# ALPHA RELEASE

**Submitted by: Divya Shrivastava & Kirubel Tadesse**

## 1. Introduction

The visualization tool provides an interactive visualization of websites and their performance parameters like speed index, load time, etc. which affects their rankings.

We used the capabilities of Bokeh - a powerful Python library, to develop our visualization tool. The tool provides an interactive interface for the users to modify the parameter values to visualize the performance of the websites. The tool bar on the right side includes panning, zooming, selection, and plot saving abilities along with the display of the plot. On the left side, we added a tool bar to select/ deselect parameters of the choice of interest. These features make the tool very interactive for the user. This tool is very helpful when we need to investigate the data and they are configurable as well. The main concept used in developing our tool is that graphs are built up one layer at a time. First, we start by creating a figure and then we add glyph elements to it layer by layer. A screenshot of our visualization tool is provided in Figure 1.

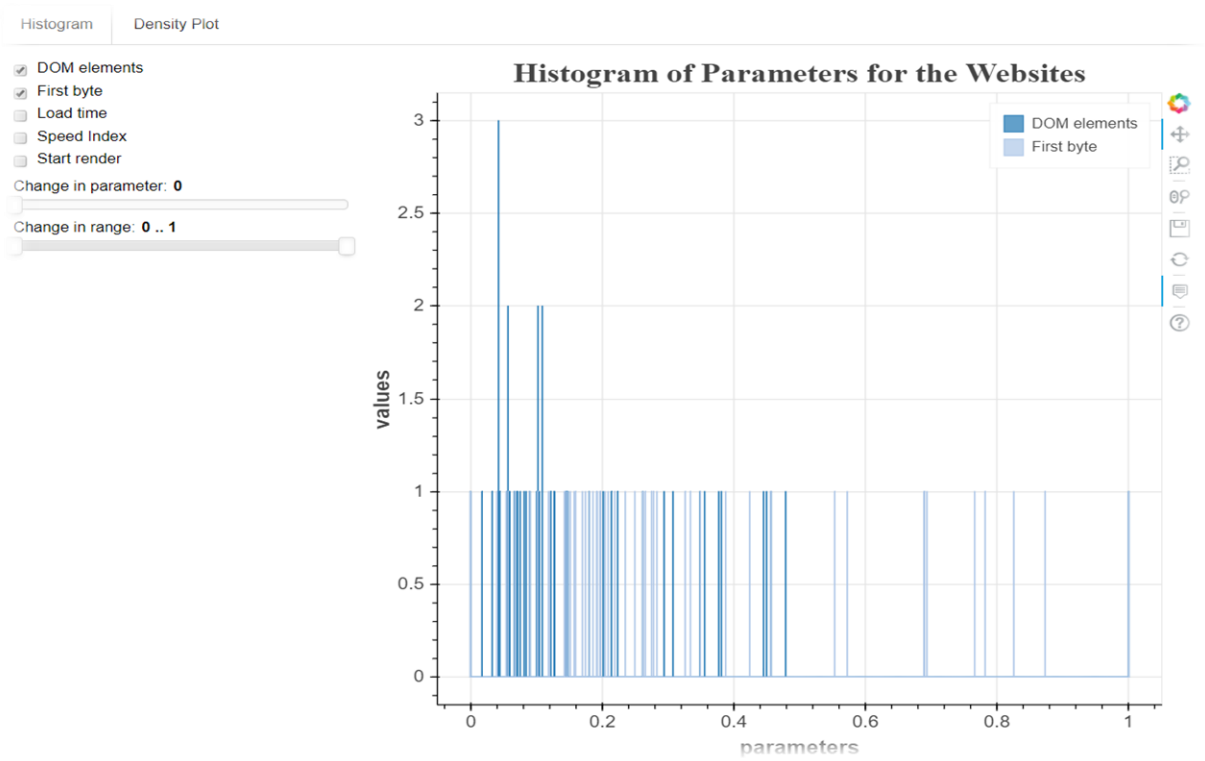


Figure 1

In our design document, we proposed to develop a model (Figure 2) to visualize the data. Therefore, for alpha release, our first step was to design the model to develop our interactive visualization tool. After the design of our model, we initiated the development of the visualization phase.

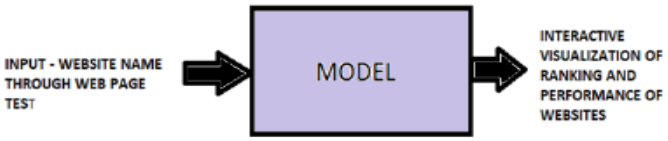


Figure 2

## 2. Current Capabilities of the Tool

Currently, our tool runs on Bokeh server which hosts the application on a local host. Our visualization tool currently has a dashboard having two different views of the websites and their performance parameters – histogram plot and density plot.

### 2.1 Histogram view

We developed a color-coded histogram view of each websites and their performance parameters. When we hover over any bar in the histogram, we would find the parameter values for each website. Currently, we are analyzing each of the parameters of top 49 websites from Alexa. For this visualization, the first step was to read the data from the csv file to a data frame. Then we analyze the data by calculating mean, standard deviation, min, max, 25%, 50% and 75%. Once we know these values for each of the parameters, we decide our range and bin width to create the histogram. We specify the bottom, top, left, and right edges of each bar the quad glyph of Bokeh.

A screenshot of the histogram plot is provided in Figure 1.

### 2.2 Density view

The density plot provides a color-coded gaussian distribution over a range for the each of the performance parameters. The user would have an interactive interface to select/deselect any of the parameters from the toolbar provided in the side. The parameters are color coded to differentiate between them. The range of the parameter can be varied with the help of a scale provided in the toolbar. The bandwidth scale can also be modified by the user and the user can click on the ‘choose bandwidth’ button to reload the new values.

A screenshot of the density plot is provided in Figure 3.

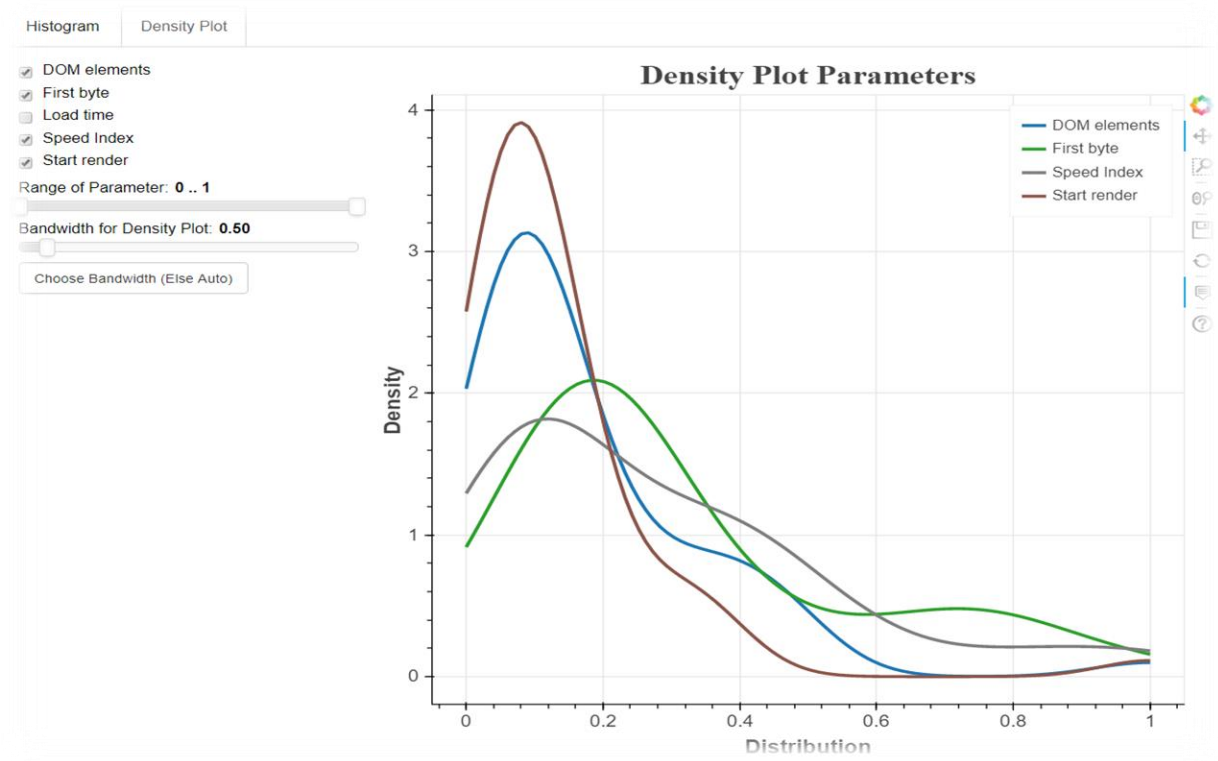


Figure 3

### 3. Interactive features

Interactive feature of our tool involves having inspectors to hover over data to visualize their values. These actions do not alter the data displayed. They help in investigating the data in more detail. The Hover Tool in Bokeh is a tool tip which is very useful inspector appearing when a user hovers the mouse over data points. Other interactive features include scaling, zoom- in, zoom-out, save and reload button to investigate the data.

### 4. Future scope

After this release, we aim to add more tabs on our dashboard providing a variety of views to visualize the data. The user will have the flexibility to select any tab in the dashboard to have an interactive visualization of the data. We also plan to integrate our designed model completely with the front-end visualization. Finally, with the help of the model integration, we aim to build a predictive model to visualize how the change in the value of the parameters would affect the performance of the website.

### 5. References

Please find below the YouTube link having the demonstration of our visualization tool- https://youtu.be/qSHbC7QEQdI