The "Data Mining" Specialization | Learn More | ✕

Programming Assignment 1: Visualize Data Using a Chart

On this page:

Overview   Goals   Time Estimation   Instructions   Submission   Evaluation   Q & A

Sample Submissions   Additional Resources for D3

Overview

This assignment will give you a chance to explore the topics covered in week 2 of the course by visualizing some data as a chart. The data set we provided deals with world temperatures and comes from NASA: http://data.giss.nasa.gov/gistemp/. Alternatively you can use any data that you would like to explore. You are not required to use D3.js, but if you would like to, we have provided some helpful resources that will make it easier for you to create a visualization. You are welcome to use the additional resources, especially if you do not want to program to complete this project.

Goals

The goal of this assignment is to give you some experience with handling and deciding how to visualize some data and for you to get a feel for the various aspects of the visualization.

This assignment will also help you to analyze other visualizations and assess their effectiveness.

Time Estimation

This is not a tricky assignment, but the amount of time that it takes will vary based on the tools you use and the amount of customization you would like to do in your chart.

Instructions

1. Take the data from the GISTEMP site, specifically the data from "Table Data: Global and Hemispheric Monthly Means and Zonal Annual Means." Alternatively you can use any data that you would like to explore instead.
2. Parse the data to a suitable format for the tools that you are using – we have provided two files (in JS, TXT, and CSV formats) that contain the data used to make the visualizations here, which is a subset of the data on the GISTEMP site.
3. Visualize the data in any meaningful way you wish, keeping in mind the requirements of the Programming Assignment 1 Rubric.
4. Click to download the .zip file for this programming assignment. Here is the original data from the GISTEMP site.

If you need them, there are free-to-use websites that you can use to write and execute D3 code:

- http://code.runnable.com (see also http://code.runnable.com/d3.js for examples)
- https://vida.io

Log of Changes to the Data

We realize that we have updated the data provided from our side multiple times, so this a quick log of the changes:

1. We updated the .zip by adding the CSV and TXT files, in response to a request.

2. There was a formatting issue (there were commas in incorrect places) in the files which caused confusion. The formatting was changed to become clearer.

3. It was pointed out that the data from the GISTEMP site did not match the data we had provided on our side. This was because there had been an update to the GISTEMP data, due to a bug which they had found, since the time we created the .zip on our side. We overlooked this update, which is why the data had differed. A small formatting change was also made for clarity. In the second DATA2 files, there are no longer spaces in the longitudinal demarcations. From "24N -90N", it is now "24N-90N", and likewise.

Submission

You must upload an image (or a URL to a website displaying an image) of your visualization for peer evaluation.

Alongside your visualizations, feel free to include a paragraph that helps explain your submission.

A few questions that your paragraph could answer:

1. What are your X and Y axes?
2. Did you use a subset of the data? If so, what was it?
3. Are there any particular aspects of your visualization to which you would like to bring attention?
4. What do you think the data, and your visualization, shows?

Submit Programming Assignment 1

Submission Deadline

Your submission for Programming Assignment 1 will be due Aug. 2.

Evaluation

Your peers will use the Programming Assignment 1 Rubric to evaluate your submissions. This assignment worth 15 points. The evaluation period will not begin until Aug. 3.

Evaluation Deadline

Your evaluations of at least 4 student submissions for Programming Assignment 1 will be due Aug. 9.
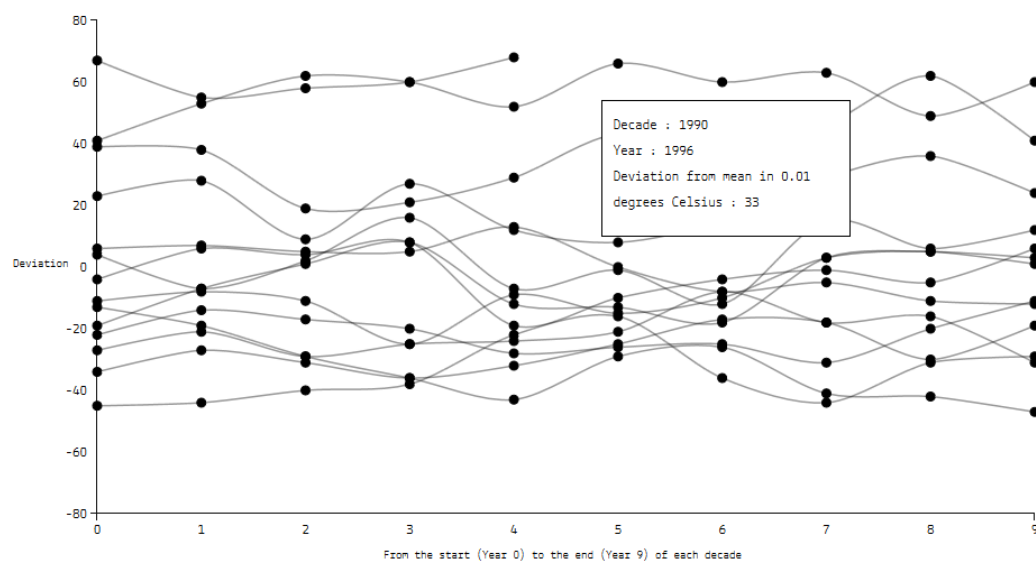
Q & A

Programming Assignment 1 Help Forum

Sample Submissions

Sample Submission 1:

Graph Submission:

Explanation:

This is a sample submission of a visualization of the data from the GISTEMP site. Every line corresponds to a decade, with every point on the line being a year. The Y Axis is the deviation from the 1951 – 1980 mean in 0.01 degrees. The X Axis then goes from the start to end of the decade.

When hovering over any of the dots, there is a tooltip that displays that year along with the particular deviation for that year. The resulting graph shows a clear increasing mean temperature over the decades.

Grading Rubric

Appropriate Chart Selection and Variable	5 points - The data is well represented by the assignment of data to variables.
Design of the Chart	4 points - Hard to follow crossing lines.

Note: The Contest Part of the Rubric is up to you.

Sample Comments from Instructor:

The easiest fix to the first chart submission would be to use color (hue) to differentiate the decades (here shown as a nominal or ordinal variable indicating which line corresponds to which decade). Since the lines cross each other so often, it may be a little hard to follow one line closely. It would be a good idea to have each line colored differently to better distinguish between them.

Sample Submission 2:

Graph Submission:
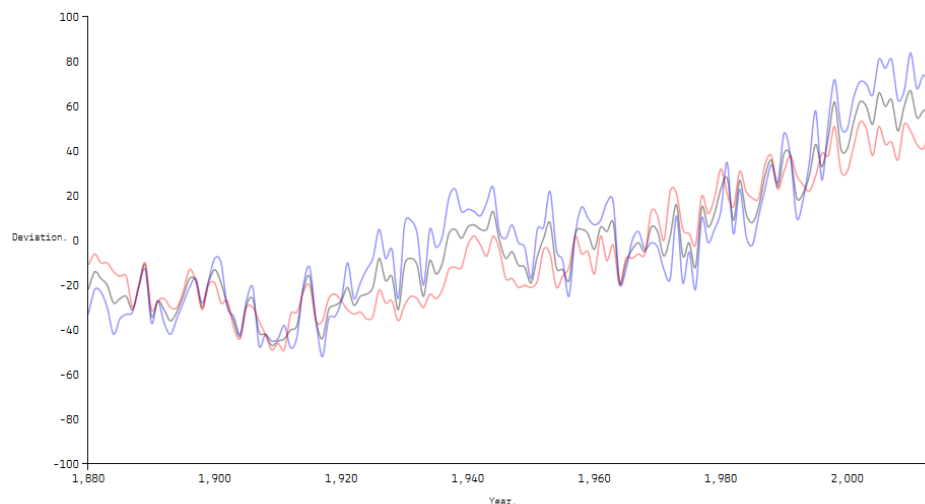
Explanation:

This graph visualizes the GISTEMP data for the Globe and the North and South Hemispheres through all the given years. The Blue line describes the data for the Northern Hemisphere, the Red for the South Hemisphere, and the Black line is for the Globe.

The resulting graph shows an increasing mean Global Temperature over the years.

Grading Rubric - Part 2:

Appropriate Chart Selection and Variable	5 points – The data is well represented by the assignment of data to variables.
Design of the Chart	4 points - The lack of a legend makes the explanation necessary and reduces the direct effect of the chart.

Note: The Contest Part of the Rubric is up to you.

Sample Comments from Instructor:

The graph would probably be better with a legend on the graph itself, so it would be easier to tell what each line describes.

Additional Resources for D3

These Sample Submissions used D3.js for the Graphs. You are not required to use D3.js, but if you would like to, the following are some helpful resources that will make it easier for you to create a visualization:

D3: http://d3js.org/

Tutorials and examples:

D3 Tutorial Book: https://www.dashingd3js.com/table-of-contents

D3 – Making a Line Graph: http://code.tutsplus.com/tutorials/building-a-multi-line-chart-using-d3js--cms-22935

Simple D3 Line Graph: http://bl.ocks.org/d3noob/b3ff6ae1c120eea654b5

D3 – Line Graph: http://www.d3noob.org/2014/07/d3js-multi-line-graph-with-automatic.html