STAT 102: Week 2

Ricky's Section

Introductions and Attendance

Introduction: Name, Hometown, Year, House/Dorm, Concentration

Question of the Week: What is something you're looking forward to this spring?

Important Information

STAT 102 Logistics and Procedures

- **Class** on Tuesdays and Thursdays
  - Thursday will be an interactive lab
- **Section** once a week
- **Quiz** once a week
  - Due Monday at noon to help keep your routine
- **P-set** due at Tuesday at midnight

Attendance

- Attendance is taken during section
- If something comes up, you can attend another TF's section (please notify us!)
- If you can't attend any section during a week, please notify me

Plans for Section

- **12-1 PM** every Thursday (I'll be staying after)
- Review important material
- Answer questions
- Live coding
- Work on p-set
- Community!

Tips

- Keep a running document of important code/definitions/concepts/images
  - Let's take 15 seconds to create a **Google Doc**!
- Go to class to keep up with pace/material
- Don't be afraid to ask questions

Content Review: Week 2

Grammar of Graphics

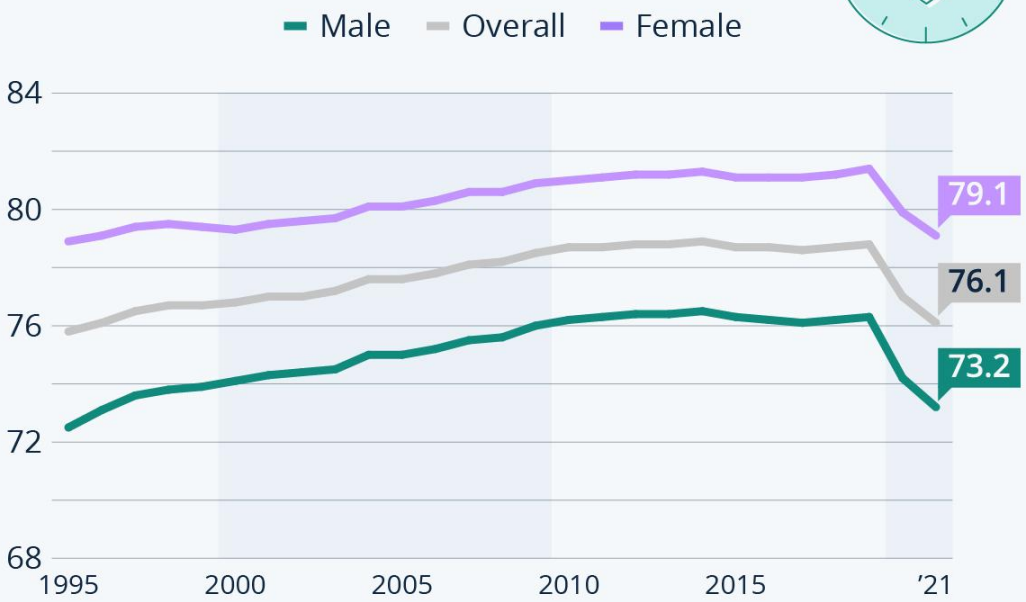
- **Dataset**: “Spreadsheet” containing data/info
- **Geom**: Geometric representation of data (as a shape), such as bars or points
- **Aesthetic**: Visual properties of geoms, such as color, scale, x-direction, or y-direction, that mean something in the context of the data

Grammar of Graphics: Example

https://drive.google.com/file/d/1y13AW7qNlvMnHGtNv1FspjUb4Vfr9Ohn/view?usp=drive_link

U.S. Life Expectancy Hits Lowest Point Since 1996

Overall life expectancy at birth in years



Source: National Vital Statistics System, National Center for Health Statistics

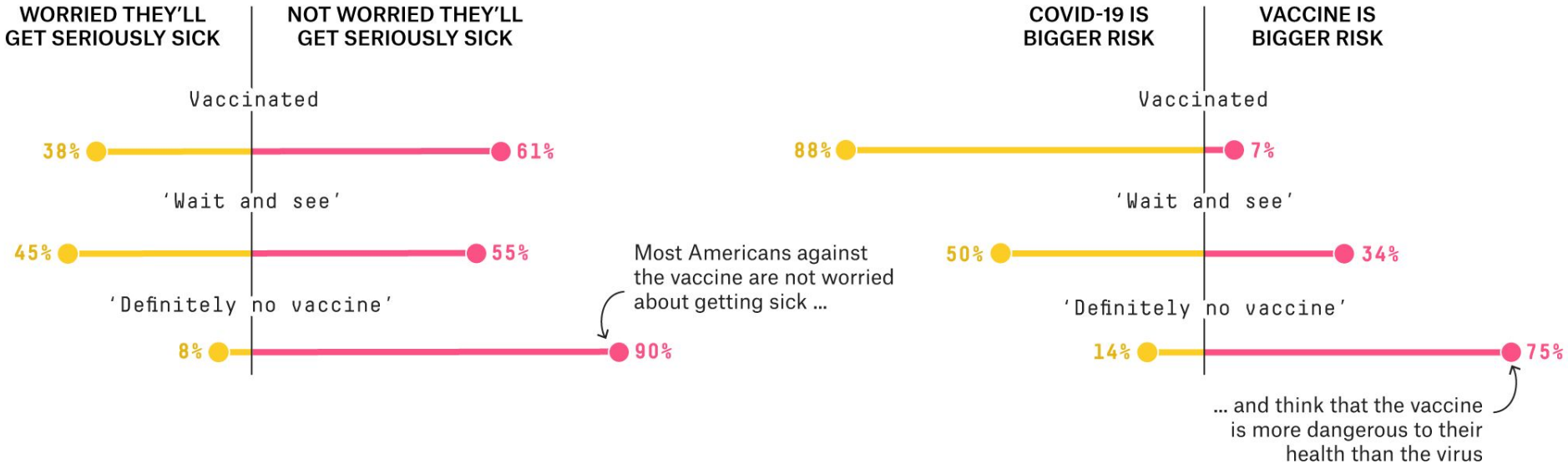
statista

Solution

- **Variables**: Year, life expectancy, gender
- **Geom**: Line
- **Aesthetics**: Year is mapped to x-location, life expectancy is mapped to y-location, gender is mapped to color

Most unvaccinated Americans think vaccine poses bigger risk than COVID-19

Share of respondents who said they are either worried or not worried from getting seriously sick of COVID-19 and who said whether they think COVID-19 or the vaccine is a bigger risk to their health, by vaccine intention



Survey conducted July 15-27 among 1,517 participants.

FiveThirtyEight

SOURCE: KAISER FAMILY FOUNDATION

Solution

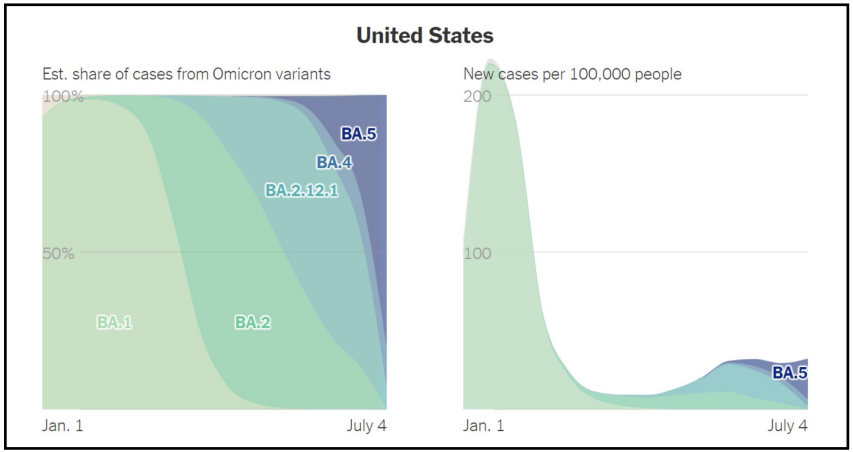
- **Variables**: Vaccine stance (e.g., “Wait and see”), response to question (e.g., “Vaccine is bigger risk”), response percentage (e.g., 34%)
- **Geom**: Line (and point)
- **Aesthetics**: Vaccine stance is mapped to y-location, response to question is mapped to color, and response percentage is mapped to x-location

Color Palettes

- **Sequential**: Color progresses from low to high value
- **Diverging**: Color splits in opposite directions, departing from a meaningful middle
- **Qualitative**: Color only distinguishes cases from each other, with no inherent order

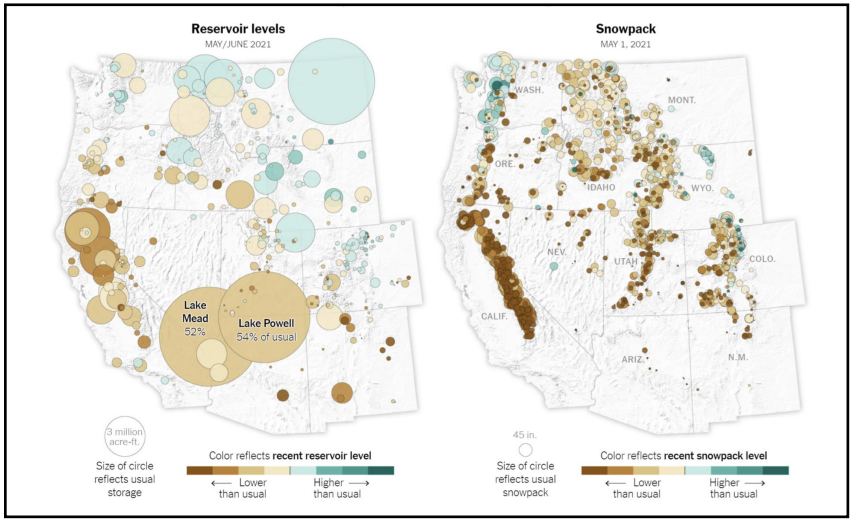
Sequential

Sequential: Color progresses from low to high value



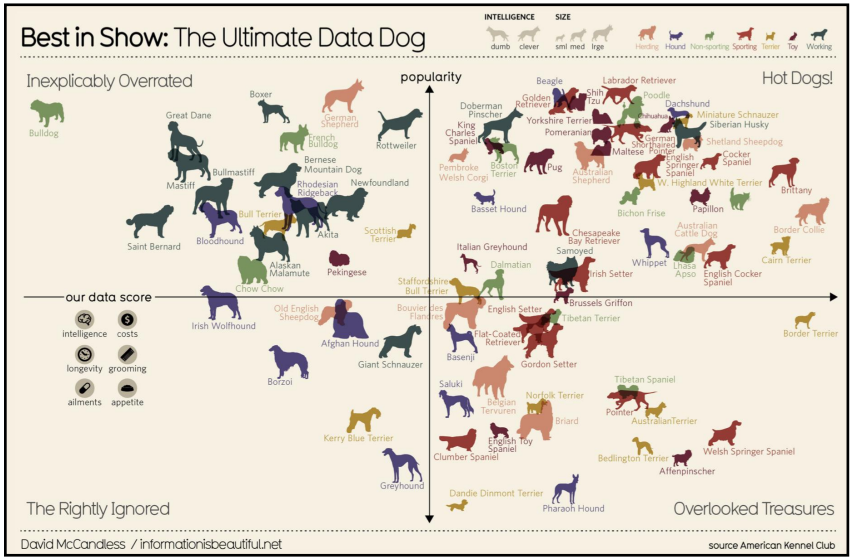
Diverging

Diverging: Color splits in opposite directions, departing from a meaningful middle



Qualitative

Qualitative: Color only distinguishes cases from each other, with no inherent order

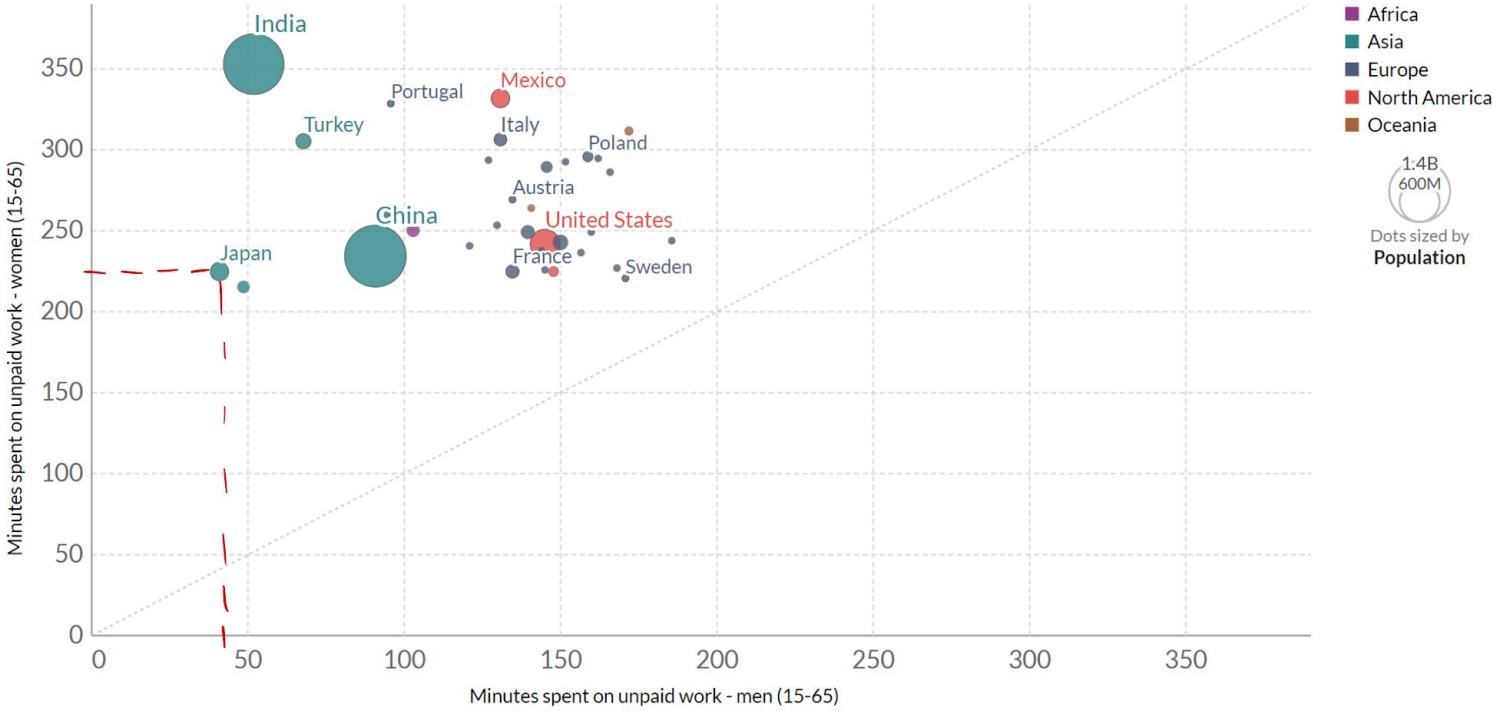


Time spent on unpaid work, per day, men vs women

Average minutes spent on unpaid work or study, per day, by sex (ages 15-65). Unpaid work activities include: routine housework; care for household members; child care; adult care; care for non-household members; volunteering; travel related to household activities; other unpaid work. Estimates come from time-use surveys and include both weekdays and weekends. The survey years differ across countries. See the source description for the survey year used for each country.

Our World in Data

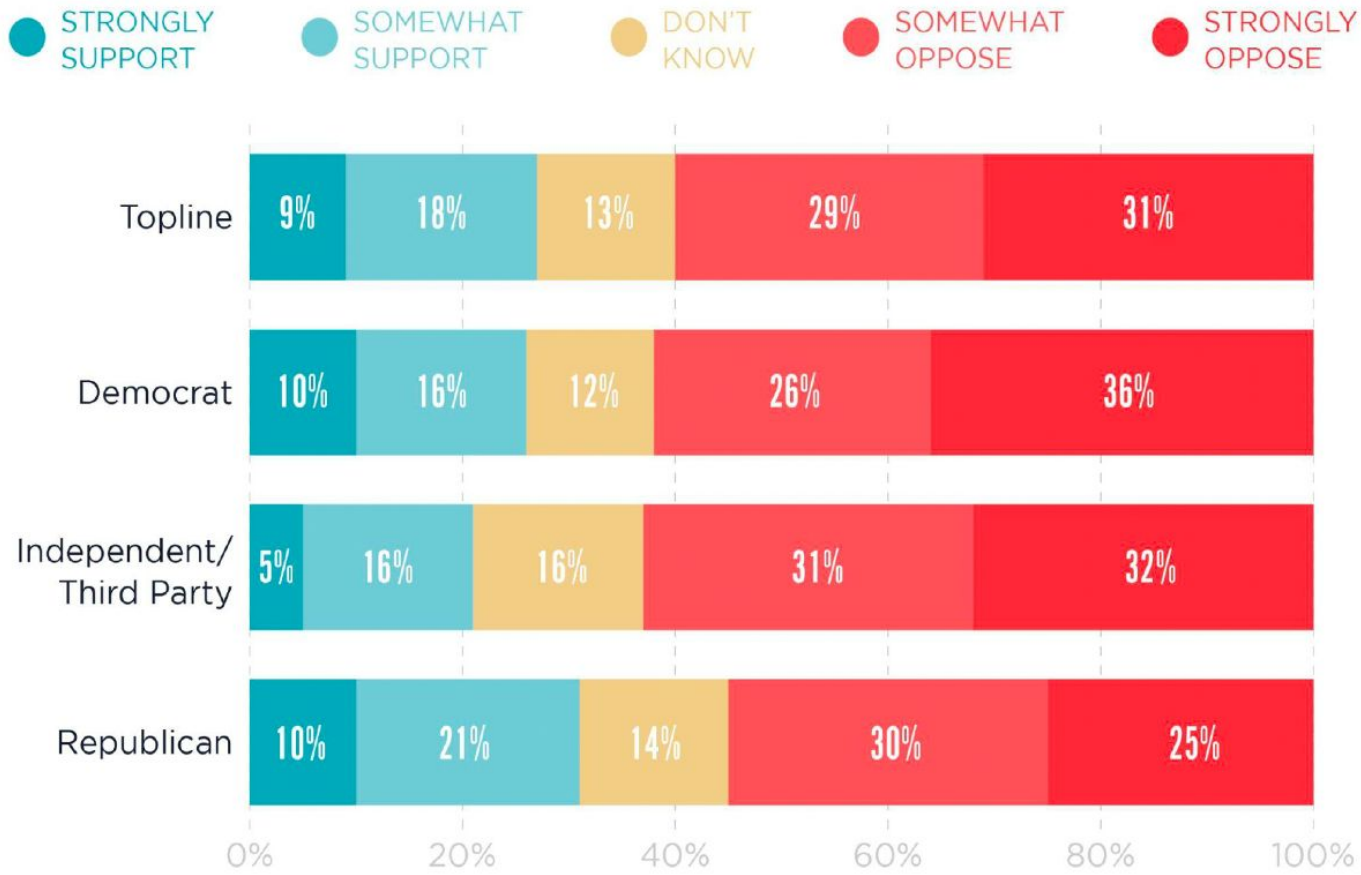
Select countries



Solution

- This is using a **qualitative** color palette
  - These colors are meant to distinguish points from one another, but they have no order
  - Europe is a darker shade of blue than Asia, but that doesn't mean Europe is "more ___" than Asia

Do you support or oppose your local prosecutor seeking a death sentence against a person with a diagnosed mental illness?



Solution

- This is using a **diverging** color palette
  - The “Don’t know” (in yellow) is our meaningful middle, from which the colors depart in opposite directions

Types of Variables

- **Numerical variables**: Take on numerical values, which you can measure and "do math"
  - *For example, salary: $100k, $50k, $70k, …*
- **Categorical variables**: Take on values that are labels, which you use to divide into groups
  - *For example, income level: low, middle, high, …*

Explanatory vs. Response

- **Explanatory variable**: Expected cause ("input")
- **Response variable**: Expected result of explanatory variable ("output")
  - *For example, measuring the effect of education level (explanatory) on salary (response)*

Choosing the Right Graph

https://drive.google.com/file/d/1GlfYFuUYMPxMgnvBtojxzQn6MGh5yl-H/view?usp=drive_link

Questions?

P-Set 1

Have a great rest of your week!