# Best Practices for Telling Stories with Data

**Arthi Padmanabhan**

Analyst, E Source

**Heather Hilgenkamp**

Lead Analyst, E Source

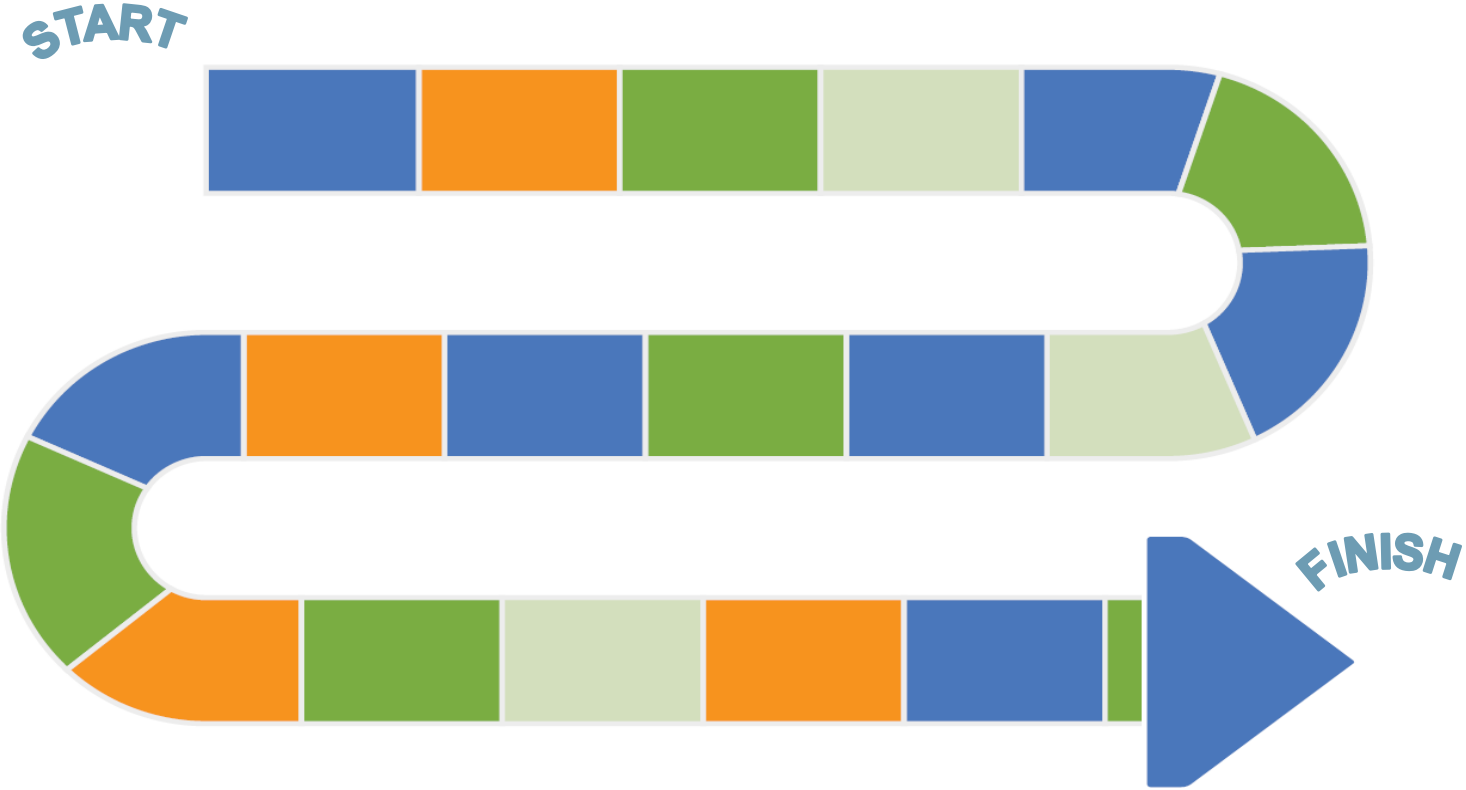
E Source

Fall 2019 E Source Market Research Leadership Council

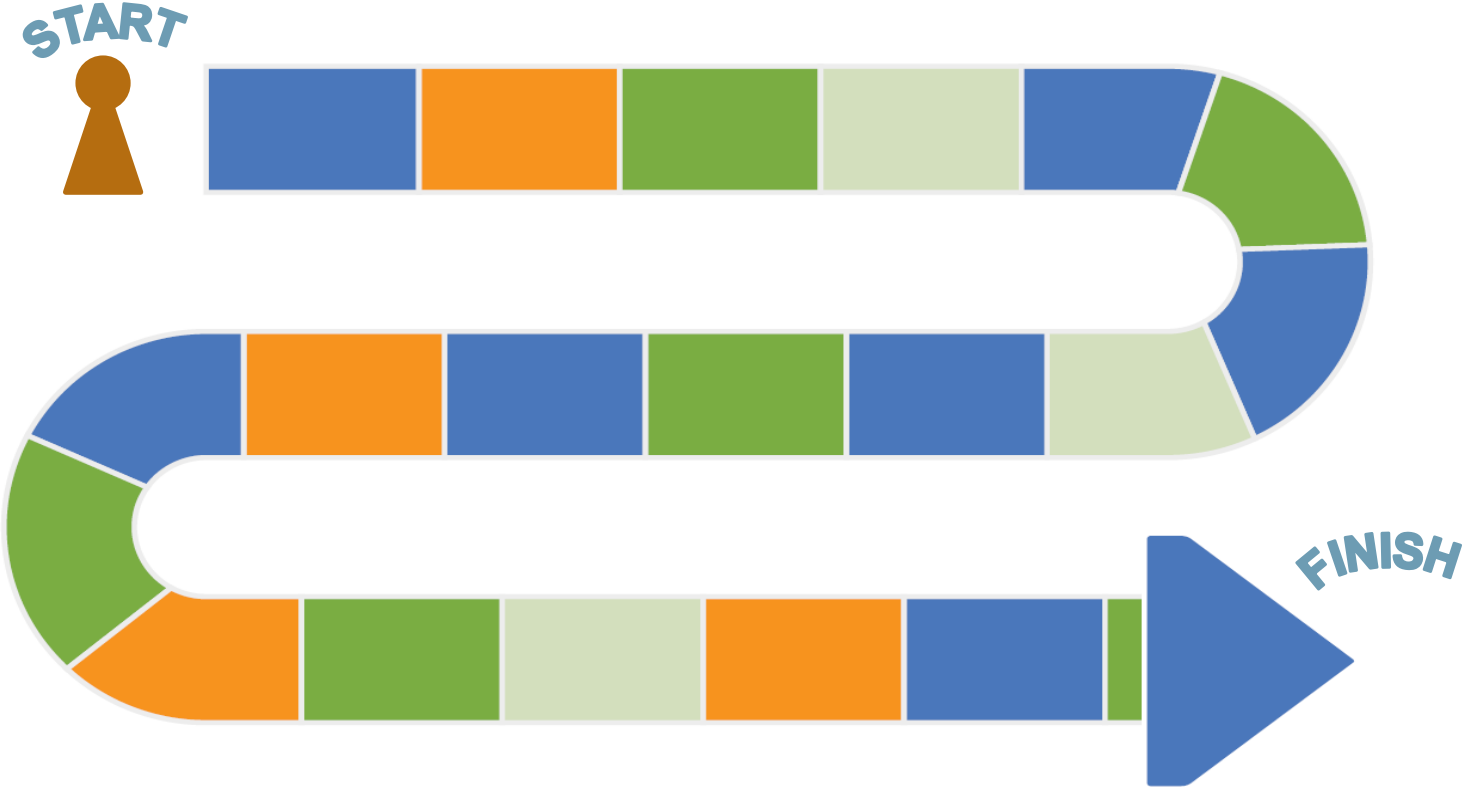
# Key points to good data storytelling

- Know what story you want to tell before you create your final graphics

# The path to insights

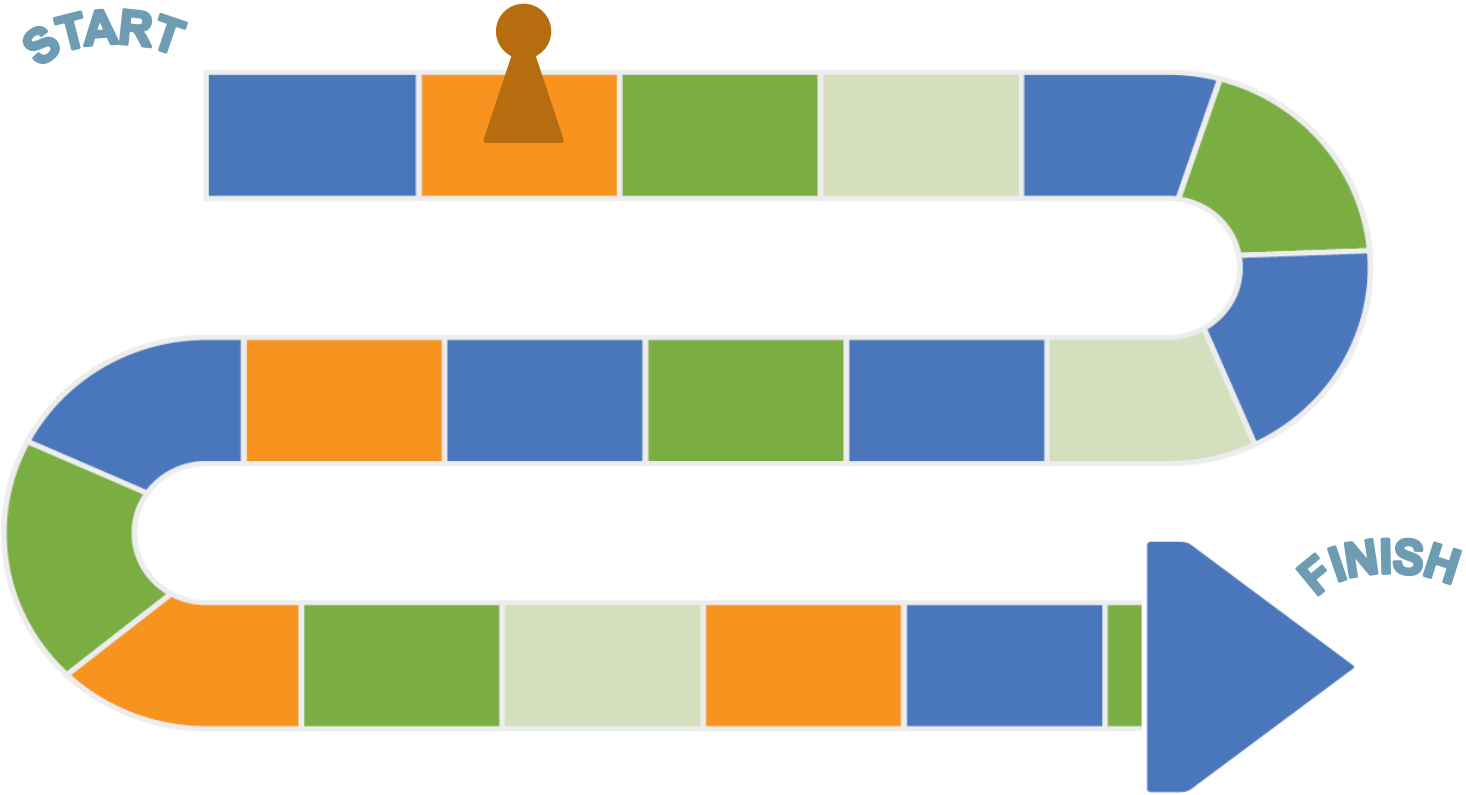


# The path to insights: researchers

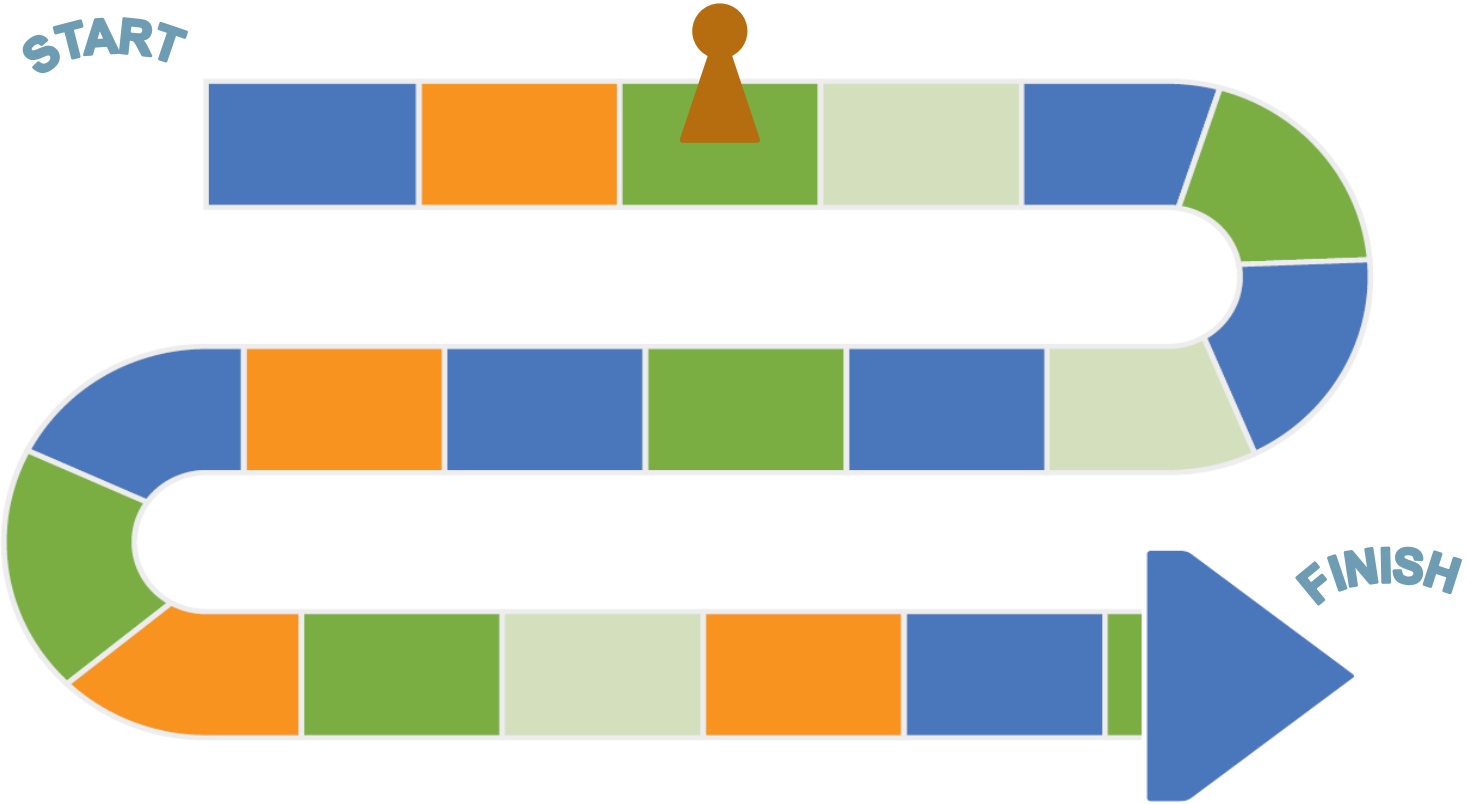


# The path to insights: researchers

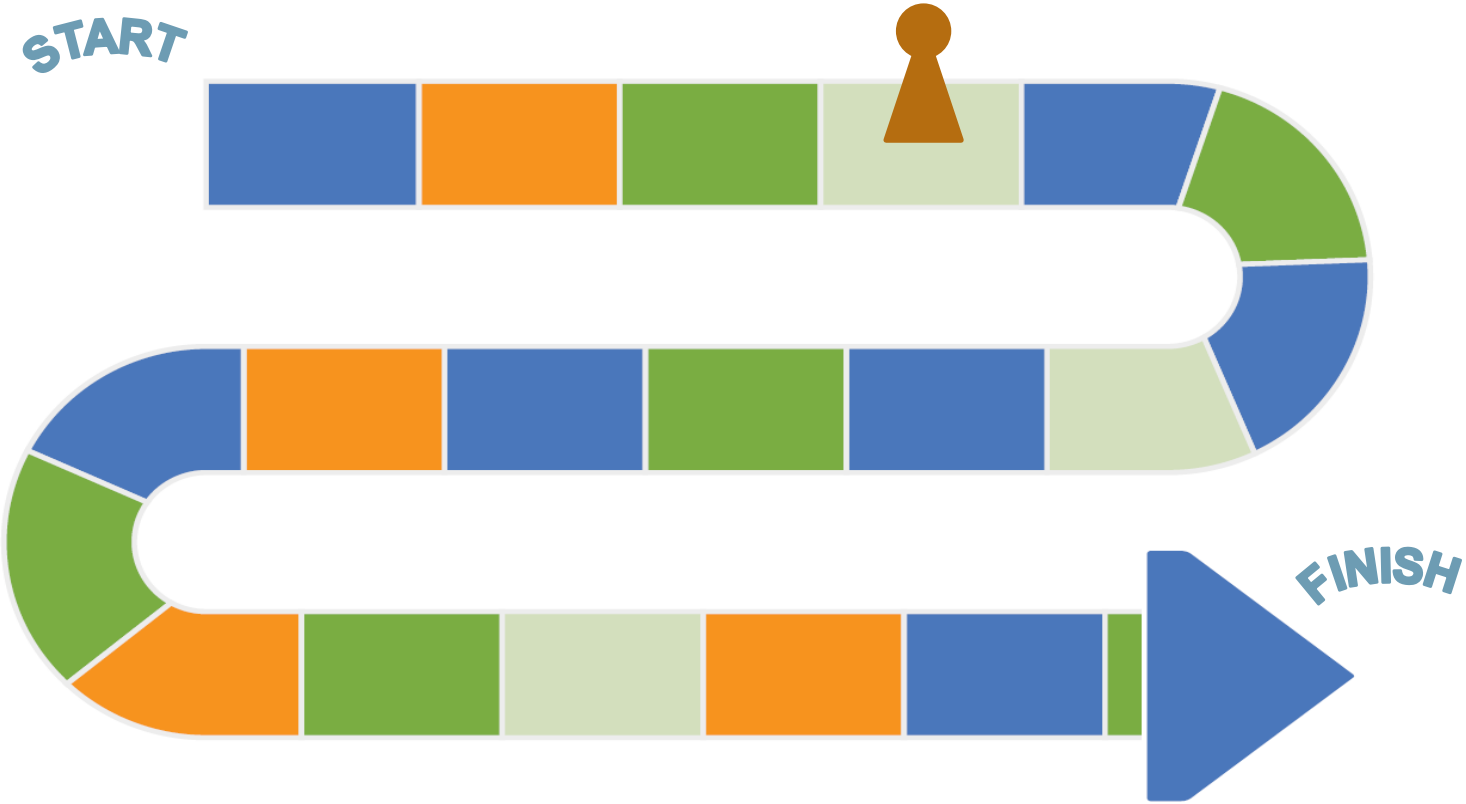
# The path to insights: researchers



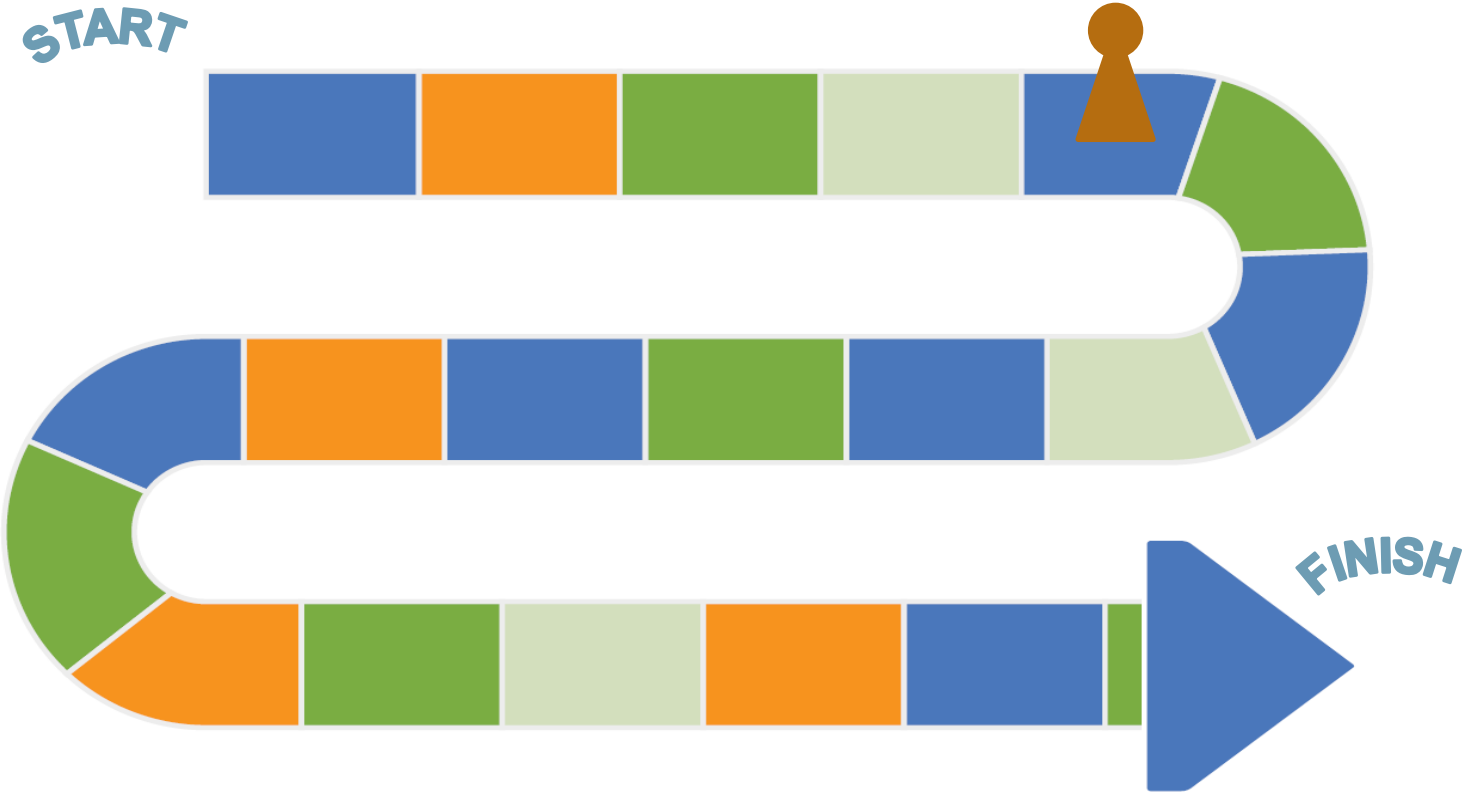
# The path to insights: researchers



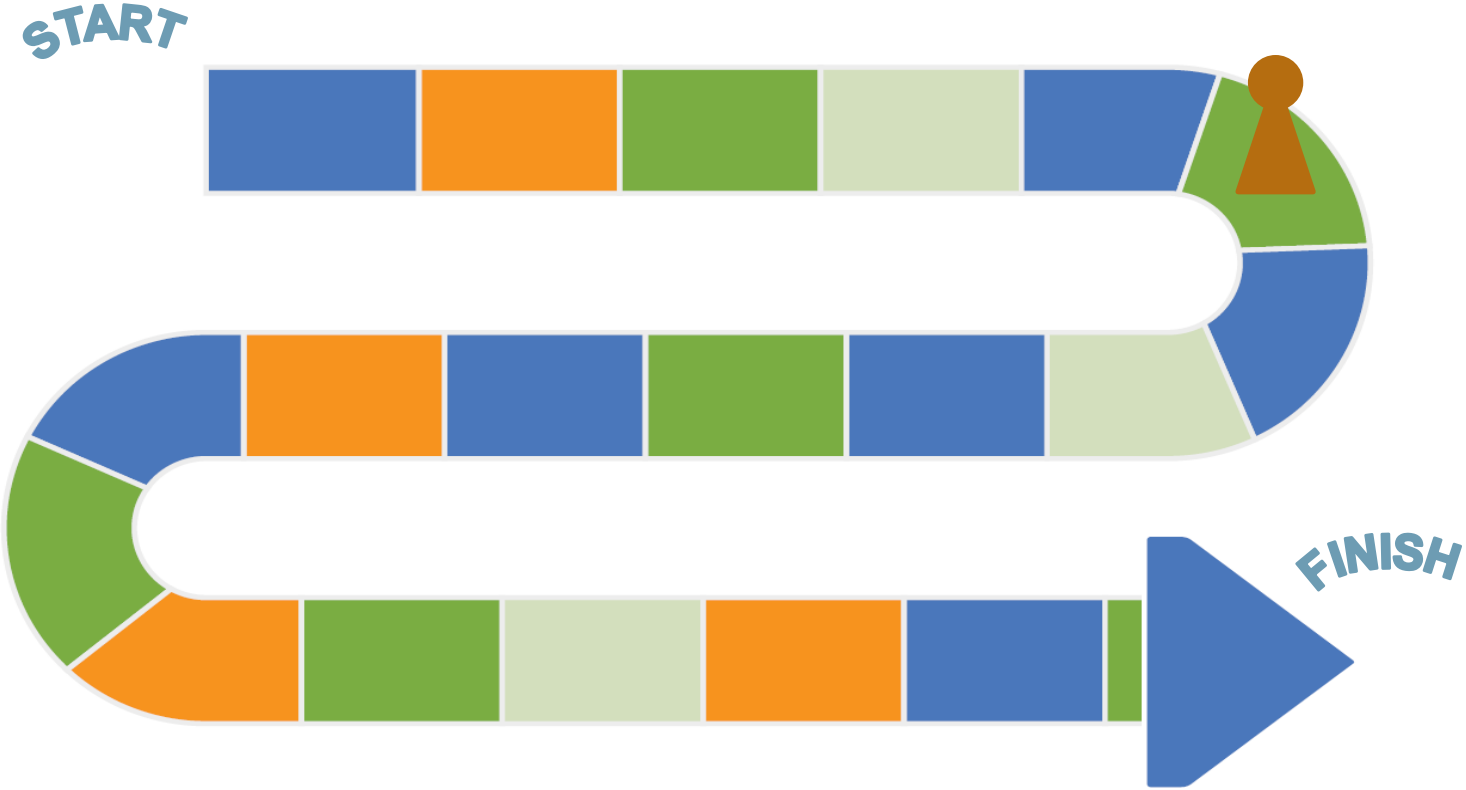
# The path to insights: researchers



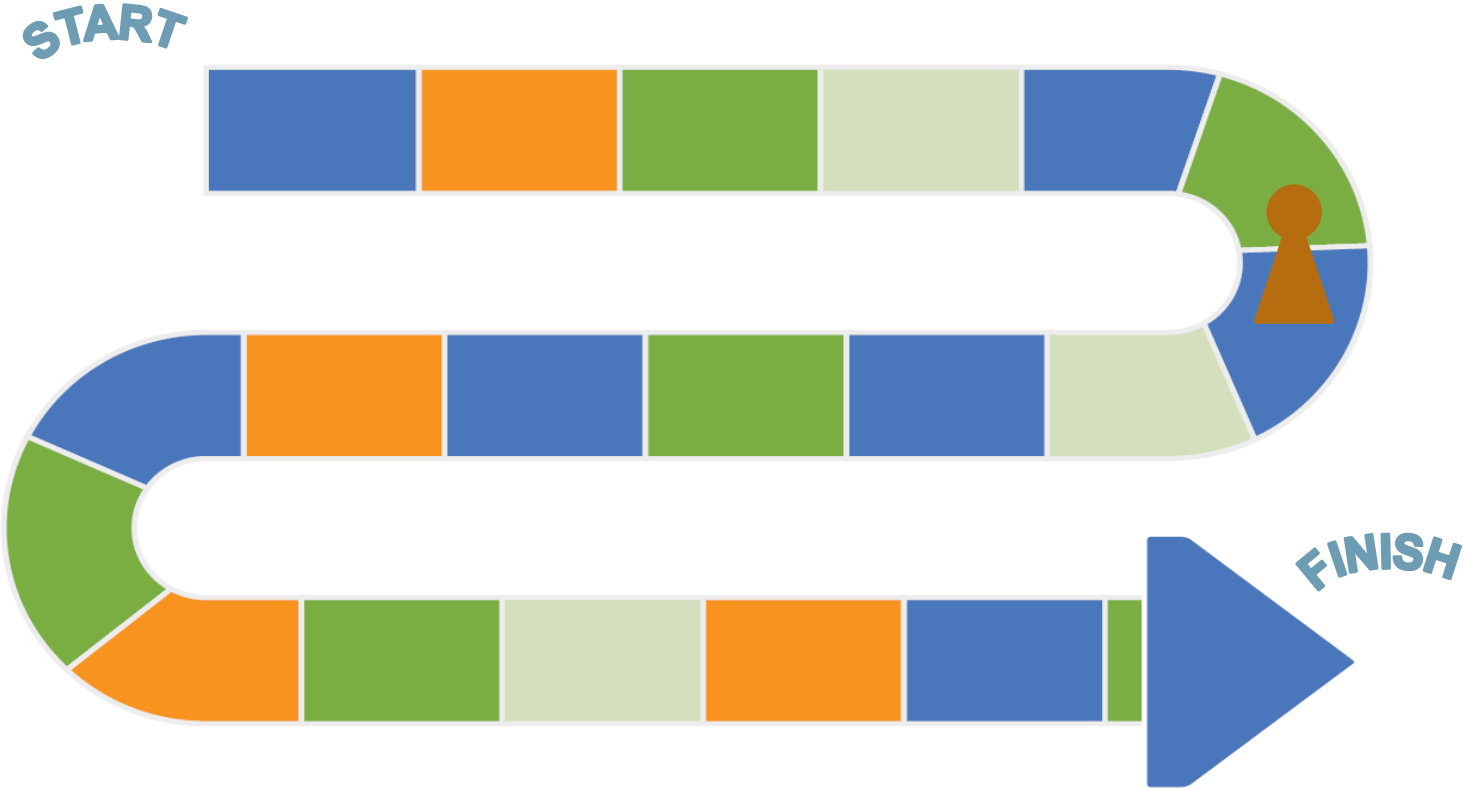
# The path to insights: researchers



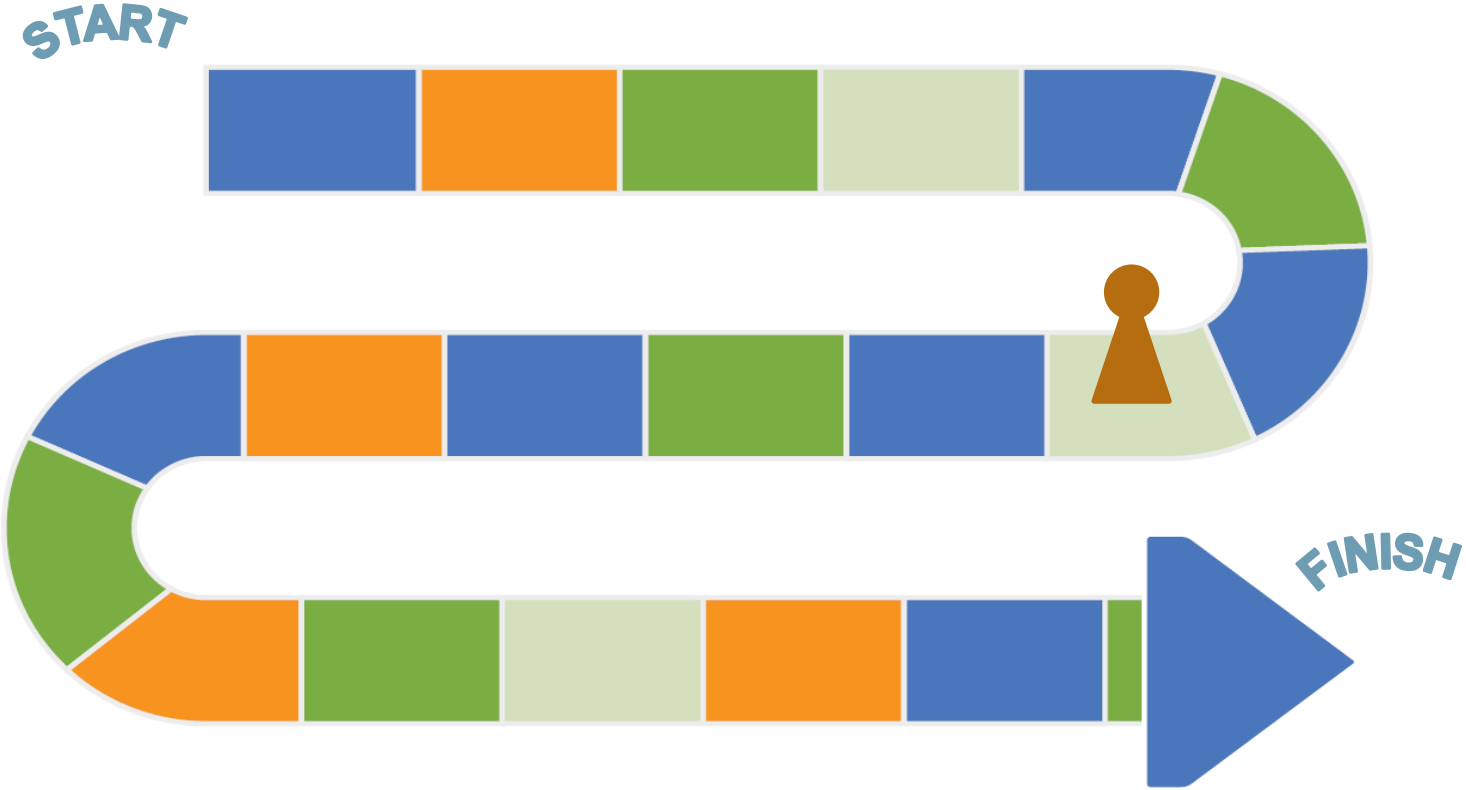
# The path to insights: researchers



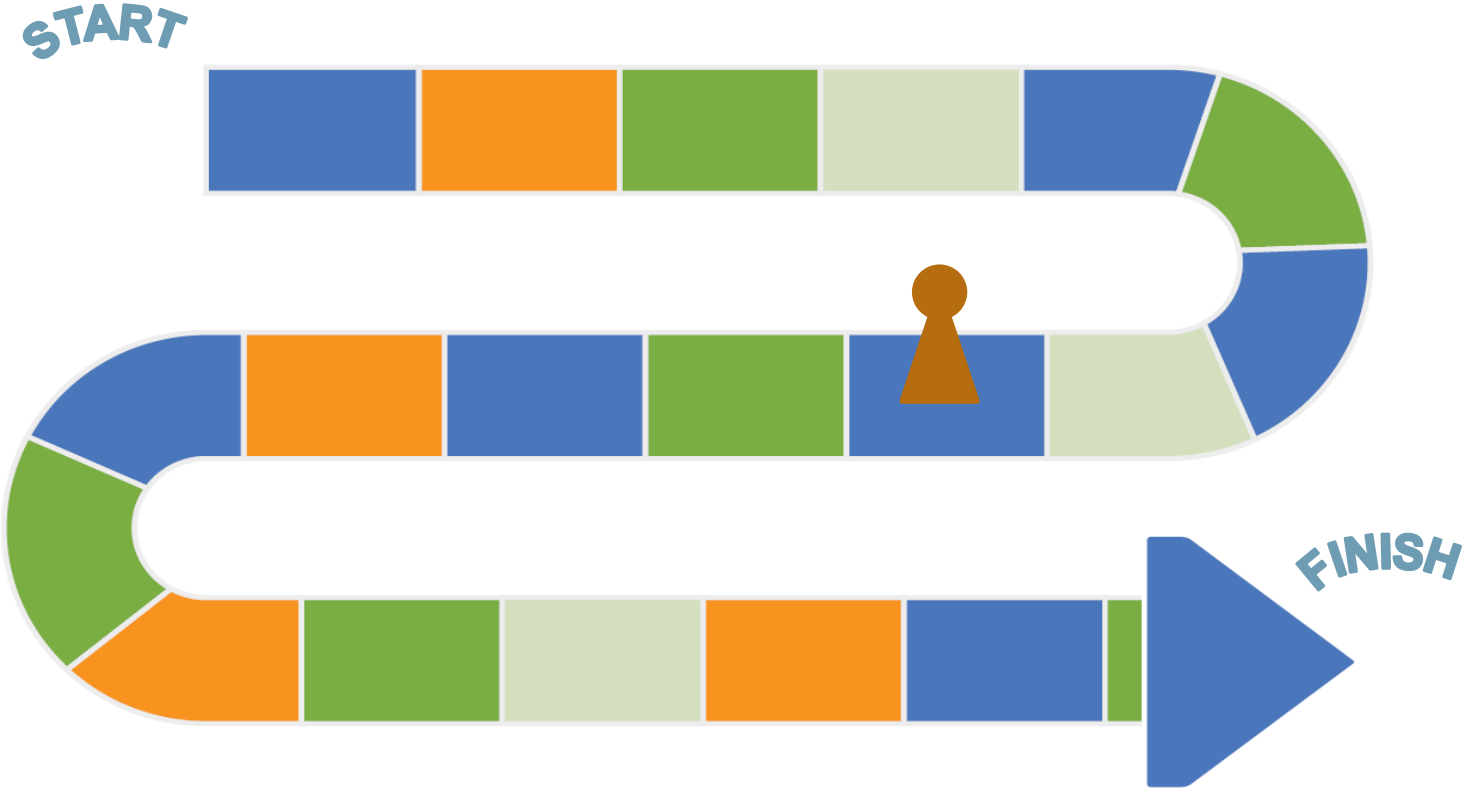
# The path to insights: researchers



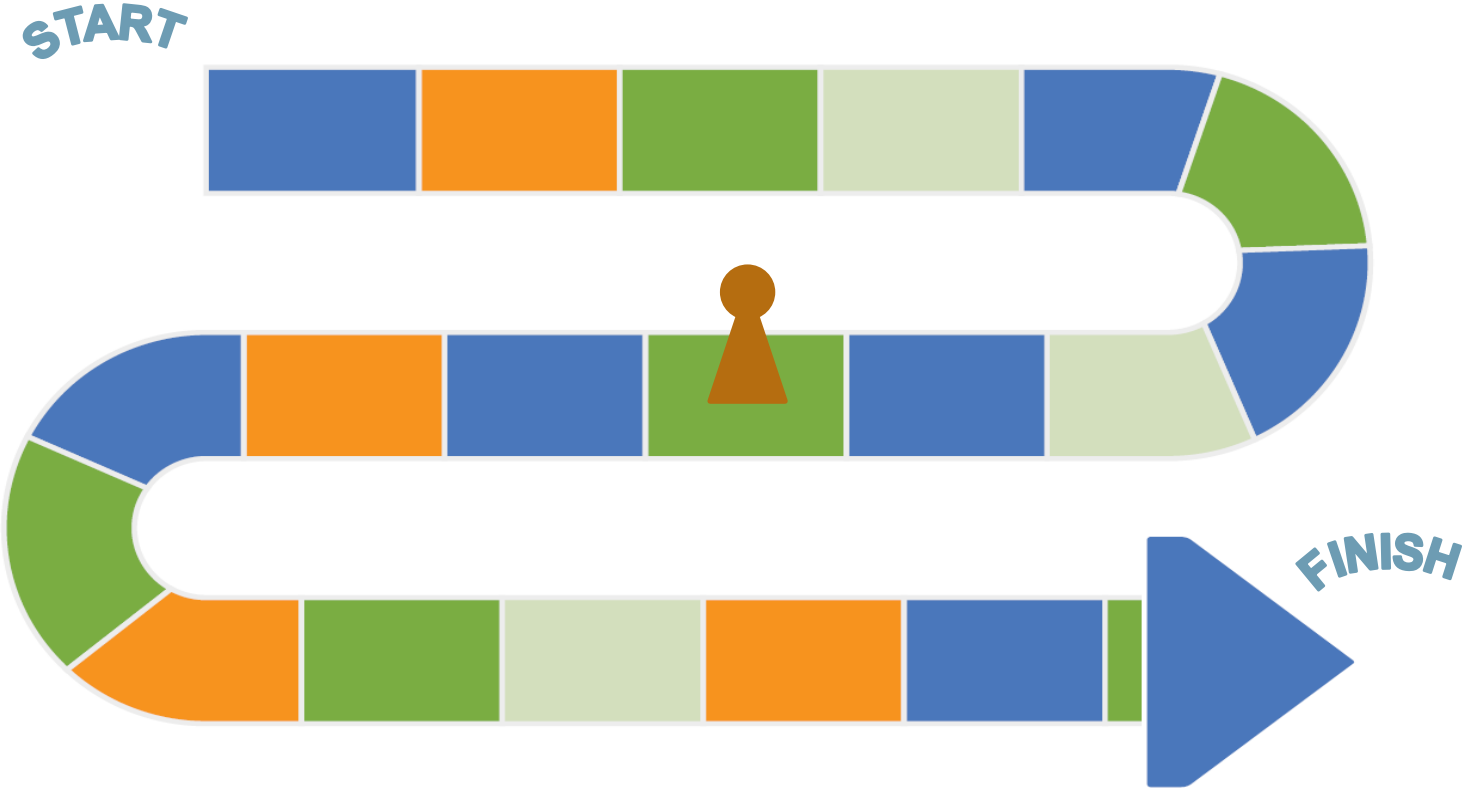
# The path to insights: researchers



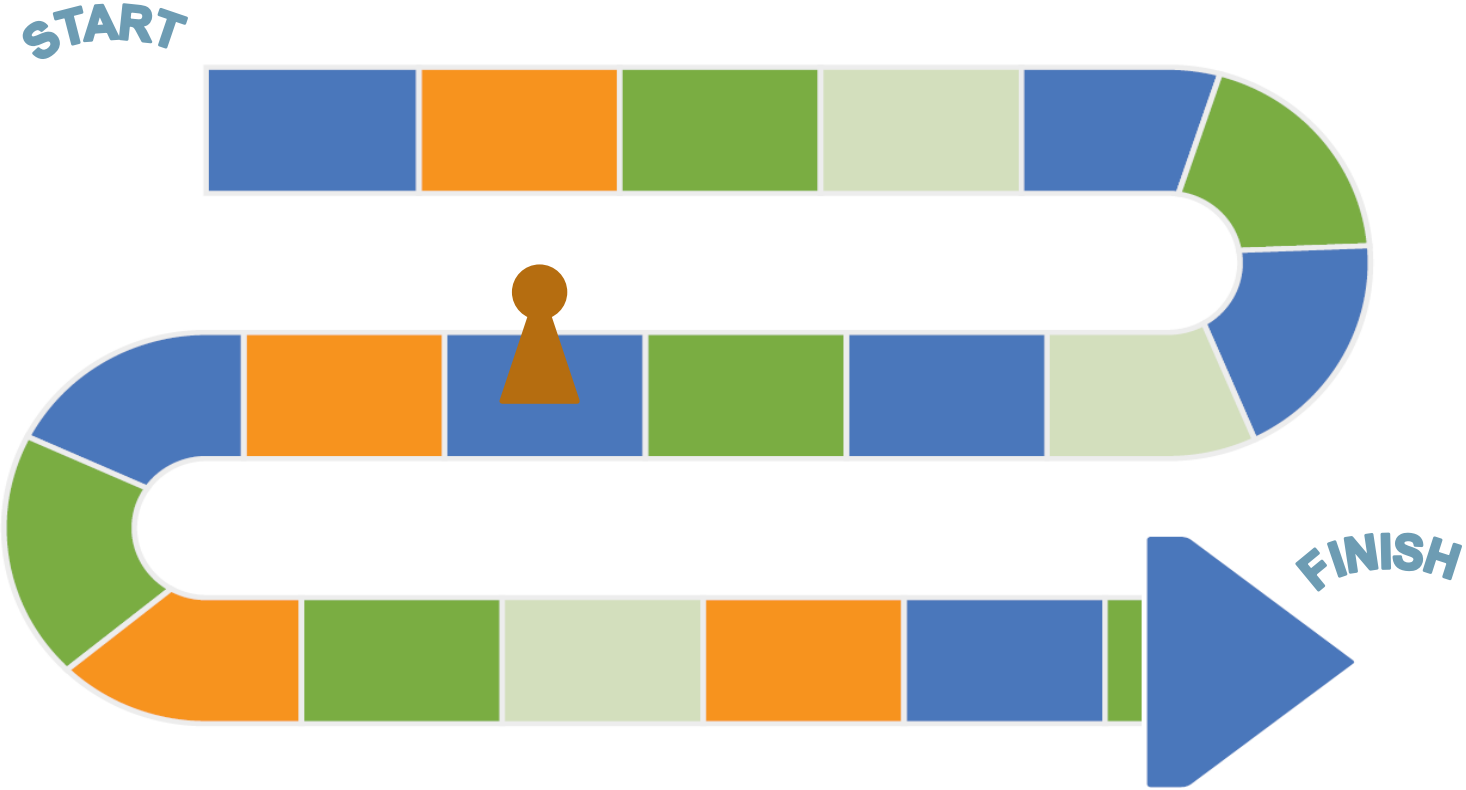
# The path to insights: researchers



# The path to insights: researchers

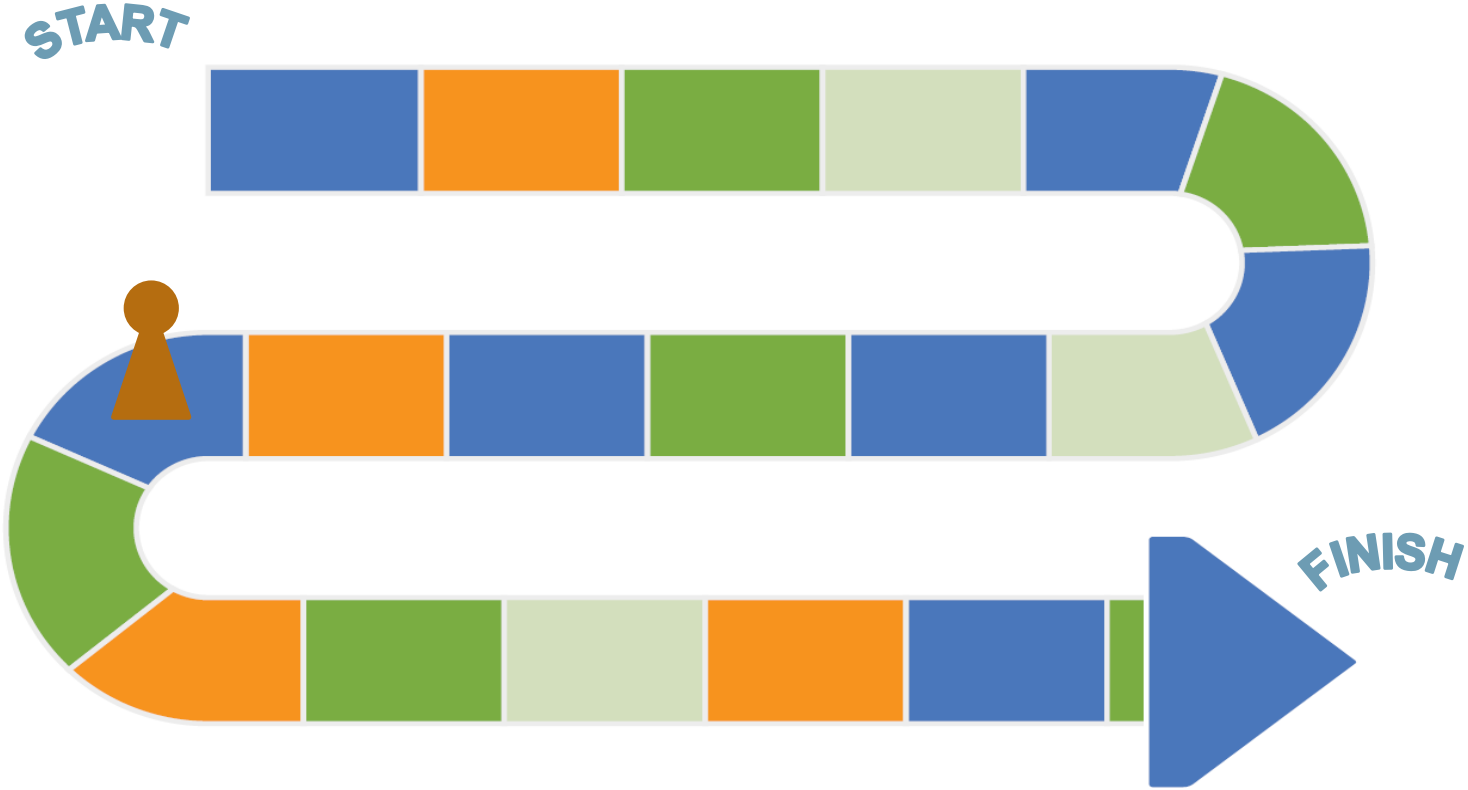


# The path to insights: researchers

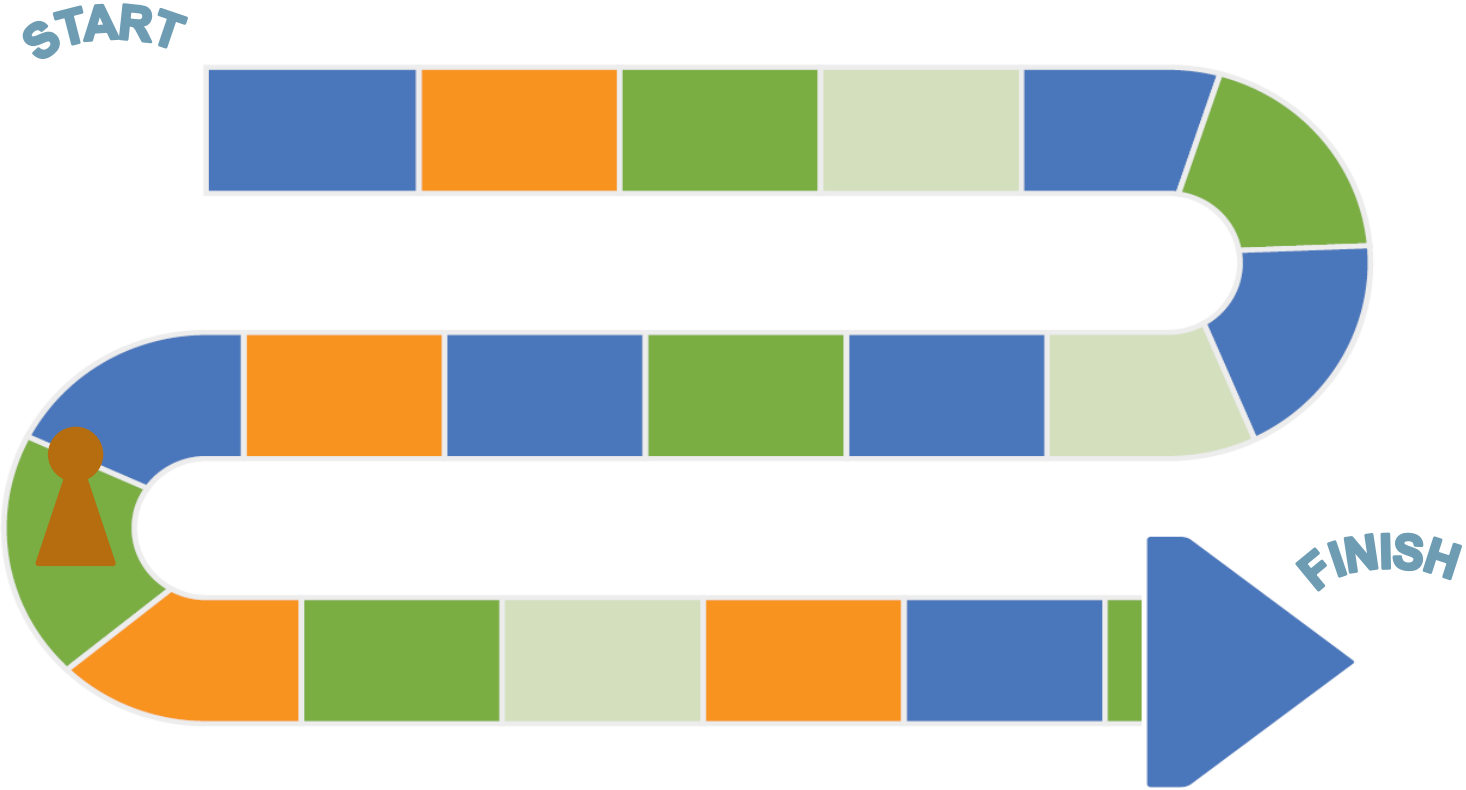


# The path to insights: researchers

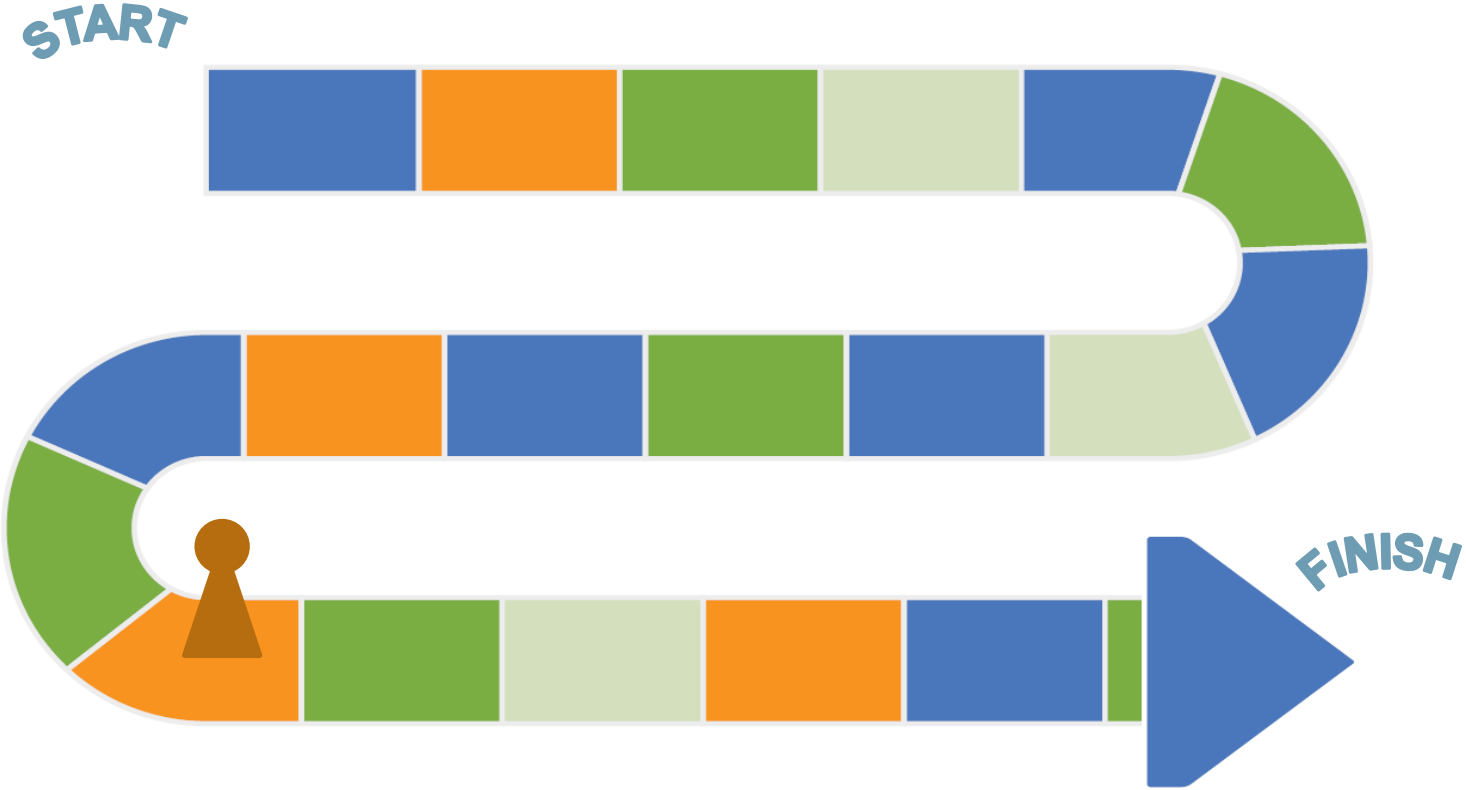
# The path to insights: researchers



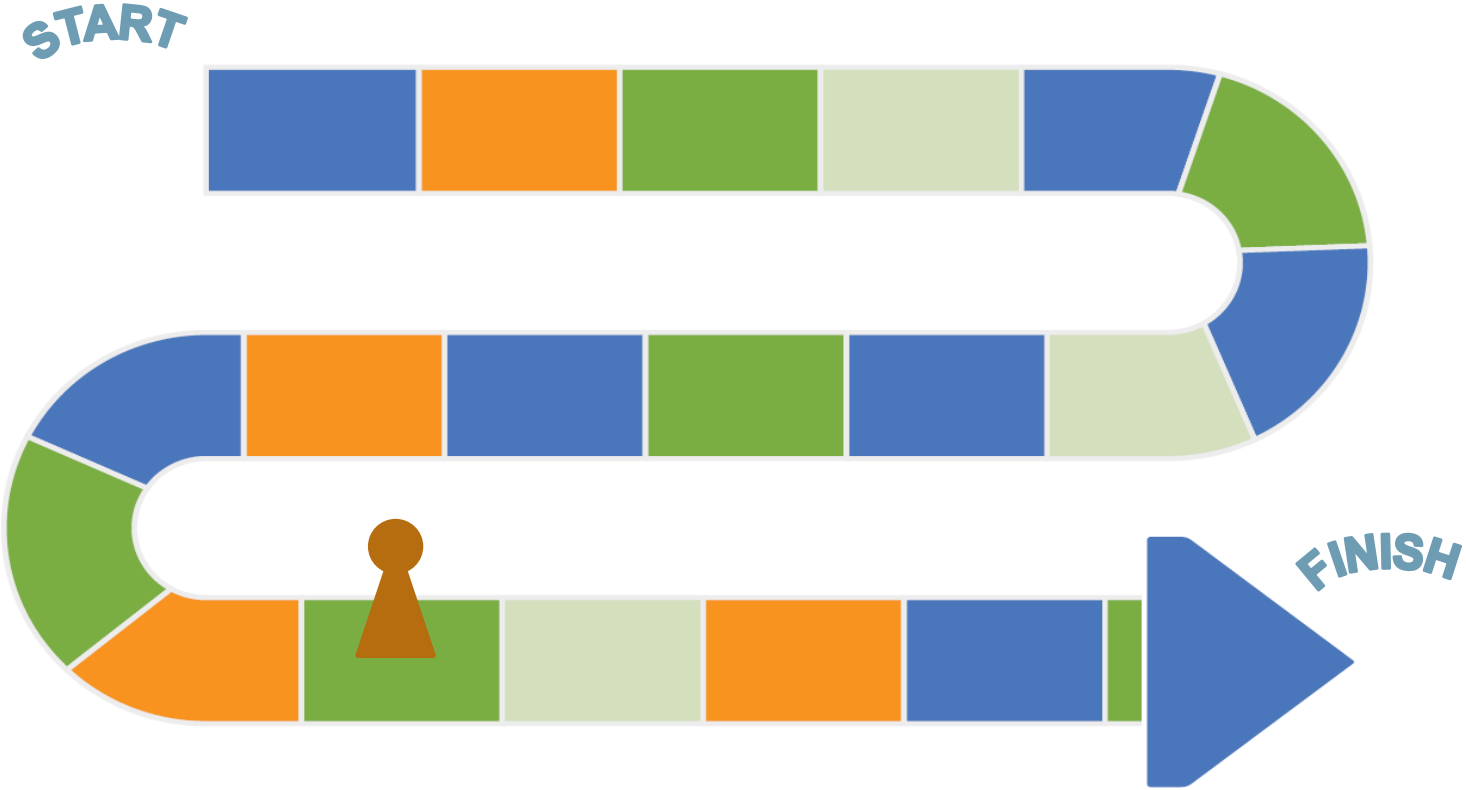
# The path to insights: researchers



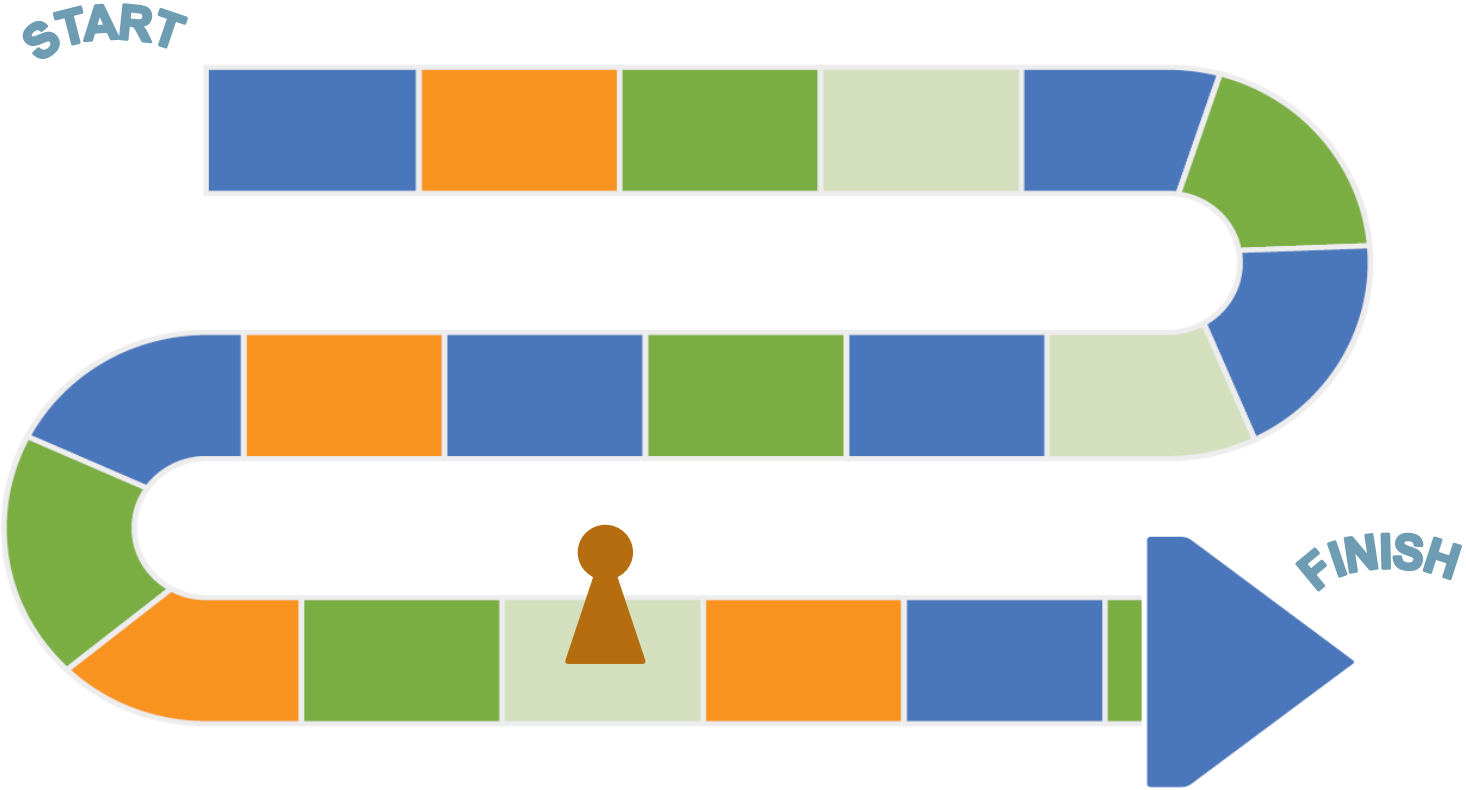
# The path to insights: researchers



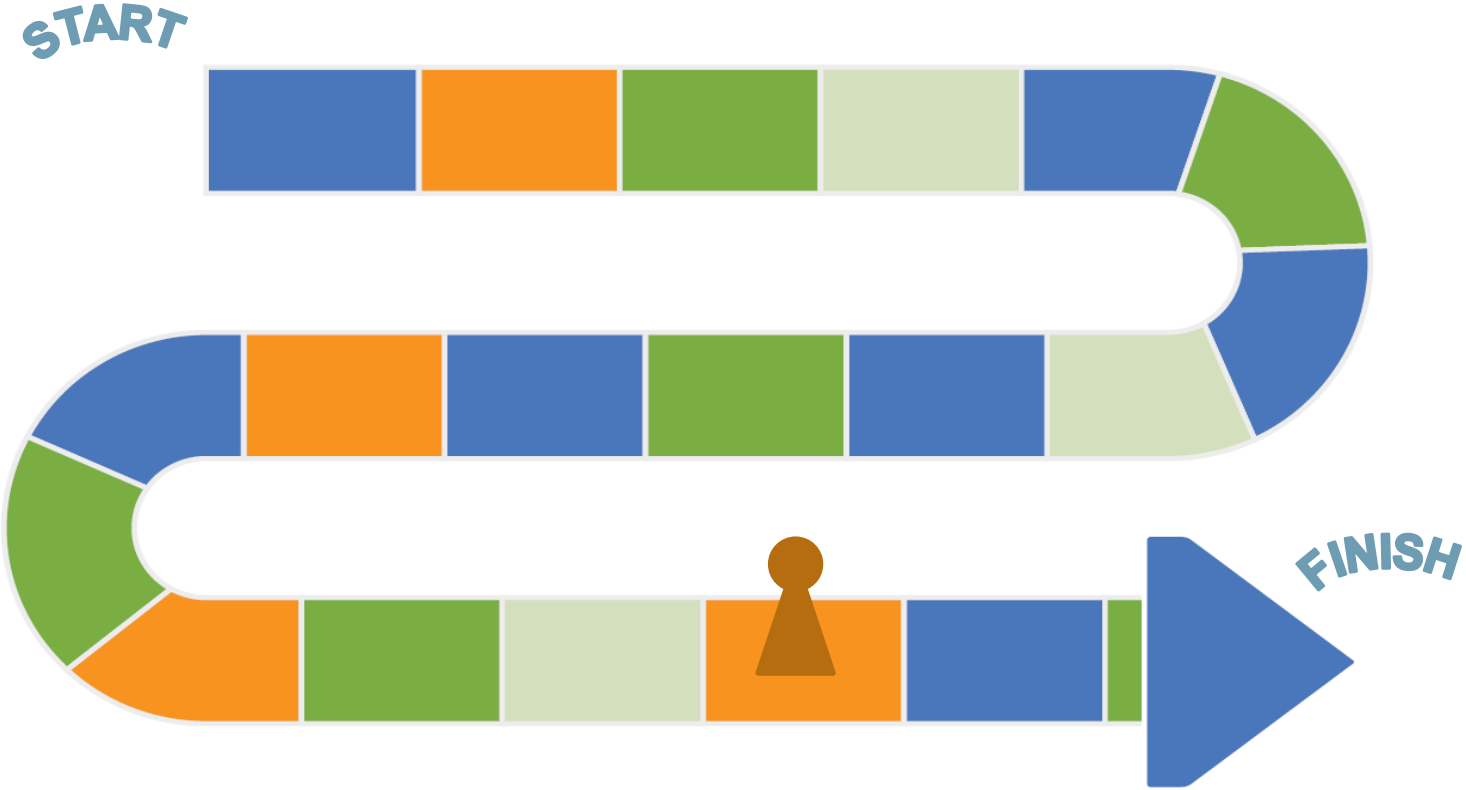
# The path to insights: researchers



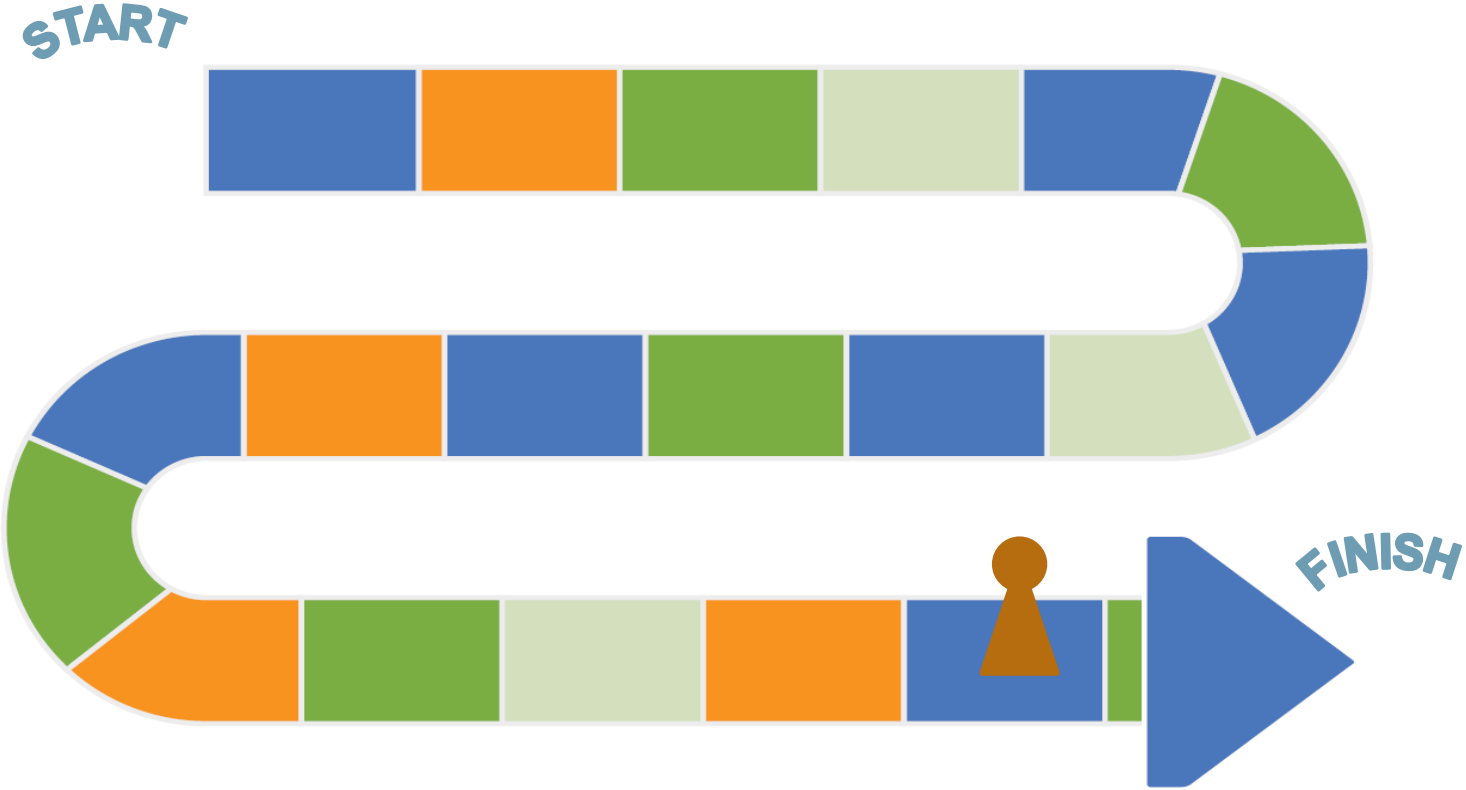
# The path to insights: researchers



# The path to insights: researchers

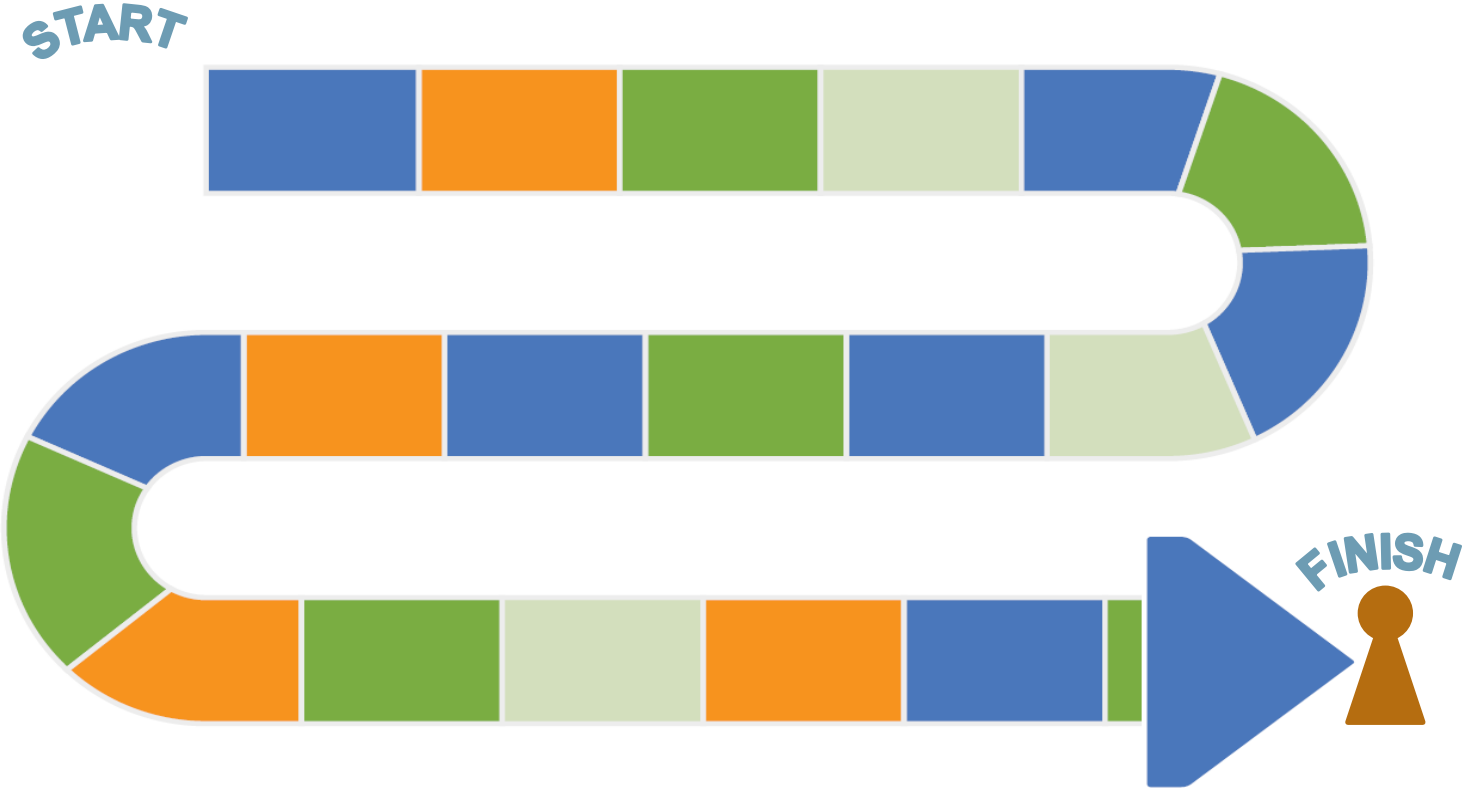


# The path to insights: researchers



# The path to insights: researchers

# The path to insights: researchers



# Key points to good data storytelling

- Know what story you want to tell before you create your final graphics
- Know your audience and tailor your story to them

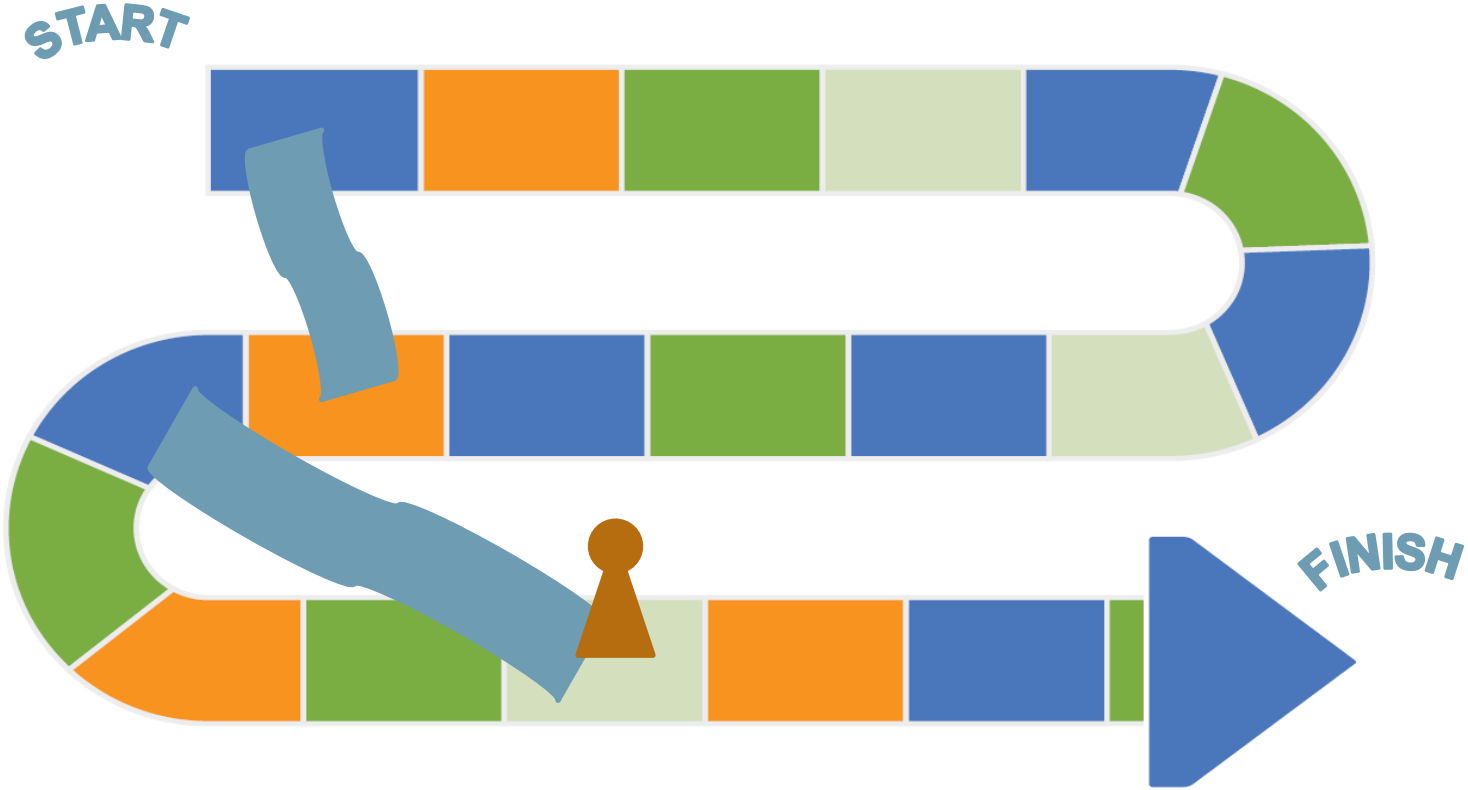
# The path to insights: stakeholders

# The path to insights: stakeholders

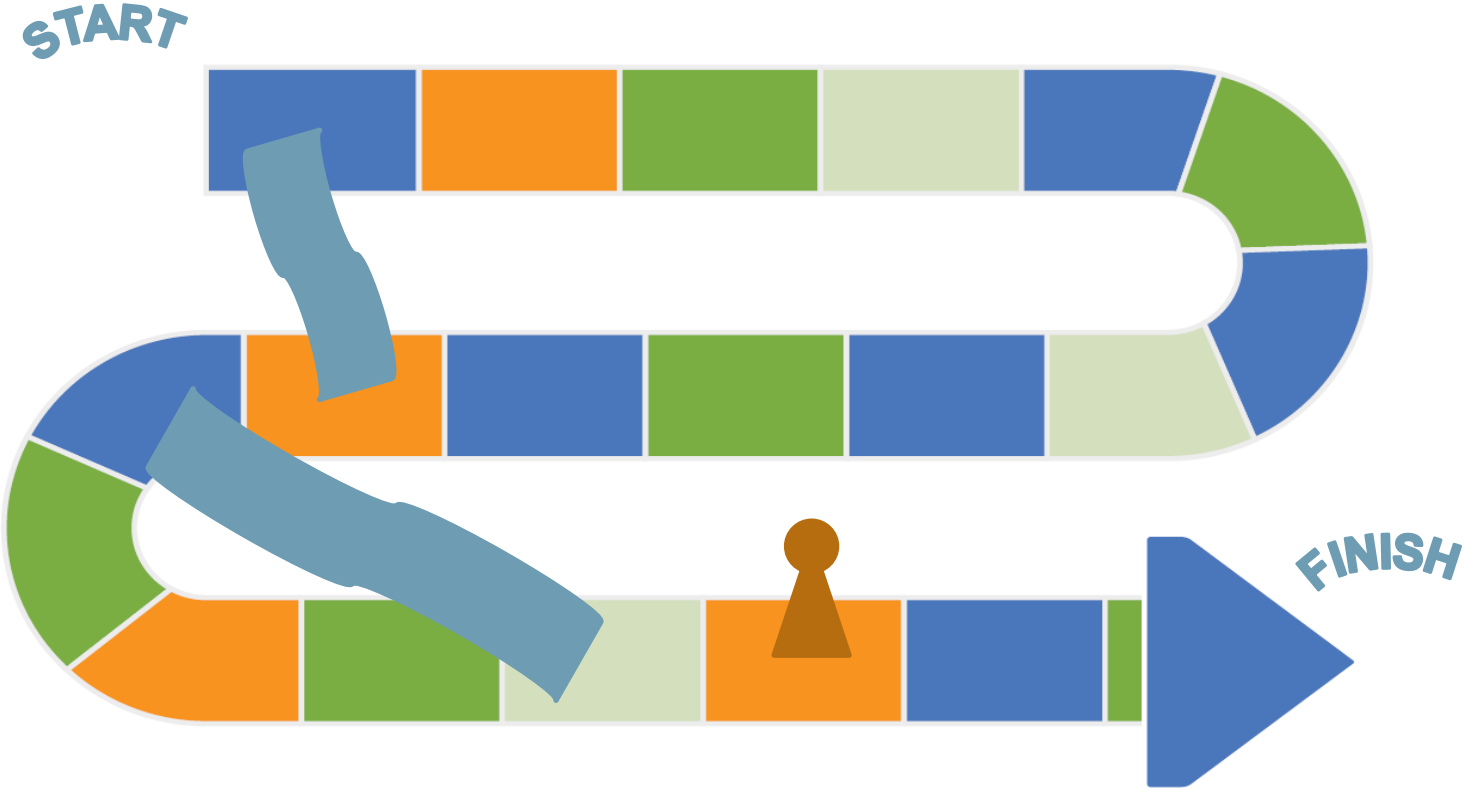
# The path to insights: stakeholders

# The path to insights: stakeholders

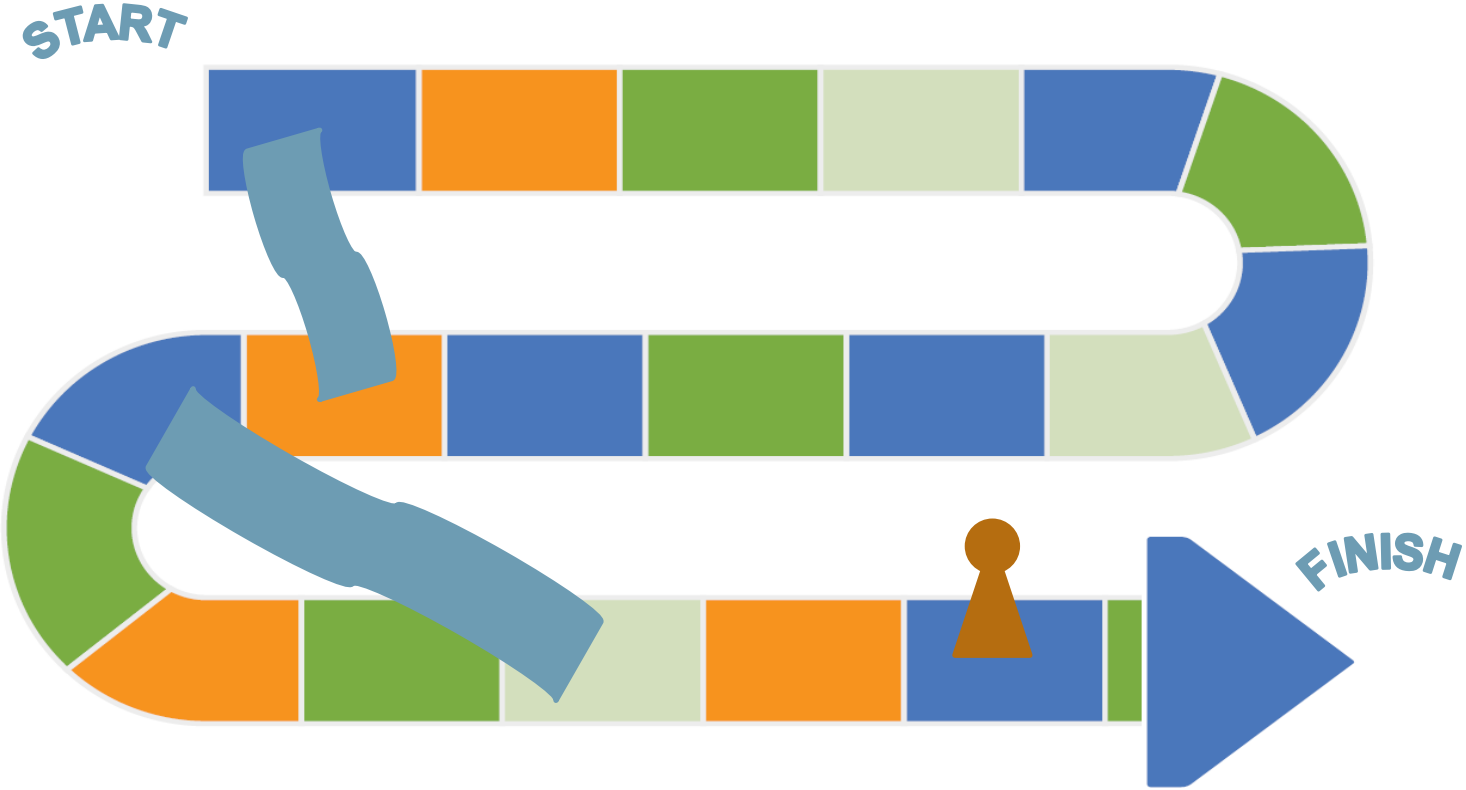
# The path to insights: stakeholders



# The path to insights: stakeholders



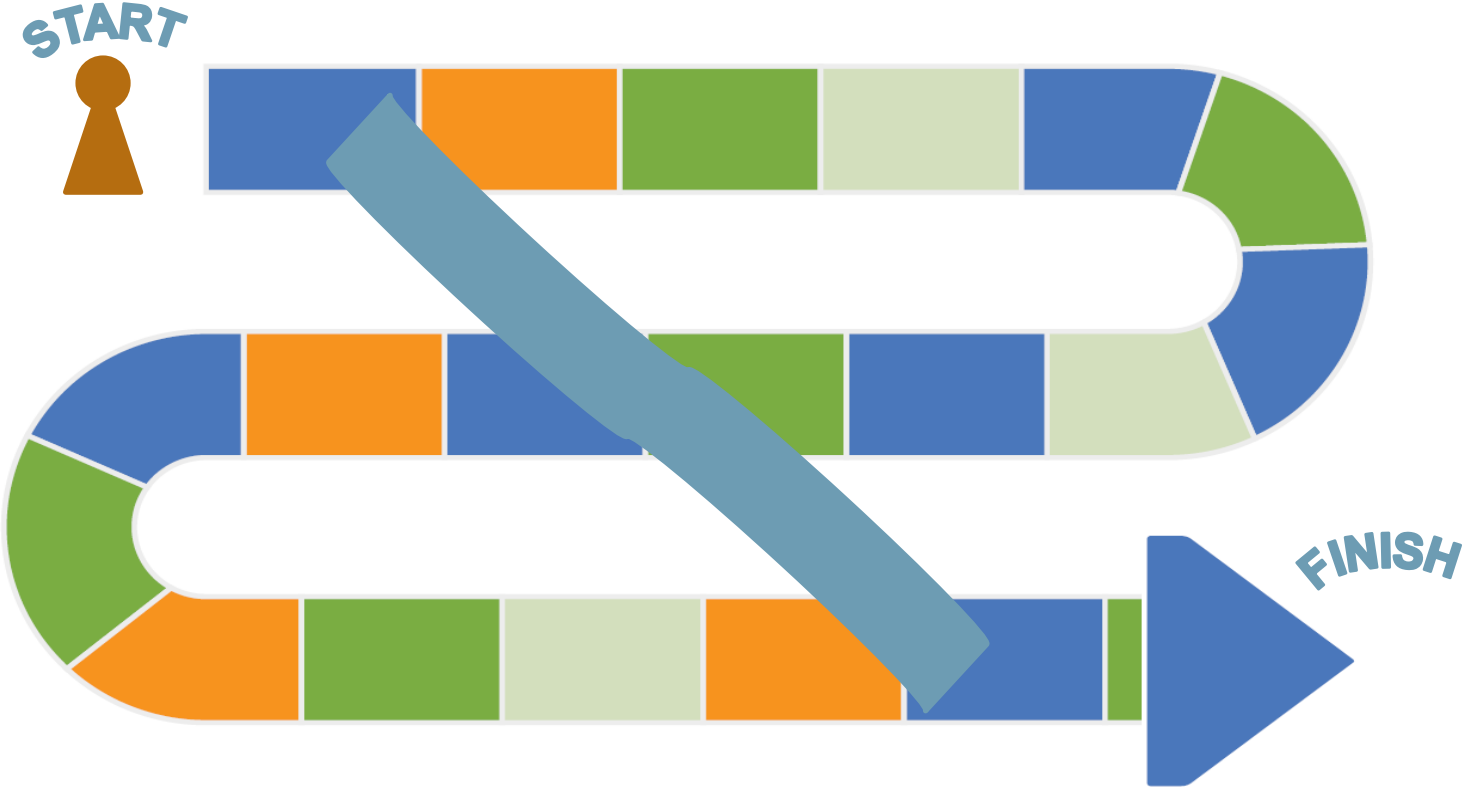
# The path to insights: stakeholders



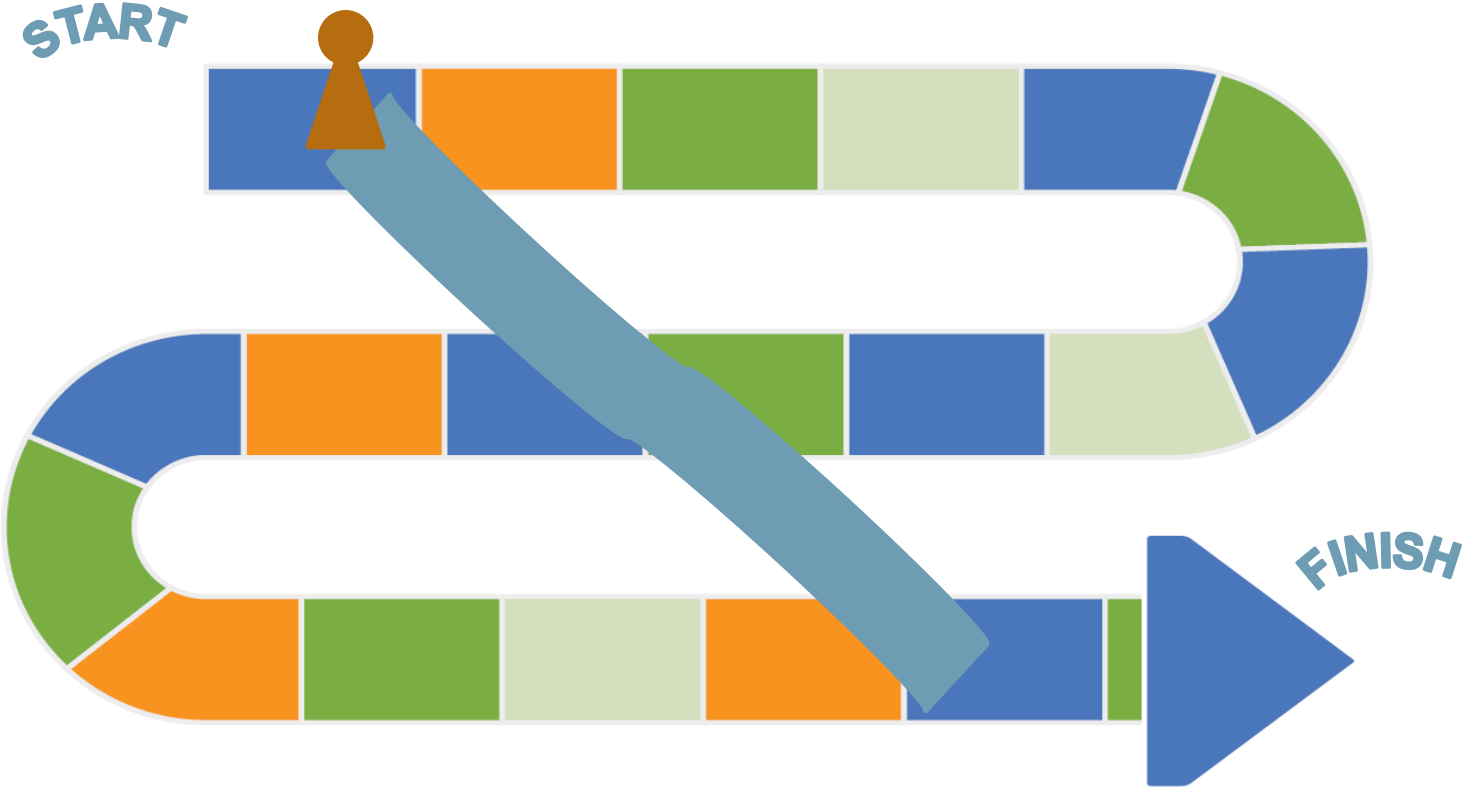
# The path to insights: stakeholders

# The path to insights: stakeholders

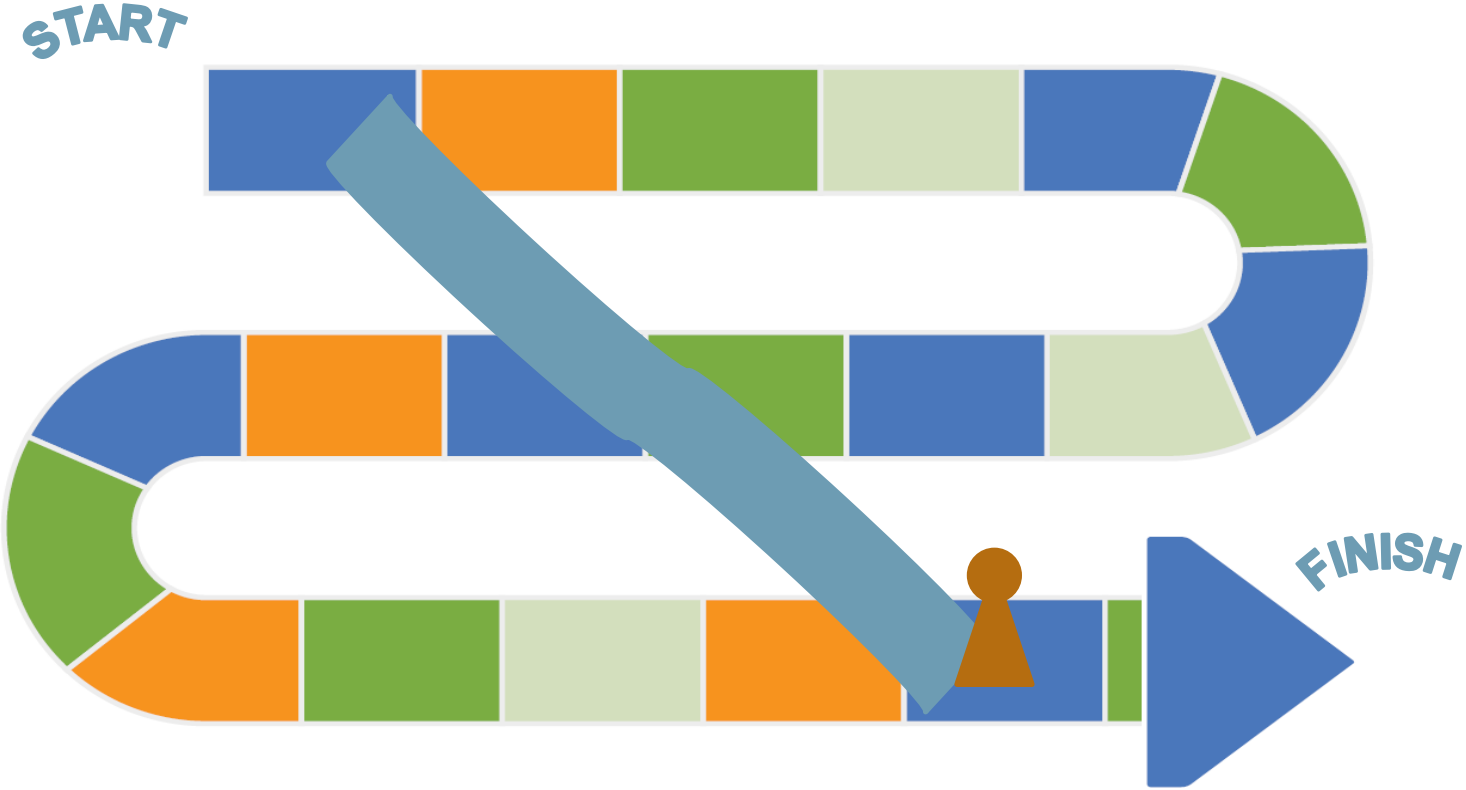
# The path to insights: executives



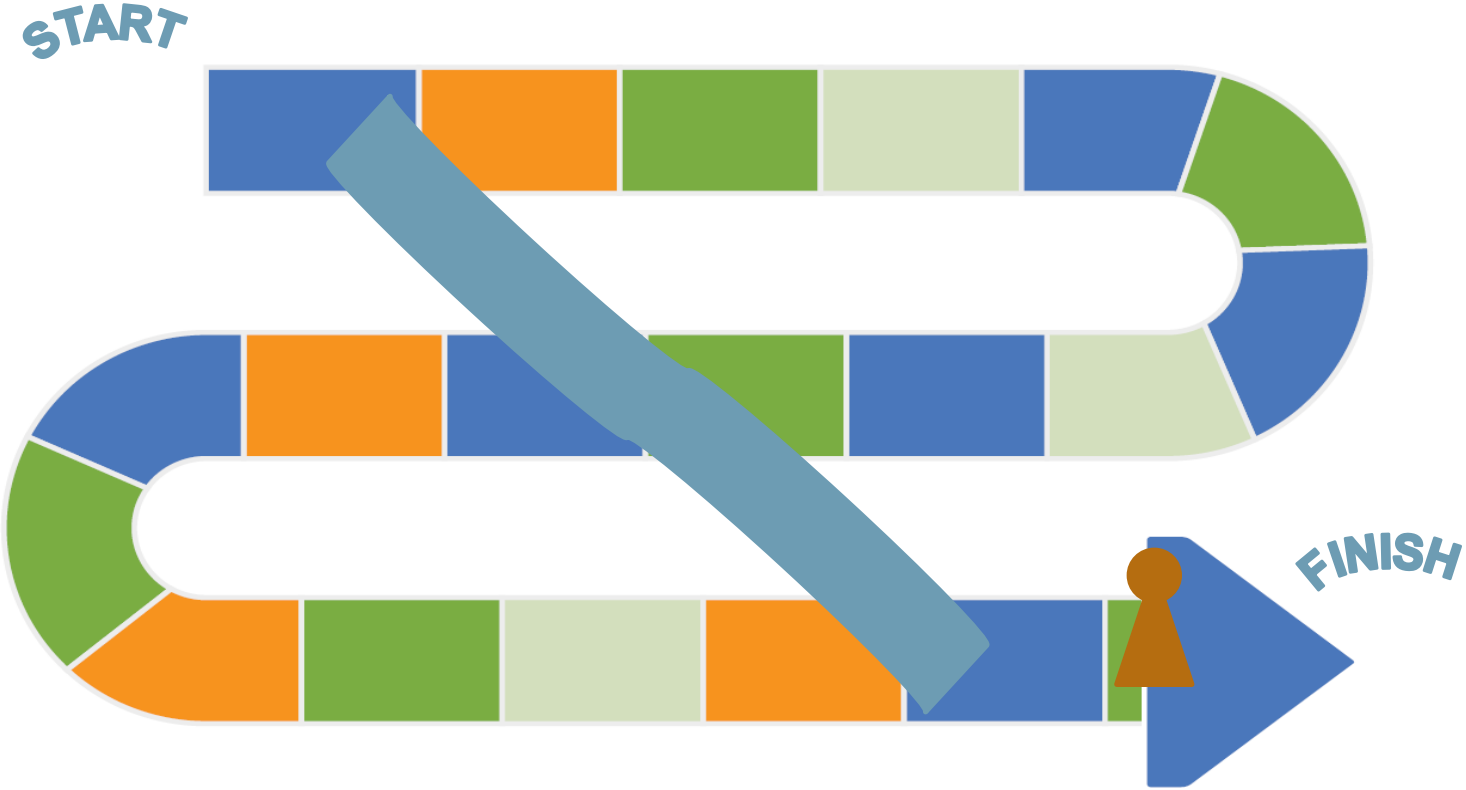
# The path to insights: executives



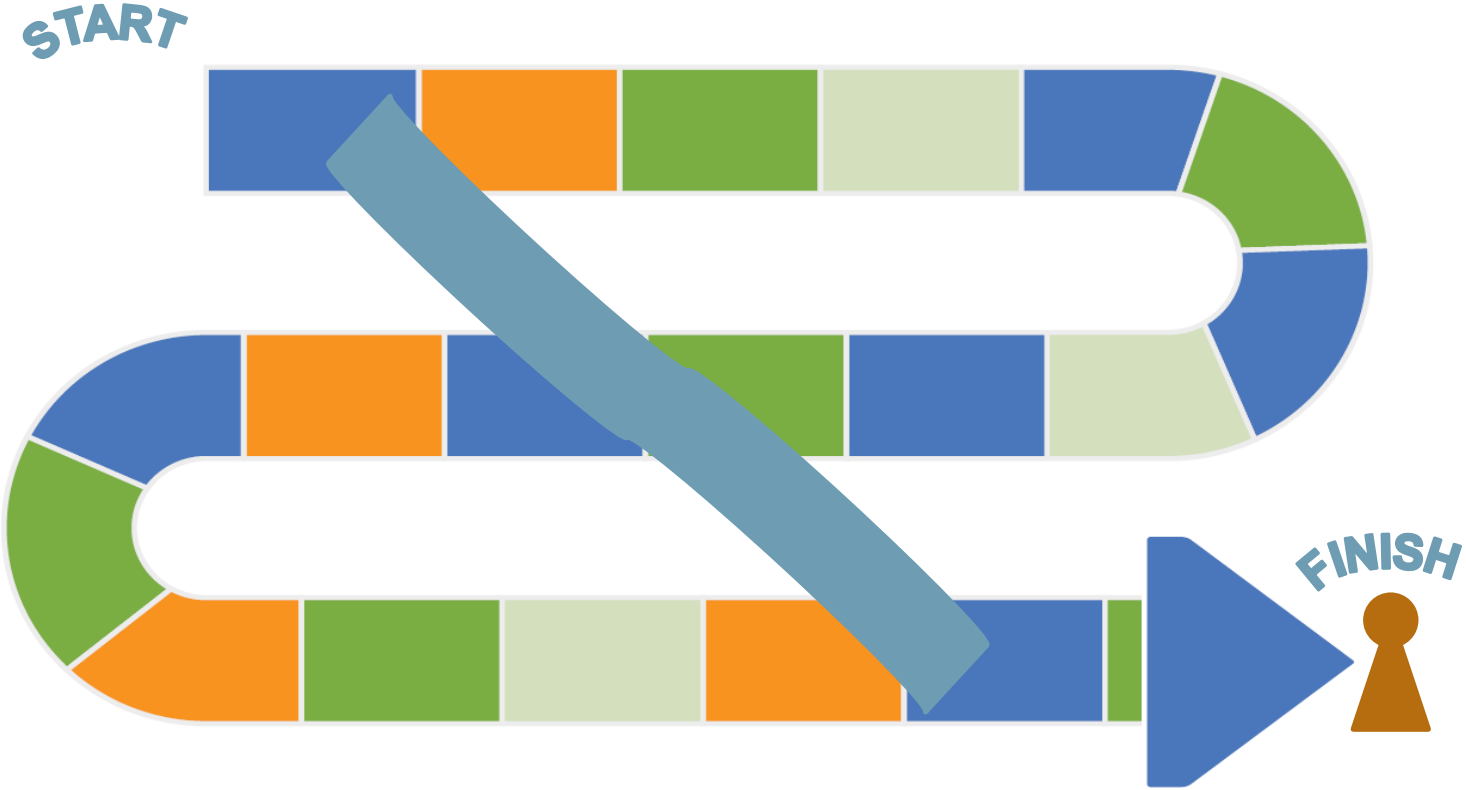
# The path to insights: executives



# The path to insights: executives



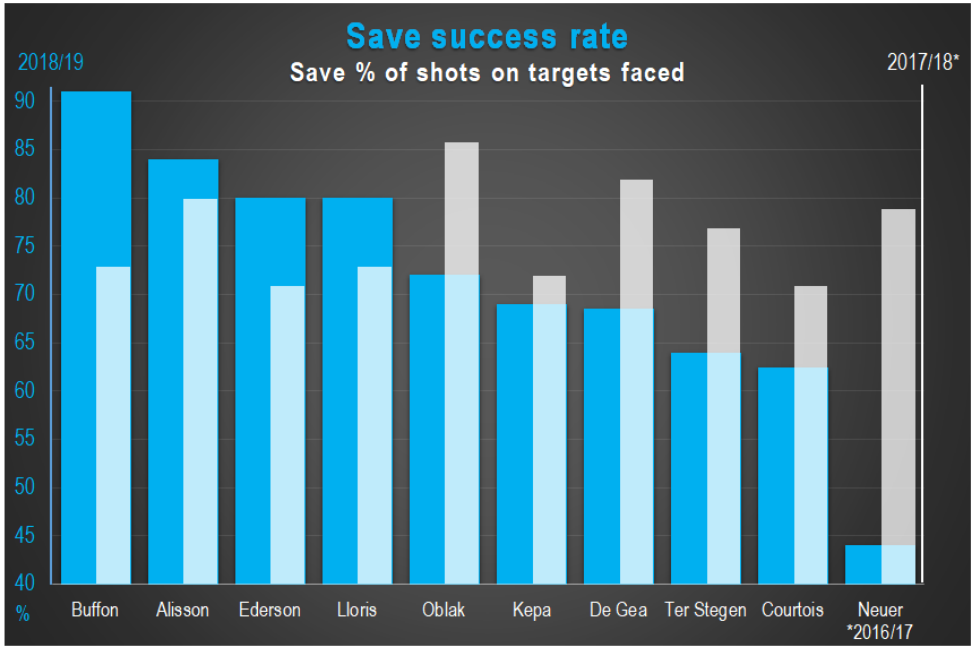
# The path to insights: executives



# Key points to good data storytelling

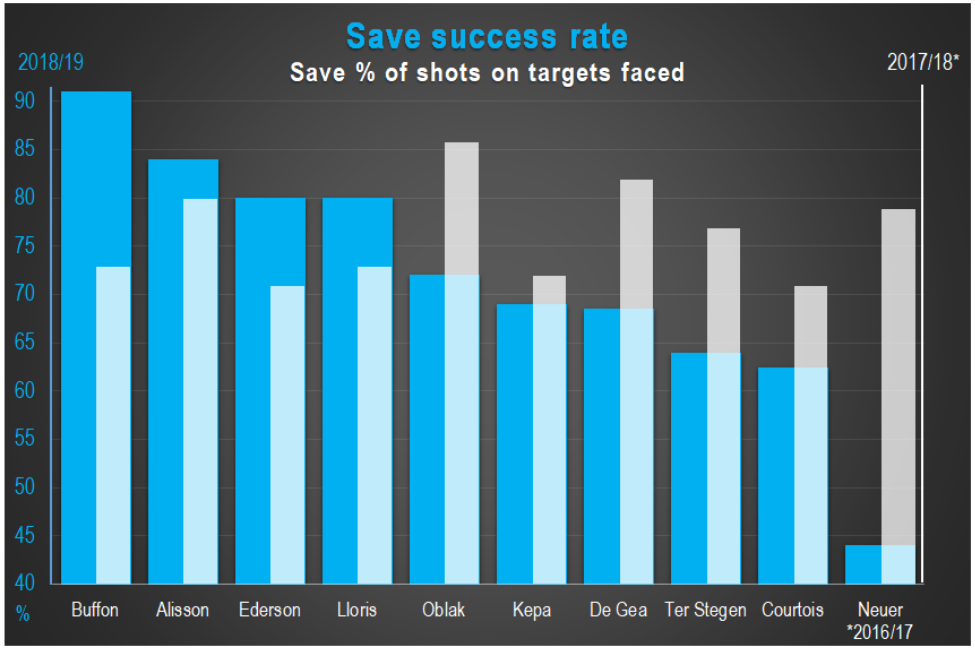
- Know what story you want to tell before you create your final graphics
- Know your audience and tailor your story to them
- Choose graphics wisely

# Discussion graph




Data from WhoScored.com; based on design from Opta and the Guardian

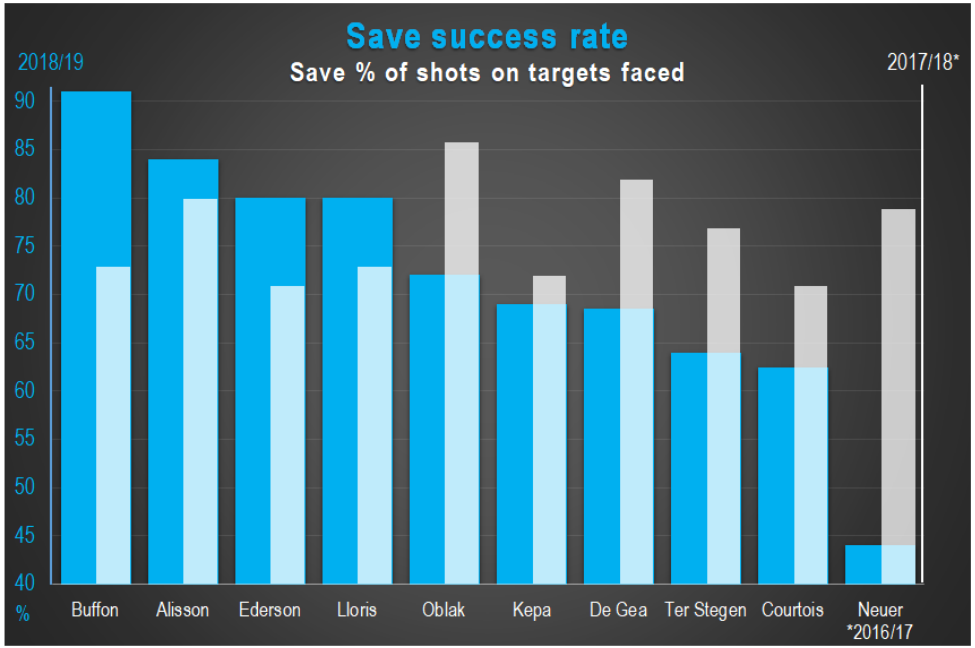
# Question




Data from WhoScored.com; based on design from Opta and the Guardian

*What does the asterisk indicate?*

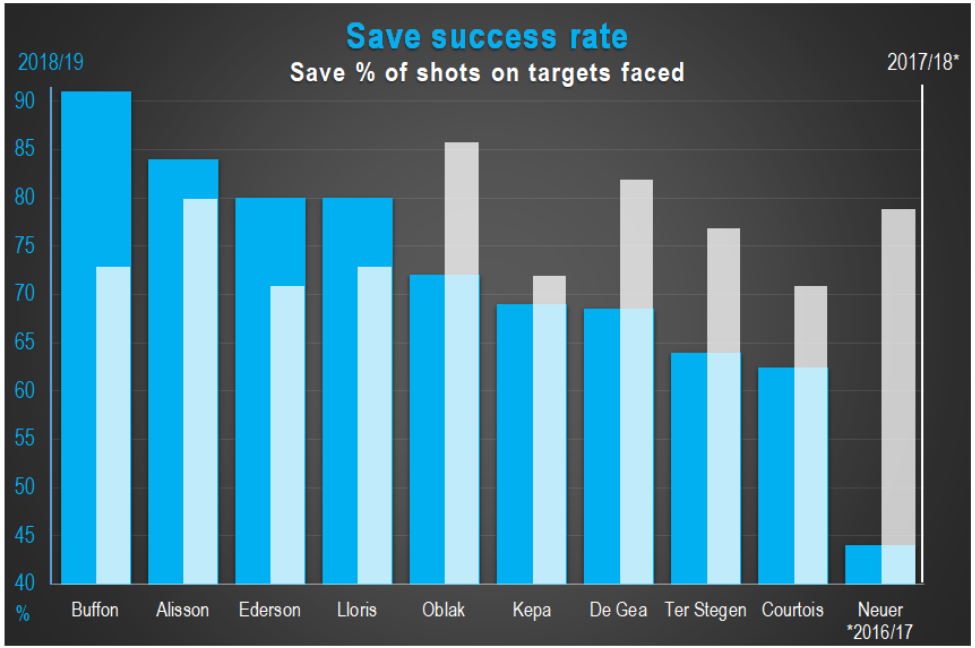
# Question




Data from WhoScored.com; based on design from Opta and the Guardian

*Which goalie's save rate was the most consistent year to year?*

# Question

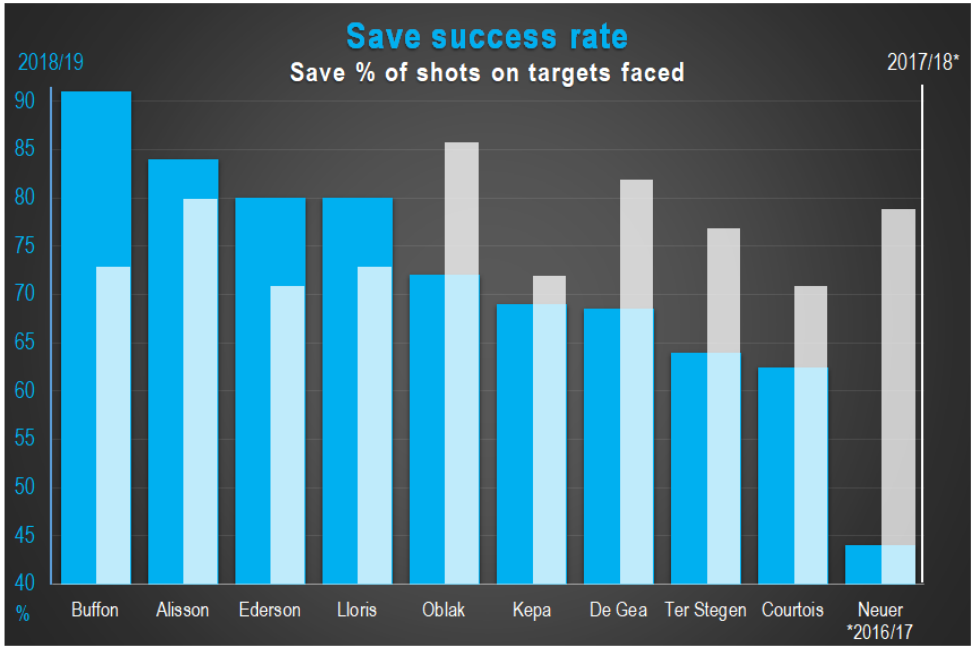



Data from WhoScored.com; based on design from Opta and the Guardian

*How many goalies' save rates improved over time?*

*How many goalies' save rates declined over time?*

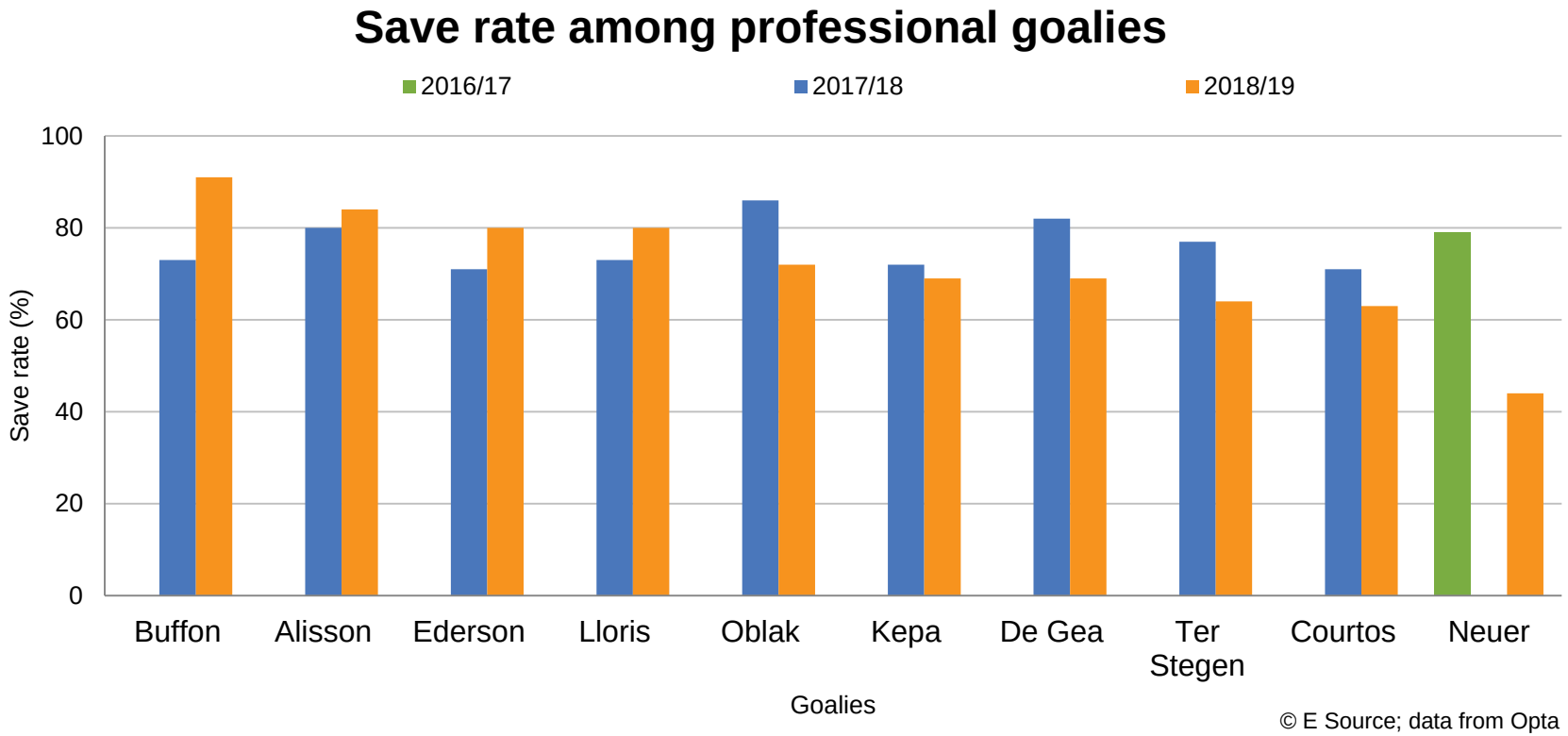
# Question




Data from WhoScored.com; based on design from Opta and the Guardian

*Whose save rate (any season) was roughly half of Buffon's 2018/19 save rate?*

# Better graphics example

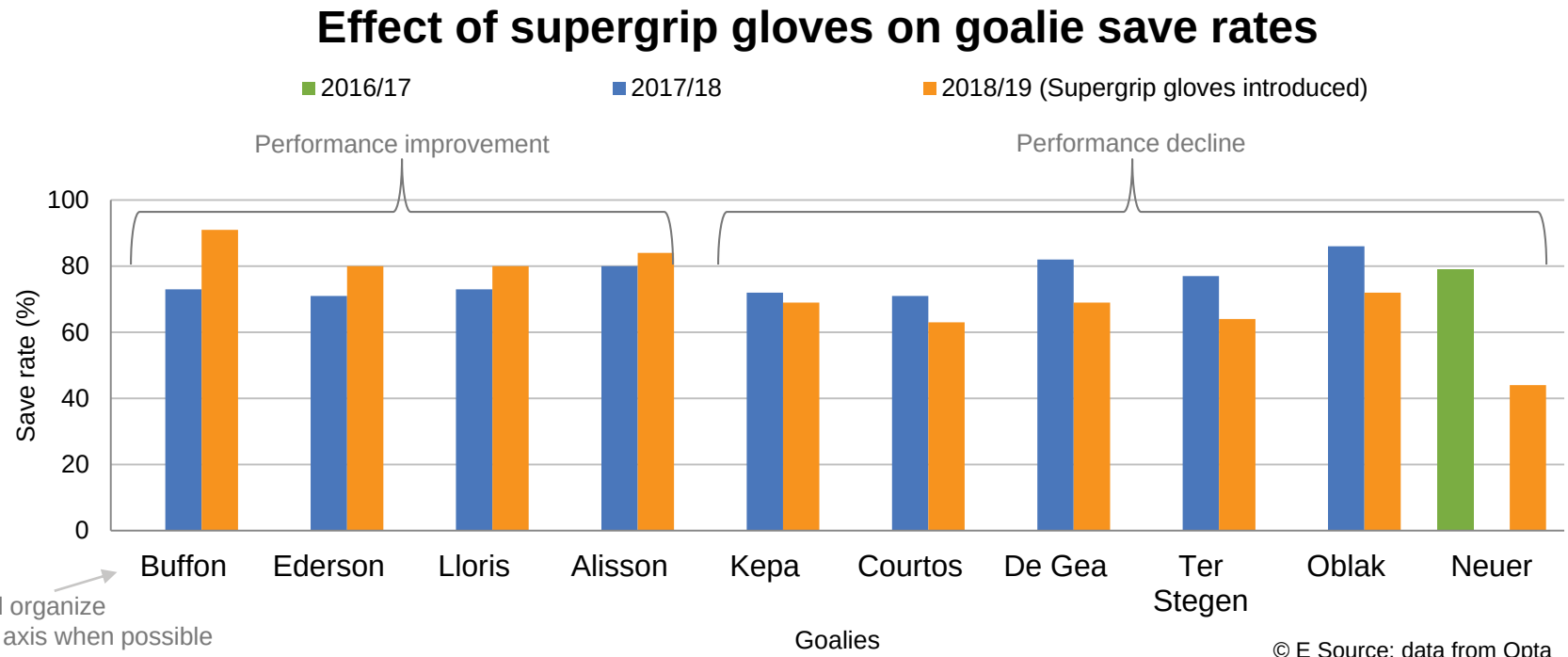


# Key points to good data storytelling

- Know what story you want to tell before you create your final graphics
- Know your audience and tailor your story to them
- Choose graphics wisely
- Guide your audience with findings

# Compelling titles grab attention

Subtext and annotation direct audience focus. Example: Supergrip gloves had no apparent effect on goalie save rates because performance improved for four participants but declined for the other six.



# Key points to good data storytelling

- Know what story you want to tell before you create your final graphics
- Know your audience and tailor your story to them
- Choose graphics wisely
- Guide your audience with findings
- Recommendations are key

# Now it's your turn

Based on the data we give you, work in small groups to decide:

- What is the story?
- How would you best tell or show that story?

# Exercise

Question from a member:

> “I've noticed a trend that walk-in centers as a payment option seem to be resurging. Have you heard of or seen a trend in the industry of utilities reopening walk-in centers?”

# Data

- E Source Contact Center Performance study: 14 utilities gave the number of walk-in centers they operated in 2015 and 2018 (see table).
- Claritas Energy Behavior Track: walk-in centers as a payment channel
  - Low-income (income less than US$25,000), Spanish speakers, and young people (18–34) are segments that report using walk-in centers almost twice as frequently as the general US population on an annual basis.

| | 2015 | 2018 |
|---|---|---|
| Utility A | 2 | 4 |
| Utility B | 2 | 3 |
| Utility C | 12 | 16 |
| Utility D | 1 | 1 |
| Utility E | 3 | 3 |
| Utility F | 6 | 6 |
| Utility G | 0 | 0 |
| Utility H | 0 | 0 |
| Utility I | 0 | 0 |
| Utility J | 2 | 1 |
| Utility K | 3 | 2 |
| Utility L | 45 | 43 |
| Utility M | 28 | 3 |
| Utility N | 16 | 0 |

# Exercise discussion

- What did you think of that exercise? Did you find it easy or difficult? Why?
- What, if any, other information or data would you have liked to have?
- How do you think the executives would react to your story or presentation?
- What would you do differently next time?

# Thank you! Questions?

Heather Hilgenkamp, PhD

Lead Analyst, Market Research, E Source

303-345-9115

heather_hilgenkamp@esource.com

Arthi Padmanabhan

Analyst, Market Research, E Source

303-345-9210

arthi_padmanabhan@esource.com



You're free to share this document in its entirety inside your company. If you'd like to quote or use our material outside of your business, please contact us at customer_service@esource.com or 1-800-ESOURCE (1-800-376-8723).

# Evaluative criteria

# Evaluative criteria for data visualization

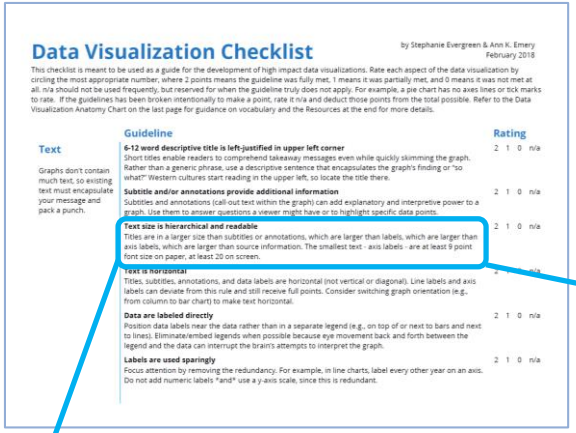
## The checklist

24 Guidelines on a 3-point scale

- Text [6]
- Arrangement [5]
- Color [5]
- Lines [4]
- Overall [4]

*Good visuals score 90% to 100% of available points*

# Data visualization checklist

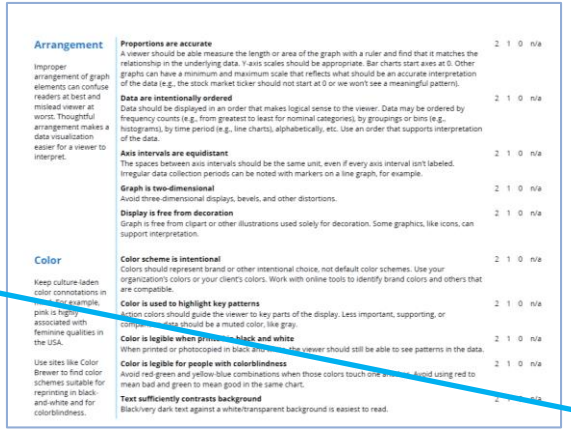


## Data Visualization Checklist

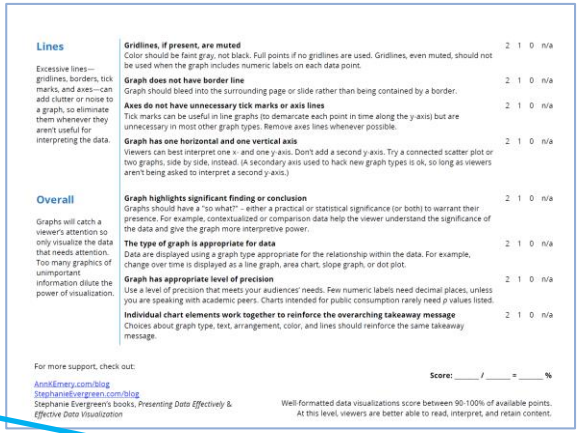
by Stephanie Evergreen & Ann K. Emery
February 2018

This checklist is meant to be used as a guide for the development of high impact data visualizations. Rate each aspect of the data visualization by circling the most appropriate number, where 2 points means the guideline was fully met, 1 means it was partially met, and 0 means it was not met at all. n/a should not be used frequently, but reserved for when the guideline truly does not apply. For example, a pie chart has no axes lines or tick marks to rate. If the guidelines has been broken intentionally to make a point, rate it n/a and deduct those points from the total possible. Refer to the Data Visualization Anatomy Chart on the last page for guidance on vocabulary and the Resources at the end for more details.

| | Guideline | Rating |
|---|---|---|
| **Text**<br>Graphs don't contain much text, so existing text must encapsulate your message and pack a punch. | **6-12 word descriptive title is left-justified in upper left corner**<br>Short titles enable readers to comprehend takeaway messages even while quickly skimming the graph. Rather than a generic phrase, use a descriptive sentence that encapsulates the graph's finding or "so what?" Western cultures start reading in the upper left, so locate the title there. | 2 1 0 n/a |
| | **Subtitle and/or annotations provide additional information**<br>Subtitles and annotations (call-out text within the graph) can add explanatory and interpretive power to a graph. Use them to answer questions a viewer might have or to highlight specific data points. | 2 1 0 n/a |
| | **Text size is hierarchical and readable**<br>Titles are in a larger size than subtitles or annotations, which are larger than labels, which are larger than axis labels, which are larger than source information. The smallest text - axis labels - are at least 9 point font size on paper, at least 20 on screen. | 2 1 0 n/a |
| | **Text is horizontal**<br>Titles, subtitles, annotations, and data labels are horizontal (not vertical or diagonal). Line labels and axis labels can deviate from this rule and still receive full points. Consider switching graph orientation (e.g., from column to bar chart) to make text horizontal. | |
| | **Data are labeled directly**<br>Position data labels near the data rather than in a separate legend (e.g., on top of or next to bars and next to lines). Eliminate/embed legends when possible because eye movement back and forth between the legend and the data can interrupt the brain's attempts to interpret the graph. | 2 1 0 n/a |
| | **Labels are used sparingly**<br>Focus attention by removing the redundancy. For example, in line charts, label every other year on an axis. Do not add numeric labels *and* use a y-axis scale, since this is redundant. | 2 1 0 n/a |



| | | |
|---|---|---|
| **Arrangement**<br>Improper arrangement of graph elements can confuse readers at best and mislead viewer at worst. Thoughtful arrangement makes a data visualization easier for a viewer to interpret. | **Proportions are accurate**<br>A viewer should be able measure the length or area of the graph with a ruler and find that it matches the relationship in the underlying data. Y-axis scales should be appropriate. Bar charts start axes at 0. Other graphs can have a minimum and maximum scale that reflects what should be an accurate interpretation of the data (e.g., the stock market ticker should not start at 0 or we won't see a meaningful pattern). | 2 1 0 n/a |
| | **Data are intentionally ordered**<br>Data should be displayed in an order that makes logical sense to the viewer. Data may be ordered by frequency counts (e.g., from greatest to least for nominal categories), by groupings or bins (e.g., histograms), by time period (e.g., line charts), alphabetically, etc. Use an order that supports interpretation of the data. | 2 1 0 n/a |
| | **Axis intervals are equidistant**<br>The spaces between axis intervals should be the same unit, even if every axis interval isn't labeled. Irregular data collection periods can be noted with markers on a line graph, for example. | 2 1 0 n/a |
| | **Graph is two-dimensional**<br>Avoid three-dimensional displays, bevels, and other distortions. | 2 1 0 n/a |
| | **Display is free from decoration**<br>Graph is free from clipart or other illustrations used solely for decoration. Some graphics, like icons, can support interpretation. | 2 1 0 n/a |
| **Color**<br>Keep culture-laden color connotations in mind. For example, pink is highly associated with feminine qualities in the USA.<br><br>Use sites like Color Brewer to find color schemes suitable for reprinting in black-and-white and for colorblindness. | **Color scheme is intentional**<br>Colors should represent brand or other intentional choice, not default color schemes. Use your organization's colors or your client's colors. Work with online tools to identify brand colors and others that are compatible. | 2 1 0 n/a |
| | **Color is used to highlight key patterns**<br>Action colors should guide the viewer to key parts of the display. Less important, supporting, or context data should be a muted color, like gray. | 2 1 0 n/a |
| | **Color is legible when printed in black and white**<br>When printed or photocopied in black and white, the viewer should still be able to see patterns in the data. | 2 1 0 n/a |
| | **Color is legible for people with colorblindness**<br>Avoid red-green and yellow-blue combinations when those colors touch one another. Avoid using red to mean bad and green to mean good in the same chart. | 2 1 0 n/a |
| | **Text sufficiently contrasts background**<br>Black/very dark text against a white/transparent background is easiest to read. | 2 1 0 n/a |



| | | |
|---|---|---|
| **Lines**<br>Excessive lines—gridlines, borders, tick marks, and axes—can add clutter or noise to a graph, so eliminate them whenever they aren't useful for interpreting the data. | **Gridlines, if present, are muted**<br>Color should be faint gray, not black. Full points if no gridlines are used. Gridlines, even muted, should not be used when the graph includes numeric labels on each data point. | 2 1 0 n/a |
| | **Graph does not have border line**<br>Graph should bleed into the surrounding page or slide rather than being contained by a border. | 2 1 0 n/a |
| | **Axes do not have unnecessary tick marks or axis lines**<br>Tick marks can be useful in line graphs (to demarcate each point in time along the y-axis) but are unnecessary in most other graph types. Remove axes lines whenever possible. | 2 1 0 n/a |
| | **Graph has one horizontal and one vertical axis**<br>Viewers can best interpret one x- and one y-axis. Don't add a second y-axis. Try a connected scatter plot or two graphs, side by side, instead. (A secondary axis used to hack new graph types is ok, so long as viewers aren't being asked to interpret a second y-axis.) | 2 1 0 n/a |
| **Overall**<br>Graphs will catch a viewer's attention so only visualize the data that needs attention. Too many graphics of unimportant information dilute the power of visualization. | **Graph highlights significant finding or conclusion**<br>Graphs should have a "so what?" - either a practical or statistical significance (or both) to warrant their presence. For example, contextualized or comparison data help the viewer understand the significance of the data and give the graph more interpretive power. | 2 1 0 n/a |
| | **The type of graph is appropriate for data**<br>Data are displayed using a graph type appropriate for the relationship within the data. For example, change over time is displayed as a line graph, area chart, slope graph, or dot plot. | 2 1 0 n/a |
| | **Graph has appropriate level of precision**<br>Use a level of precision that meets your audiences' needs. Few numeric labels need decimal places, unless you are speaking with academic peers. Charts intended for public consumption rarely need *p* values listed. | 2 1 0 n/a |
| | **Individual chart elements work together to reinforce the overarching takeaway message**<br>Choices about graph type, text, arrangement, color, and lines should reinforce the same takeaway message. | 2 1 0 n/a |

For more support, check out:

AnnKEmery.com/blog
StephanieEvergreen.com/blog
Stephanie Evergreen's books, *Presenting Data Effectively* & *Effective Data Visualization*

**Score:** _____ / _____ = _____ %

Well-formatted data visualizations score between 90-100% of available points. At this level, viewers are better able to read, interpret, and retain content.

**Text size is hierarchical and readable**

Titles are in a larger size than subtitles or annotations, which are larger than labels, which are larger than axis labels, which are larger than source information. The smallest text - axis labels - are at least 9 point font size on paper, at least 20 on screen.

Source: Evergreen & Emery, 2018

# Evaluative criteria for data visualization

## The checklist

24 Guidelines on a 3-point scale

- Text [6]
- Arrangement [5]
- Color [5]
- Lines [4]
- Overall [4]

*Good visuals score 90% to 100% of available points*

## The effectiveness profile

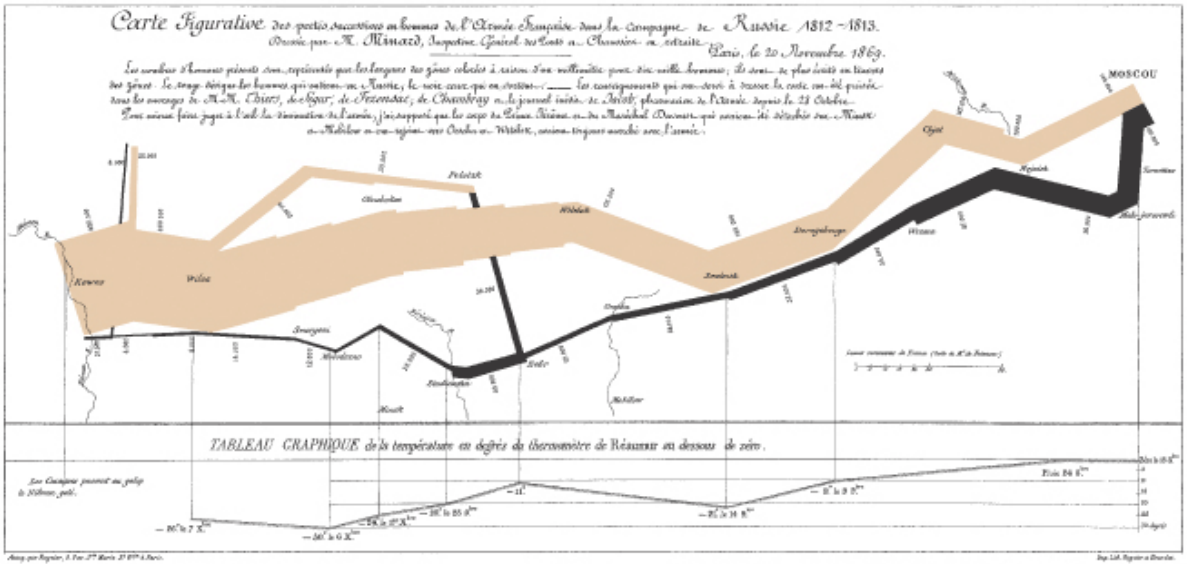
Seven target qualities on a line scale:

- Informative
  - Usefulness
  - Completeness
  - Perceptibility
  - Truthfulness
  - Intuitiveness
- Emotive
  - Aesthetics
  - Engagement

# Visualization how-to

# An old visualization

## Map of Napoleon's Russian campaign of 1812

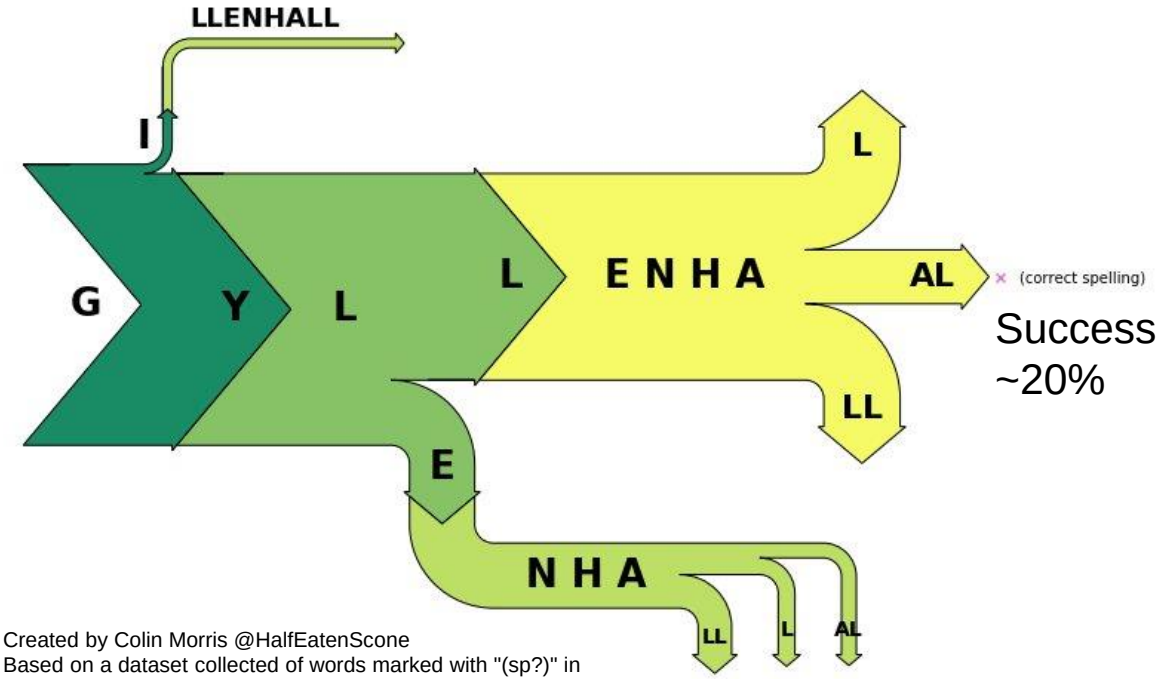



Charles Joseph Minard, 1869

*This Sankey diagram shows losses of the French army in the Russian campaign.*

# A modern Sankey diagram

## Jake Gyllenhaal spelling success rate among unsure spellers




Success rate ~20%

Created by Colin Morris @HalfEatenScone
Based on a dataset collected of words marked with "(sp?)" in Reddit comments.

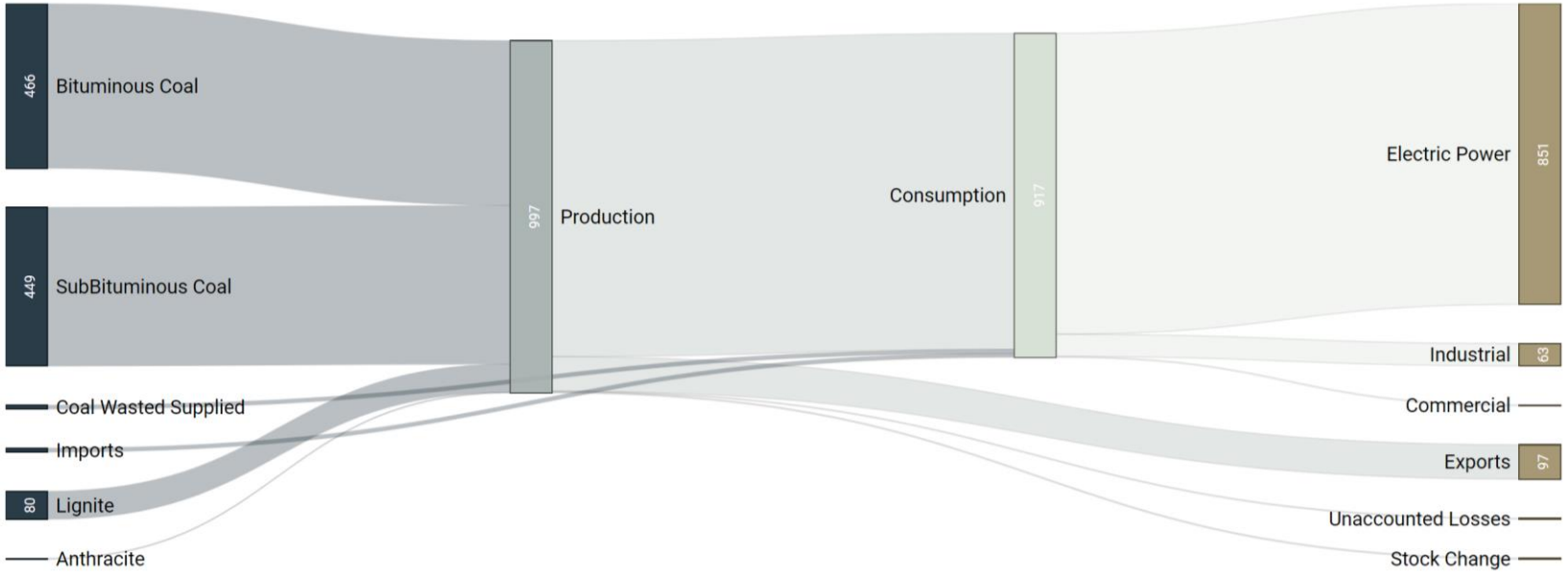
*This modern Sankey diagram uses arrows and flow to express spelling divergences.*

*Sankey diagrams work well for data than funnels or narrows.*

*How many people are aware, interested, and have adopted a product or program? How satisfied or dissatisfied are adopters?*

# Web-based Sankey diagram generator

## Coal flows in the US




Source: Tonnes, 2014

Visualization by Acquire Procurement Services

http://sankey-diagram-generator.acquireprocure.com/