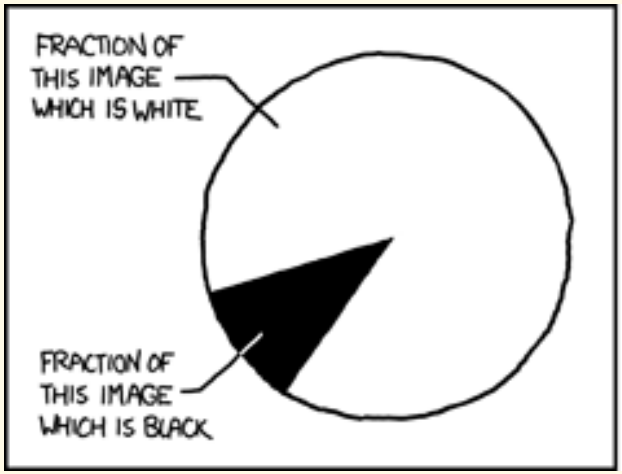
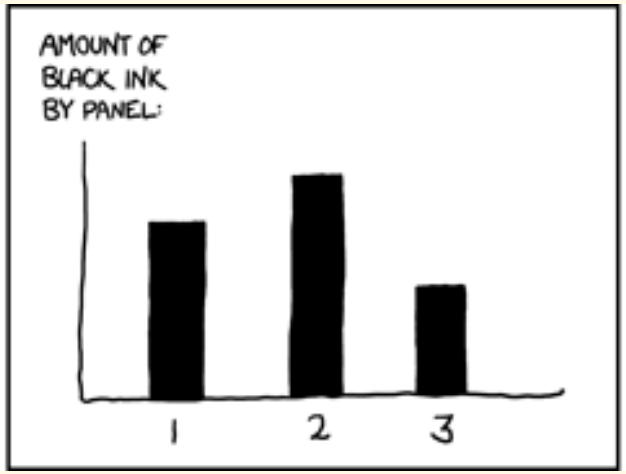
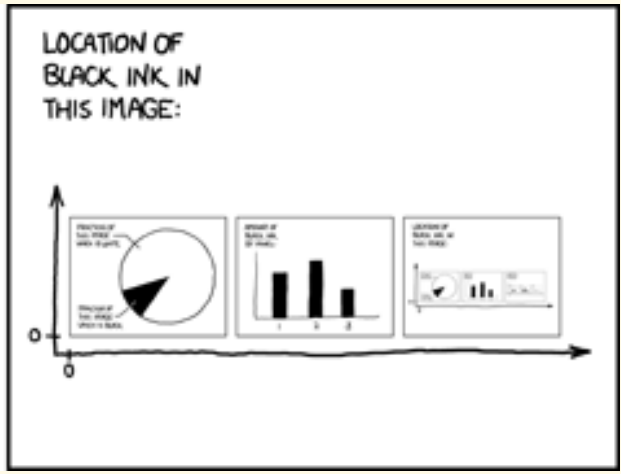
Visualization & Visual Analytics 1

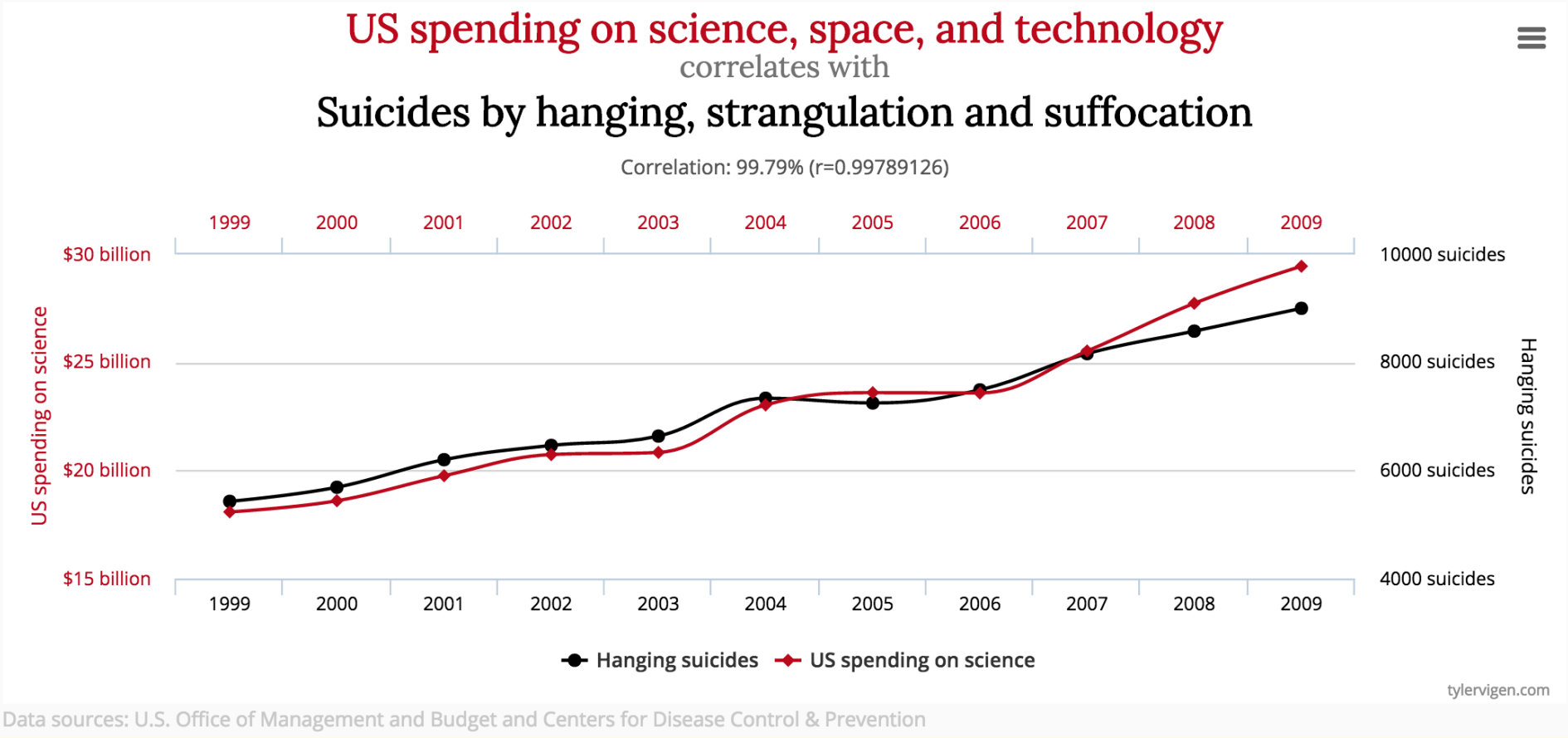
Angus Forbes

creativecoding.evl.uic.edu/courses/cs424

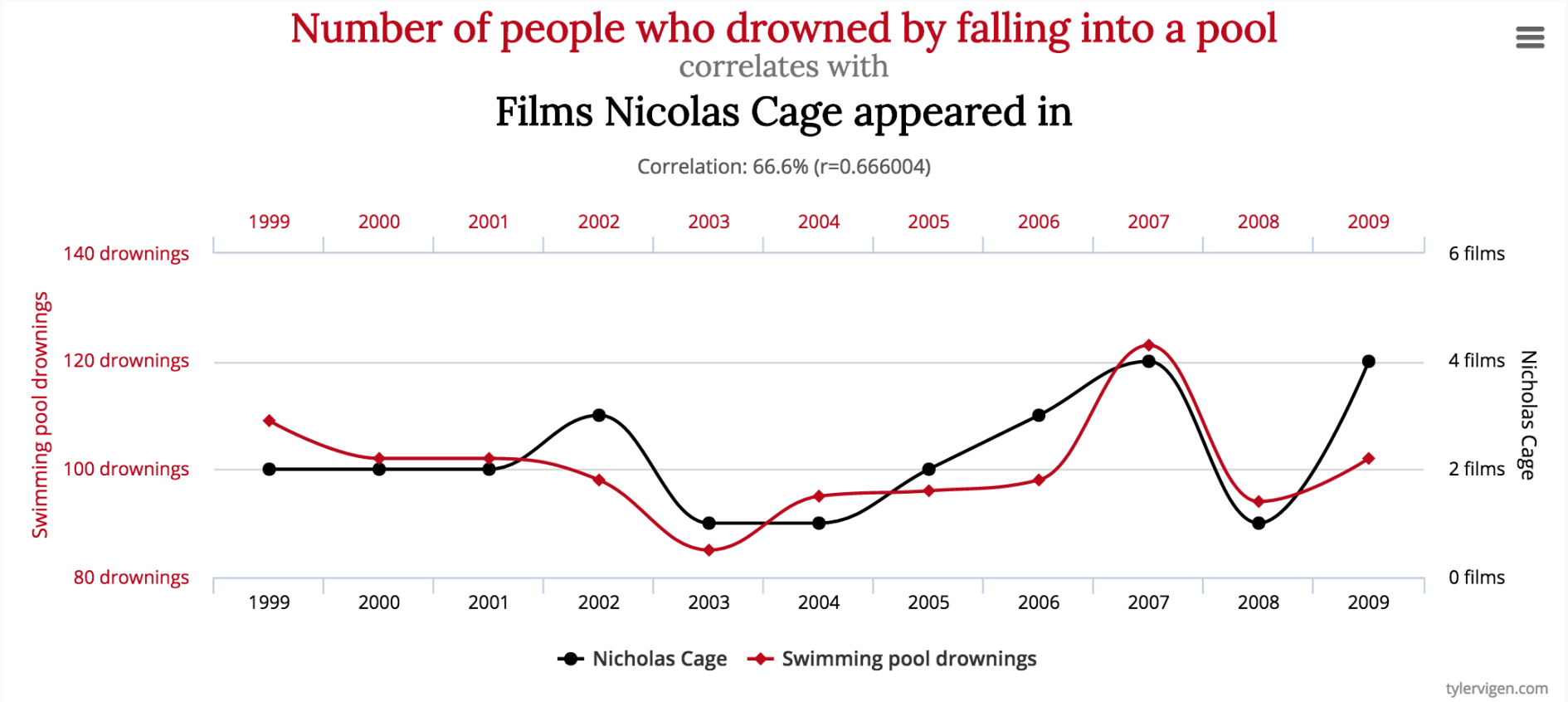

FRACTION OF THIS IMAGE WHICH IS WHITE
FRACTION OF THIS IMAGE WHICH IS BLACK


AMOUNT OF BLACK INK BY PANEL:
1
2
3

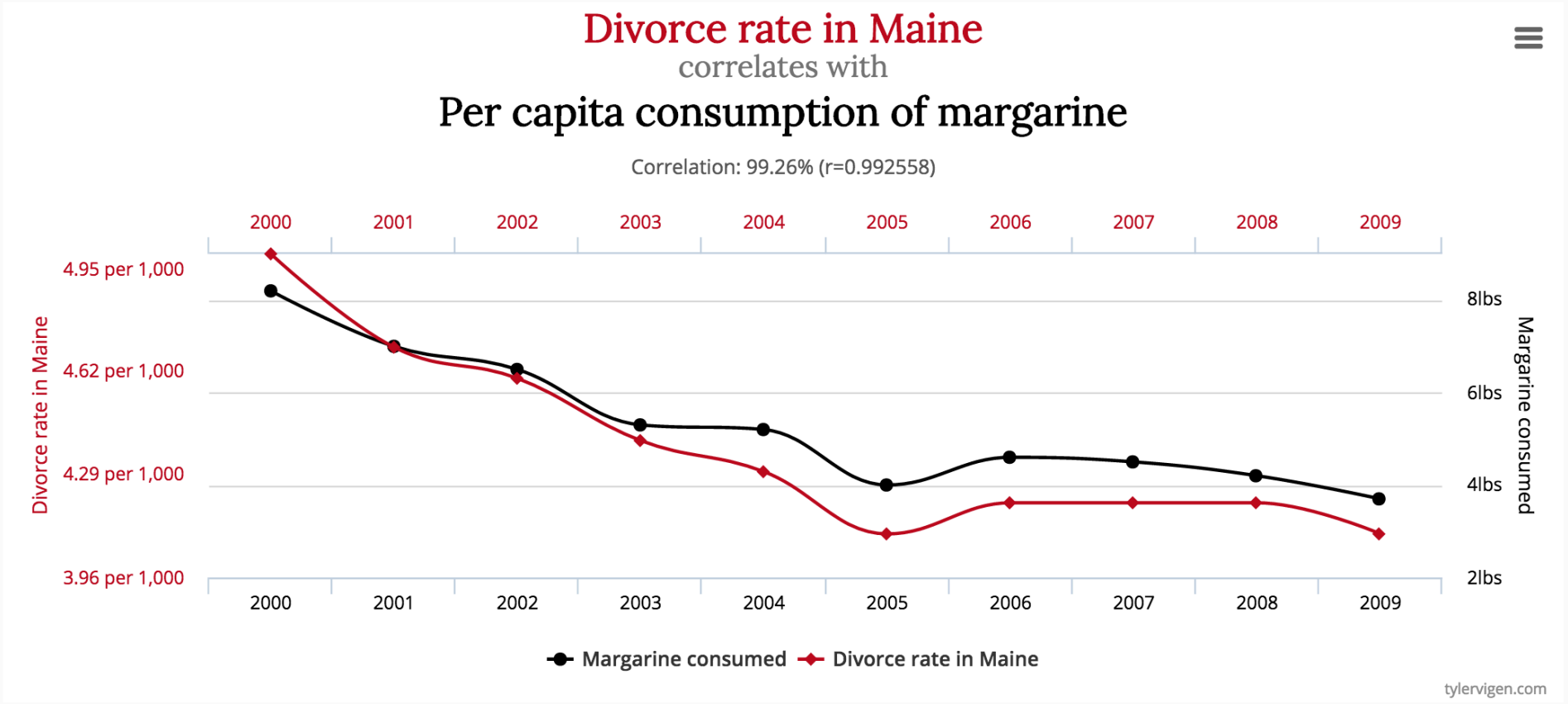

LOCATION OF BLACK INK IN THIS IMAGE:
0
0



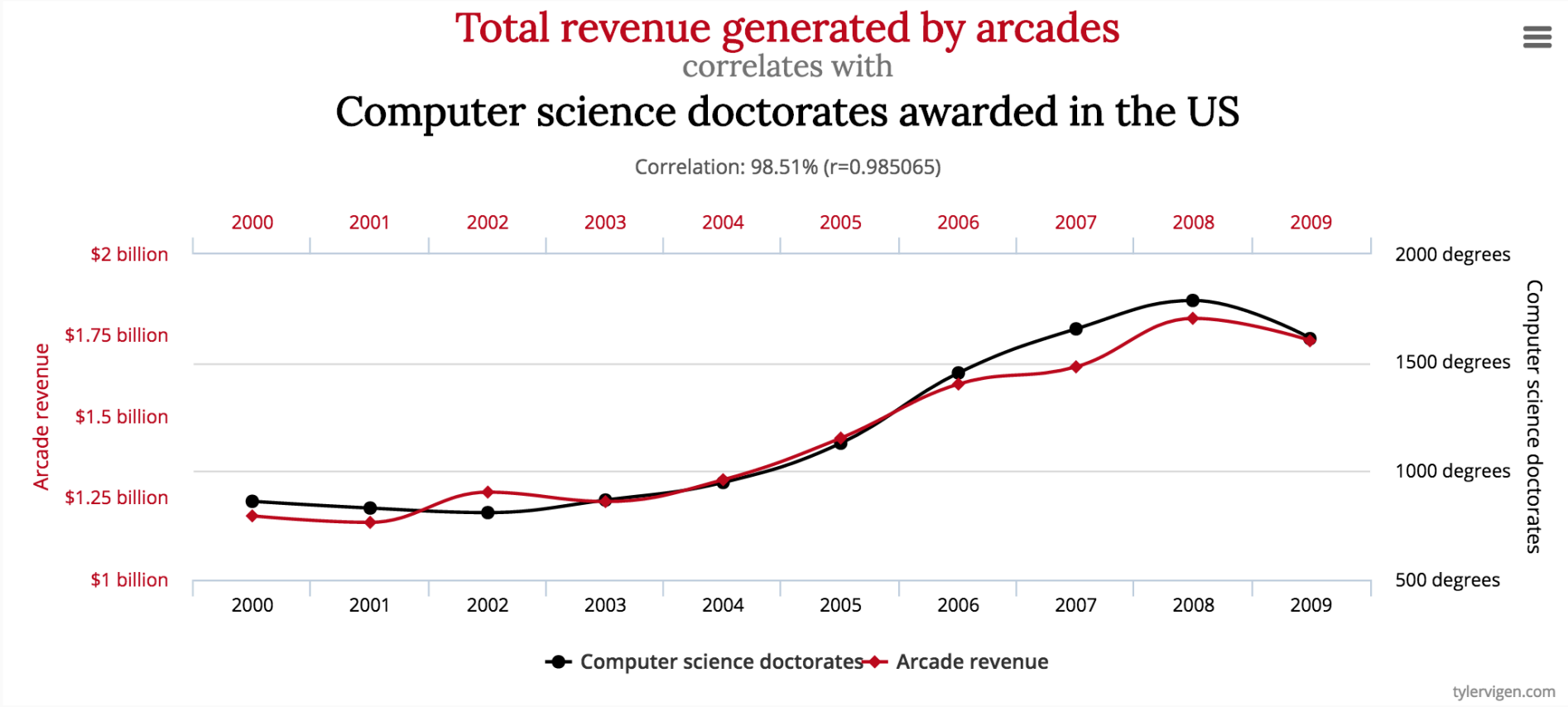
Data sources: U.S. Office of Management and Budget and Centers for Disease Control & Prevention



Data sources: Centers for Disease Control & Prevention and Internet Movie Database



Data sources: National Vital Statistics Reports and U.S. Department of Agriculture



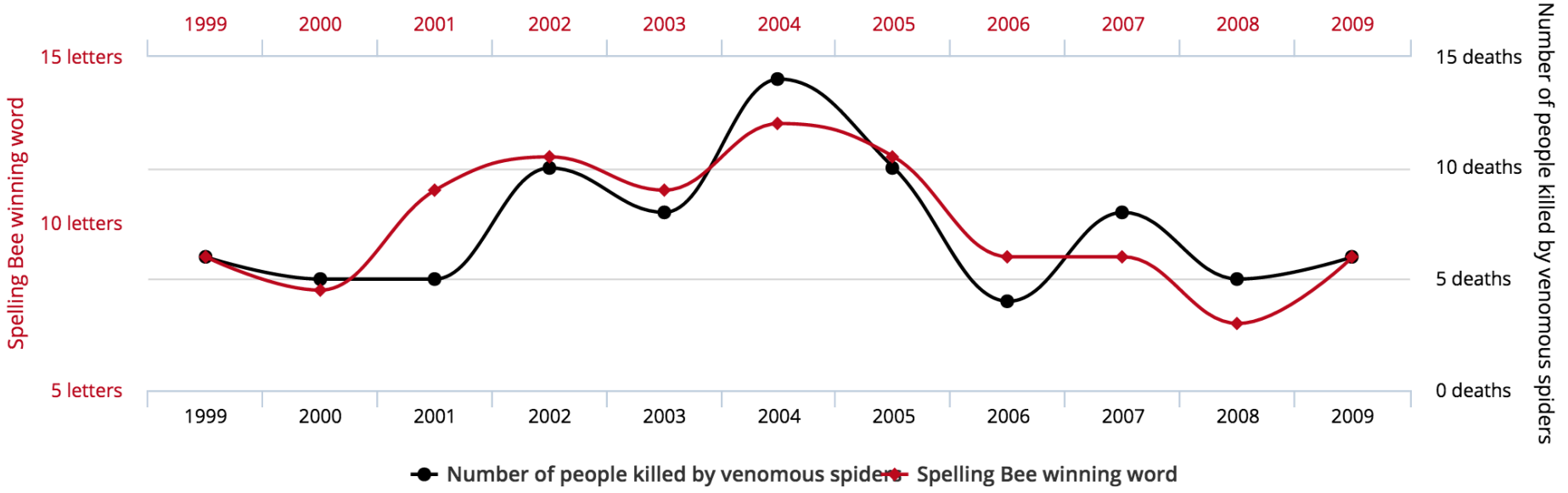
Data sources: U.S. Census Bureau and National Science Foundation

Letters in Winning Word of Scripps National Spelling Bee

correlates with

Number of people killed by venomous spiders

Correlation: 80.57% (r=0.8057)



tylervigen.com

Data sources: National Spelling Bee and Centers for Disease Control & Prevention

How – What - Why

Why (why is it important or interesting?)

What (what visualization tasks will you support?)

How (how do the visual encodings / interaction idioms enable effective visualization tasks?)

Project 2

Build a Visual Analytics Web Application

- Group Project
  - groups of 3 or 4
- Multiple components
  - research
  - application development
  - evaluation
  - documentation
- Project presentations on 11/1

Project 2

- Your visualization will integrate *two or more datasets* of your choice.

The datasets must represent different dataset types. For example, if one of your datasets includes geospatial data, then your other dataset should include, say, network data or temporal data, etc.

Project 2

- Your visualization will support one or more *visualization tasks*. That is, you will let a user filter, sort, and/or query the data in order to view subsets of your datasets and to find and analyze relationships between them.

Project 2

1. Choose datasets and determine tasks for your project

- Find some examples of interesting datasets

- What are some interesting/intriguing/ important questions that these datasets could help you to answer?

- What specific visualization tasks could you perform using these datasets?

Project 2

1. Choose datasets and determine tasks for your project

Census data, geographical data, satellite imagery, co-authorship network data, opinions on political issues, polling data, bird migration data, bee extinction data, river pollution data, weather patterns, music charts, website popularity, crime data, traffic data, etc.

Project 2

1. Choose datasets and determine tasks for your project

- compare demographics in different cities, compare universities, find correlations between spending and results, find trends over time, show geospatial shifts in research activity, analyze popularity of books, music, films for different age groups.

Tasks should be interesting, but also enable a user to answer high-level questions

Project 2

1. Choose datasets and determine tasks for your project

- A visualization to explore the relationships between crime, income, and proximity to different services (hospitals, stores, trains, etc).

- A visualization to explore how the population of the US has shifted over the last few decades, along with the rise and fall of particular industries.

- A comparison between demographic trends in the U.S. and another country.

- A visualization to identify influential people in a particular social network.

Project 2

1. Choose datasets and determine tasks for your project

Tasks should be interesting, but also enable a user to answer high-level questions.

Project 2

2. Determine appropriate visual encodings and interaction idioms for your project

- Research potential visualization techniques.

Based on the dataset type and the tasks you are interested in, find example visualizations and visualization papers that other people have used to explore similar datasets and visualization tasks. E.g., blogs, surveys, STAR reports, academic articles, etc.

Project 2

2. Determine appropriate visual encodings and interaction idioms for your project

- See syllabus for example “Survey websites”
- Syllabus has links to main visualization journals
- Google scholar is your friend

Project 2

2. Determine appropriate visual encodings and interaction idioms for your project

- Implement at least **three** visualizations that emphasize different aspects of your data.

Each visualization can look at a single dataset, or at an integration of the different datasets.

Project 2

2. Determine appropriate visual encodings and interaction idioms for your project

- Implement *brushing-and-linking*, such that a change to one visualization affects the other visualizations.

Project 2

3. Evaluate your project

- Create a user study that shows that your choices of visual encodings are appropriate ways to present your data.

- Create a qualitative study that shows that users are able to perform visualization tasks effectively using your integrated visual analytics dashboard.

Project 2

4. Document your project

- Describe your project in a short, but thorough write-up (between 4-8 pages), following the style of a paper for the VAST or InfoVis conferences

Project 2 - Example

Areas I'm interested in:

- impact of funding on academic research
- high-energy physics breakthroughs
- patterns of scientific collaboration

Project 2 - Example

Questions:

- What are popular topics in high-energy physics, and how have these topics changed over the last 20 years?

- What geographic locations are most of the research in specific areas being produced?

- Which universities / labs have the highest impact for specific sub-areas of high-energy physics?

Project 2 - Example

- How much NSF funding has been allocated to each of these sub-areas?
- I want to study in a particular sub-area, but my GREs aren't perfect - is there a lab in a less renowned university that produces great research but that I might have shot of being accepted to?

Project 2 - Example

Find potentially useful databases that could give you information to help you answer these questions:

- Download SNAP collaboration network of high-energy physicists
- Find a database about total research expenditures at universities and labs in the US
- Find database about rankings of different physics departments

Project 2 - Example

Find existing examples of how collaboration networks, research expenditures, university rankings have been visualized:

- Do a keyword search on Google Scholar
- Skim through recent proceedings of TVCG or CGF
- Look through visualization blogs
- Ask colleagues and instructor if they know of relevant techniques

In class exercise, part 1:

- What are some **topic areas** you are interested in exploring?
- What **types of questions** would you be interested in answering about these areas?
- What are some attractive / compelling / new **visualization techniques** that you are interested in exploring?
- What are some interesting **data science** tools or **statistical analysis** methods that you are interested in?

In class exercise, part 2:

- What **datasets** are available that could help answer these questions?
- Are there datasets that could help provide **context** for this data?
- What specific **visualization tasks** could be used to explore and analyze this data?
- How could **visualization techniques** be used to represent this data in order to help with your visualization tasks?

Assignment for Thursday

- Who is in your group? (Groups of 3 or 4).
- What are you broad research questions & potential visualization tasks?
- What are some example datasets that you think you will use for the project?
- What are some initial ideas for how the visualization might look?
- How might a user interact with your visualization in order to find out detailed information about particular data points or relationships between data points?

Assignment for Thursday

Think about:

Why (why is it important or interesting?)

What (what visualization tasks will you support?)

How (how do the visual encodings / interaction idioms enable effective visualization tasks?)

Tuesday's class

- Programming Lab

- What are some D3 questions you would like Shiwangi to cover?

- Maps? Animation? Interaction? Loading data from an SQL database? User interface elements? Network layouts?

Aspects of the Project

Domain situation

Data/task abstraction

Visual encoding/interaction idiom

Algorithm