Federal Contracting in Hampton Roads

Shawn Jones, *Old Dominion University Computer Science Department* and Valentina Neblitt-Jones, *Old Dominion University Libraries*

Abstract—We present a visualization for United States federal government contracting awards in the Hampton Roads region of the State of Virginia. Our visualization consists of a map (choropleth) displaying different colors based on the funding each locality receives. To the right of the map is a bar chart displaying a bar indicating how much funding each industry received. On top of the map and the bar chart is a small line chart (sparkline) showing the trend in funding. Our visualization allows the user to select a year, agency, or locality with the Hampton Roads area, updating the choropleth, bar chart, and sparkline as appropriate. We detail issues with acquiring and processing data from usaspending.gov. We also discuss how our visualization fits within Munzner's What-Why-How framework.

1 Introduction

Understanding the influx of federal contracting money for a region is important for municipal planning. The site usaspending.gov is the official United States Government web site providing data on federal monetary awards [19]. This data can be mined to find information for specific localities individually. A visualization comparing these localities to one another is important for understanding how funding affects a given region. We present our visualization, developed in D3 [2], in Figure 1, consisting of coordinated multiple views [10] allowing the user to compare the funding for each locality in Hampton Roads as well as comparing the top 5 industries for each agency or locality depending on which options the user chooses for agency, year, or locality. A choropleth [9], [12] on the bottom right allows the user to compare funding values for each locality. A bar chart on the bottom left allows a user to compare funding values for each industry. A sparkline [18] on top shows the trend in funding. The user may interact with the visualization by altering the year, agency and locality to acquire more specific information.

Our region is Hampton Roads in the United States state of Virginia. As defined by the Hampton Roads Planning District Commission (HRPDC), Hampton Roads consists of the City of Norfolk, the City of Virginia Beach, the City of Chesapeake, the City of Portsmouth, the City of Suffolk, the City of Franklin, Isle of Wight County, Southampton County, Town of Smithfield, Surry County, James City County, the City of Williamsburg, Gloucester County, York County, the City of Newport News, the City of Hampton, and the City of Poquoson. Figure 2 shows the map as defined by the HRPDC. We use this definition of Hampton Roads in our visualization as well.

Existing visualizations of this data focus on the whole United States, or summarize data for types of funding. We seek to create a visualization that is more specific to federal contracting in the Hampton Roads region, such that the

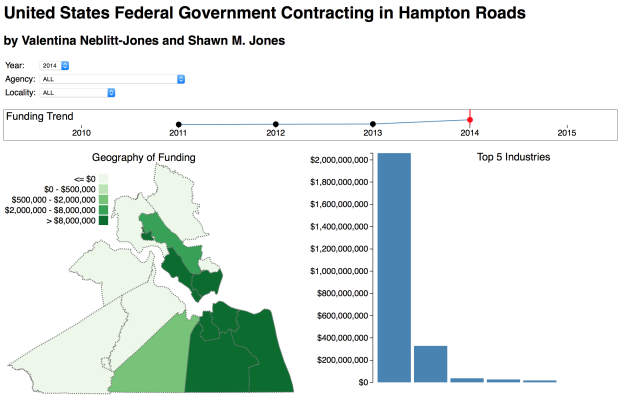

Fig. 1. The United States Federal Contracting in Hampton Roads visualization in its default state

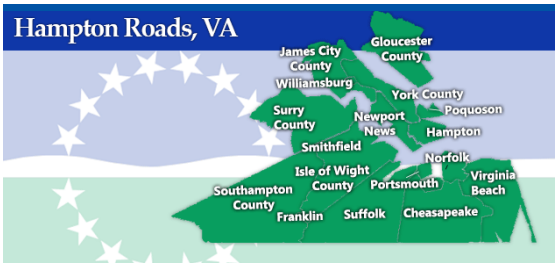

Fig. 2. Hampton Roads, as defined by the HRPDC

Hampton Roads Planning District Commission can utilize the visualization for decision making. We seek to help answer questions like the following:

- What industries are being funded in Hampton Roads?
- What is the trend of funding in Hampton Roads (e.g. increasing, decreasing, stable)?
- What knowledge areas should employees have to support work now and in the future in Hampton Roads?
- What localities are receiving funding from which agencies?

- *S. M. Jones is with the Department of Computer Science, Old Dominion University, Norfolk, VA, 23529. E-mail: sjone@cs.odu.edu*
- *V. Neblitt-Jones is with University Libraries, Old Dominion University, Norfolk, VA 23529. E-mail: vneblitt@odu.edu*

Manuscript received May 4, 2015.

- How much funding does each locality receive from which agency?
- Which agencies are providing funding to which industries in the region?

Using this visualization, we hope that the HRPDC can help guide decisions on such topics as funding for education and enticement for companies looking to move into the area.

Others, however, have tried to visualize this data before.

2 Related Work

The usaspending.gov site does contain several ways to visualize the existing data. These visualizations suffer from being too general for our needs, but we highlight them here in order to discuss for what they might be useful, and all of the ways in which they fail to meet our needs. Each visualization appears on a different web page with separate controls, hence they are displayed separately.

Figure 3 shows the summary of awards for the current year, broken down into contracts, grants, loans, and other financial assistance. It shows a bar chart, with each bar demonstrating the amount of money awarded for each category in the given year. A user can generate a tooltip by hovering the mouse pointer over each bar. This tooltip contains the value awarded for each award type, along with the number of transactions. This visualization aggregates data for the whole United States, and does not break down the data for each locality in Hampton Roads. We are also interested in contracting dollars, of which we are not interested. It is also fixed to the current year, rather than allowing the user to select a given year to see the data.

Figure 4 shows the summary of awards for each state for the current year. It is a choropleth coloring each state based on the amount of money awarded to each state. Just like the bar chart, a user can generate a tooltip by hovering over each state. This tooltip contains the value awarded to each state, along with the number of transactions. This visualization provides data for a given state, such as Virginia, getting us closer to Hampton Roads, but it aggregates data of all types of awards; we are only interested in contracting data. Just like the bar chart, it is fixed to the current year, not satisfying our need to see data from different years.

Figure 5 is a more interactive visualization, providing data more in line with what we are looking for. A user can use several drop down selectors to select the fiscal year, state, county, agency, ZIP code, congressional district, award type, amount, and even restrict the the visualization to a specific recipient. It provides information for the place of performance selected by the user, in this case the City of Virginia Beach. It displays the contracts awarded to the right of the map, allowing the user to view each contract awarded. The map provides squares pointing out where each contracting company is located. A tooltip can be generated by the user by clicking on each square. The tooltips contain information on the awarding agency, location, amount, and type of award. This satisfies some of our needs, but does not contrast the different regions of Hampton Roads. It also only displays information about prime award data, leaving out subcontractors, which are an important part of the work performed. We also see no trend data, nor do we get industry data for comparison.

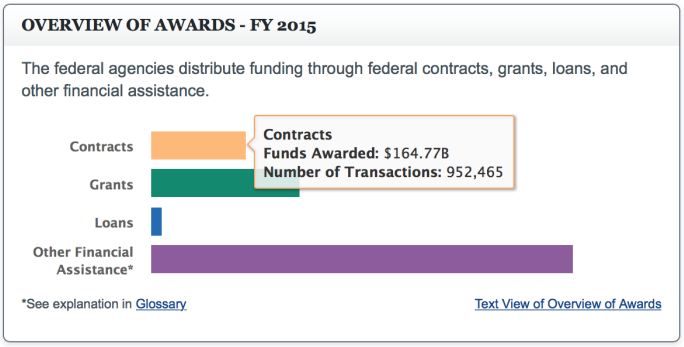

Fig. 3. USA Spending.gov visualization showing a summary of awards by year for the entire United States

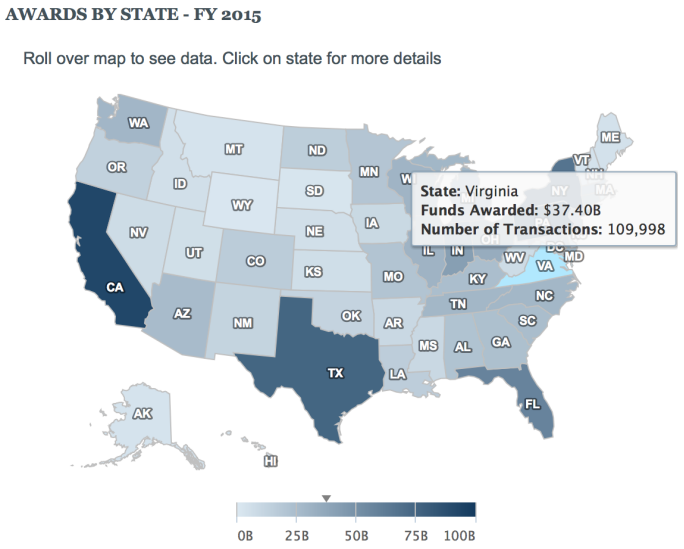

Fig. 4. USA Spending.gov visualization showing a summary of awards by state for the entire United States

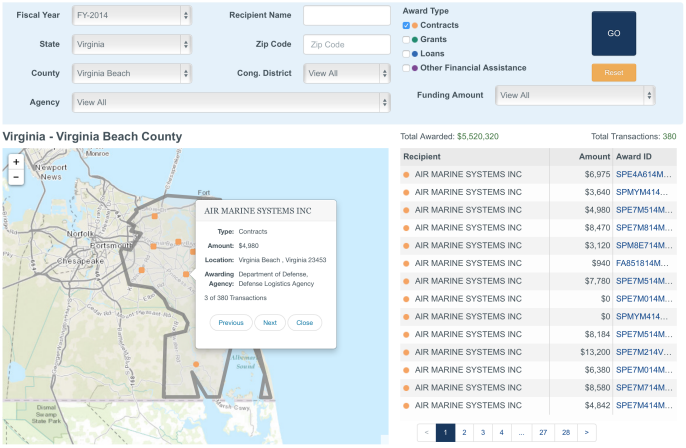

Fig. 5. USA Spending.gov visualization showing a summary of awards by locality for the entire United States

We want to show data for only Hampton Roads. We also want to display industry information to address the questions we wished to answer with our visualization. We want to allow the user to compare funding to individual localities based on agency and year. Finally, we want to show trends across all years based on agency and locality. The existing visualizations at usaspending.gov do not meet these needs.

3 Data Processing

Remember that our visualization consists of a choropleth, bar chart, and line chart. Our data for this visualization comes from two sources. The first source is usaspending.gov, providing the financial data for the visualization. The second source is the United States Census geographic data site, from which we needed shapefiles for rendering the choropleth in our visualization.

3.1 Financial Spending Data

Data from usaspending.gov can be downloaded in CSV format [16] using the form shown in Figure 6. Using this form, one can choose Prime or Sub Award data, the spending type (we wanted only contracts), the agency (we wanted all agency data), the fiscal year (we downloaded a single CSV for each year), the recipient state (we left this at all), place of performance state (we chose Virginia). select a date range (we left this at the default for the year selected), and chose the type of file to download (we chose CSV).

Not all data is available. As shown in Figure 7, some data combinations produced a download that could not be acquired via the web site. Fortunately, we were only concerned with contracts where the work was performed in Virginia. We tried to download data for Prime Award contracts in Virginia, but were told by the site that 0 records existed (in stark contrast to the results of their own visualizations), so we concentrated on Sub-Award data instead.

We downloaded the data for years 2006 - 2014 on April 1, 2015. Being relegated to Sub-Award data, we attempted to download records for all of the years available. We found that the number of records decreases precipitously, as shown in Figure 8. For example, we found it hard to believe that the entire state of Virginia only had 2 contracts awarded in 2006. These outliers forced us to consider any data from 2006 - 2010 as suspect, so we limited the year range of the data to only 2010 - 2014, ensuring that our visualization would feature the most accurate data we could find. The usaspending.gov site has changed in recent months, and has produced a number of complaints relating to the completeness of its data and usefulness of its visualizations [5].

The CSV files consisted of 101 fields, providing information on each contract awarded. From these files, and also using the provided data dictionary, we identified five useful fields for our endeavor. Table 1 lists these fields and how we mapped them during our data processing. We use the `subaward_principle_place_city` field to indicate the locality in which the work took place. We use the `subaward_principle_naics_desc` field for information on which industry the contract applies to. We use the `subaward_major_agency_name` field to identify

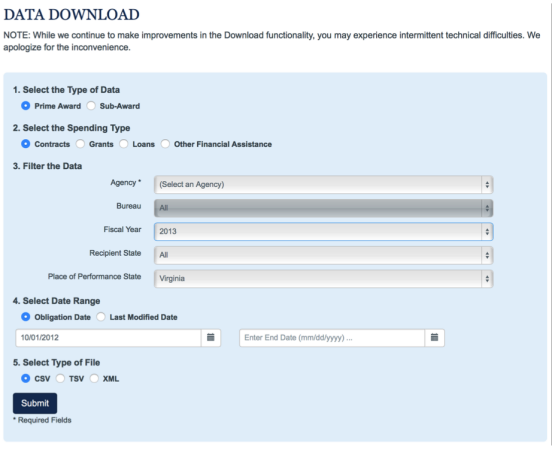

Fig. 6. For for downloading data from usaspending.gov

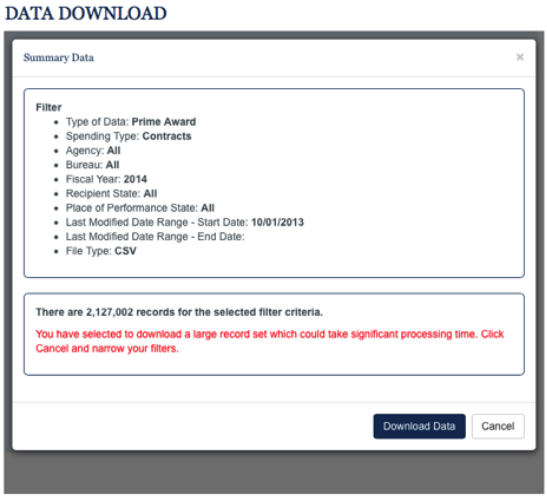

Fig. 7. Downloading large data sets is impossible via the web form at usaspending.gov

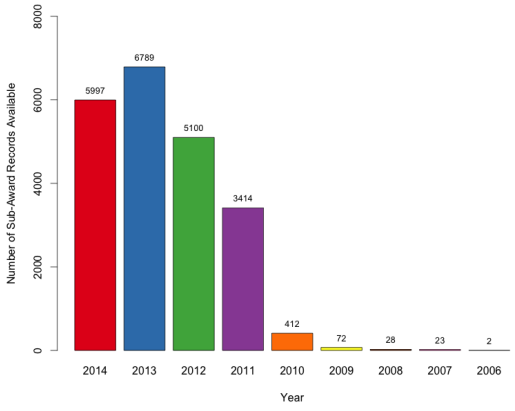

Fig. 8. Number of Sub-Award records available from usaspending.gov for Virginia for each year; note the precipitous falloff prior to 2011

TABLE 1
Mapping of fields from usaspending.gov into fields useful for our visualization

Original field name from usaspending.gov	Field Name for our visualization
subaward_principal_place_city	locality
subaward_principal_naics_desc	industry
subaward_federal_agency_name	agency
subaward_amount	funding
subaward_fiscal_year	year

the agency each contract applies to. Contracts are complex, and have many pieces. We had 3 agency fields to choose from, but `subaward_major_agency_name` indicates funding for the particular subcontractor performing work in the region, not necessarily the agency that controls the overall contract across the United States. We use `subaward_amount` to indicate the amount of funding for the contract. We chose this rather than the prime award amount because we are only interested in funding that ends up in Hampton Roads. We use `subaward_fiscal_year` to determine which data ends up in our visualization for a given year. Most of the time, this field matched the download, but there were a few cases where it failed to, meaning that the subaward was dispersed in a separate year than the contract was awarded. We were concerned about when the money ended up in Hampton Roads, so we use this field to ensure the data is correct.

Once we had acquired the CSV files, we then needed to clean the data and ensure that it only contained records for Hampton Roads. For this step we used OpenRefine [15], as shown in Figure 9. OpenRefine allowed us to correct the many spelling mistakes present for each entry in the `subaward_principle_place_city` field. It also allowed us to standardize capitalization for the different localities. Oddly, as one can see in Figure 9, the data contained addresses in this city field. We had to use resources, such as Google Maps, to identify which localities these addresses corresponded to. Thanks to this step, we were able to include more data about contracting in the region that would have been left out if we had just honored the existing city entries. Also, the file contained cities that do not exist in Virginia, such as San Diego, even though we requested a data download only containing records for Virginia. To resolve these inconsistencies, we used the `subaward_principle_place_zip` field to determine if the ZIP code is in Virginia, and, more importantly in Hampton Roads. The `subaward_federal_agency_name` field appeared to have a controlled vocabulary, but we converted the agency names so that they would support consistency throughout the visualization; for example, "DEPT OF DEFENSE" was changed "Dept of Defense, but also listed was "COMMERCE, DEPARTMENT OF", which we changed to "Dept of Commerce" for consistency of nomenclature for sorting and viewing. We left the controlled vocabulary field of `subaward_principle_naics_city` alone, seeing as we required no consistency from its values for our bar chart.

Once we had cleaned CSV files for each year, we then used a Python script, as shown in Figure 10 to compute all possible combinations of Agency, Year, and Locality, summing up the funding amounts for each. The Python

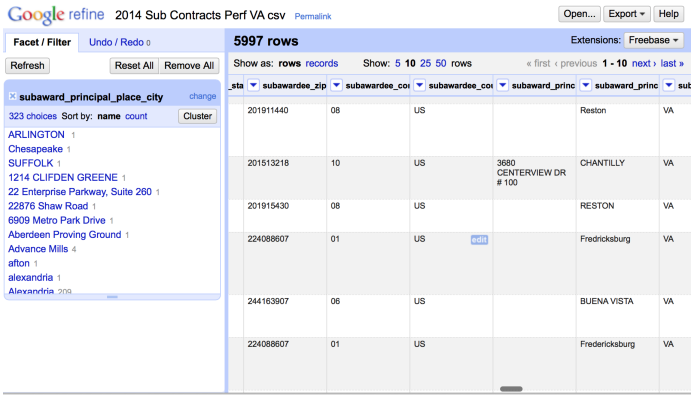

Fig. 9. Data from usaspending.gov loaded into OpenRefine (formerly Google Refine)

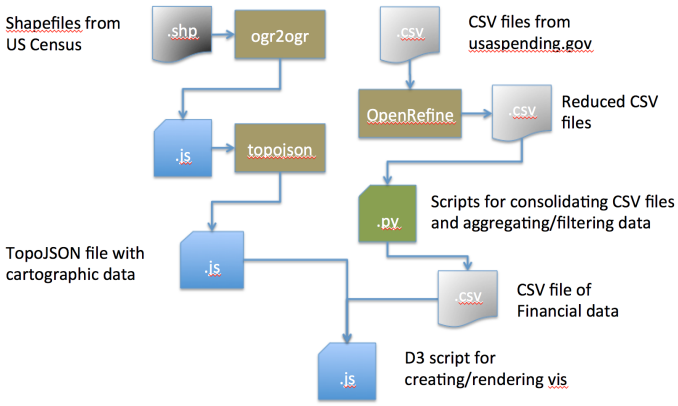

Fig. 10. Data flow to produce CSV and JSON map data for visualization

script calculates every combination of choice, and uses the value "ALL" in a field to assist in generating the data for each part of our visualization. Table 2 shows how the use of "ALL" in individual fields results in aggregated values and which part of the visualization to which they pertain. Any entry in our financial data file not containing "ALL" as a value in a particular field represents a filtering of data, allowing the user to get to specific funding values for year, agency, locality, and industry.

As shown in Figure 10, these steps produced half of the data needed for the visualization. For the other half, we needed to extract and use geographic data from the US Census.

3.2 Geographic Data

The HRPDC web site provides a series of maps for the Hampton Roads area, allowing the user to view information, such as the Regional Transportation Network, and the location of federal facilities [11]. They even provide a map of all of the localities in Hampton Roads. Unfortunately, all of these maps are provided in Portable Document Format, which, though well suited for printing, is not suitable for consumption by other tools, such as Tableau, or D3.

To produce a map for use with D3, we required the data to be in the TopoJSON format [1]. TopoJSON can be

TABLE 2
Use of the "ALL" value in various fields allows our vis to aggregate data for different combinations of user selections

"ALL" in field	Available User Selections	Part of Vis
none	locality, agency, year	Industry Bar Chart for specific industries
industry	locality, agency, year	Choropleth for user selection, including values for all industries; also Sparkline for user selection
locality	agency, year	Industry Bar Chart for all localities
agency	locality, year	Industry Bar Chart for all agencies
agency, locality	year	Industry Bar Chart for all localities and agencies
agency, locality, industry	year	Sparkline for all localities and agencies
locality, industry	year, agency	Sparkline for all localities, but selected agency
agency, industry	year, locality	Sparkline for single locality, all agencies

produced from GeoJSON format [4] using the `topojson` JavaScript command line tool [17] available via the npm package manager [14]. GeoJSON can be produced using the `ogr2ogr` tool provided by the Geospatial Data Abstraction Library (GDAL) [7]. The format consumed by `ogr2ogr` is the shapefile, a complex GIS format that contains metadata about each locality as well as vector graphics information necessary to render it [6].

The United States Census provides shapefiles for all counties and other localities within the United States. Upon attempting to utilize these shapefiles with D3, we were surprised to find that the localities in Hampton Roads did not display as we had expected. There appeared to be no water separating the Eastern Shore from the rest of Virginia, or many of the other parts of Virginia. We used the ERDAS IMAGINE tool to visualize the GIS data so we could verify that we were seeing a correct rendering. Figure 11 displays the Hampton Roads area, as surrounding localities, as rendered by ERDAS IMAGINE using this first shapefile.

Through this process discovered that the Census provides different resolutions of GIS data. Further research on the web site, brought us to a file at a resolution of 20m, which is an abbreviation of 1:20,000,000. We loaded this file into ERDAS IMAGINE and it rendered as shown in Figure 12. This was still considered unacceptable seeing as Gloucester County has a waterway separating it from York County, but this was not rendered at this resolution. The best resolution we could find was 500k, which is an abbreviation of 1:500,000, shown in Figure 13. The 500k shapefile is what is used in our choropleth.

Once we had the proper resolution, we needed to produce a map just containing the localities of Hampton Roads. Fortunately, the shapefile format contains metadata fields STATEFP for identify states and COUNTYFP for identifying counties. The STATEFP field contains globally unique values, but the COUNTYFP field does not, so both are needed to filter the data from the shapefile. We produced a GeoJSON file as shown in line 1 from Listing 1, filtering out just the localities from Hampton Roads. We then produced a TopoJSON file for use in D3 using line 2 from Listing 1.

```
1 ogr2ogr -f GeoJSON -where "STATEFP IN ( '51' ) AND
      COUNTYFP IN ( '550', '620', '073', '650',
     '093', '095', '700', '710', '735', '740',
     '175', '800', '810', '181', '830', '199' )"
     hrCounties.json cb_2013_us_county_500k.shp
2 topojson -o hrCounties-topo.json --id-property
     COUNTYFP --properties name=NAME -- hrCounties.
     json
```

Listing 1. Command for converting shapefile into GeoJSON format

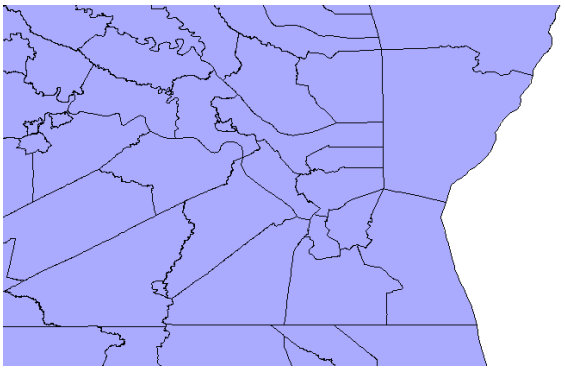

Fig. 11. Display of Hampton Roads using default United States Census shapefile using ERDAS IMAGINE

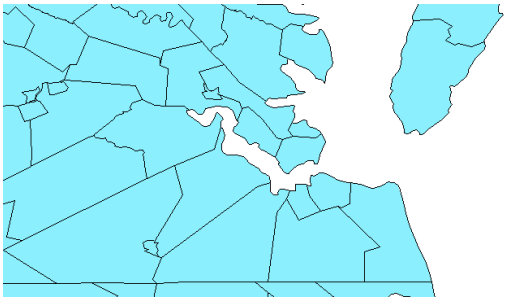

Fig. 12. Display of Hampton Roads using 20m United States Census shapefile using ERDAS IMAGINE

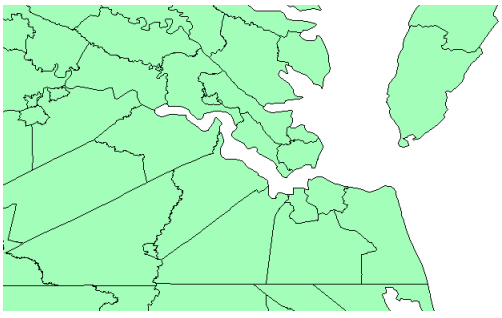

Fig. 13. Display of Hampton Roads using 500k United States Census shapefile using ERDAS IMAGINE

Once we had the geographic data, we could complete the left part of the data flow shown in Figure 10. With the

TABLE 3
Overview of User Actions

	Choropleth	**Bar Chart (Top 5)**	**Line Chart**
Default	Year: 2014 Agency: All Locality: All (sticky)	Year: 2014 Agency: All Locality: All	Year: all (sticky) Agency: All Locality: All
User selects locality A	Year: Selected Agency: Selected Locality: All (sticky)	Year: Selected Agency: Selected Locality: A	Year: All (sticky) Agency: Selected Locality: A
User selects agency B	Year: Selected Agency: B Locality: All (sticky)	Year: Selected Agency: B Locality: Selected	Year: All (sticky) Agency: B Locality: Selected
User selects year C	Year: C Agency: Selected Locality: All (sticky)	Year: C Agency: Selected Locality: Selected	Year: All (sticky) Agency: Selected Locality: Selected

TABLE 4
What: Attributes and Attribute Types

Attribute	**Attribute Type**
agency	categorical
funding	ordered quantitative - sequential
industry	categorical
locality	categorical
year	ordered ordinal

TABLE 5
Why: Targets

Target	**Result**
all data	trends, outliers
spatial data	shape
attributes	many: correlation, many: similarity

financial data and geographic data, we could then move on to producing the visualization.

4 Our Visualization

4.1 System Description

Our visualization contains three idioms - choropleth, sparkline, and bar chart. The sparkline shows the trend of spending across the four years. The bar chart shows a top five of industries to which the funding is directed. The choropleth shows the following:

- Hampton Roads localities (cities and counties as defined by the Hampton Roads Planning District Commission
- Single hue (greens) to show funding levels - darker greens (higher funding) to lightest green (lowest to negative funding)

Table 3 shows the possible user actions. "Sticky" means that the visualization stays the same on that attribute regardless of user selections.

4.2 What-Why-How Framework

Munzner describes a framework that designers should use as a guideline to develop visualizations [13]. Our design choices should map to parts of that what-why-how framework. This section is an analysis of our visualization using that framework.

4.2.1 What: Datasets

Our visualization uses data types attributes, item, and positions. We used dataset types - table and geometry. For the table our row represents a contract (item) and the columns represent attributes (e.g. locality, funding, year) of the contract. The choropleth uses the geometry dataset type with the shapefiles providing spatial position information, but also utilizes the table information (i.e. attributes) for encoding. As mentioned in previous sections, the unclean nature of the data provided by USASpending.gov requires manual pre-processing before it can be used by our visualization. Therefore the dataset availability is static instead of dynamic.

4.2.2 What: Attributes

Our attributes are categorical, quantitative ordered, and sequential. Refer to Table 4 to see a breakdown of the attributes mapped to attribute type. Categorical is the most common along with two different types of ordered data.

4.2.3 Why: Actions & Abstract Tasks

Of the three types of actions (analyze, query, search) our visualization utilizes analyze and query. It supports the analyze-consume sub-functions of discover, present and enjoy. The viewer can use it to discover or enjoy contracting information about the Hampton Roads area. After either of these actions, the viewer can use the visualization to present viewpoints and suggest a course of action based on the visualized data.

A viewer can also query the visualization. It supports identification, comparison and summarizing actions. The sparkline allows the viewer to see funding trends. The bar chart and choropleth supports comparing year to year. The view can also identify a single piece of information based on choices in the dropdown (e.g. top industry in 2012 in Virginia Beach).

This is a specialized vis tool that supports the following abstract tasks:

- Create or support hypotheses about government funding in the Hampton Roads area
- Instruct interested parties on the economy of the area
- Support planning and decision-making for housing, education, commercial endeavors, etc.

4.2.4 Why: Targets

Table 5 outlines the targets in our visualization. Viewers can see trends and outliers in our visualization. Using the "ALL" field as described in Table 2, one can aggregate data and see trends. TopoJSON data for the choropleth is spatial data resulting in the shape of each locality. Coupled with an industry and/or agency, the funding attribute can be used to see correlation or similarity. As an example of an outlier, Gloucester in 2014 has a negative funding amount and shipbuilding is usually but not always one of the top industries.

4.2.5 How: Encode

Shown in Figure 14, the choropleth uses luminance with ColorBrewer scheme 5-class Greens [8]. It is colorblind safe, but not print friendly nor photocopy safe. This choropleth has the following colors mapped to the following five categories of funding:

- for $\leq \$0$
- for $\$0 - \$500,000$
- for $\$500,000 - \$2,000,000$
- for $\$2,000,000 - \$8,000,000$
- for $> \$8,000,000$

Our choropleth also uses the given geometry of map shapes.

The bar chart shown in Figure 15 relies on size (i.e. length of the bar) for encoding therefore the color is uniform. Funding is expressed via vertical position. The bars are aligned on the x-axis to provide a uniform starting point improving length/area judgments. They are also ordered highest to lowest and separated along the x-axis.

The sparkline shown in Figure 16 is using angle to indicate funding trend. Is the line moving up, down or staying level? The dot is red to indicate the selected year. All years for which we have data are shown, but the colored dot provides the user feedback about the year chosen since it does change other parts of the visualization. Vertical position expresses funding and it is ordered by year.

4.2.6 How: Manipulate/Facet/Reduce

The visualization provides the following three options, shown in Figure 17, for manipulation or change via drop-down lists:

- Year
- Agency
- Locality

Choosing any of these options reduces the dataset shown. Choosing the All option for any of these aggregates the data. The idioms are juxtaposed - choropleth and bar chart are side by side and sparkline is above.

The choropleth supports selection with a tooltip that provides the exact dollar amount of funding when the user hovers over a locality with their mouse.

Initial analysis using Tableau [3] with the industry data demonstrated that industry names were too long and most of them were empty. There were more than 72 options and less than 10 of them were utilized at any given time. Therefore, the decision was made to show only the top five

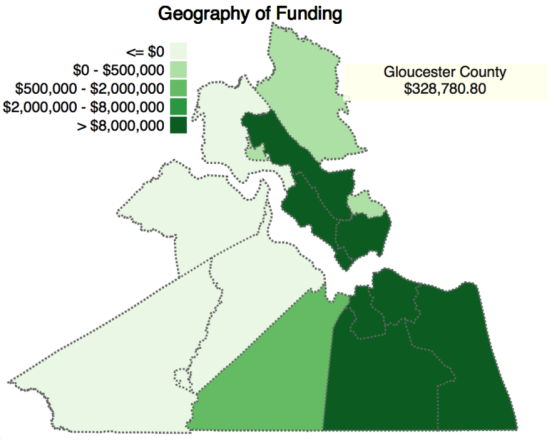

Fig. 14. The choropleth portion of our visualization

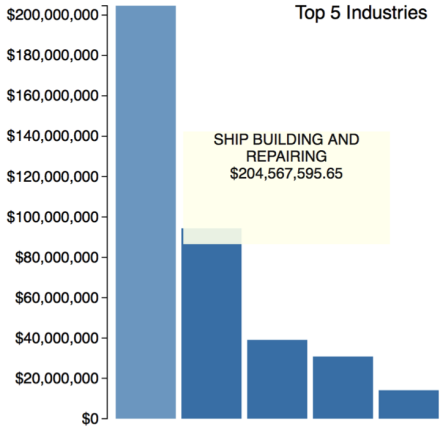

Fig. 15. The bar chart portion of our visualization

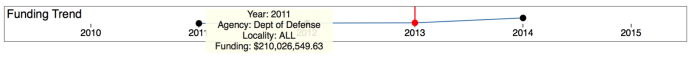

Fig. 16. The sparkline portion of our visualization

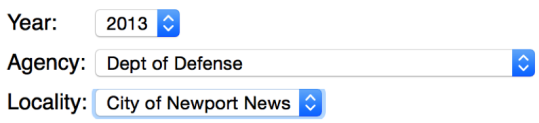

Fig. 17. Our visualization selections of year, agency, and locality

categories in the bar chart. Selection of a bar via mouse over shows the industry name and exact funding amount.

The line of the sparkline changes based on agency and locality selection. Selection of a circle provides tooltip showing year, agency, locality and total year funding.

5 Conclusion

Using data from usaspending.gov and the United States Census, we have been able to produce a visualization that displays federal contracting information about the localities specific to Hampton Roads. Our visualization consists of a connected sparkline, choropleth, and bar chart. A user may manipulate these using 3 drop down controls, opting to see funding for a specific year, agency, and/or locality. The sparkline shows the trend in federal contracting funding based on the agency and locality chosen. The bar chart shows the top 5 industries funded with these federal awards, based on the options chosen by the user. The choropleth allows the user to compare the amount of funding different localities receive based on their awards, also responding to the options chosen by the user.

We have shown how the data from usaspending.gov can be aggregated and converted into a form that allows users to query it based on agency, year, and/or locality. We have shown how data from the United States Census can be processed to produce spatial data just for the Hampton Roads area. We have discussed how the issues with the resolution of this data when used specifically for the localities in the Hampton Roads region, favoring the resolution of 1:500,000 as the best fit available. We posit that other regions may benefit from higher resolution data, just like Hampton Roads. We have discussed our struggles converting the data from usaspending.gov, and identified issues with its completeness.

We have discussed out our visualization fits into Munzner's What-Why-How Framework, showing how our design choices reflect the intentions of that framework. Our data consists of categorical, ordinal, quantitative and spatial attributes. Our visualization is a specialized tool for answering questions about federal contracting awards in Hampton Roads, supporting the actions of analyze and query. The viewer can use it to discover new information and derive new hypothesis, while also enjoying the data in a casual manner. Viewers can also view trends and outliers, while also comparing funding between localities and industries.

Using our visualization, the HRPDC and others can answer questions about federal contracting for industries and localities, allowing them to further make decisions about funding education or courting new businesses and industries.

In the future, we would like to add support for subagencies, allowing the user to filter the data further. We would also like to incorporate information about contracting companies in addition to industries. Once usaspending.gov has more complete data, we would not only like to load data from prior to 2011, but also would like to incorporate prime contracting data as well. We also would like to improve the visualization so that it supports a wider variety of browsers, to include mobile platforms. Finally, we want to rework the existing data handling process so that updated data can be acquired from usaspending.gov and incorporated into the visualization for more up to date information.

6 Final Thoughts

In working on this visualization, we came to the conclusion that data processing was the most difficult part. On the spectrum of data complexity, from unstructured to fully structured, the data was mostly structured, which made working with OpenRefine easy. Unfortunately, our confidence in the data from usaspending.gov is not as high as we would like, considering some of the records appeared to come from cities outside of Virginia, even though we explicitly chose Virginia.

We did, however, find that Munzner's What-Why-How framework was useful in making us consider the aspects of the visualization in a general way before we looking at domain-specific applications.

Acknowledgments

The authors would like to thank James Clary, Senior Economist for the Hampton Roads Planning District Commission, for an interesting project idea. We would also like to thank Topher Lawton, Science Reference Services Librarian for Old Dominion University, for introducing us to OpenRefine.

References

[1] M. Bostock and C. Metcalf. The TopoJSON Format Specification. https://github.com/topojson/topojson-specification/blob/master/README.md, 2013.
[2] M. Bostock, V. Ogievetsky, and J. Heer. D^3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, Dec 2011.
[3] Business Intelligence and Analytics — Tableau Software. http://www.tableau.com, 2015.
[4] H. Butler, M. Daly, A. Doyle, S. Gillies, T. Schaub, and C. Schmidt. The GeoJSON Format Specification. http://geojson.org/geojson-spec.html, 2008.
[5] C. S. Clark. Refreshed usaspending website irks some transparency advocates. *Government Executive*, page 1, 2015.
[6] Environmental Systems Research Institute, Inc. ESRI Shapefile Technical Description. http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf, 1998.
[7] GDAL - Geospatial Data Abstraction Library. http://www.gdal.org/index.html.
[8] M. Harrower and C. A. Brewer. Colorbrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37, 2003.
[9] J. Heer, M. Bostock, and V. Ogievetsky. A Tour Through the Visualization Zoo. *Communications of the ACM*, 53(6):59–67, June 2010.
[10] J. Heer and B. Shneiderman. Interactive Dynamics for Visual Analysis. *Communications of the ACM*, 55(4):45–54, Apr. 2012.
[11] HRPDC: Maps and GIS. http://www.hrpdcva.gov/page/maps-and-gis/.
[12] A. MacEachren. *How Maps Work: Representation, Visualization, and Design*. Guilford Press, 2004.
[13] T. Munzner. *Visualization Analysis & Design*. CRC Press, Boca Raton, FL, USA, 2014.
[14] npm. https://www.npmjs.com.
[15] OpenRefine. http://openrefine.org.
[16] Y. Shafranovich. RFC 4180: Common Format and MIME Type for Comma-Separated Values (CSV) Files. https://tools.ietf.org/html/rfc4180, 2005.
[17] Topojson. https://github.com/mbostock/topojson, 2015.
[18] E. R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT, USA, 2001.
[19] USA Spending.gov. https://www.usaspending.gov/Pages/Default.aspx.

Shawn M. Jones is a computer scientist with Space and Naval Warfare Systems Center Atlantic in Norfolk, VA. He completed his B.S. and M.S. in computer science from Old Dominion University. He is a Ph.D. student in computer science at Old Dominion University.

Valentina Neblitt-Jones is the Head of Systems Development at Old Dominion University Libraries in Norfolk, VA. She completed her B.A. in international studies at Virginia Tech, B.S. in computer science at Old Dominion University, and M.S in library science at Florida State University.