**Programming languages:**

In 'A Study of External Community Contribution to open-source Projects on GitHub; This article USES the boxplot analysis of data, of which this article use the data features for developers that use the programming language, in our 2 data set as the **‘language’**, the data reflects the language used by the project development, and whether the developer as the core members, including developers can be divided into three kinds of article, However, we provide only one feature in the data set to determine whether the developer is a core member, namely **'core\_member'.**

In the article 'An Insight into the Pull Requests of GitHub', a bar chart was used to measure the number of developers using different languages and whether the merge was successful. In our data set, we have the **"“merged\_or\_not‘"** feature to determine whether the request was successfully merged.

In the article 'Acceptance Factors of Pull Requests inopen-source Projects', association rules are used to quantify the influence of project language on Pull Requests, so as to calculate the correlation between project language and Pull Requests. And use a bar chart to display.

**Popularity of project:**

The **"open\_issue\_num"** and **"fork\_num"** features of the dataset can be used to determine the popularity of items, and the relationship between the number of branches and the pull success rate of each item can be determined by the number of branches and the success of the pull. You can also use the **'open\_issue\_num'** feature to determine whether the repository is more likely to accept pull requests if it has more problems.