



Decisions in Visualization Design

Alberto Cairo




Generative art by Nadieh Bremer

U MIAMI SCHOOL of COMMUNICATION

ABOUT ADMISSIONS PROGRAMS STUDENTS OUR WORK FRIENDS Q

Alberto Cairo



Alberto Cairo is a journalist and designer with many years of experience leading graphics and visualization teams in several countries. He joined the School of Communication in January 2012. He teaches courses on infographics and data visualization. He is also director of the [Center for Visualization, Data Communication & Information Design](#) at UM's [Institute for Data Science and Computing](#), and a Faculty Fellow at the [Abess Center for Ecosystem Science and Policy](#).

Cairo has been described by *Microsoft* as always "in the vanguard of visual journalism". He is author of the books *How Charts Lie: Getting Smarter About Visual Information* (W.W. Norton, 2019), *The Truthful Art: Data, Charts, and Maps for Communication* (Peachpit Press, 2016), and *The Functional Art: an introduction to Information Graphics and Visualization* (Peachpit Press 2012). His next book, which deals with ethics and moral reasoning in visualization design, will be published by Wiley in 2021.

Cairo has also written for *The New York Times* and *Scientific American* magazine.

Cairo has a popular visualization weblog, <http://www.thefunctionalart.com/>, and his Twitter account, <https://twitter.com/AlbertoCairo>, is followed by more than 50,000 people interested in visualization and data journalism.

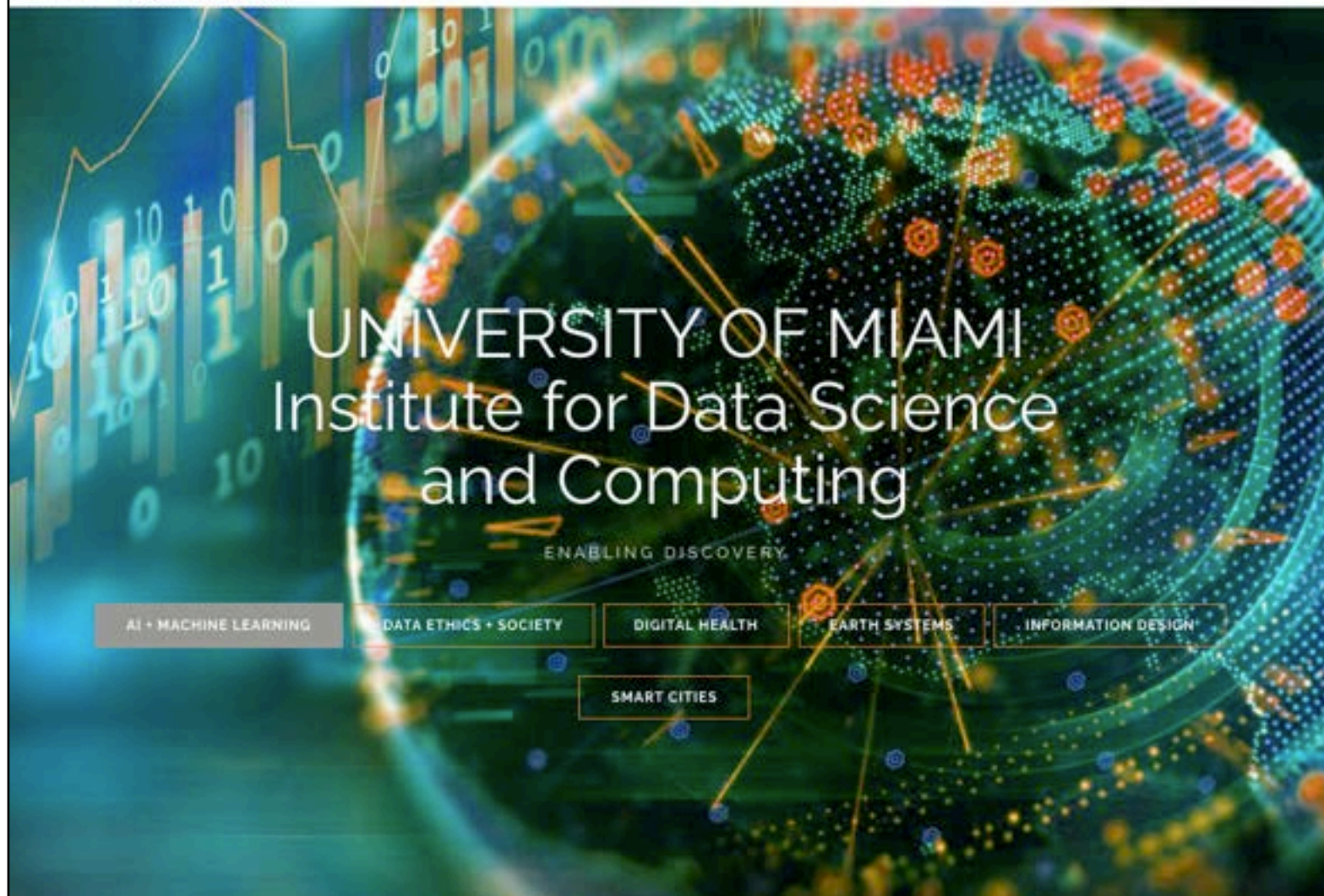
Associate Professor, Knight Chair in Visual Journalism

Office
CIB 5051-L

Phone

U MIAMI INSTITUTE for DATA SCIENCE & COMPUTING

HOME - RESEARCH - EDUCATION+OUTREACH - INNOVATION TECHNOLOGY PLATFORMS - ABOUT - EVENTS - QUICKLINKS Q



UNIVERSITY OF MIAMI Institute for Data Science and Computing

ENABLING DISCOVERY

AI • MACHINE LEARNING DATA ETHICS • SOCIETY DIGITAL HEALTH EARTH SYSTEMS INFORMATION DESIGN SMART CITIES

Open Visualization Academy by alberto cairo

Home About News Articles f i t



Open Visualization Academy

by alberto cairo

Coming in 2025

The Open Visualization Academy is an open source repository of knowledge about information design and data visualization.

Stay in the loop!

<https://com.miami.edu/profile/alberto-cairo/>

<https://idsc.miami.edu/>

<https://openvisualizationacademy.org/>

Consulting, freelancing, art direction

Google Trends interface showing search interest for "Aaron Rodgers" in the United States. The main heading reads "Explore what United States is searching for right now". A search bar contains "Aaron Rodgers" and an "Explore" button. Below, a line graph shows search interest over time. A section titled "Why is Aaron Rodgers trending?" lists three recent news items: "Aaron Rodgers to Derwin James after Chargers-Jets: 'Give me a..." (7 hours ago, NBC Sports - NFL), "Aaron Rodgers hints at Jets return after loss to Chargers" (7 hours ago, Fox News), and "Aaron Rodgers Dropped Another Hint About His Future After Jets..." (8 hours ago, Newsweek).

Microsoft AI for Earth website. The header includes "Microsoft AI" and navigation links for "Why AI", "Copilot products", "Solutions", "Responsible AI", and "Resources". The main content area features a large image of a blue butterfly on a green leaf. Text reads "AI for Good: Providing technology and resources to empower organizations working to solve global challenges to the environment, humanitarian issues, accessibility, health, and cultural heritage." Below is a "Play AI for Good video" button. At the bottom, the "AI for Earth" logo is displayed with the tagline "AI for Earth put AI technology and cloud software in the hands of those working to solve global environmental challenges." and a "Learn about AI for Earth" link.

McMaster-Carr website showing a grid of hardware categories. The "Fastening & Joining" section includes items like bolts, nuts, washers, and screws. Other sections include "Adhesives & Tape", "Welding, Brazing & Soldering", "Pipe, Tubing, Hose & Fittings", and "Power Transmission". Each category is represented by small icons of the respective hardware items.

The World Bank website featuring a large image of the Earth. The headline reads "A world free of poverty on a livable planet." Below the headline, text states: "A landmark speech by Ajay Banga on our ambitious new vision. One challenge cannot be solved without solving the others. They demand an expanded vision of the World Bank." Two buttons are visible: "WATCH SPEECH" and "READ SPEECH".

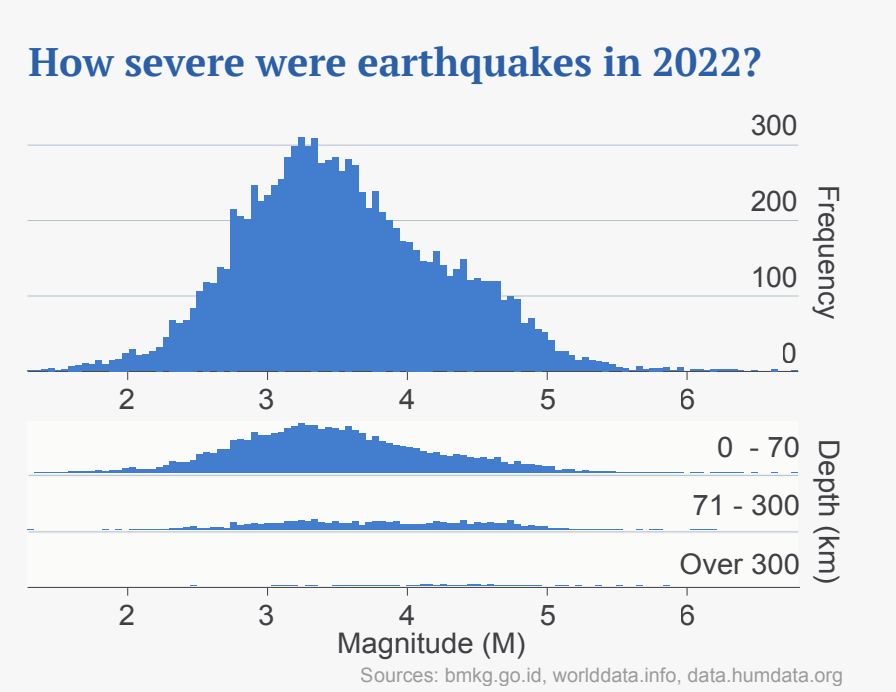
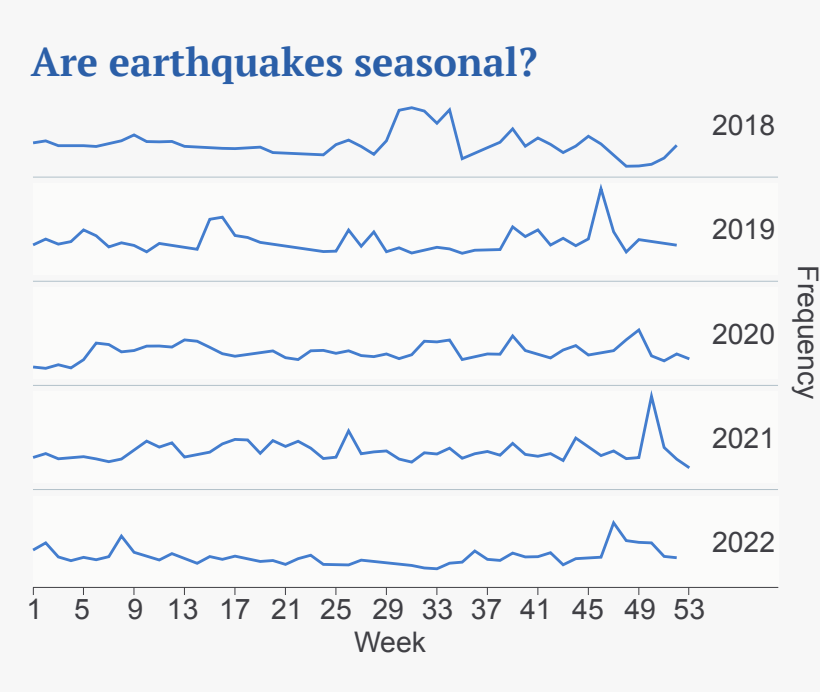
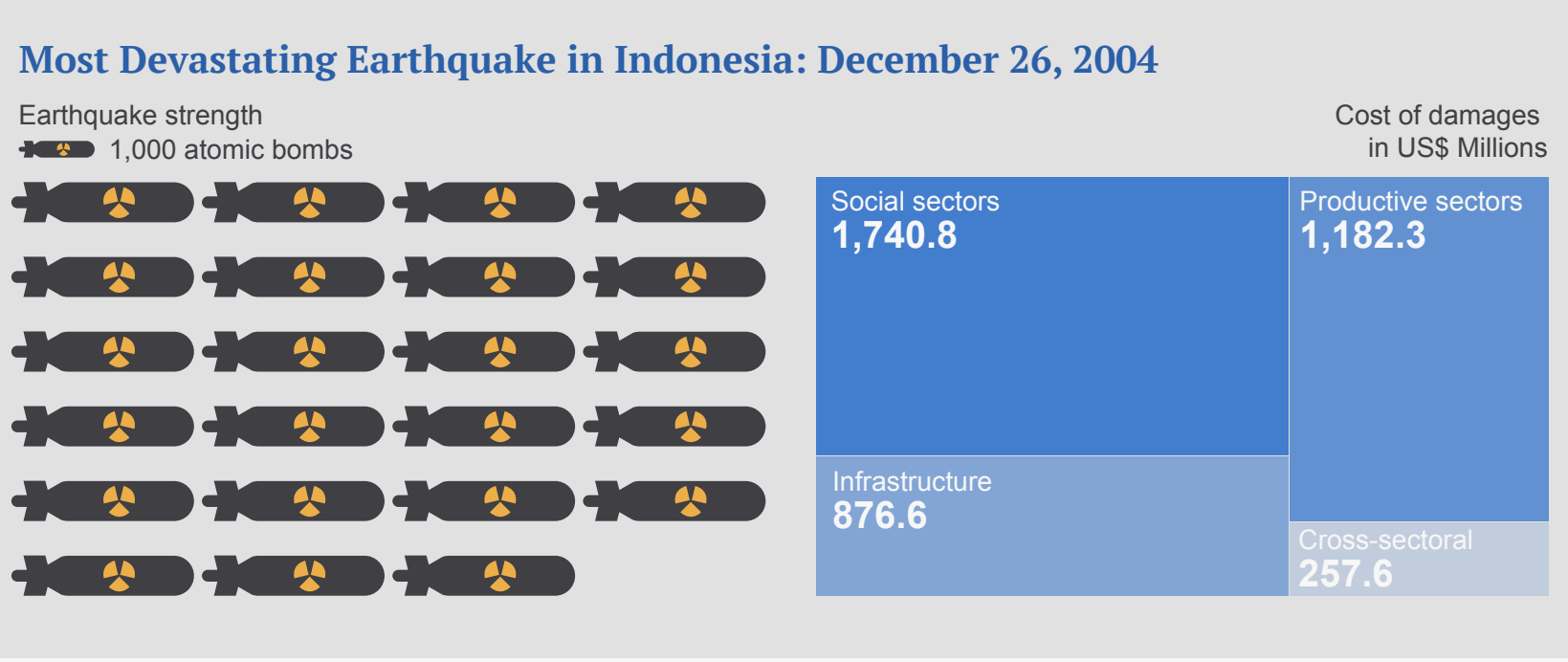
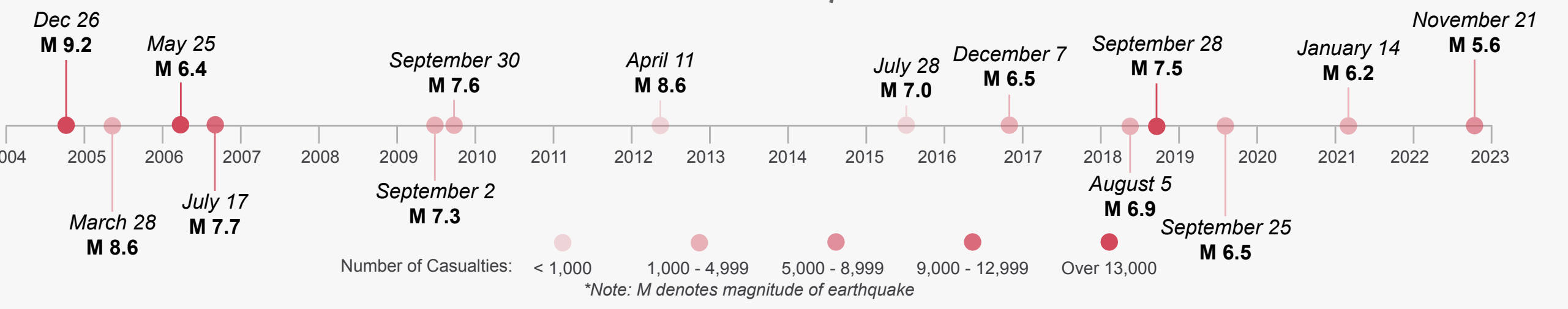
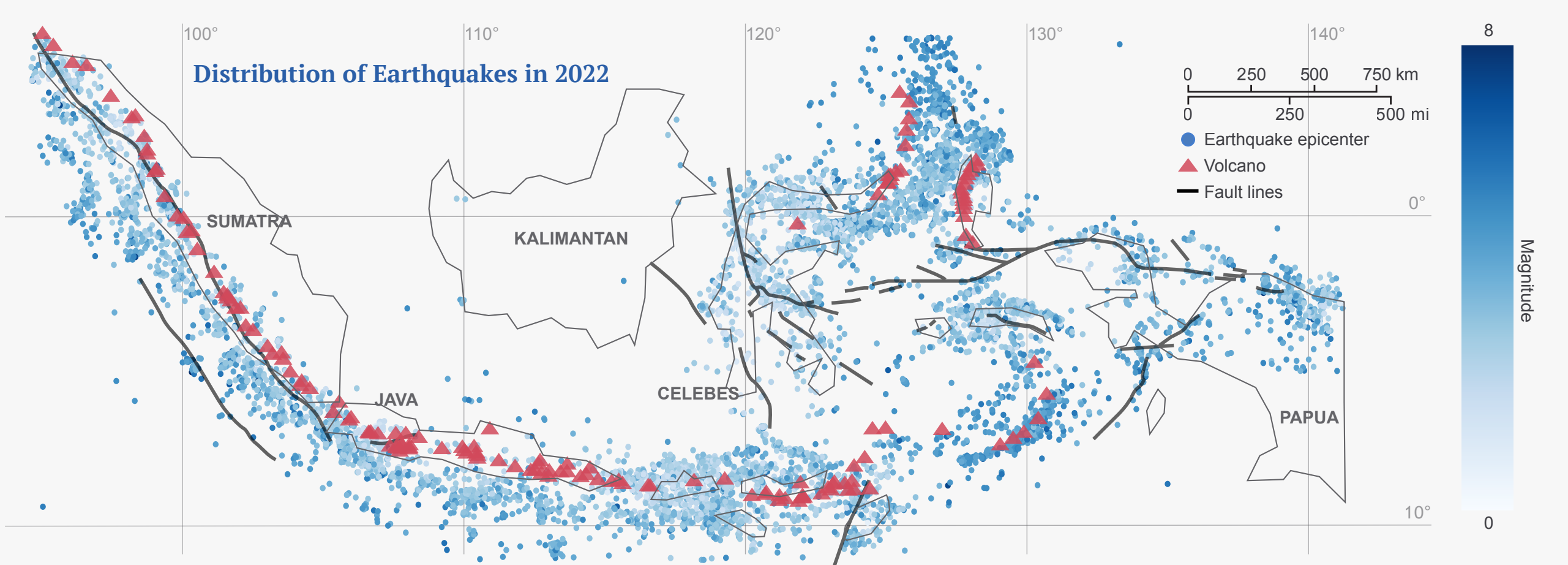
Congressional Budget Office website. The header includes "Congressional Budget Office" and navigation links for "About CBO", "Topics", "Cost Estimates", "Data", "Interactives", "FAQs", and "Blog". The main content area features a large image of a hand holding a glowing orb with data charts. Below the image, the text reads "Short-Term Spending and Long-Term Dynamic Effects" and "Federal policies related to health care, education, infrastructure, the environment, and research and development often involve short-term expenditures that give rise to economic and budgetary effects far in the future. CBO's conventional cost estimates project the budgetary effects of proposed legislation over a 10-year period and incorporate the assumption that nominal gross domestic product (GDP) remains unchanged." To the right, a "FEATURED" section lists several articles: "Long-Term Implications of the 2024 Future Years Defense Program", "Estimated Budgetary Effects of the Israel Security Supplemental Appropriations Act, 2024", "Monthly Budget Review: September 2023", "Federal Subsidies for Health Insurance: 2023 to 2033", "Federal Budgetary Effects of the Activities of the Center for Medicare & Medicaid Innovation", and "Communities at Risk of Flooding".

A core idea in my classes is that **anyone** can learn to design well

NAVIGATING INDONESIA'S EARTHQUAKES

Indonesia is positioned on active fault lines within the Ring of Fire, a semi-circle of volcanoes and shifting plates along the Pacific Ocean's edge. There have been numerous earthquakes, volcanic activities, and tsunamis in the region due to ongoing geological pressures. In 2022, The Meteorology, Climatology, and Geophysics Agency (BMKG.go.id) recorded 10,792 earthquakes. The severity of these earthquakes varies, with the most intense occurring in 2004, reaching a magnitude of 9.2.

What impact do earthquakes have?



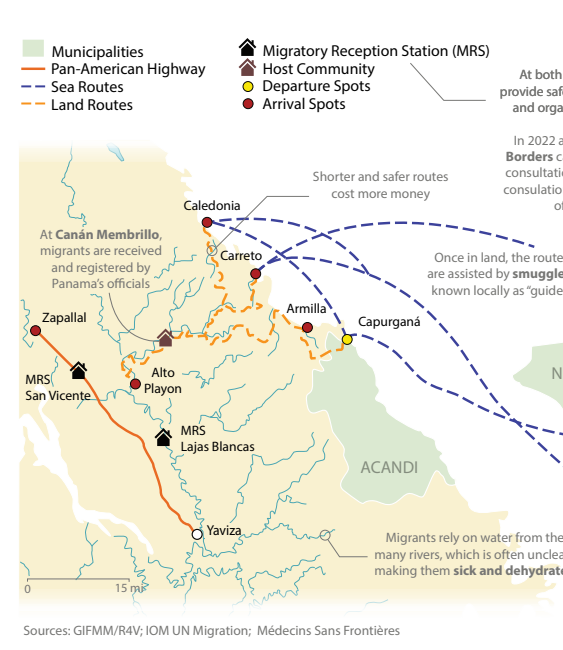
Poster by Michela Effendi

Defying death

This year, over 400,000 migrants and asylum seekers have traversed the challenging expanse of the Central American jungle known as the Darién Gap, as reported by the government of Panama. Despite the formidable dangers and obstacles within the jungle, it stands as the sole terrestrial route linking South America to Central America. Those undertaking this journey face numerous human rights violations, including sexual violence, homicides, disappearances, trafficking, robbery, and intimidation perpetrated by organized criminal groups.



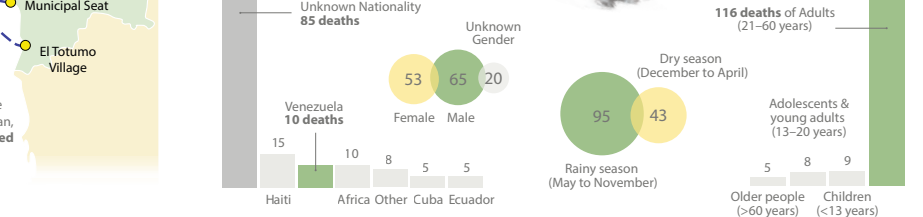
Major routes taken to cross the Darién Gap



Transit throughout the years



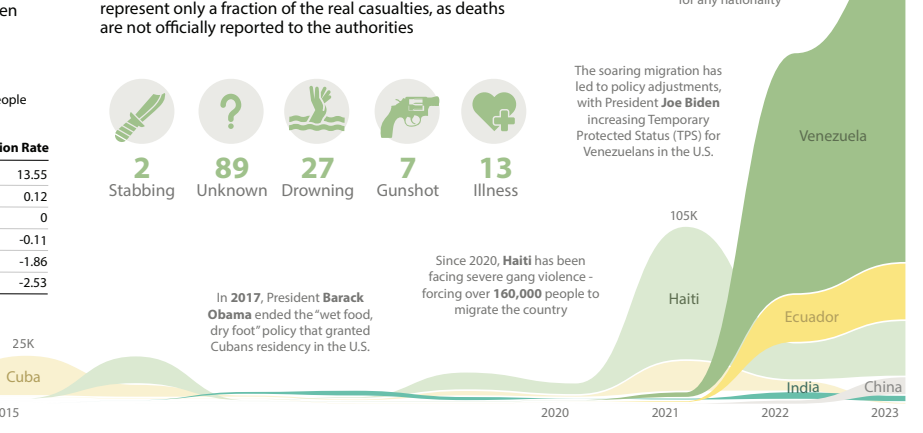
Reported deaths



Venezuela's migration rates lead the Darién



Main death causes



Poster by Daniela González

A puzzling question;

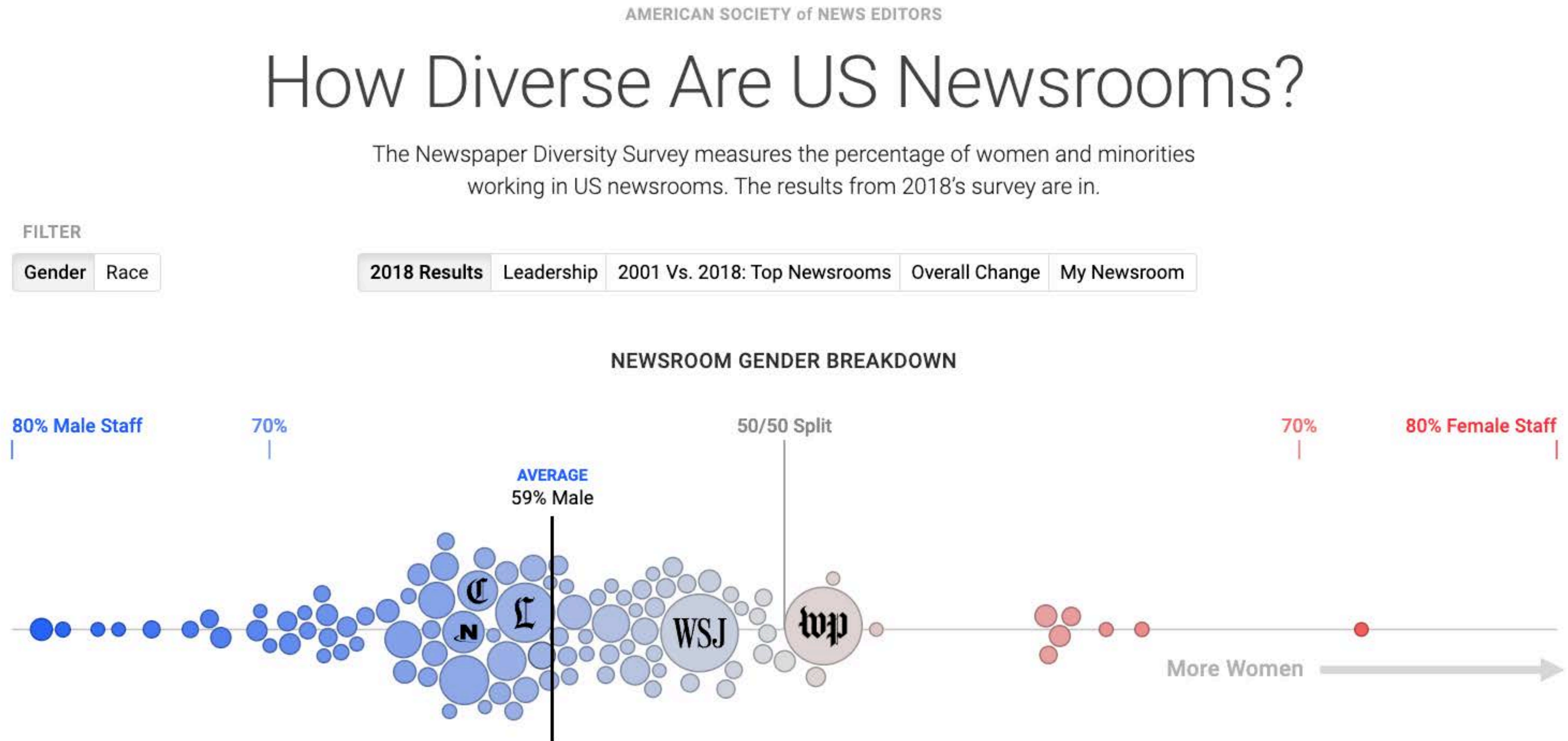
“What are the **rules** of data visualization design”?

Like **writing**, visualization design is **more a craft than a science**.

(Although some guidance and heuristics can be derived from scientific research on different areas, such as human-computing interaction, visualization, accessibility, perception, or cognitive science).

A good exercise: To reverse-engineer existing visualizations

“If I were the designer who created this, what choices would lead me to this solution?”



<https://googletrends.github.io/asne/>

How Diverse Are US Newsrooms?

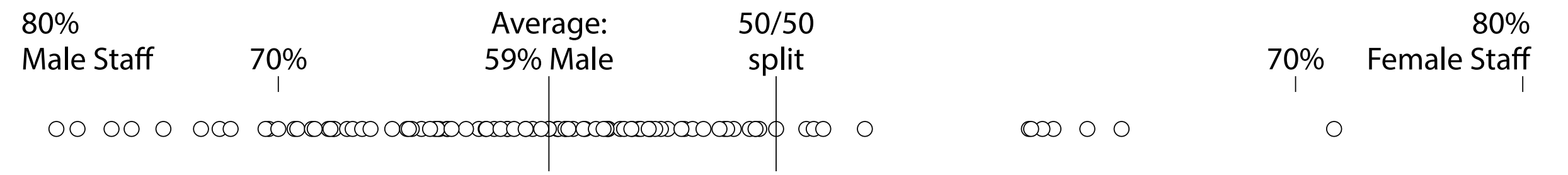
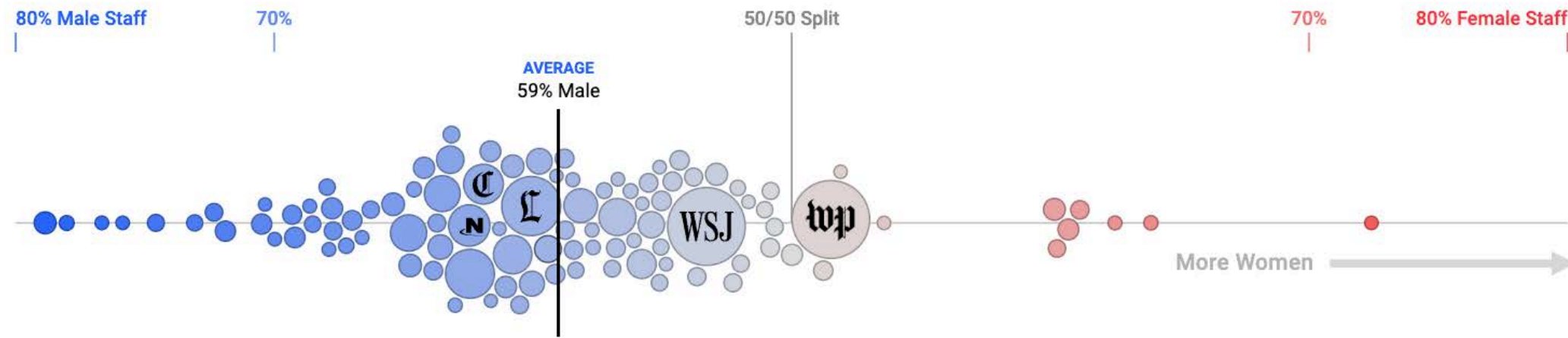
The Newspaper Diversity Survey measures the percentage of women and minorities working in US newsrooms. The results from 2018's survey are in.

FILTER

Gender Race

2018 Results Leadership 2001 Vs. 2018: Top Newsrooms Overall Change My Newsroom

NEWSROOM GENDER BREAKDOWN



Designing information graphics and data visualizations doesn't consist of applying **rules**, but of **reasoning** about choices, and **justifying** them.

Every choice in design is **subjective**, and therefore debatable, but it should never be **arbitrary**.

How Diverse Are US Newsrooms?

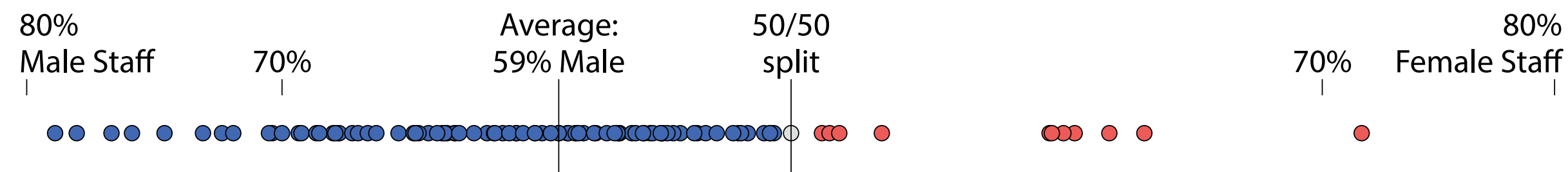
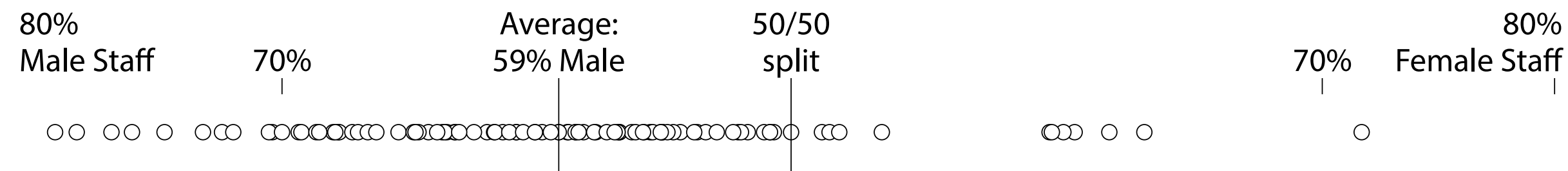
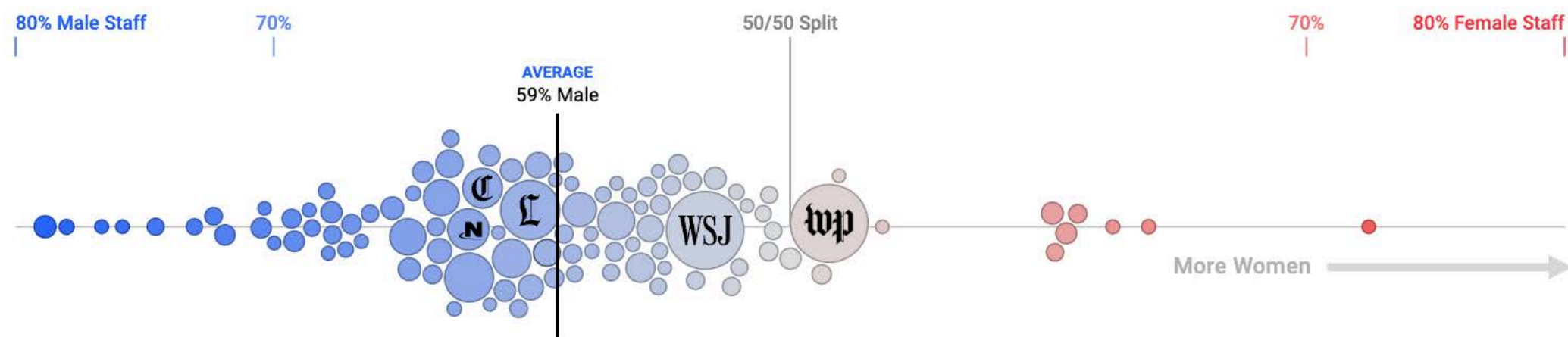
The Newspaper Diversity Survey measures the percentage of women and minorities working in US newsrooms. The results from 2018's survey are in.

FILTER

Gender Race

2018 Results Leadership 2001 Vs. 2018: Top Newsrooms Overall Change My Newsroom

NEWSROOM GENDER BREAKDOWN



Designing information graphics and data visualizations doesn't consist of applying **rules**, but of **reasoning** about choices, and **justifying** them.

Every choice in design is **subjective**, and therefore debatable, but it should never be **arbitrary**.

How Diverse Are US Newsrooms?

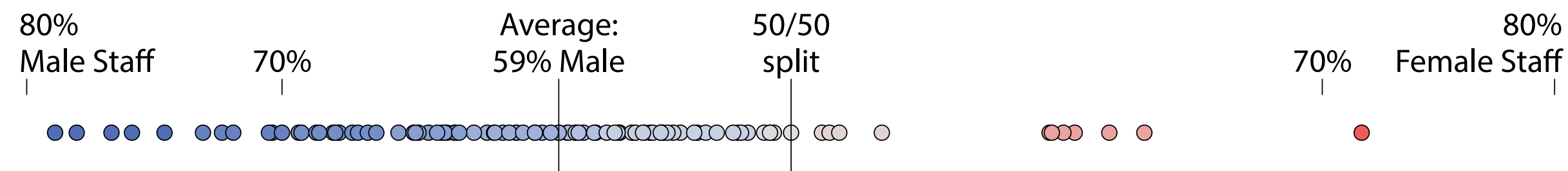
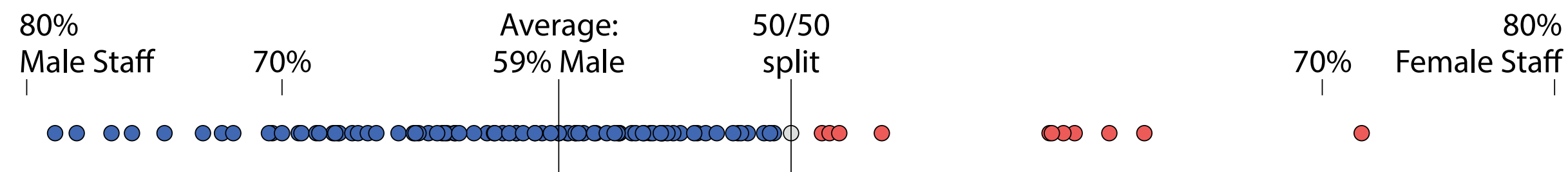
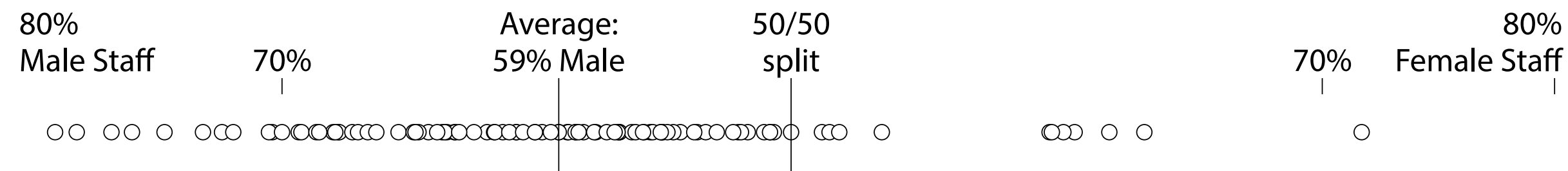
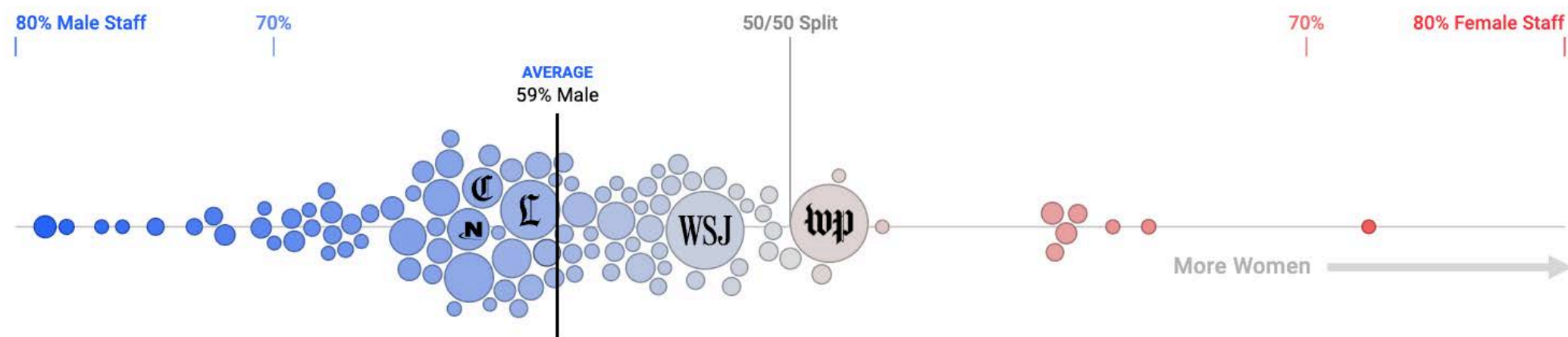
The Newspaper Diversity Survey measures the percentage of women and minorities working in US newsrooms. The results from 2018's survey are in.

FILTER

Gender Race

2018 Results Leadership 2001 Vs. 2018: Top Newsrooms Overall Change My Newsroom

NEWSROOM GENDER BREAKDOWN



Designing information graphics and data visualizations doesn't consist of applying **rules**, but of **reasoning** about choices, and **justifying** them.

Every choice in design is **subjective**, and therefore debatable, but it should never be **arbitrary**.

How Diverse Are US Newsrooms?

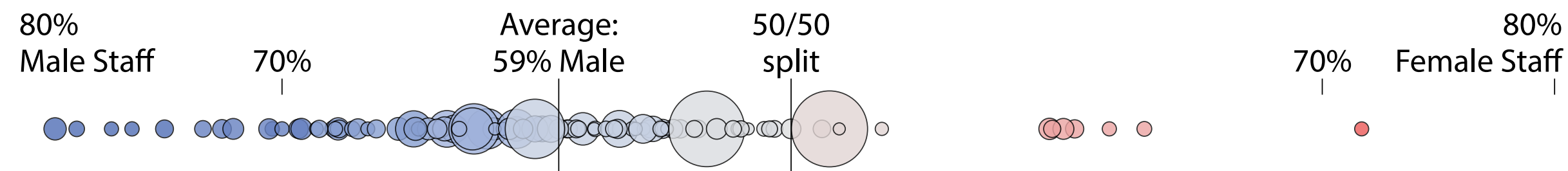
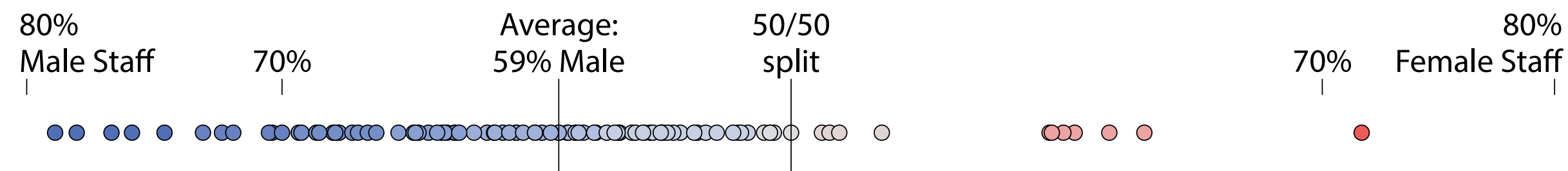
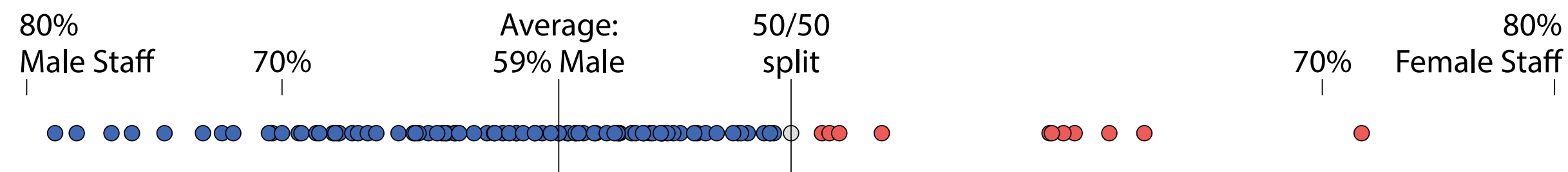
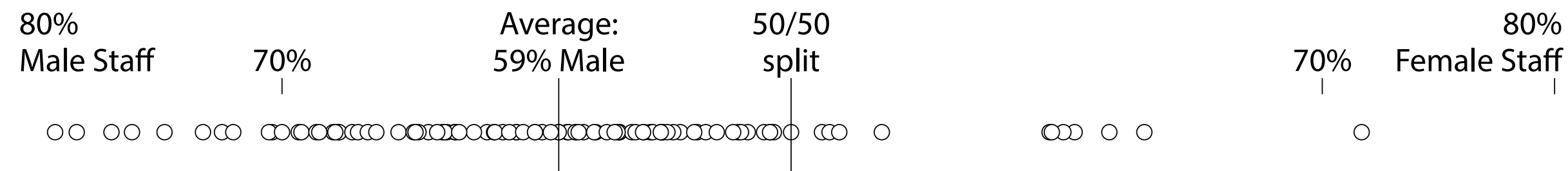
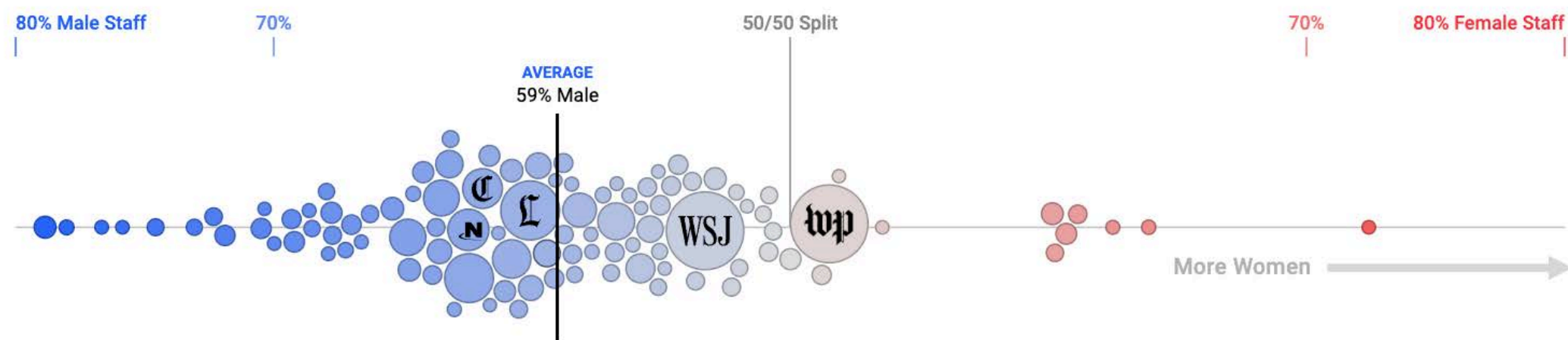
The Newspaper Diversity Survey measures the percentage of women and minorities working in US newsrooms. The results from 2018's survey are in.

FILTER

Gender Race

2018 Results Leadership 2001 Vs. 2018: Top Newsrooms Overall Change My Newsroom

NEWSROOM GENDER BREAKDOWN



Designing information graphics and data visualizations doesn't consist of applying **rules**, but of **reasoning** about choices, and **justifying** them.

Every choice in design is **subjective**, and therefore debatable, but it should never be **arbitrary**.

How Diverse Are US Newsrooms?

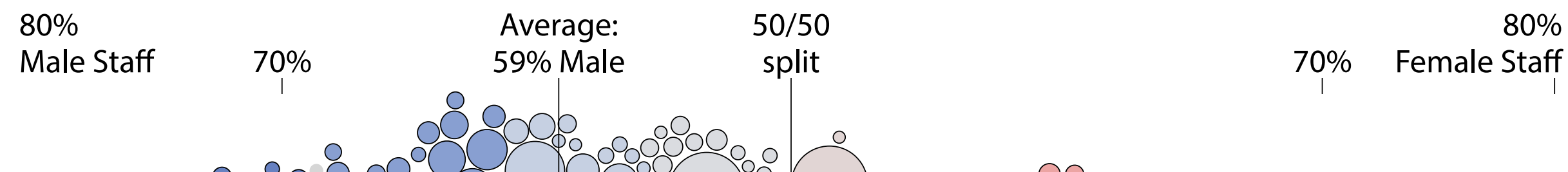
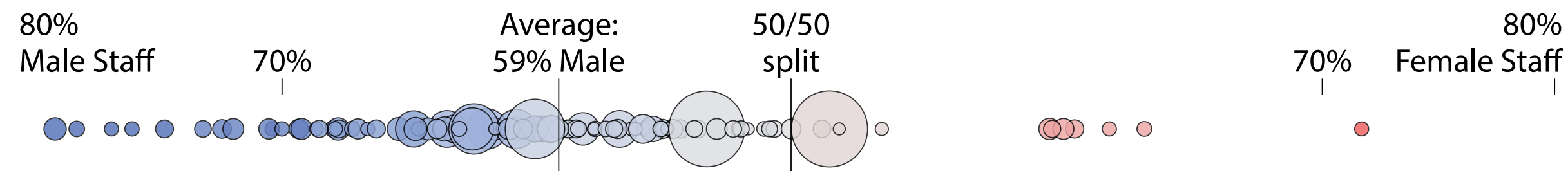
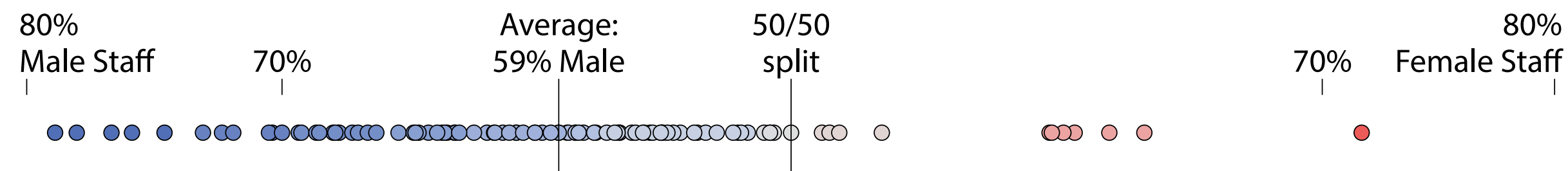
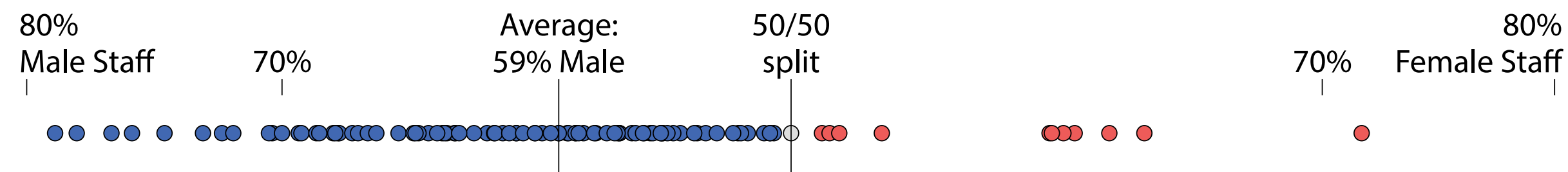
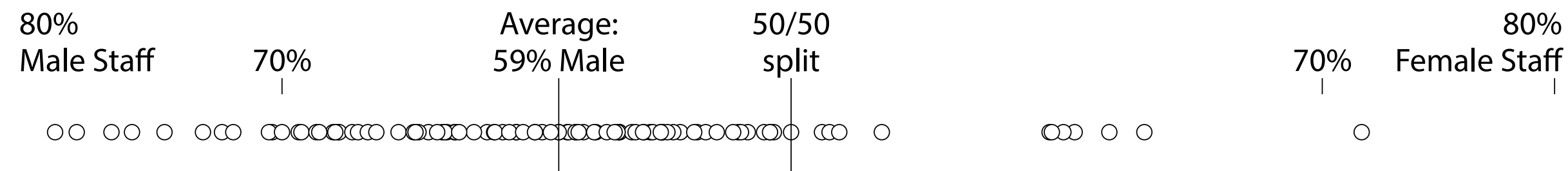
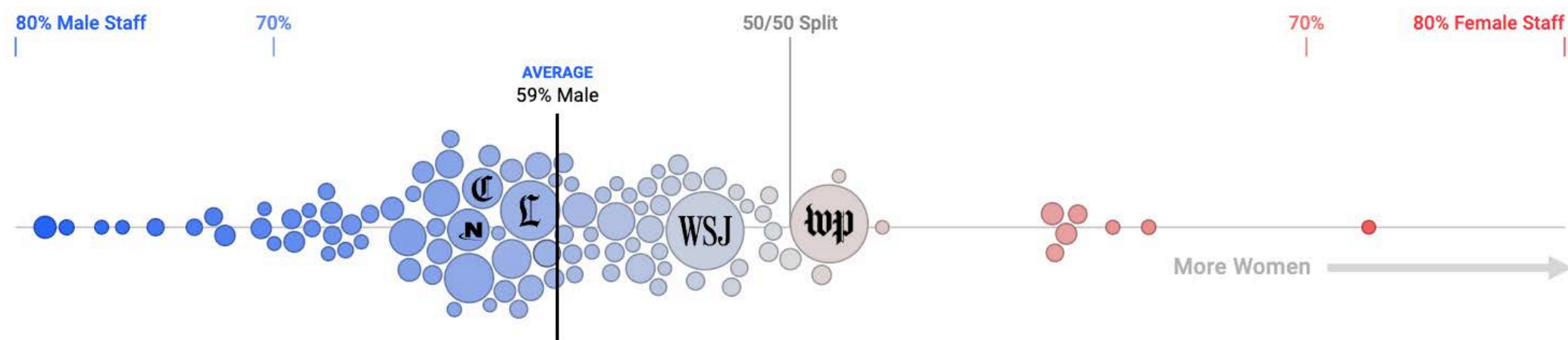
The Newspaper Diversity Survey measures the percentage of women and minorities working in US newsrooms. The results from 2018's survey are in.

FILTER

Gender Race

2018 Results Leadership 2001 Vs. 2018: Top Newsrooms Overall Change My Newsroom

NEWSROOM GENDER BREAKDOWN

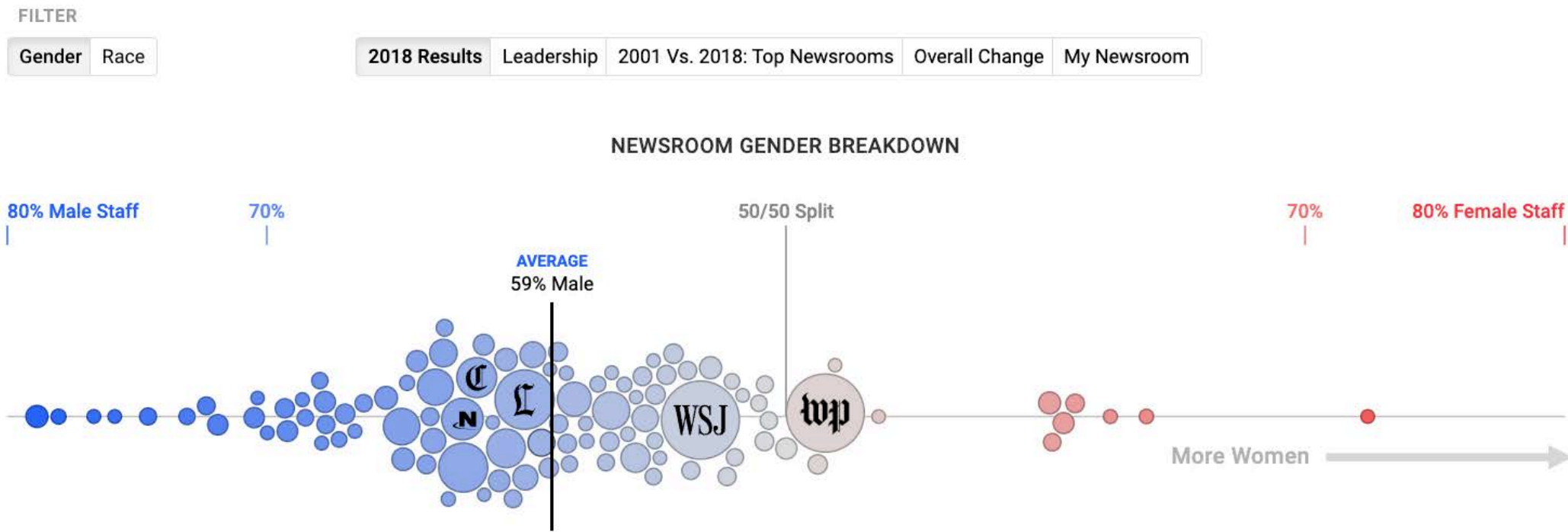


Designing information graphics and data visualizations doesn't consist of applying **rules**, but of **reasoning** about choices, and **justifying** them.

Every choice in design is **subjective**, and therefore debatable, but it should never be **arbitrary**.

How Diverse Are US Newsrooms?

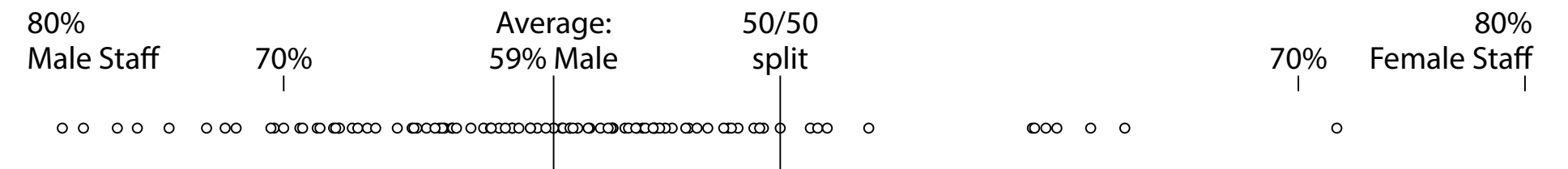
The Newspaper Diversity Survey measures the percentage of women and minorities working in US newsrooms. The results from 2018's survey are in.



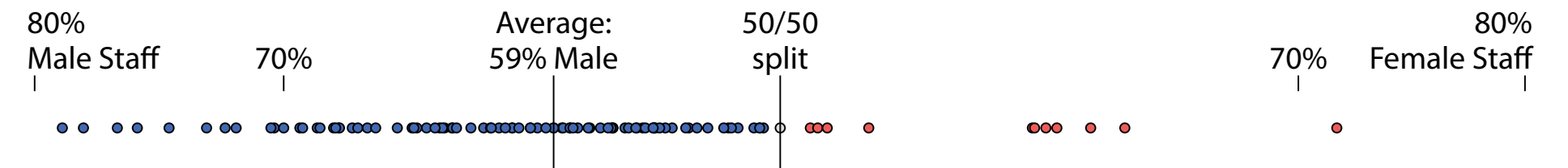
Designing information graphics and data visualizations doesn't consist of applying **rules**, but of **reasoning** about choices, and **justifying** them.

Every choice in design is **subjective**, and therefore debatable, but it should never be **arbitrary**.

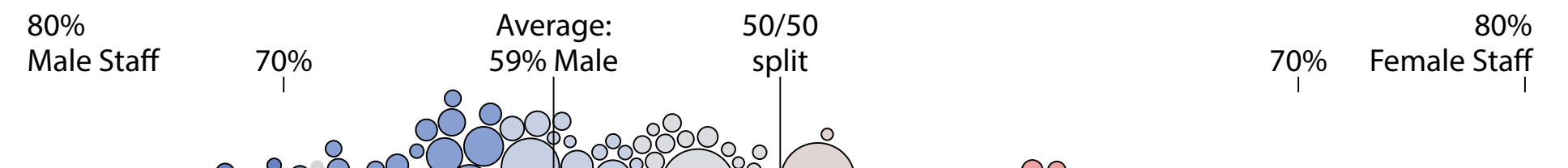
Decisions related to axis scales and encodings



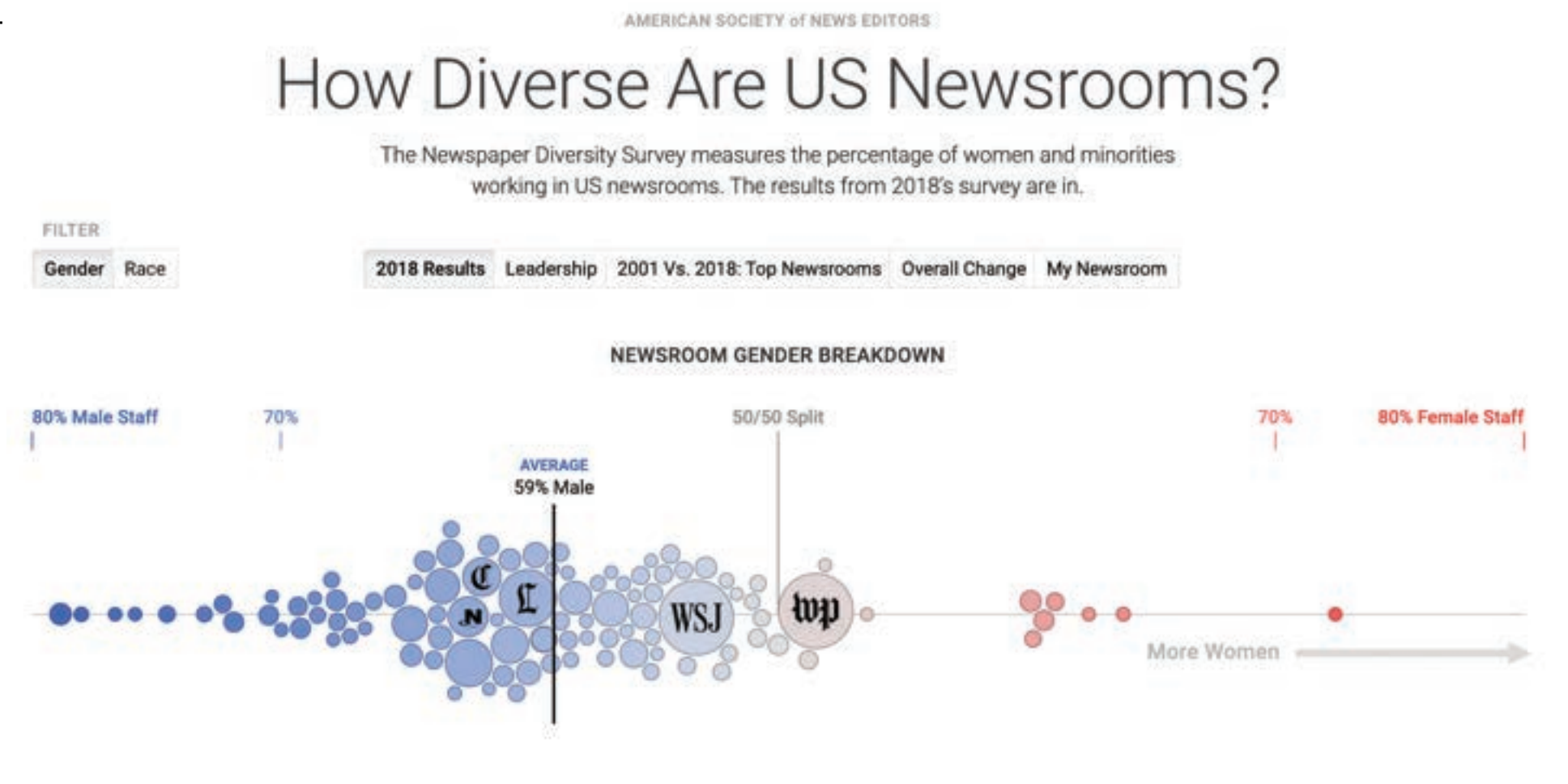
Decisions related to emphasis and increasing attractiveness



Decisions to improve legibility

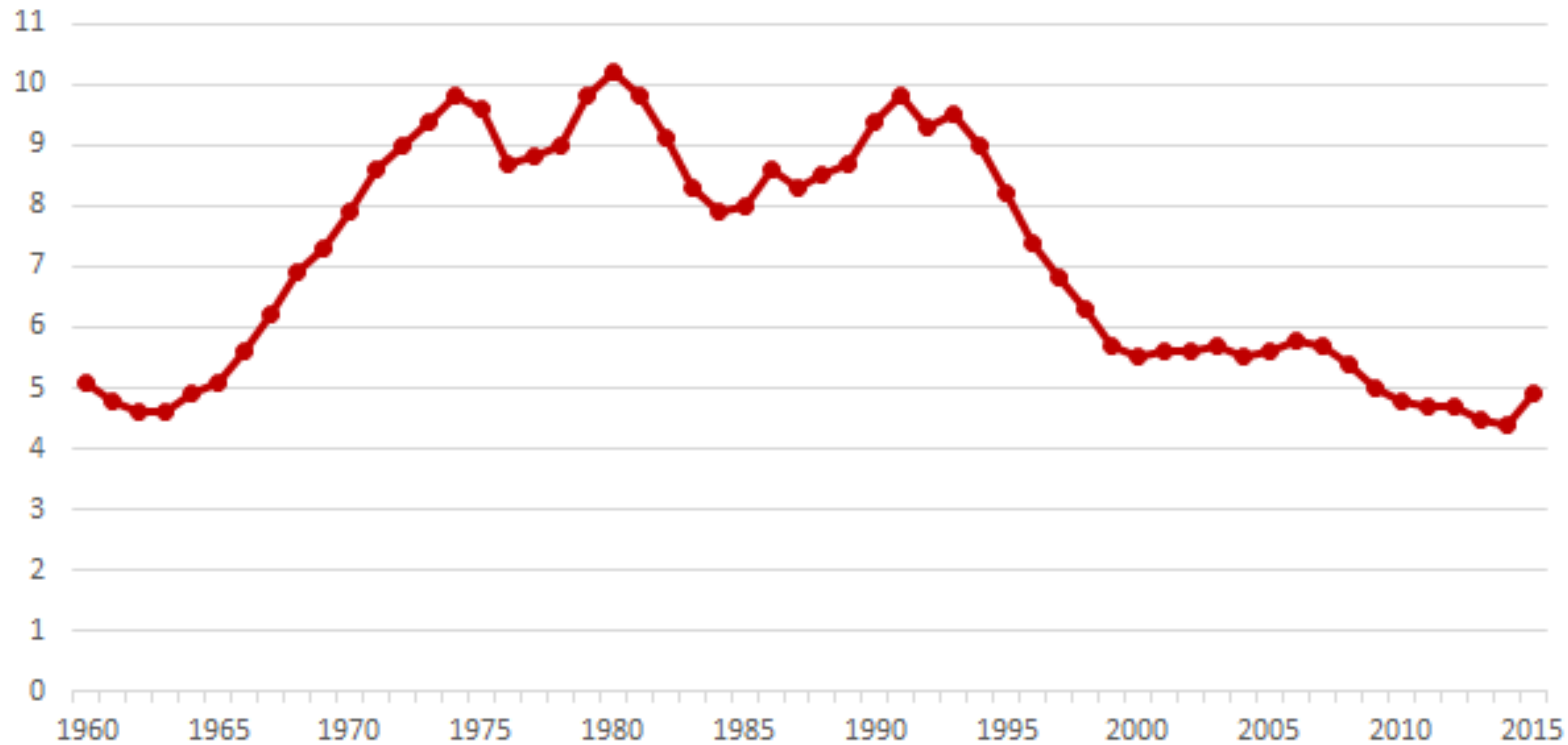


Decisions related to layout and annotations



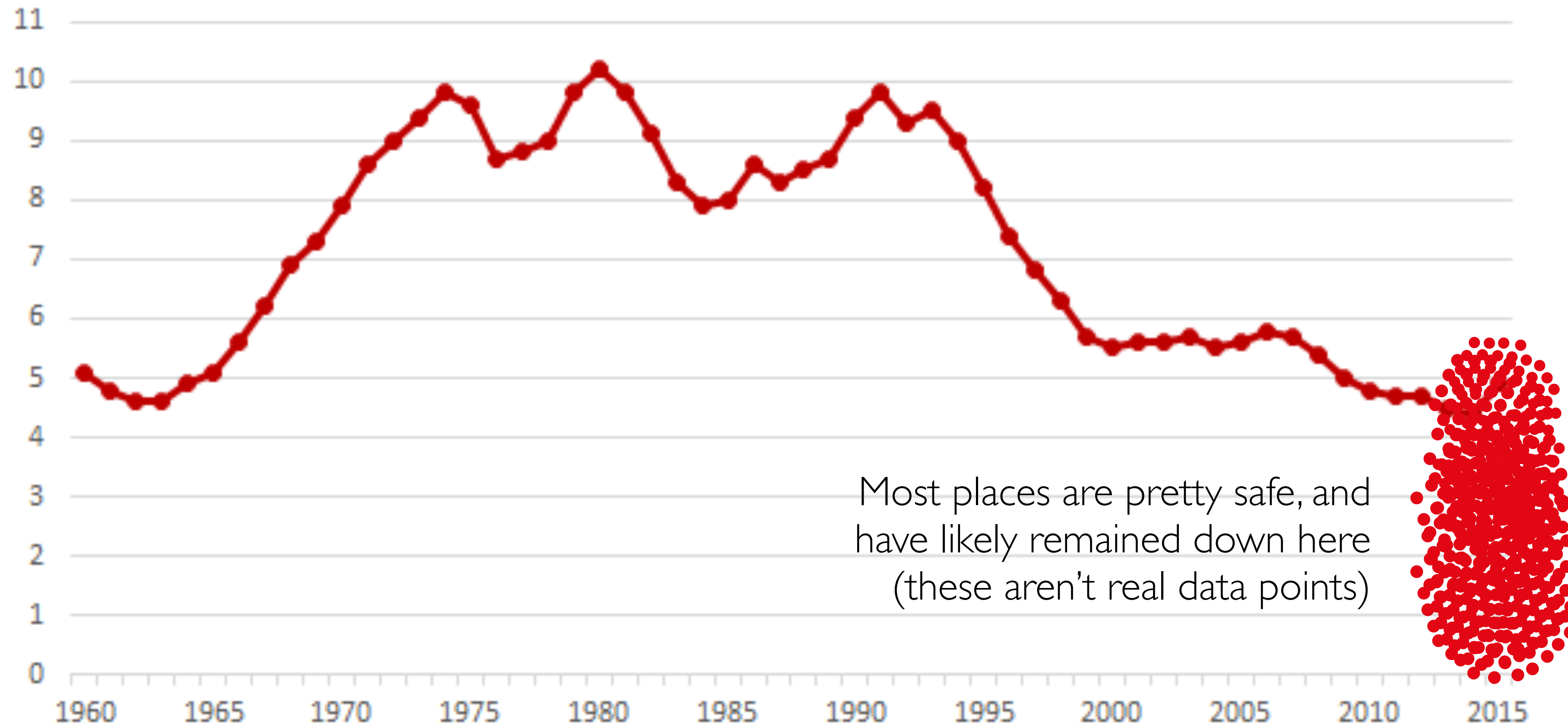
What to show and how much to show?

The U.S. Murder Rate (per 100,000 people)



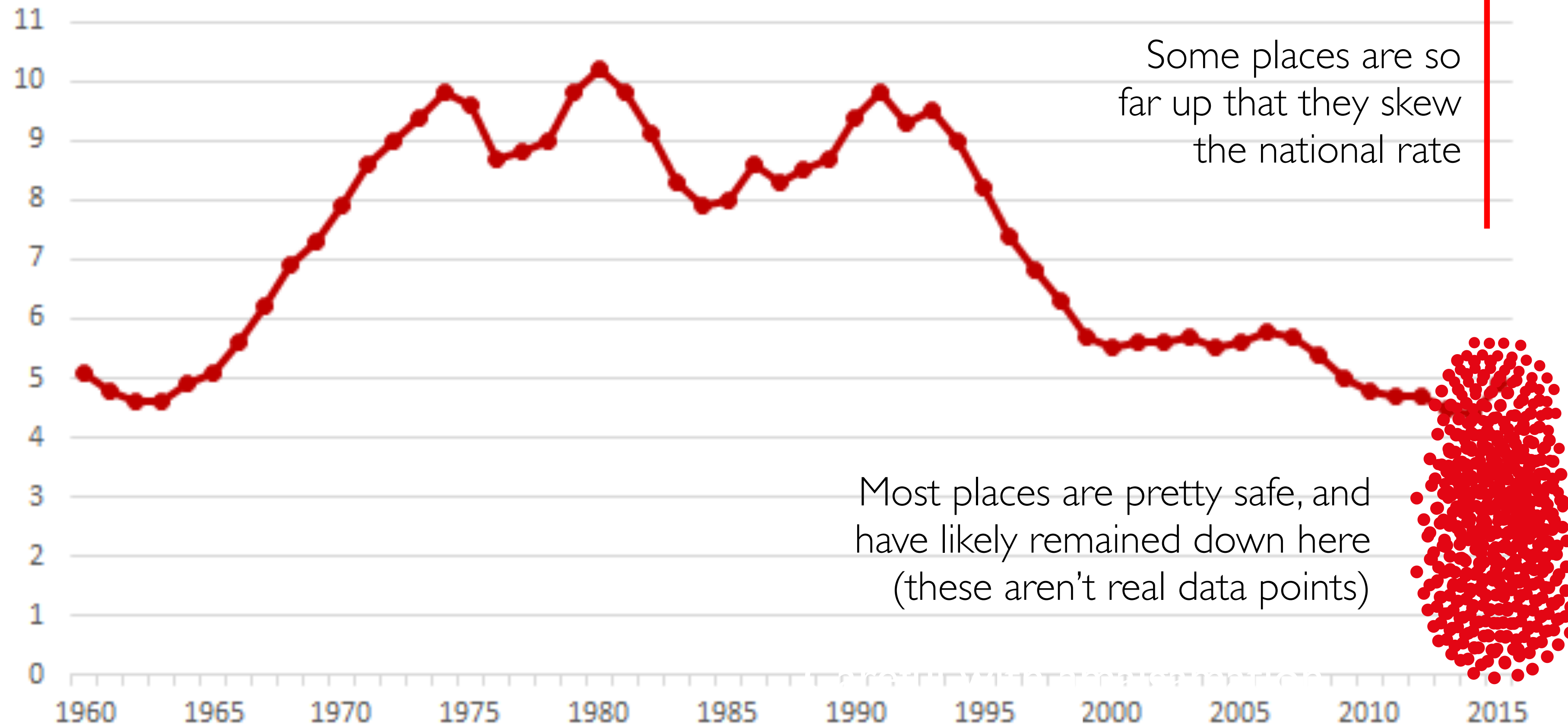
FactCheck.org chart based on FBI data

The U.S. Murder Rate (per 100,000 people)



Most places are pretty safe, and have likely remained down here (these aren't real data points)

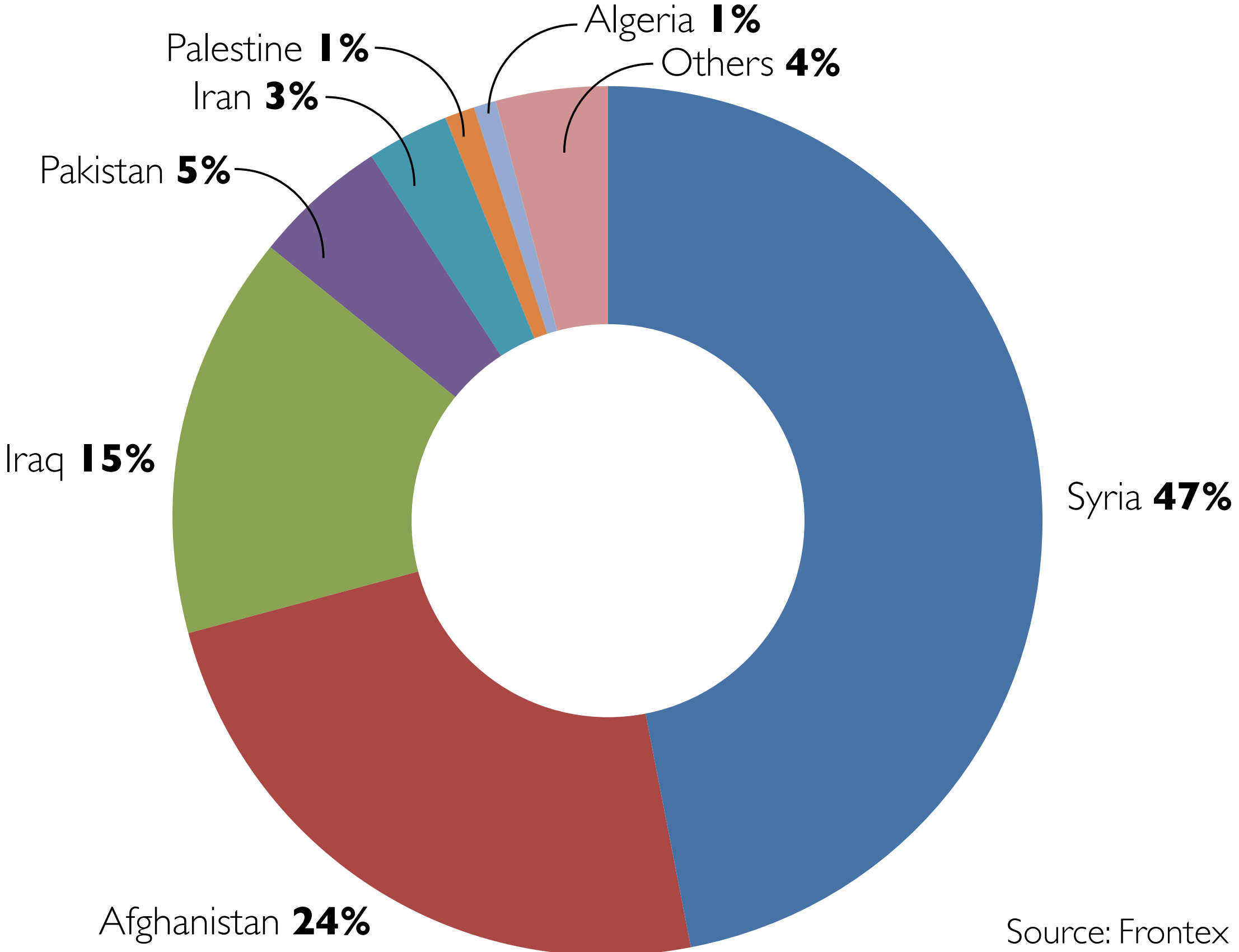
The U.S. Murder Rate (per 100,000 people)



FactCheck.org chart based on FBI data

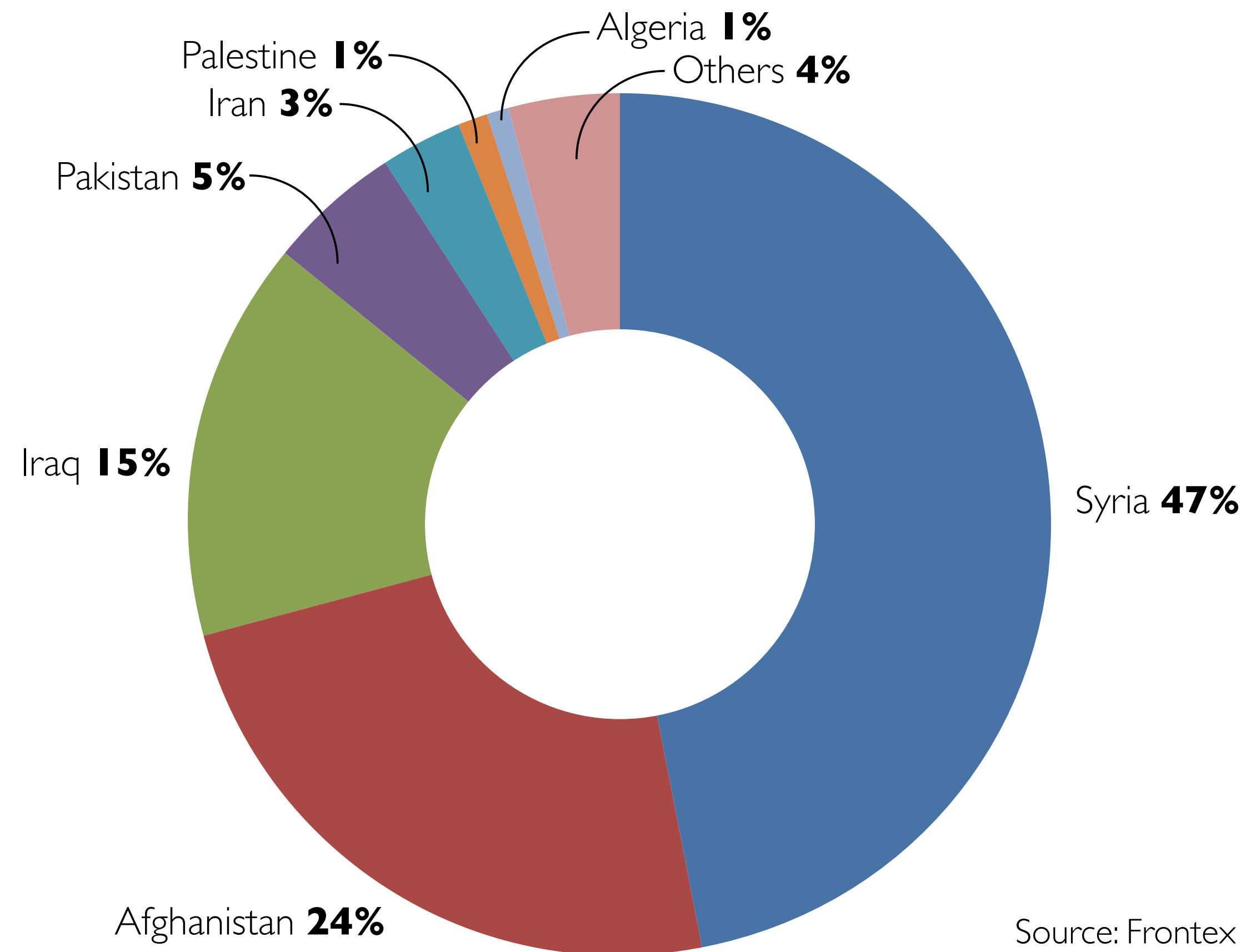
How to show it?

GREECE - Main nationalities of arriving migrants in 2016

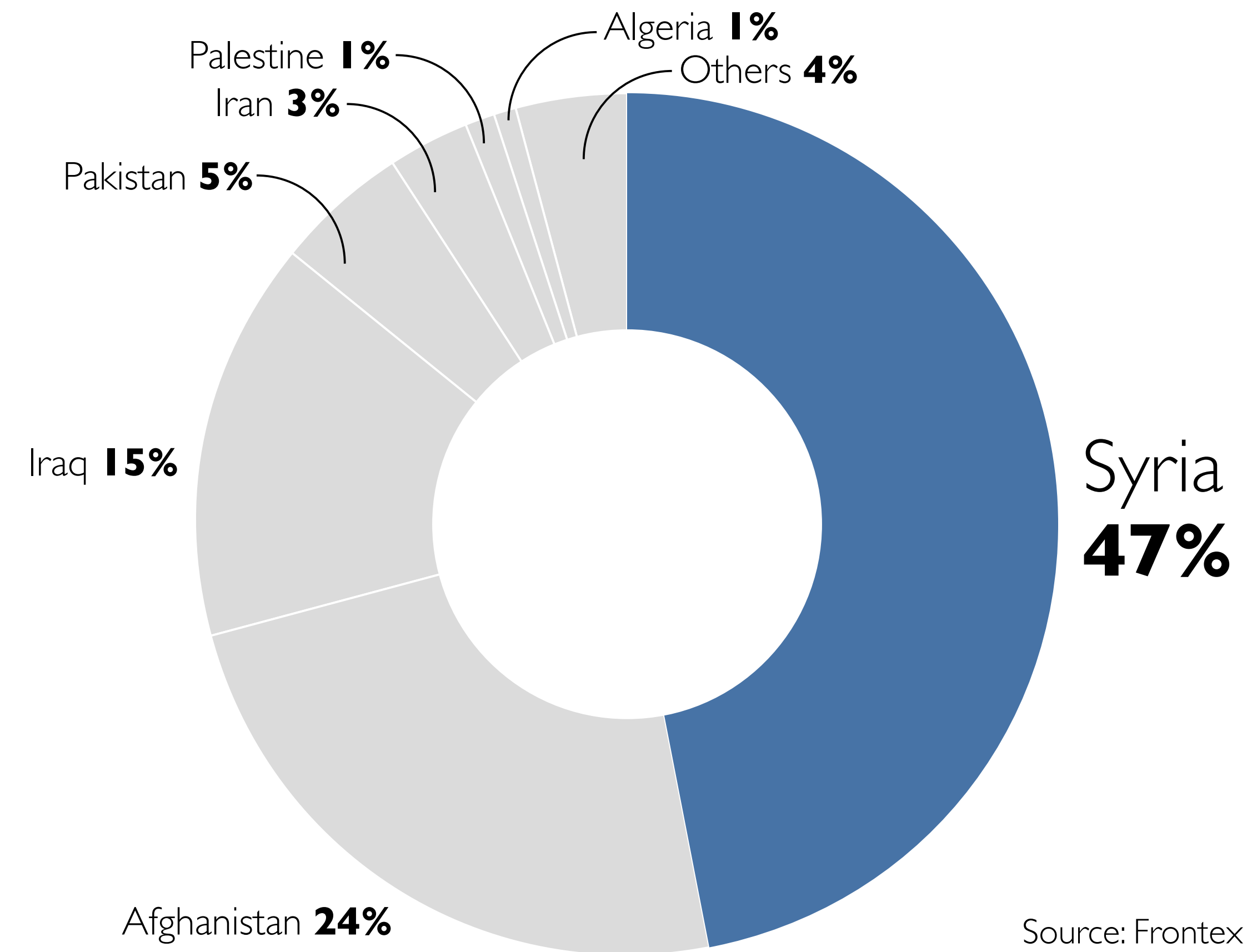


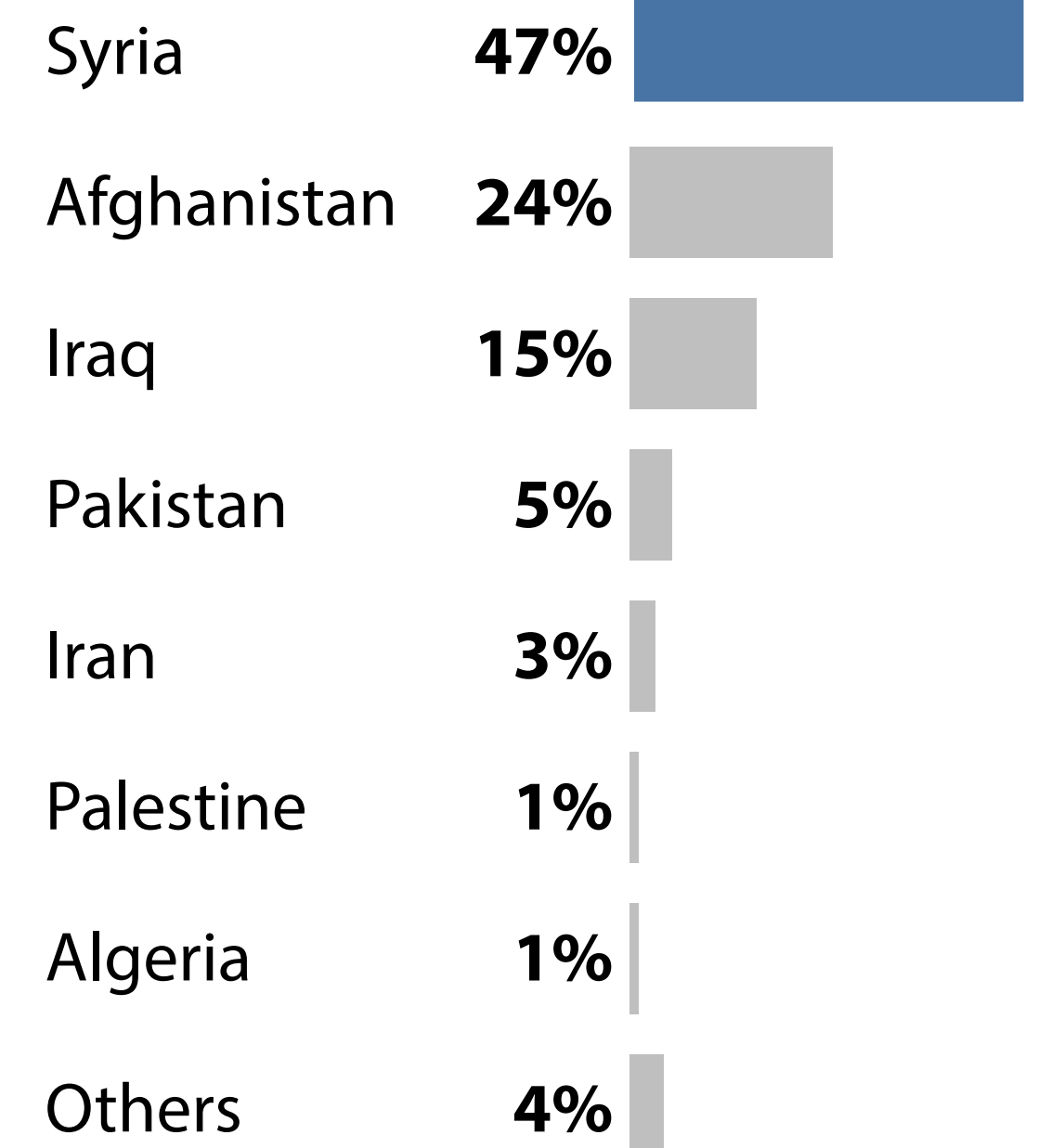
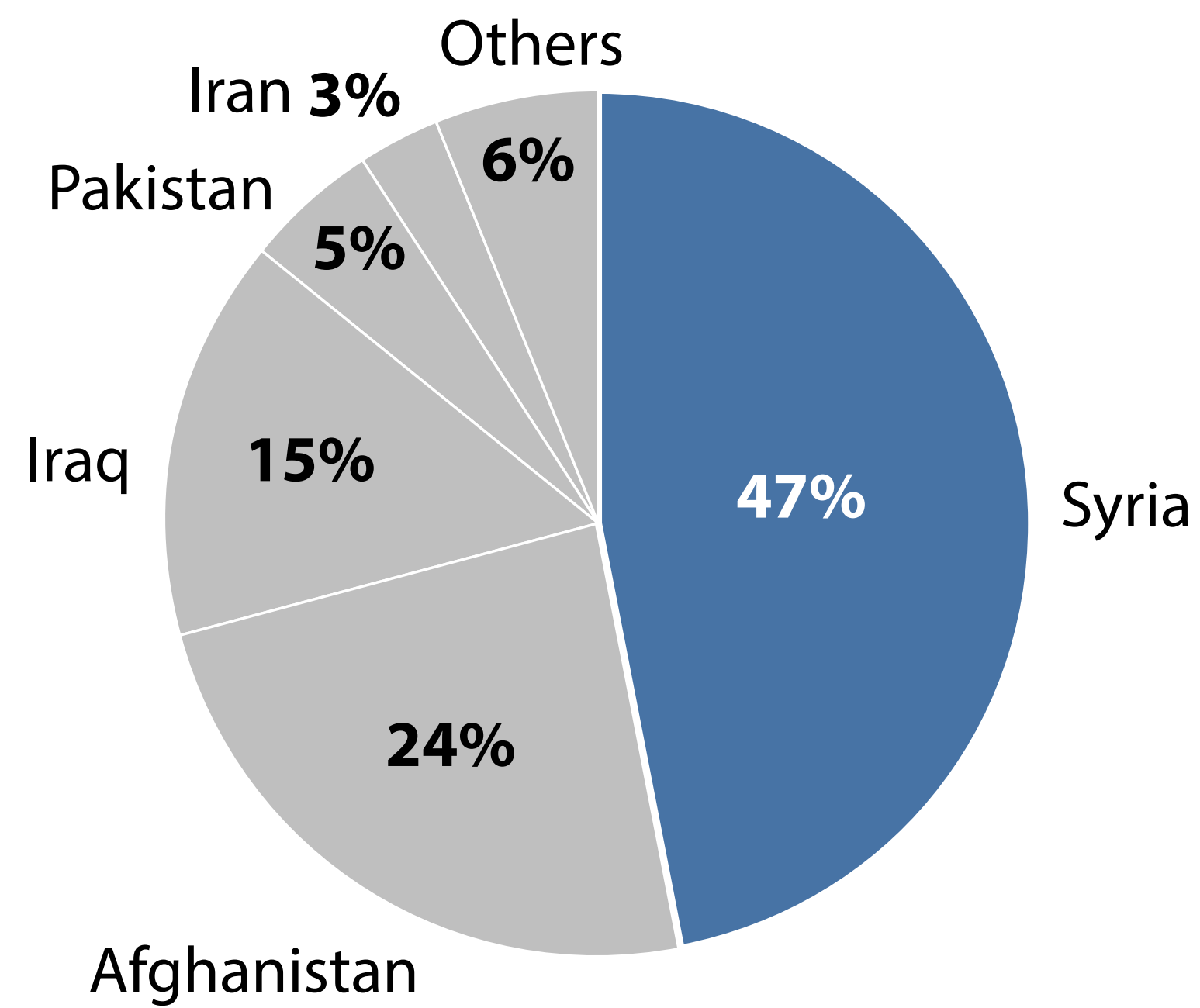
Source: Frontex

GREECE - Main nationalities of arriving migrants in 2016



GREECE - Main nationalities of arriving migrants in 2016

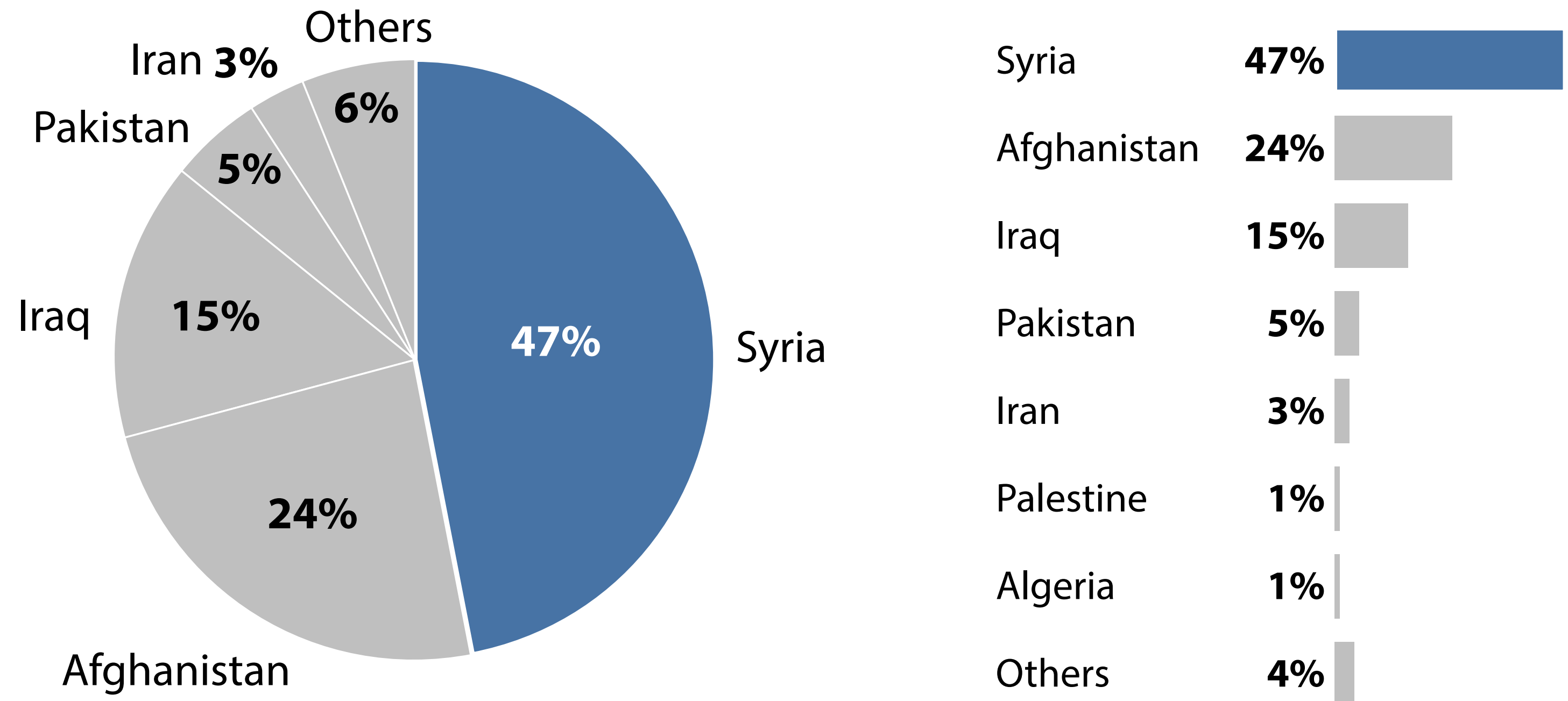
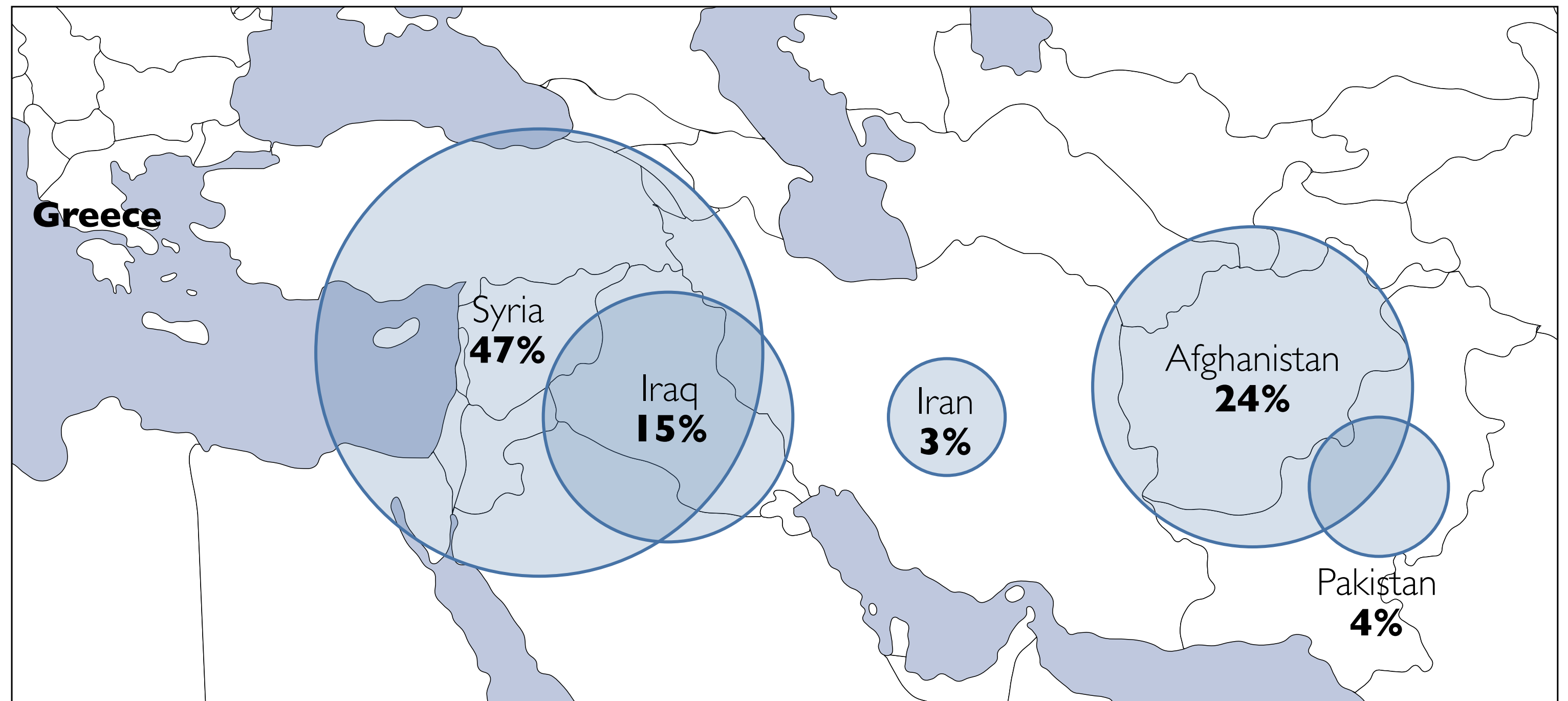




None of these charts is *right* or *wrong*, *good* or *bad* per se.

They are just **better** or **worse** depending on our **intent**:

What it is that we want to communicate, what we want our reader to **be able to see in the data**.



Deviation	Correlation	Ranking	Distribution	Change over Time	Part-to-whole	Magnitude	Spatial	Flow
<p>Emphasise variations (+/-) from a fixed reference point. Typically the reference point is zero but it can also be a target or a long-term average. Can also be used to show sentiment (positive/neutral/negative).</p> <p>Example FT uses Trade surplus/deficit, climate change</p>	<p>Show the relationship between two or more variables. Be mindful that unless you tell them otherwise, many readers will assume the relationships you show them to be causal (i.e. one causes the other).</p> <p>Example FT uses Inflation & unemployment, income & life expectancy</p>	<p>Use where an item's position in an ordered list is more important than its absolute or relative value. Don't be afraid to highlight the points of interest.</p> <p>Example FT uses Wealth, deprivation, league tables, constituency election results</p>	<p>Show values in a dataset and how often they occur. The shape (or 'skew') of a distribution can be a memorable way of highlighting the lack of uniformity or equality in the data.</p> <p>Example FT uses Income distribution, population, (age/sex) distribution</p>	<p>Give emphasis to changing trends. These can be short (30-90-day) movements or extended series, traversing decades or centuries. Choosing the correct time period is important to provide suitable context for the reader.</p> <p>Example FT uses Share price movements, economic time series</p>	<p>Show how a single entity can be broken down into its component elements, if the reader's interest is solely in the size of the components, consider a magnitude-type chart instead.</p> <p>Example FT uses Fiscal budgets, company structures, national election results</p>	<p>Show size comparisons. These can be relative (just being able to see larger/smaller) or absolute (need to see fine differences). Usually these show a 'counted' number (for example, barrels, dollars or people) rather than a calculated rate or per cent.</p> <p>Example FT uses Commodity production, market capitalisation</p>	<p>Used only when precise locations or geographical patterns in data are more important to the reader than anything else.</p> <p>Example FT uses Locator maps, population density, natural resource locations, natural disaster risk impact, catchment areas, variation in election results</p>	<p>Show the reader volumes or intensity of movement between two or more states or conditions. These might be logical sequences or geographical locations.</p> <p>Example FT uses Movement of funds, trade, migrants, lawsuits, information, relationship graphs</p>
<p>Diverging bar A simple standard bar chart that can handle both negative and positive magnitude values.</p>	<p>Scatterplot The standard way to show the relationship between two continuous variables, each of which has its own axis.</p>	<p>Ordered bar Standard bar charts display the ranks of values much more easily when sorted into order.</p>	<p>Histogram The standard way to show a statistical distribution - keep the gaps between columns - use to highlight the 'shape' of the data.</p>	<p>Line The standard way to show a changing time series. If data are irregular, consider markers to represent data points.</p>	<p>Stacked column A simple way of showing part-to-whole relationships but can be difficult to read with more than a few components.</p>	<p>Column The standard way to compare the size of things. Must always start at 0 on the axis.</p>	<p>Basic choropleth (date/ratio) The standard approach for putting data on a map - should always be rates rather than totals and use a sensible base geography.</p>	<p>Sankey Shows changes in flows from one condition to at least one other, good for tracing the eventual outcome of a complex process.</p>
<p>Diverging stacked bar Perfect for presenting survey results which involve sentiment (eg disagree/neutral/agree).</p>	<p>Line + Column A good way of showing the relationship between an amount (columns) and a rate (line).</p>	<p>Ordered column See above.</p>	<p>Boxplot Summarise multiple distributions by showing the median (centre) and range of the data.</p>	<p>Column Columns work well for showing change over time - but usually best with only one series of data at a time.</p>	<p>Proportional stacked bar A good way of showing the size and proportion of data at the same time - as long as the data are not too complicated.</p>	<p>Bar See above. Good when the data are not too time series and labels have long category names.</p>	<p>Proportional symbol (count/magnitude) Use for totals rather than rates - be wary that small differences in data will be hard to see.</p>	<p>Waterfall Designed to show the sequencing of data through a flow process, typically budgets. Can include +/- components.</p>
<p>Spine chart Splits a single value into 2 contrasting components (eg Male/Female).</p>	<p>Connected scatterplot Usually used to show how the relationship between 2 variables has changed over time.</p>	<p>Ordered proportional symbol Use when there are big variations between complex distributions (data that cannot be summarised with simple averages).</p>	<p>Violin plot Similar to a box plot but more effective with complex distributions (data that cannot be summarised with simple averages).</p>	<p>Line + column A good way of showing the relationship over time between an amount (columns) and a rate (line).</p>	<p>Pie A common way of showing part-to-whole data - but be aware that it's difficult to accurately compare the size of the segments.</p>	<p>Paired column As per standard columns but allows for multiple series. Can become tricky to read with more than 2 series.</p>	<p>Flow map For showing movement across a map.</p>	<p>Chord A complex but powerful diagram, which can illustrate 2-way flows (and not always in a matrix).</p>
<p>Surplus/deficit filled line The shaded area of these charts allows a balance to be shown - either against a baseline or between two series.</p>	<p>Bubble Like a scatterplot, but adds additional detail by sizing the circles according to a third variable.</p>	<p>Dot strip plot Dots placed in order on a strip are a space-efficient method of laying out ranks across multiple categories.</p>	<p>Population pyramid A standard way for showing the age and sex breakdown of a population distribution; effectively, back to back histograms.</p>	<p>Stack price Usually focused on showing opening/closing and high/low points of each day.</p>	<p>Donut Similar to a pie chart - but the centre can be a good way of making space to include more information about the data (eg totals).</p>	<p>Paired bar See above.</p>	<p>Contour map For showing areas of equal value on a map. Can use deviation colour schemes for showing +/- values.</p>	<p>Network Used for showing the strength and inter-connectness of relationships of varying types.</p>
	<p>XY heatmap A good way of showing the patterns between 2 categories of data, less good at showing fine differences in amounts.</p>	<p>Slope Perfect for showing how rates have changed over time or vary between categories.</p>	<p>Dot strip plot Good for showing individual values in a distribution, can be a problem when too many dots have the same value.</p>	<p>Slope Good for showing changing data as long as the data can be simplified into 2 or 3 points without missing a key part of story.</p>	<p>Treemap Use for hierarchical part-to-whole relationships, can be difficult to read when there are many small segments.</p>	<p>Proportional stacked bar A good way of showing the size and proportion of data at the same time - as long as the data are not too complicated.</p>	<p>Equalised cartogram Converting each unit on a map to a regular and equally-sized shape - good for representing voting regions with equal value.</p>	<p>Scalped cartogram (value) Stretching and shrinking a map so that each area is sized according to a particular value.</p>
		<p>Lollipop chart Lollipops draw more attention to the data value than standard bar/columns and can also show rank and value effectively.</p>	<p>Dot plot A simple way of showing the change or range (min/max) of data across multiple categories.</p>	<p>Area chart Use with care - these are good at showing changes to total, but seeing change in components can be very difficult.</p>	<p>Voronoi A way of turning points into areas - any point within each area is closer to the central point than any other centroid.</p>	<p>Proportional symbol Use when there are big variations between values and/or seeing fine differences between data is not so important.</p>	<p>Dot density Used to show the location of individual events/locations - make sure to annotate any patterns the reader should see.</p>	
		<p>Barcode plot Like dot strip plots, good for displaying all the data in a table, they work best when highlighting individual values.</p>	<p>Cumulative curve A good way of showing how unequal a distribution is; y axis is always cumulative frequency, x axis is always a measure.</p>	<p>Fan chart (projections) Use to show the uncertainty in future projections, usually this grows the further forward to projection.</p>	<p>Sunburst Another way of visualising hierarchical part-to-whole relationships. Use sparingly (if at all) for obvious reasons.</p>	<p>Isotype (pictogram) Excellent solution in some instances - use only with whole numbers. Do not slice off an arm to represent a decimal.</p>	<p>Heat map Grid-based data values mapped with an intensity colour scale. As choropleth map - but not snapped to an administrative unit.</p>	
		<p>Calendar heatmap A great way of showing temporal patterns (daily, weekly, monthly) - at the expense of showing precision in quantity.</p>	<p>Priestley timeline Great when date and duration are key elements of the story in the data.</p>	<p>Connected scatterplot A good way of showing changing data for two variables wherever there is a relatively clear pattern of progression.</p>	<p>Art A hemicycle, often used for visualising political results in parliaments.</p>	<p>Lollipop chart Lollipop charts draw more attention to the data value than standard bar/columns - does not HAVE to start at zero (but preferable).</p>	<p>Radar chart A space-efficient way of showing value of multiple variables - but make sure they are organised in a way that makes sense to reader.</p>	
		<p>Circle timeline Good for showing discrete values of varying size across multiple categories (eg earthquakes by continent).</p>	<p>Circle timeline Another alternative to the circle timeline for showing series where there are big variations in the data.</p>	<p>Area chart Use with care - these are good at showing changes to total, but seeing change in components can be very difficult.</p>	<p>Gridplot Good for showing % information, they work best when used on whole numbers and work well in multiple layout forms.</p>	<p>Parallel coordinates An alternative to radar charts - again, the arrangement of the variables is important. Usually benefits from highlighting values.</p>		
				<p>Waterfall Can be useful for showing part-to-whole relationships where some of the components are negative.</p>				

Visual vocabulary

Designing with data

There are so many ways to visualise data - how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list is not meant to be exhaustive, nor a wizard, but is a useful starting point for making informative and meaningful data visualisations.

FT graphics: Alan Smith; Chris Campbell; Ian Bott; Liz Faunce; Graham Parrish; Billy Ehrenberg; Paul McCallum; Martin Stabe
Inspired by the Graphic Continuum by Jon Schwabish and Severino Ribeca

ft.com/vocabulary



Search by Function

View by List



Arc Diagram



Area Graph



Bar Chart



Box & Whisker Plot



Brainstorm



Bubble Chart



Calendar



Chord Diagram



Choropleth Map



Circle Packing



Donut Chart



Dot Matrix Chart



Flow Map



Histogram



Illustration Diagram



Line Graph



Marimekko Chart



Multi-set Bar Chart



Nightingale Rose Chart



Non-ribbon Chord Diagram



Parallel Sets



Pictogram Chart



Pie Chart



Population Pyramid



Proportional Area Chart



Radar Chart



Radial Bar Chart



Sankey Diagram



Scatterplot



Span Chart

<http://www.datavizcatalogue.com/>

Who is the audience?



11:00 PM

Hurricane Lee

15.4°N, 47.7°W

Max Winds: **80 mph**

Moving: **WNW at 14 mph**

Pressure: **989 mb**



What we design...

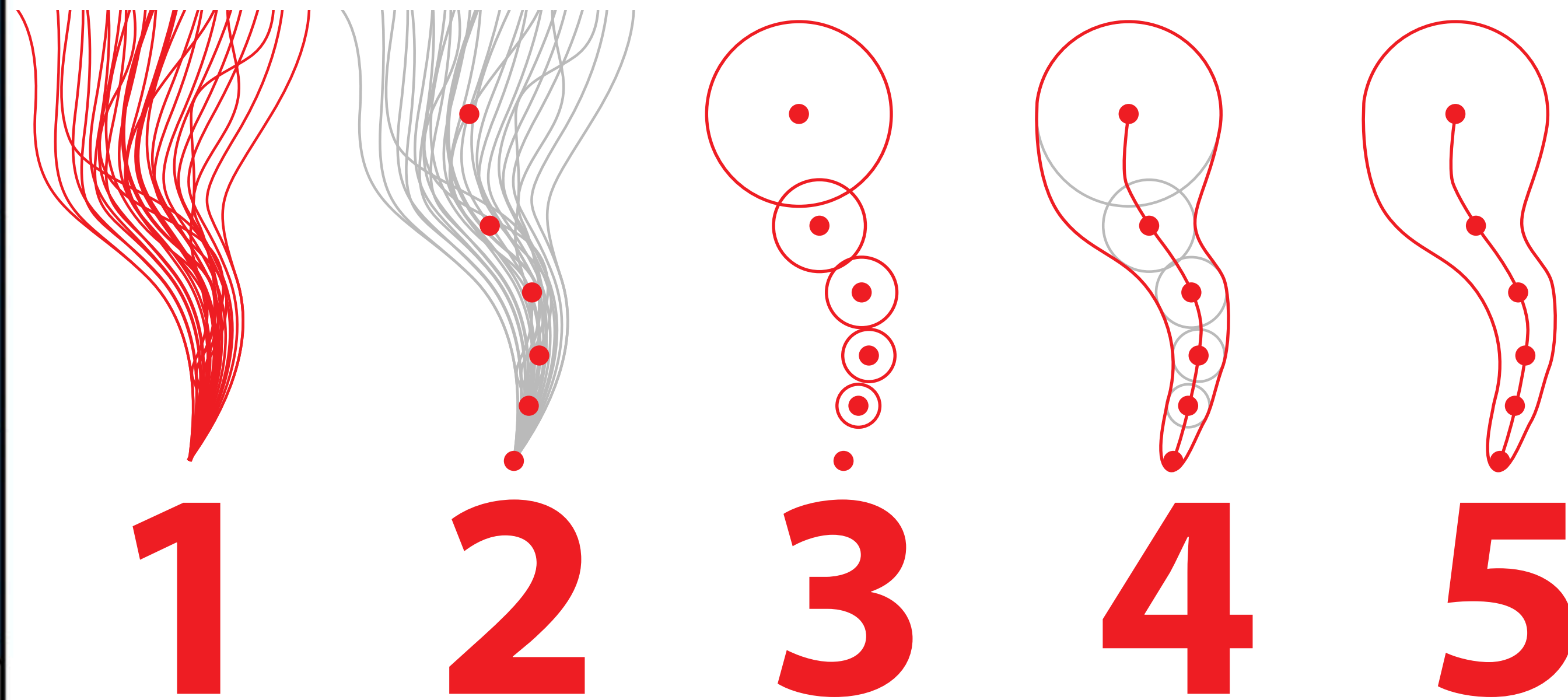


What we design...

...is not what people see



How to read the NHC cone of uncertainty



Opinion

Those Hurricane Maps Don't Mean What You Think They Mean

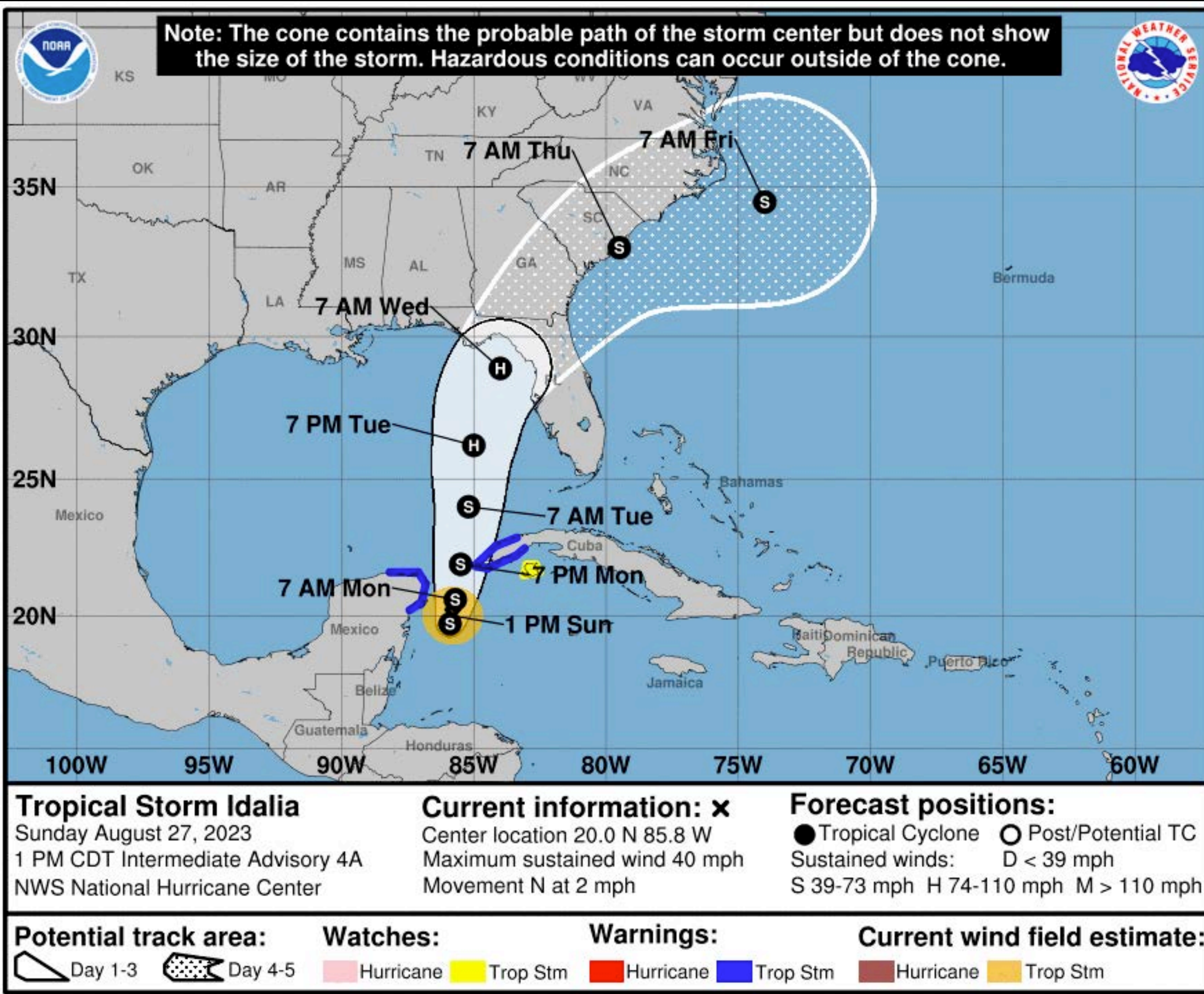
We use hurricane forecasts to warn people. Why do we misinterpret them so often?

By Alberto Cairo
With Tala Schlossberg

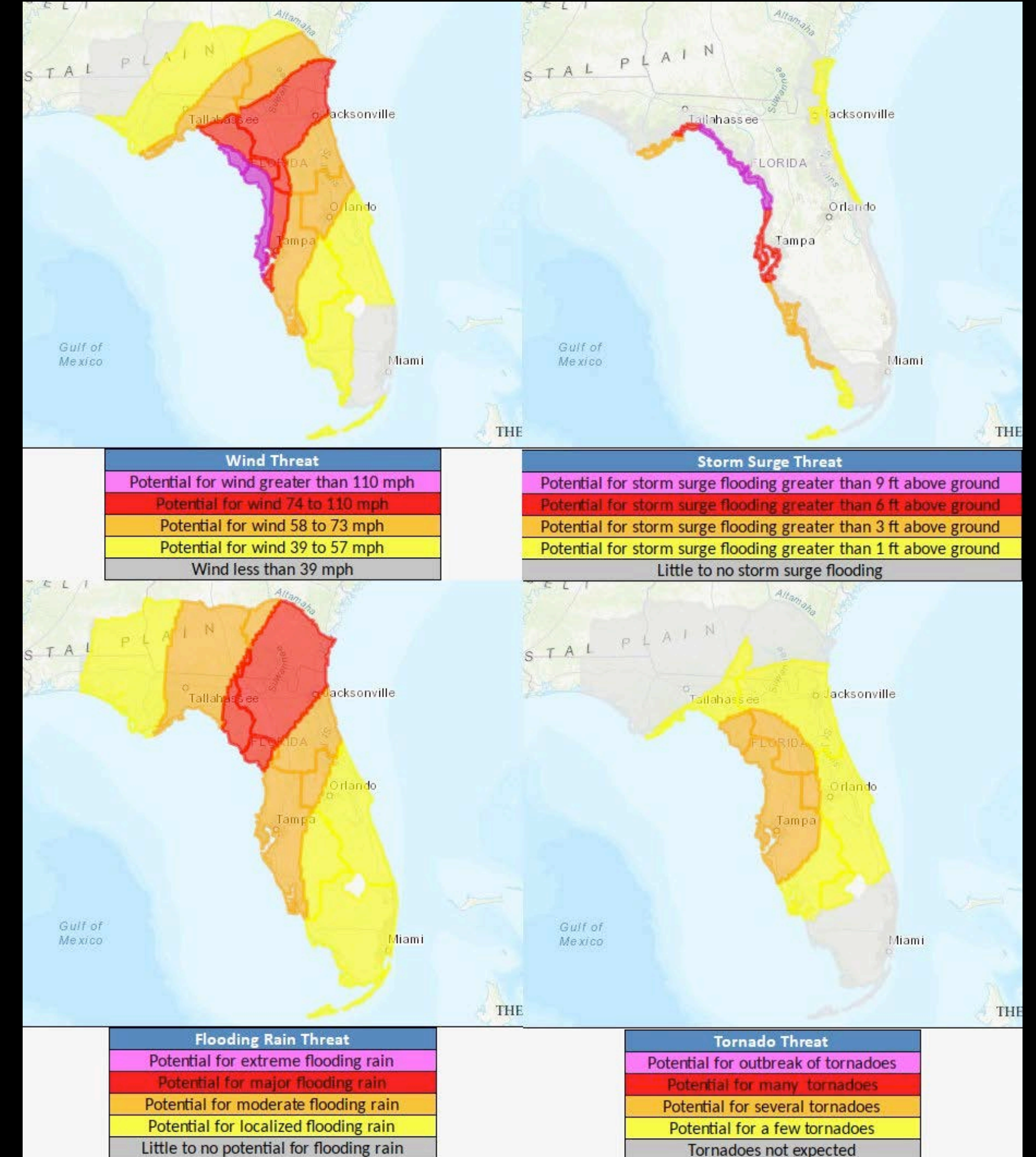
<https://www.nytimes.com/interactive/2019/08/29/opinion/hurricane-dorian-forecast-map.html>

What do readers really need to know?

The possible path?

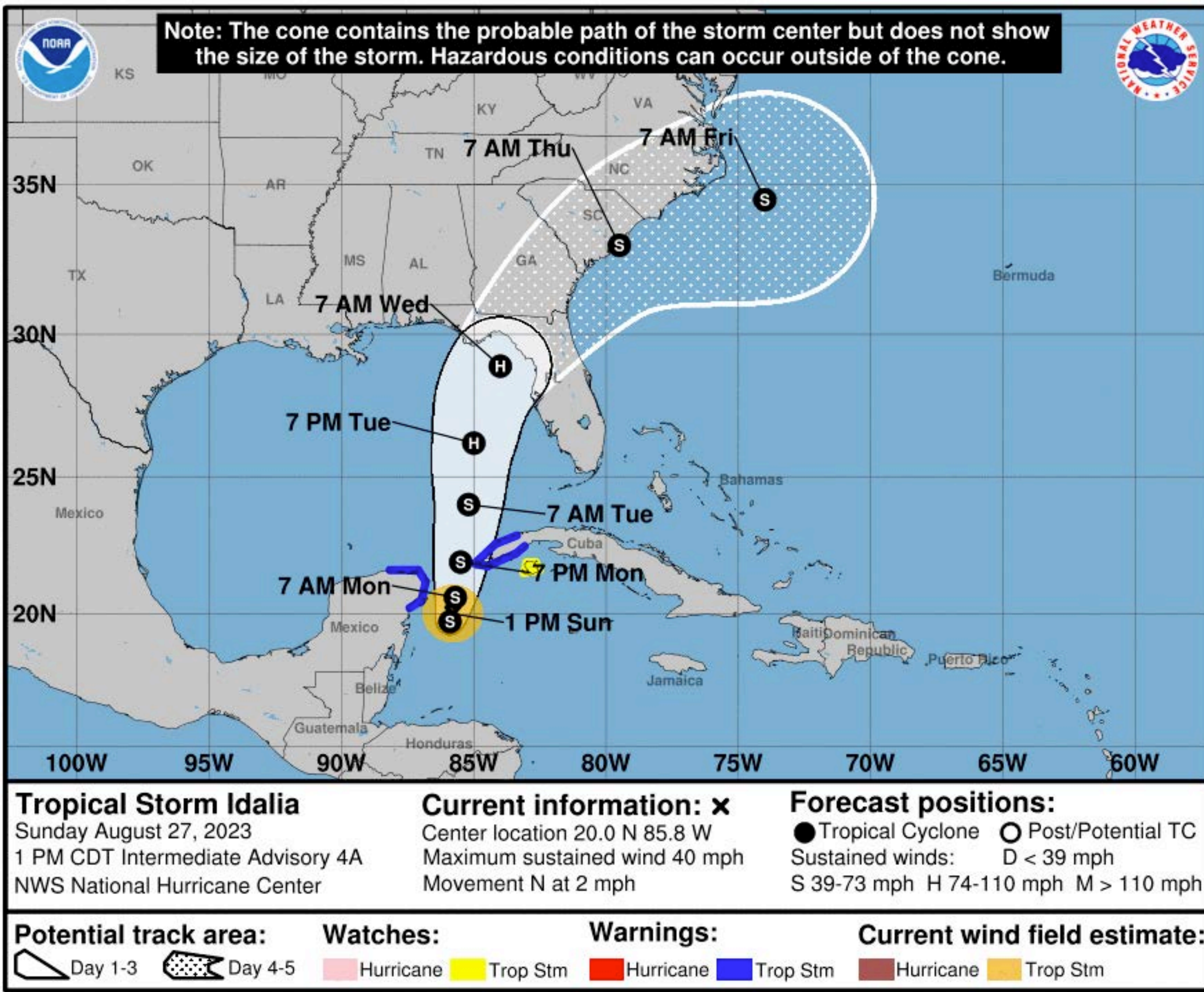


Or a hurricane's threats and impacts (HTI)?

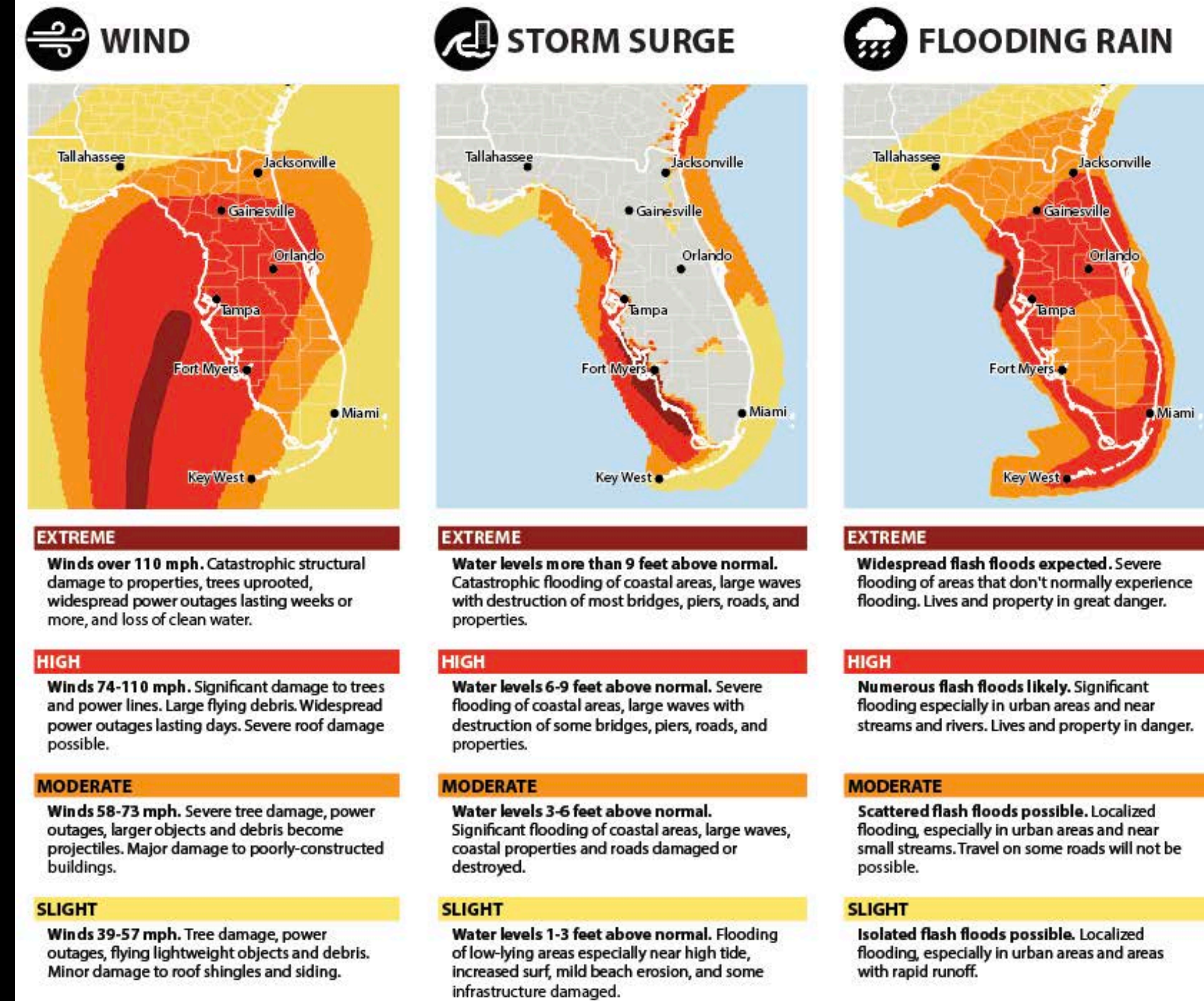


What do readers really need to know?

The possible path?



Or a hurricane's threats and impacts (HTI)?



Designs by the University of Miami's 'Hurakan' research group

How to make it look good?

Multi-scale Modeling and Assessment of Malaria Risk in Northern South America



Alimi, T. O.¹; Fuller, D. O.^{1,2} and Beier, J.C.^{1,3}

¹ Abess Center for Ecosystem Science and Policy; ² Department of Geography and Regional Studies; ³ Department of Epidemiology and Public Health, University of Miami

1. Introduction

The public health problem posed by malaria has made it a top priority for control efforts and the general consensus globally, is that its elimination is crucial for continued international development. Consequently, there is ongoing research in different regions including South America (SA) to better understand the disease dynamics with the intent that findings may establish scientific framework that would support the development of new intervention strategies for malaria elimination in areas with seasonal malaria. One of such investigations is undertaken by the International Centers of Excellence in Malaria Research (ICEMR) under a National Institutes of Health (NIH) grant.

While only about 3% of the global malaria burden is borne by SA¹, undertaking malaria research in the region is currently important because an estimated 23million people are still at risk² and approximately about 80% of clinical cases are found in Northern South America (NSA)³. A key factor limiting effective control is lack of data and uneven implementation of control measures, including use of bed-nets, sprays, early diagnosis, and treatment. As part of the ICEMR investigation, this project seeks to model the spatial patterns of malaria risk in NSA through vector distribution and land-use changes. Furthermore, I intend to investigate the perceptions of malaria risk in order to identify barriers to adoption and how they can be circumvented.

2. Significance

Spatial distribution of malaria risk is still perceived as broadly categorized by the WHO's traditional risk maps which are highly generalized, of low resolution and have broad categories with uncertain boundaries (see da Nunes-Silva et al. 2012). There is need for up- to-date high resolution risk maps which can aid malaria control efforts. Secondly, modeling distribution of principal malaria vectors and land use changes which may explain the observed distribution and risk are useful tools which would guide future management strategies. Finally, understanding the perceptions of at risk populations may help address barriers to adoption of interventions and influence policies. Overall, findings will empower NMCPs to achieve effective control and move them closer to elimination.

3. Specific Aims

- Specific Aim 1: Model the spatial patterns of malaria risk through vector distribution and land use changes
 - Hypothesis 1.1: GIS-based Multi-Criteria Evaluation (MCE) model can accurately predict spatial extent of malaria risk areas. **Objective:** Generate risk maps that represent risk of malaria transmission.
 - Hypothesis 1.2: The Maximum Entropy (Maxent) model can accurately depict actual and predict potential distribution of three *Anopheles* species. **Objective:** Model observed and potential spread of *An. albimanus*, *An. darlingi*, and *An. nuneztovari*.
 - Hypothesis 1.3: Land- use changes can explain the variations in predicted malaria risk. **Objective:** Characterize land use land cover (LULC) and investigate changes in areas of risk.
- Specific Aim 2: Investigate the perceptions of malaria risk in order to identify barriers to adoption and how they can be circumvented.
 - Hypothesis 2.1: Knowledge of perception of malaria risk can aid design of malaria control strategies. **Objective:** Obtain and analyze data on subjective perceptions of risk.
 - Hypothesis 2.2: Identification of barriers to adoption of malaria control interventions provide means of tackling them. **Objective:** Analyze data addressing perceived barriers and policy implications

*Only ongoing work on Hypothesis 1.1 in presented here

4. Materials and Methods

- **Study Area:** is NSA comprising of ten countries- Bolivia, Brazil, Colombia, Ecuador, French Guiana, Guyana, Panama, Peru, Suriname and Venezuela. These countries account for approximately 90% of clinical cases in the region hence, the choice as study area (Fig. 1).



Figure1: Map of study area

- **Research Approach:** Due to the complexity of malaria problem, I'm employing an interdisciplinary approach to address the problem (Fig. 2).

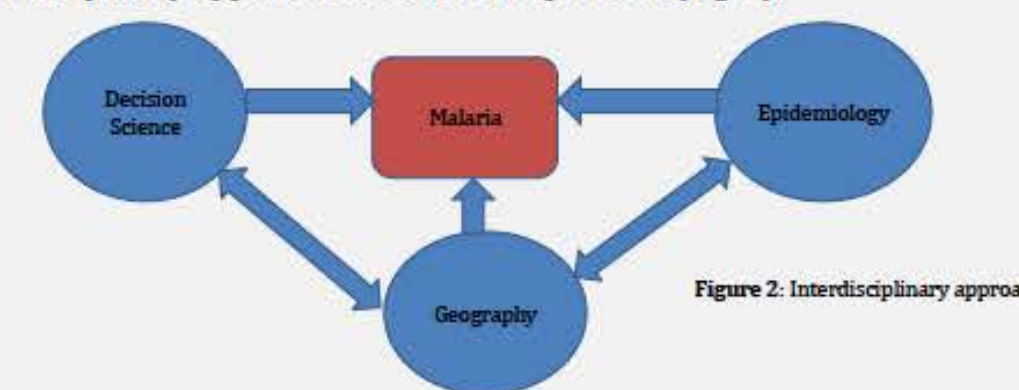


Figure 2: Interdisciplinary approach

- **Materials:** Raster data layers of environmental, climatic and anthropogenic parameters from satellite imageries, weather monitoring stations, global land cover and population data were collected from Worldclim, Digital Charts of the World, Globcover and Landsat. Vector data was collected from field sampling by our collaborators and the Walter Reed Biosystematics Unit. Sociological data would be collected through questionnaires to be administered in one of the study area. Other data will be collected as needed.
- **Procedure:** To test hypothesis 1.1, raster data of parameters that influence mosquito distribution (rivers, wetlands, urban areas, roads, population and elevation) were combined using a Multi-Criteria Evaluation in Idrisi GIS package. This produced a map of potential exposure to malaria vectors which is used as a proxy for risk of malaria transmission. All the data layers were gridded at 1km spatial resolution. A set of distance layers had been created for discrete factors using standard GIS operations. All factors were subsequently standardized into a continuous common numeric range on a byte 0-255 probability scale using a fuzzy function based on knowledge of mosquito interaction with the factor. Weights were generated for each factor based on the importance of the factor to malaria transmission by expert opinions and then assigned using Analytical Hierarchy Process. The risk maps produced were validated statistically using data on *An. darlingi* distribution and malaria case data from some parts of the study area. See preliminary results (Fig. 3,4,5)

5. Preliminary Results

- Areas of high to moderate risk corresponded with locations of some of the anophelines collected.

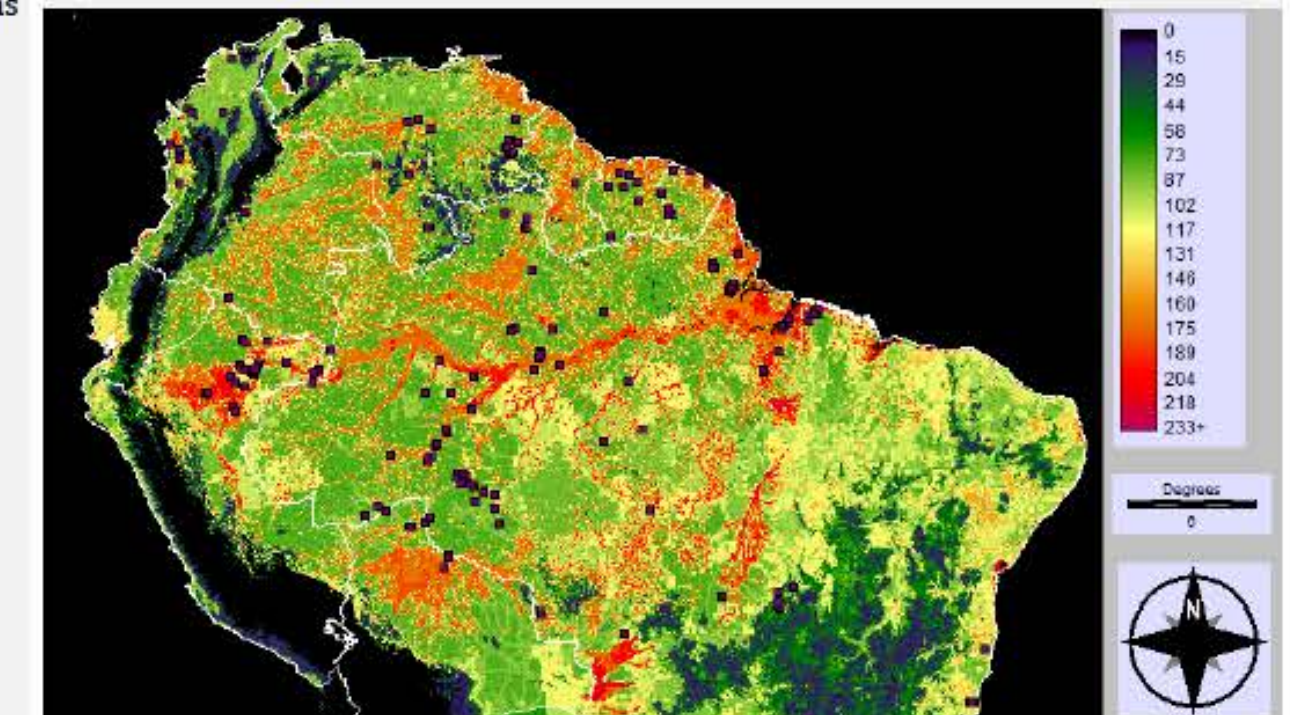


Figure 3: Potential risk of exposure to malaria vectors across NSA (0 indicate little or no risk while 233 indicate high risk)

- Risk scores for mosquito occurrence points were significantly higher than those generated randomly (Fig. 4).

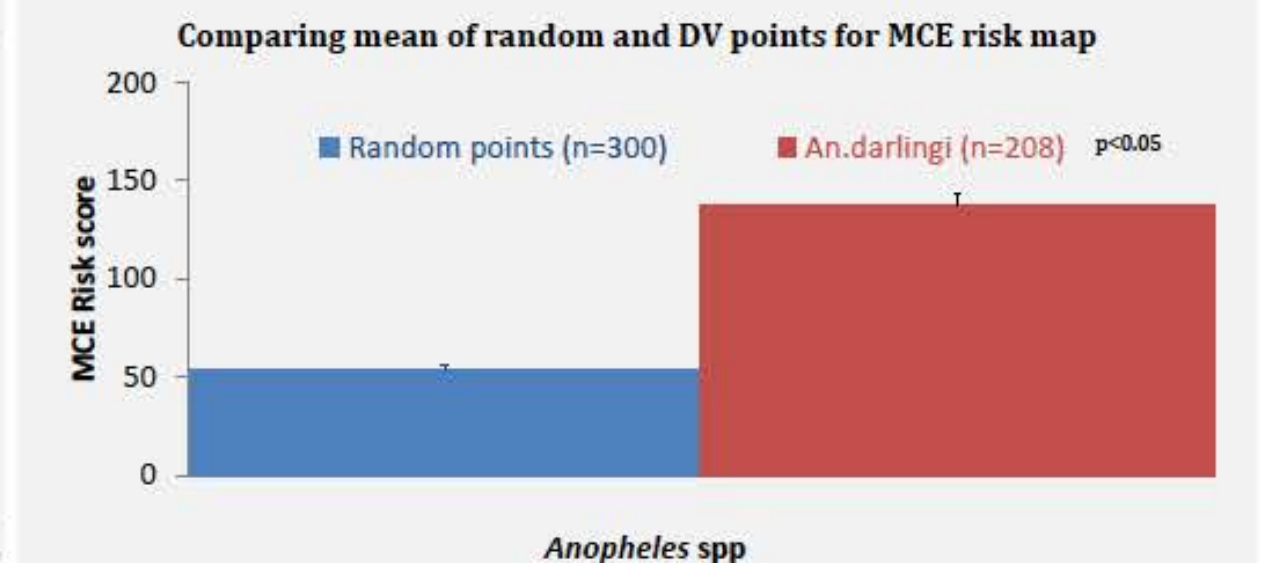


Figure 4: Plot showing the MCE risk values for randomly sampled points and for occurrence points of a DV, *An. darlingi*

6. Conclusion

Findings from preliminary results suggest that the MCE approach is a viable method to modeling spatial risk. The high resolution risk map produced aligned well with sampled vector points and may therefore be used to plan control of malaria vectors. Further analysis is planned to generate and validate risk maps with actual measures of malaria transmission, results of which could be used to plan containment of future outbreaks.

References

1. WHO. (2007). MALARIA ELIMINATION: A field manual for low and moderate endemic countries
2. PAHO (2012) PAHO Honors 2012 Malaria Champions of the Americas. Available: http://new.paho.org/hq/index.php?option=com_content&view=article&id=7429&Itemid=39639
3. Gusmao R. (1999) Overview of malaria control in the Americas. Parasitologia 41:355-60.
4. Da Silva-Nunes, M., Moreno, M., Conn, J.E., Gamboa, D., Abeles, S., Vinetz, J.M., and Ferreira, M.U. (2012) Amazonian malaria: Asymptomatic human reservoirs, diagnostic challenges, environmentally driven changes in mosquito vector populations, and the mandate for sustainable control strategies. Acta Tropica 121 (3): 281-29

Polishing
our designs

Multi-scale Modeling and Assessment of Malaria Risk in Northern South America

Alimi, T. O.¹; Fuller, D. O.^{1,2} and Beier, J.C.^{1,3}



INTRODUCTION

Malaria as a public health problem has become a priority for control efforts worldwide. The global consensus is that its elimination is crucial for continual development. Ongoing research projects in different regions, including South America (SA), try to improve our understanding of the disease dynamics. Their goal is to establish a new framework that would lead to new intervention strategies for malaria elimination in areas where the disease is seasonal. One of such investigations is undertaken by the International Centers of Excellence in Malaria Research (ICEMR) under a National Institutes of Health grant.

While only about 3% of the global malaria burden is borne by SA, undertaking malaria research in the region is currently important because an estimated 23 million people are still at risk and approximately about 80% of clinical cases are found in **Northern South America (NSA)**. A key factor limiting effective control is lack of data and uneven implementation of control measures, including use of bednets, sprays, early diagnosis, and treatment. As part of the ICEMR investigation, this project seeks to model the spatial patterns of malaria risk in NSA through vector distribution and land-use changes. Furthermore, I intend to investigate the perceptions of malaria risk in order to identify barriers to adoption and how they can be circumvented.

SIGNIFICANCE

Spatial distribution of malaria risk is still perceived as broadly categorized by the WHO's traditional risk maps which are highly generalized, of low resolution and have broad categories with uncertain boundaries (see da Nunes-Silva et al. 2012). There is need for up-to-date high resolution risk maps which can aid malaria control efforts. Secondly, modeling distribution of principal malaria vectors and land use changes which may explain the observed distribution and risk are useful tools which would guide future management strategies. Finally, understanding the perceptions of at risk populations may help address barriers to adoption of interventions and influence policies. Overall, findings will empower NMCPs to achieve effective control and move them closer to elimination.

AIMS

Specific Aim 1: Model the spatial patterns of malaria risk through vector distribution and land use changes

- **Hypothesis 1.1:** GIS-based Multi-Criteria Evaluation (MCE) model can accurately predict spatial extent of malaria risk areas. **Objective:** Generate risk maps that represent risk of malaria transmission.
- **Hypothesis 1.2:** The Maximum Entropy (Maxent) model can accurately depict actual and predict potential distribution of three Anopheles species. **Objective:** Model observed and potential spread of *An. albimanus*, *An. darlingi*, and *An. nuneztovari*.
- **Hypothesis 1.3:** Land-use changes can explain the variations in predicted malaria risk. **Objective:** Characterize land use land cover (LULC) and investigate changes in areas of risk.

Specific Aim 2: Investigate the perceptions of malaria risk in order to identify barriers to adoption and how they can be circumvented.

- **Hypothesis 2.1:** Knowledge of perception of malaria risk can aid design of malaria control strategies. **Objective:** Obtain and analyze data on subjective perceptions of risk.
- **Hypothesis 2.2:** Identification of barriers to adoption of malaria control interventions provide means of tackling them. **Objective:** Analyze data addressing perceived barriers and policy implications.

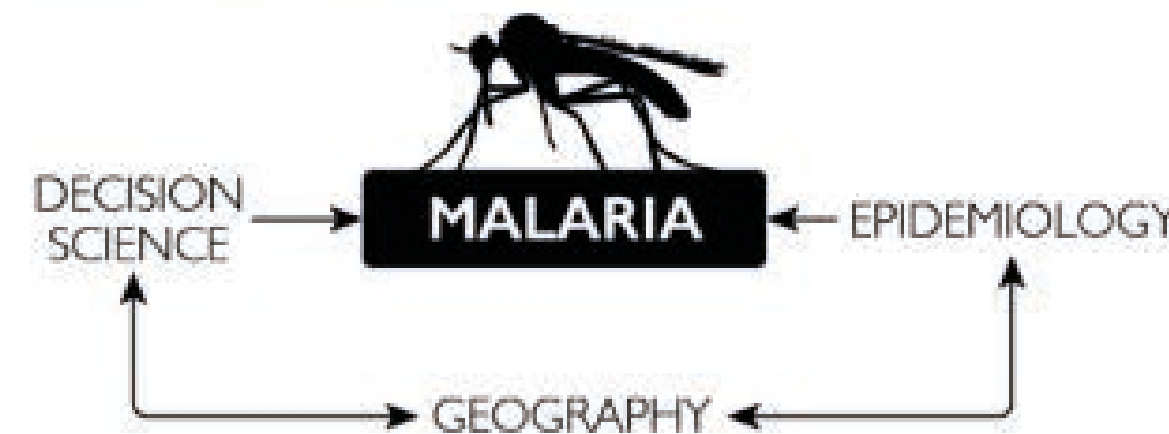
*Only ongoing work on Hypothesis 1.1 is presented here

MATERIALS AND METHODS

NSA comprising of ten countries - Bolivia, Brazil, Colombia, Ecuador, French Guiana, Guyana, Panama, Peru, Suriname and Venezuela. These countries account for approximately 90% of clinical cases in the region.



Research approach: Due to the complexity of malaria problem, I'm employing an interdisciplinary approach to address the problem.

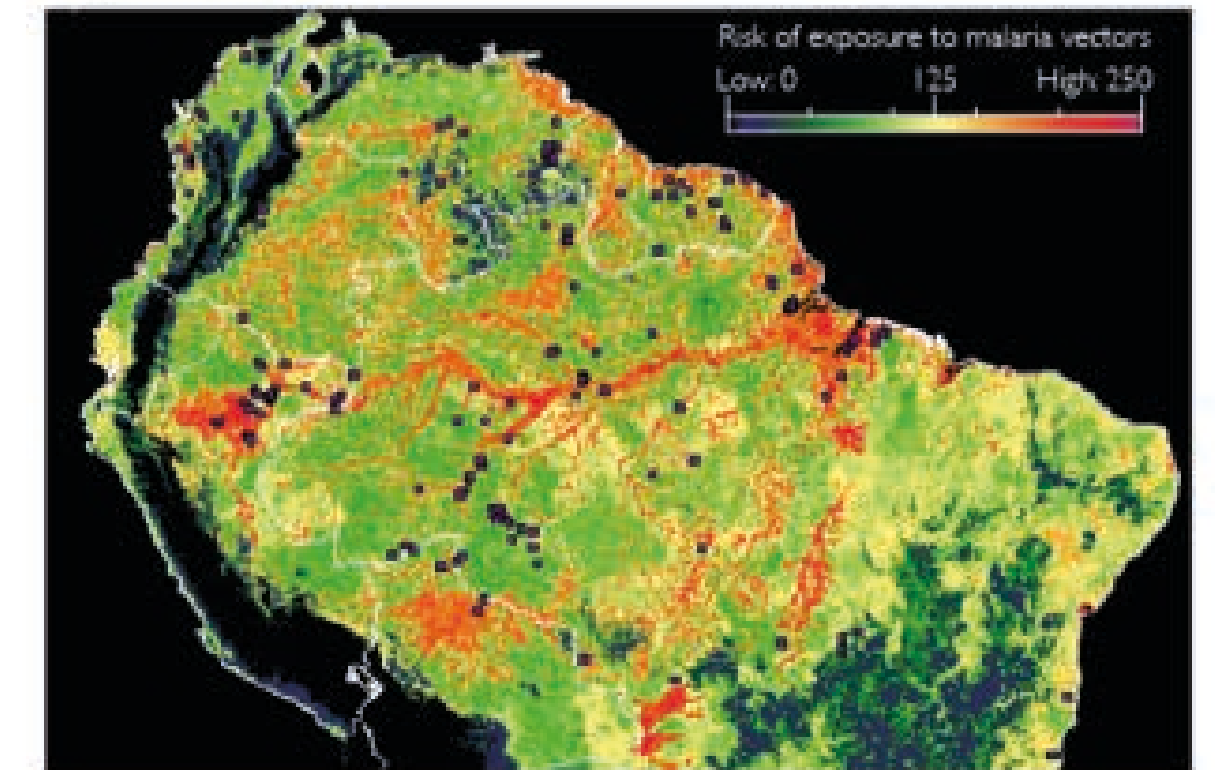


Materials: Raster data layers of environmental, climatic and anthropogenic parameters from satellite imagery, weather monitoring stations, global land cover and population data were collected from WorldClim, Digital Charts of the World, Globcover and Landscan. Vector data was collected from field sampling by our collaborators and the Walter Reed Biosystematics Unit. Sociological data would be collected through questionnaires to be administered in one of the study area. Other data will be collected as needed.

Procedure: To test hypothesis 1.1, raster data of parameters that influence mosquito distribution (rivers, wetlands, urban areas, roads, population and elevation) were combined using a Multi-Criteria Evaluation in Idrisi GIS package. This produced a map of potential exposure to malaria vectors which is used as a proxy for risk of malaria transmission. All the data layers were gridded at 1km spatial resolution. A set of distance layers had been created for discrete factors using standard GIS operations. All factors were subsequently standardized into a continuous common numeric range on a byte 0-255 probability scale using a fuzzy function based on knowledge of mosquito interaction with the factor. Weights were generated for each factor based on the importance of the factor to malaria transmission by expert opinions and then assigned using Analytical Hierarchy Process. The risk maps produced were validated statistically using data on *An. darlingi* distribution and malaria case data from some parts of the study area. See preliminary results

RESULTS

Areas of high to moderate risk corresponded with locations of some of the anophelines collected.



Risk scores for mosquito occurrence points were significantly higher than those generated randomly.

Comparing mean of random and DV points for MCE risk map - $p < 0.05$



CONCLUSION

Findings from preliminary results suggest that the MCE approach is a viable method to modeling spatial risk. The high resolution risk map produced aligned well with sample vector points and may therefore be used to plan control of malaria vectors. Further analysis is planned to generate and validate risk maps with actual measures of malaria transmission, results of which could be used to plan containment of future outbreaks.

References

1. WHO. (2007). MALARIA ELIMINATION: A field manual for low and moderate endemic countries.
2. PAHO. (2012). PAHO Honors 2012: Malaria Champions of the Americas. Available: http://www.paho.org/hq/index.php?option=com_content&view=article&id=7420&Itemid=20632
3. Guisao R. (1999) Overview of malaria control in the Americas. Parasitologia 41: 855-80.
4. Da Silva-Nunes, M., Moreno, M., Conn, J.E., Gamboa, D., Abalos, S., Venz, J.M., and Ferreira, A.U. (2012) Amazonian malaria: Asymptomatic human reservoirs, diagnostic challenges, environmentally driven changes in mosquito vector populations, and the mandate for sustainable control strategies. Acta Tropica 121 (3): 281-29.

Polishing
our designs

Designing information graphics and data visualizations doesn't consist of applying **rules**, but of **reasoning** about choices, and **justifying** them.

Every choice in design is **subjective**, and therefore debatable, but it should never be **arbitrary**.

THANK YOU

www.albertocairo.com

www.openvisualizationacademy.com