Tax Data 2014 – Returns, Exemptions, and AGI

Week 4 Project – 11/20/16

Michael Sloan – Student MSDS Program

Computer Science Department - Lewis University

Romeoville, Illinois

michaeldsloan@lewisu.edu

*Abstract*—This document provides an overview of the work completed to create the TaxStats interactive data visualization. This visualization provides a view of specific statistics monitored and distributed by the United States Internal Revenue Service.

# Introduction

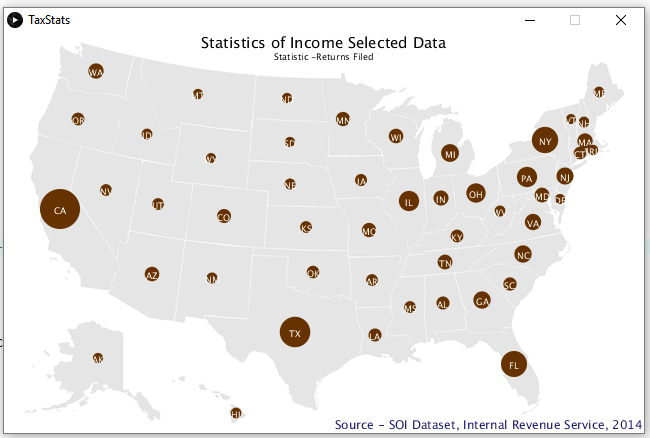
This document provides a list of features required, features provided, and a brief review of the dataset produced in the document. The document also provides a description of the components provided with this report to access and install the visualization.

# Visualization Features

## Required Features

This visualization required the following:

* Create a Processing Sketch.
* Display the chosen map.
* Display a title describing the type of data visualized.
* Display a subtitle describing the specific data visualized.
* Visualize one statistic from the dataset that is described by the subtitle.
* Allow the user to toggle using the space part to display different a different statistic and appropriately change the title.
* There should be a smooth transition for displaying one statistic to another.
* Use the Design Considerations learned in the previous weeks.
* Visualize Three Different Statistics.



* Run the Code without Errors.

## Provided Features

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors*).

The visualization created addresses all the essential features.

* The image from Fig 1. comes from the TaxStats Processing 3 sketch.
* The chosen map and location data come from <http://benfry.com/writing/map/map.png> [1][2].
* The subtitle in Fig 1 shows the statistic “Returns Field” for 2014.
* TaxStats displays three statistics. [3]

Figure – Sample First Page

* + Returns Filed (roughly approximates households).
  + Exemptions (taken) – Roughly approximates the populations.
  + AGI – Adjusted Gross Income – shows unweighted average incomes by state by AGI code.
* The reader only transitions from page to page by pressing the space bar after selecting the visualization. Pressing other keys show no effect.
* The reader smoothly transitions from screen to screen. Screens reset when the reader pushes the space bar. Existing values crisply replace old ones.
* Specific design considerations impacted the design.
  + The developer selected a suitable method for proportional relationships. - Circles work well for showing relative proportions.
  + Proximity associated values (quantities) with States.
  + Integration of common state codes effectively supported the association of states to geographic values.
  + High contrast and darker colored bubbles reinforced the idea of data “overlaying” the map.
  + Colors linked population statistics between pages (Returns Filed – *households*, and Exemptions – *population*).
  + Customized scaling allowed enhancing the obviousness of relationships.
* The code contains no known functional errors.
* The code is liberally commented, with comments at document, method, and line.

# Dataset

| **Field Name** | **Schema Metadata** | | |
| --- | --- | --- | --- |
| ***Type*** | ***Category*** | ***Purpose*** |
| FIPS | String – Class | Dimension | 2 Character FIPS Code |
| RETURNS FILED | Numeric – Continuous | Measure | Primary energy consumption less the total energy consumption for the observation period. |
| EXEMPTIONS | Numeric – Continuous | Measure | Total number of exemptions filed per state. |
| AGI | Numeric – Continuous | Measure | Adjusted Gross Income per state. |
| TOTAL INCOME | Numeric – Continuous | Measure | Total income per return (before adjustments). |
| AVERAGE TOTAL INCOME PER RETURN | Numeric – Continuous | Measure | Calculated averages per return at a state level. |
| STATE | String – Class | Dimension | 2 Character FIPS Code |
| XLOC | Numeric – Continuous | Dimension | Location (X coordinate) |
| YLOC | Numeric - Continuous | Dimension | Location (Y coordinate) |

The dataset consolidates information from two sources: the *locations.tsv* file [2] and the *SOI Tax Stats - Historic Table 2 Data Tax Year 2014* table [2].

The *locations.tsv* file provided fields STATE, XLOC, and YLOC. The *SOI Tax Stats - Historic Table 2 Data Tax Year 2014* file provides fields: FIPS, RETURNS FILED, EXEMPTION, AGI, TOTAL INCOME, AVERAGE TOTAL INCOME PER RETURN.

# Files Provided

1. Files in Zip File

| File | ZIP File Content |
| --- | --- |
| File Description |
| Table.pde | Table Class PDE File. |
| TaxStats.pde | The Processing 3 PDE File |
| 20162011 – Tax Returns 2014 – Week 4 Project.docx | This report. |
| frymap.png | Image file used by TaxStats.pde |
| TaxStats.tsv | Consolidated data file. |

##### References

1. Ben Fry, MAP.PNG (2007)[HTML Download], http://benfry.com/writing/map/map.png
2. Ben Fry, locations.tsv (2007)[HTML Download], http://benfry.com/writing/map/locations.tsv
3. US Government, SOI Tax Stats - Historic Table 2 Data Tax Year 2014, (2014) [Download], https://www.irs.gov/uac/soi-tax-stats-historic-table-2
4. US Government, SOI Tax Stats - Historic Table 2 Data Tax Year 2014 Documentation Guide, (2014) [Download DOCx], https://www.irs.gov/uac/soi-tax-stats-historic-table-2
5. Ben Fry, MAP.PNG (2007)[HTML Download], http://benfry.com/writing/map/Table.pde.