Data Visualization Using Spreadsheets

Stanford University

Google

INSTABASE

Google Cloud Platform

acm Association for Computing Machinery

Very Large Data Bases Endowment Inc.

amazon web services™

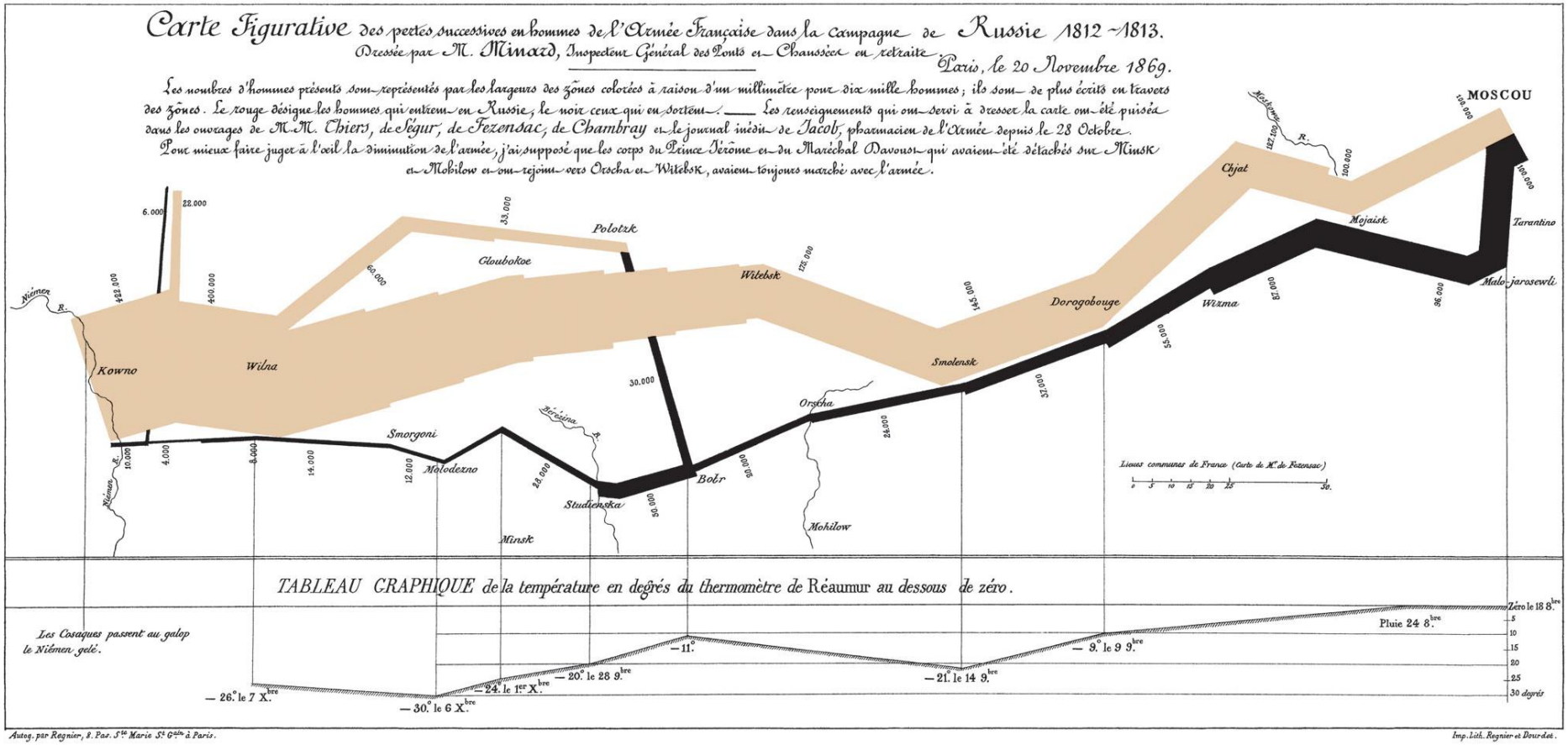
Data Visualization

"A picture is worth a thousand words"

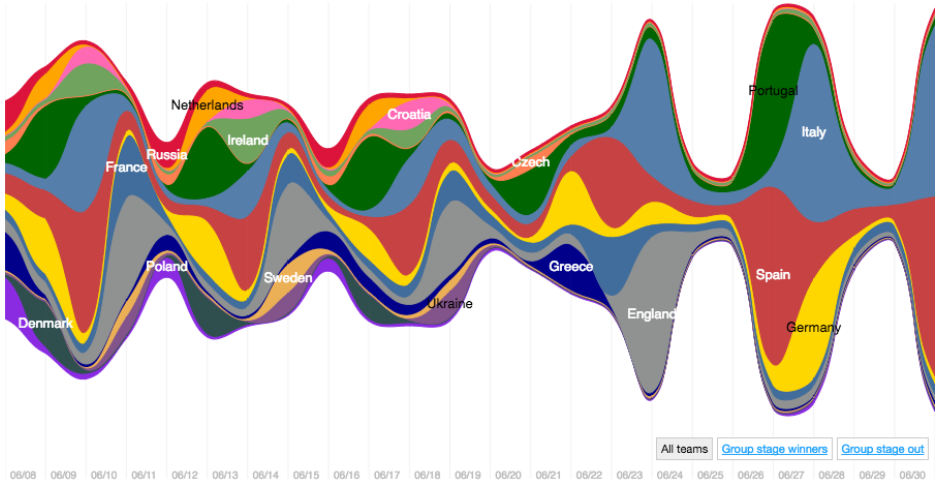
- Data visualizations can be enlightening and powerful
- Creating good visualizations is a science and an art

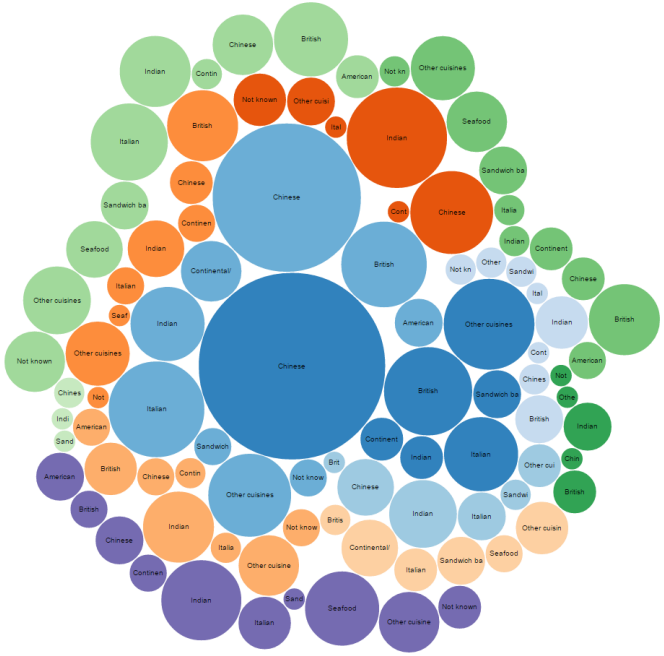
Early Data Visualization

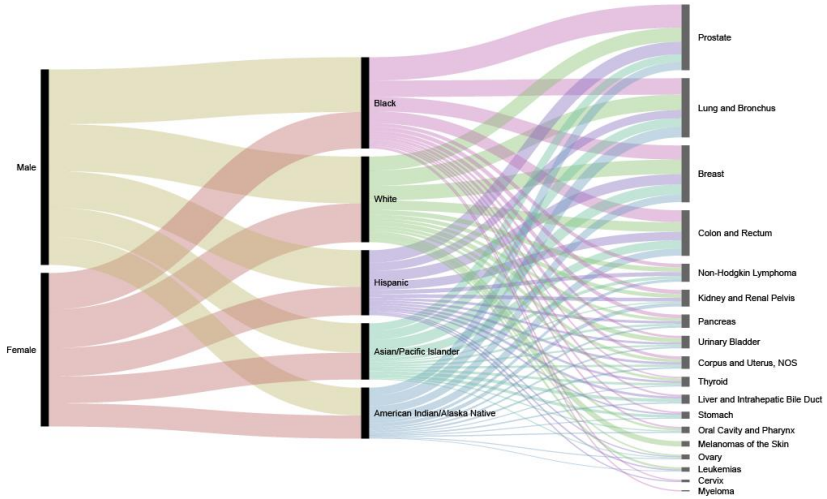
Napoleon's Army



Modern Data Visualization







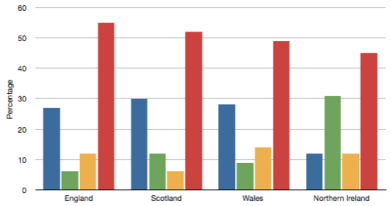
Dynamic Data Visualization

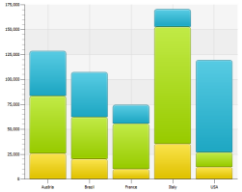
https://www.youtube.com/watch?v=jbkSRLYSojo

Basic Data Visualization

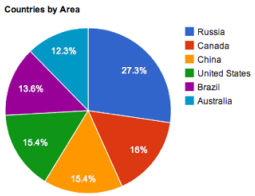
Don't underestimate the power of basic visualizations

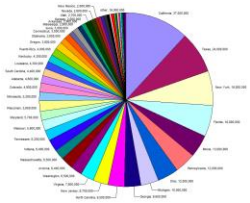
- Bar charts



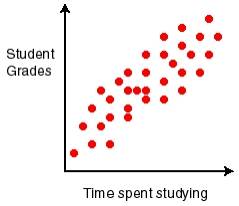


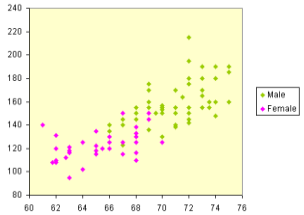
- Pie charts



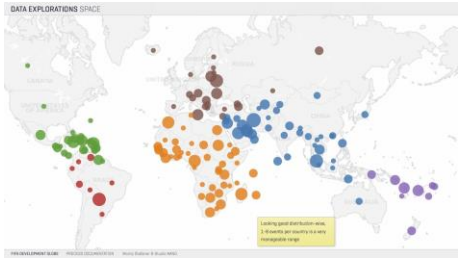


- Scatterplots





- Maps



Bar Charts

When one axis is categories
and the other is numeric

Your Turn

Ten countries with the highest population, bar chart showing populations

Pie Charts

To compare relative sizes of categories

Your Turn

Pie chart showing relative number of cities with negative longitude and positive longitude

Label the two slices “west” for west of the Prime Meridian (negative longitude), and “east” for east of the Prime Meridian (positive longitude)

Scatterplots

When both axes are numeric

Your Turn

Switch to CitiesExt.csv

Use a scatterplot to explore whether there is any relationship between the latitude of cities in a country (x-axis) and the population of that country (y-axis)

Why are there lines of dots?

Maps

Values by geographic region

Your Turn

Map of European countries colored by whether or not the country is in the European Union

Data Visualization with Spreadsheets

Convenient and useful

- Suggested charts are often good ones
- Google Sheets has fewer features than Microsoft Excel (for now)
  Though some people prefer the simplicity

For help while working with charts

- Web search, sometimes
- Or just keep fiddling ☺

World Cup Data Visualization

1. Create a **bar chart** showing the average number of minutes played by players in each of the four positions.
2. Create a **stacked bar chart** for teams that played more than 4 games, showing their number of wins, draws, and losses (not percentages).
3. Create a **pie chart** showing the relative percentage of teams with 0, 1, and 2 red cards. *Note: the pie should have three slices.*
4. Create a **scatterplot** of players showing passes (y-axis) versus minutes (x-axis). (Why are there some lines of dots?)
5. Create a **map** of countries colored light to dark blue based on how many goals their team made ("goalsFor").
6. Create a **pie chart** showing the relative percentage of players making <= 0.25 passes per minute, >= 0.5 passes per minute, and between 0.25 and 0.5. Hint: =countif() and =countifs() ✪

Titanic Data Visualization (Extra)

1. Create a **bar chart** showing the average fare paid by passengers in each class, with the three bars labeled "first", "second", "third".
2. Create a **stacked bar chart** showing the number of passengers in each class, divided into male and female (three bars). Then reverse roles and show the number of passengers of each gender, divided into class (two bars).
3. Create a **pie chart** showing the relative number of male survivors, male non-survivors, female survivors, and female non-survivors (four slices).
4. For passengers in 3rd class, create a **scatterplot** of fare (y-axis) versus age (x-axis). Don't worry about missing ages.