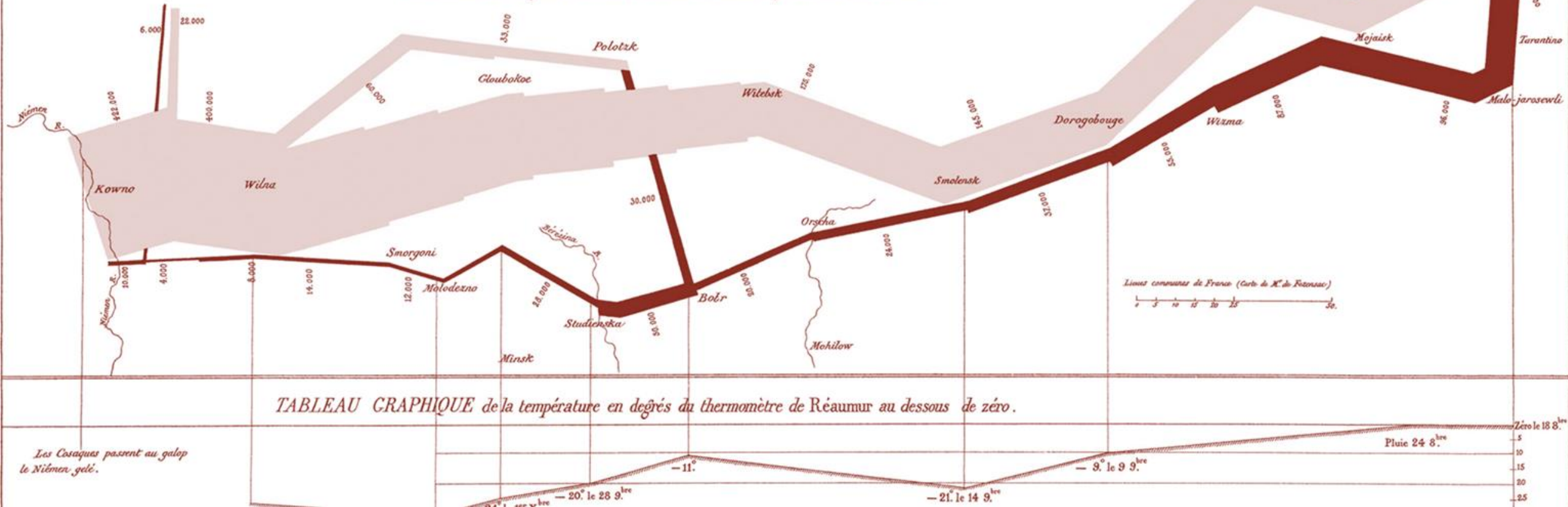


Data Visualization Concepts

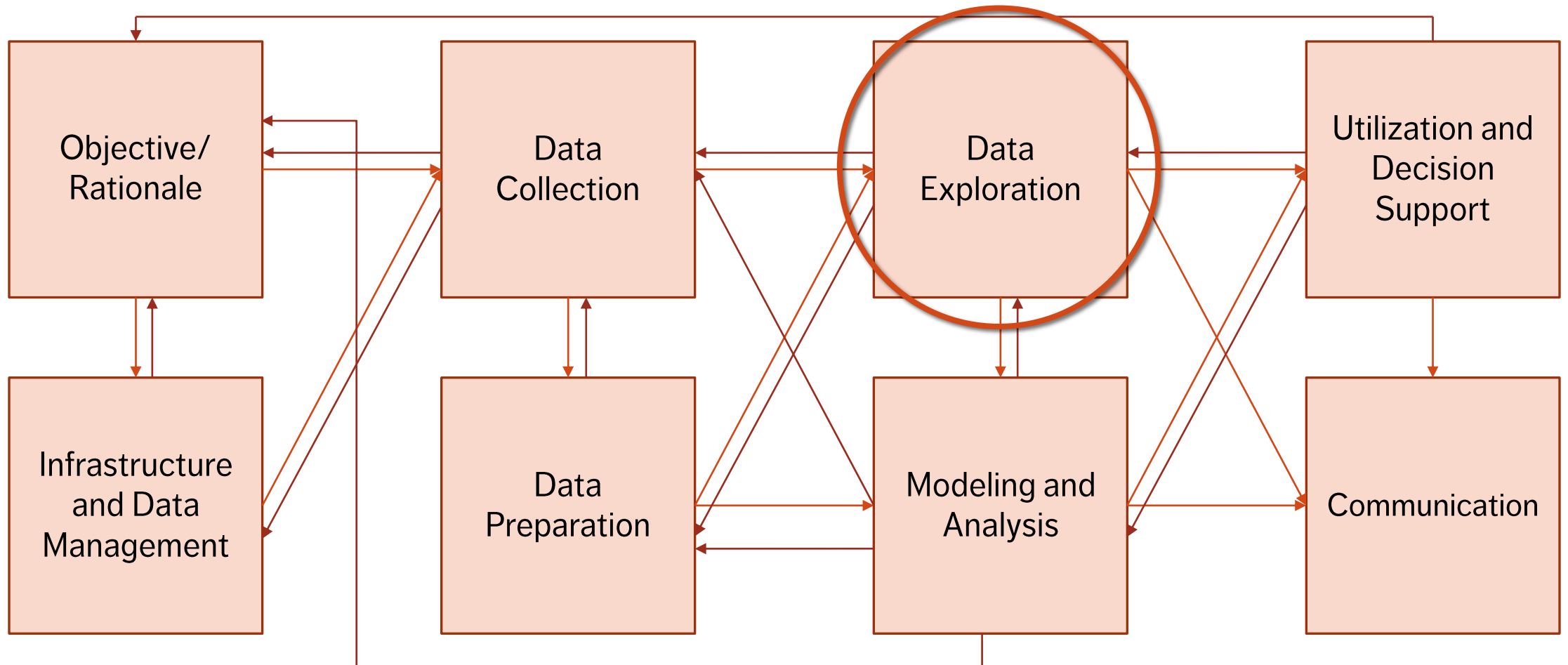
DATA VISUALIZATION AND DASHBOARDS

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Thiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.



1. Exploratory Data Analysis

The (Messy) Analysis Process



Some Basic Questions

What **system** does your data represent – **objects, attributes, relationships**?

How does it represent this system – i.e., the **data model**?

Who made this dataset? When? For what purpose?

Assuming a flat file – what do the rows and columns represent?

Do you have enough information (e.g., **metadata**) to answer these questions?

Where can you find more information?

Non-Visualization Summaries

Cl	N03	NH4
Min. : 0.222	Min. : 0.000	Min. : 5.00
1st Qu.: 10.994	1st Qu.: 1.147	1st Qu.: 37.86
Median : 32.470	Median : 2.356	Median : 107.36
Mean : 42.517	Mean : 3.121	Mean : 471.73
3rd Qu.: 57.750	3rd Qu.: 4.147	3rd Qu.: 244.90
Max. : 391.500	Max. : 45.650	Max. : 24064.00
NA's : 16	NA's : 2	NA's : 2

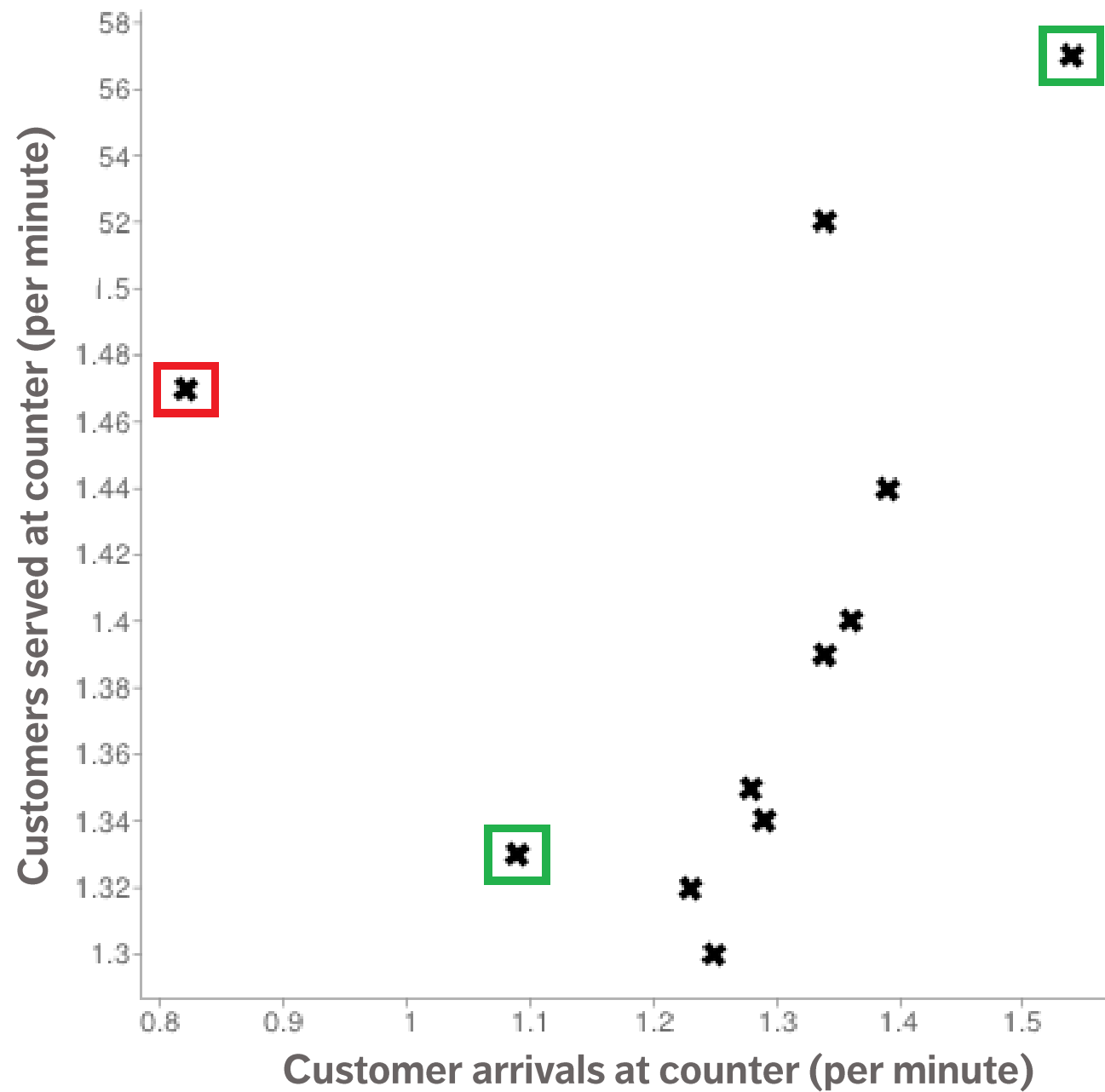
season
 Length: 340
 Class : character
 Mode : character

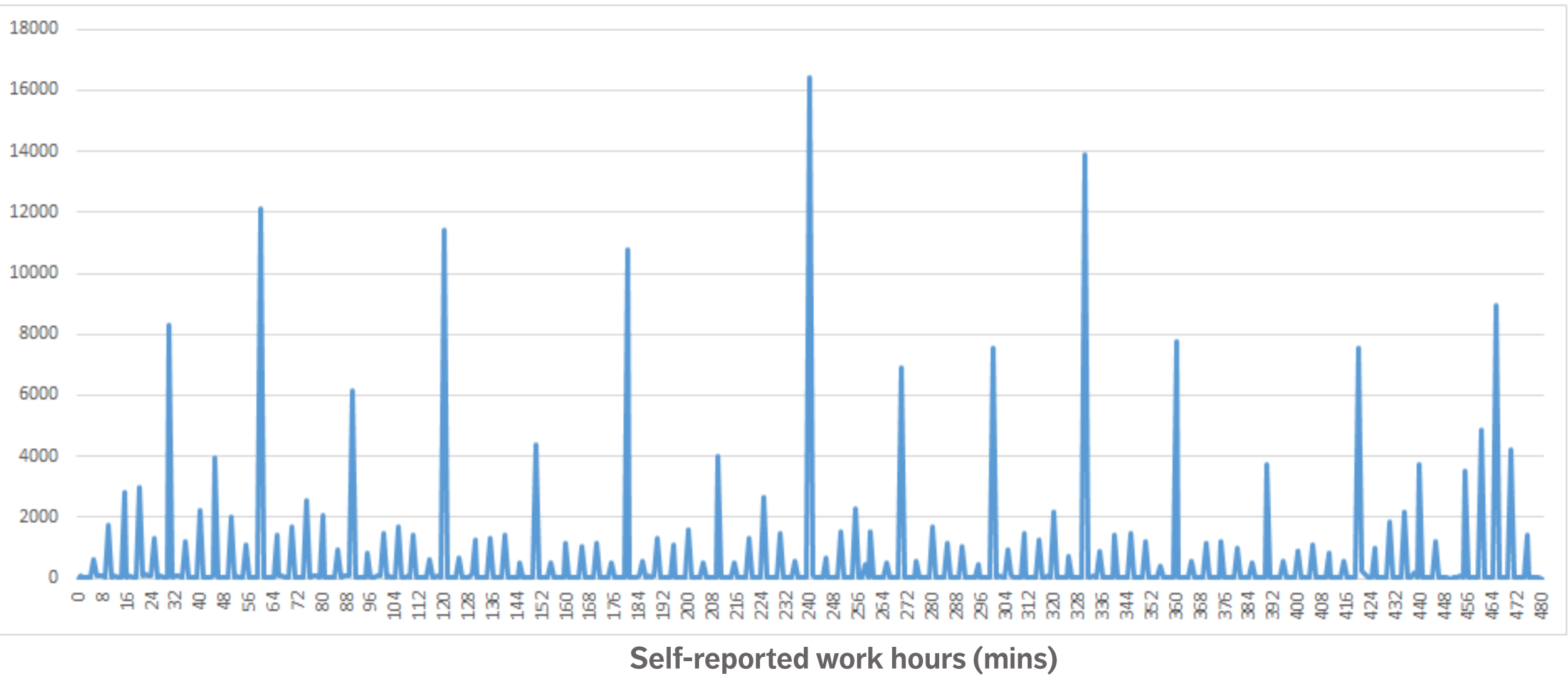
autumn spring summer winter
 80 84 86 90

Pre-Analysis Use

Data visualization can be used to set the stage for analysis:

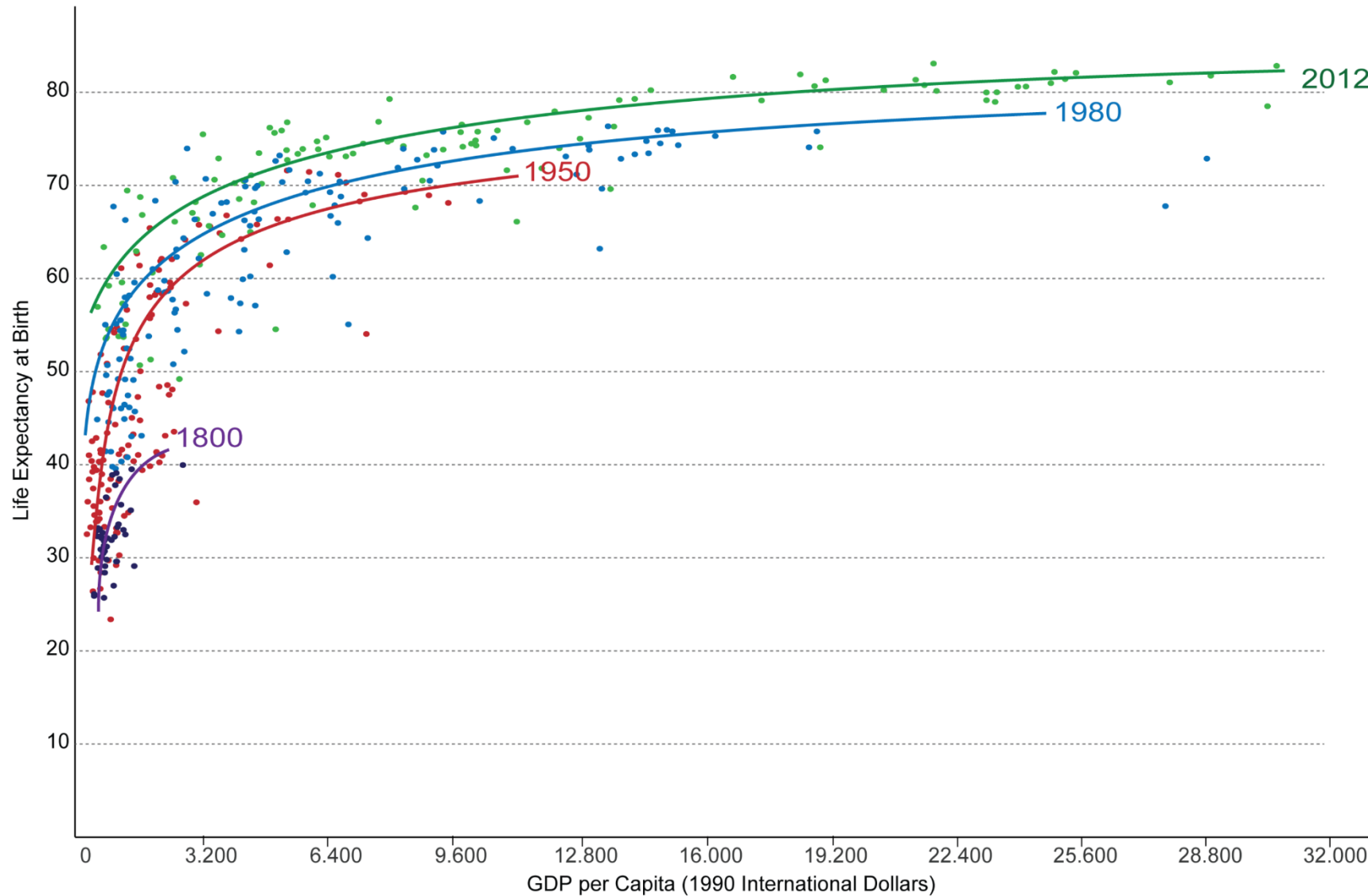
- **detecting anomalous entries**
invalid entries, missing values, outliers
- **shaping the data transformations**
binning, standardization, Box-Cox transformations, PCA-like transformations
- **getting a sense for the data**
data analysis as an art form, exploratory analysis
- **identifying hidden data structure**
clustering, associations, patterns informing the next stage of analysis





Life Expectancy vs. GDP per Capita from 1800 to 2012 – by Max Roser

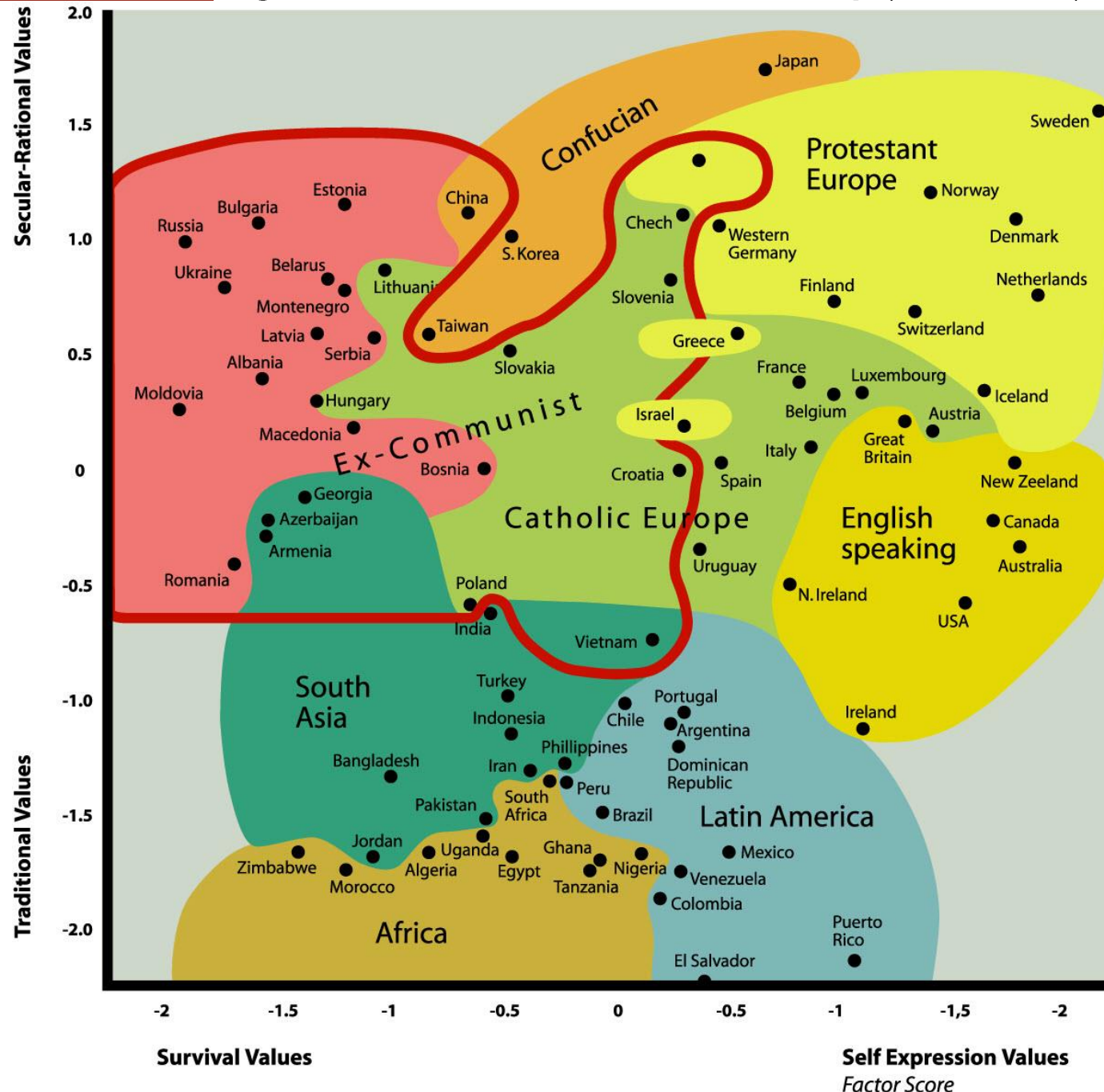
GDP per capita is measured in International Dollars. This is a currency that would buy a comparable amount of goods and services a U.S. dollar would buy in the United States in 1990. Therefore incomes are comparable across countries and across time.



This graph displays the correlation between life expectancy and GDP per capita.

Countries with higher GDP have a higher life expectancy, in general.

The relationship seems to follow a logarithmic trend: the unit increase in life expectancy per unit increase in GDP decreases as GDP per capita increases.



Traditional values

importance of religion, parent-child ties, deference to authority and traditional family values.

Secular-rational values

less emphasis on religion, traditional family values and authority.

Survival values

emphasis on economic and physical security.

Self-expression values

high priority to environmental protection, growing tolerance of foreigners, gays and lesbians and gender equality

Workhorse Data Exploration Charts

Text and Tables

Rug Charts/Number Lines

Histograms/Bar Charts

Boxplots

Line Graphs

Scatterplots

Line Chart/Rug Chart

Gaps in the number line: **absence** of those numeric values in the data.

Remember: this is (possibly) different from the order that values appear in the dataset – since it is a number line, it shows where the values fall numerically.

If some values are identical, they lie on top of each other (use **jitter**?).

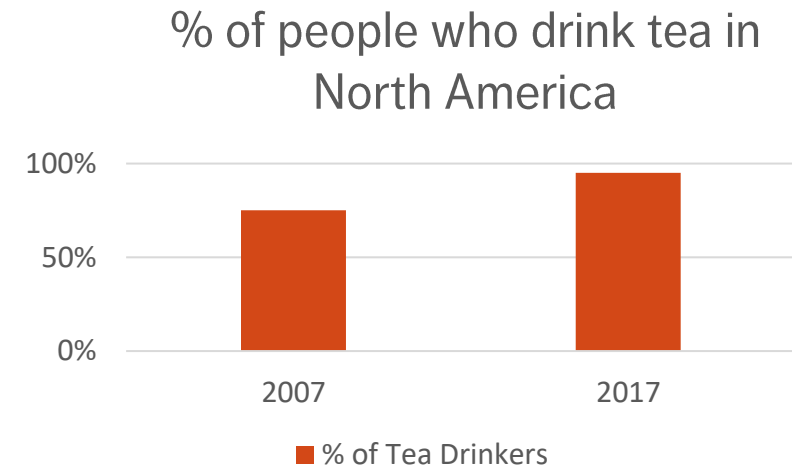


Simple Text

One or two numbers to focus on.

Good at “setting the scene”.

Draws focus to an area of the report.



95% of the population
drinks tea today compared to
75% in 2007

Tables

Tables interact with our **verbal** system, which means we **read** them:

- used to compare values
- audiences will look for their rows

Table design needs to **blend** into background

- the data should stand out, not the borders
- dense table/data: use alternating row colour

Name	Last Year	This Year
Bob	20	30
Fred	30	40
George	10	15

Name	Last Year	This Year
Bob	20	30
Fred	30	40
George	10	15

Table Heatmaps

Leverage colour to convey magnitude

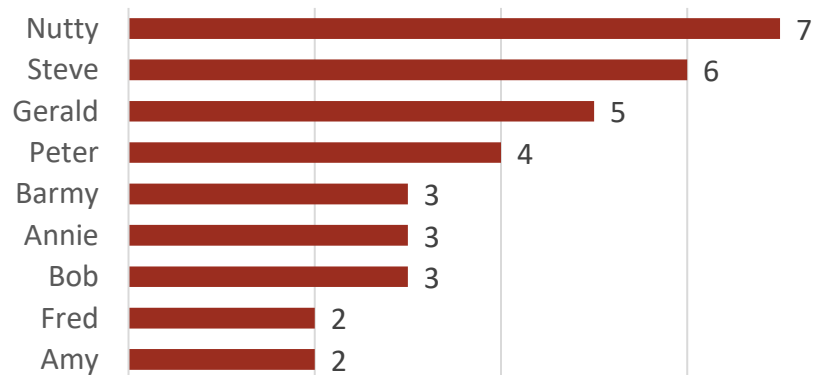
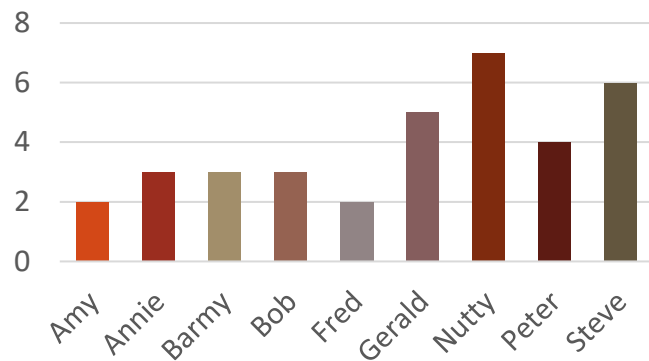
- use **single colour saturation** rather than differentiation (different colours)
- with a legend (white = low, blue = high), numbers can be removed without altering the message

	Last Year	This Year	Next Year	Optimum
George	20	20	20	20
Peter	40	35	30	25
John	10	10	5	5
Sandra	25	30	35	40

	Last Year	This Year	Next Year	Optimum
George	20	20	20	20
Peter	40	35	30	25
John	10	10	5	5
Sandra	25	30	35	40

	Last Year	This Year	Next Year	Optimum
George				
Peter				
John				
Sandra				

Bar Charts



Very versatile and useful.

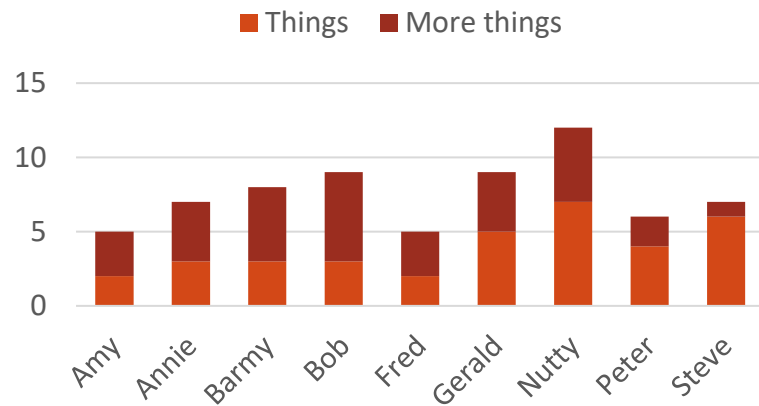
ALWAYS (?) have a zero baseline.

Use graph axis OR data labels. Axis for broad statements, data labels for more detail.

Horizontal charts are apparently **easier to read** (according to many studies).

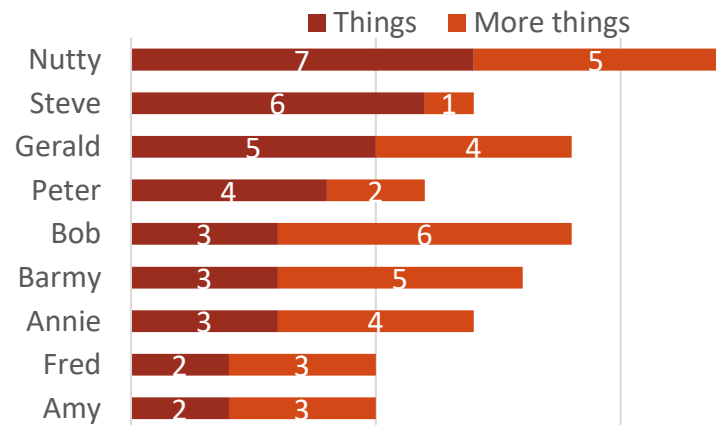
Think about the ordering of categories.

Stacked Bar Charts



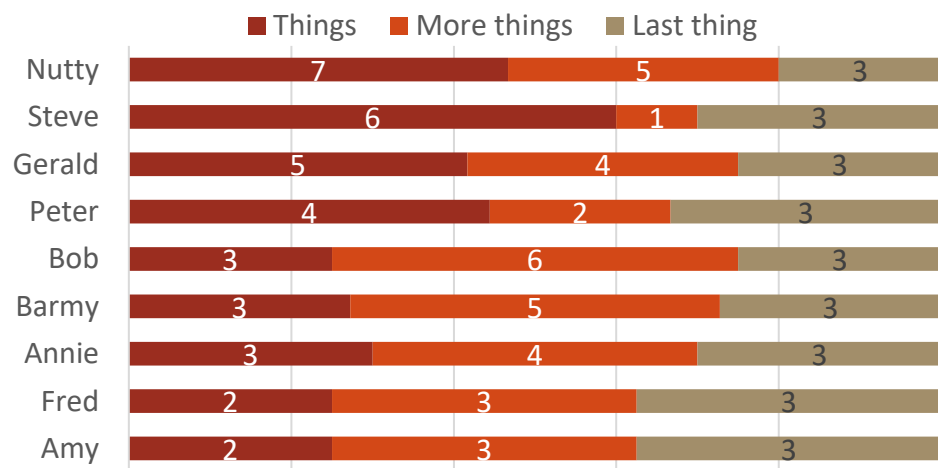
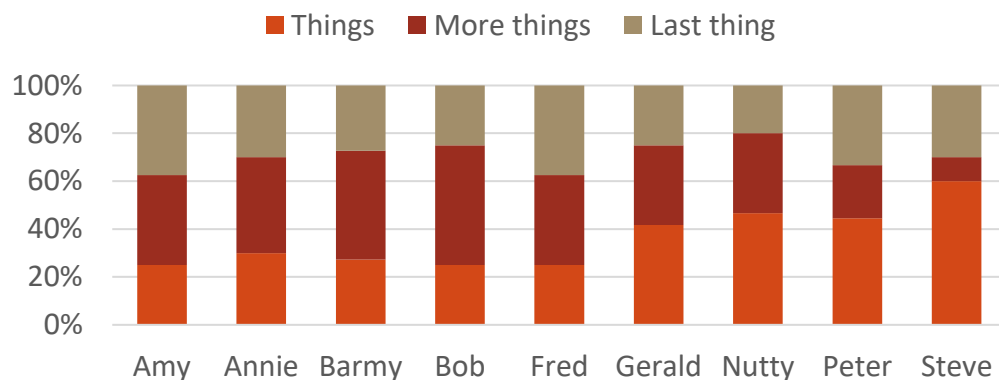
Designed for **comparing totals**, but can quickly become **overwhelming**.

Hard to sort / order.



Filtering is complicated in Power BI (what do you click on & how the chart responds when filter is clicked on?)

100% Bar Charts



Work well for visualizing **proportions** of a whole on a scale from negative to positive.

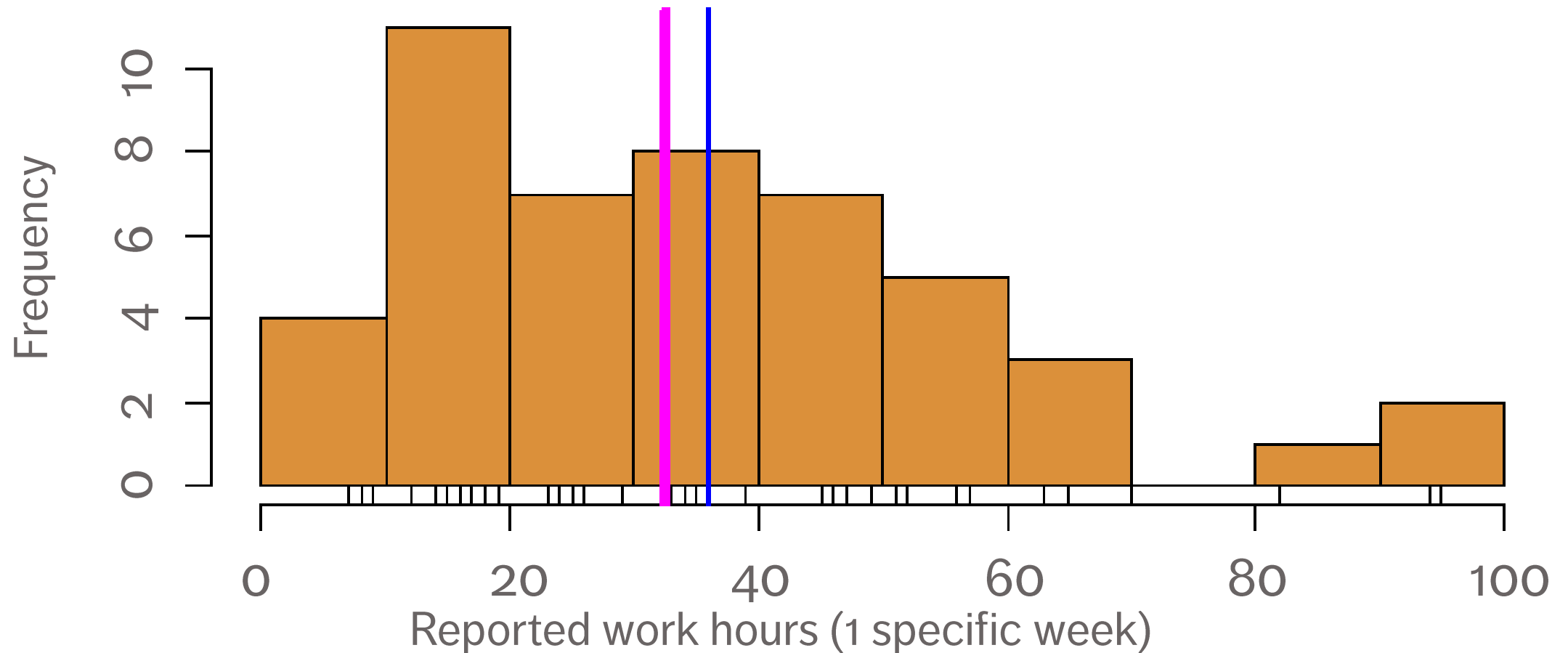
Consistent baseline on far left and right.

Easy to compare.

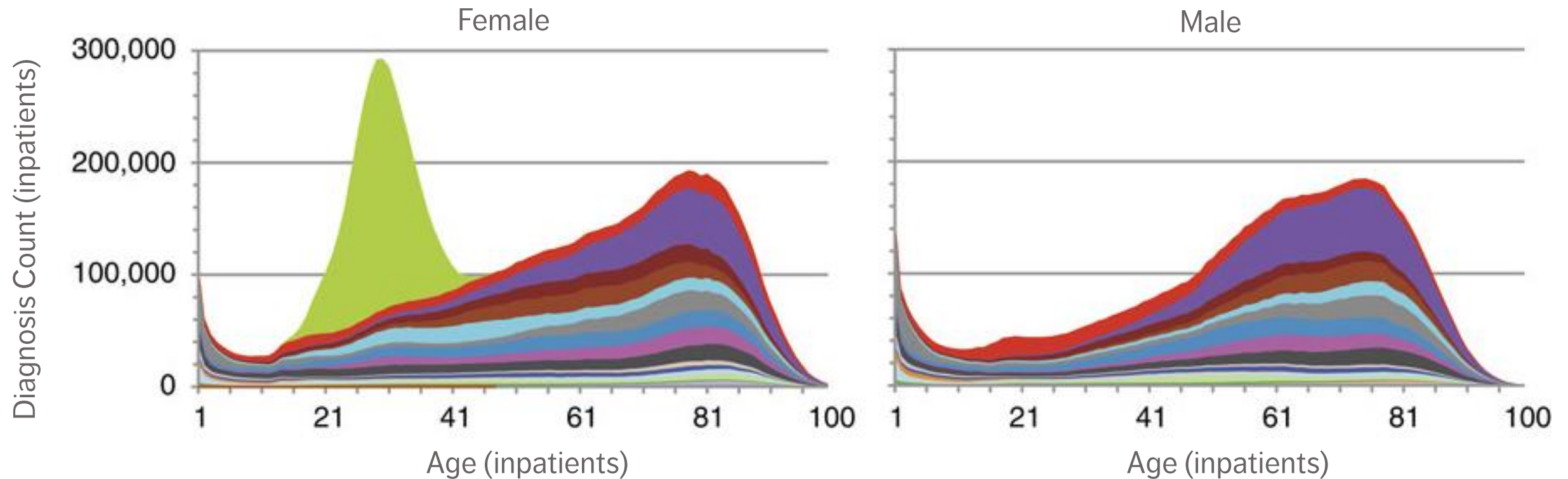
No relative measure to **magnitude** of data.

Research shows that horizontal is easier to process than vertical.

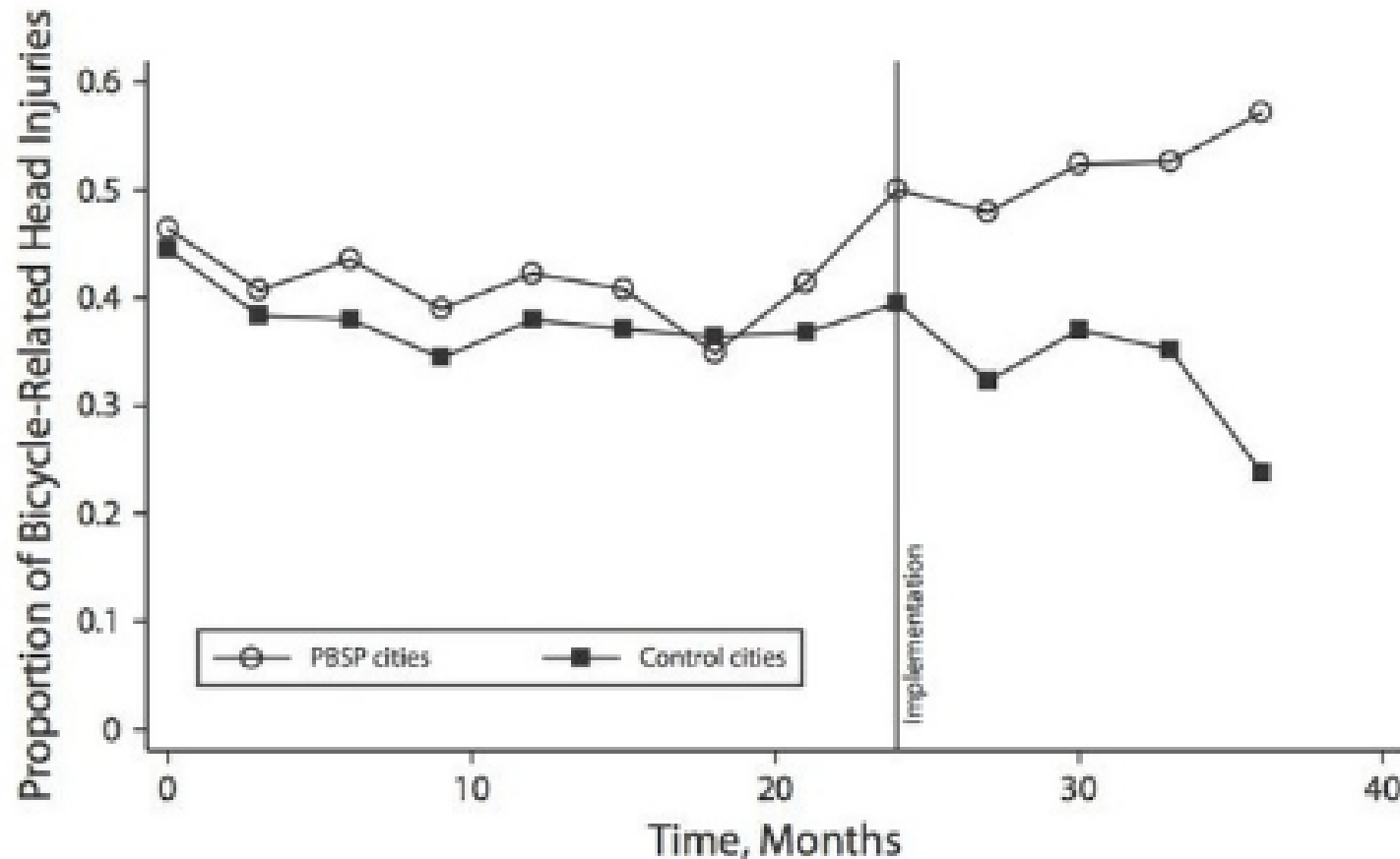
Histogram



Stacked Histograms

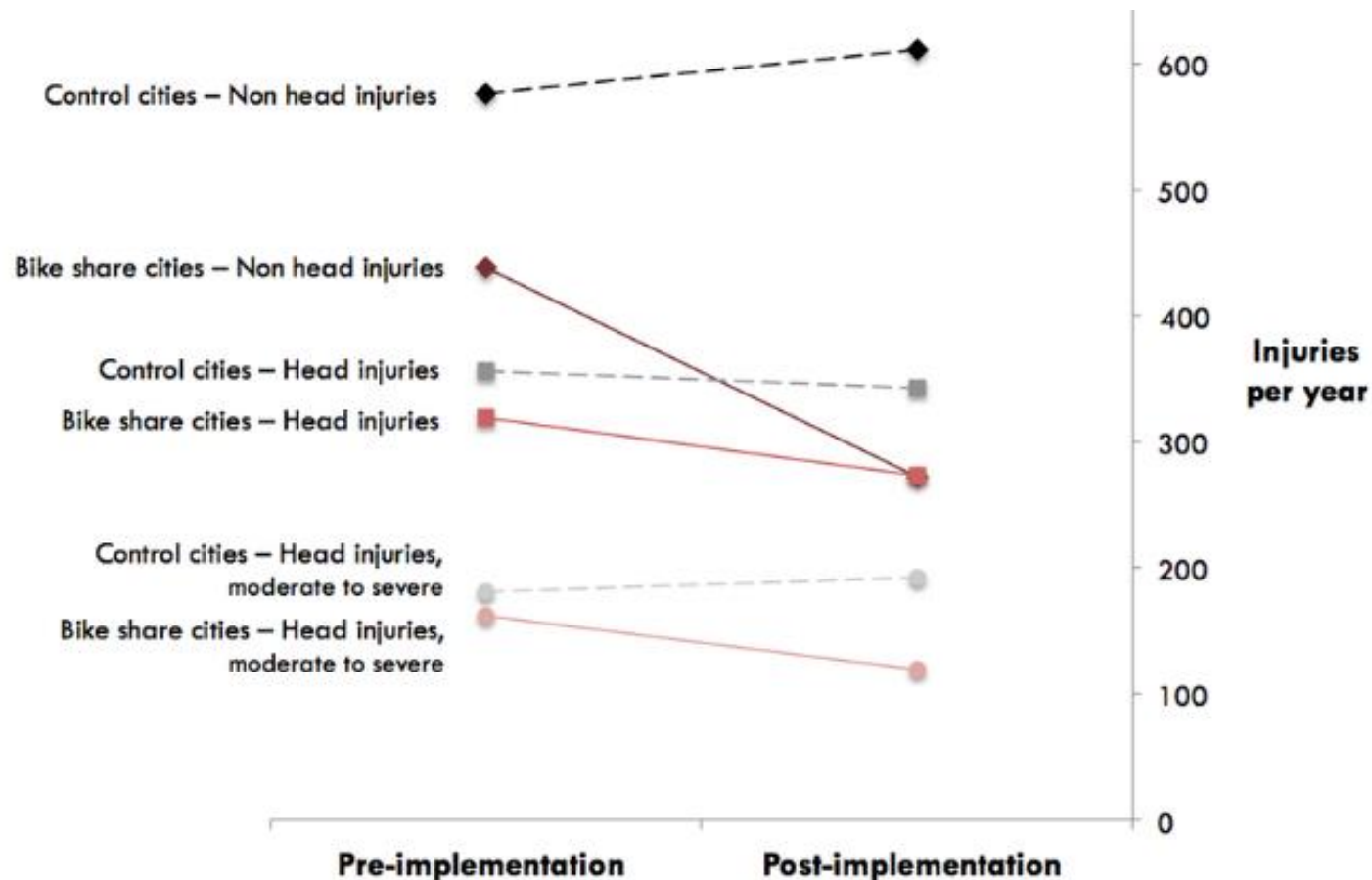


Line Graphs



Proportion of all bicycle-related injuries that were classified as head injuries among cities with public bike share programs and control cities, centered on intervention date (vertical line); North America.

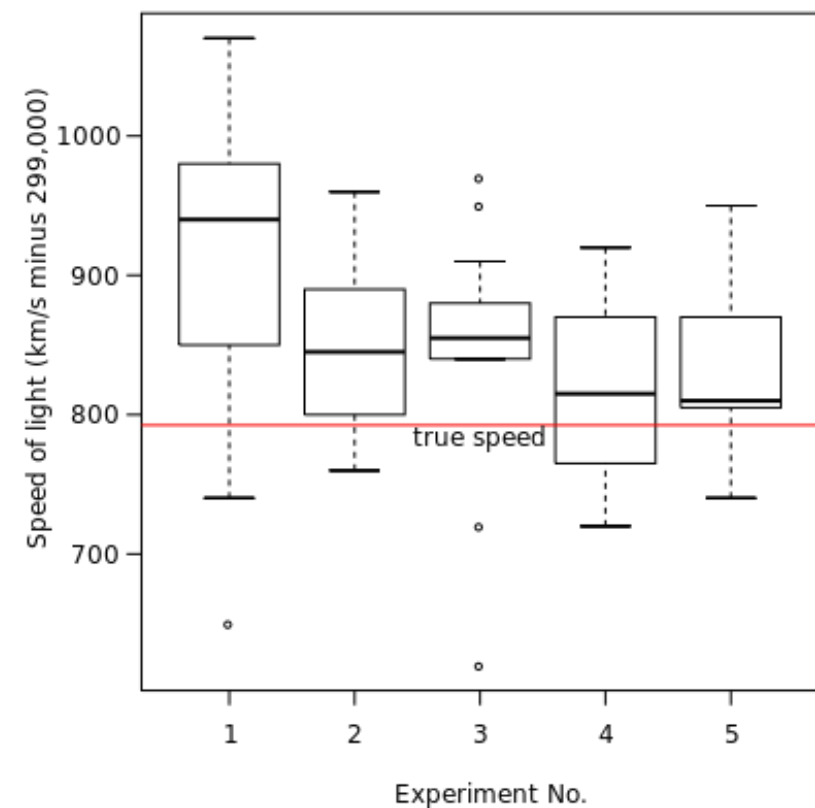
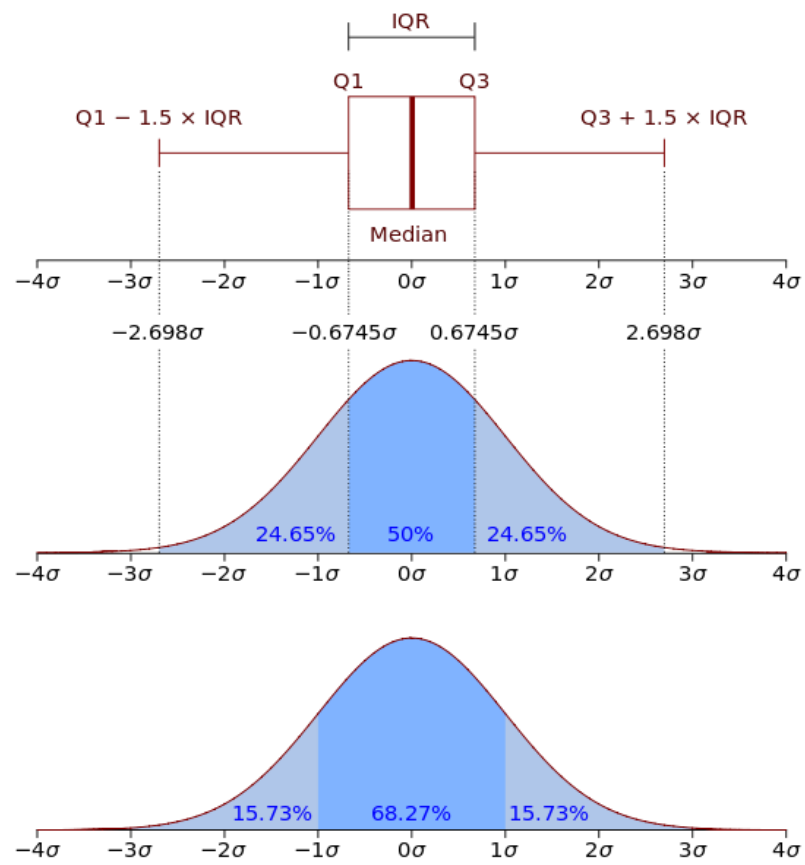
Line Graphs



Data from new study show declines in all injuries, including head injuries after bike share system implemented.

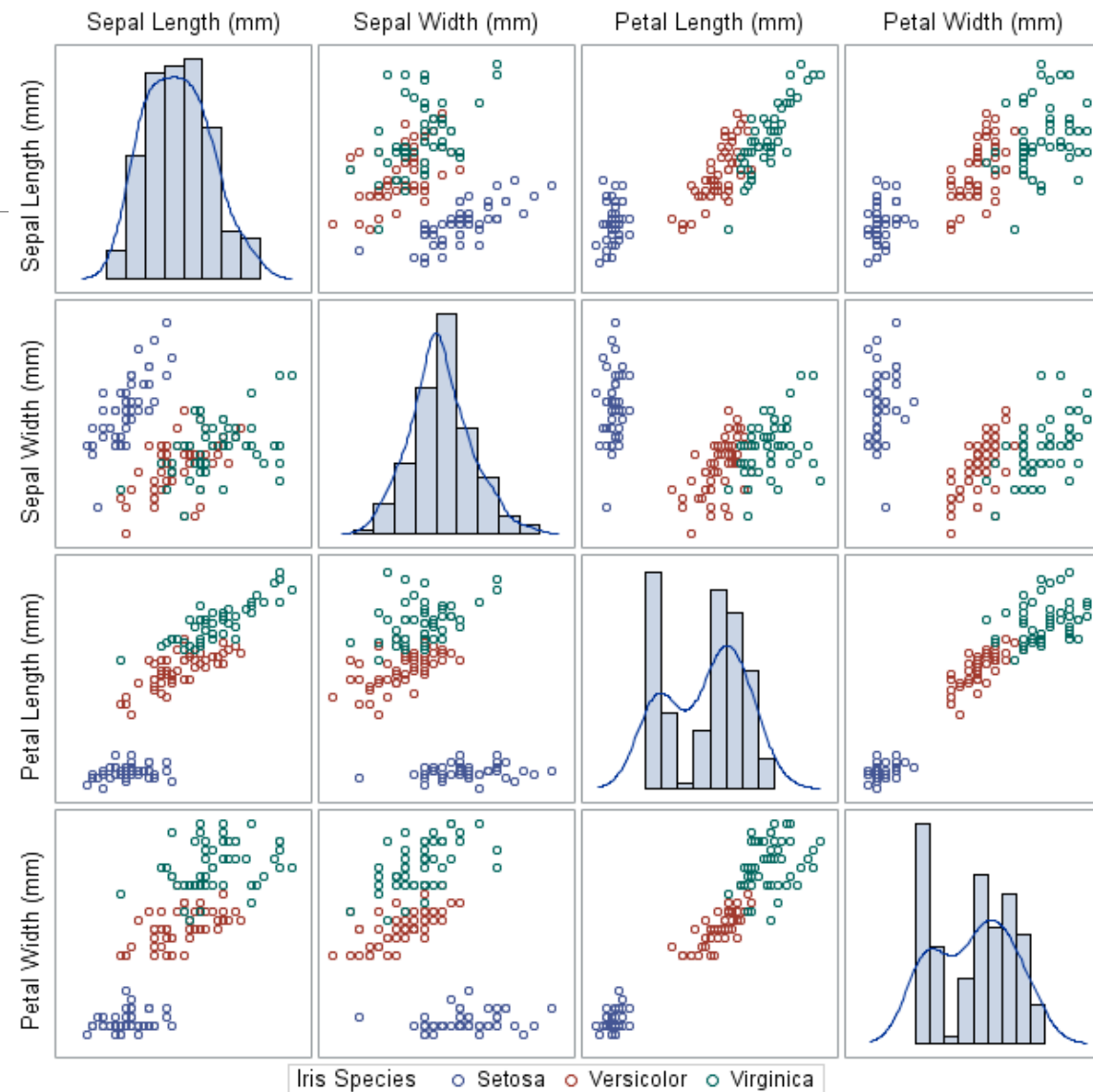
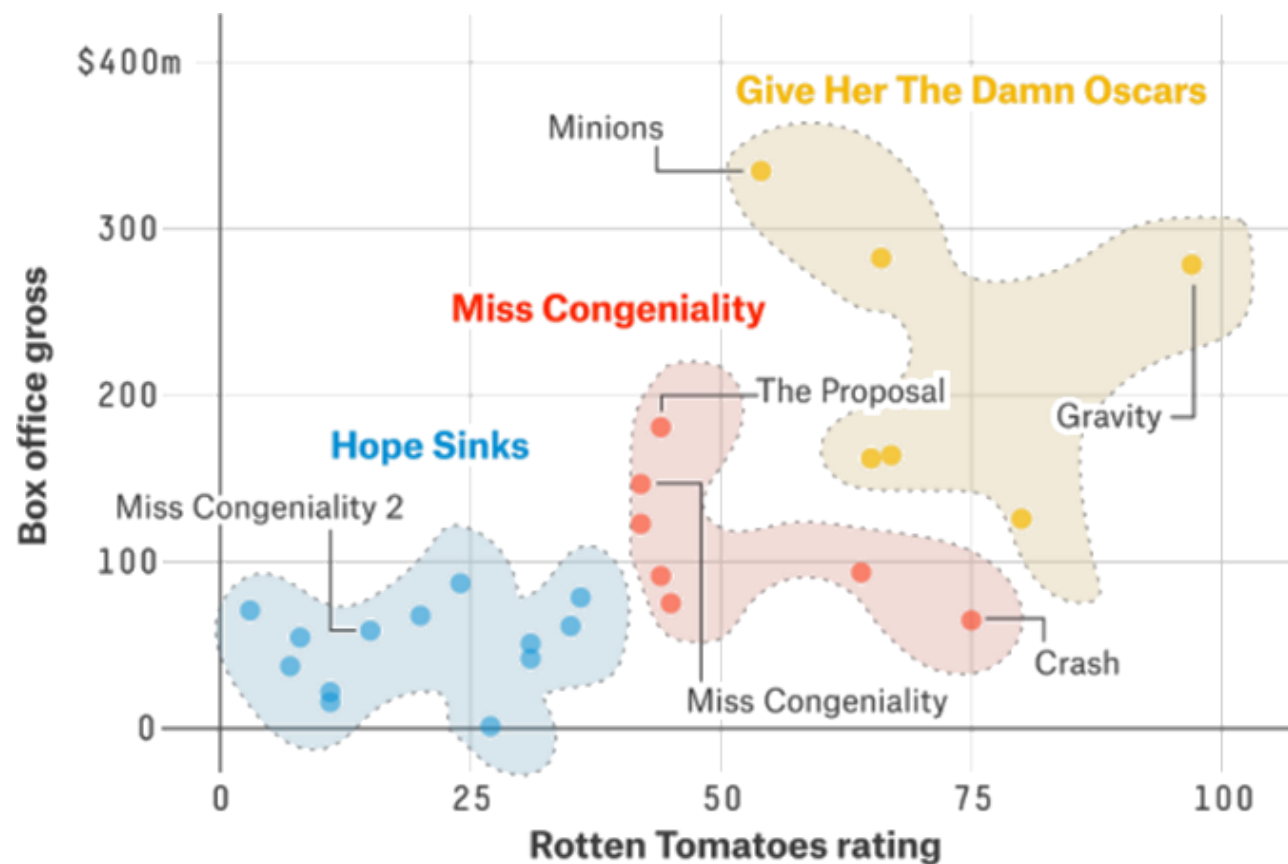
Because head injuries decline less than other injuries, they are now a larger proportion of all injuries.

Boxplots



Michelson-Morley Experiment to Determine the Speed of Light

Scatterplots



Suggested Reading

Exploratory Data Analysis

Data Understanding, Data Analysis, Data Science **Volume 2: Fundamentals of Data Insight**

18. Data Exploration and Data Visualization

18.1 Data and Charts

- Pre-Analysis Uses

The Practice of Data Visualization **Part I: A Data Exploration Primer**

2. Data Exploration and Visualization

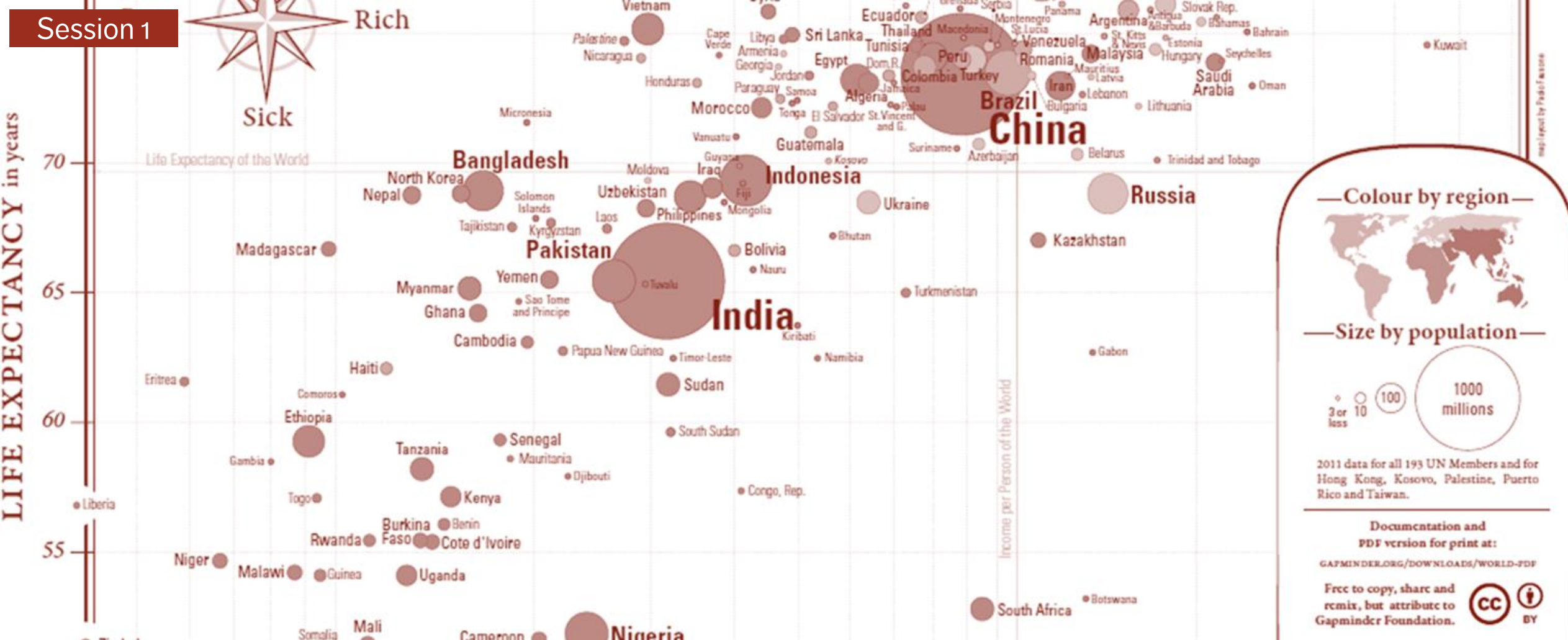
2.1 Exploratory Data Analysis

2.2 Workhorse Data Visualizations

Exercises

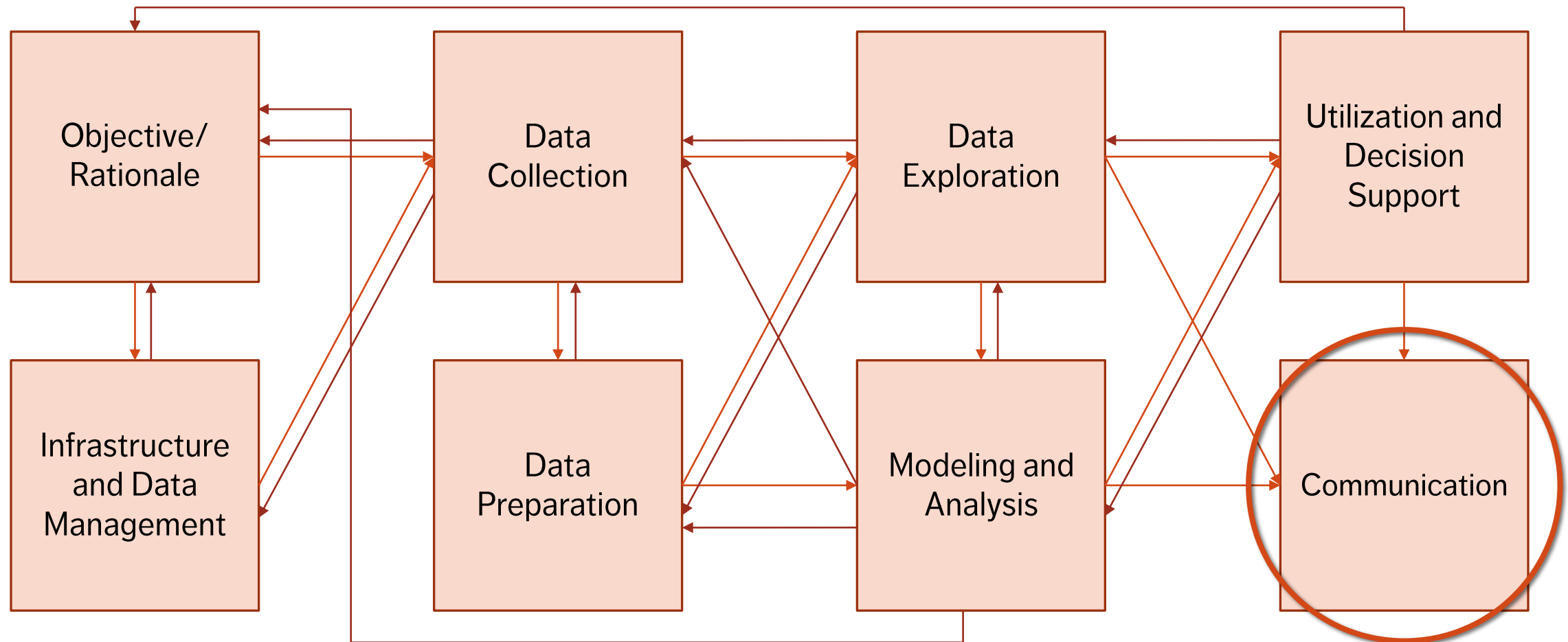
Exploratory Data Analysis

1. Find examples of data presentations that you consider to be particularly insightful and/or powerful. Discuss their strengths/weaknesses.
2. Find examples of data presentations that you consider to be particularly misleading and/or useless. Discuss their strengths/weaknesses.
3. How do you think new technologies (e.g. virtual or augmented reality, 3D-printing, wearable computing) will influence data presentations?



2. Data Visualization for Communication

The (Messy) Analysis Process



Principles of Analytical Design

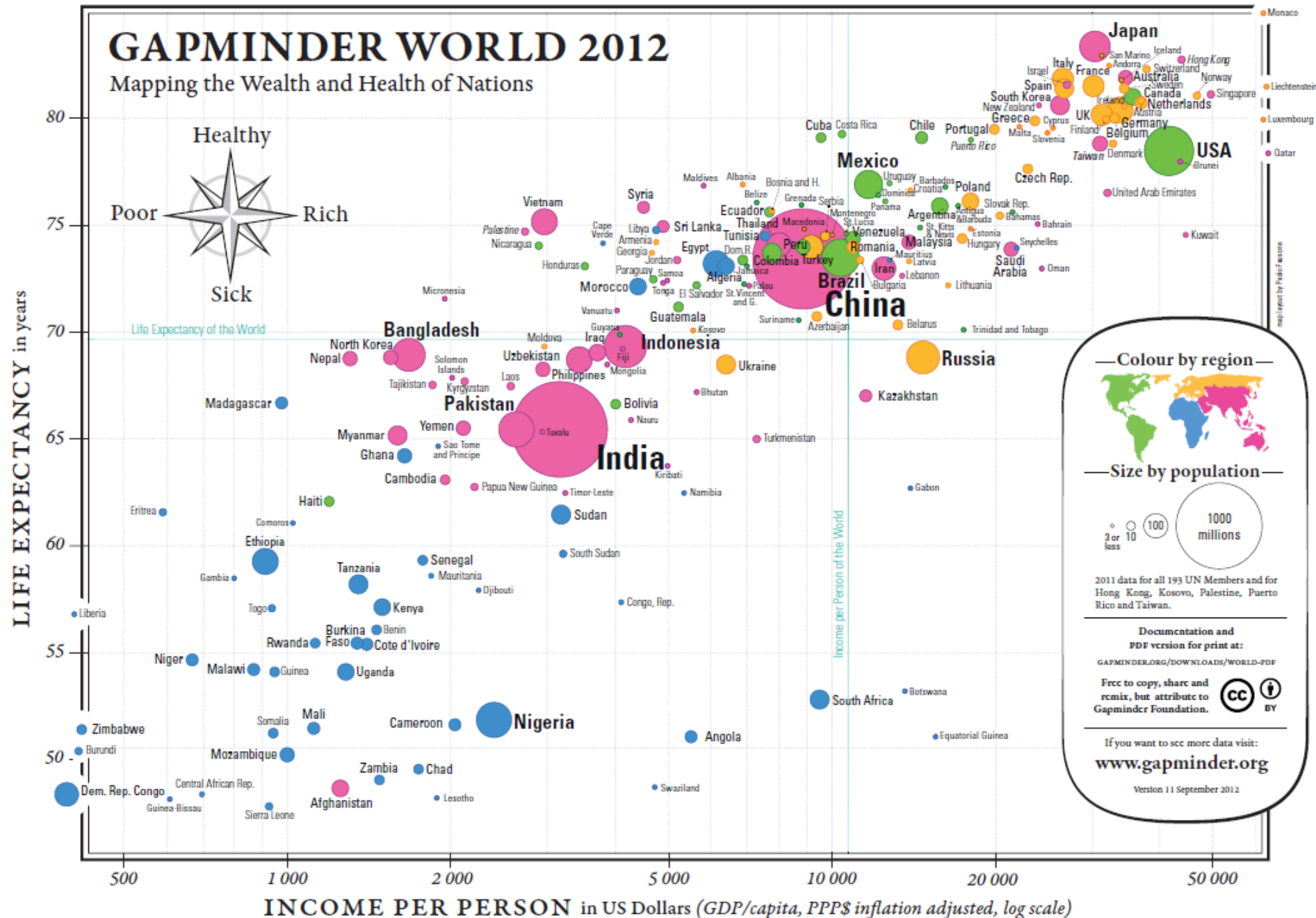
Reasoning and **communicating** our thoughts are intertwined with our lives in a causal and dynamic multivariate Universe.

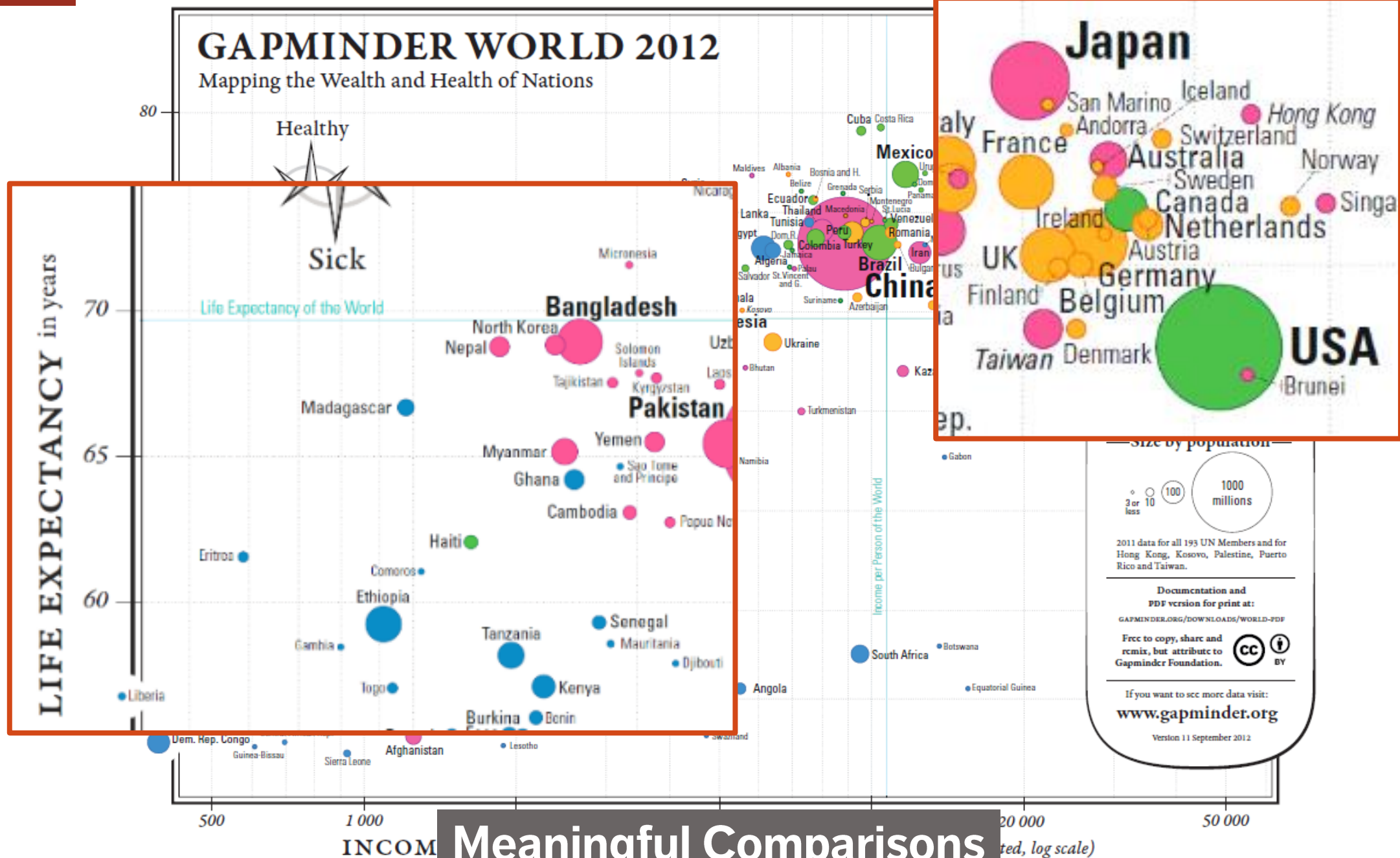
There is a **symmetry** to visual displays of evidence: consumers should be seeking exactly what producers should be providing, namely

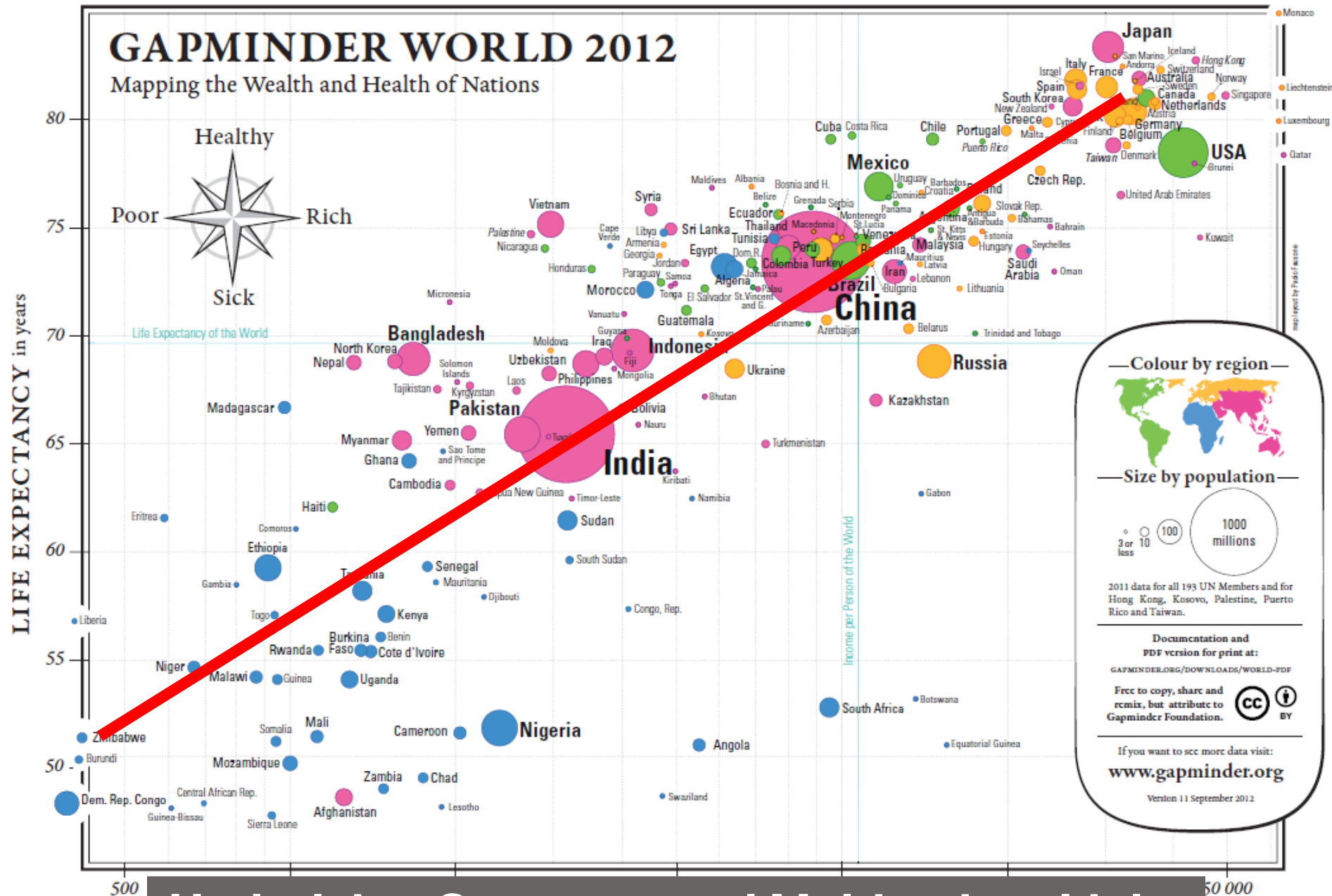
- meaningful comparisons
- potential causal networks and underlying structure
- multivariate links
- integrated and relevant data
- honest documentation
- primary focus on content



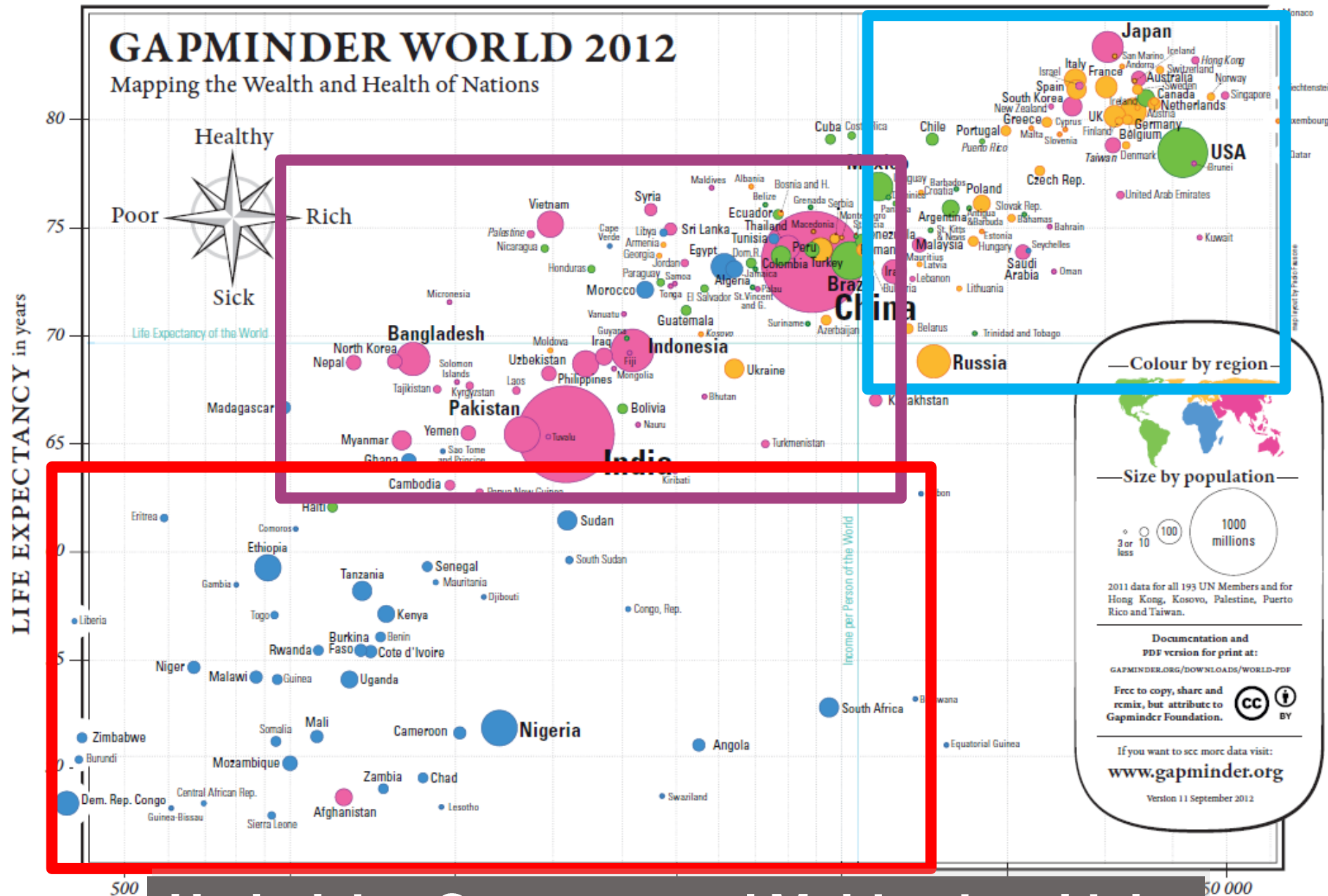
Non-Integrated Data

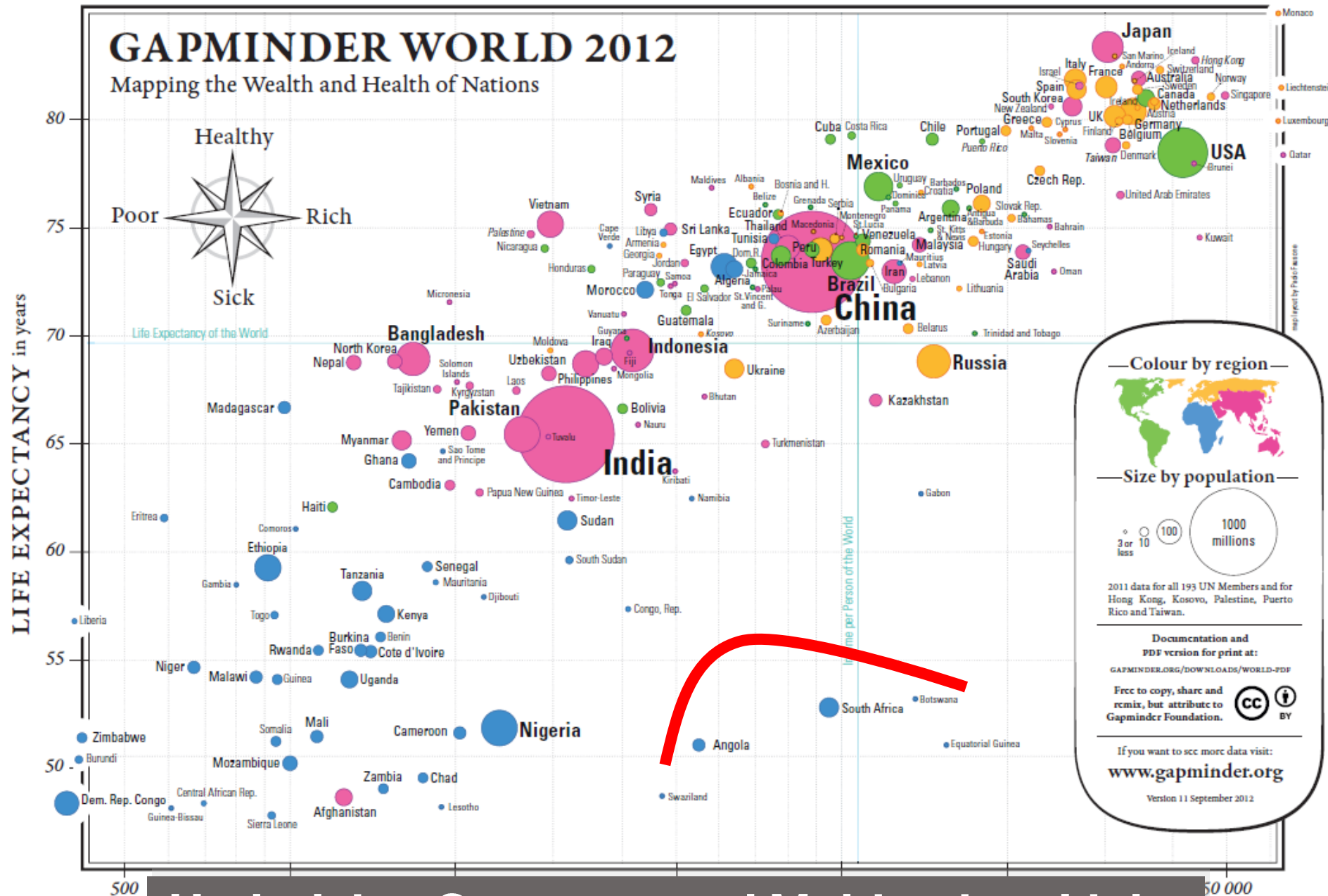


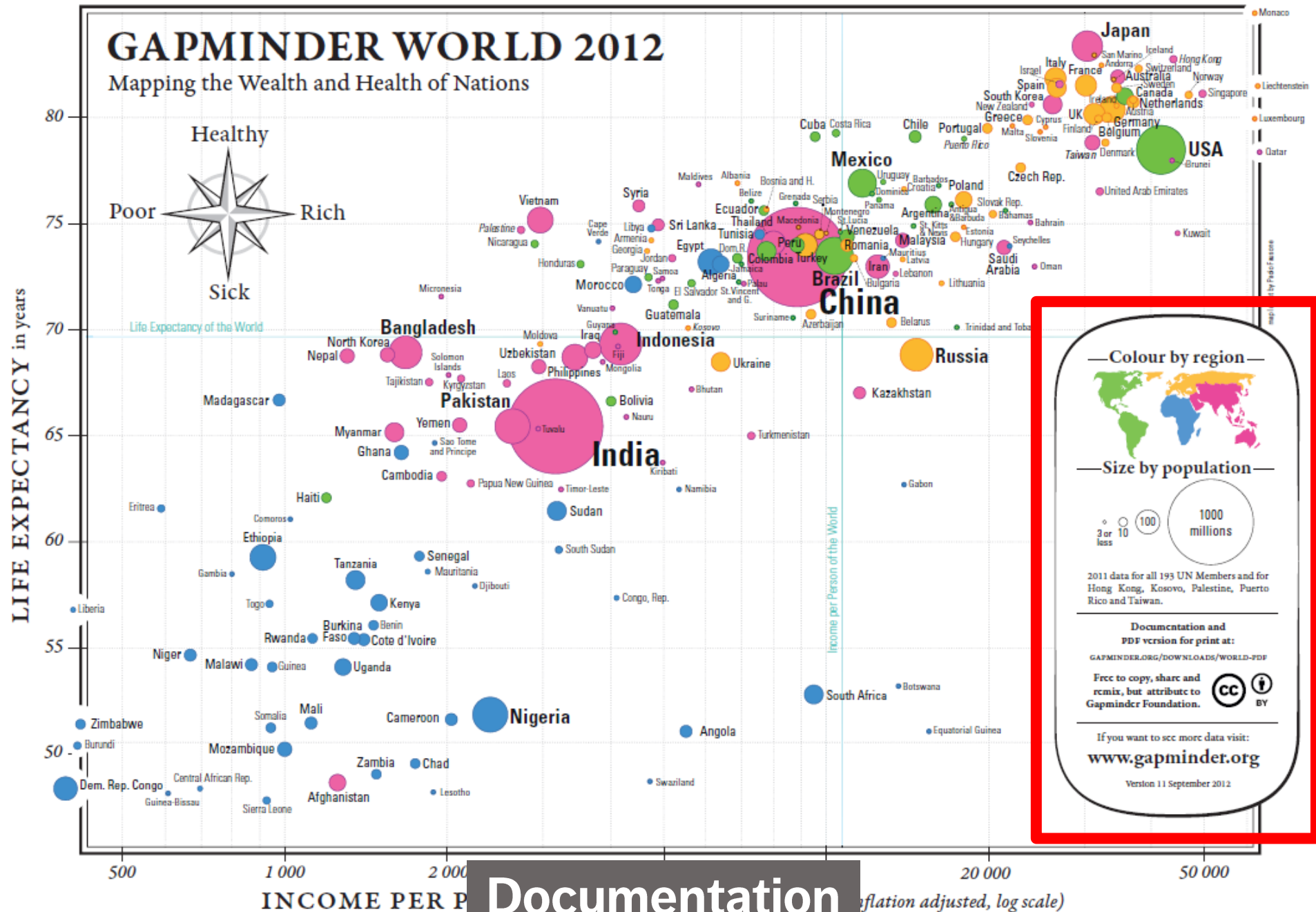




Underlying Structure and Multivariate Links





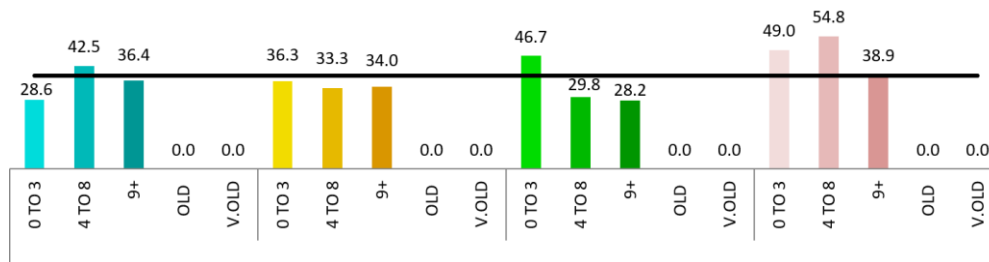


Basic Rules

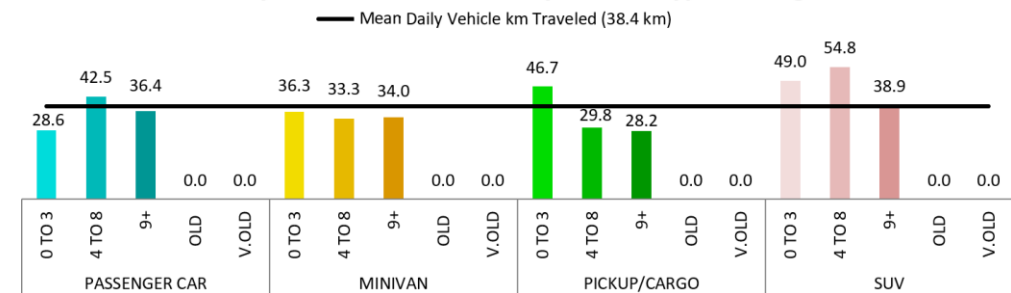
1. Check the data
outliers, spikes, anomalies

2. Explain encoding
don't assume the reader knows what everything means

Daily Vkt by Type and Age



Daily Vehicle km Traveled by Vehicle Type and Age

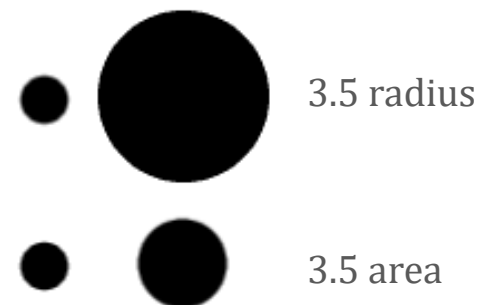


3. Label axes
knowing the scale is important

Basic Rules

4. Include units

eliminate the need for guesswork



5. Keep your geometry in check

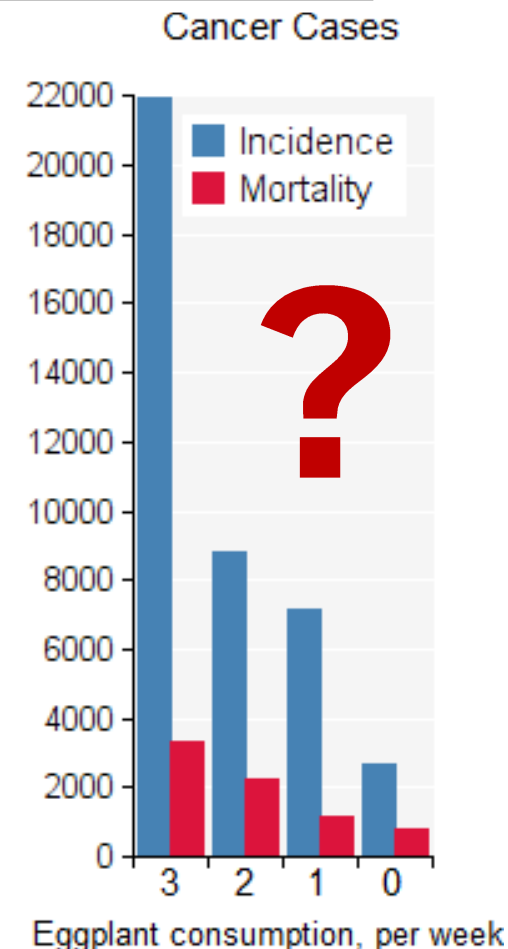
circles and 2D shape are sized by area, bar charts by length

6. Include your sources

protect yourself, and let those who want to dig deeper do so

7. Consider your audience

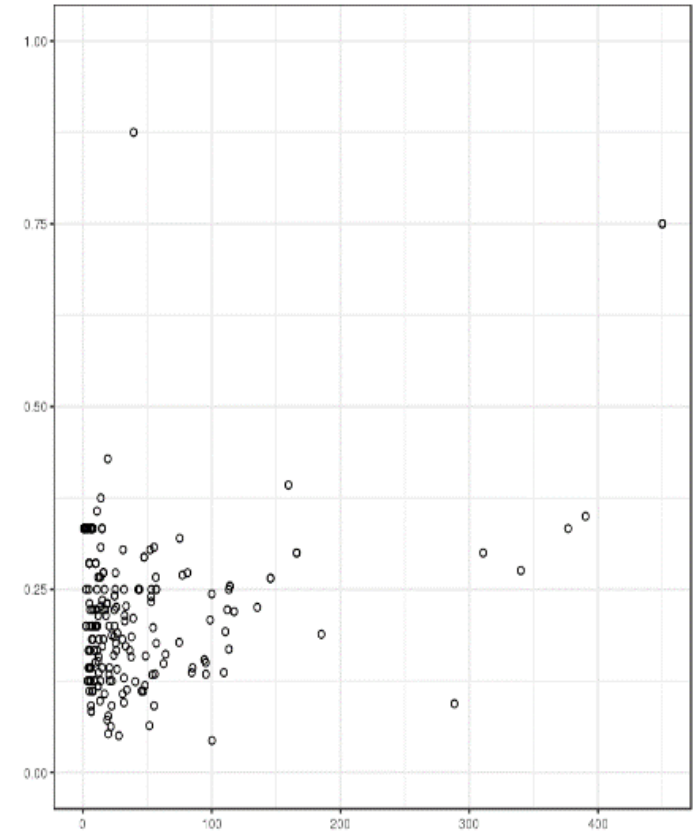
a poster can be wordy, a presentation should be minimalist



Representing Multivariate Data

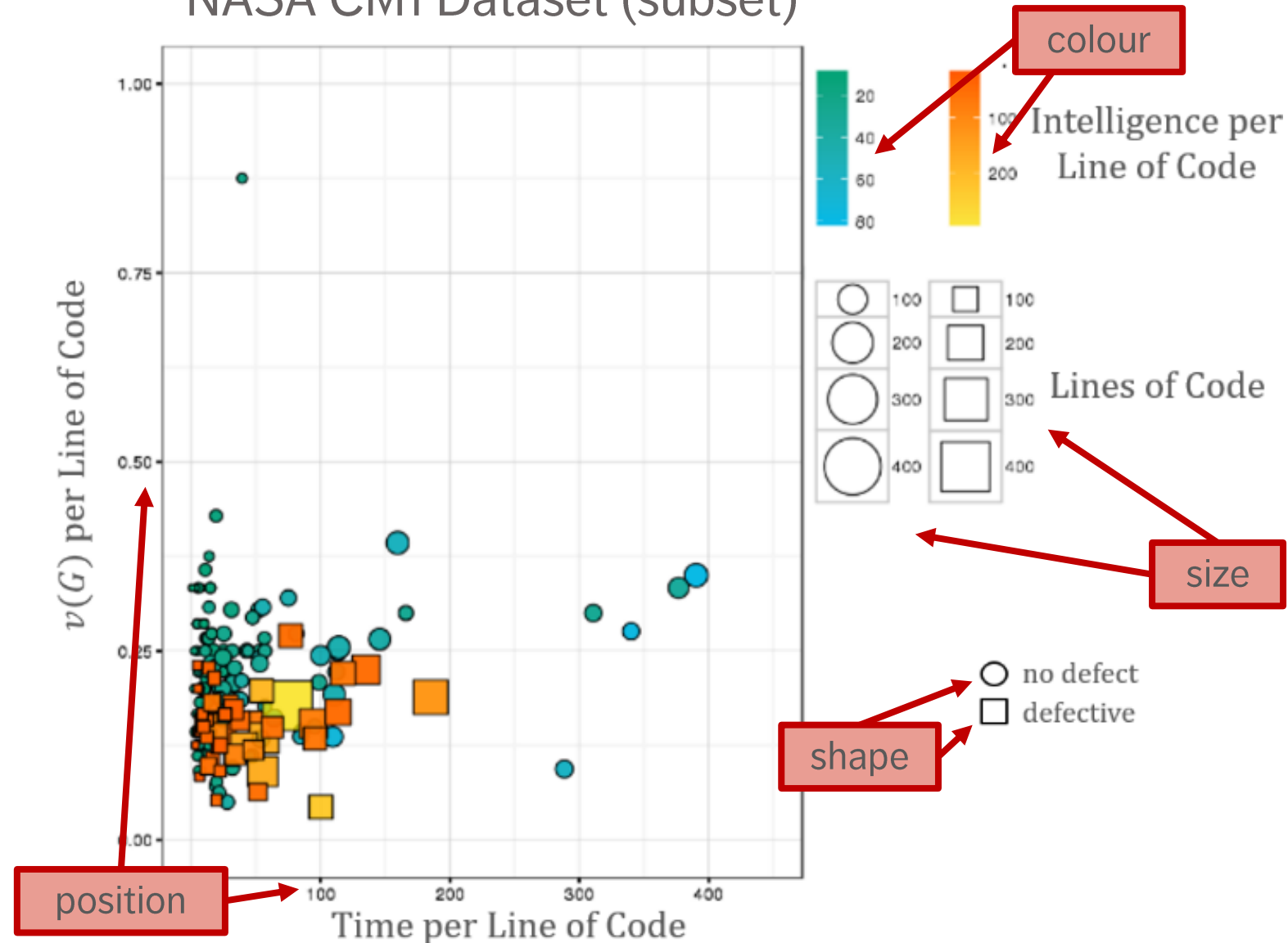
2 variables can be represented by position in the plane. Additional factors can be depicted with:

- size
- color
- value
- texture
- line orientation
- shape
- (motion?)



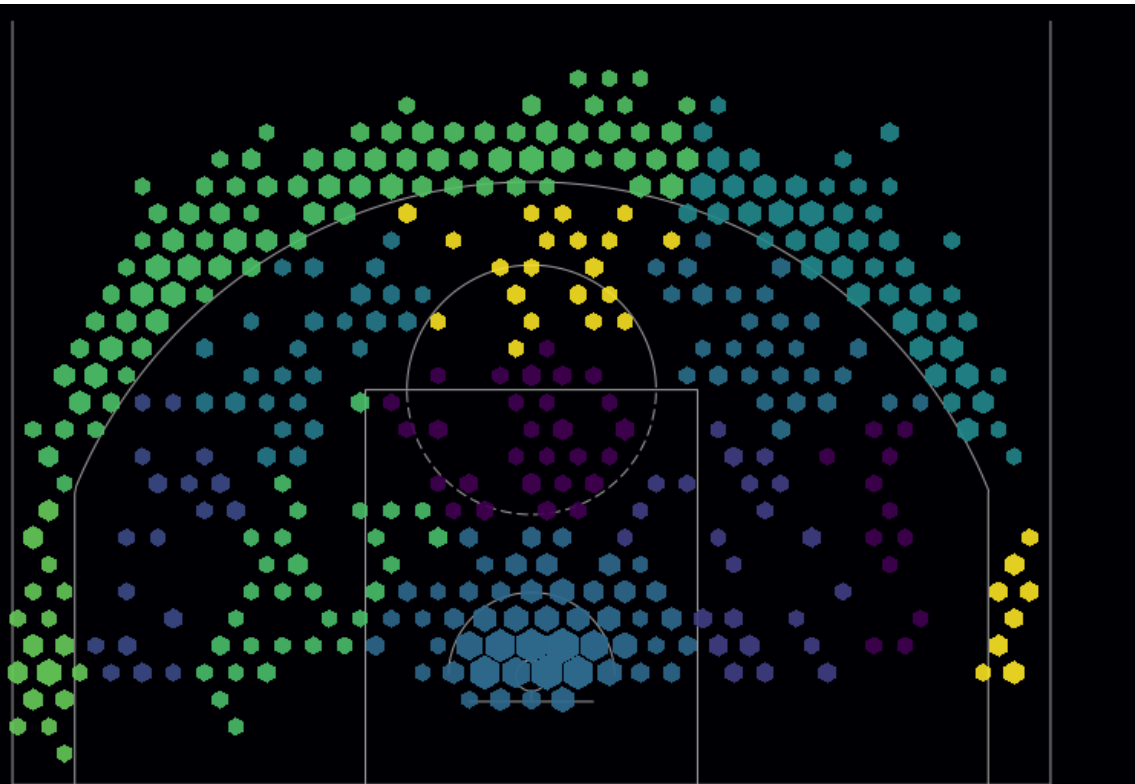
NASA CM1 Dataset (subset)

NASA CM1 Dataset (subset)

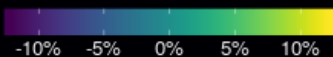


NBA FG% Against League Average ('15-'16)

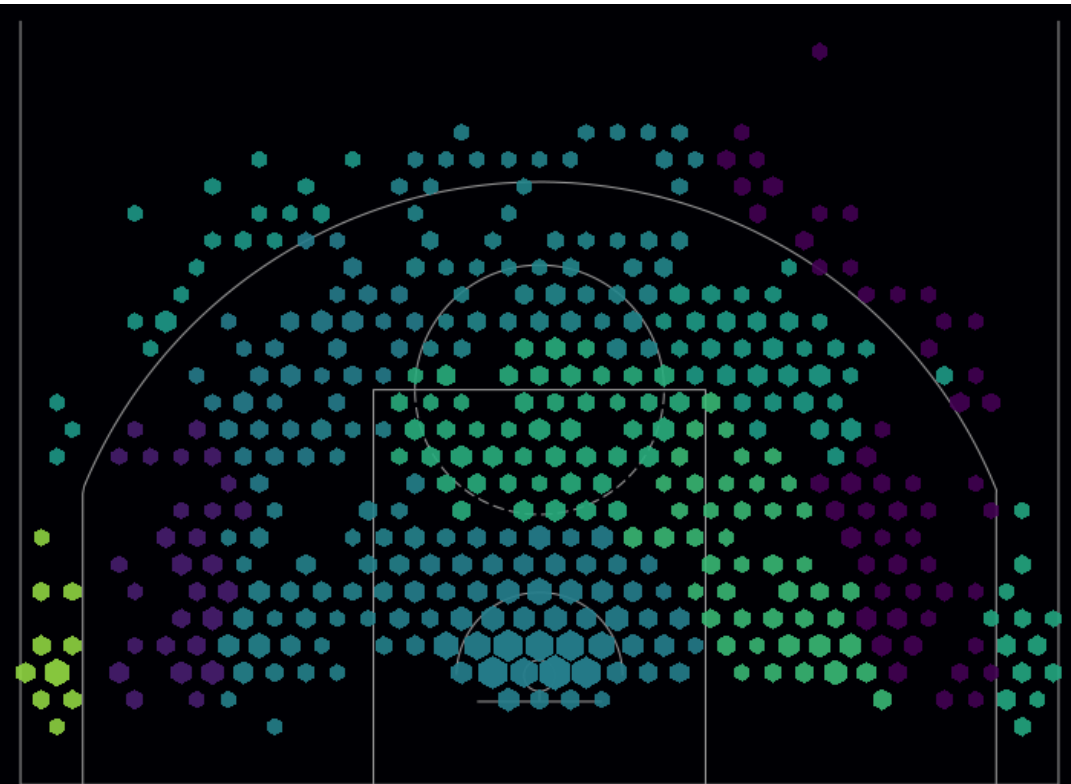
Kyle Lowry



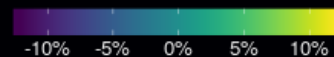
FG% vs. League Avg



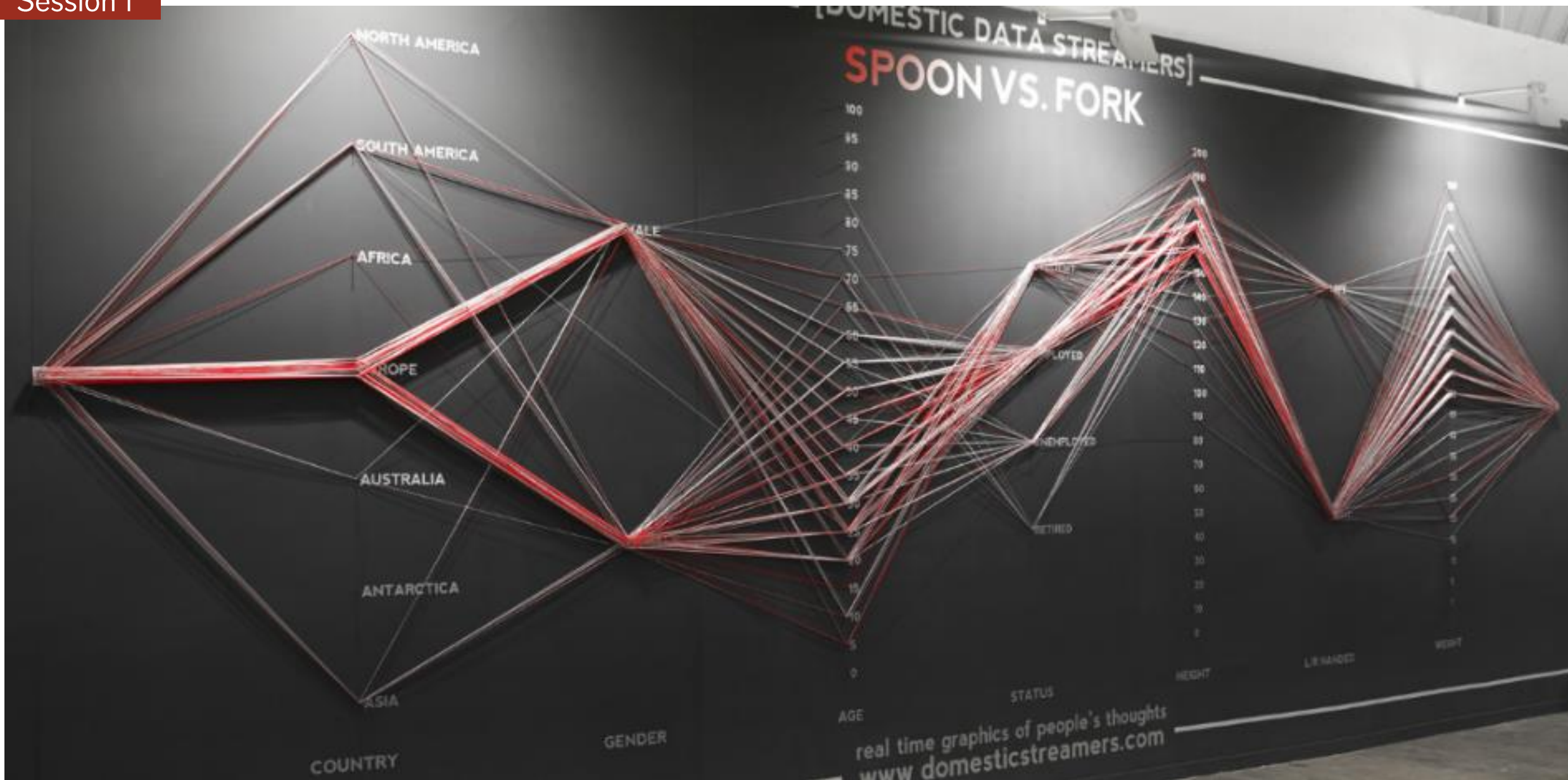
DeMar DeRozan



FG% vs. League Avg

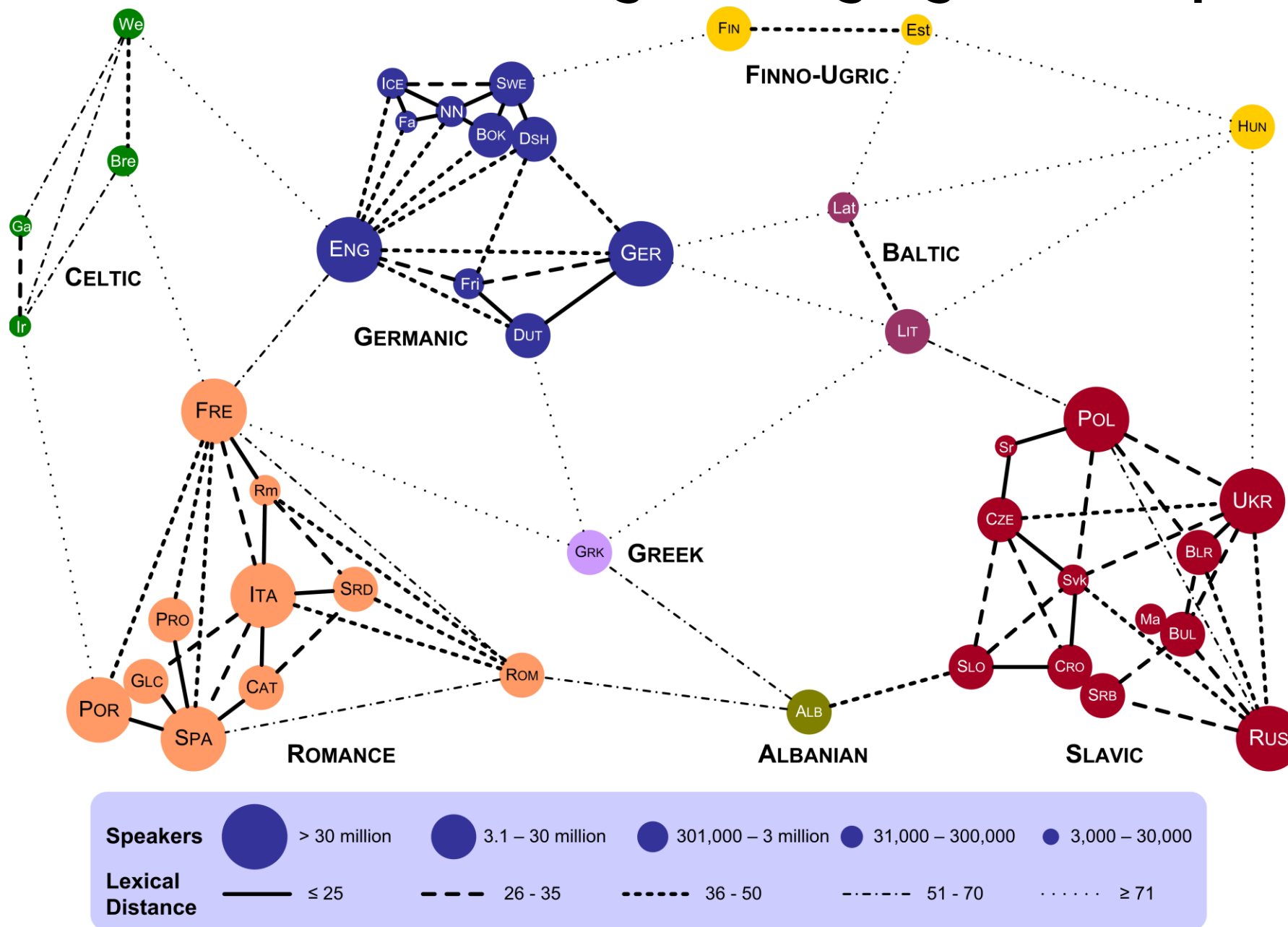


What comparisons can you make? Do you understand the encoding? The context?

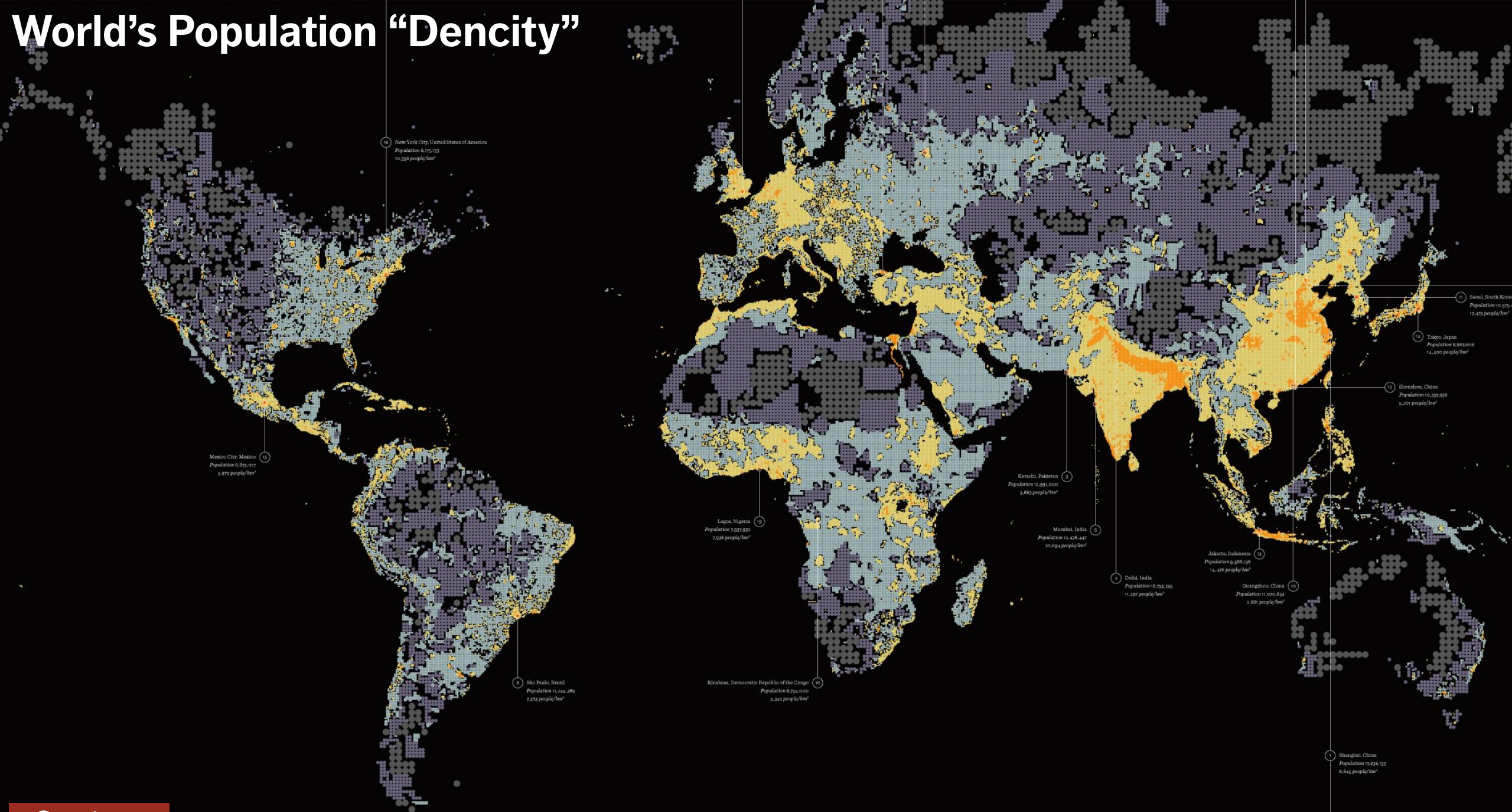


Are there any issues with data collection? Where do you think this event took place?
Is the spoon/fork question a red herring?

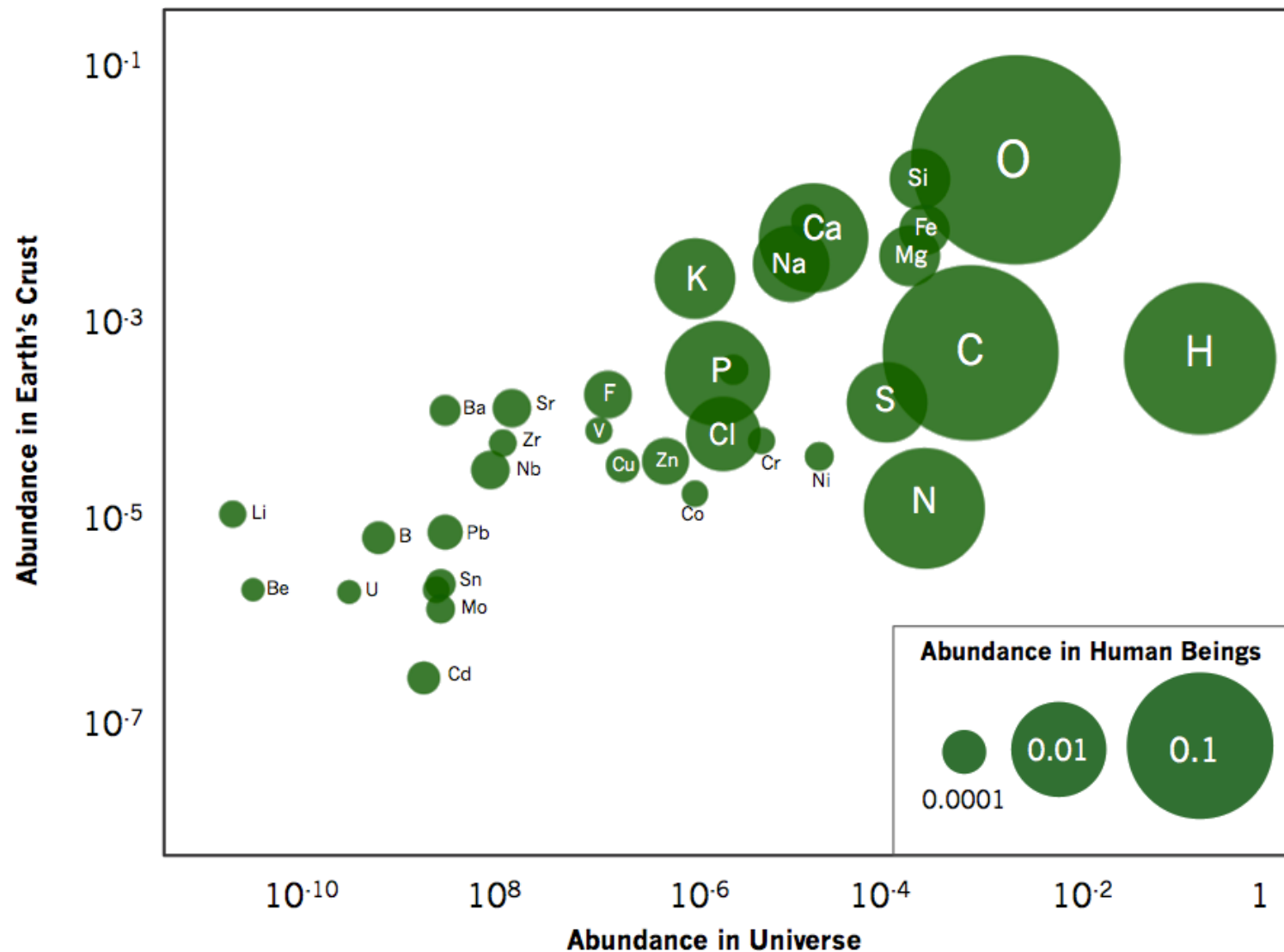
Lexical Distance Among the Languages of Europe



World's Population "Dencity"



Abundance of Chemical Elements



Suggested Reading

Data Visualization for Communication

Data Understanding, Data Analysis, Data Science **Volume 2: Fundamentals of Data Insight**

18. Data Exploration and Data Visualization

18.1 Data and Charts

- Presenting Results
- Multivariate Elements in Charts

The Practice of Data Visualization **Part I: A Data Exploration Primer**

2. Data Exploration and Visualization

2.3 Representing Multivariate Observations

2.4 Communicating Analysis Results

3. Hall-of-Fame / Hall-of-Shame

Exercises

Data Visualization for Communication

1. Study the charts presented in this section (or any other chart of your choosing) using the principles of analytical design. Are the basic rules followed?
2. What chart/design elements could be added to enhance the presentations?



Bubble Map



Bullet Graph



Calendar



Candlestick Chart



Chord Diagram



Choropleth Map



Circle Packing



Connection Map



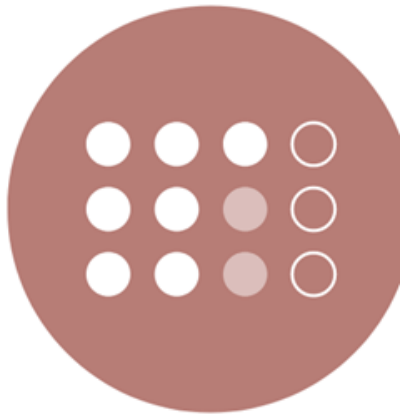
Density Plot



Donut Chart



Dot Map



Dot Matrix Chart

3. Visualization Catalogue

Data comparison charts

Data reduction charts

Comparison

Composition

Distribution

Evolution

Relationship

Profiling

Bars



Pie



Histogram



Line



Scatterplot



Grouped bars



Dot plot



Bullet



Pareto



ID Scatterplot



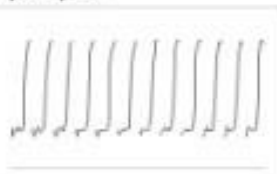
Horizon



Connected Scatterplot



Cycle plot



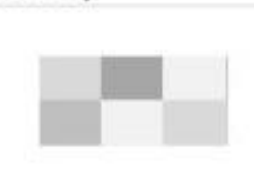
Scatterplot matrix



ID Scatterplot



Heat map



Multidimensional Pie



Boxplot



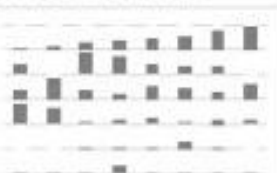
Step



Bubble



Reorderable matrix



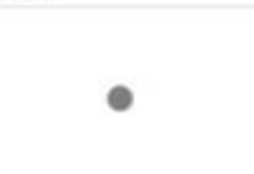
Horizon



Slope



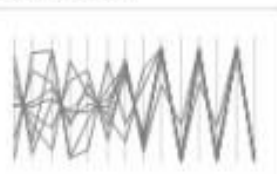
Alert



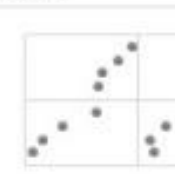
Connected Scatterplot



Parallel Plot



Trellis



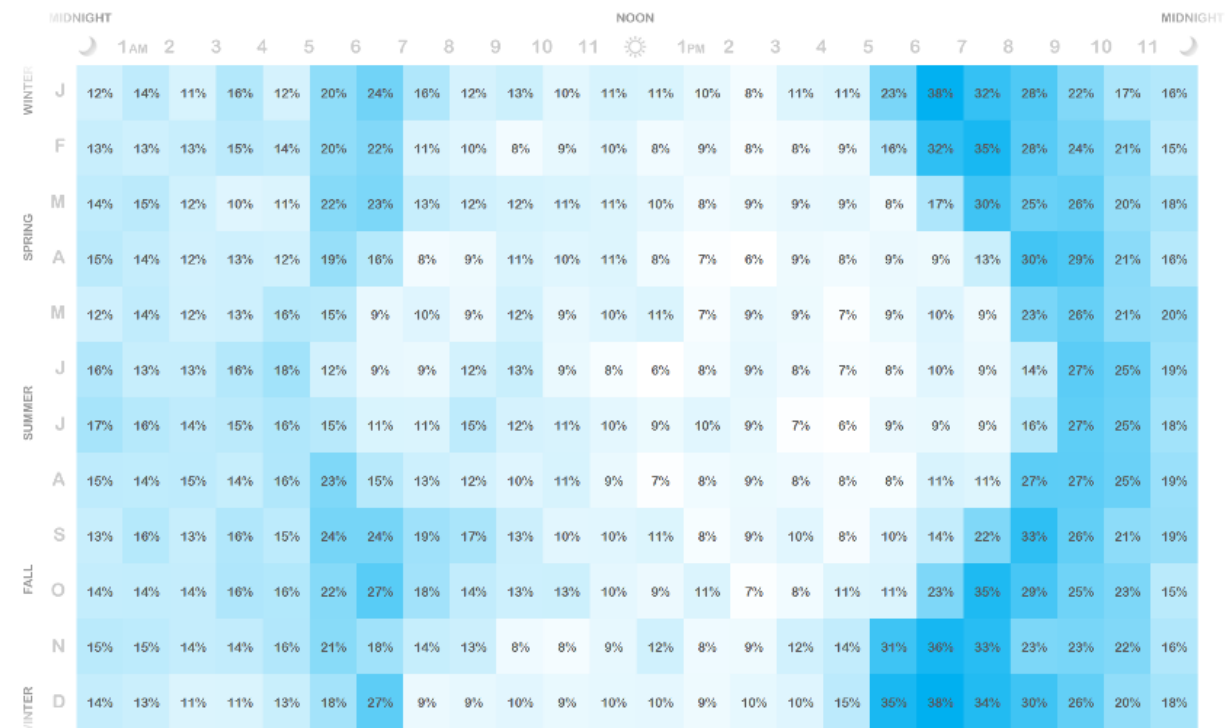
Heat Maps

The Horizon of Pedestrian Risk

The rate of fatal traffic incidents involving pedestrians, each hour of the day, throughout the seasons of the year.

The seasonal shift of our setting sun traces an ark of elevated risk – an echo of the curve of the Earth, itself (Note: ???).

Source: Fatality Analysis Reporting System (NHTSA 2006-2010)

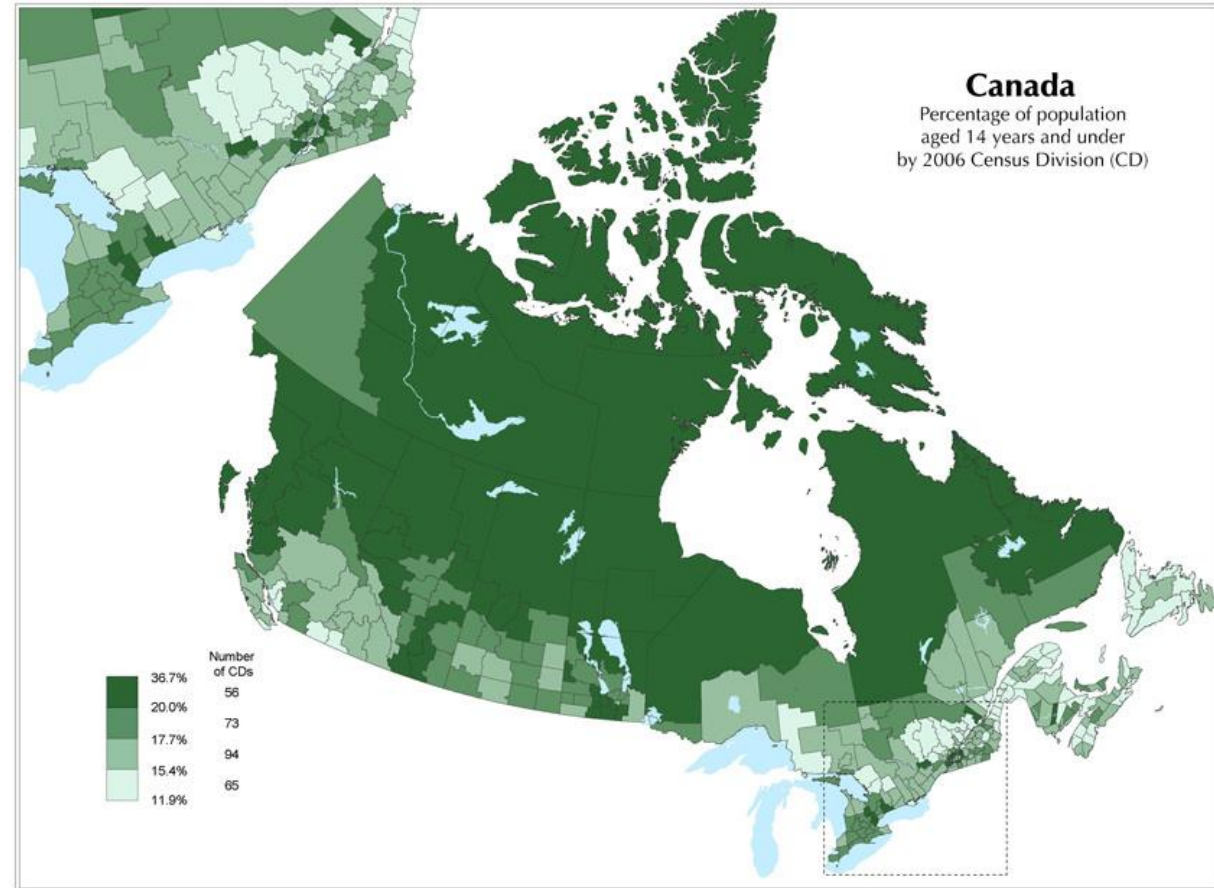
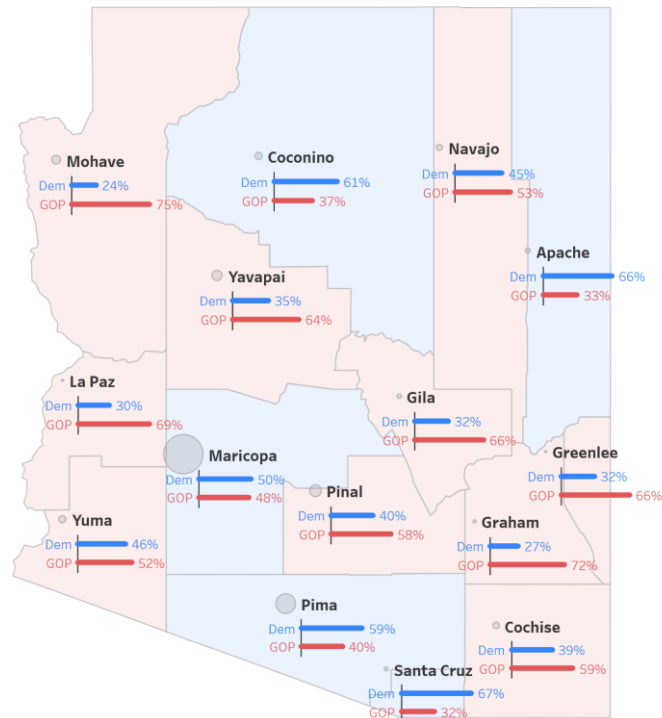


Heat Maps (Choropleths)

2020 Arizona Presidential Election Results by County

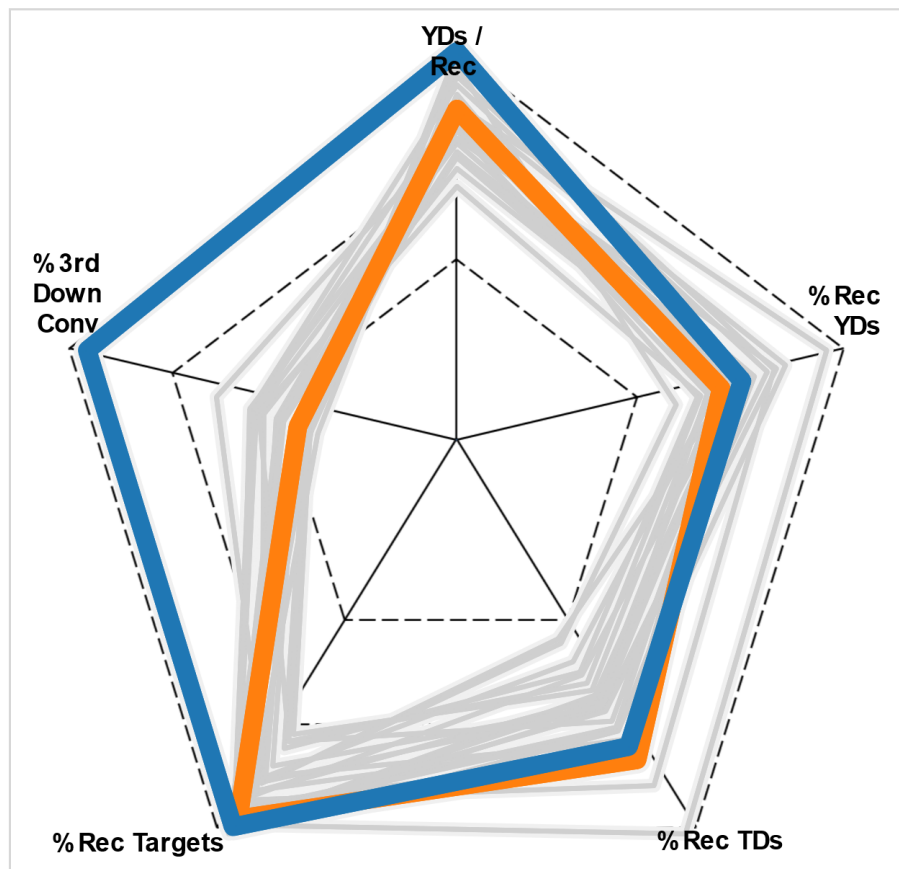
County color = 2020 winner. Blue = Dem. Red = GOP. Source NY Times

Change Year —●— 2020

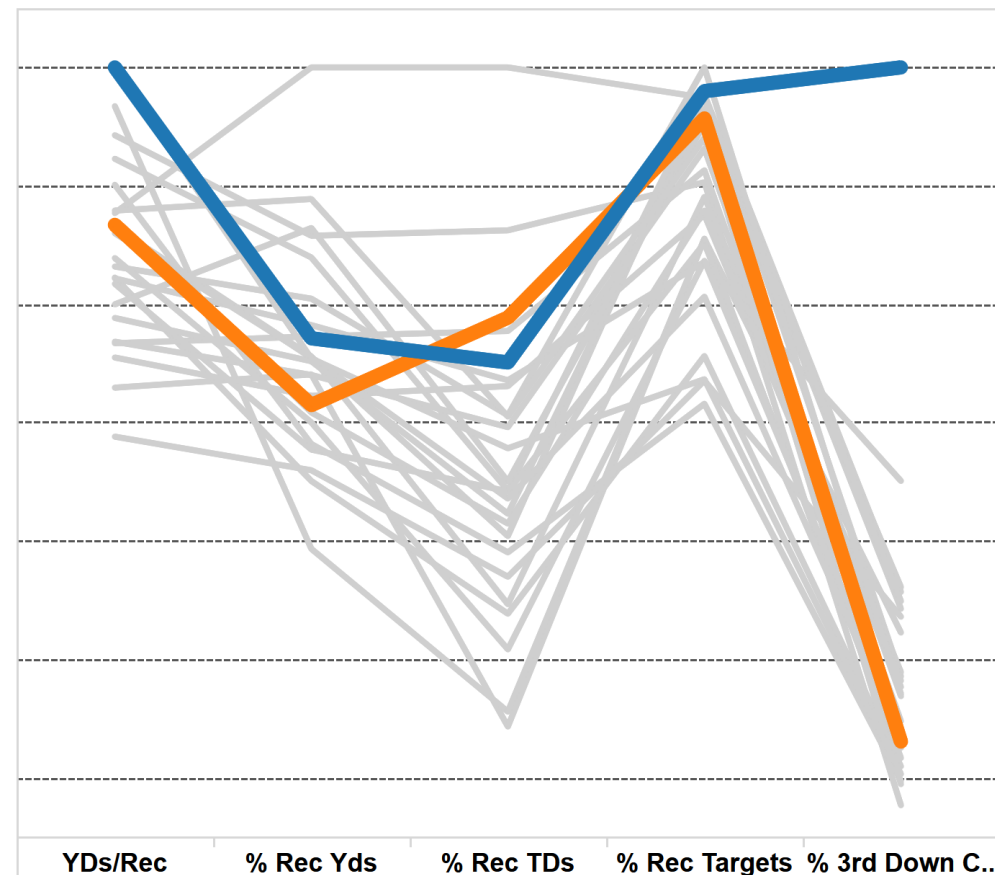


Parallel Coordinates / Radar Charts

Radar Chart



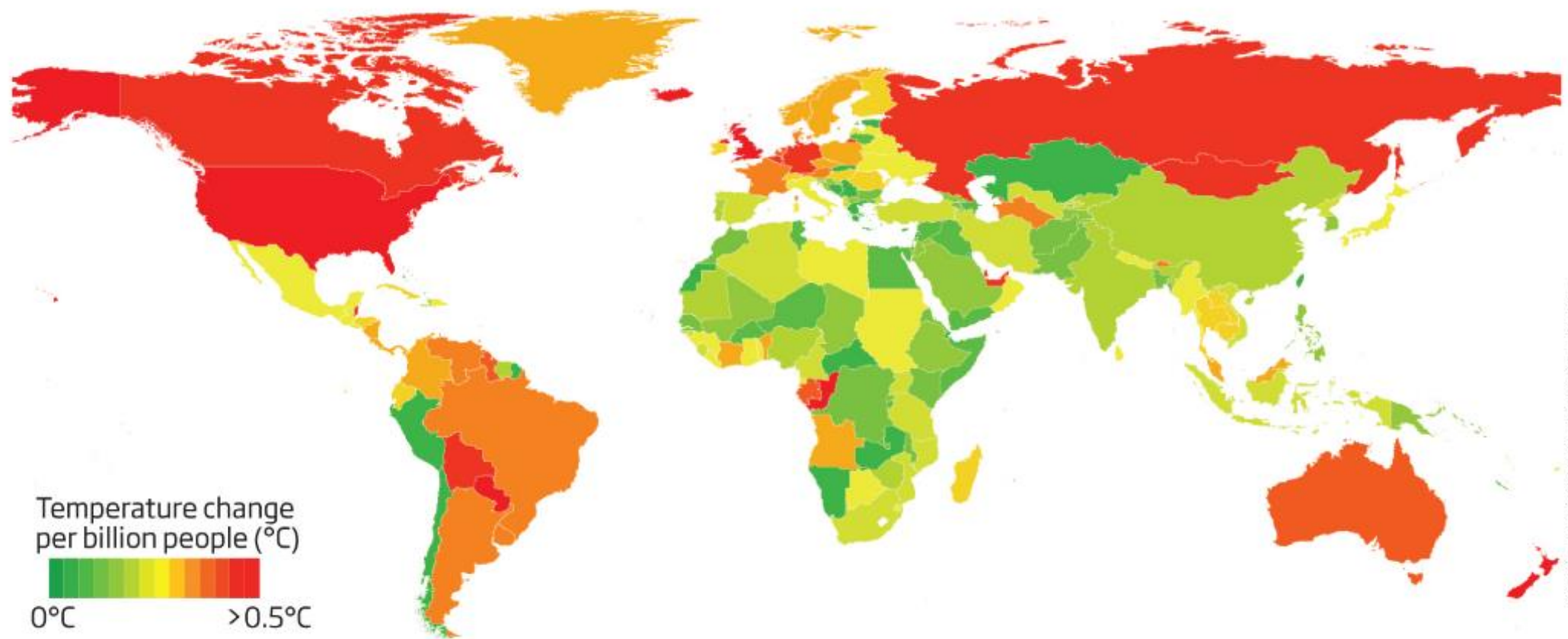
Parallel Coordinate



Geographical Maps

Global warming culprits, judged by population

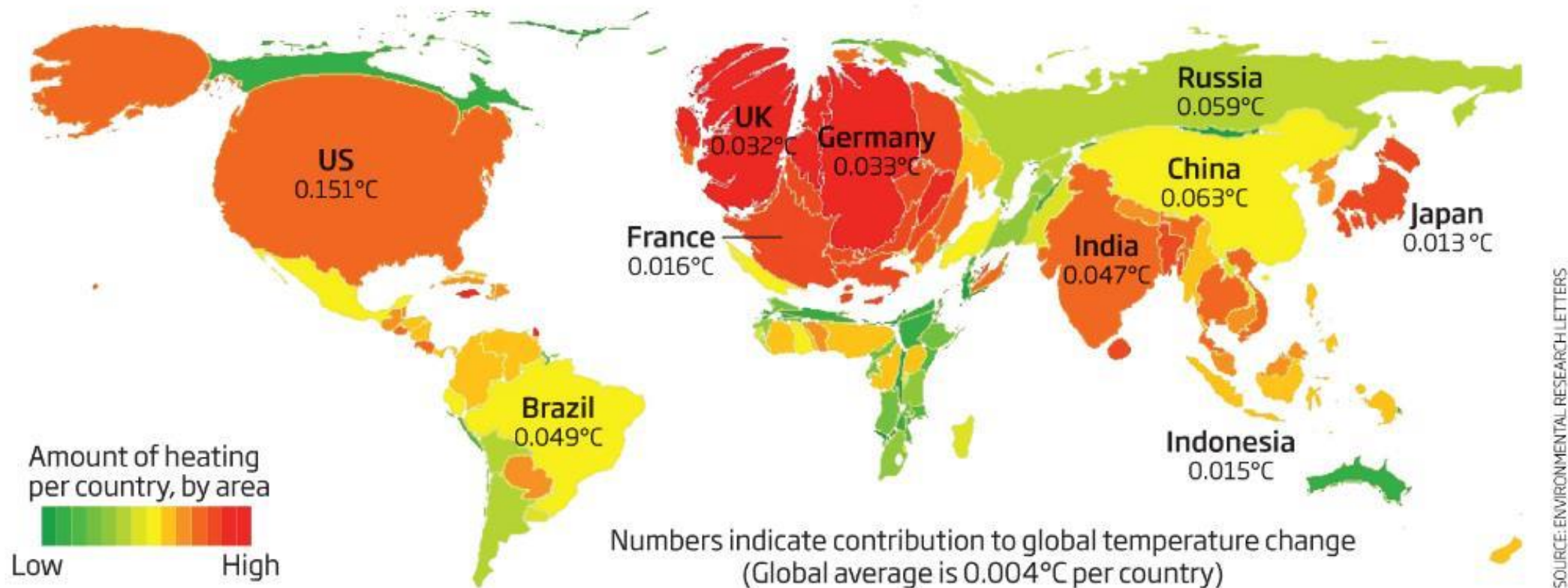
Countries that have caused more global warming per billion people are coloured red and low-emitters are dark green



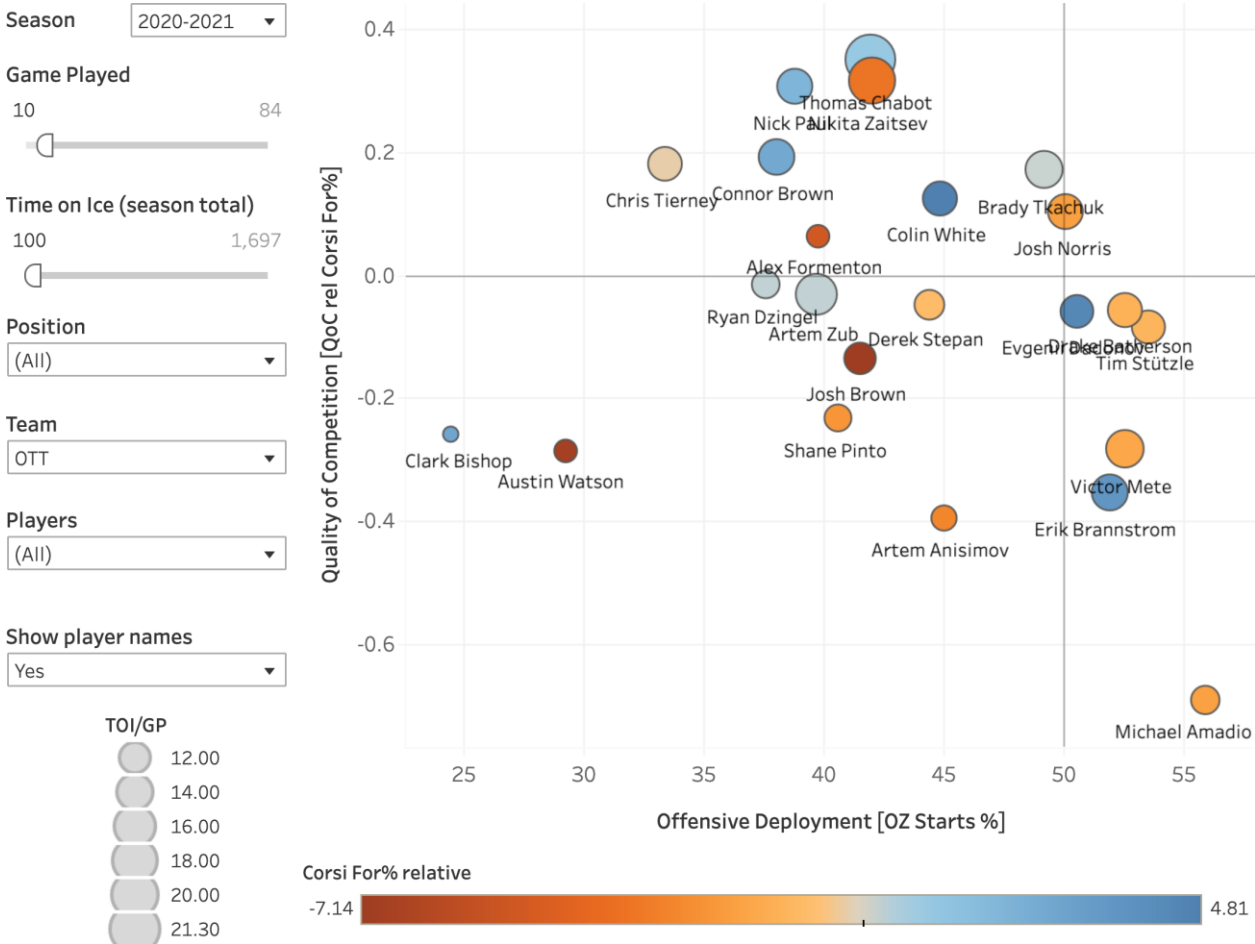
Distorted Geographical Maps

Global warming culprits, judged by size

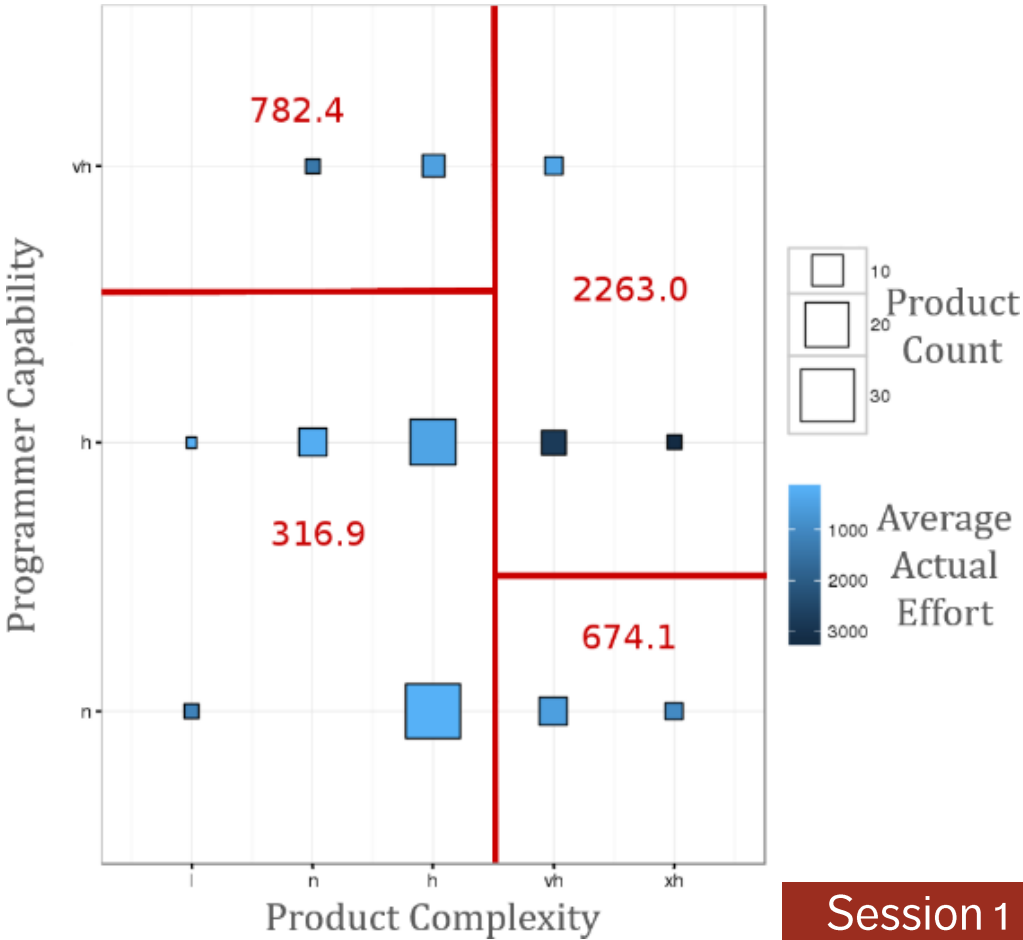
Countries that have caused disproportionately more global warming than their area would suggest are shown swollen, while low-emitters in relation to their size are shrunk



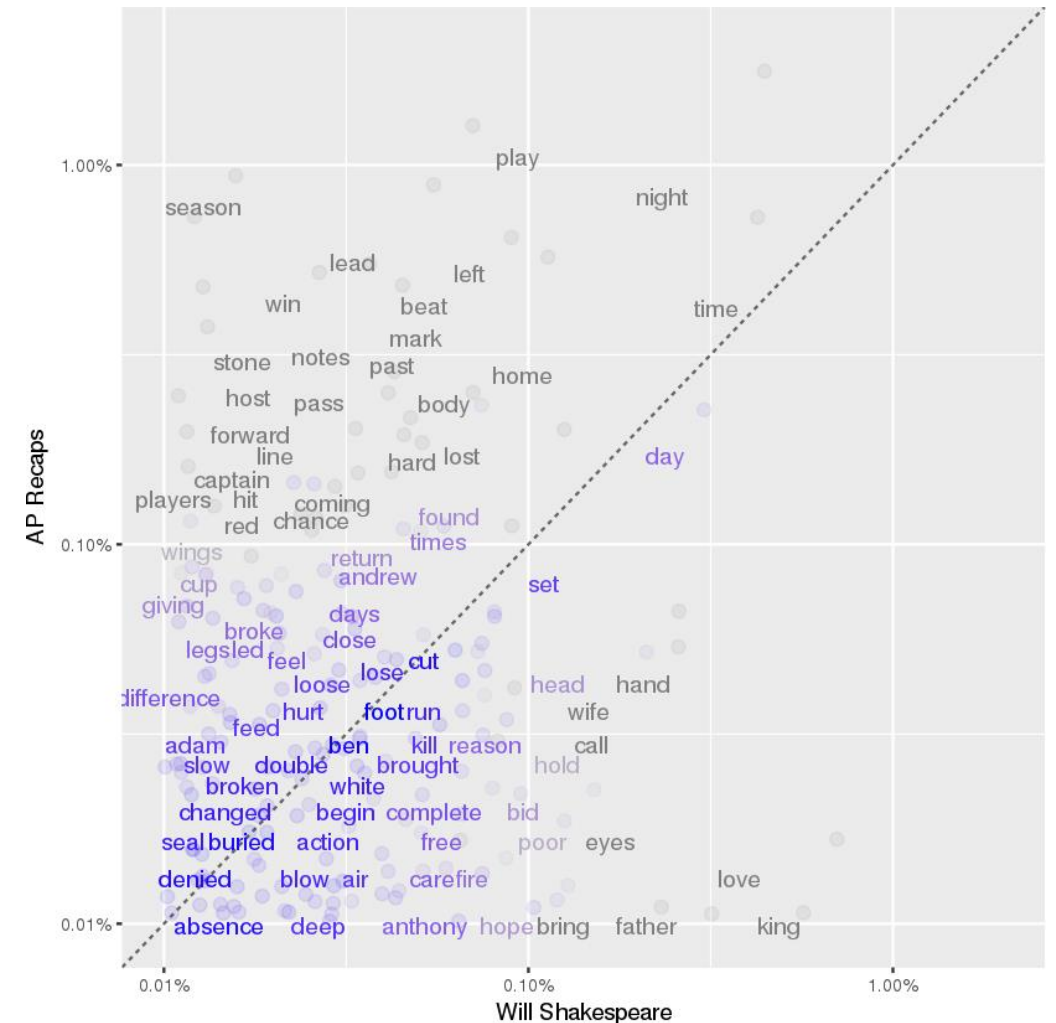
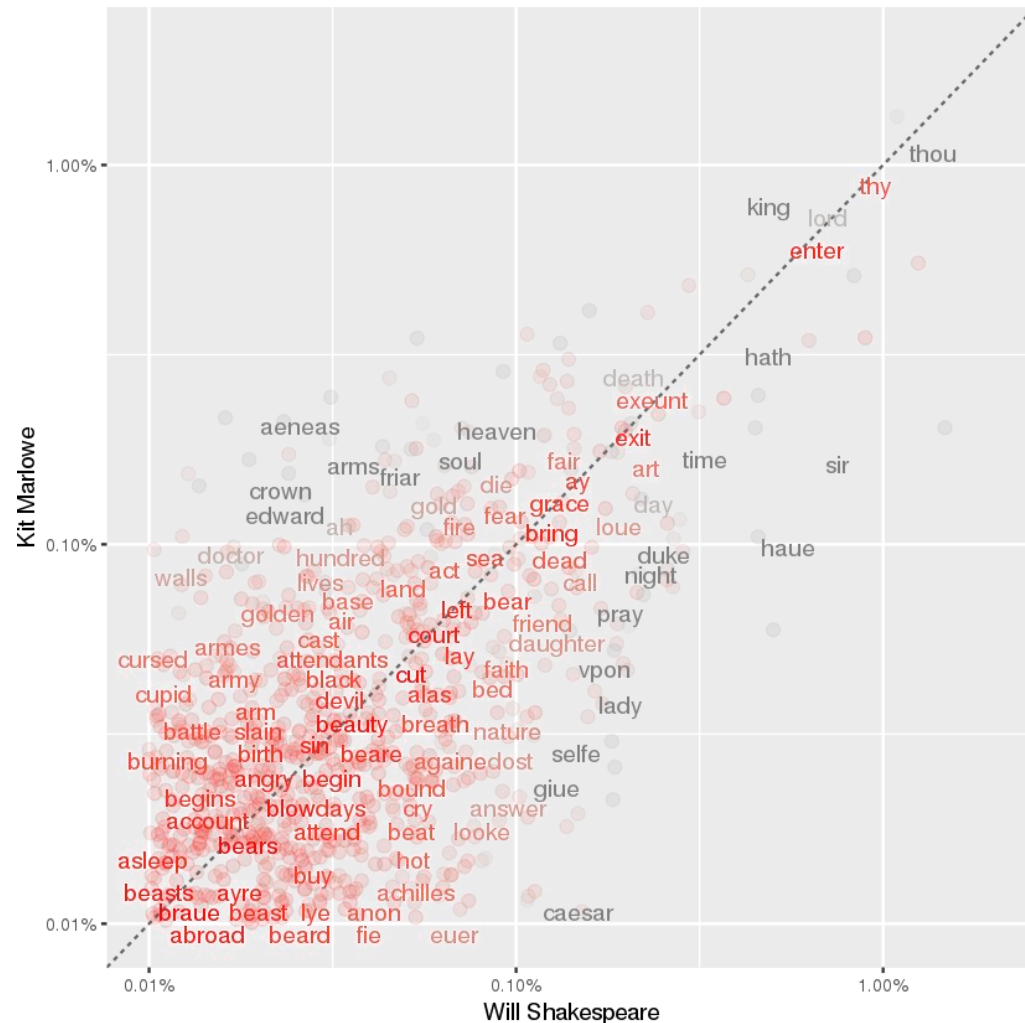
Bubble Charts



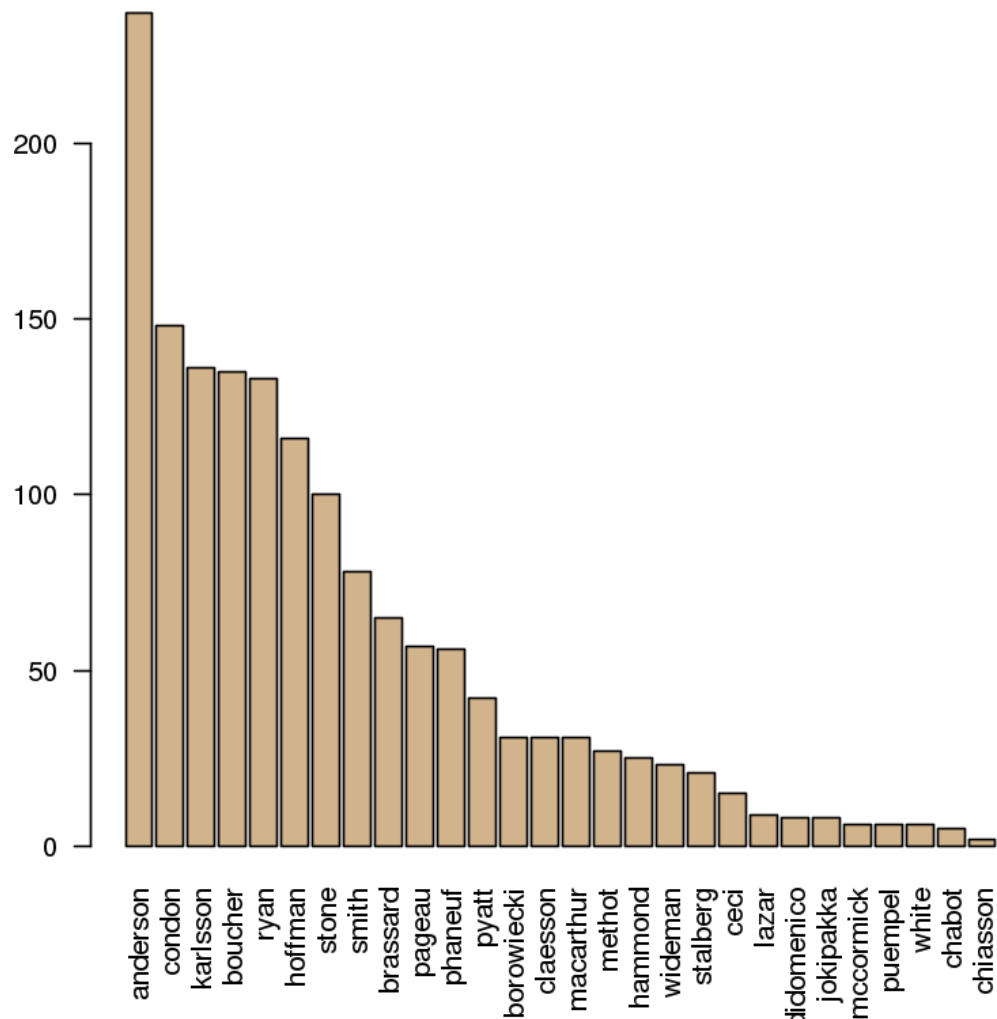
NASA COCOMO Dataset



Text Visualization & Representation

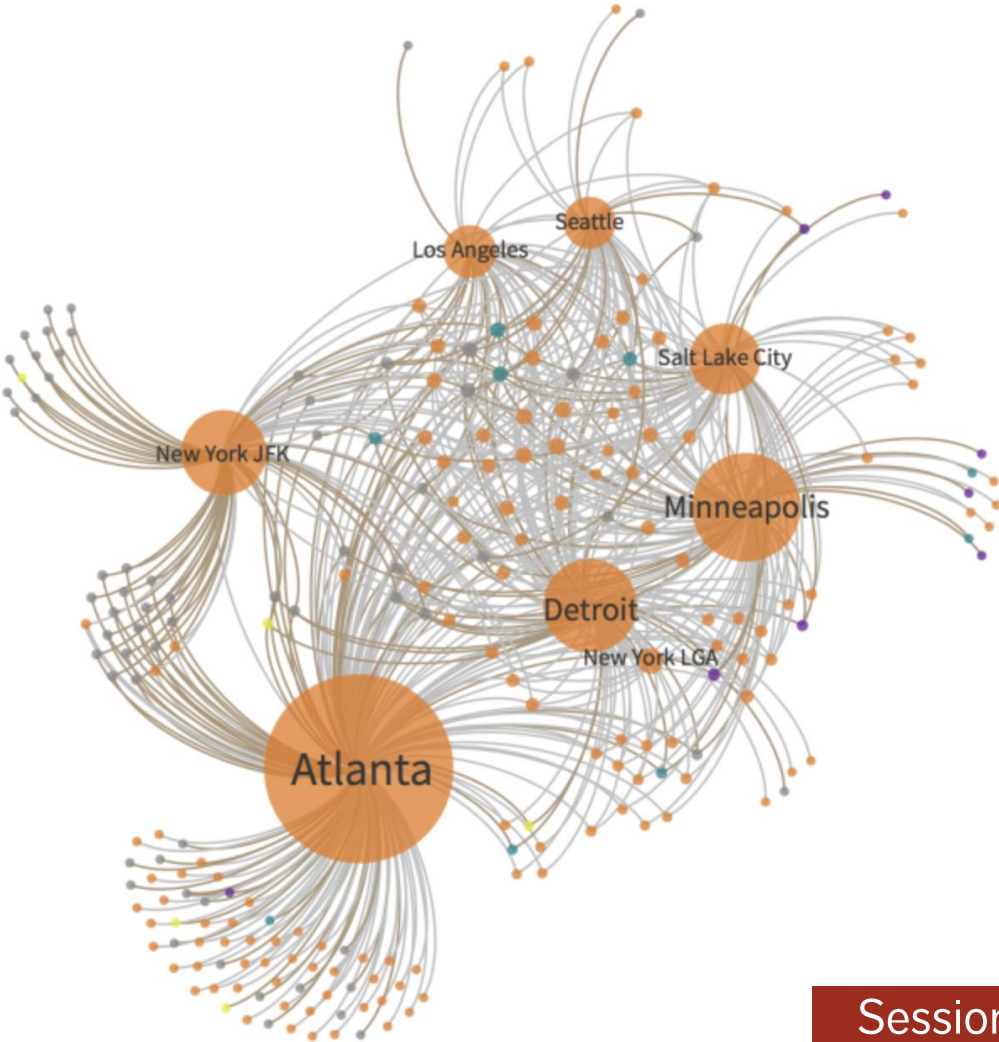
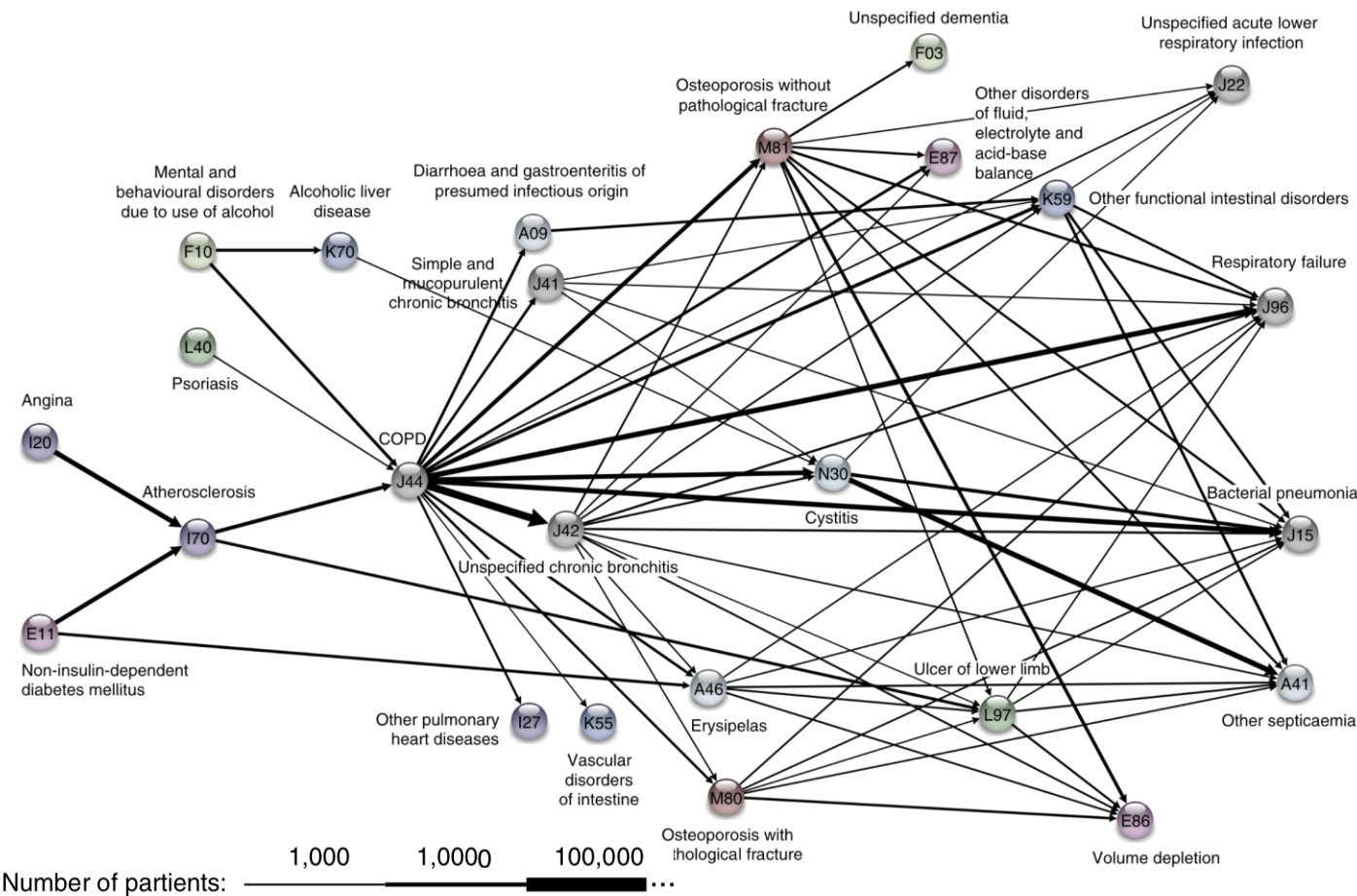


Text Visualization & Representation



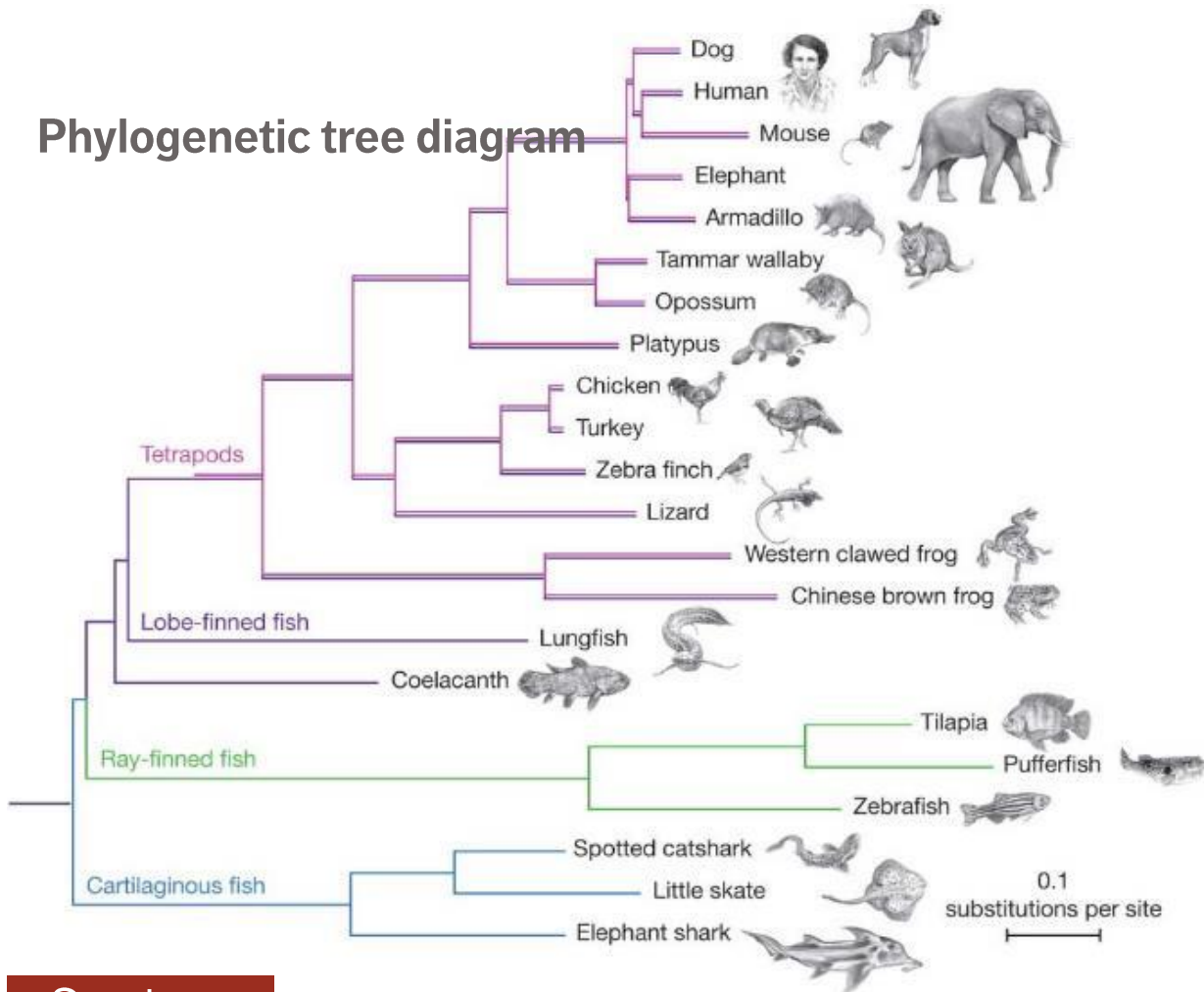
Network Diagrams

Diagnosis path around COPD in Danish population

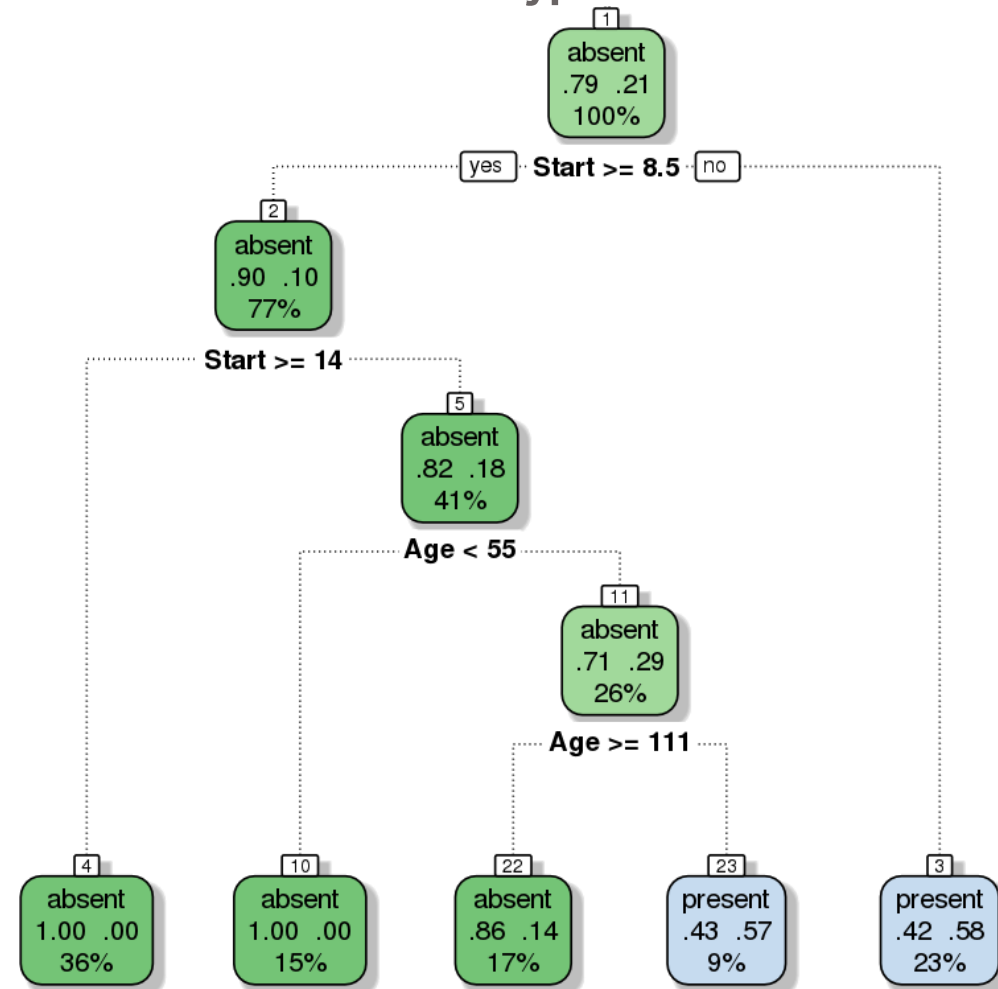


Decision Trees and Dendrograms

Phylogenetic tree diagram



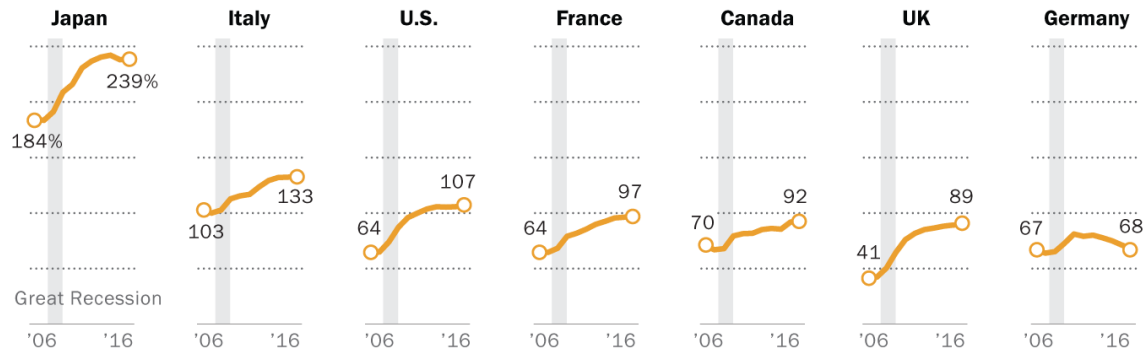
Classification of kyphosis dataset



Small Multiples

After Great Recession, debt increased substantially in most G-7 economies

Total gross debt as a share of GDP in the Group of Seven nations

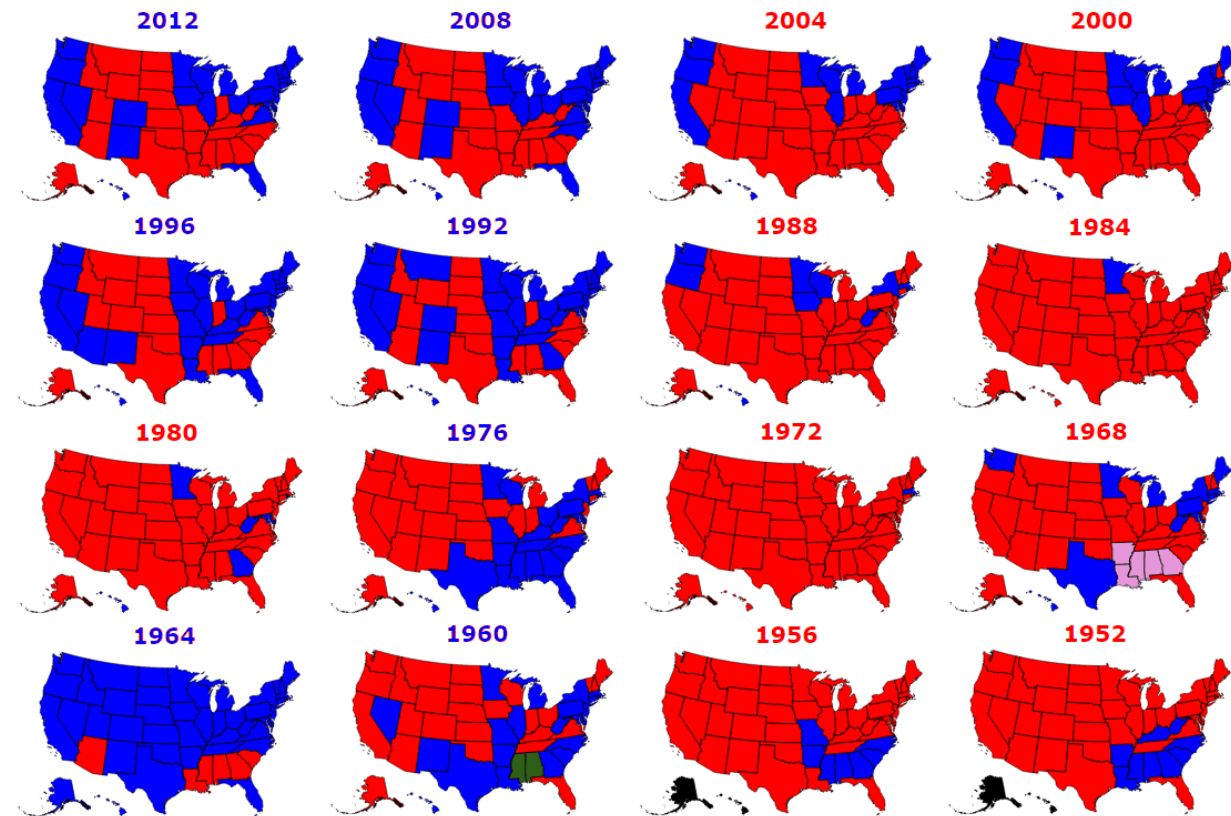


Note: Gross debt represents total liabilities of all levels and units of government — national, state/provincial and local — less liabilities held by other levels or units of government, unless otherwise noted by source.

Source: The International Monetary Fund, World Economic Outlook, accessed Sept 7, 2017.

PEW RESEARCH CENTER

U.S. Electoral College Results 1952 – 2012



Sparklines

TOTAL

Hospital #1

Hospital #2

Hospital #3

Hospital #4

Hospital #5

Hospital #6

Hospital #7

Hospital #8

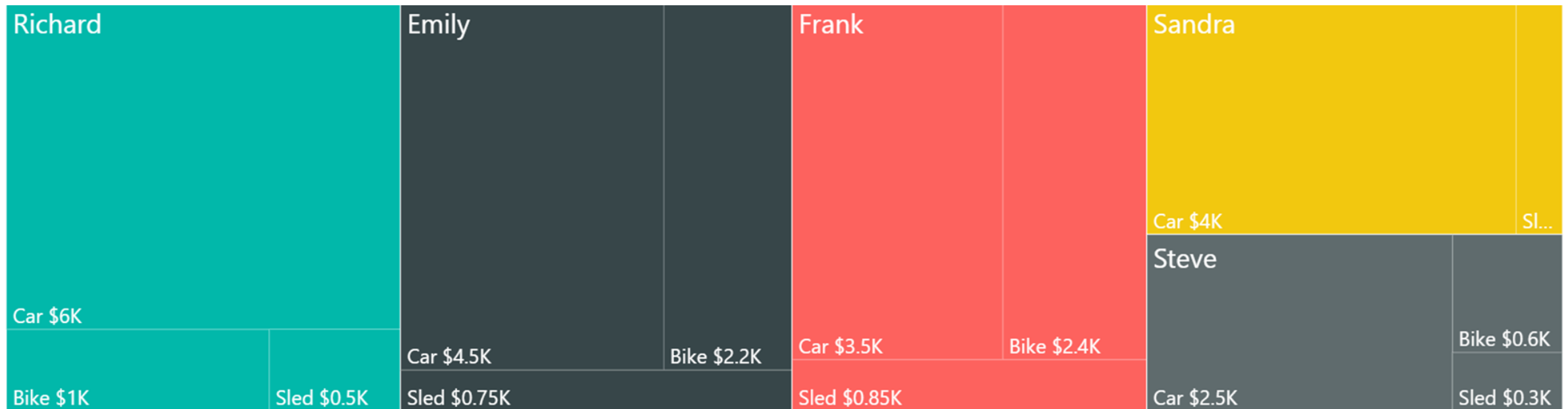
Start	Monthly Number of Cases	End	Low	High	Mean	Std Dev	Blanks	Zeros	Trend
19502		17265	15150	25072	19903	2612	0.0	0.0	379.2
46		19	3	46	19	9	0.0	0.0	-1.6
156		240	101	326	194	60	0.0	0.0	9.7
16		11	2	76	15	15	0.0	0.0	-2.9
3		13	0	105	9	15	0.0	0.4	-1.8
42		50	25	91	61	16	0.0	0.0	1.2
48		53	34	169	67	25	0.0	0.0	0.6
0		N.A.	0	0	0	0	2.2	9.8	0.0
56		104	34	150	73	25	0.0	0.0	4.6

Treemaps

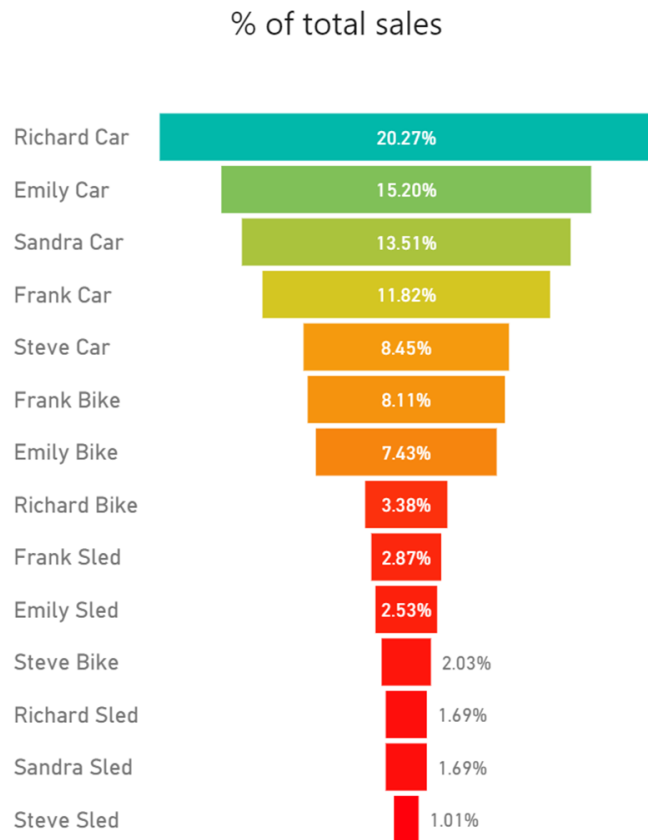
Simultaneously show big picture and can compare related easily. Easy to process data sub-categories.

Useful to prioritize “big ticket items” in dynamic dashboards.

Labeling and colouring are tricky.



Funnel Charts

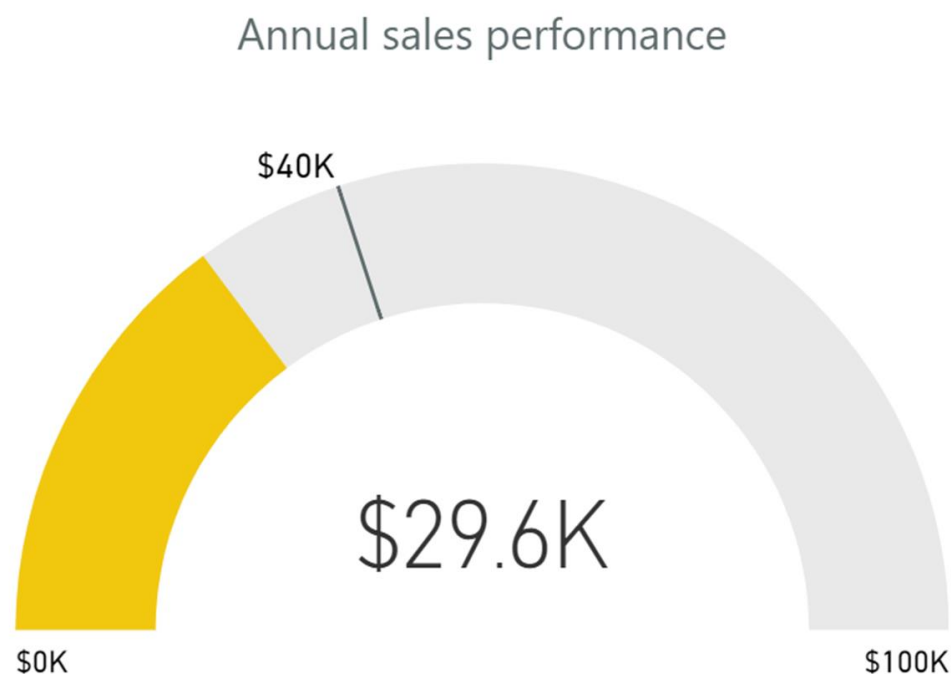


Typically represents **decreasing proportions** amounting to 100% total (not always though).

If this is not the default sort, users should manually sort from high to low (otherwise, plot looks messy).

VERY useful to help audience quickly prioritize items without having to actively filter.

Gauge Charts



Often used as a dashboard component (with or without needle).

Displays single value measures towards goal / KPI.

Great to show progress (a bit of a management fad, though...)

Displays information that can be quickly **scanned** and **understood**.

Interactive & Animated Visualizations

Animation **does not always** improve a visualization. What insights can interactivity provide? That depends on the data, and on the visualization method.

Examples:

- [The Clubs That Connect the World Cup](#), NY Times, 2014
- [Who Marries Whom](#), Bloomberg, 2016
- [Hipparcos Star Mapper](#), European Space Agency, 2016
- [The Internet of Things – a Primer](#), Information is Beautiful, 2016
- [The Genealogy and History of Popular Music Genres](#), Musicmap, 2016

Interactive & Animated Visualizations

Examples (continued):

- [Sequences Sunburst](#), Kerry Rodden, 2015
- [Health and Wealth of Nations](#), Gapminder Foundation
- [Mobius Transformations Revealed](#), Arnold D.N, Rogness, J, 2007
- [Visualizing the Riemann \$\zeta\$ Function and Analytic Continuation](#), 3Blue1Brown, 2016
- [Small Arms and Ammunition – Imports and Exports](#), Google, 2012
- [The Evolution of the Web](#), Google, Hyperakt, Vizzuality, 2012
- [peoplemovin](#), Carlo Zapponi, 2012

Interactive & Animated Visualizations

“There is always a danger that if certain types of visualization techniques take over, the kinds of questions that are particularly well-suited to providing data for these techniques will come to dominate the landscape, which will then affect data collection techniques, data availability, future interest, and so on.”
(P. Boily)

Even when done well, 85% of users don't bother with interactive viz (NY Times).

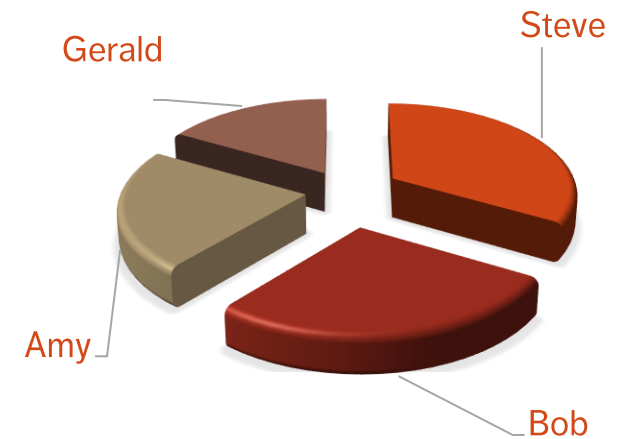
Take-Away: explore the data and try different methods

Charts to Avoid

ANYTHING with an arc (except gauge)

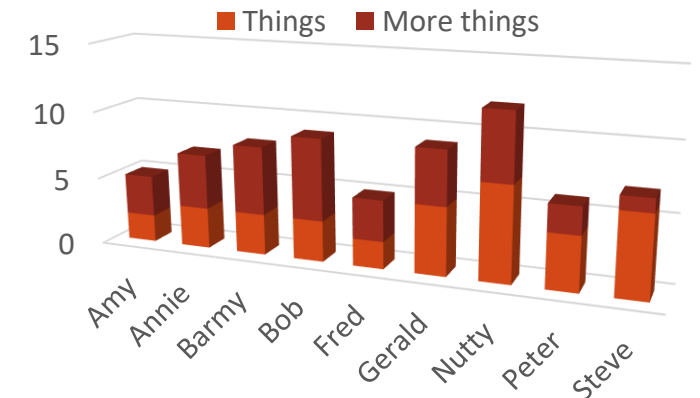
- pie
- donut

Brain cannot compare arcs and they can be misleading:
how different are Steve & Bob in the pie chart?



ALL 3D IS EVIL!

- as with arc, we cannot easily visually compare data series
- adds way too much clutter



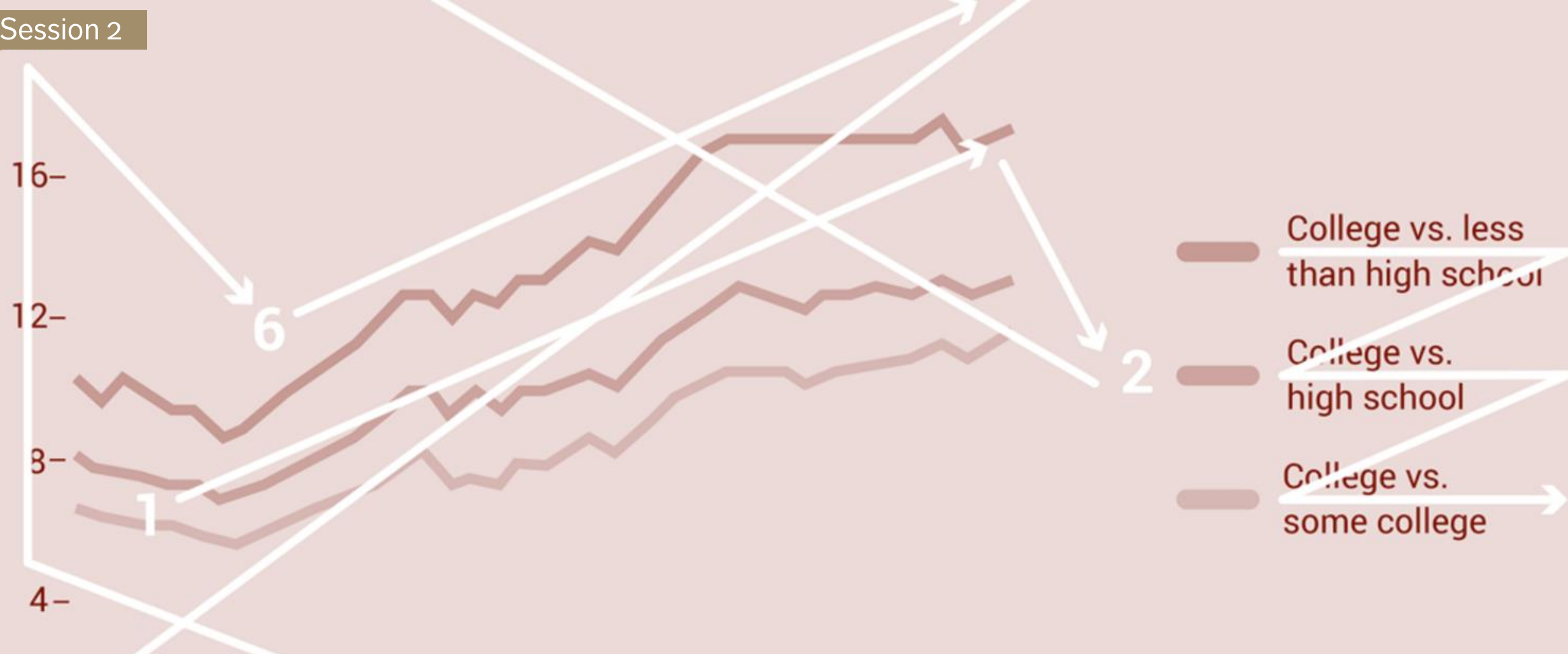
Suggested Reading

Visualization Catalogue

The Practice of Data Visualization

Part IV: Some Practical Aspects of Data Visualization

9. Visualization Toolbox



4. Chart Aesthetics

Gestalt Principles

The **Gestalt principles** are the “laws” of human perception.

They describe how humans group similar elements, recognize patterns and simplify complex images when they perceive objects.

Designers use them to organize content on charts, dashboards, websites, and other interfaces so that they be **aesthetically pleasing/easy to understand**.

Gestalt Principles

“Gestalt” is German for “unified whole”.

The first principles were devised in the 1920s by German psychologists Wertheimer, Koffka (“the whole is greater than the sum of the parts”), Kohler.

Aim: understand how we gain meaning from the chaotic stimuli around us.

The Gestalt principles are a set of “laws” which address the natural compulsion to find order in disorder. According to this, the mind “informs” what the eye sees by **perceiving a series of individual elements as a whole**.

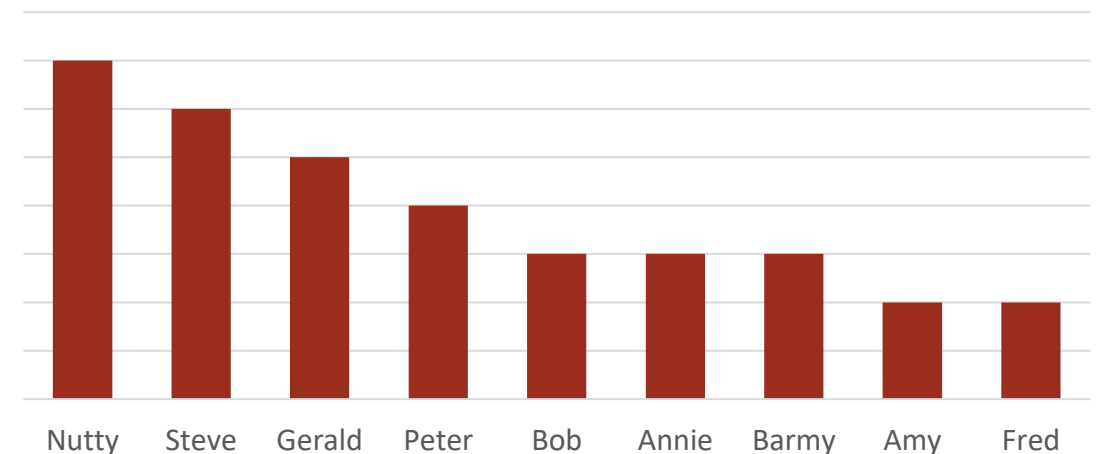
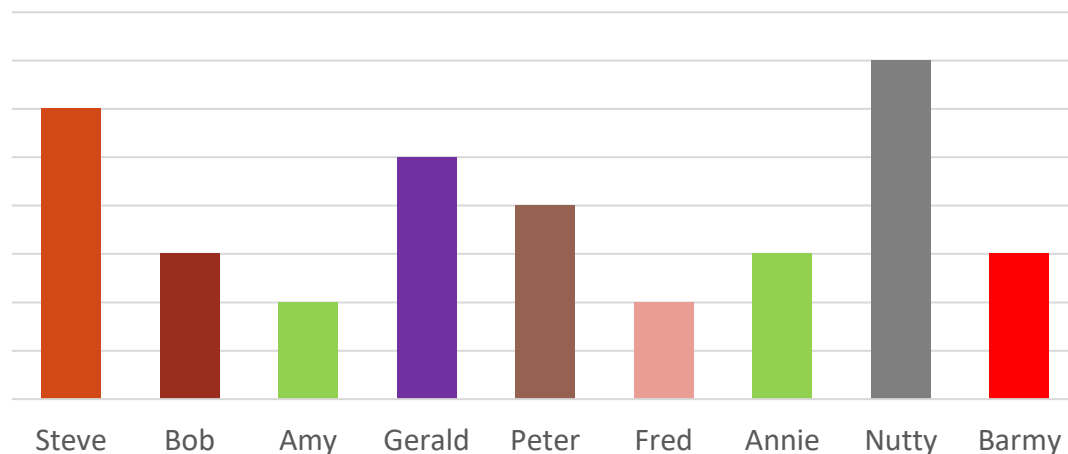
Gestalt Principles

- simplicity
- continuation
- proximity
- similarity (invariance)
- focal point
- isomorphic correspondence
- figure / ground duality
- common fate*
- closure*
- uniform connectedness*

Gestalt Principles – Simplicity

The brain has a preference for **simplicity** – it tends to process simple patterns faster than patterns that are more complex.

Lesson: arrange data simply and logically wherever possible.

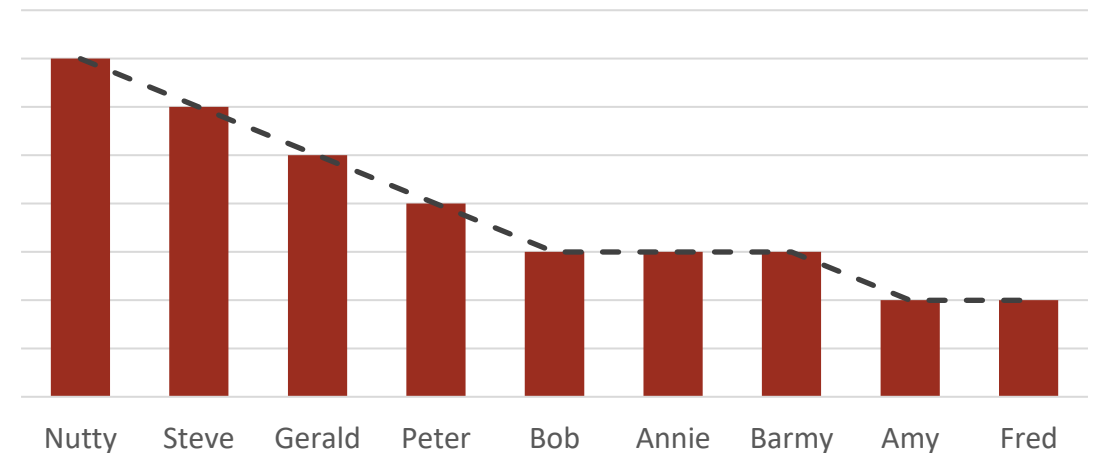
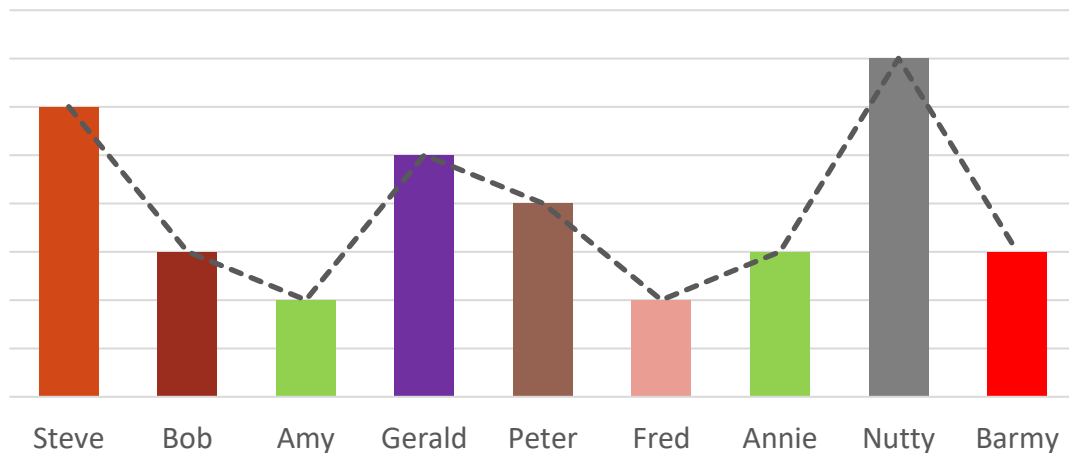


Gestalt Principles – Continuation

Our eyes group things that are **aligned** (e.g., sorted from high to low).

In the chart on the right the eyes follow a **continuous path**; it makes the whole chart more readable because of the continuous downward direction

Lesson: arrange objects in a line to facilitate grouping and comparison.



Gestalt Principles – Proximity

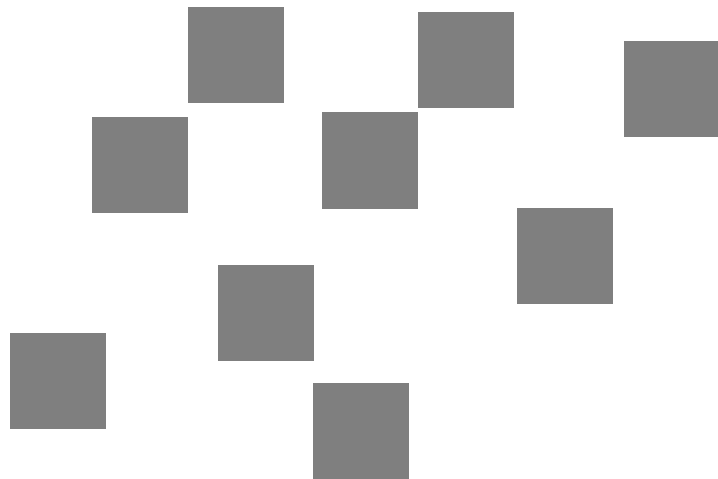
Objects/shapes that are in **proximity** (close) appear to form **groups**.

The effect generated by the collected group is more “powerful” than that generated by separate elements.

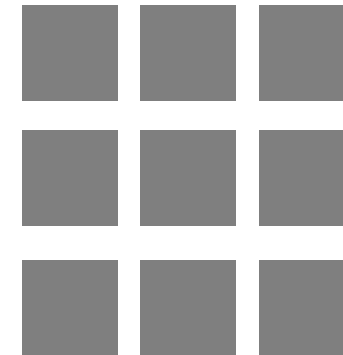
Elements which are grouped together create the **illusion** of shapes/planes in space, even if the elements are not touching.

Lesson: understand the chart’s priorities and create groupings through proximity that support those priorities.

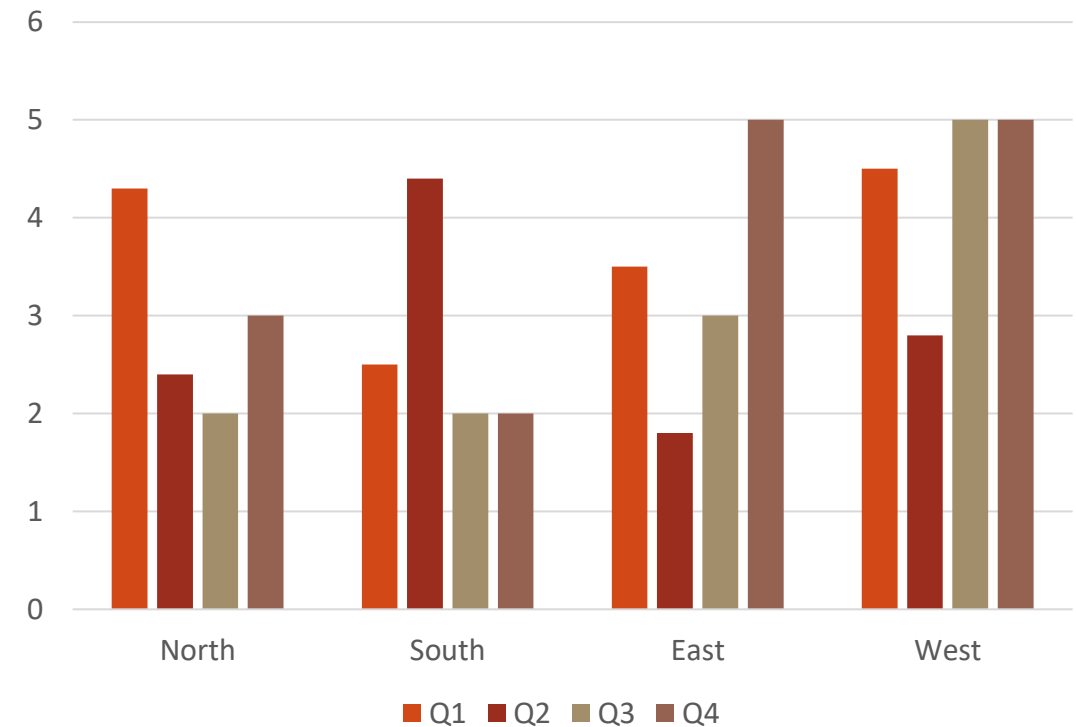
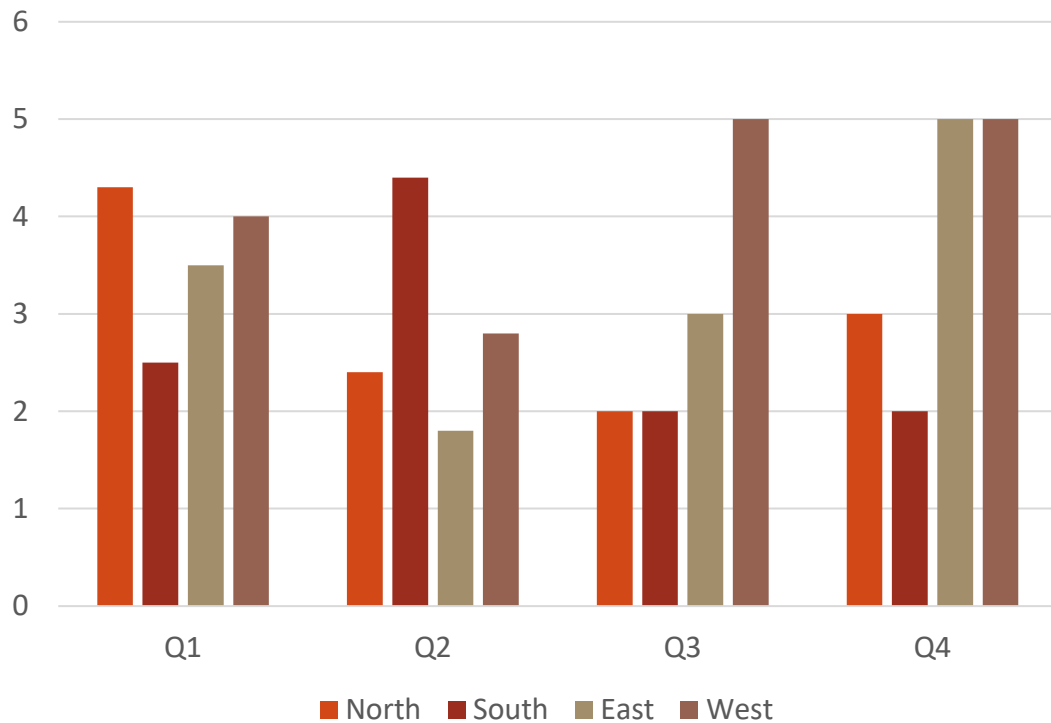
Gestalt Principles – Proximity



VS



Gestalt Principles – Proximity



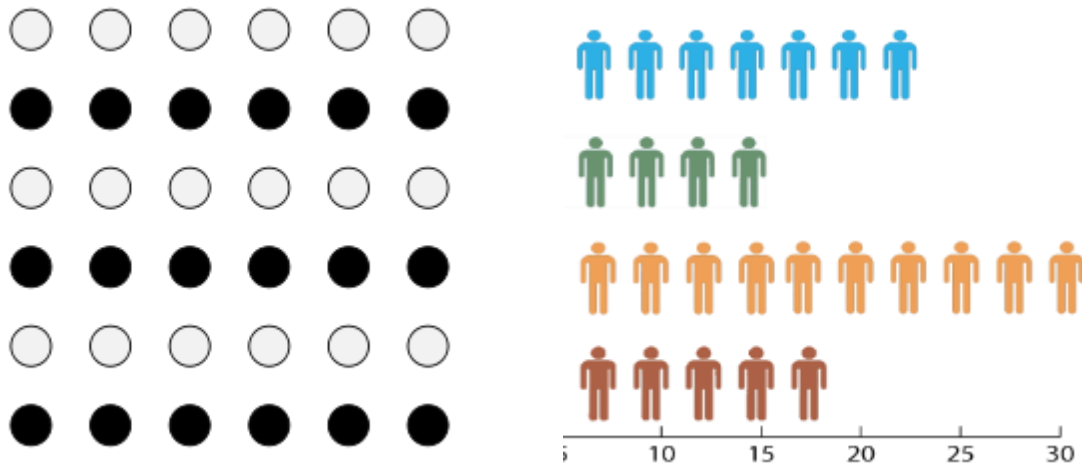
Gestalt Principles – Similarity

Stimuli that physically resemble each other are viewed as **part of the same object**; stimuli that don't are viewed as part of a different object.

Similarity and proximity often come together to form a **visual hierarchy**. Either principle can dominate the other, depending on their application and combination.

Lesson: use similar characteristics to establish relationships and to encourage groupings of objects.

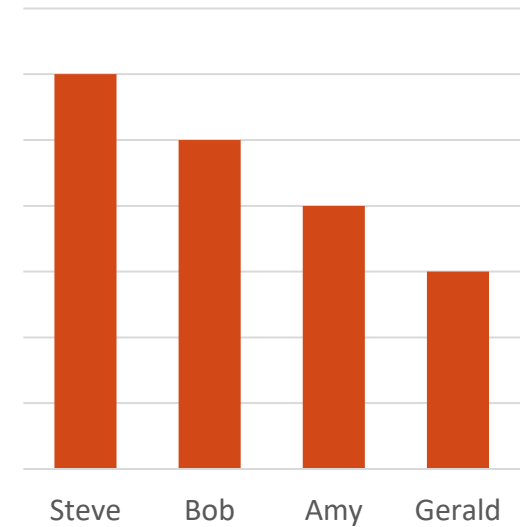
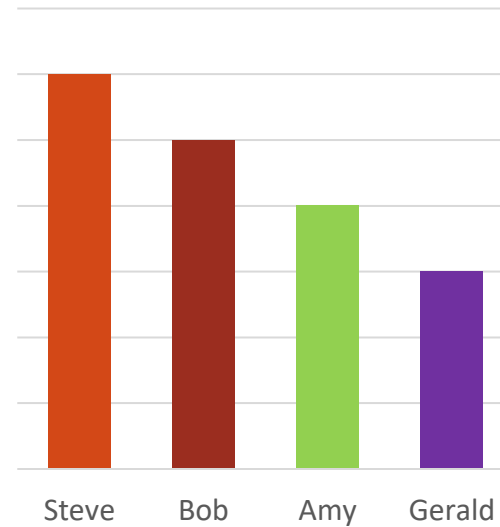
Gestalt Principles – Similarity



In these examples, similarity dominates over proximity: we see rows before we see columns.

Gestalt Principles – Similarity

Making things similar can reduce cognitive load (colour).

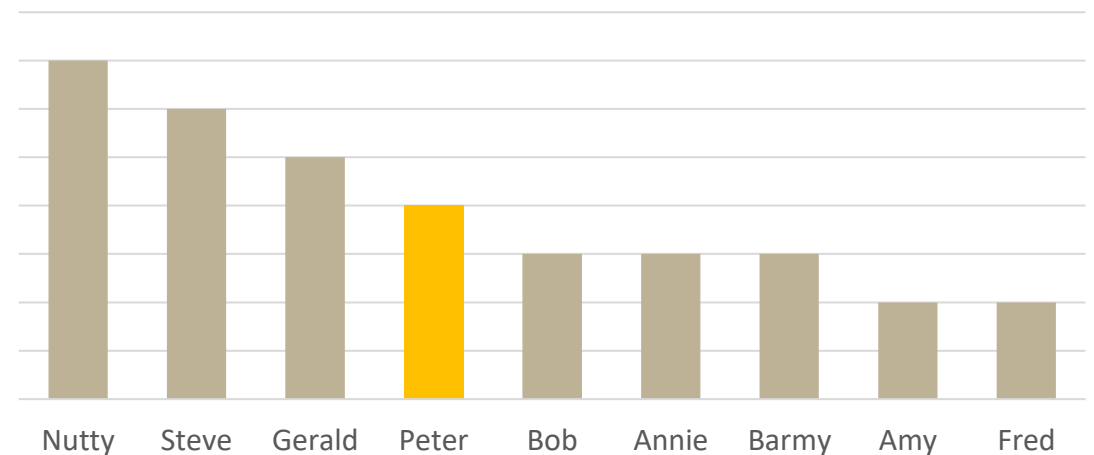
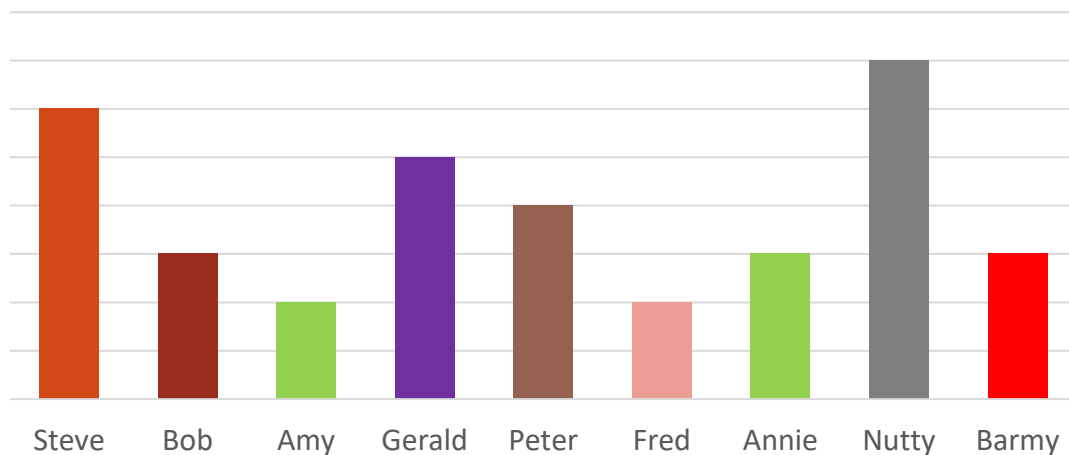


Gestalt Principles – Focal Point

In opposition to similarity, the **focal point** principle states that distinctive-looking objects can create a focal point.

To highlight one salesperson's performance, make their bar graph color different.

Lesson: use different characteristics to highlight and create focal points.

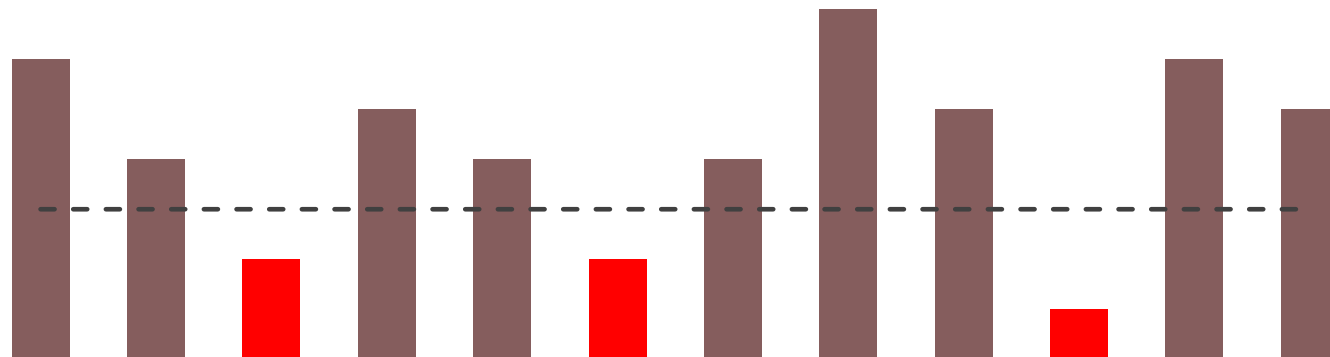


Gestalt Principles – Correspondence

People interpret and respond to images based on past/shared experiences (in particular, for the selection of chart colours).

Red is often associated with “**bad**” and **green** with “**good**” (colour-blindness?). We can colour-code charts accordingly.

Lesson: stick to well-established conventions/best practices (even if boring!)



Gestalt Principles – Duality

Chart elements are either perceived as **figures** (focus) or as (back)**ground**.

Foreground objects are **promoted** by the brain, background objects are **demoted**.

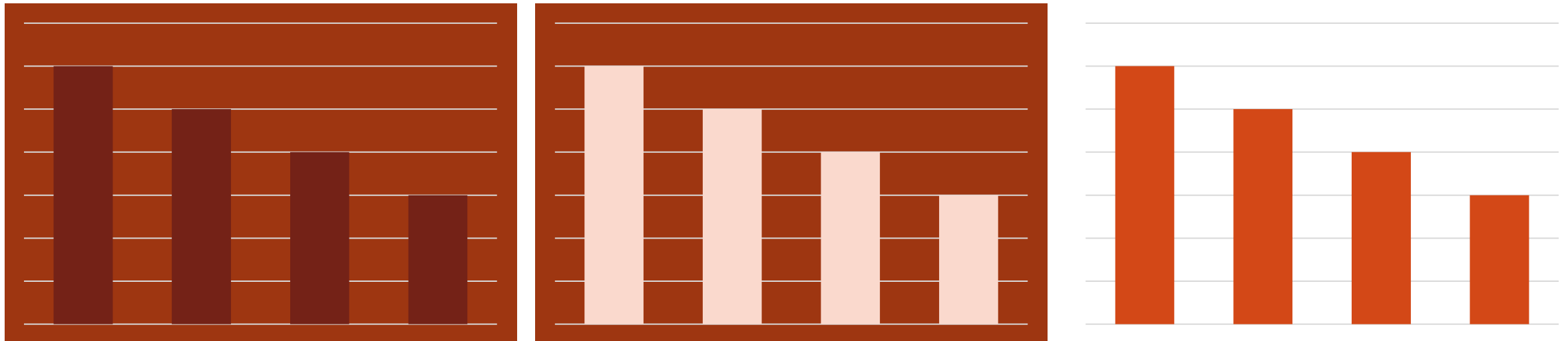
Strong contrast makes it easier to distinguish between the two types of objects.

Lesson: ensure there is enough contrast between the chart foreground (figures) and their background.

Gestalt Principles – Duality

Because of the **low contrast** between the figure and background in the chart on the left, there is an **additional cognitive load**.

Increasing the contrast on the right improves readability.



Decluttering

Clutter is the enemy!

Every element on a page adds cognitive load

- identify and remove anything that isn't adding value
- think of cognitive load as mental effort required to process information (lower is better)

Tufte refers to the **data-to-ink ratio** – “the larger the share of a graphic's ink devoted to data, the better”

In *Resonate*, Duarte refers to this as “**maximizing the signal-to-noise ratio**” where the signal is the information or the story we want to communicate.

Decluttering

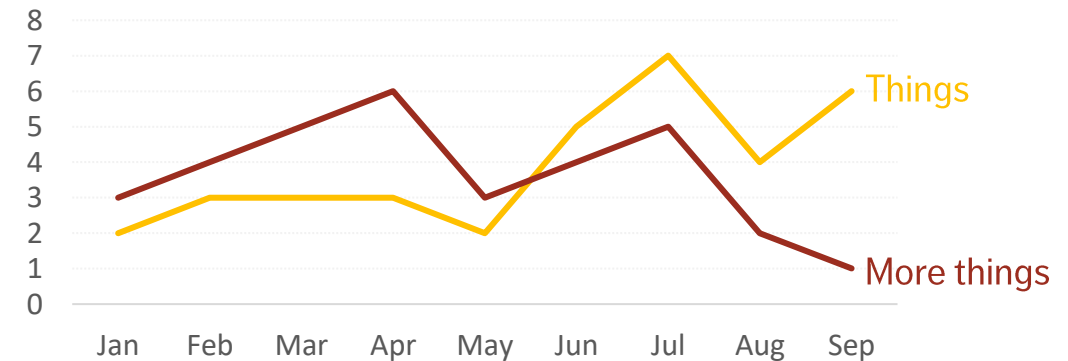
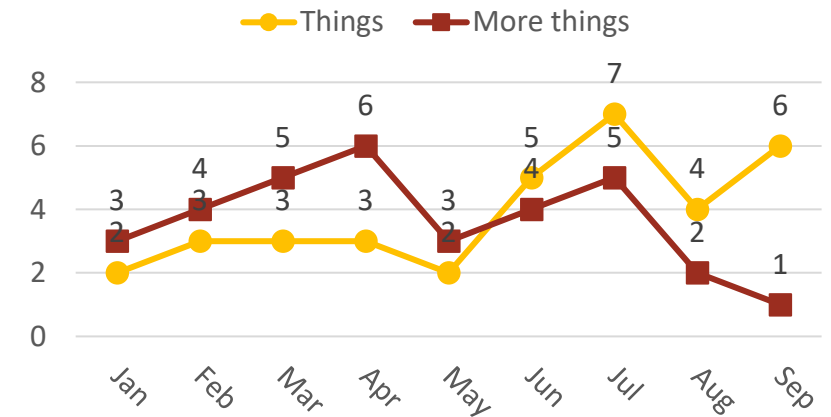
Use the Gestalt Principles to **organize/ highlight** data in the chart.

Align all elements (graphs, text, lines, etc.):

- don't rely on eye, use position boxes and values

Charts:

- remove border, gridlines, data markers
- clean up axis labels
- label data directly



Decluttering

Use **consistent** fonts, font size, colour and alignment.

Don't rotate text to anything other than 0 or 90 degrees (however: English/French incompatibility with vertical text).

Use **white space**:

- margins should remain free of text and visuals
- don't stretch visuals to edge of page or too close to other visuals
- think of white space as a border

Colour Theory

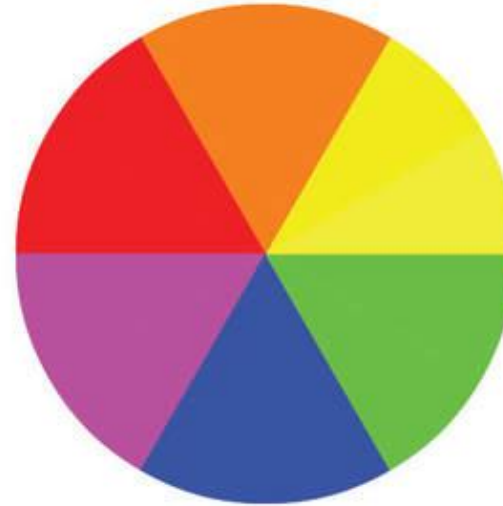
It is a complicated topic – here is a start:

- <http://www.deanenettles.com/webexamples/colorexamples/>
- <https://www.sessions.edu/color-calculator/>

Colour wheels:



Primary Colours



Secondary Colours



Tertiary Colours

Colour Schemes

Achromatic (colourless, using only blacks, whites and grays)



Monochromatic (1-colour schemes)



Complementary (colours directly across from each other on the colour wheel)



Split complementary (2 of the 3 colors are adjacent; 1 of the colours is opposite)



Colour Tips

When it comes to colour, **less is more**: use it sparingly (graphic designers are taught to “get it right, in black and white”).

Based on the Gestalt Principles, **monochrome** schemes can be effective.

When appropriate, pick corporate identity scheme (this maximizes buy in).

Create a **template** (and stick to it).

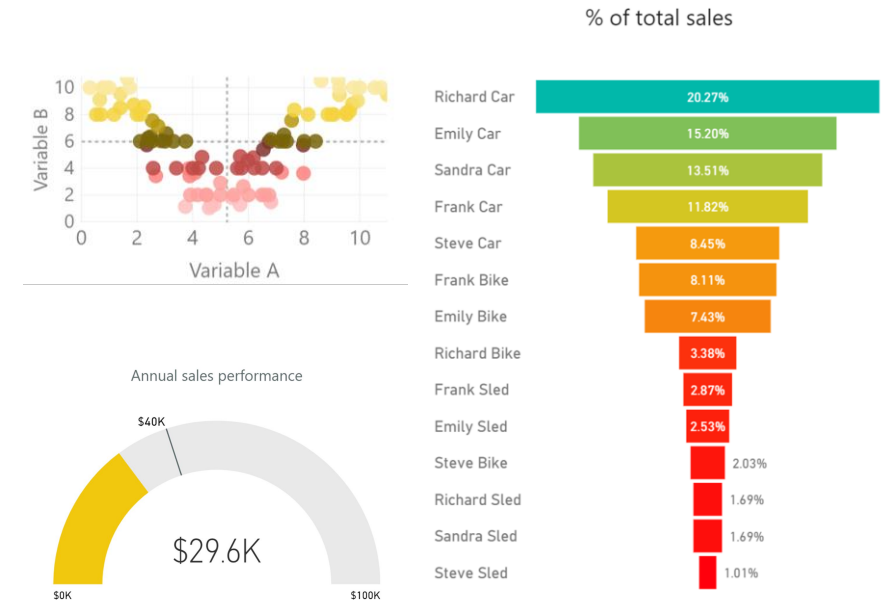
Upload images to see what charts look like for flavours of colourblindness:

- <https://www.color-blindness.com/coblis-color-blindness-simulator> (not the only tool)

Size Considerations

Assuming that the charts has been decluttered:

- things of equal importance size similarly
- other things scale to importance



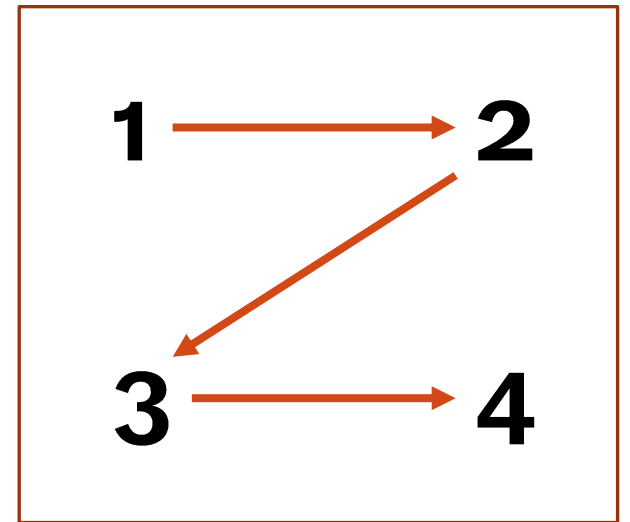
Position Considerations

How should the elements be placed in a chart/dashboard?

In the West, most people start at the **top left** and zig-zag all the way to the **bottom right**.

Simple rule: don't make people work too hard

- main message: top left/top right
- info in order of preference
- people concentrate less as they scan so get less complex as you move to bottom corner



Session 2

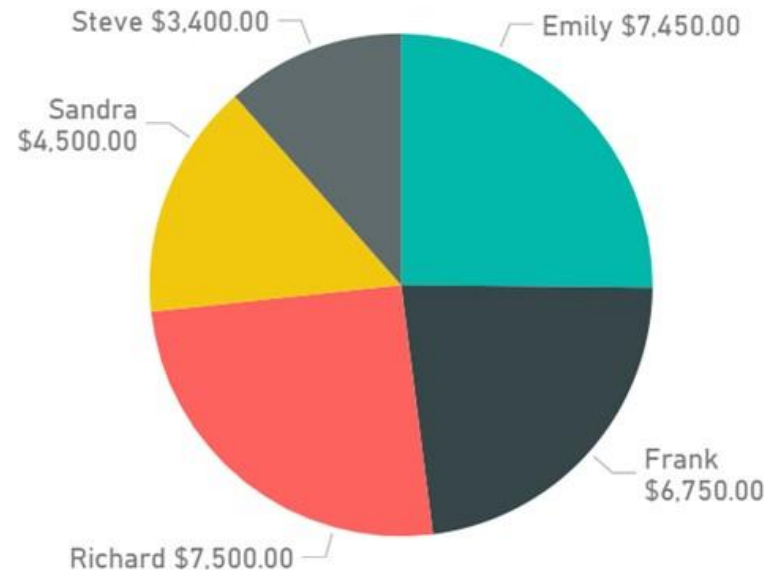
Sales Dashboard

\$ sales



\$ sales by Salesperson

Salesperson Emily Frank Richard Sandra Steve



\$ sales by Month and Salesperson

Salesperson Emily Frank Richard Sandra Steve



\$ sales by Product and Salesperson

Product Car Bike Sled



Sales Dashboard

Annual Sales for 2017

Total Sales

\$29.6K



Suggested Reading

Chart Aesthetics

The Practice of Data Visualization **Part II: The Foundations of Visual Design**

4. The Mechanics of Visual Perception

5. Visual Design and Data Charts

5.1 Visualization and Memory

5.2 Aesthetic Considerations

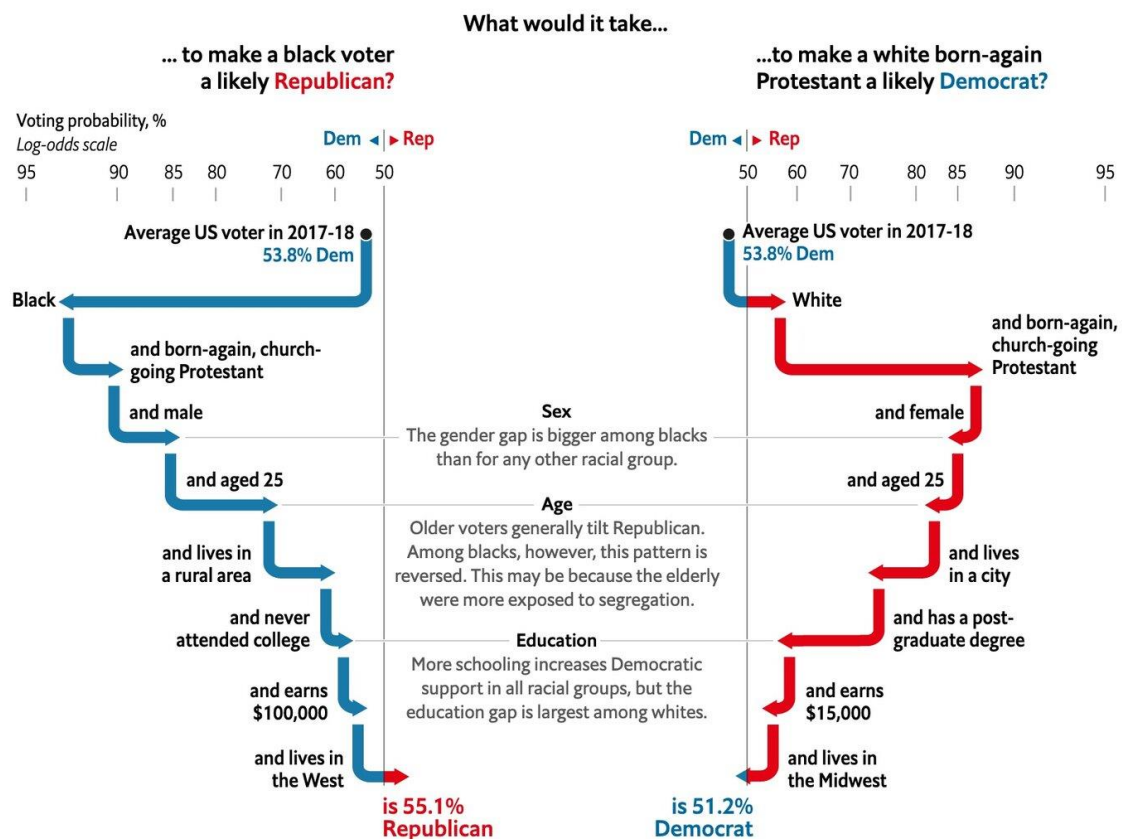
5.3 Decluttering

Exercises

Chart Aesthetics

Comment on the aesthetics of the following charts, according to:

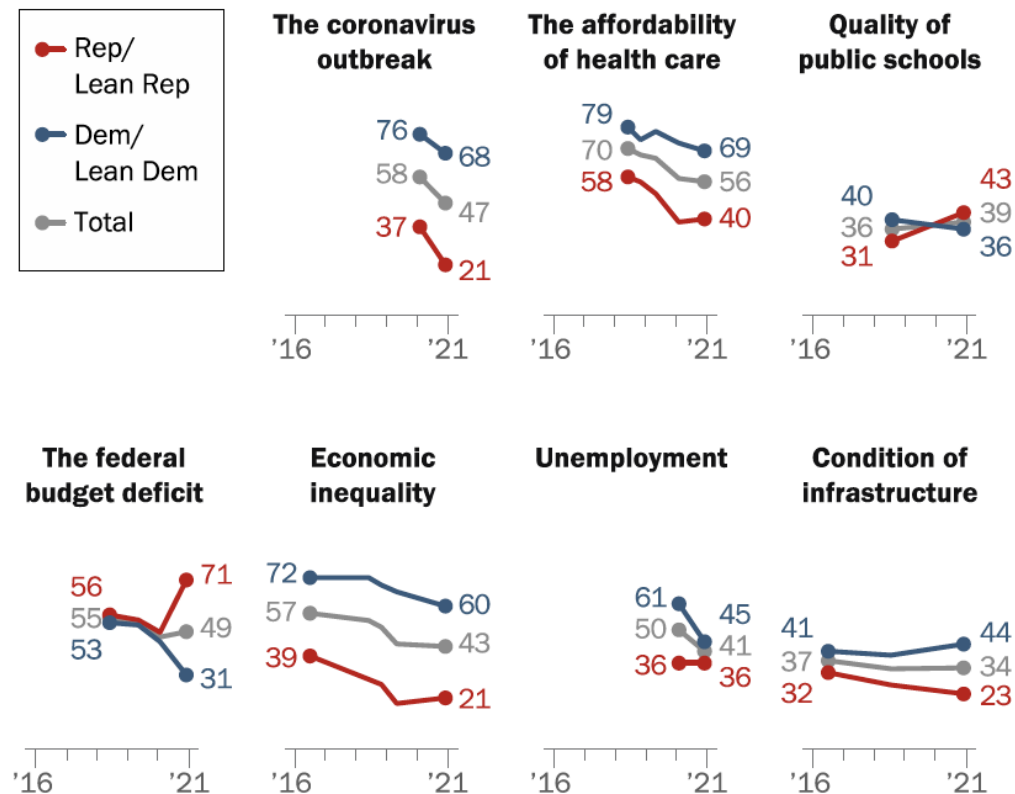
- Gestalt principles
- use of colours
- lack of clutter
- size and position
- etc.



Sources: YouGov; *The Economist*

Republican concern about the budget deficit increases sharply; Democratic concern declines

% who say ____ is a *very big problem* in the country today

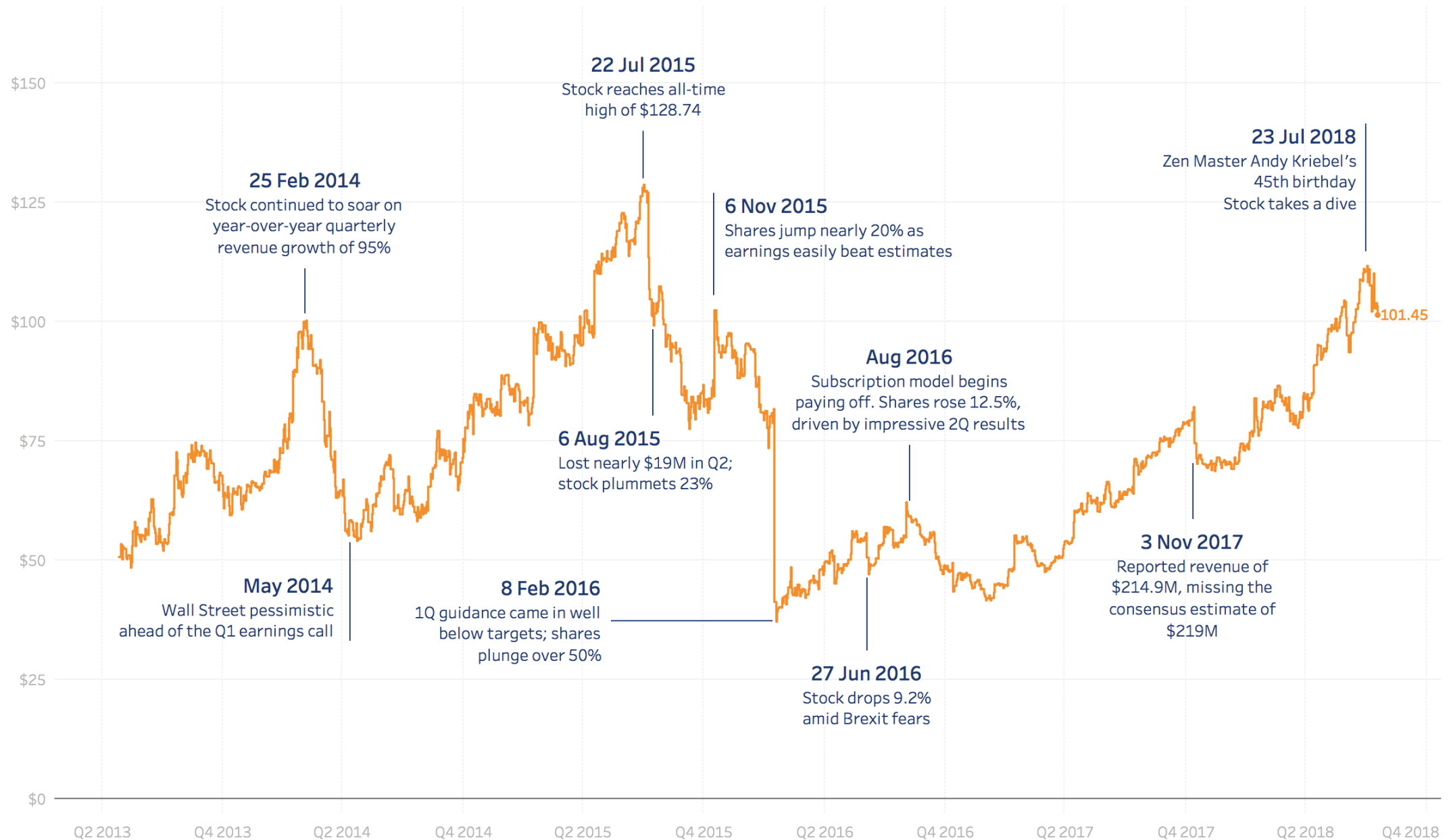


Note: March 2019 and earlier wording for economic inequality was "The gap between the rich and poor." See topline for details.

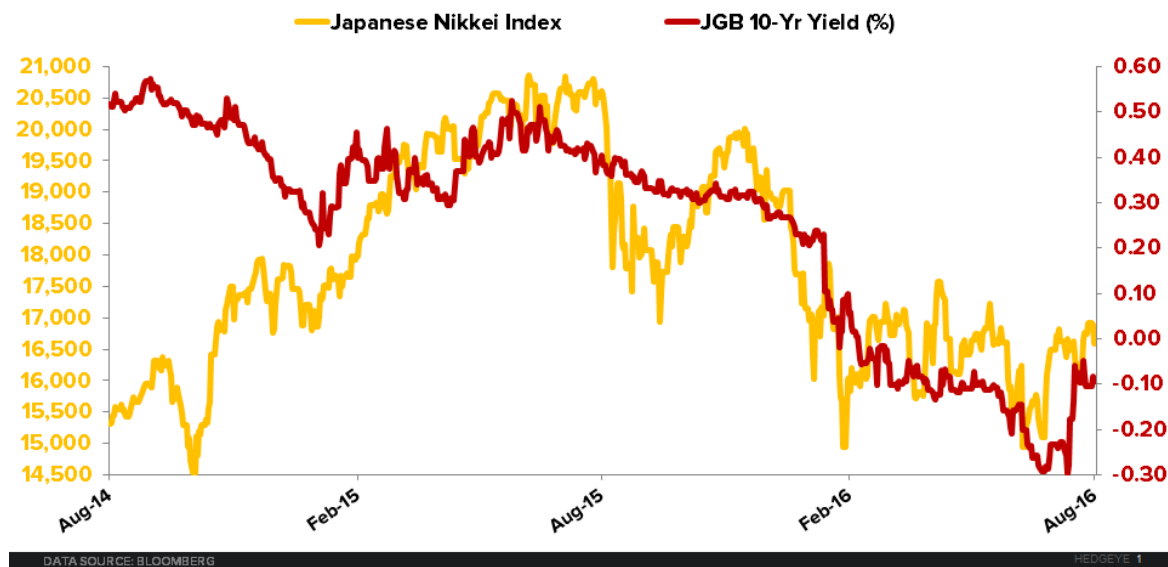
Source: Survey of U.S. adults conducted April 5-11, 2021.

PEW RESEARCH CENTER

The Roller Coaster Ride of Tableau's Stock

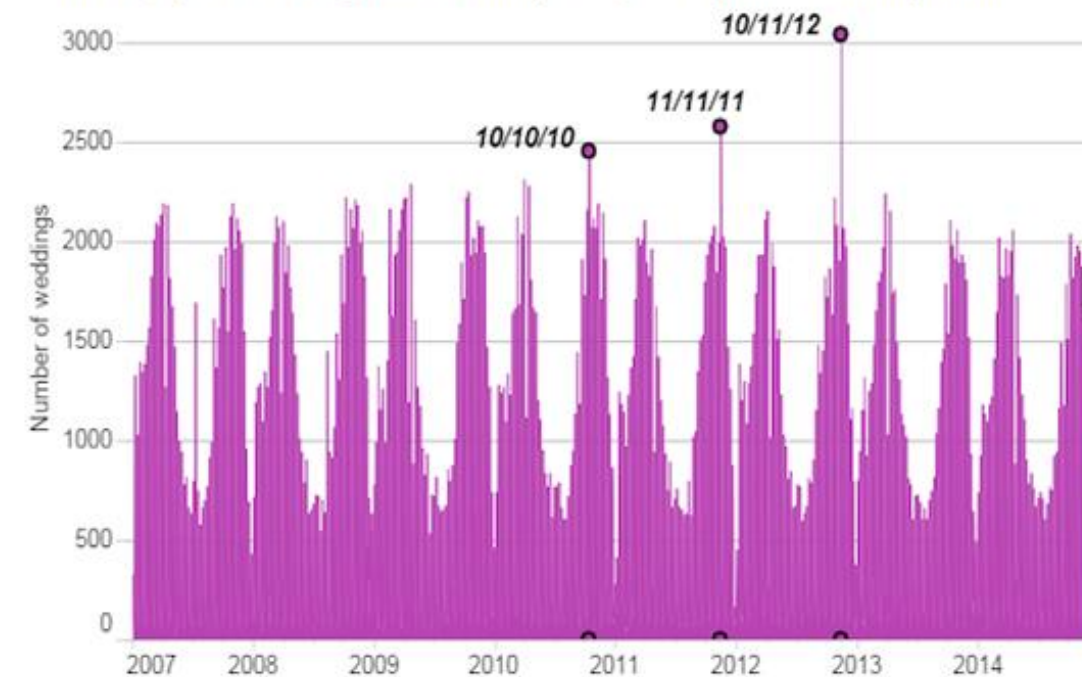


BAD ISN'T GOOD FOREVER: JAPAN

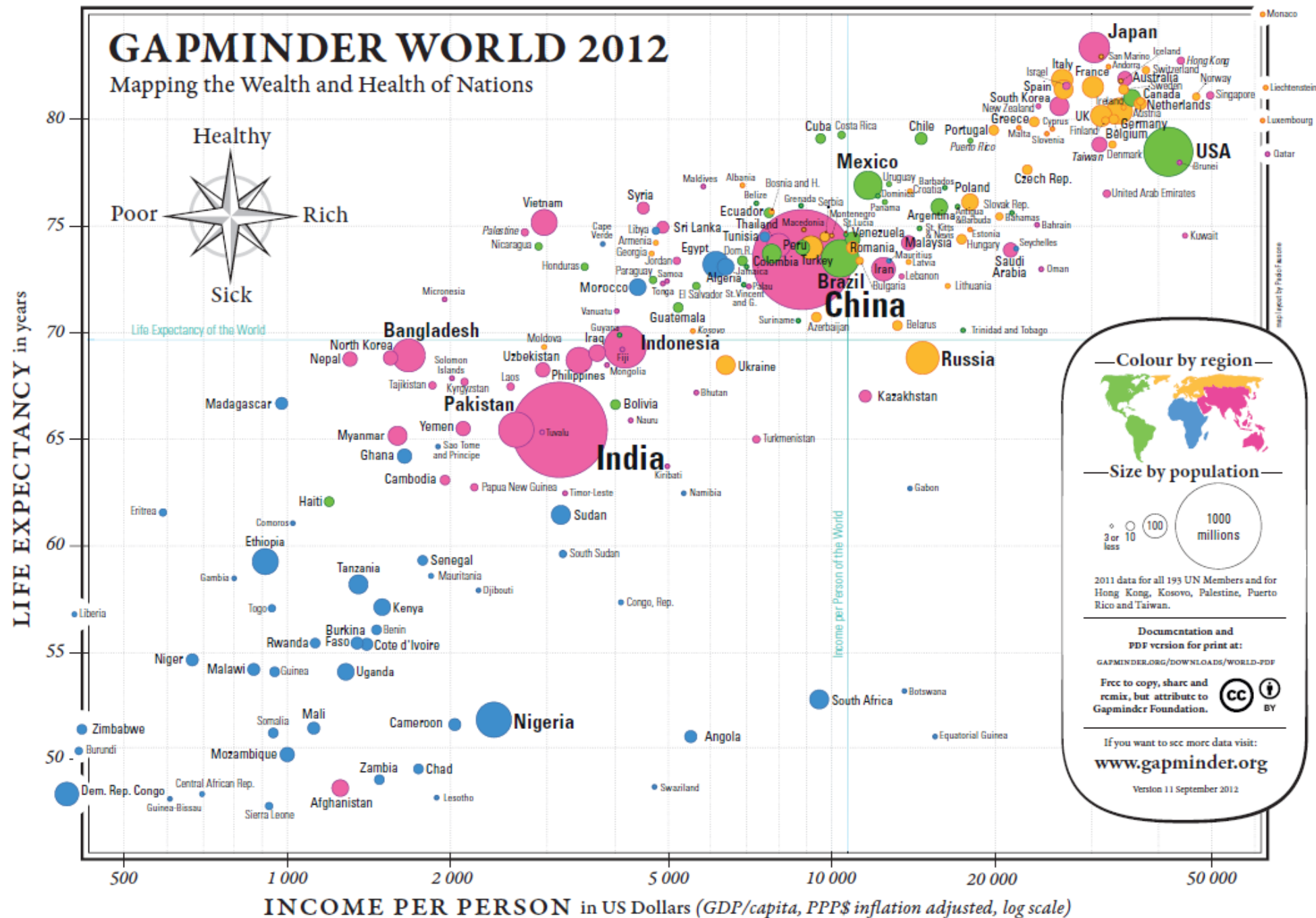


Weddings in Australia

The most popular wedding dates form repeating or sequential number patterns.



Graphic: Inga Ting | Source: ABS 2015





1097
Total Students in five years

687
Total Students in 2015-2016

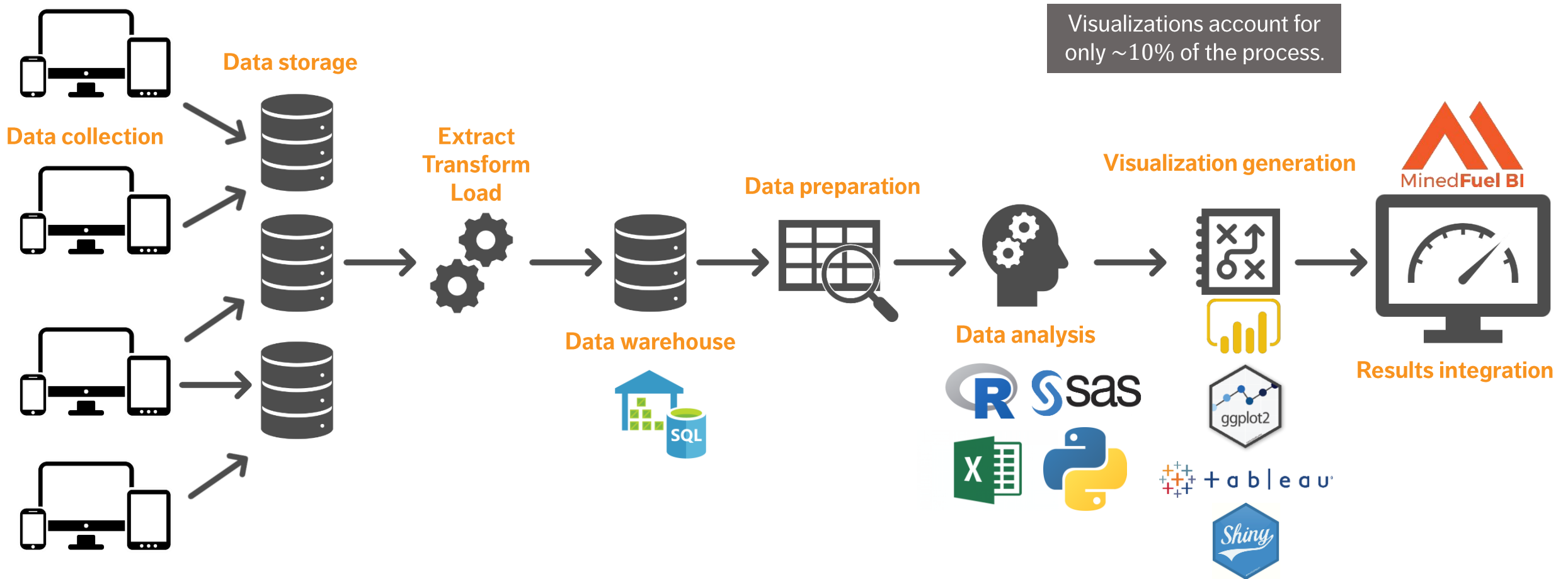
21
Total Classes in five years

7.7 of 8
Most recent instructor rating (out of 8.0)



5. Dashboards

Data Environment

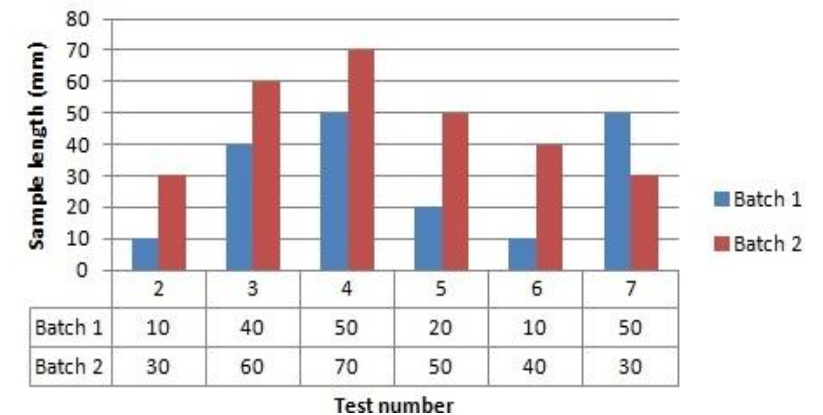


Overview

The past is **data-driven**:

- mostly Excel (or reporting tools like Cognos)
- mostly numbers, tables and non-interactive graphs
- distributed on desktop computers, by email, in PowerPoint presentation
- static, mostly backwards looking (lagging indicators)
- KPIs and dashboards were somewhat contrived

Region	Jan-06	Feb-06	Mar-06	Apr-06	May-06	Jun-06	Jul-06
Actuals							
Seattle	111	653	1,598	3,411	3,972	5,092	5,298
Boise	26,779	27,867	29,153	30,557	33,402	35,400	35,450
Portland	33,078	34,401	37,535	39,916	41,357	45,306	46,670
Spokane	25,417	26,669	28,092	29,020	29,674	30,501	30,830
North Region	199,841	211,053	226,789	242,957	256,605	273,640	277,770
Plan							
Seattle	693	468	790	1,383	2,205	3,180	4,210
Boise	29,525	26,062	27,088	28,269	29,536	30,821	32,160
Portland	32,276	34,708	36,737	38,857	41,066	43,364	45,750
Spokane	30,500	26,644	27,987	29,430	30,994	32,594	34,230
North Region	191,783	203,916	216,524	230,474	246,390	263,378	281,220
Variance							
Seattle	-582	185	808	2,029	1,767	1,912	1,078
Boise	-2,746	1,805	2,064	2,288	3,866	4,578	3,288
Portland	802	-307	798	1,059	291	1,942	920
Spokane	-5,082	25	105	-410	-1,320	-2,093	-3,398
North Region	-8,057	7,437	10,265	12,483	10,215	10,264	-3,410



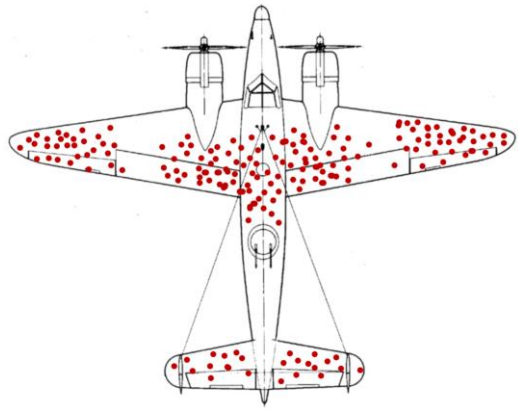
Overview

The future is **story-driven**:

- new tools: Power BI, Tableau, Qlickview, Shiny, etc.
- mostly visualizations, occasional numbers and tables
- distributed on the web (internal and external)
- dynamic and both backwards and forwards looking (leading and lagging indicators)
- data for everyone

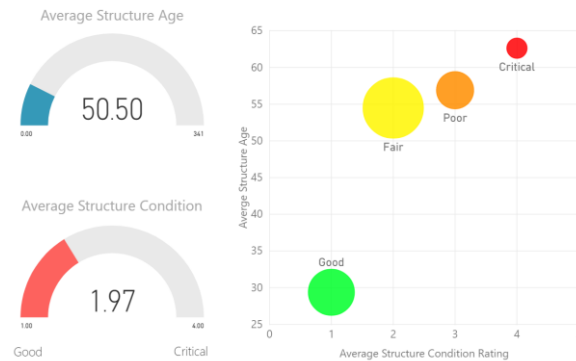


Defining Context



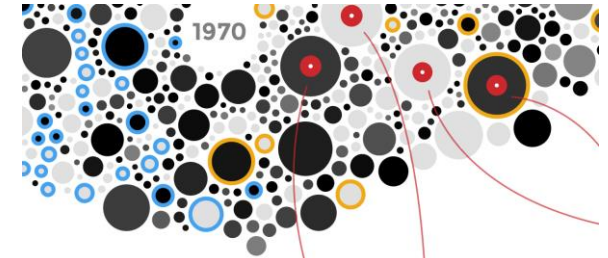
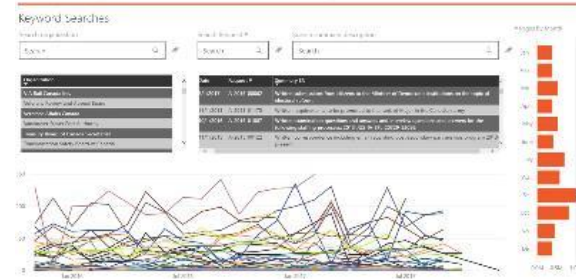
Directory of Federal Real Property (DFRP) Dashboard

You have selected 20,186 properties that contain 35,148 structures



Access to Information and Privacy (ATIP) search

You have currently selected 28,711 requests totaling 6,597,612 pages of information



The Beatles

No other artist or band has more songs in the Top 2000 as the Beatles. With 38 songs they are responsible for 14% of all titles before 1970. Nonetheless, only 5 years ago they still had 50 songs in the list.

- 4 Piano Man
Billy Joel: 1974
- 5 Child in Time
Deep Purple: 1972

Seconds

Minutes

Fraction of Hour

Hours

Infographics/Data Viz

Dashboards

Reports and Exploration

Data Art

Dashboards

A **dashboard** is any visual display of data used to monitor conditions and/or facilitate understanding.

Examples:

- interactive display that allows people to explore motor insurance claims by city, province, driver age, etc.
- PDF showing key audit metrics that gets e-mailed to a Department's DG on a weekly basis.
- wall-mounted screen that shows call centre statistics in real-time.
- mobile app that allow hospital administrators to review wait times on an hourly- and daily-basis for the current year and the previous year.

Some Questions To Consider

In a car's dashboard, a small number of **key indicators** (speed, gasoline level, etc.) need to be understood **at a glance**. Dashboard designs that do not take these characteristics under consideration have catastrophic consequences.

The following questions need to be answered prior to the dashboard being designed:

- who is the dashboard's consumer?
- what story does the dashboard tell?
- what data (categories) will be used?
- what will appear on the dashboard?
- how can the dashboard help the consumer?



Dashboard Design Guidelines

Nick Smith suggests the following 6 Golden Rules:

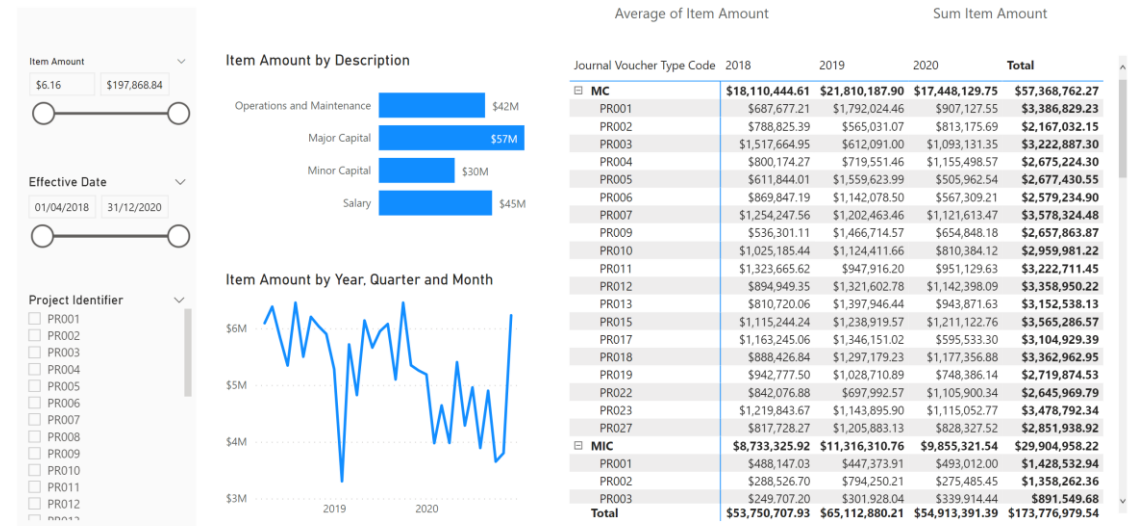
- **Consider the audience** (who are you trying to inform? does the DG really need to know that the servers are operating at 88% capacity?)
- **Select the right type of dashboard** (operational, strategic/executive, analytical)
- **Group data logically, use space wisely** (split functional areas: product, sales/marketing, finance, people, etc.)
- **Make the data relevant to the audience** (scope and reach of data, different dashboards for different departments, etc.)
- **Avoid cluttering the dashboard** (present the most important metrics only)
- **Refresh your data at the right frequency** (real-time, daily, weekly, monthly, etc.)

Dashboard Types

Exploration: using visualizations as a tool to explore/understand the data

- high level of interactivity
- high level of detail
- **all** aspects of data should be represented (tables, columns, calculations etc.)
- no annotations or explanations required

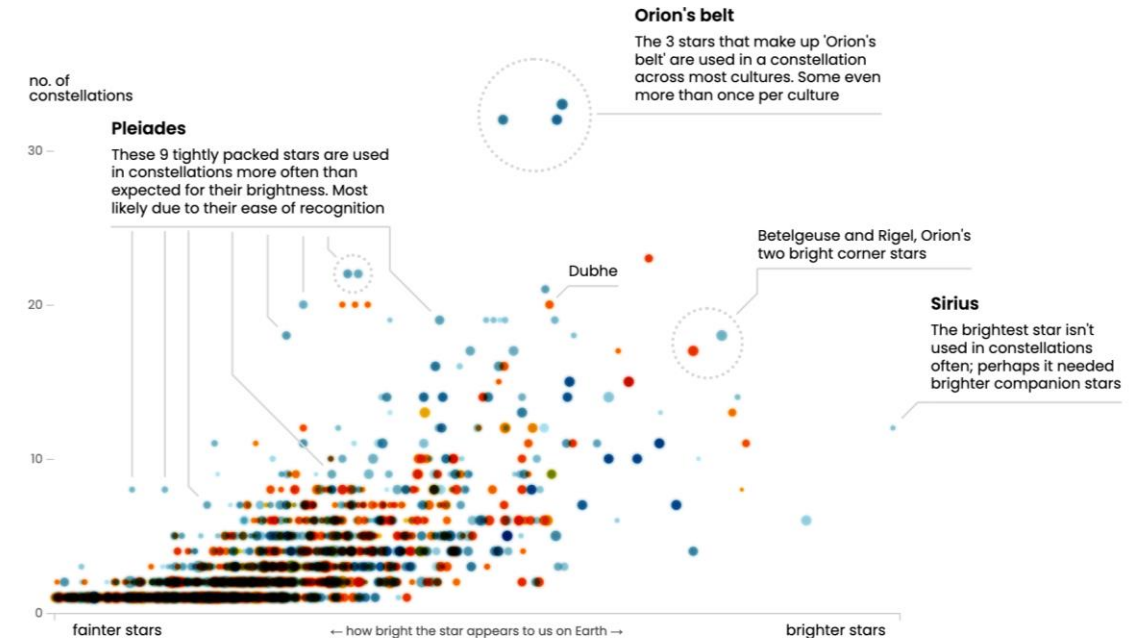
Financial Data Exploration



Dashboard Types

Storybook: using visualizations as a tool to explain the data and communicate the story

- low level of interactivity
- low level of detail
- key aspects of data should be represented
- annotations and explanations drive the “story”

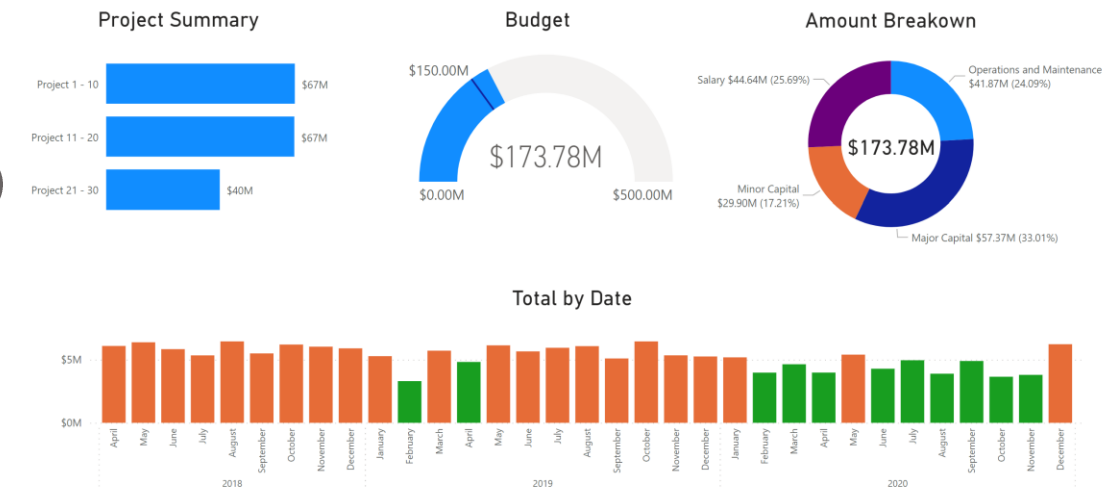


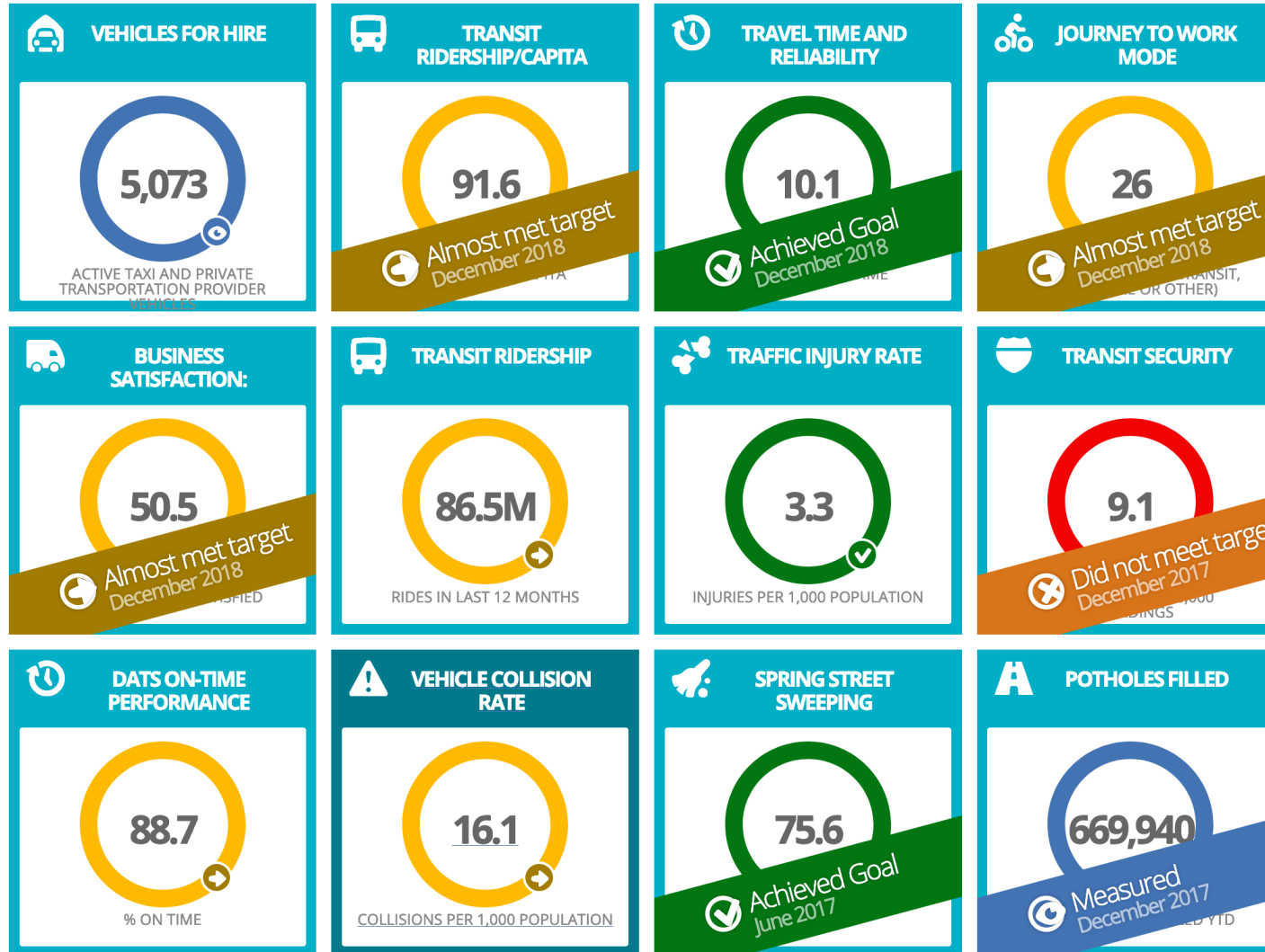
Dashboard Types

Situational Awareness: using visualizations as a tool to provide a snapshot of the data

- medium level of interactivity
- not “scripted” but well organized (e.g., categorized)
- summary data should be represented
- anomalies are highlighted
- often used for internal presentations

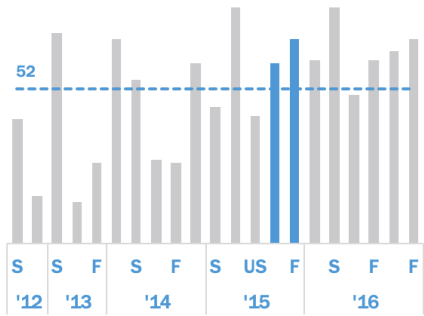
Financial Snapshot





✓ Meets or Exceeds Target
 ➕ Near Target
 ✖ Needs Improvement
 🕒 Measuring
 📊 Collecting Data

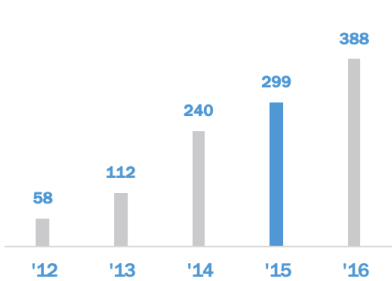
Students



1097

Total Students in five years

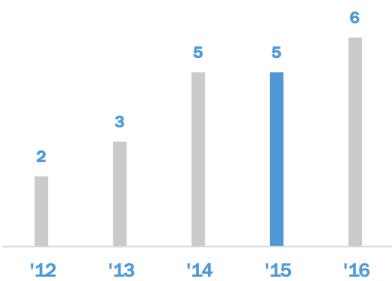
Enrollments



687

Total Students in 2015-2016

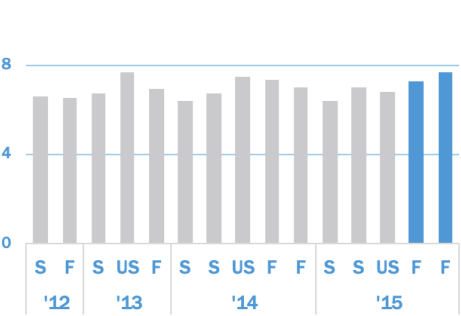
Classes



21

Total Classes in five years

Ratings



7.7 of 8

Most recent instructor rating (out of 8.0)

Semesters

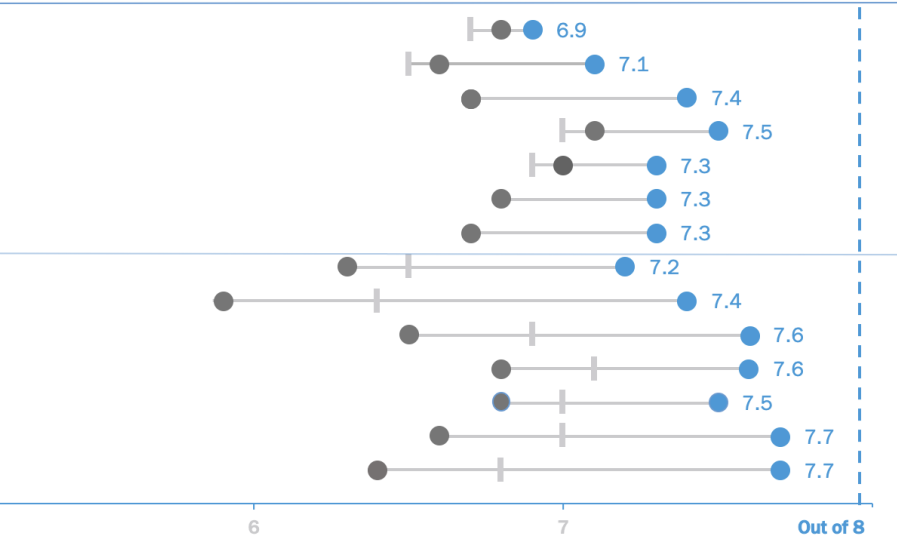
2015 Fall Semester 001	I developed specific skills and competencies
	Overall, this was an excellent course
	The instructor communicated clearly
	The Instructor graded fairly
	The instructor was well organized
2015 Fall Semester 002	The instructor interacted well with students
	Overall, this instructor was excellent
	I developed specific skills and competencies
	Overall, this was an excellent course
	The instructor communicated clearly
	The Instructor graded fairly
	The instructor was well organized
	The instructor interacted well with students
	Overall, this instructor was excellent

Questions

I developed specific skills and competencies
Overall, this was an excellent course
The instructor communicated clearly
The Instructor graded fairly
The instructor was well organized
The instructor interacted well with students
Overall, this instructor was excellent
I developed specific skills and competencies
Overall, this was an excellent course
The instructor communicated clearly
The Instructor graded fairly
The instructor was well organized
The instructor interacted well with students
Overall, this instructor was excellent

● BANA ● College ● Shaffer

Ratings



Suggested Reading

Dashboards

Data Understanding, Data Analysis, Data Science
Volume 2: Fundamentals of Data Insight

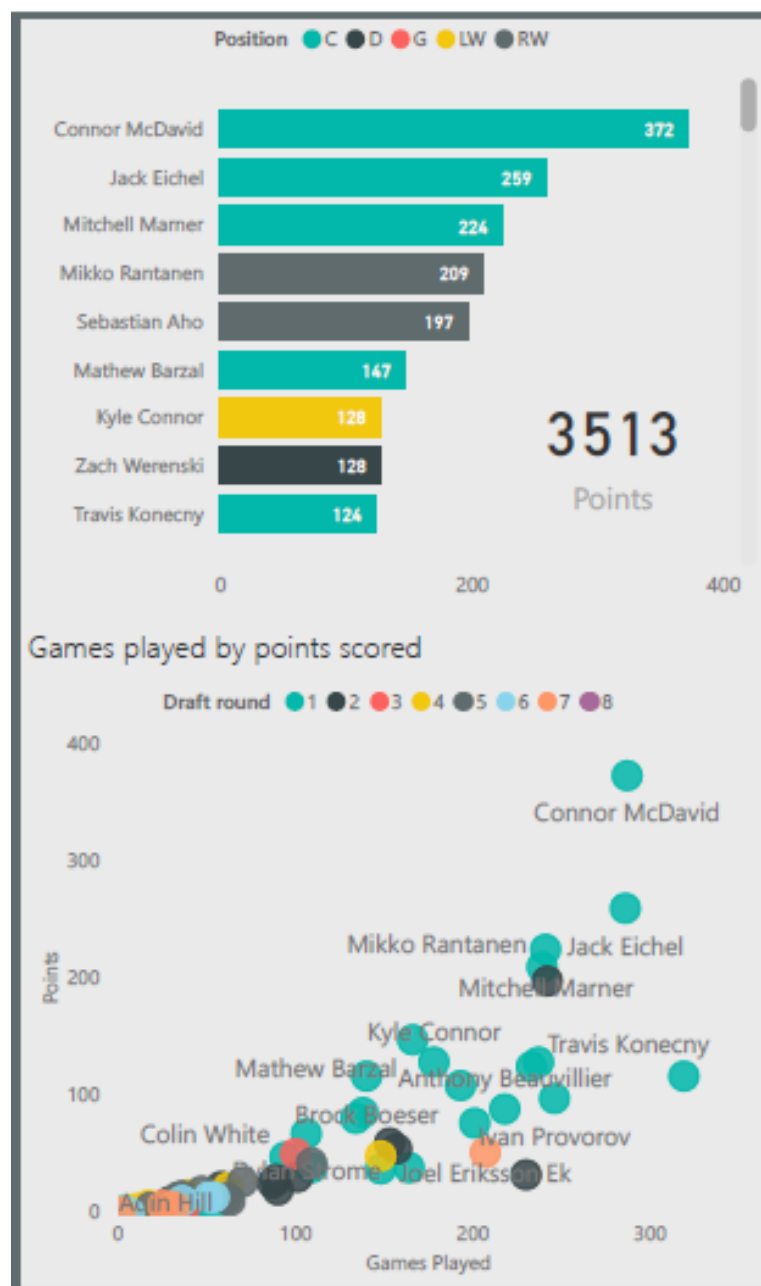
18. Data Exploration and Data Visualization
18.3 Dashboards

Exercises

Dashboards

1. In teams or individually, identify a few data visualizations that appeal to you. What is the story being told by the visualization? What kind of data is needed to build these visualizations?
2. In teams or individually, identify work scenarios for which data visualization could prove useful. What insight could be drawn from such visualizations? Would such visualizations get a buy-in from your supervisors/employers? How much work would be required to get from design to completion? Are the obstacles mostly of a technical nature? Related to data procurement?
3. Consider the following dashboards. Can you figure out, at a glance, who their audience is? What are their types? Their strengths and limitations? How could you improve them?

Top Scorers

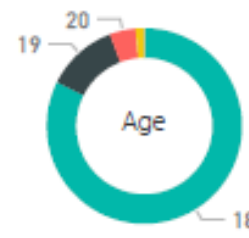
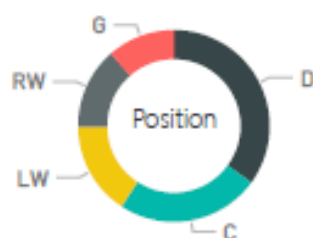
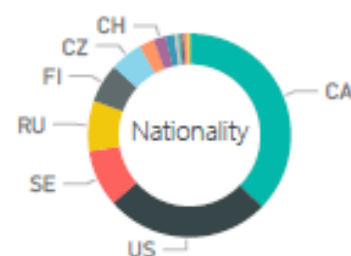
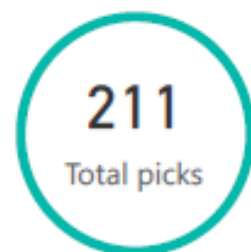


2015 NHL Draft Class

Select NHL team

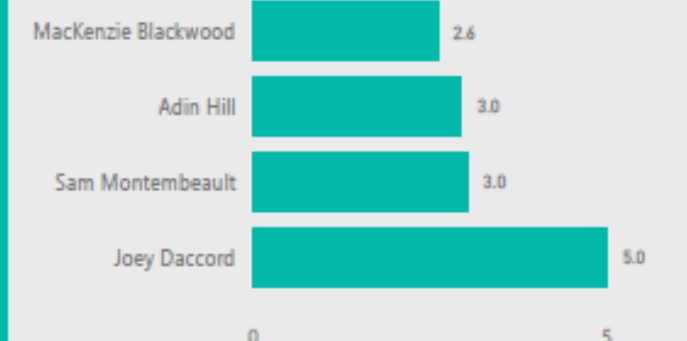
All

Overall	Player	Team	Position	Games Played
1	Connor McDavid	Edmonton Oilers	C	287
2	Jack Eichel	Buffalo Sabres	C	286
3	Dylan Strome	Arizona Coyotes	C	106
4	Mitchell Marner	Toronto Maple Leafs	C	241
5	Noah Hanifin	Carolina Hurricanes	D	319
6	Pavel Zacha	New Jersey Devils	C	201
7	Ivan Provorov	Philadelphia Flyers	D	246
8	Zach Werenski	Columbus Blue Jackets	D	237
9	Timo Meier	San Jose Sharks	RW	193
10	Mikko Rantanen	Colorado Avalanche	RW	239

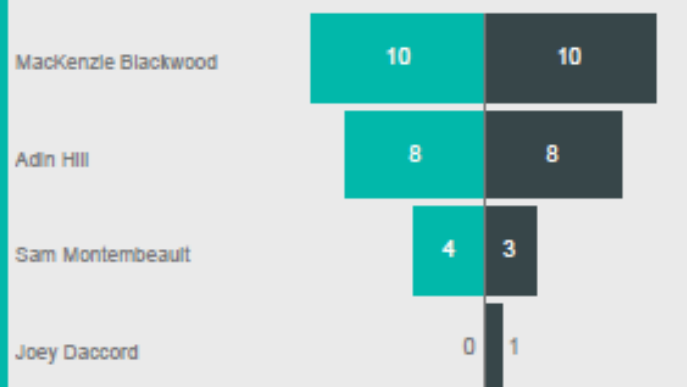


Goalies

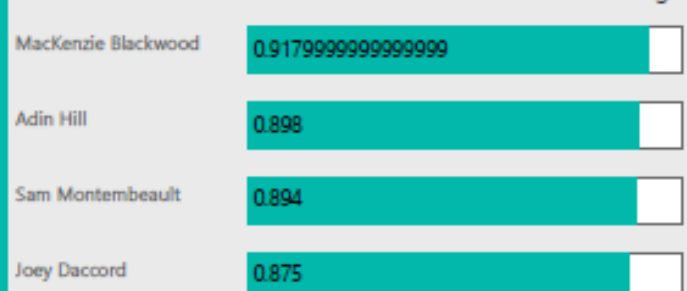
Goals Against Average



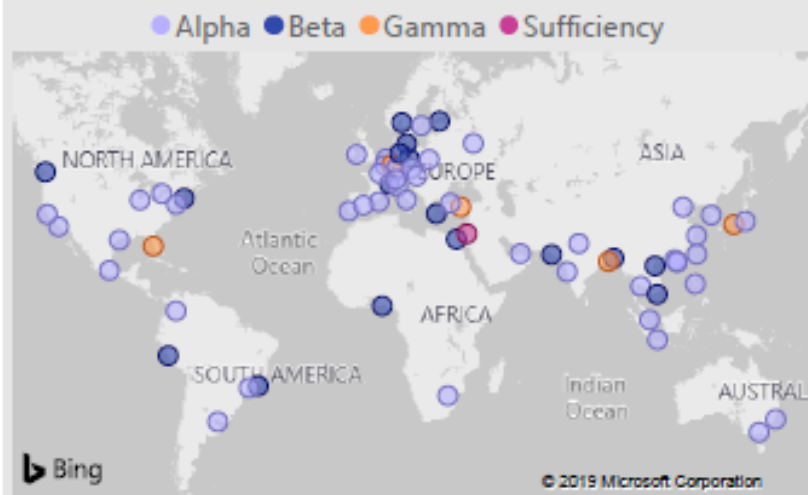
Wins VS Losses



Save Percentage



Global Cities Overview



68

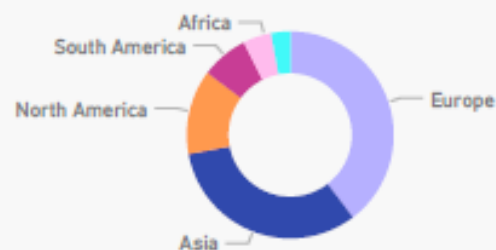
Cities

44

Countries

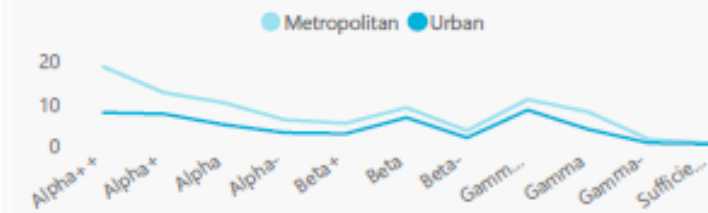
City	Country	Rating
Bangkok	Thailand	Alpha
Brussels	Belgium	Alpha
Buenos Aires	Argentina	Alpha
Chicago	USA	Alpha
Frankfurt	Germany	Alpha
Guangzhou	China	Alpha
Istanbul	Turkey	Alpha
Jakarta	Indonesia	Alpha
Los Angeles	USA	Alpha
Madrid	Spain	Alpha
Melbourne	Australia	Alpha
Mexico City	Mexico	Alpha
Milan	Italy	Alpha
Moscow	Russia	Alpha

Summary

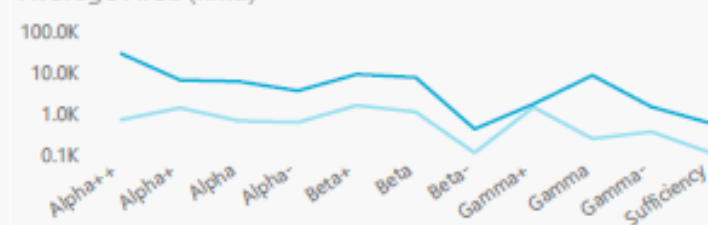


Metro vs Urban

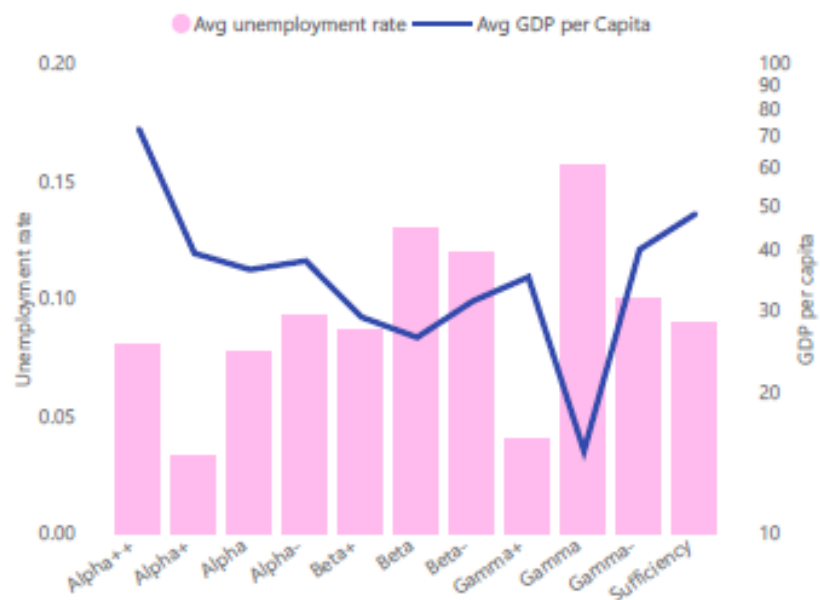
Average Population (millions)



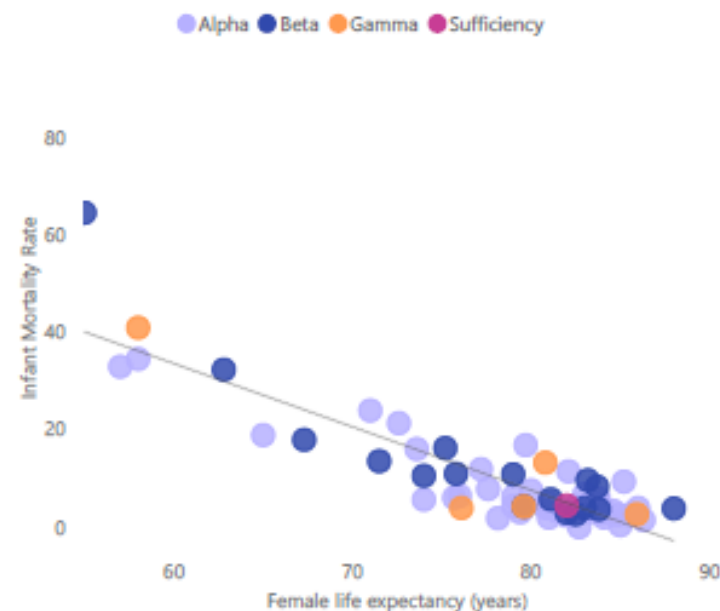
Average Area (km2)

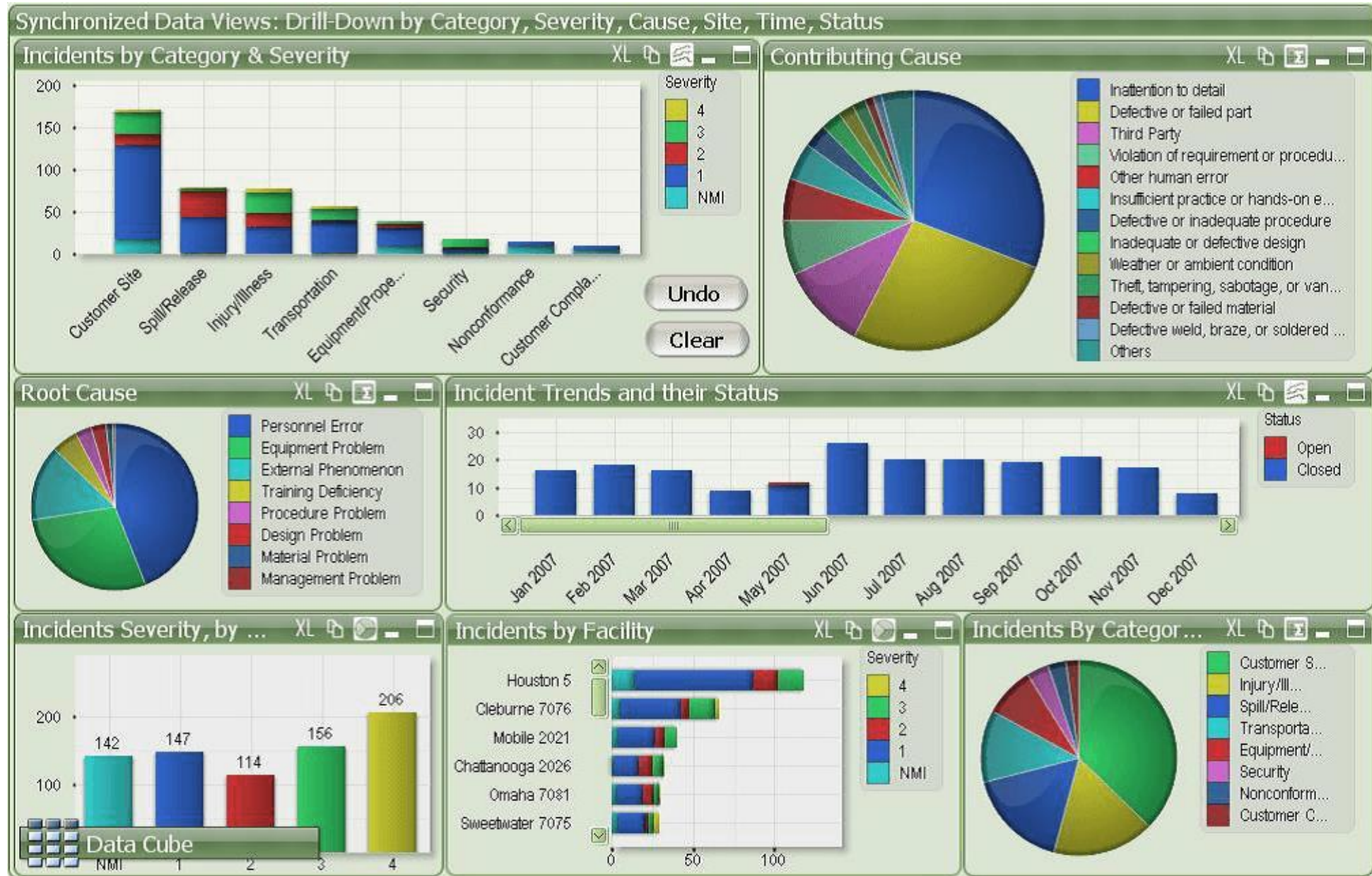


Average unemployment rate vs GDP per capita



Female life expectancy vs Infant Mortality



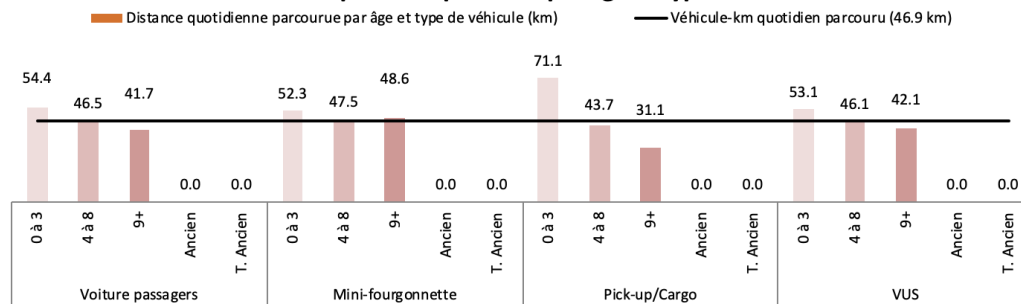




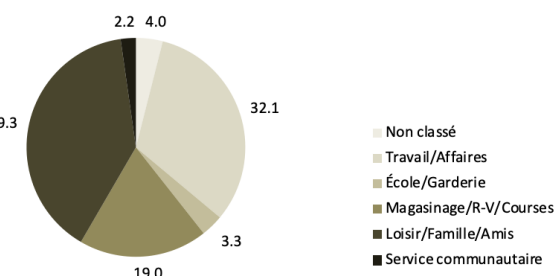
Ontario – 1er trimestre 2012

Caractéristiques des déplacements

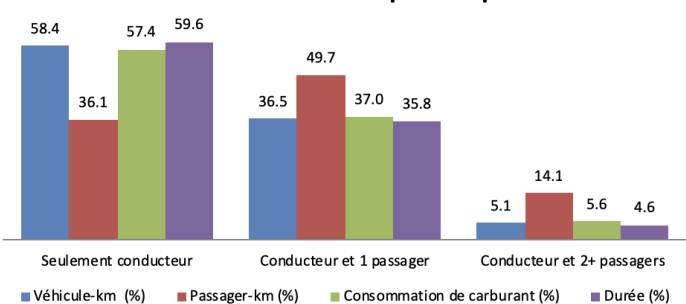
Véhicule-km quotidien parcouru par âge et type de véhicule



Passager-km quotidien parcouru par but des déplacements (%)



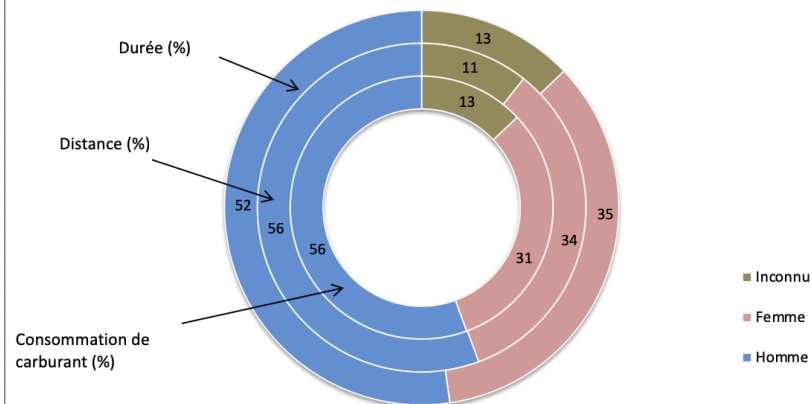
Distance, passager-km parcouru, durée et consommation de carburant par occupation



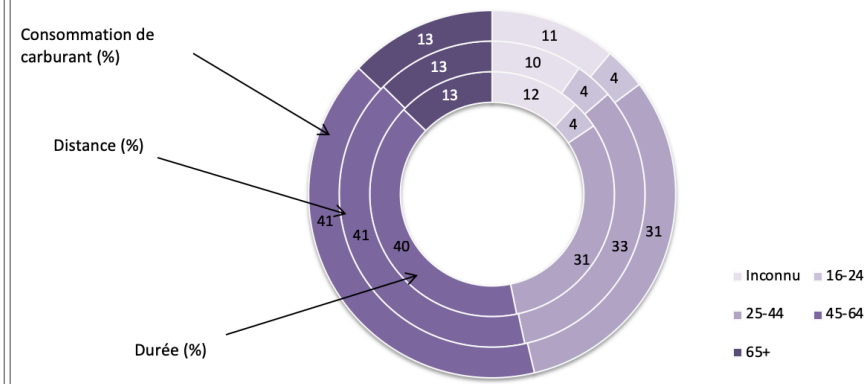
Proportion de déplacements par segments de distance



Durée, distance et consommation de carburant par sexe

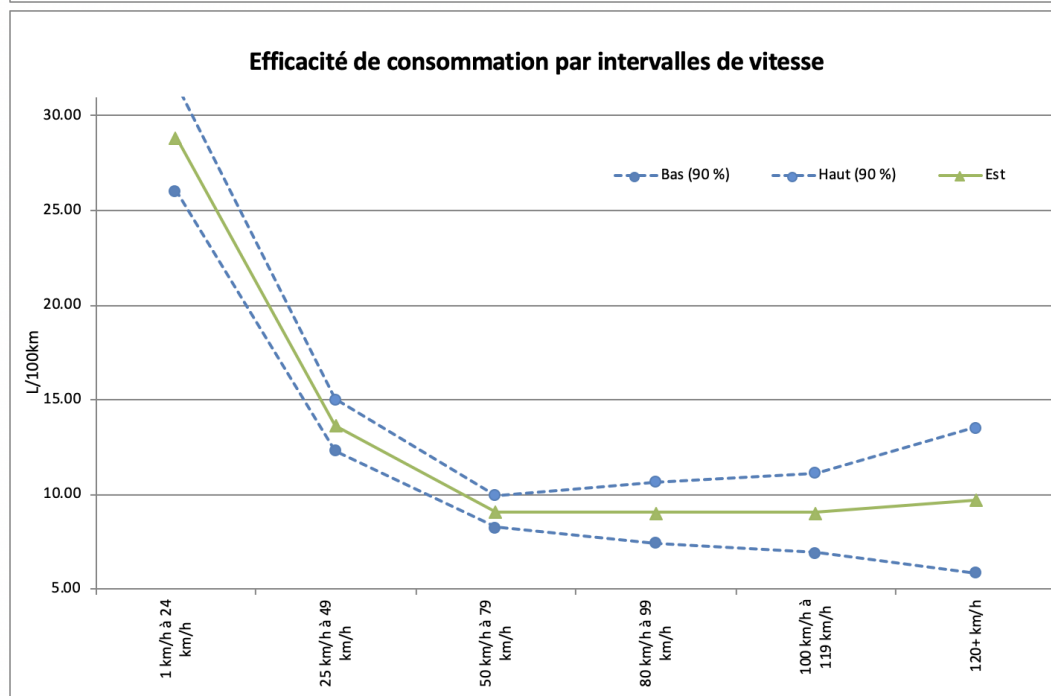
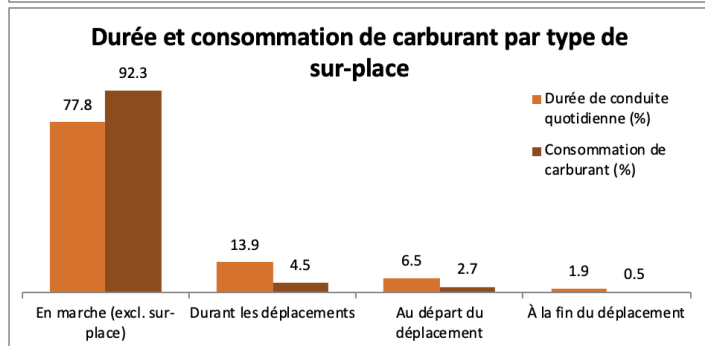
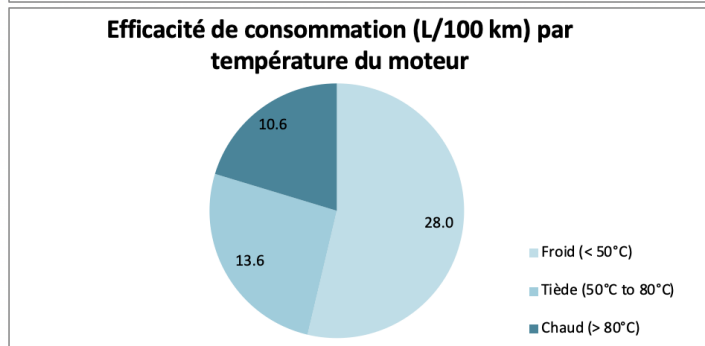
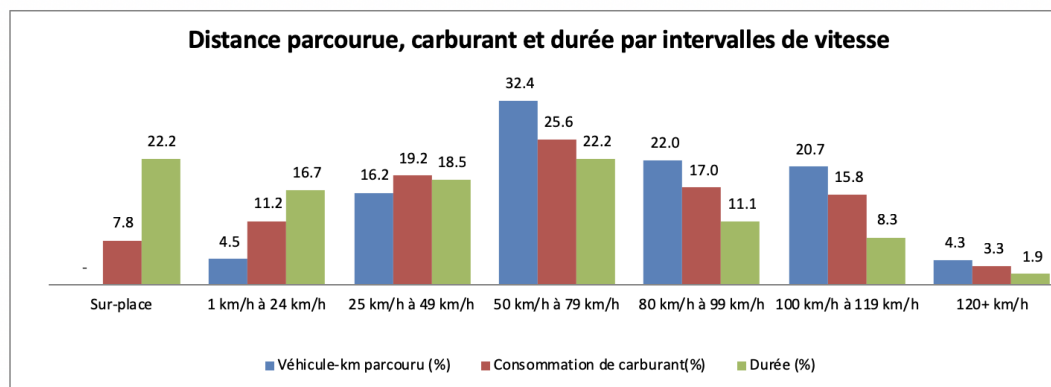
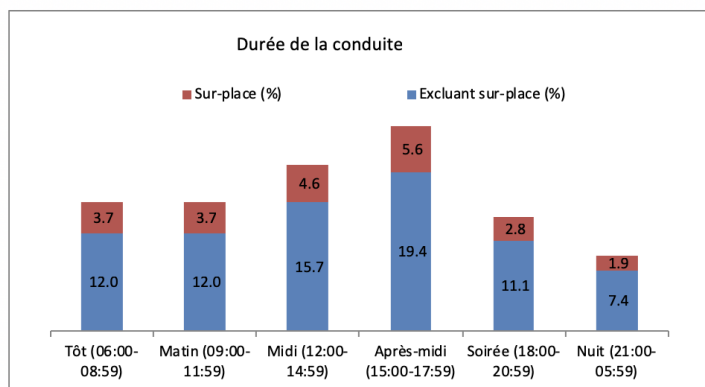


Consommation de carburant, distance et durée par âge des conducteurs



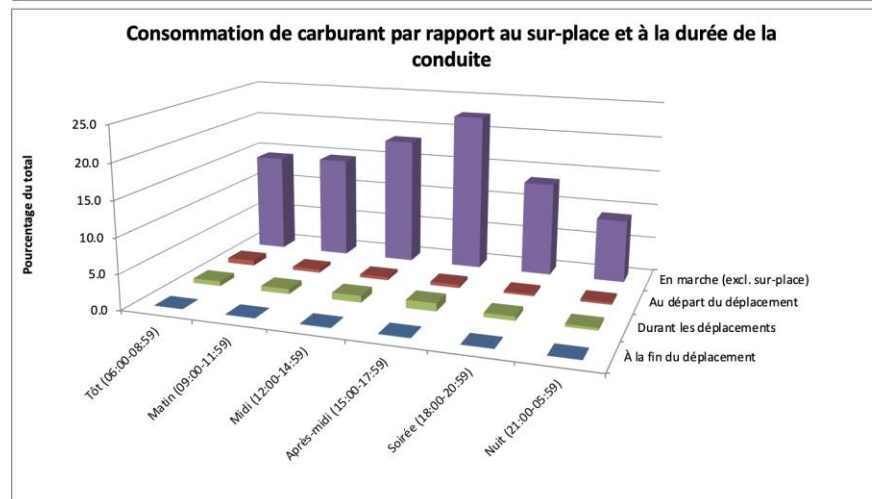
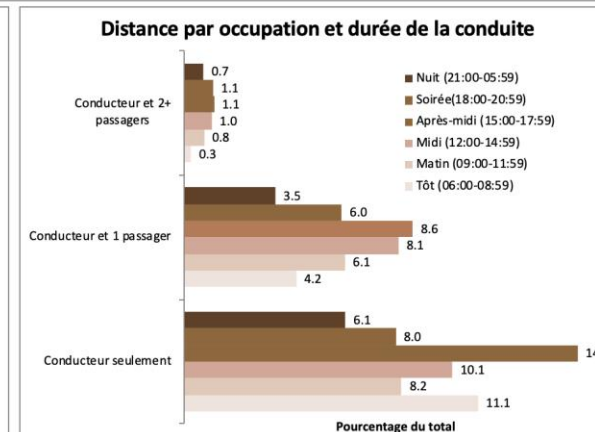
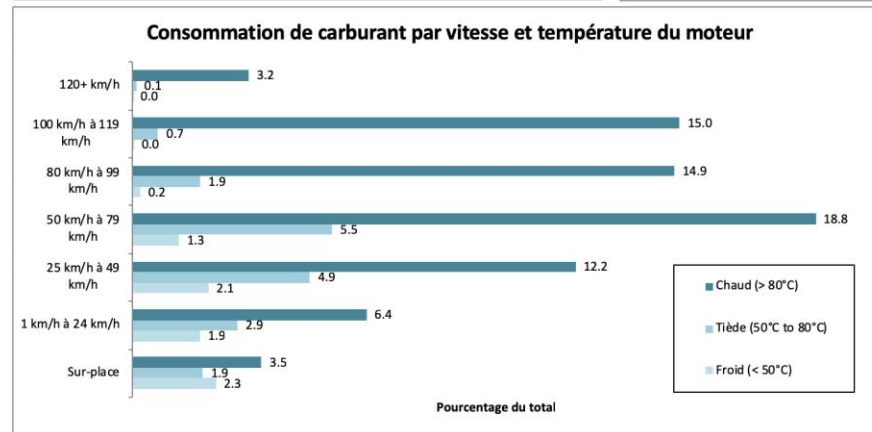
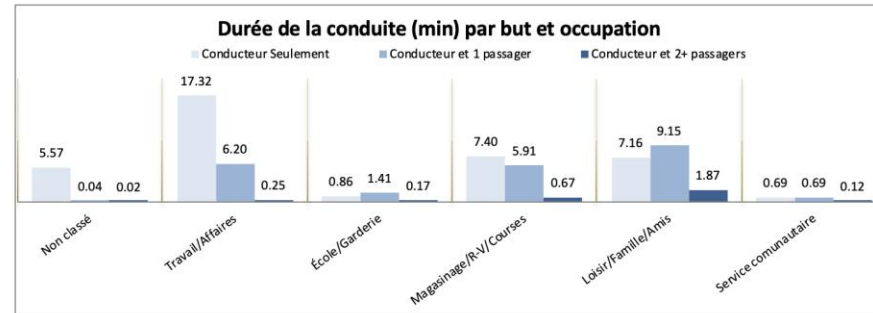
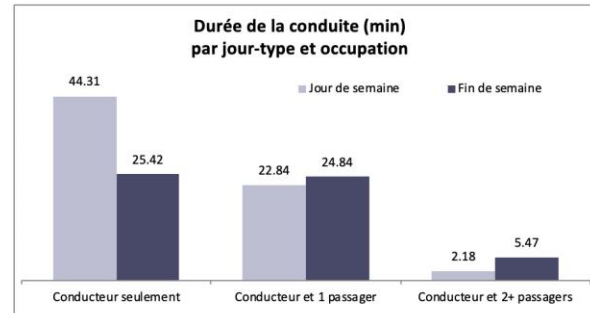
Ontario – 1er trimestre 2012

Sous-caractéristiques des déplacements



Ontario – 1er trimestre 2012

Caractéristiques mixtes sur les déplacements



What-If Analysis: Impact of Minimum Wage

[<https://bigbookofdashboards.com/dashboards.html>]



Proposed Minimum Wage

Required Service

Developed by Matt Chambers
<http://sirvialot.blogspot.com/>

