About:

GitHub is one of the most popular sites for online code repositories. With a large user base GitHub generates lot of activity. This activity, in the form of commits, contains the relations between repositories and contributors. With the use of appropriate visualization techniques we can effectively map the relation between projects and contributors onto visual graphs. These graphs can reveal some interesting trends and insights that are not apparent using text and tables. This has motivated us to create visualizations using commit activity on public repositories of GitHub.

**Existing Visualizations:**

GitHub itself has some interesting graphs that map the activity occurring on each repository. Most of these graphs, however, are mostly standalone and provide limited interactivity. Except the contributor view, there is a lack of connected views. This limits the usefulness of these graphs. We decided to answers these issues by interconnecting all graphs and concurrently updating them based on user interactions. This allowed us to maintain flow consistency and also provides more information to the user.

Design:

The data generated on GitHub is essentially temporal in nature. However, we also wanted to provide aggregate information in a single view. For this we chose tree-maps that effectively incorporated the data on a single page. The relation between users and repositories was mapped using overlapping bar-charts. Finally a heat map was used to provide detailed activity of each contributor.

Ideation and initial sketches:

Prototype:

The final prototype implemented using D3.js. The data is from public repositories of company called Balanced.

Initially the graphs show all of the commit activity. When the user selects a time period the graphs zoom in to accommodate the selected data.

Video