Educational Data Visualization Project

Amy Gray, Everett Moser, Jewell Simon, John Hogan, Billy Musgraves, Laura Vavrek, Mary Wang, Wayne Picou

Math 4020 Capstone Course, Fall 2016, Department of Mathematics, Louisiana State University

MOTIVATION AND MISSION

Motivation
- There exists a large amount of public data collected by Louisiana about the performance of its schools.
  - This data is divided into data sets and distributed yearly on the department of education's website. LouisianaBelieves.com
  - The data is in raw formats of various excel files and PDF documents.
  - This makes the data hard to analyze without having some expertise in mathematics.

Mission
- To create a simple and comprehensive data analysis tool to better understand a variety of data centered around the performance of school in Louisiana
- To increase the potential uses and demographic of this tool, it must serve multiple functions
- The project must be created using the programming language R

SIMPLIFYING THE DATA

- The data on the website louisianabelieves.com came in all different formats. They were in PDF file and Word documents and had to be turned into files that the R program can read
- Needed consistency across the data so that when uploaded into R, the program will be able to work the same across all data
- Required to clean and reformat the data so that it is consistent from year-to-year and school-to-school and thus can be analyzed collectively.

ONE NUMBER SUMMARY

- **Problem.** We want to rank the schools and if possible to assign a single number to each school as an indicator of the aggregate performance of its students on each test.
- **Solution I.** Assuming the scores in a school are (approximately) normally distributed, it is possible to estimate the mean and standard deviation from the percentages in the performance bands. We can compute the probability of the observed band percentages as a function of m and s (and the number of students). The Maximum Likelihood estimate is the value of (m, s) that maximizes this probability. It is independent of the number of students. (the number of students enrolled would be useful in determining confidence intervals) We can use the Maximum Likelihood estimate of m as an indicator of school performance (in the given test).
- **Solution II.** Other one-number summaries are linear functions of the band percentages. The optimal linear one-number summary is that from which we can most accurately recover the band percentages of all schools. Experiments show that this tends to be about

  (0.8)(%Excellent) + (0.3)(%Good) + (0.2)(%Fair) - (0.3)(%Needs Improvement).

DATA SIMPLIFICATION TOOL

- We use Rstudio and Shiny to create the Data Simplification Tool and present the information to the public.
- This tool is used to filter your desired results.
- The different filters would be:
  - List of schools to choose from
  - Upper and lower bound of a population of certain schools
  - Percentage of students who are receiving free lunch
  - District filter (Rural, Urban, or Either)
  - Percentage of minority at different schools
  - You could select a range of a certain score percentile of a test or overall school performance of a school.
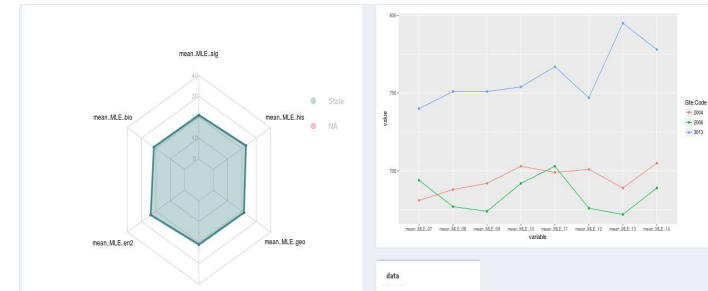

GRAPHIC USER INTERFACE

- We used Rstudio and Shiny to create our GUI to present to the public
- What is Rstudio?
  - Wikipedia says: "Rstudio is a free and open-source integrated development environment (IDE) for R, a program language for statistical computing and graphics"
- What is Shiny?
  - "Shiny is an open source R package that provides an elegant and powerful web framework for building web applications using R"
- What graphical representation should we use?
  - Radar chart and line graph are examples of possible ways to visualize the results of the search.

CONCLUSION

- What will the project do?
  - Data that was once scattered and unorganized will be useable and meaningful to the public
- What we hope it accomplishes?
  - Administrators can make connections between similar schools and learn how they can improve their own schools.
  - Parents can choose to place their child in the school that is best suited for their needs.