

Spring 5-4-2017

Using Predictive Analysis for Meals on Wheels

Dylan Snyder

Valparaiso University, dylan.snyder@valpo.edu

Adali Johnson

Valparaiso University, adi.johnson@valpo.edu

Omar Wasim

Valparaiso University, omar.wasim@valpo.edu

Jason Henry

Valparaiso University, jason.henry@valpo.edu

Follow this and additional works at: <https://scholar.valpo.edu/cus>

Recommended Citation

Snyder, Dylan; Johnson, Adali; Wasim, Omar; and Henry, Jason, "Using Predictive Analysis for Meals on Wheels" (2017). *Symposium on Undergraduate Research and Creative Expression (SOURCE)*. 628.

<https://scholar.valpo.edu/cus/628>

This Poster Presentation is brought to you for free and open access by the Office of Sponsored and Undergraduate Research at ValpoScholar. It has been accepted for inclusion in Symposium on Undergraduate Research and Creative Expression (SOURCE) by an authorized administrator of ValpoScholar. For more information, please contact a ValpoScholar staff member at scholar@valpo.edu.

Using Predictive Analysis for Meals on Wheels

Dylan Snyder, Adali Johnson, Jason Henry, Omar Wassim, Prof. Karl Schmitt

Valparaiso University



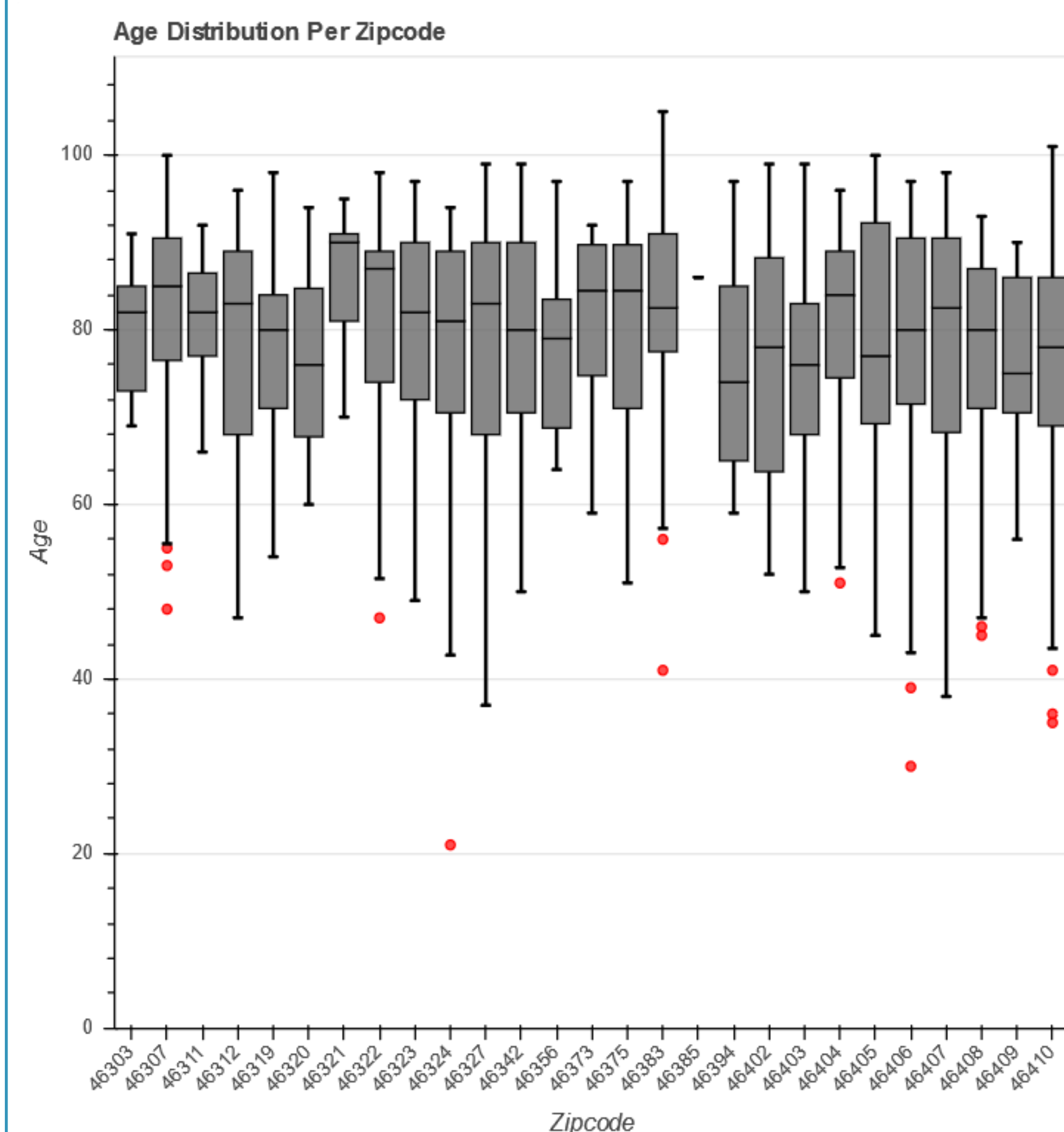
ABSTRACT

We are living in the age of data. With the vast multitude of information flowing around us, we have an urgency to keep it, analyze it, and pull what we can from it. For our project, we are helping Meals on Wheels of Northwest Indiana do just that with the information they have. In our project, we will be helping the Meals on Wheels team sift through their data to find the characteristics of their clients, especially those that have signed up to pay for their food, so that they may better market to those individuals. We will be doing this by using Python scripts to clean the data, explore that data with manual review, and then create visualizations and statistics using Bokeh graphing tool, Excel, and other various tools.

OBJECTIVES

Our objects are as followed:

- Clean and organize the data that has been given to us.
- Interpret and explore the data using EDR (exploratory data analysis)
- Do statistical analysis on the data to provide meaningful information on the raw data
- Create visualizations to better interpret that data
- Find the characteristics of the clients, especially the private pay clients, for marketing purposes.



METHODS

For the exploratory data analysis in this project, we used:

- Manual review
- Orange Datamining Tool

For data cleaning and correcting, we used

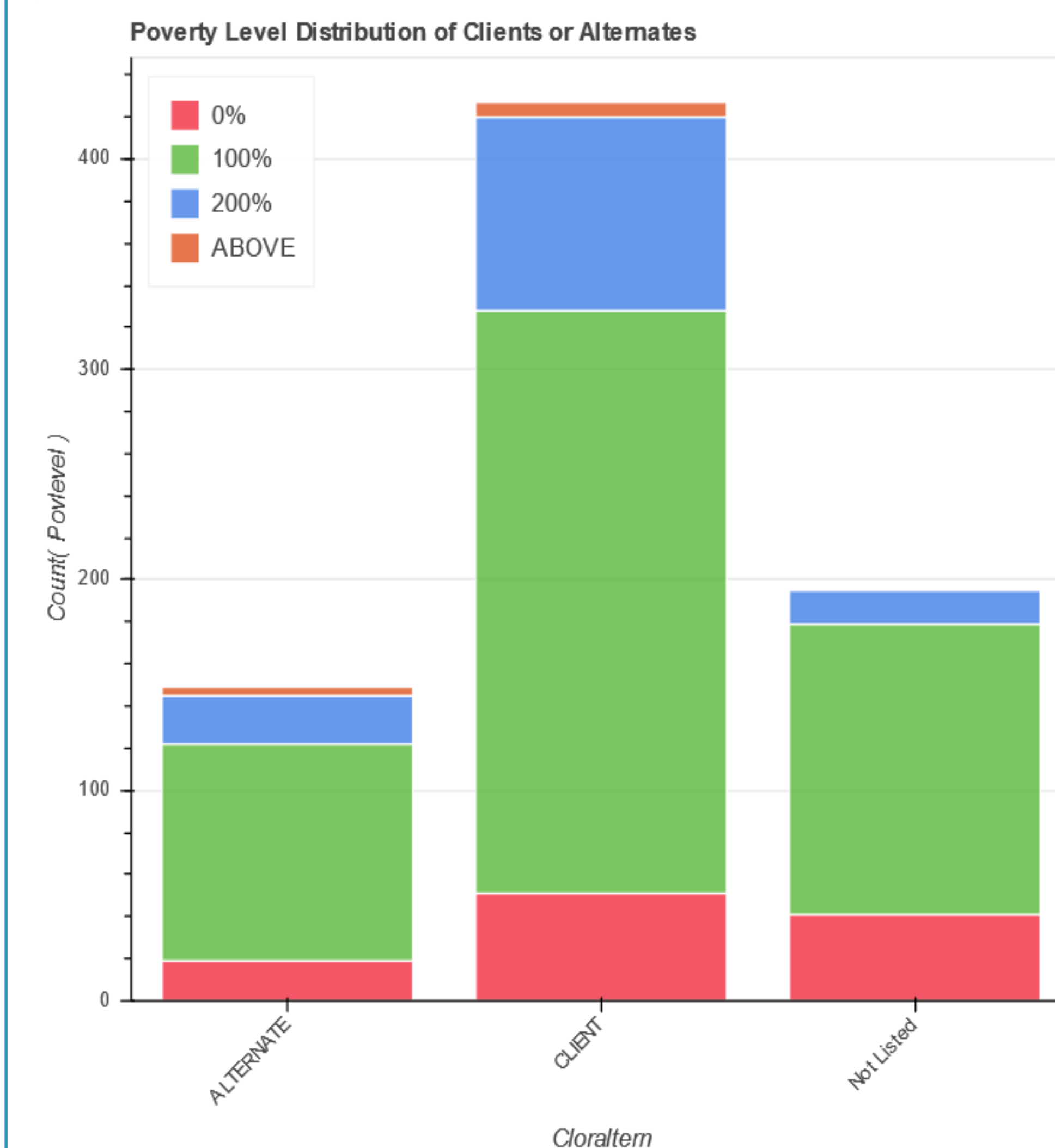
- Manual revision
- Microsoft Excel tools
- Multiple student-written Python 3 programming scripts

For data analysis and visualization, we used:

- Python 3
- Bokeh (Python visualization tool)
- Pandas (Python data analysis tool)
- Microsoft Excel

Here are the steps we used to clean the data. Each step has an associated script:

- Split data into two files: current clients, terminated clients.
- Removed the invalid dates from the data.
- Changed all the null level values in the povlevel column to a zero.
- Found spelling mistakes in cities by printing out all unique city names.
- Replaced all invalid city names with real city names.
- Went through the birthyear column and calculated each person's age



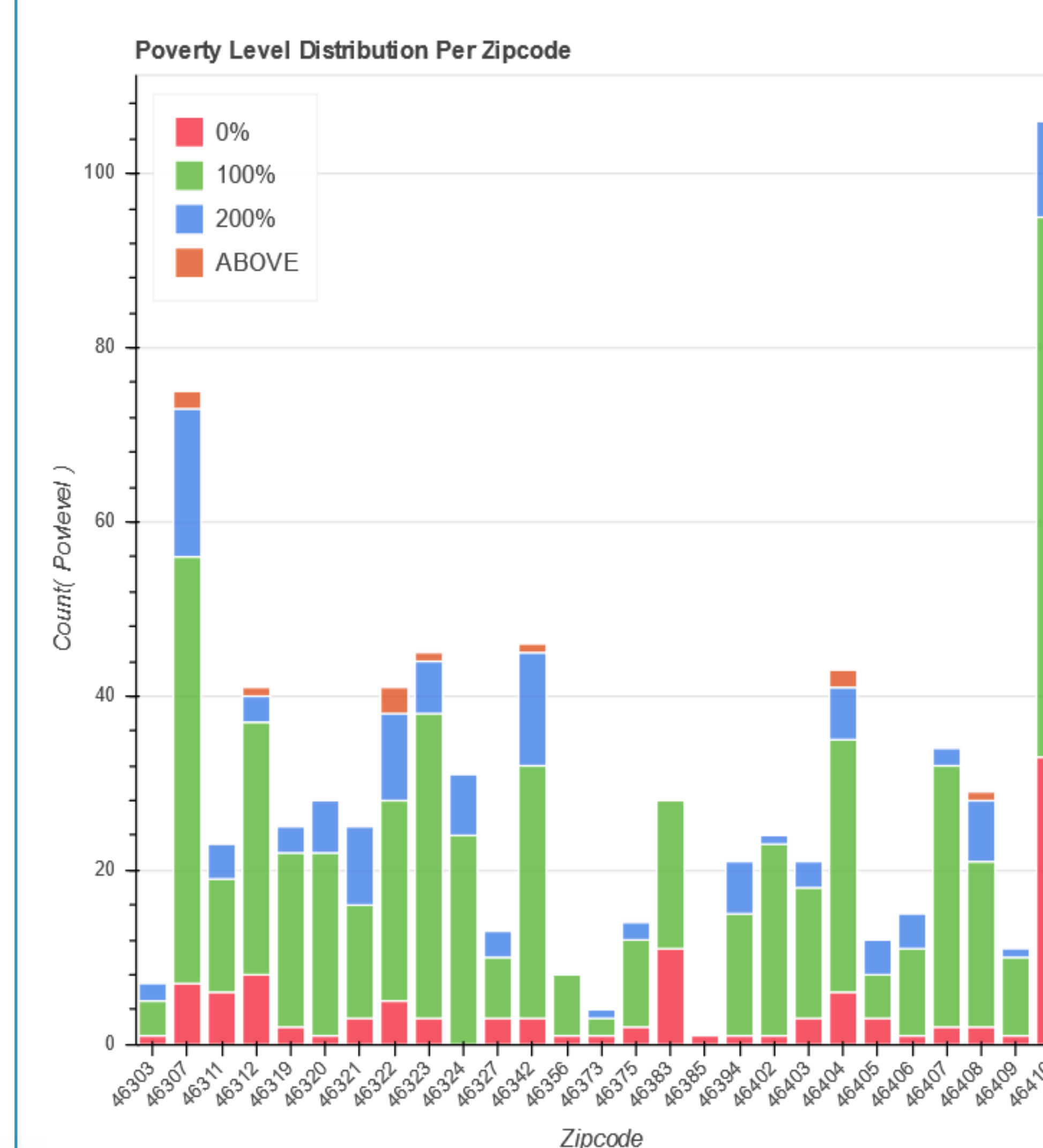
RESULTS

For our results, we found a few things. The first that we found is that the Meals on Wheels client base is mostly elderly people, in the 60+ range. We also found that a majority of people were in the private pay category. The private pay category had people of all income levels, with a majority at the poverty line. However, the private pay client category also had the most amount of people that had income levels above the poverty line.

The location data also provided interesting results. The majority of people that participate in Meals on Wheels are centered in the zip codes 46410 and 46307. The people in these counties are also typically at or at double the poverty line.

CONCLUSIONS

From these results, we can draw some conclusions. One is that the Meals on Wheels team would probably be best suited to advertising in the areas with the highest concentration of lower-middle class elderly peoples. The areas with a high population concentration also seem to have the most amount of people that are available to pay for all of or part of their meal.



FUTURE IMPROVEMENTS

As with every project, there are always many things that could have been done differently or better. What we believe could be done in the future for this project is:

- Taken in other data about population from third party sources
- Done more analysis on the full dataset, which includes the people that are deceased or otherwise terminated, rather than just the people currently in the program
- Improved documentation quality

REFERENCES

Special thanks to Meals on Wheels marketing manager of Northwest Indiana Jodi Bella.

- Python 3 programming languages. <https://www.python.org/>.
- Bokeh Visualization. <http://bokeh.pydata.org/en/latest/>

