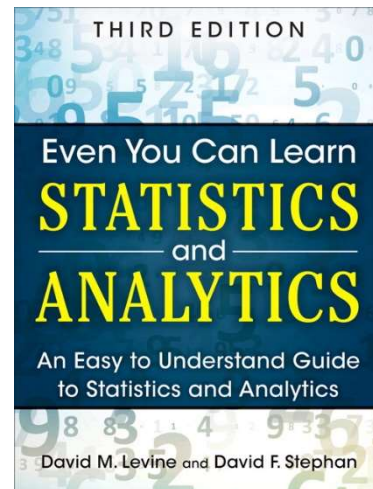


IMGD 2905

Presenting Data

Chapter 2

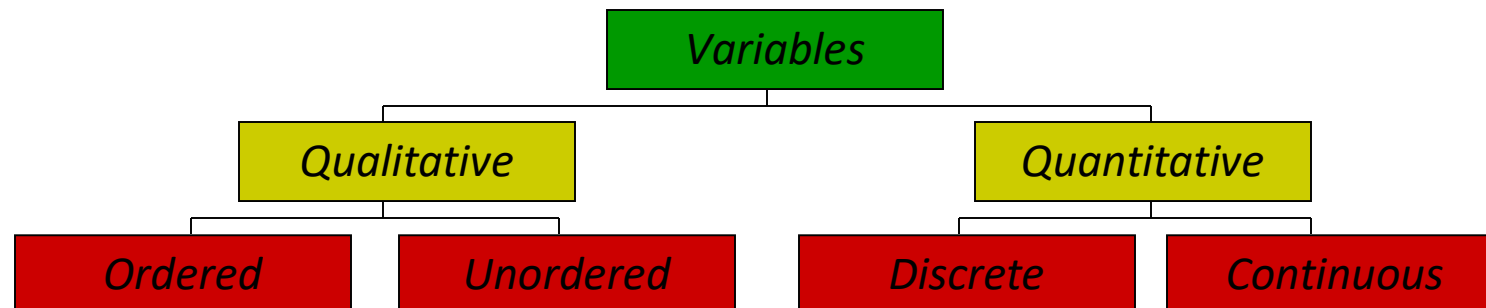


Outline

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

“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**
- **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

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

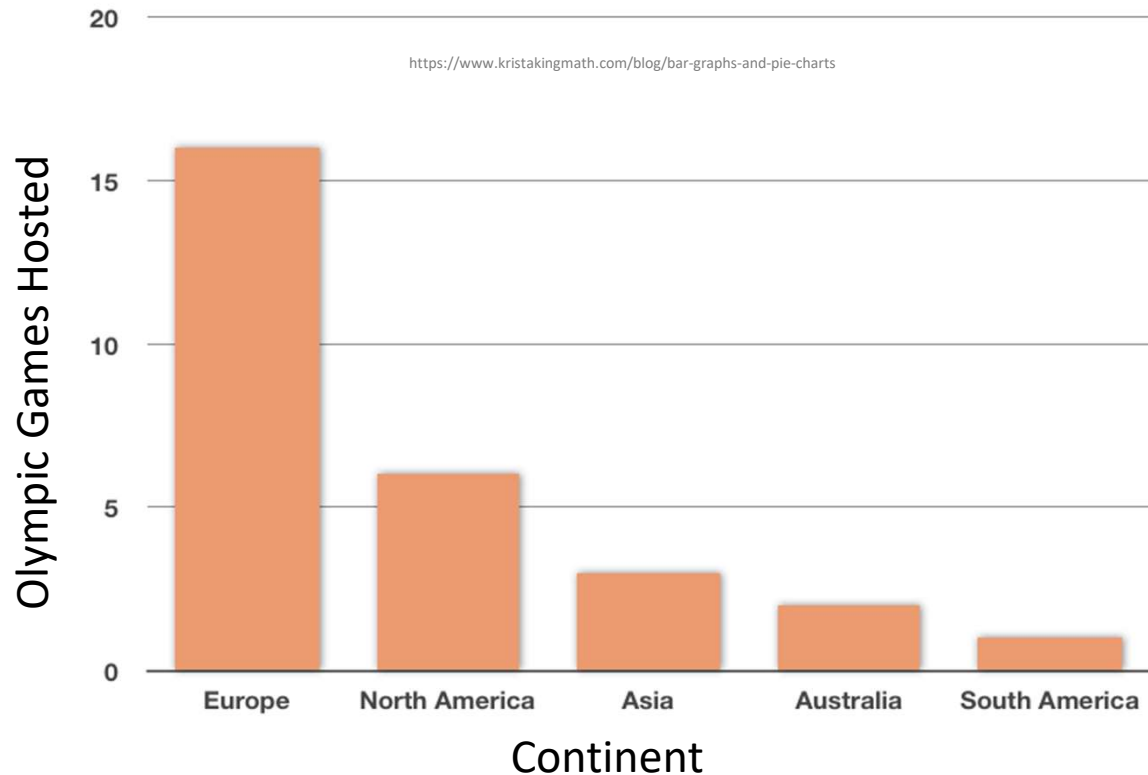
		Do you live on campus?	
		Off-campus	On-campus
Class rank	Freshman	37	100
	Sophomore	42	48
	Junior	90	8
	Senior	62	1
Total		231	157

- Number and caption
- Units labeled (as appropriate)
- Minimal vertical lines (or none)
- Lines only to break apart areas (or use Bold)

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

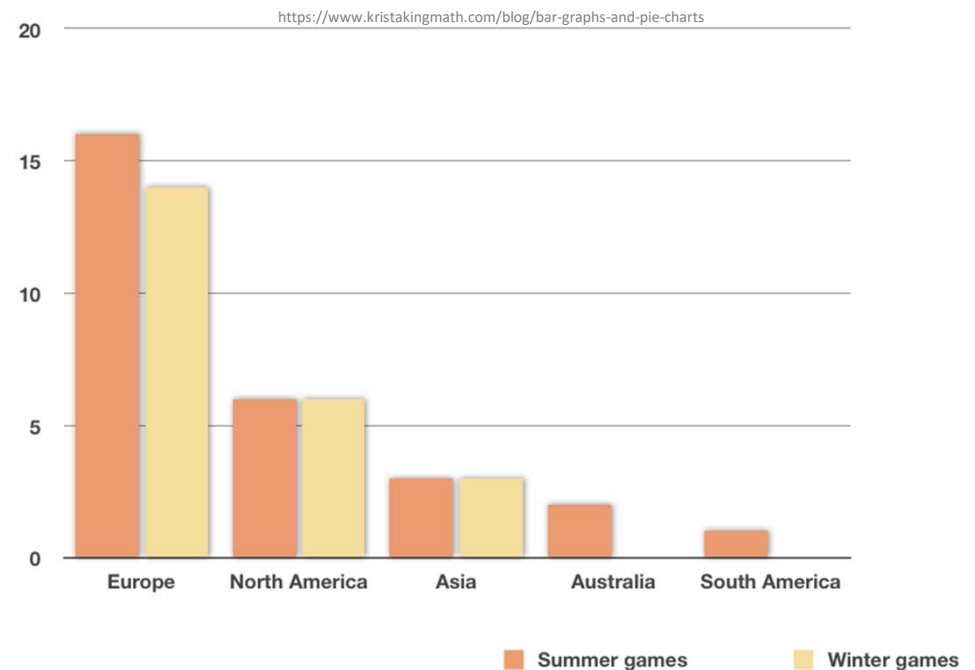
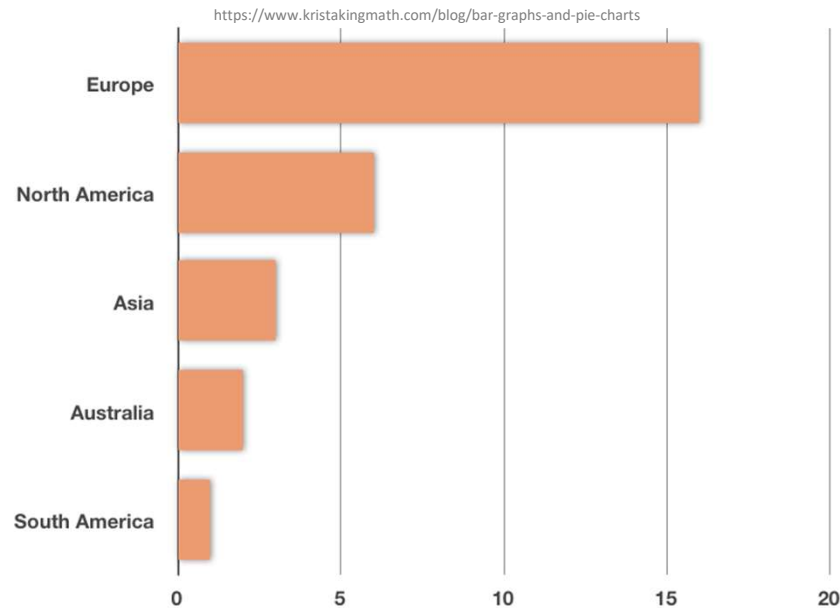
Categorical: Bar Chart (1 of 2)

- 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 2)

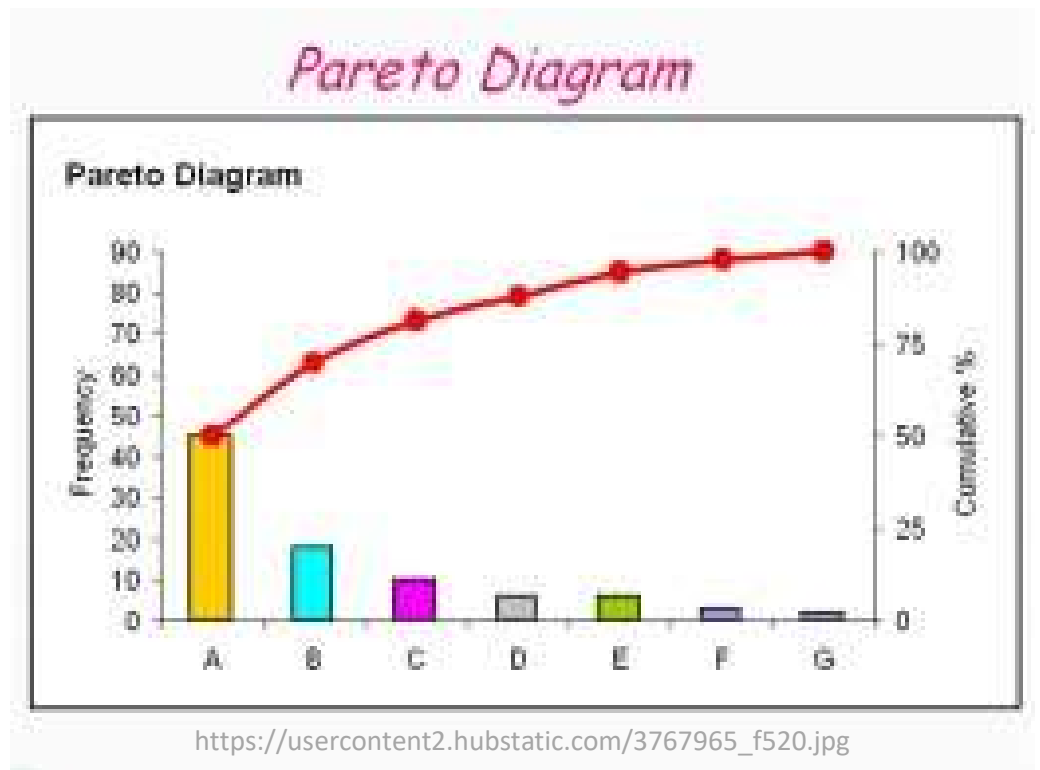
- Chart containing rectangles (“bars”) where length represents count, amount, or percent



Categorical: Pareto Chart

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

Demo: [imgdpops.xlsx](#)



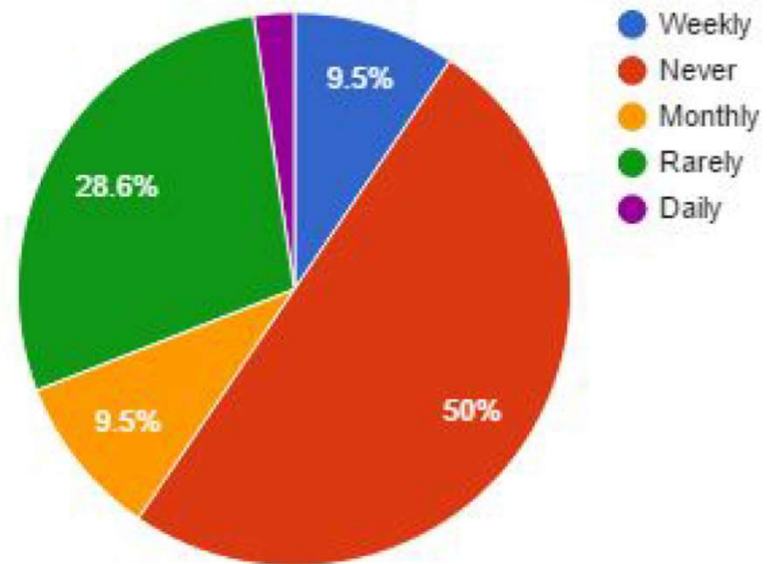
Sort by column D (Data -> Sort high to low)
New column E for percent [=D2/SUM(D\$2:D\$11)]
Note: \$ "locks" value in (e.g., D\$2 versus D2)
New column F for running [=SUM(E\$2:E2)]
Select B, D and F. Insert "combo chart"

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](#)

Time playing with Team Fortress 2



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

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

Cumulative Distribution

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

Demo: [lol-patches.xlsx](#)

Select column R (Bug Fixes)

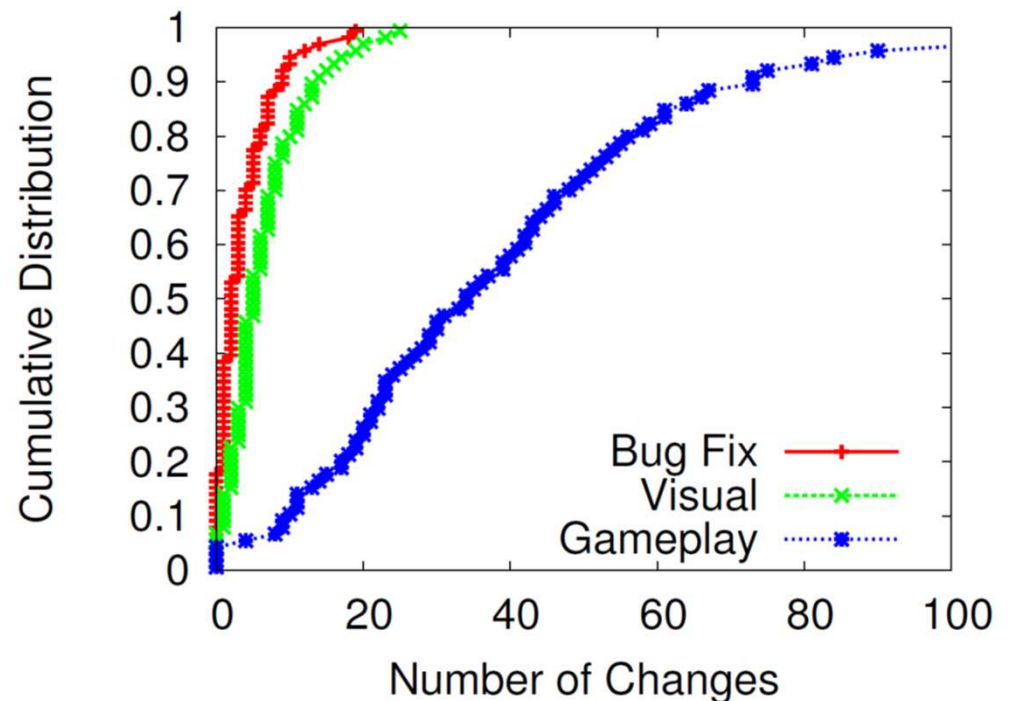
Sort low to high

New column S for percent $[=ROW()/164]$

Select column → paste down all

Select both column R and S

Insert → Scatter plot with lines

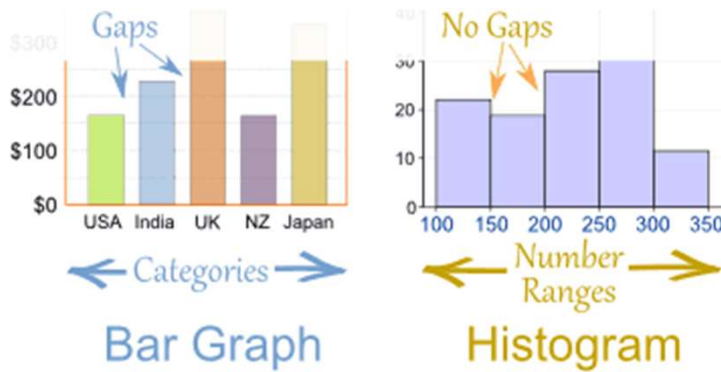


“Nerfs, Buffs and Bugs - Analysis of the Impact of Patching on League of Legends”

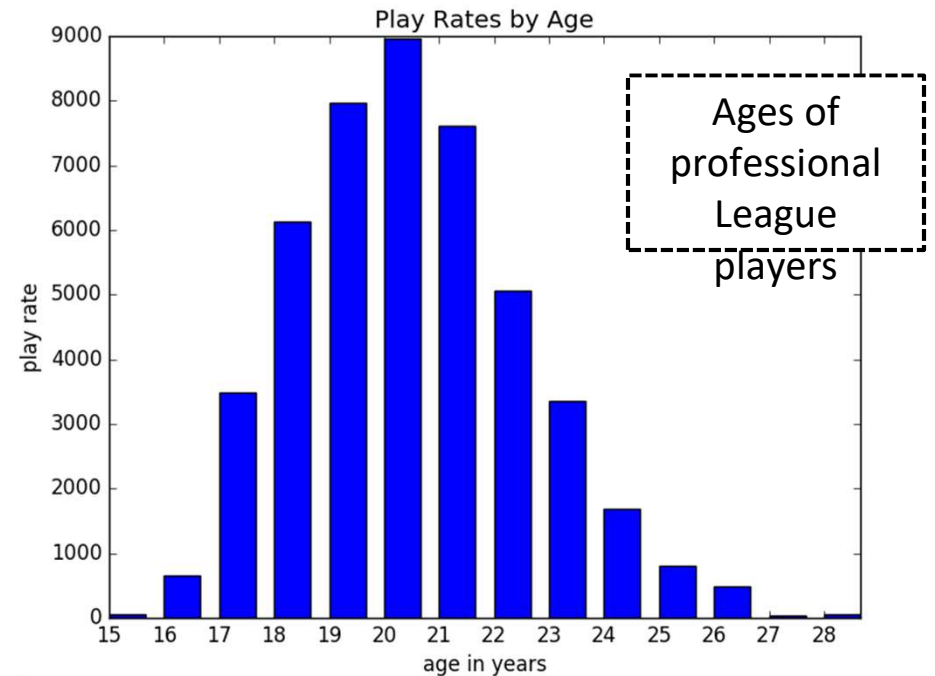
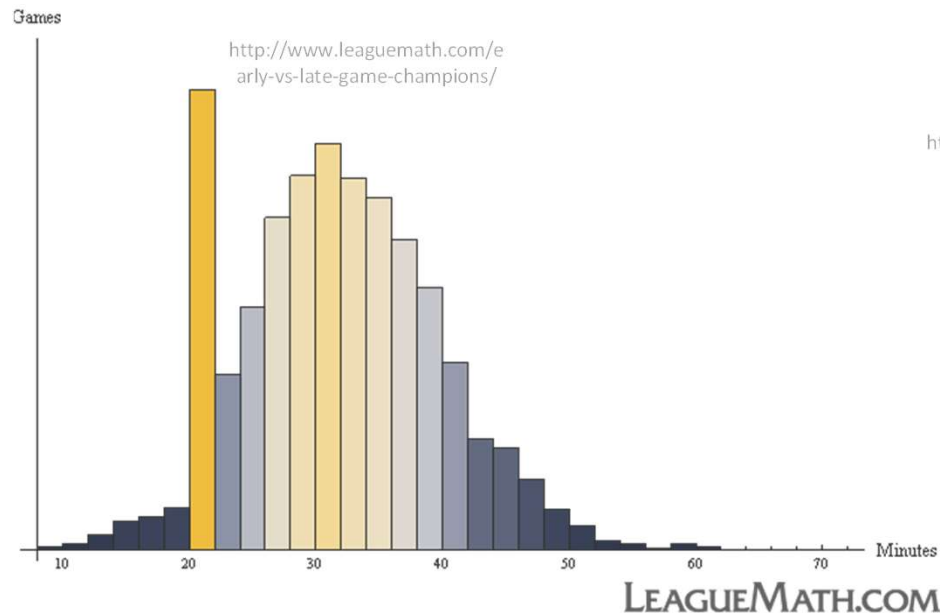
<http://www.cs.wpi.edu/~claypool/papers/lol-crawler/>

Histogram

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



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



https://www.reddit.com/r/leagueoflegends/comments/4x5s9m/analysis_of_age_in_league_of_legends/

Demo: [grades.xlsx](#)

Select GPA data

Insert → Statistics Chart → Histogram

Can adjust bins, overflow/underflow

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

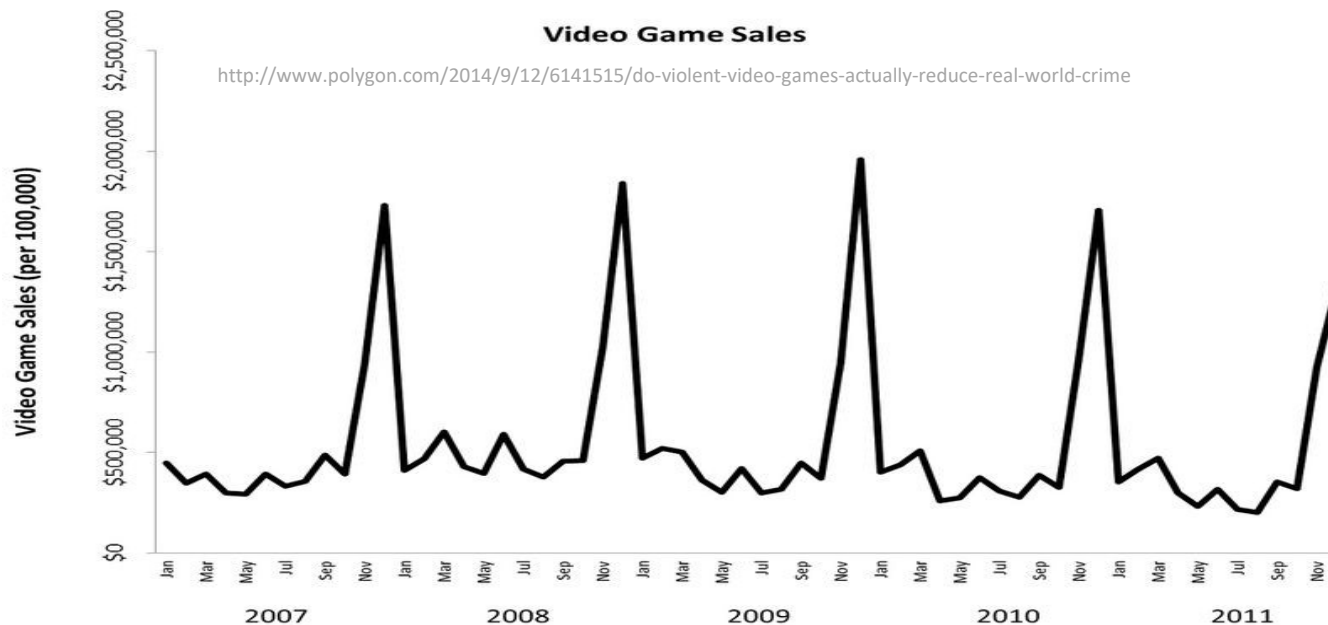
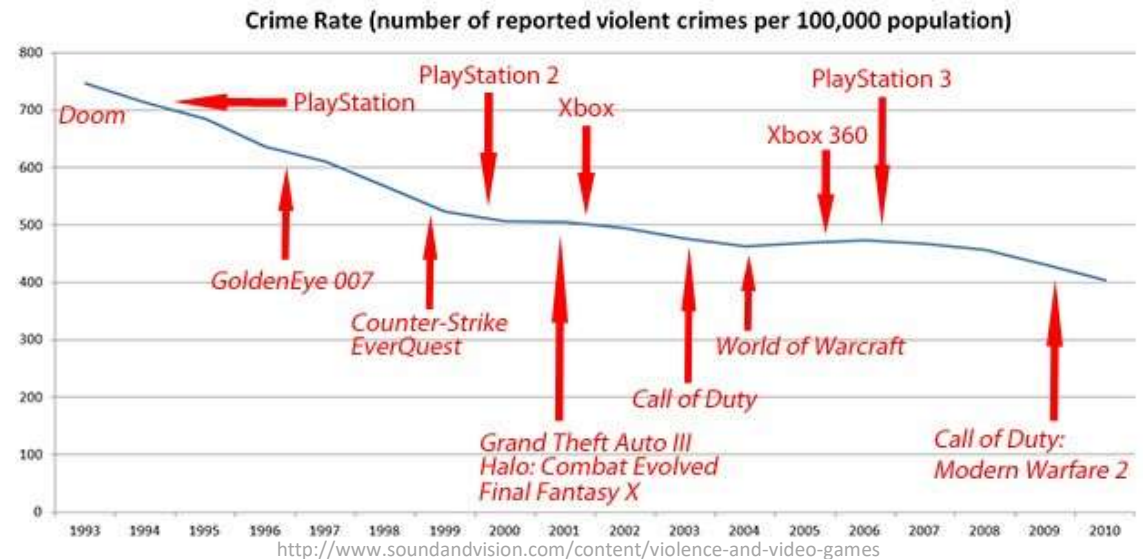
Stem	Leaf
1	5 6
2	1 3 3 6 6
3	0 2
4	1

*how to
place "32"*

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

Time Series Plot

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

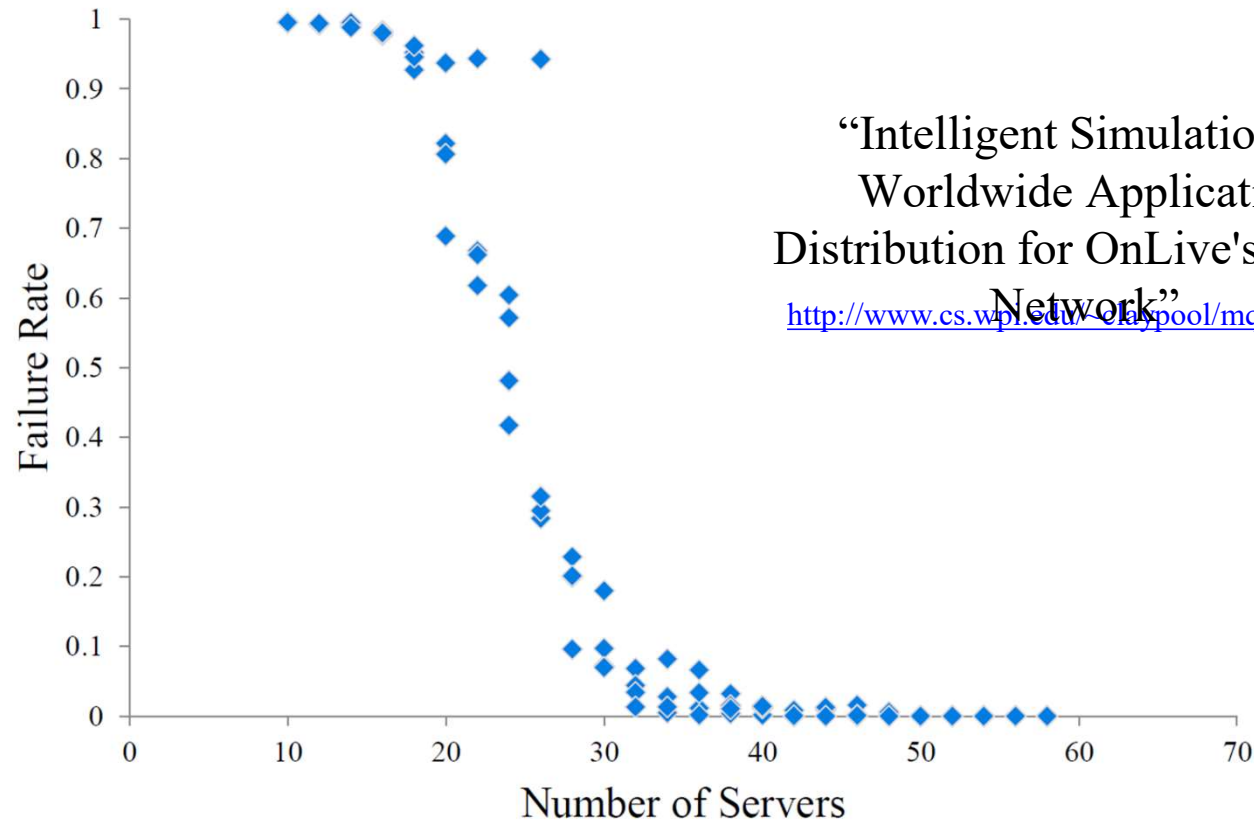
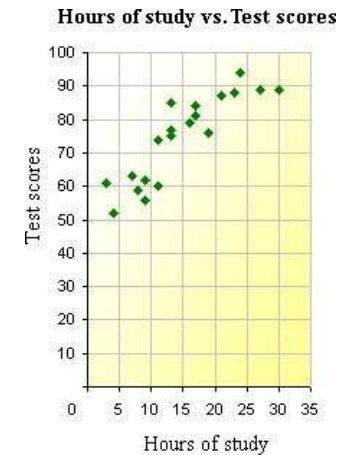


Demo: [majors.xlsx](#)

Sel. year and majors
Insert → Line Chart
→ More Line Charts

Scatter Plot

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



Demo: [lol-rates.xlsx](#)

Select two of {win,
pick, ban}
Insert → scatter plot

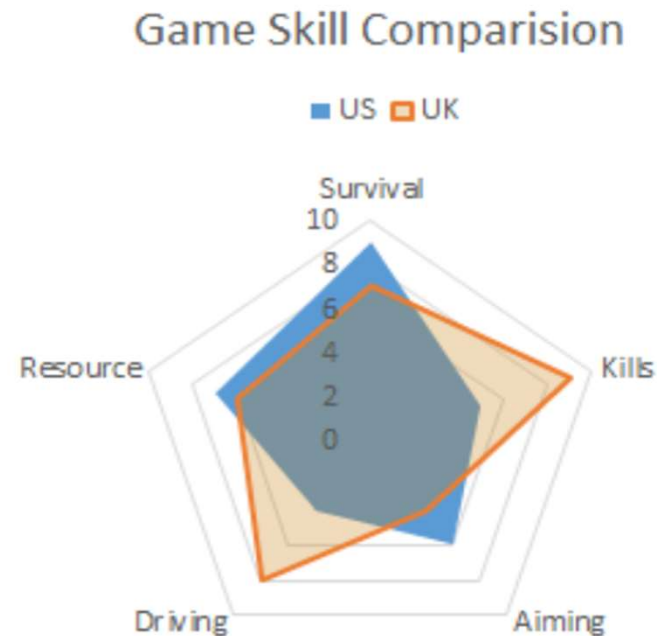
Radar Plot

PUBG game comparison

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

Demo: [lol-rates.xlsx](#)

Select top line {win, pick, ban} + 3 rows (Ctrl-select)
Insert → Other → Radar scatter plot



<https://www.exceltip.com/wp-content/uploads/2019/11/00213.png>

Note: will need to normalize data to scale Axes
Insert column E (“B Norm”)
 $=E2/\text{MAX}(E\$2:E42)$
Copy and paste down

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

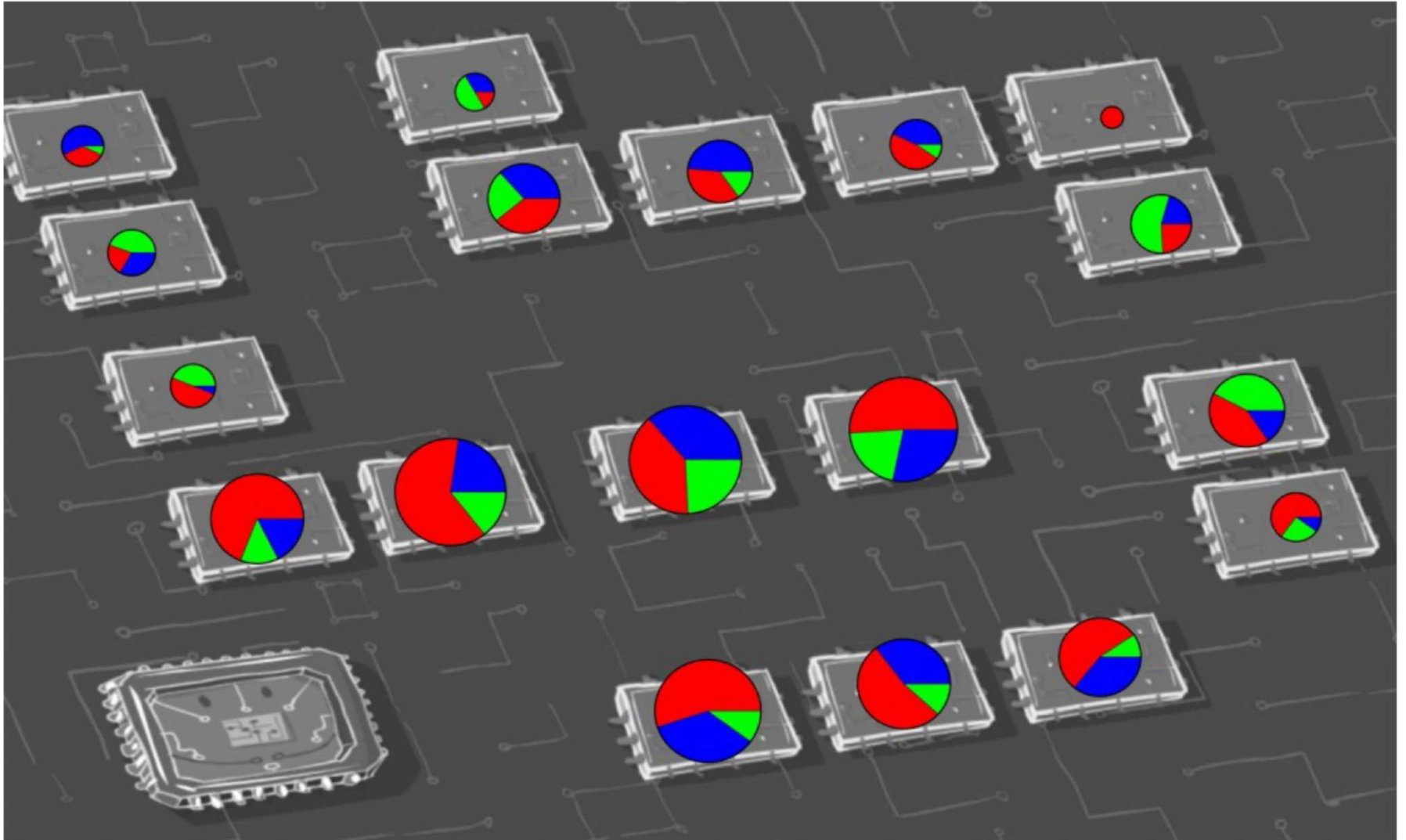


Figure 1. Pie-charts show which types of towers have been built on the different building lots. The radius of the pie-chart is proportional to the number of towers built (Kayali, et al., 2014).

(Custom game, comparative study)

Player Location – Heat Map (1 of 2)

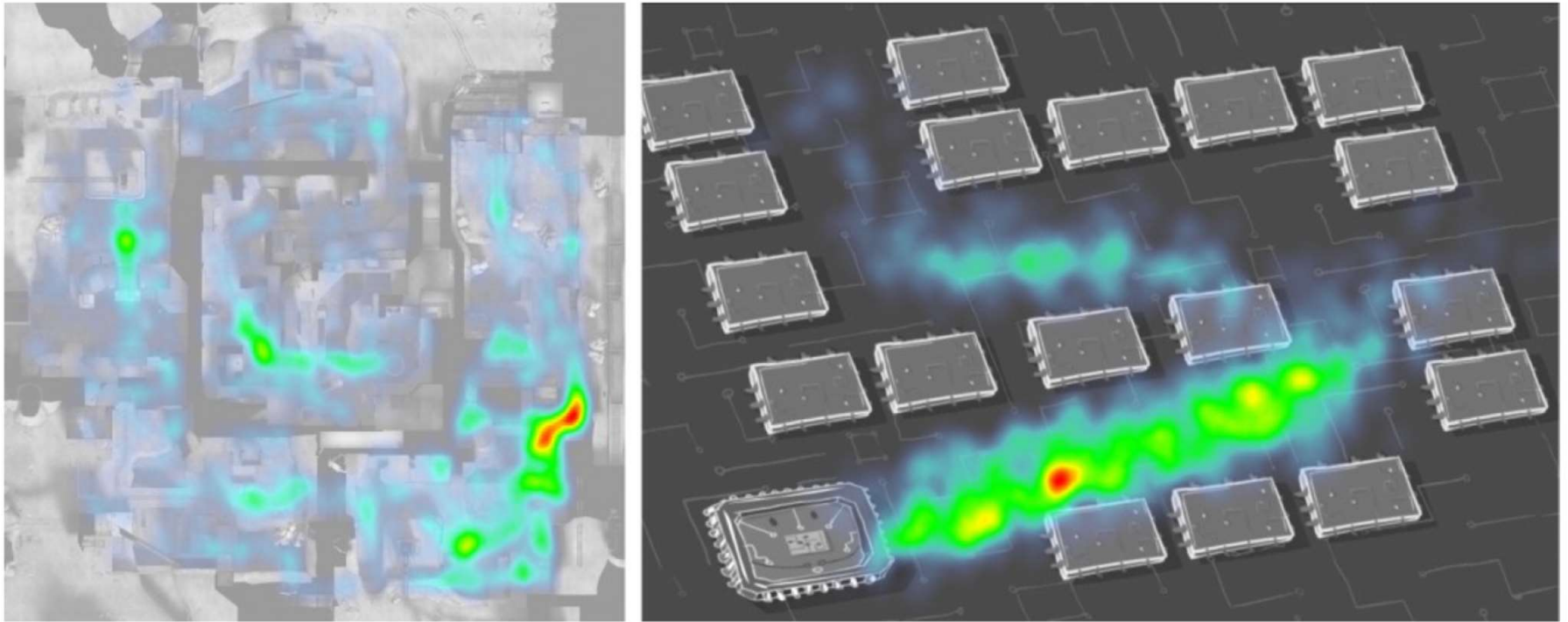
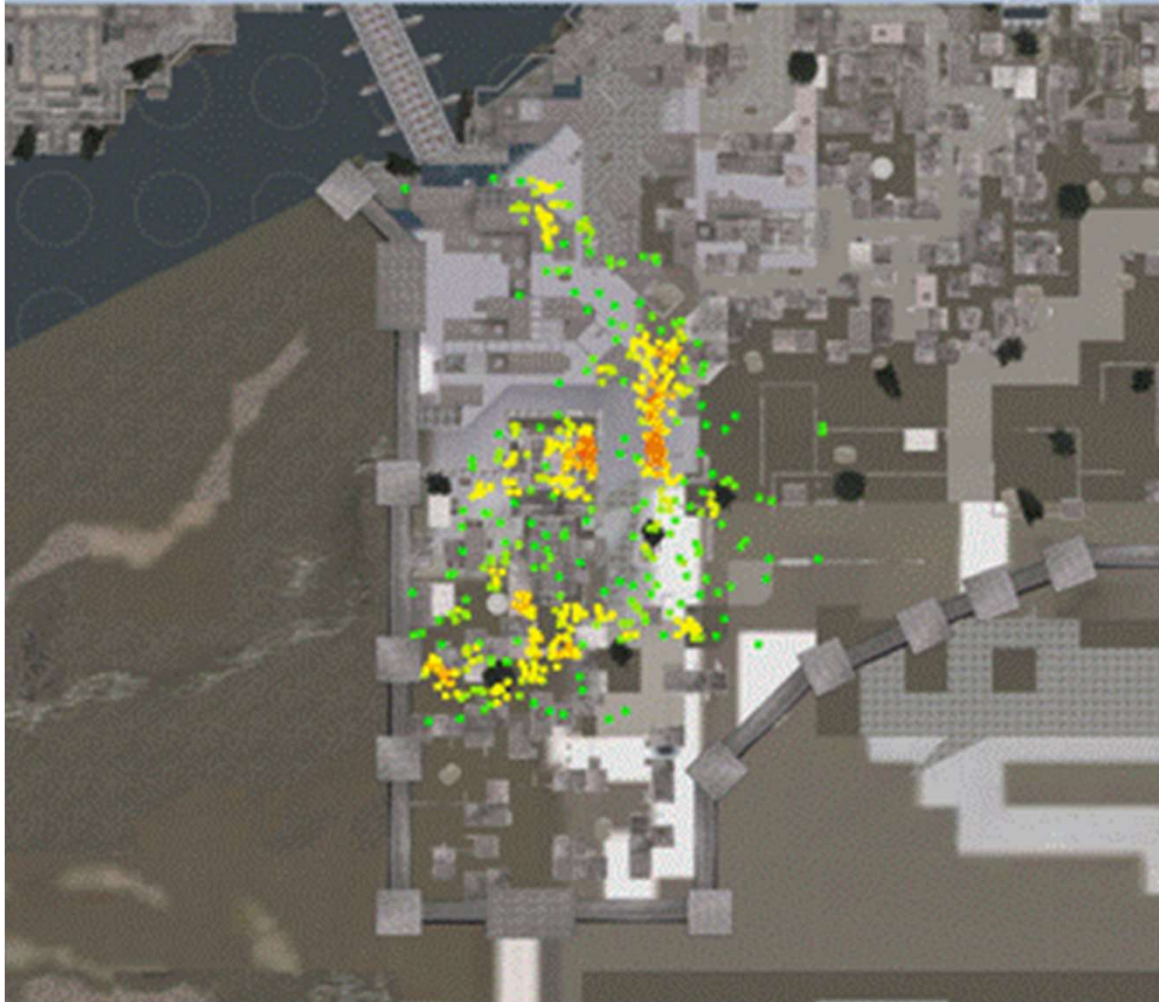


Figure 2. (a) Heatmap of death locations on the Team Fortress 2 map Goldrush. (b) Heatmap showing locations where players of a tower defense game collected coins dropped by defeated enemies (Kayali, et al., 2014).

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](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 2)

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

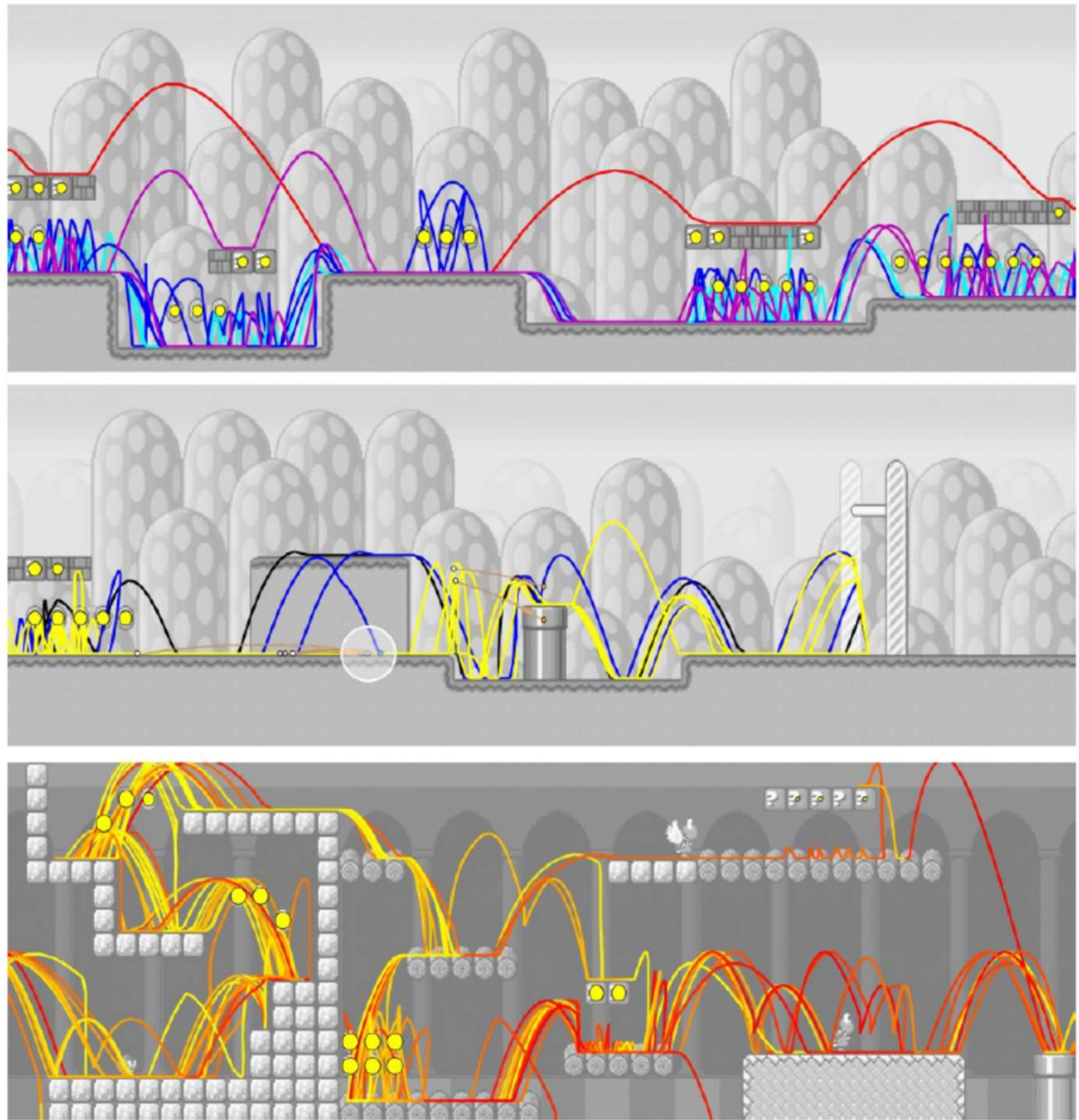


Figure 4. Examples of path visualizations coupled with color-coding to communicate additional information. Top: color coding reflects the reported expertise of players obtained through a pre-game survey. Middle: colors depict the state in which the player's character currently resides in. Bottom: the color-gradient reflects physiological data measured in the form of galvanic skin response (Mirza-Babaei, et al., 2014).

Movement (2 of 2)

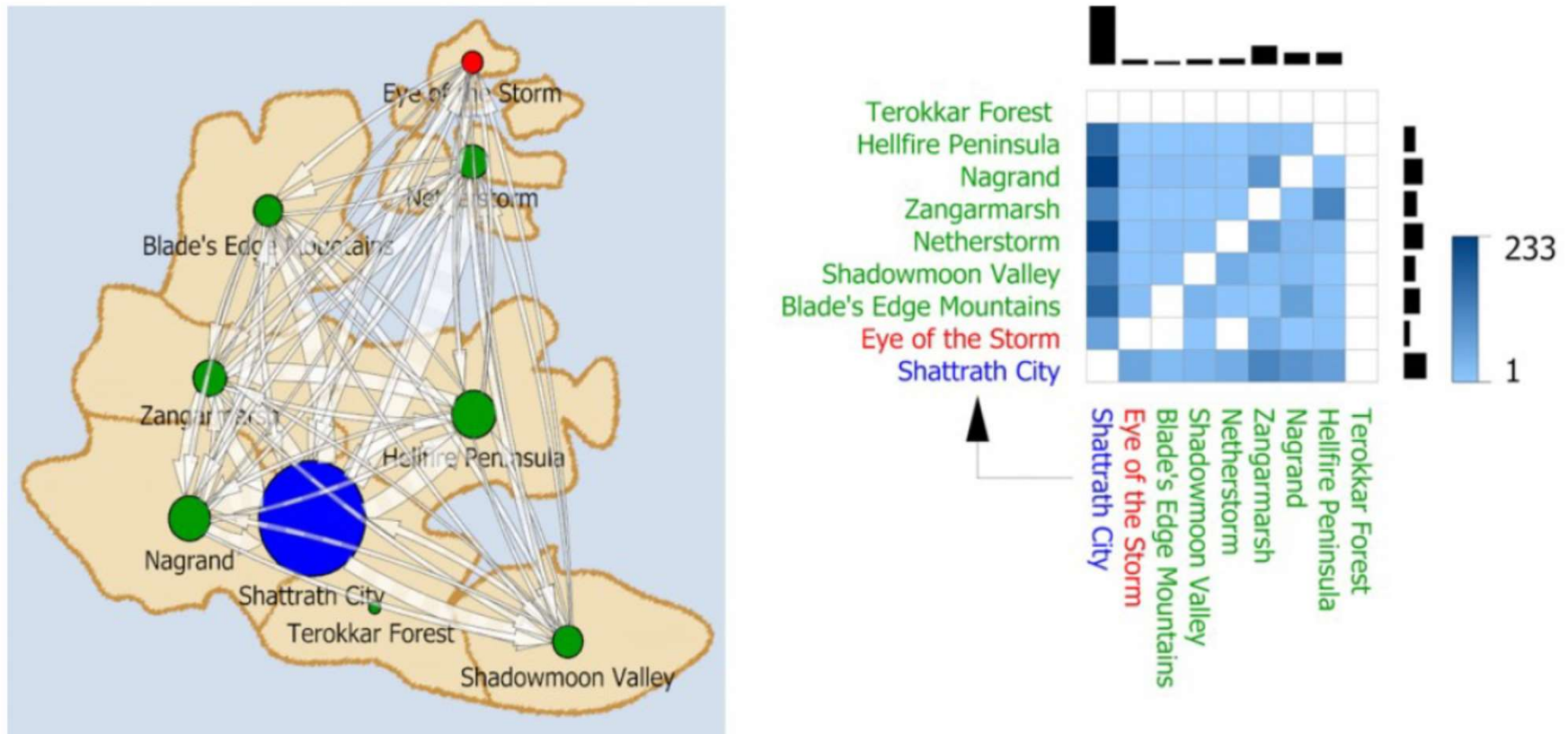
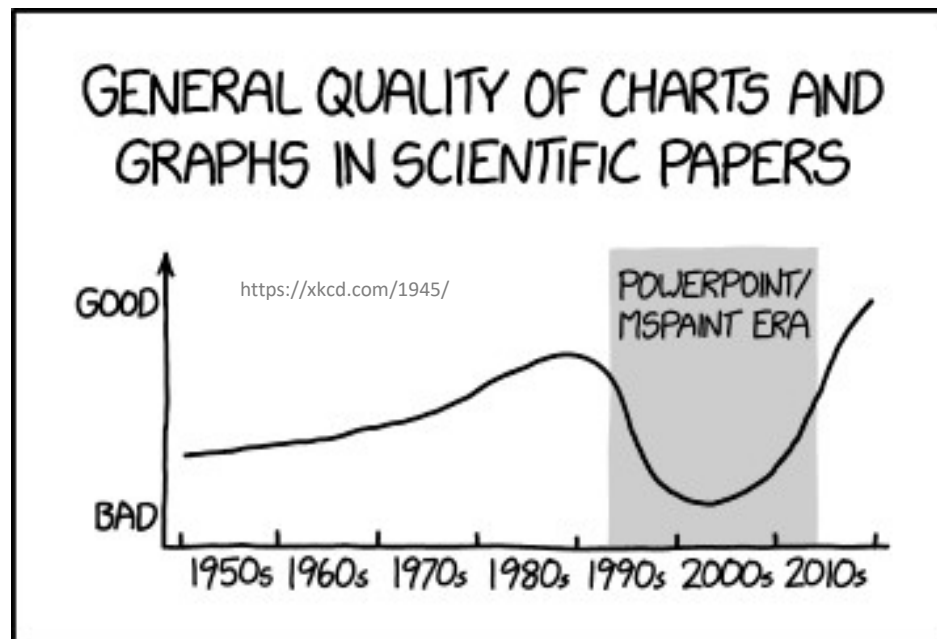


Figure 5. Left: Player movement between regions, cities, and battlegrounds on the World of Warcraft continent Outland. Right: Corresponding matrix view with cells colored according to the number of players moving from one area to another.

Outline

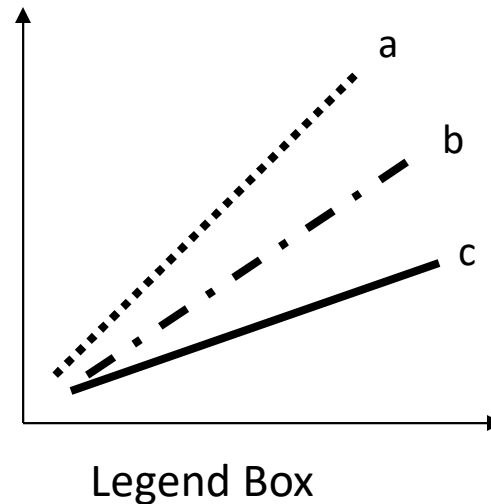
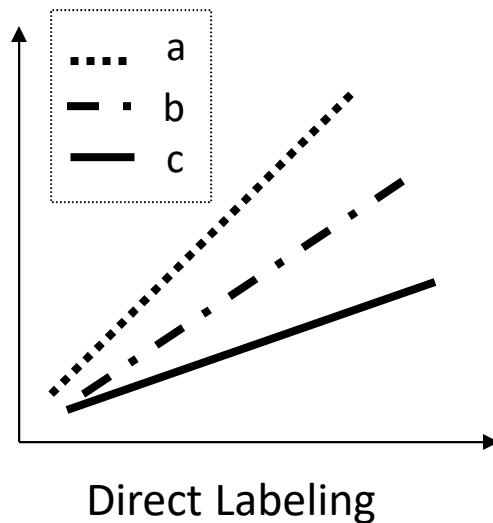
- Types of Charts (done)
- Game Analytics Examples (done)
- Guidelines for Charts (next)



Guidelines for Good Charts (1 of 7)

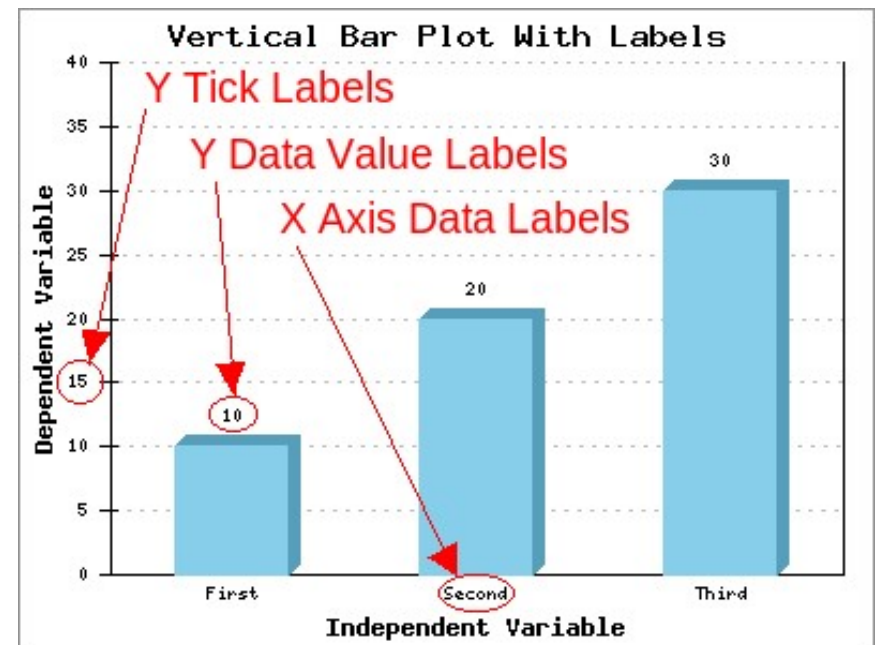
- Require minimum effort from reader
 - Perhaps *most* important metric
 - Given two, can pick one that takes less reader effort

e.g.,



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”



<http://www.phplot.com/phplotdocs/conc-labels.html>

Guidelines for Good Charts (3 of 7)

- Minimize ink

Remove
to improve
(the **data-ink** ratio)

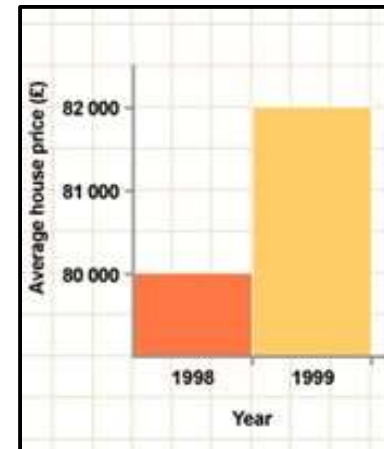
Created by Darkhorse Analytics

www.darkhorseanalytics.com

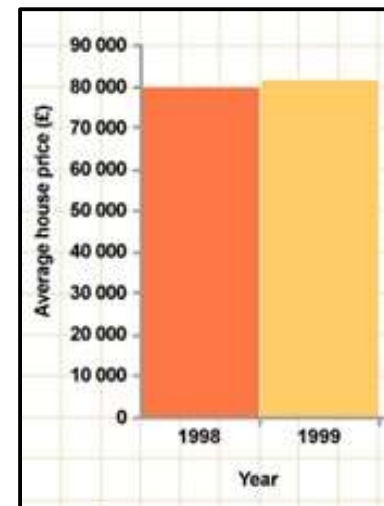
<https://www.slideshare.net/NicoleMarinsek/darkhorse-line-chart>

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 are permitted but require extra effort from reader → so use sparingly!

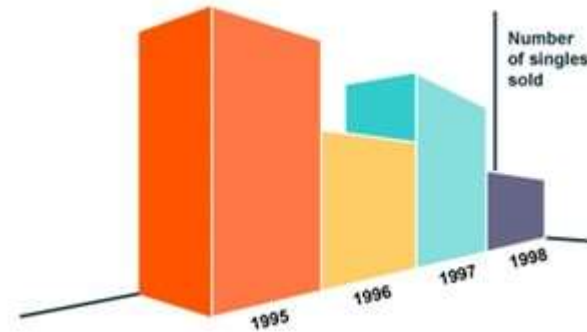


vs.

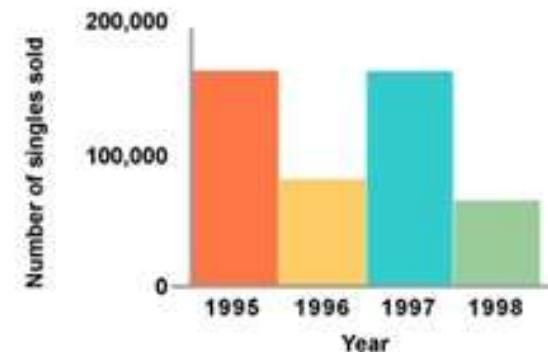


Guidelines for Good Charts (5 of 7)

- 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



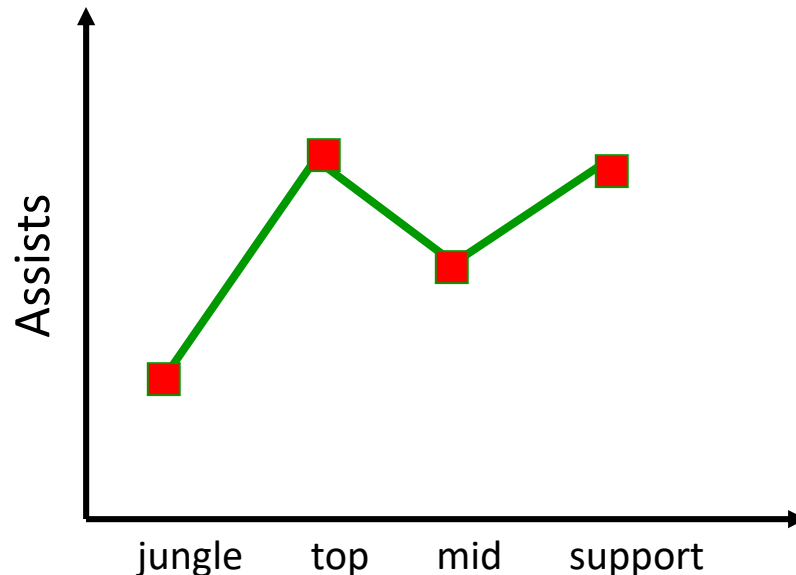
vs.



<http://www.carltonassociatesinc.com/images/confusion-new.jpg>

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 champion types
- Instead, use column chart
- Don't connect with lines

Guidelines for Good Charts (6 of 7)

- Can deceive as easily as can convey meaning

EXAMPLE 2: Amount of Land Planted with Grapes for the Wine Industry.

... they're growing more ...

Amount of land planted with grapes for the wine industry

1997-1998
243,644
acres

1979-1980
172,075
acres

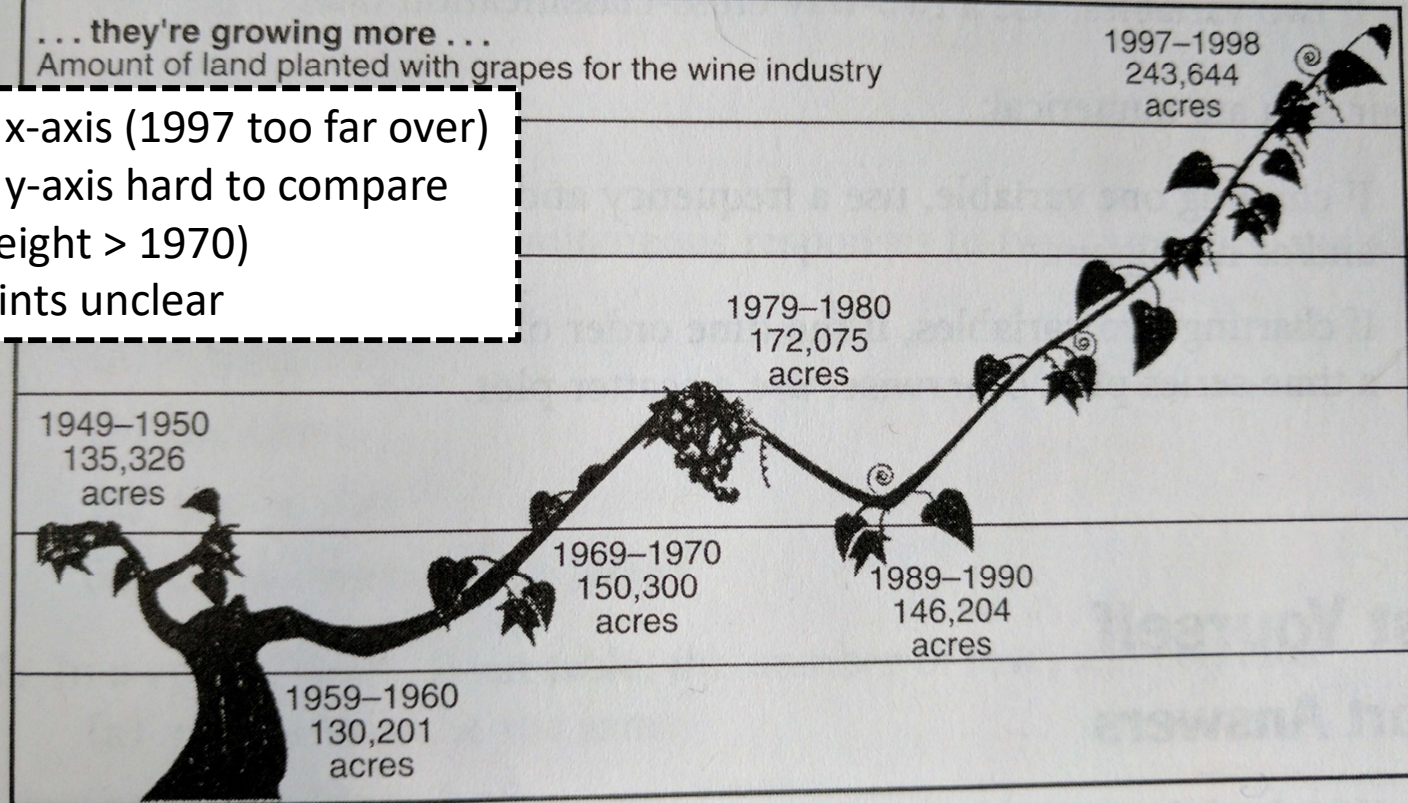
1949-1950
135,326
acres

1969-1970
150,300
acres

1989-1990
146,204
acres

1959-1960
130,201
acres

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

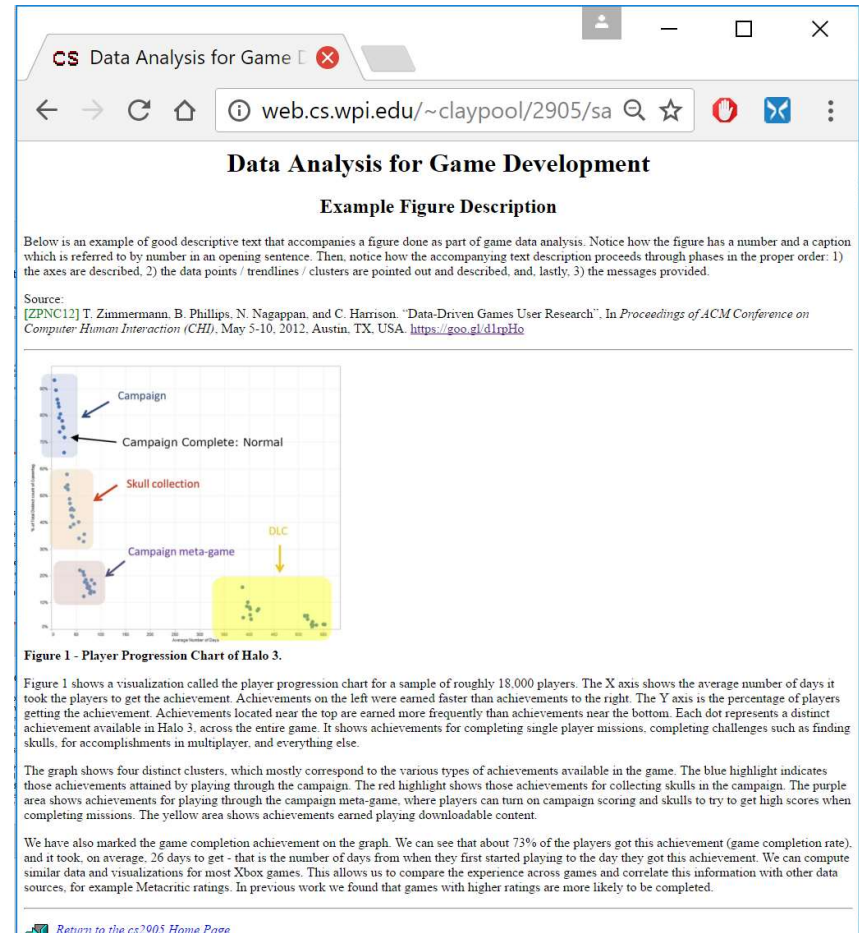


Checklist for Good Charts

- **Axes**
 - Are both axes labeled?
 - Are the axis labels self-explanatory and concise?
 - Are the scale and divisions shown on both axes?
 - Are the min and max ranges appropriate?
 - Are the units indicated?
- **Lines/Curves/Points**
 - Is the number of lines/curves reasonably small?
 - Are curves labeled?
 - Are all symbols clearly distinguishable?
 - Is a concise, clear legend provided?
 - Does the legend obscure any data?
- **Information**
 - If the y-axis is variable, is an indication of spread (error bars) shown?
 - Are grid lines required to read data (if not, then remove)?
- **Scale**
 - Are units increasing left to right (x-axis) and bottom to top (y-axis)?
 - Do all charts use the same scale?
 - Are the scales contiguous?
 - Is bar chart order systematic?
 - Are bars appropriate width, spacing?
- **Overall**
 - Does the whole chart add information to reader?
 - Are there no curves/symbols/text that can be removed and still have the same information?
 - Does the chart have a title or caption (not both)?
 - Is the chart self-explanatory and concise?
 - Do the variables plotted give more information than alternatives?
 - Is chart referenced and discussed in any accompanying report?

Describing Chart in Report & Presentation

- “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 ...”
- Example on Web page



<http://web.cs.wpi.edu/~imgd2905/d20/samples/analysis-example.html>

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

