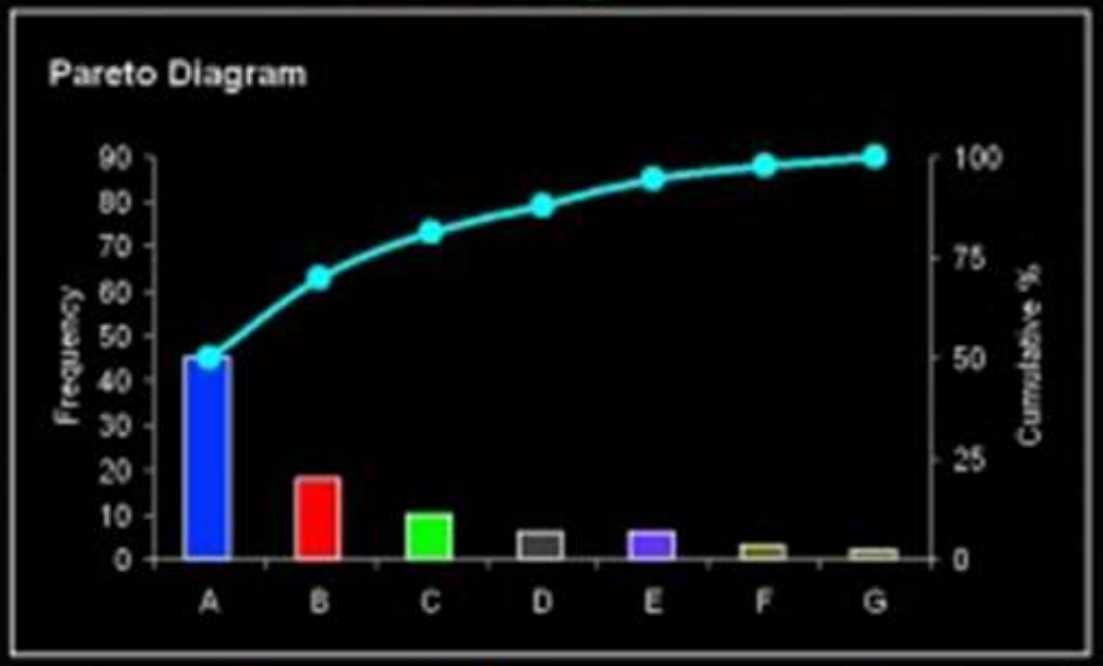
- Compare per observations



Categorical: Pareto Chart

- Bar chart, arranged most to least frequent
- Line showing cumulative percent
- Helps identify most common, and by how much

Demo: imgdpops.xlsx



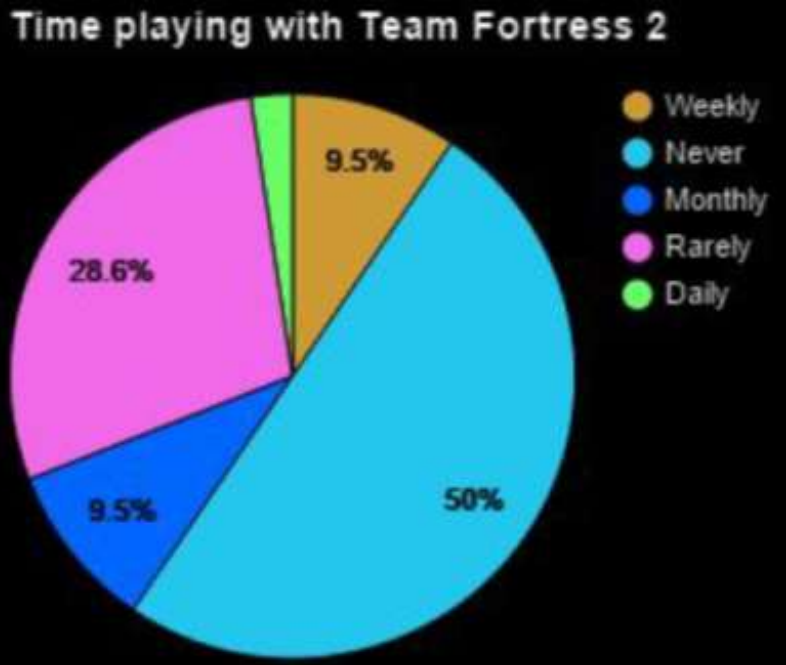
https://usercontent2.hubstatic.com/3767965_f520.jpg

Sort by column D (Data -> Sort high to low)
New column E for percent [=D2/SUM(D$2:D$11)]
Note: $ “locks” value in (e.g., D$2 versus D2)
New column F for running [=SUM(E$2:E2)]
Select B, D and F. Insert “combo chart”

Categorical: Pie Chart

- Wedge-shaped areas ("pie slices") – represent count, amount or percent of each category from whole
- Compare relative amounts at a glance
- Best if few slices since quantifying "size" of pie difficult
- Comparing pies also difficult

Demo: imgdpops.xlsx

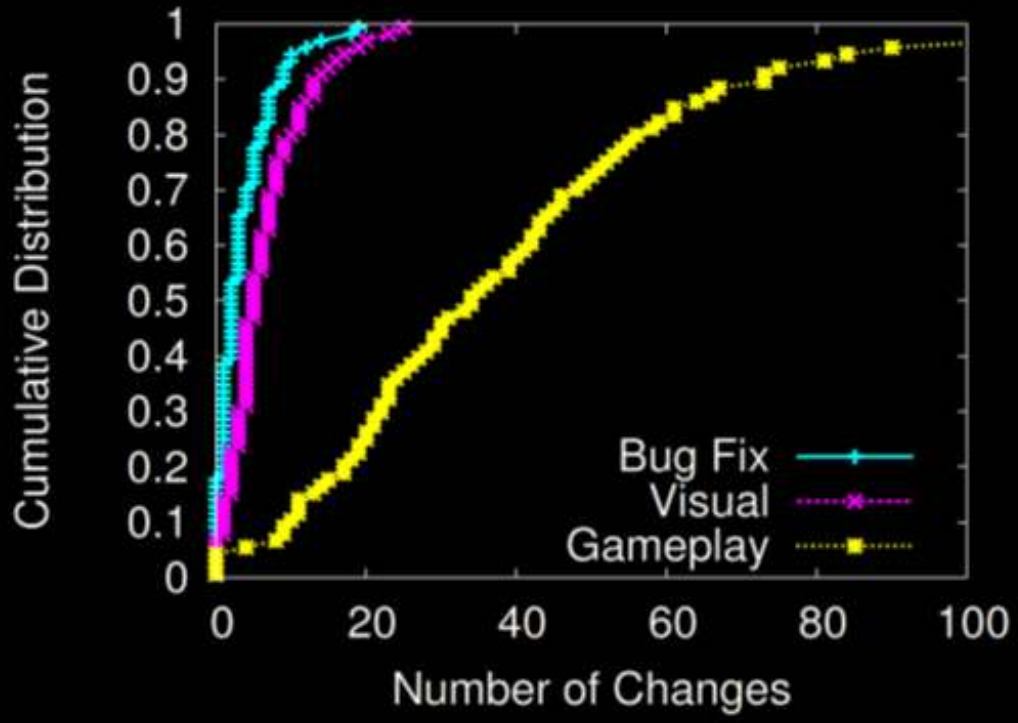


"The Effects of Latency and Jitter on a First Person Shooter: Team Fortress 2"

http://www.cs.wpi.edu/~claypool/iqp/tf2/

Cumulative Distribution GA

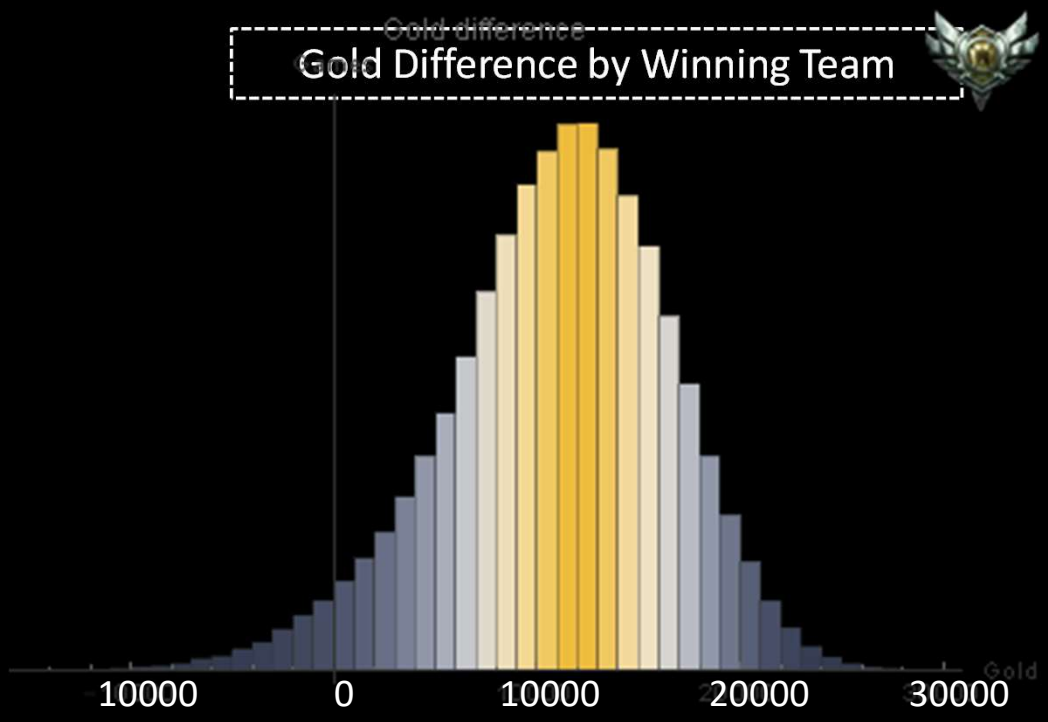
- Cumulative amount of data with value or less
- Easy to see min, max, median
- Compare shapes of distributions



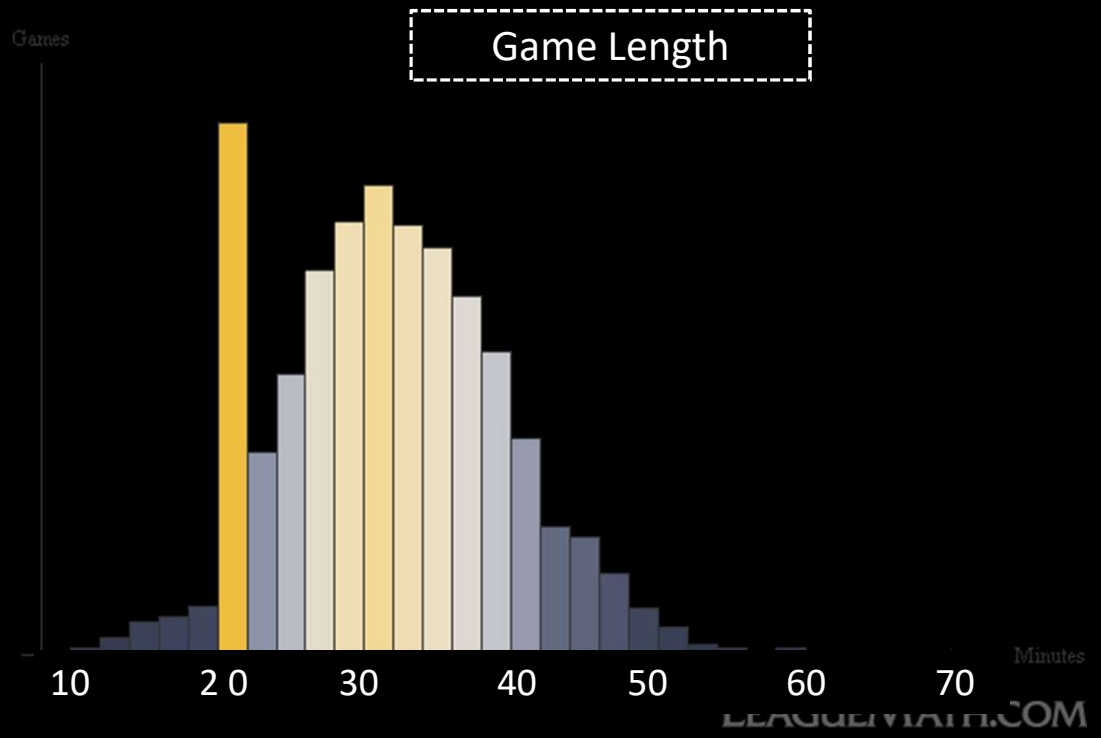
Demo: lol-patches.xlsx

Sort low to high
New column S for percent [=ROW()/164]
Select column → paste down all
Select both column R and S
Insert → Scatter plot with lines

Histogram (1 of 3)

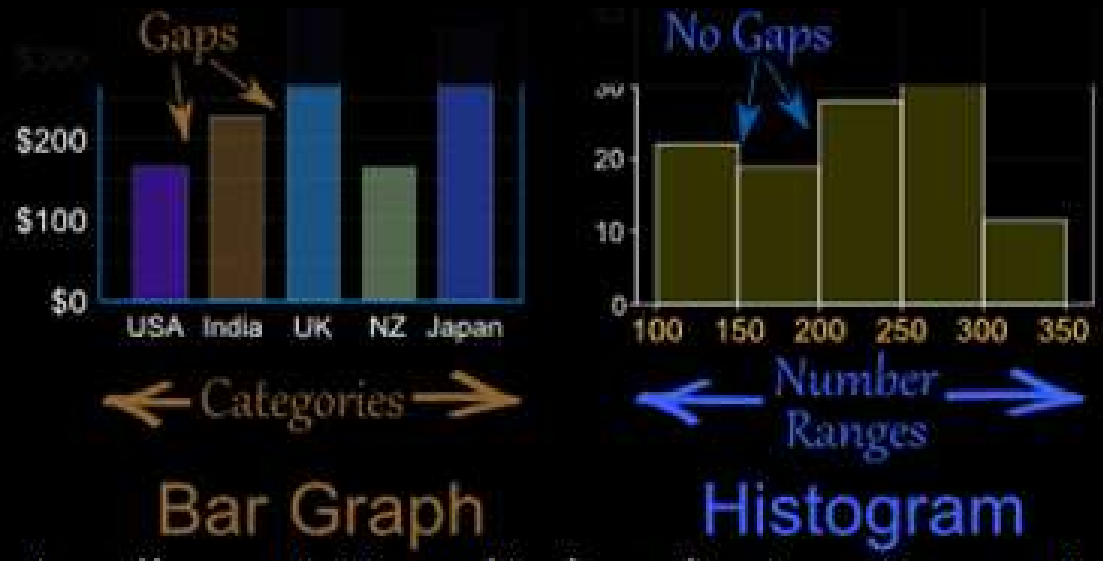


Histogram (2 of 3)



Histogram (3 of 3)

- Bar chart for grouped numerical data
  - No (or small) gaps btwn adjacent bars



https://www.mathsisfun.com/data/images/bar-chart-vs-histogram.gif

Stem and Leaf Display

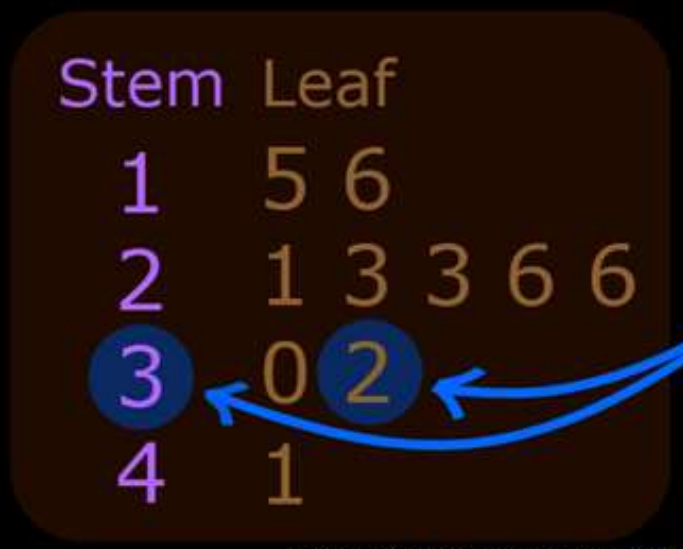
- "Histogram-lite" for analysis w/out software
  - e.g., points on homework

15,16,21,23,23,26,26,30,32,41

Stem and Leaf Display

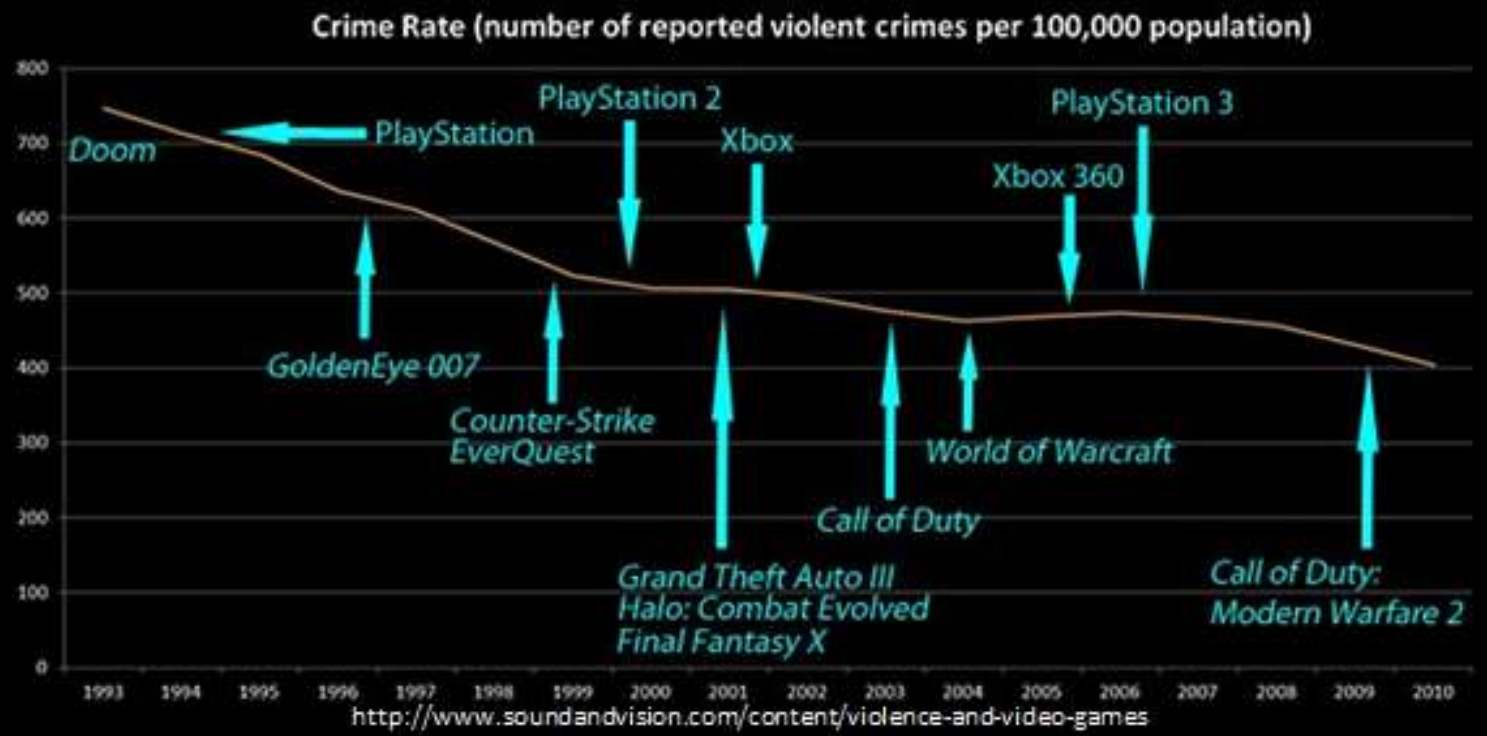
- "Histogram-lite" for analysis w/out software
  - e.g., points on homework

15,16,21,23,23,26,26,30,32,41



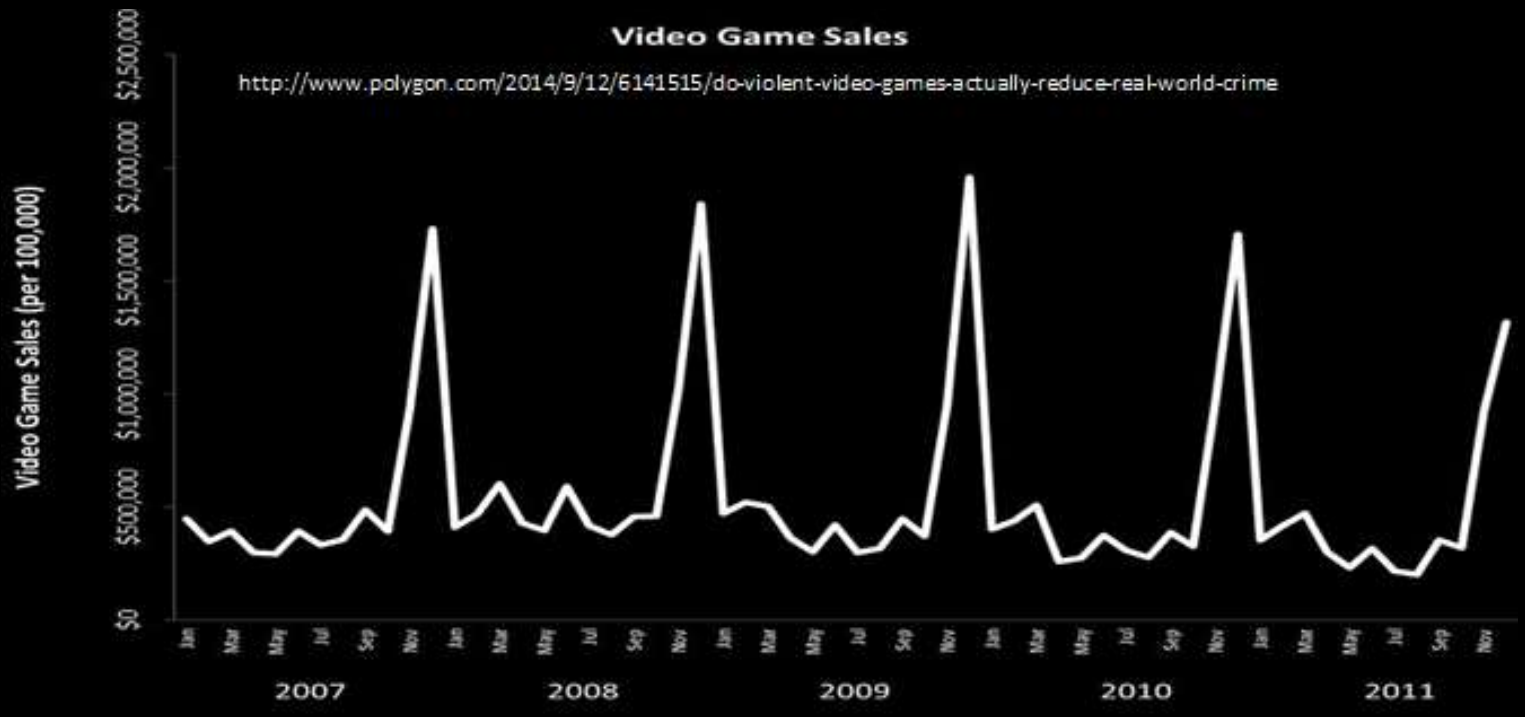
https://www.mathsisfun.com/data/stem-leaf-plots.html

Time Series Plot



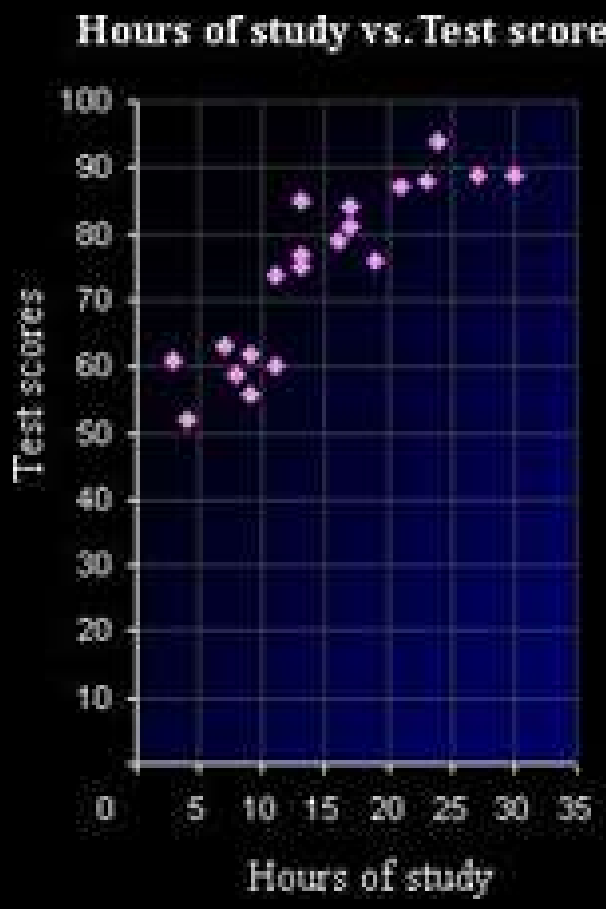
- Associate data with date / time
- Line graph with dates (proportionally spaced!)

Time Series Plot



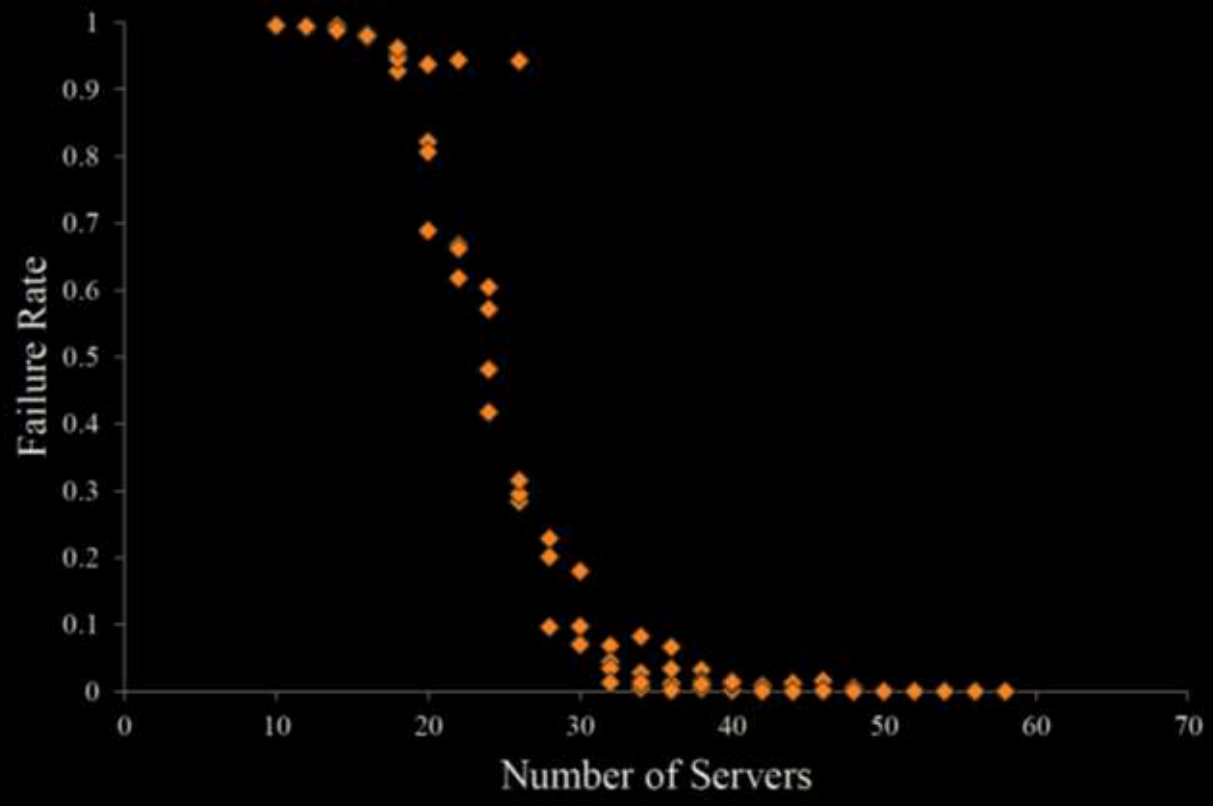
- Associate data with date / time
- Line graph with dates (proportionally spaced!)

Scatter Plot



- Two numerical variables, one on each axis
- Reveal 2D patterns in relationship
- Setup “right” models (later)

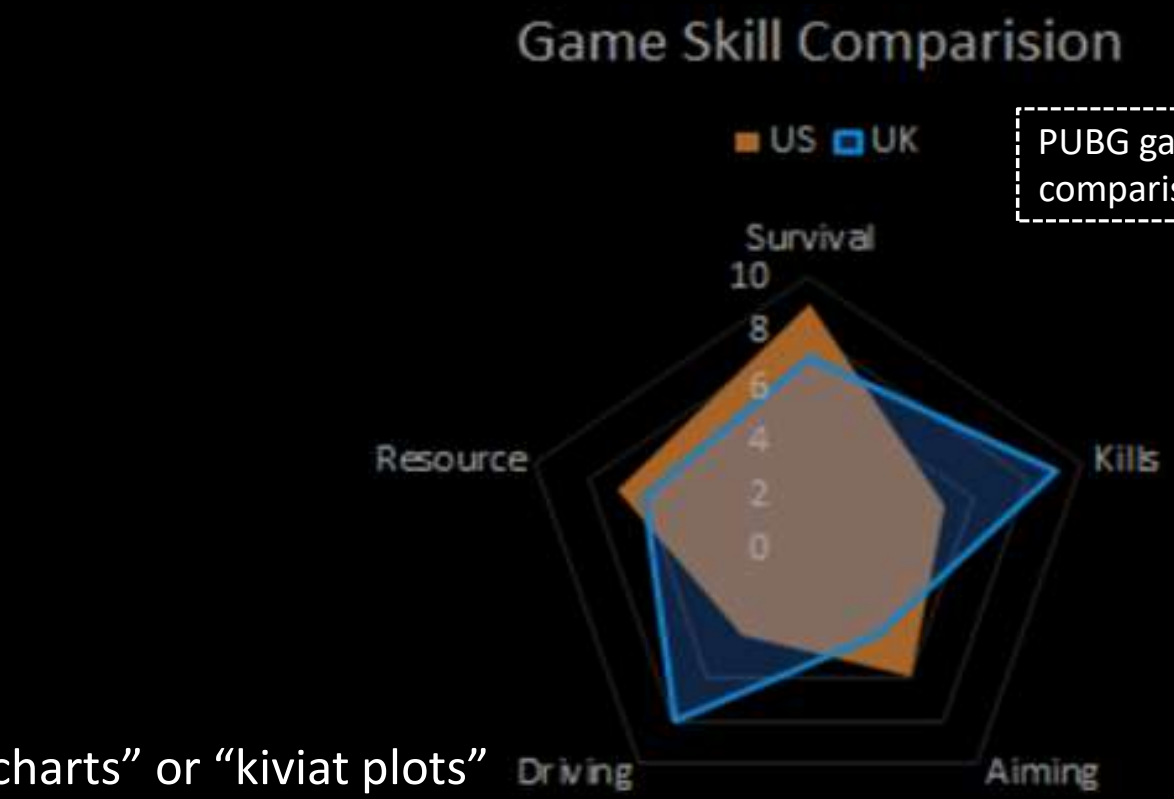
Scatter Plot



"Intelligent Simulation of Worldwide Application Distribution for OnLive's Server Network"

http://www.cs.wpi.edu/~claypool/mqp/onlive/

Radar Plot



PUBG game comparison

- Aka “star charts” or “kiviat plots”
- Good for quick visual comparison, especially when axes unequal

Many More Charts!

https://en.wikipedia.org/wiki/Chart

- Bubble
- Waterfall
- Tree
- Gap
- Polar
- Violin
- Candlestick
- Kagi
- Gantt
- Nolan
- Pert
- Smith
- Skyline
- Vowel
- Nomogram
- Natal

- If common chart effective for message, use
- Otherwise, learn/use other charts as needed
- But remember – may need to explain how to read

Outline

- Types of Charts (done)
- Game Analytics Examples (next)
- Guidelines for Charts

Game Analytics Charts

Gunter Wallner and Simone Kriglstein. "An Introduction to Gameplay Data Visualization", *Game Research Methods*, pages 231-250, ETC Press, ISBN: 978-1-312-88473-1, 2015. http://dl.acm.org/citation.cfm?id=2812792

- Player choices (e.g., build units)
- Density of activities (e.g., where spend time on map)
- Movement through levels

Player Choices – Pie-Chart Overlay

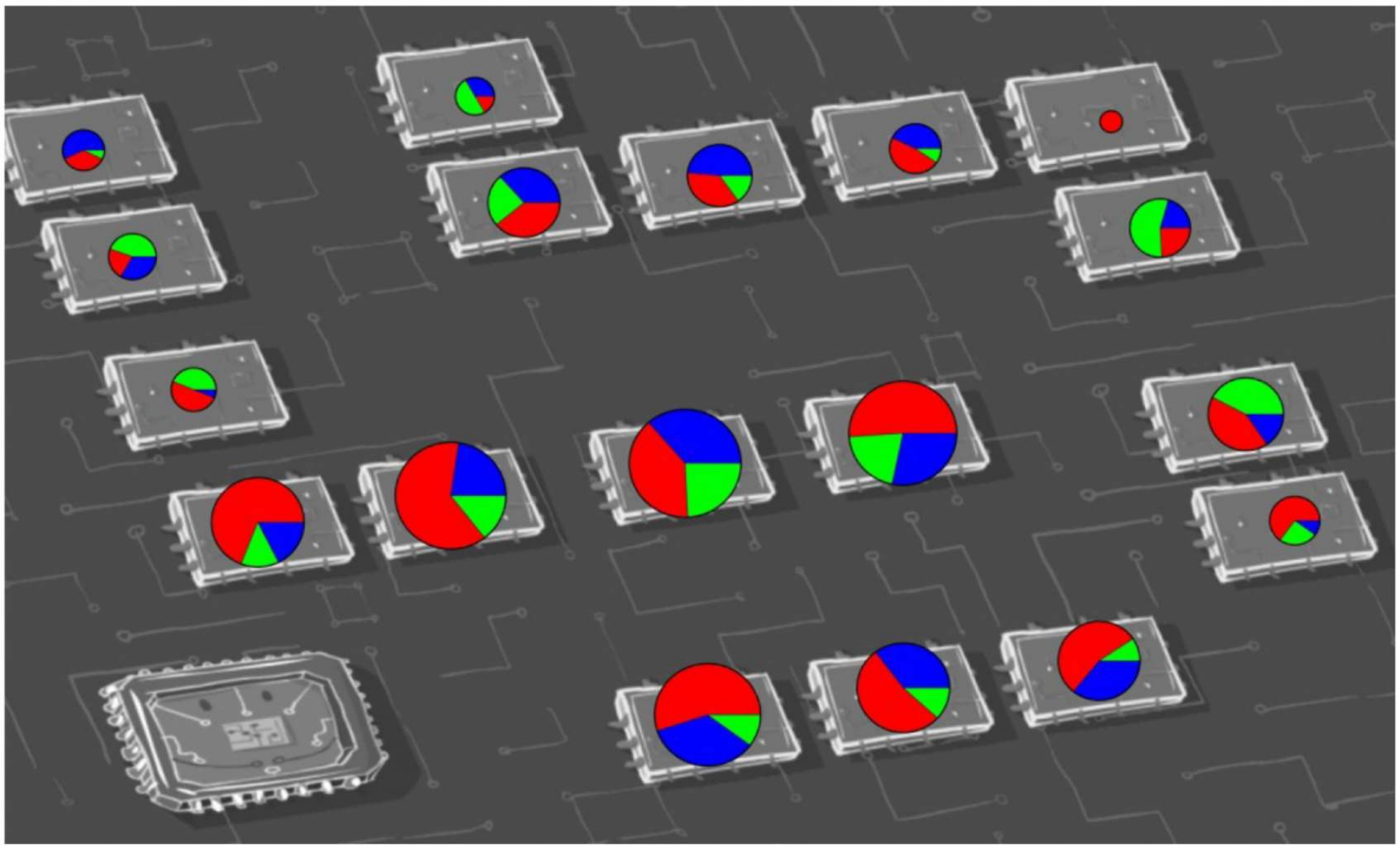
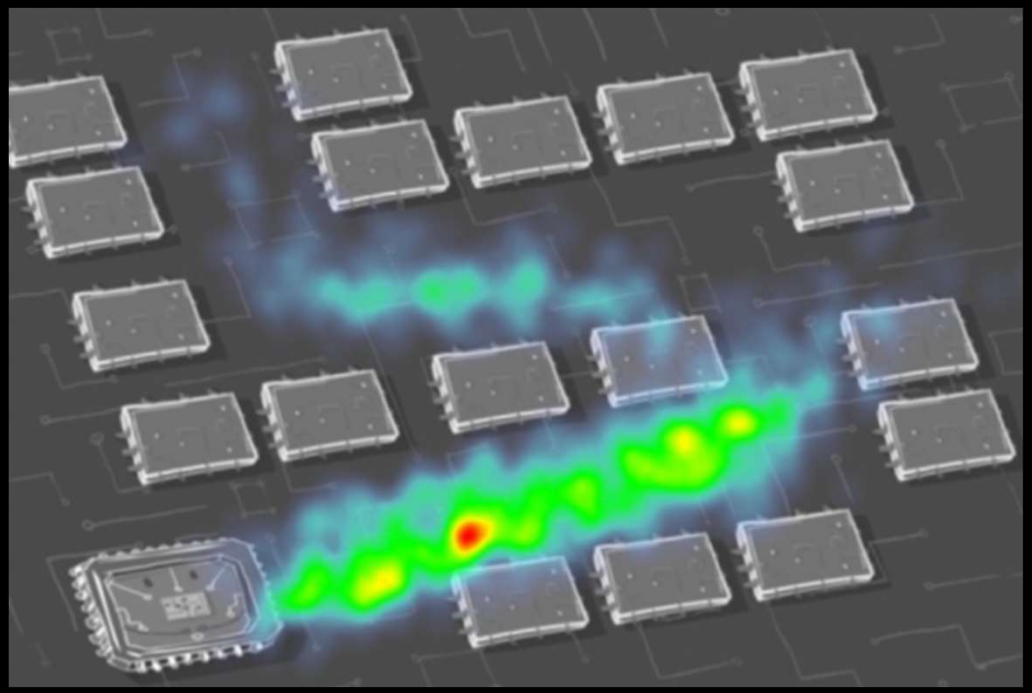


Figure 1. Pie-charts show which types of towers have been built on the different building lots. The radius of the pie-chart is proportional to the number of towers built (Kayali, et al., 2014).

Player Location – Heat Map (1 of 2)



Player Location – Heat Map (2 of 2)

Assassin's Creed

Where play testers failed

Result: Make red areas easier

http://www.gamasutra.com/blogs/JonathanDankoff/20140320/213624/Game_Telemetry_with_DNA_Tracking_on_Assassins_Creed.php

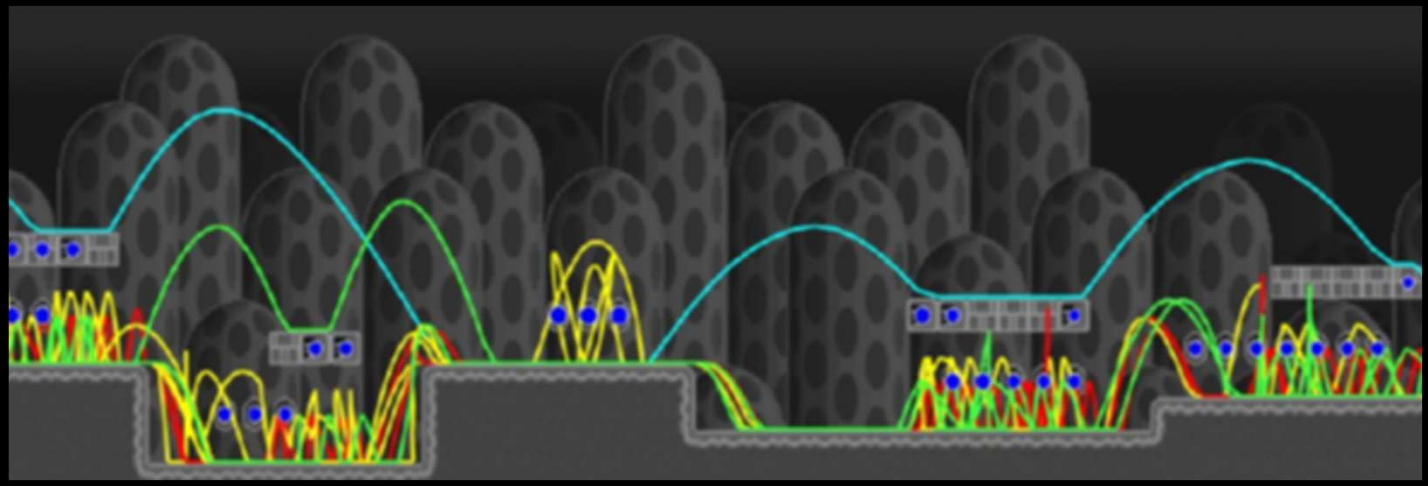
Note, Heat Map for Tables, Too!

	A	B	C	D
1		2014	2015	2016
2	January	600	708	594
3	February	607	984	749
4	March	901	886	908
5	April	608	615	835
6	May	715	833	734
7	June	520	663	618
8	July	731	521	950
9	August	709	663	987
10	September	633	863	979
11	October	533	651	841
12	November	996	958	749
13	December	792	717	875

Red means sales are low

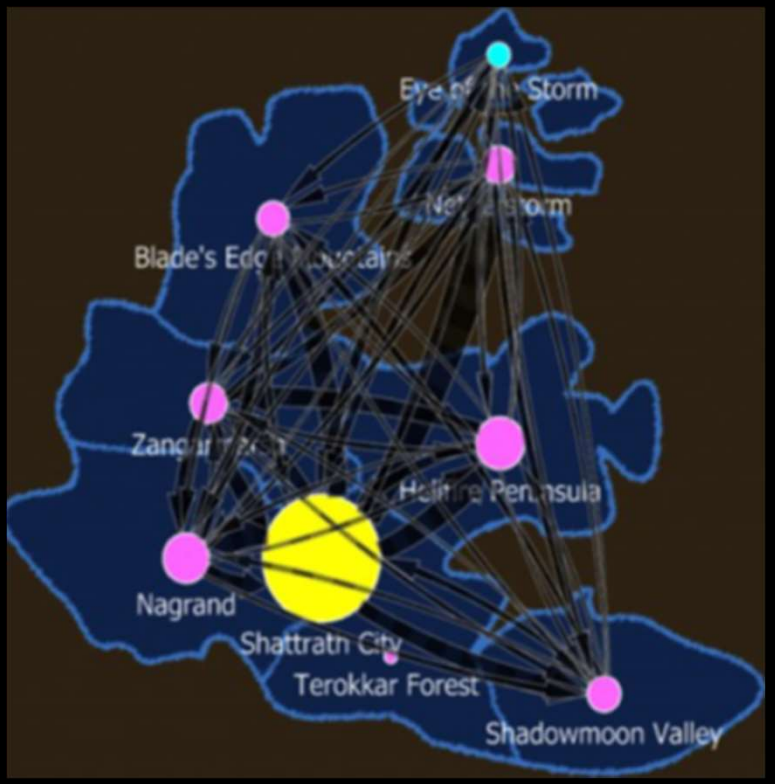
Excel tutorial at: https://trumpexcel.com/heat-map-excel/

Movement (1 of 3) GA



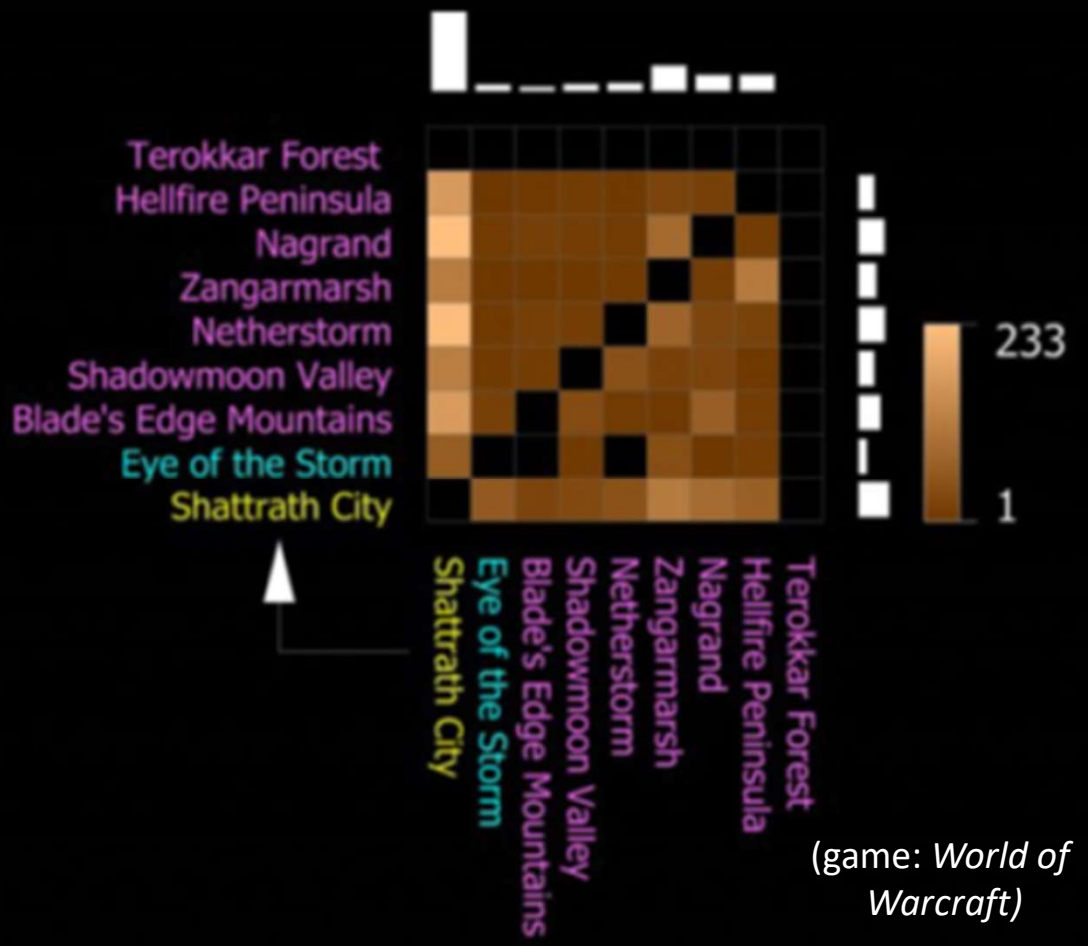
(game: *Infinite Mario*, clone of Super Mario Bros.)

Movement (2 of 3)



(game: *World of Warcraft)*

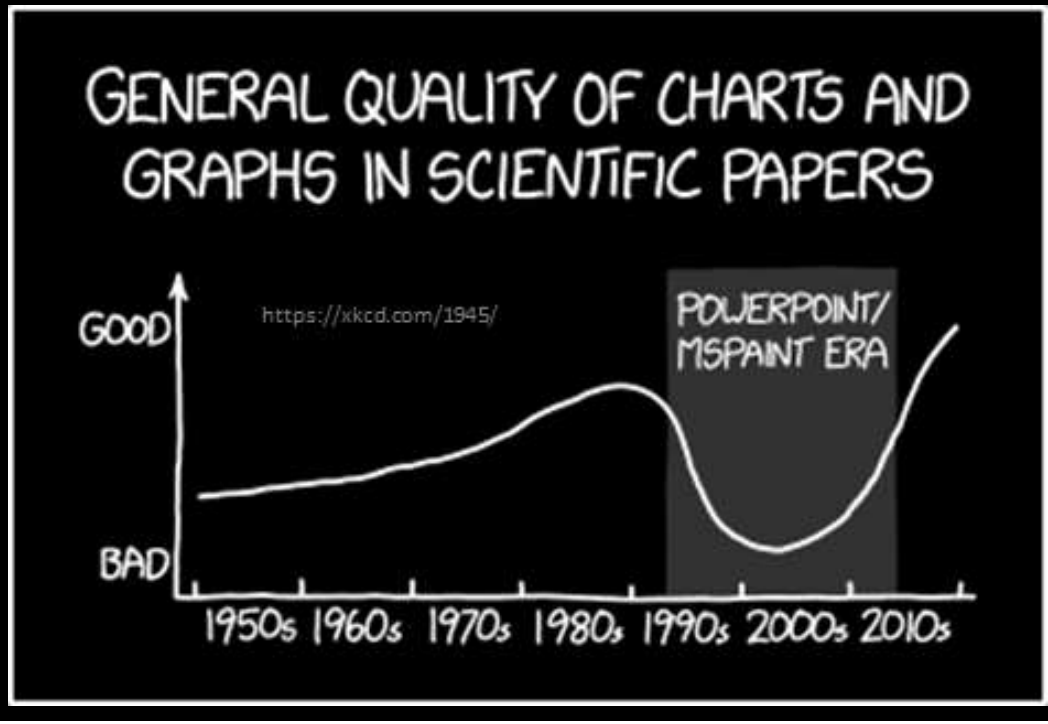
Movement (3 of 3)



(game: *World of Warcraft)*

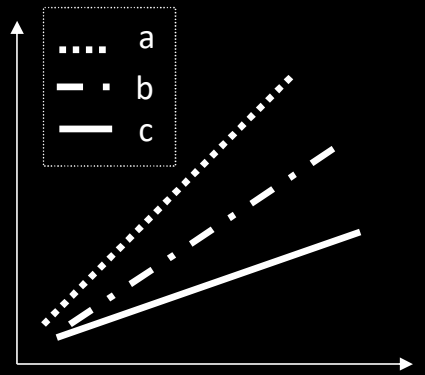
Outline

- Types of Charts (done)
- Game Analytics Examples (done)
- Guidelines for Charts (next)



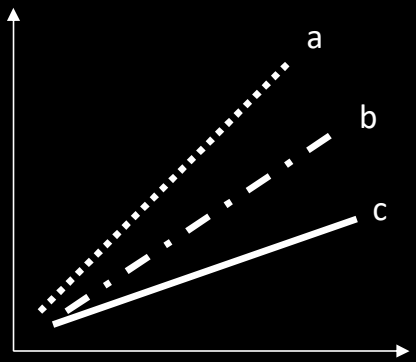
Guidelines for Good Charts (1 of 7)

- Require minimum effort from reader
  - Perhaps *most* important metric
  - Given two, can pick one that takes less reader effort



Direct Labeling

e.g.,



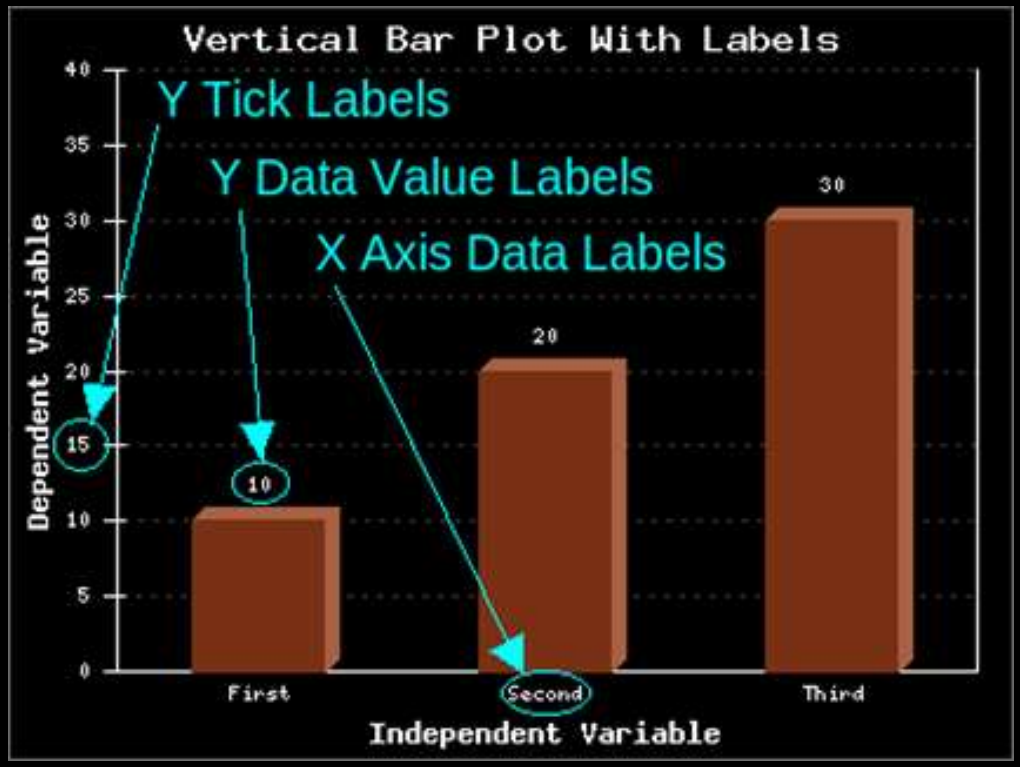
Legend Box

Guidelines for Good Charts (2 of 7)

- Maximize information
  - Make self-sufficient
  - Key words in place of symbols
    - e.g., "Gold IV" and not "Player A"
    - e.g., "Daily Games Played" not "Games Played"
  - Axis labels as informative as possible
    - e.g., "Game Time (seconds)" not "Game Time"
  - Help by using captions (or title, if stand-alone)
    - e.g., "Game time in seconds versus player skill in total hours played"

Guidelines for Good Charts (2 of 7)

- Maximize information



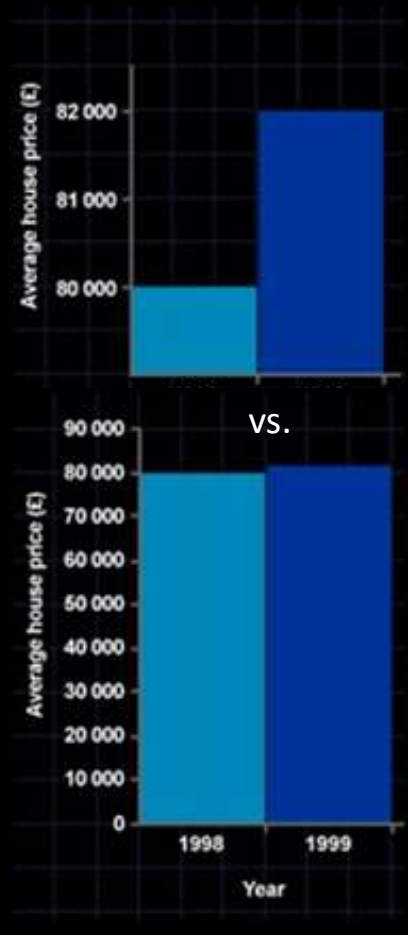
Guidelines for Good Charts (3 of 7)

- Minimize ink

https://www.slideshare.net/NicoleMarinsek/darkhorse-line-chart

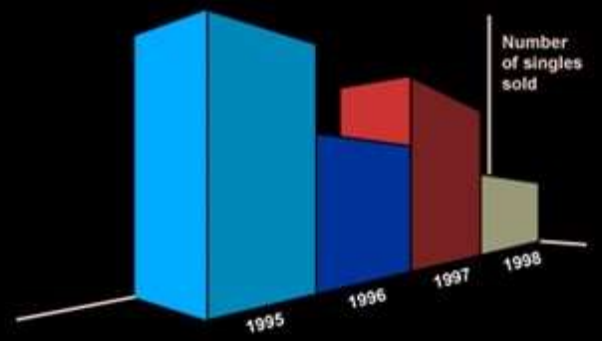
Guidelines for Good Charts (4 of 7)

- Use commonly accepted practices
  - Present what people expect
  - e.g., origin at (0,0)
  - e.g., independent (cause) on x-axis, dependent (effect) on y-axis
  - e.g., x-axis scale is linear
  - e.g., increase left to right, bottom to top
  - e.g., scale divisions equal, proportional
- Departures permitted but require extra effort from reader → so use sparingly!

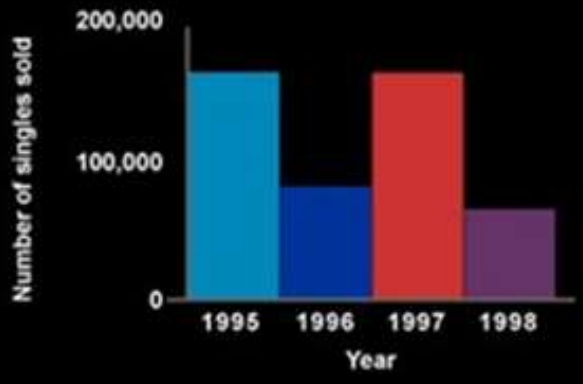


Guidelines for Good Charts (5 of 7)

- Avoid ambiguity
  - Show coordinate axes
    - at right angles
  - Show origin
    - usually at (0,0)
  - Identify individual curves and bars
    - With key/legend or label
  - Do not plot multiple variables on same chart
    - Single y-axis



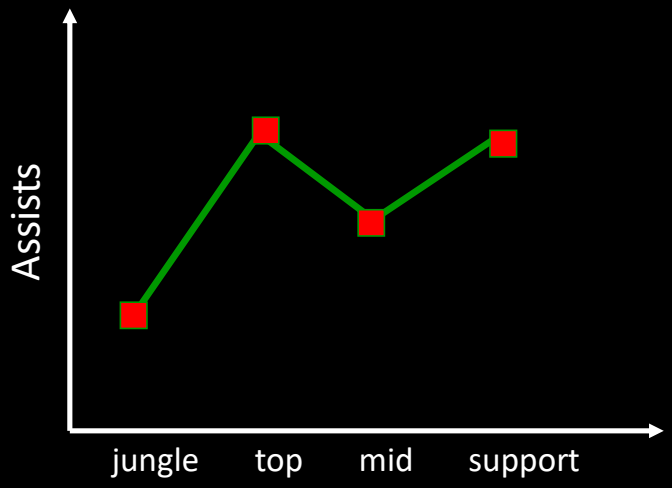
vs.



Guidelines for Good Charts (6 of 7)

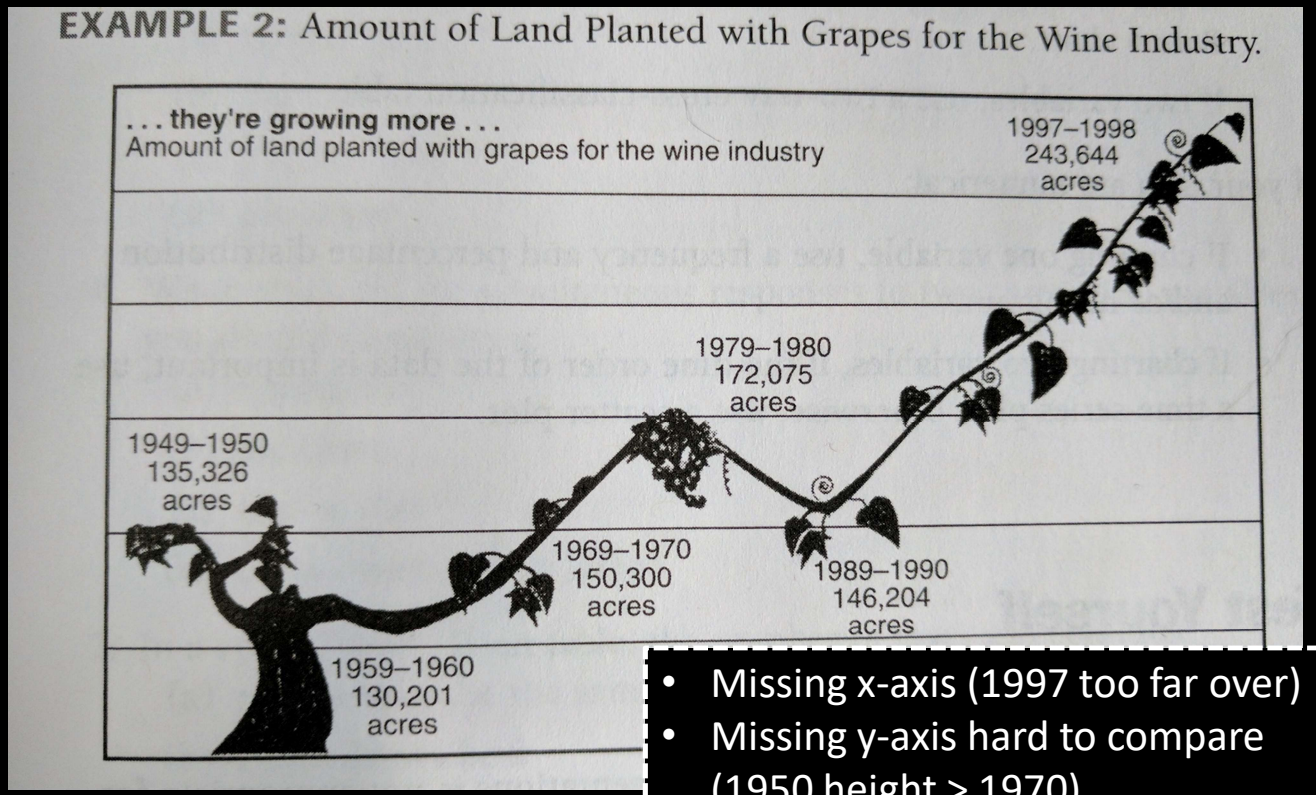
- Don’t connect categorical data with lines
  - Lines joining successive points signify that they can be approximately interpolated
  - If don’t have meaning, should not use line chart

- No linear relationship between e.g, champion types
- Instead, use column chart
- Don’t connect with lines



Guidelines for Good Charts (7 of 7)

- Can deceive as easily as can convey meaning

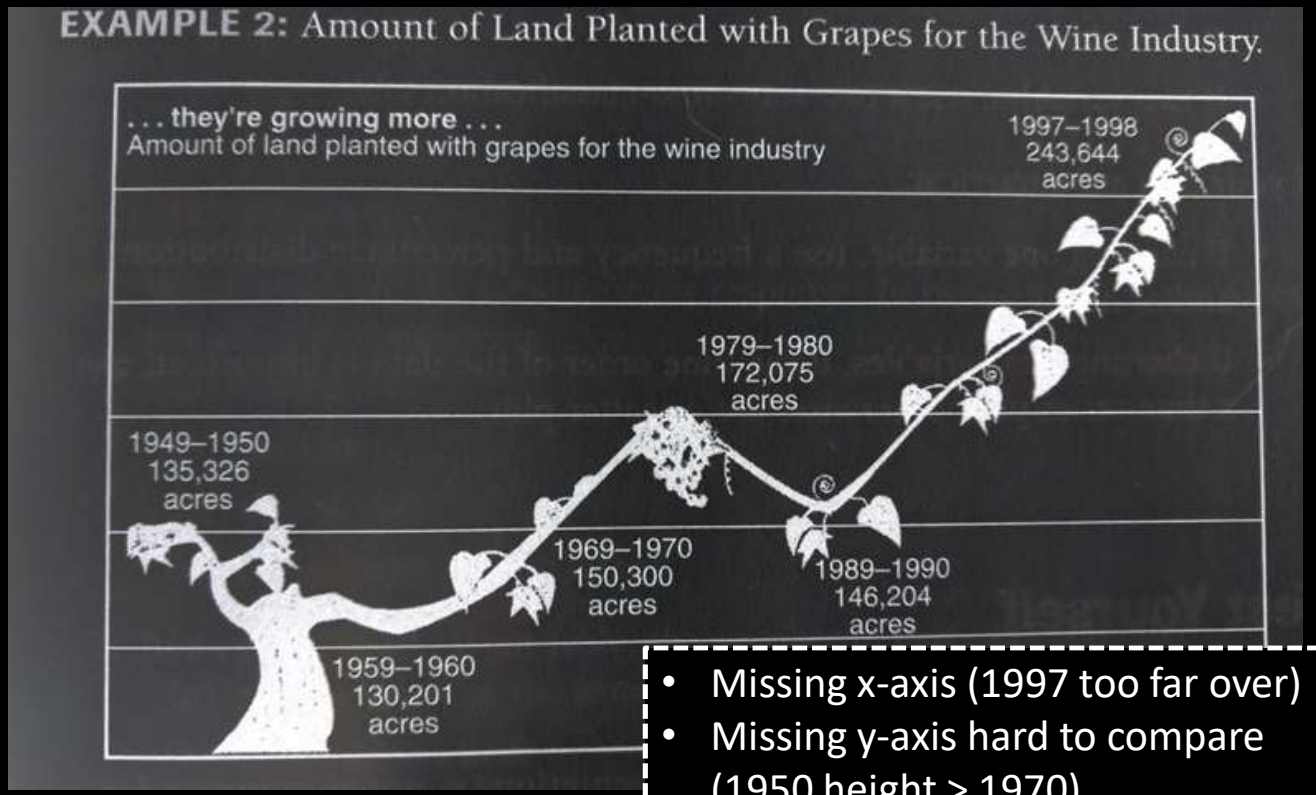


- Missing x-axis (1997 too far over)
- Missing y-axis hard to compare (1950 height > 1970)
- Data points unclear

Guidelines for Good Charts (7 of 7)

GA

- Can deceive as easily as can convey meaning



- Missing x-axis (1997 too far over)
- Missing y-axis hard to compare (1950 height > 1970)
- Data points unclear

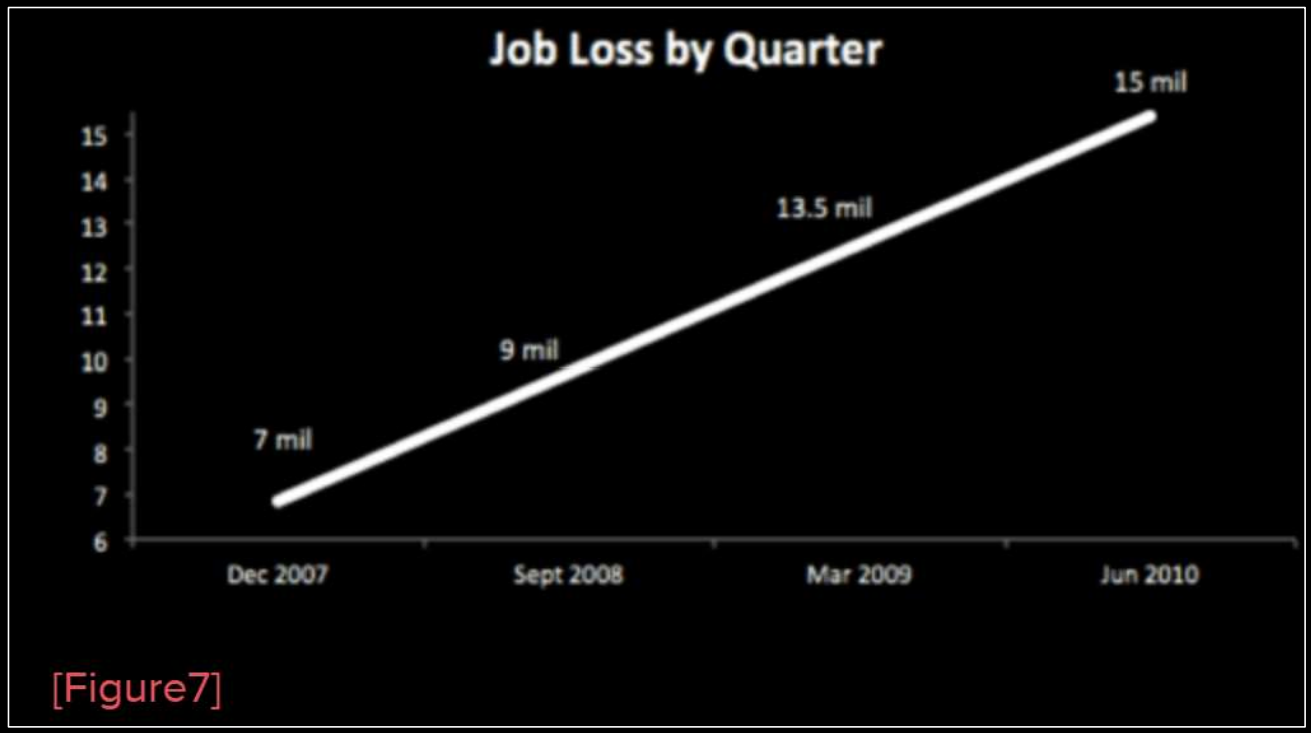
Breakout 2

1. Identify problems. Write them down.
2. After 2 minutes, compare. Discuss differences.
3. Write down combined set

Icebreaker, Groupwork, Questions

https://web.cs.wpi.edu/~imgd2905/d20/breakout/breakout-1.html

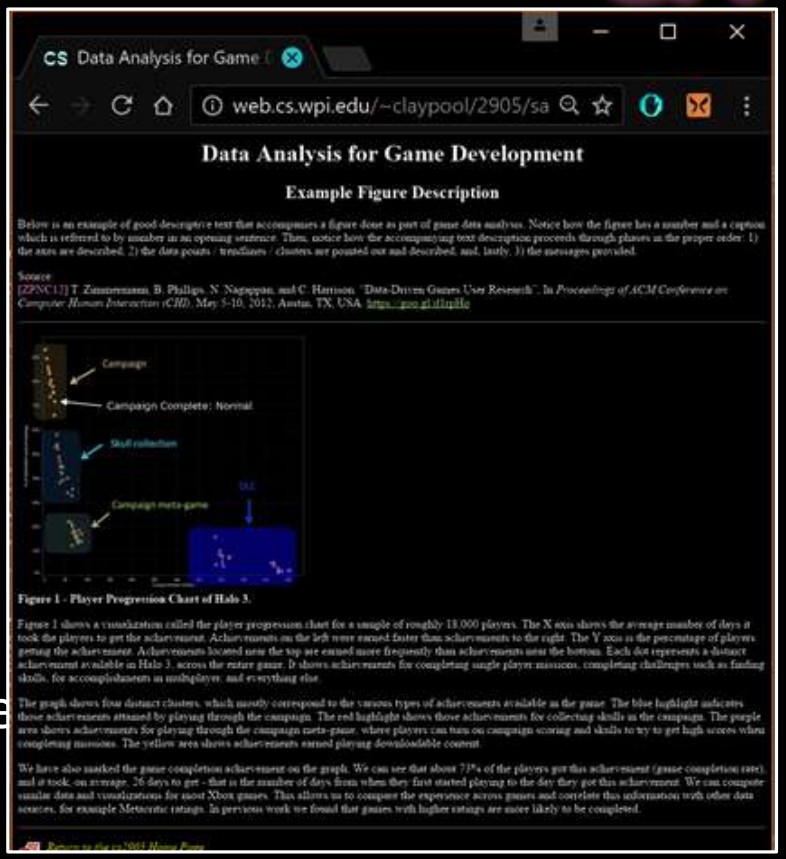
Breakout 2



[Figure7]

Describing Chart

- "Formula"
  - Describe all axes
    - E.g., "The x-axis is time since game began, in seconds"
  - Describe data sets/trendlines
    - E.g., "The blue dots are the average maze completion time"
  - **Then** provide message
    - E.g., "Notice how the red bar is higher than the blue, indicating that …"



http://web.cs.wpi.edu/~imgd2905/d21/samples/analysis-example.html

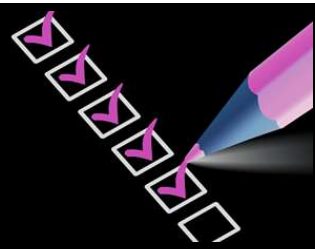
Checklist for Good Charts (1 of 2)

- Axes
  - Are both axes labeled?
  - Are axis labels self-explanatory and concise?
  - Are scale and divisions shown on both axes?
  - Are min and max ranges appropriate?
  - Are units indicated?
- Lines/Curves/Points
  - Is number of lines/curves reasonably small?
  - Are curves labeled?
  - Are all symbols clearly distinguishable?
  - Is concise, clear legend provided?
  - Does the legend obscure any data?
- Information
  - If y-axis is variable, is indication of spread (error bars) shown?
  - Are grid lines required to read data (if not, then remove)?

Checklist for Good Charts (2 of 2)

- Scale
  - Are units increasing left to right (x-axis) and bottom to top (y-axis)?
  - Do all charts use the same scale?
  - Are scales contiguous?
  - Is bar chart order systematic/deliberate?
  - Are bars appropriate width, spacing?
- Overall
  - Does whole chart add information to reader?
  - Are there no curves/symbols/text that can be removed and still have same information?
  - Does the chart have title or caption (not both)?
  - Is chart self-explanatory and concise?
  - Do variables plotted give more information than alternatives?
  - Is chart referenced and discussed in any accompanying report?

Guidelines for Good Charts (Summary)



- For each chart, go over "checklist"
- The more "yes" answers, the better
  - Remember, while guidelines, art and not science
  - So, may consciously decide not to follow these guidelines if better without them → but have *good* reason!
- In practice, takes several trials before arriving at "best" chart
- Want to present message the most: accurately, simply, concisely, logically
- Accompany with description! Text or verbal
  - Remember, audience/reader has *not* seen! – Make sure to introduce