



Project: Visualization on Technologies

CSE564 - Visualisation Project Proposal

Team: 66

Ranjith Reddy Bommidi

114241300

Manish Reddy Vadala

114190006

Instructor: Klaus Mueller

Department of Computer Science

State University of New York at Stony Brook

April 2021

1 Background

The data that we chose has diverse representation and it's one of the largest data of developer community that stack overflow has collecting over the past decade. Every year stackoverflow conducts survey and collects data all around the world. In the year 2020 stack overflow has collected data from over 65000 people in the developer community. This diverse data contains the information given by people from different parts of the world with different ethnic groups, with different levels of experience, their interests in coding languages, their salaries, level of education and most popular/desired technologies among the developer community and many other things. In total the number of attributes that are present in each dataset is 60.

Dataset: <https://insights.stackoverflow.com/survey>

1.1 Why this Dataset?

We have been in love with technology, since childhood. Every tech enthusiast would like to know the current world stats about the technologies in the industry and some insight over that. Here in our project we would like to provide correlations between our attributes and meaningful insights from the stats. Hoping this might help many upcoming developers including myself.

2 Problem

- Developers in the tech industry need to keep themselves updated on latest technologies so they might have some insights about what technologies are used across different parts of the globe.
- People might not have enough insights about the technologies used in the evolving market and the level of expertise they need, pay-scale they get for the positions they apply.
- People should also know how technology trends have changed over the years, because learning a stable technology might help people to survive at least few years.

3 Approach

3.1 Why do we need to visualise the data?

With the current number of technologies that are in market, we need to get enough insights on technologies and other attributes. So with the help of data visualisation tools in this project we will provide meaningful insights about the latest market trends that helps developers to choose their path and many other relevant things.

3.2 What are we planning to do with the data?

We are Planning to use the datasets of the past 4 years and show how the trends have changed over the years, and show the important trends in each and

every year, if possible we would also like to predict the trend for 2021. More information regarding the dashboard and it's content is mentioned below.

4 What does the Dashboard Contain?

- **Barchart:** First part of our Dashboard is barchart which shows the selected attribute. This shows the stats of many attributes, providing meaningful insight to the user. Selecting the attribute as Technology shows the user which technologies has been used by most of the people in the developer community.
- **WorldMap:** Second part of our Dashboard shows the worldmap. Here we are going to show the difference in countries in terms of attributes. Such as average pay-scale of developers, average age of a developer, average number of hours they put in week, etc from country to country, from year to year.
- **PCP Plot:** Third part of our Dashboard shows the PCP plot. Here we are going to show the Correlation between various features like technologies, pay-scale, level of experience, work satisfaction e.t.c. From this we can deduce some interesting visualizations like which set of people are more satisfied with their work.
- **Scatter Plot:** Last part of our Dashboard shows the Scatter Plot. Here we are going to show the correlation between two attributes and would also like to add a third attribute as a filter. Ex: Technologies vs salaries, Roles vs Salaries, Roles vs Experience, we will also add a third attribute such as country as the filter.

5 What are we getting out of this Visualization?

The Visualizations which we show helps us to know about the developers around the world and many more attributes about them.

- What technologies are majorly used around the world?
 - From the respondents response on their current technologies use, we can deduce the technologies used around the world. This can be shown in Barchart.
- What technologies are most loved around the world?
 - From the respondents response on their desired technologies, we can deduce the technologies they love to learn by majority. This can also be shown in Barchart.
- What set of technologies are getting highest salaries?
 - From the attributes Technologies and their salaries, we can find which set of technologies are getting paid more. This can be shown in ScatterPlot.

- Which set of people are more satisfied with their work?
 - From multiple attributes such as work hours, technologies and work satisfaction, we can find out which set of people are more satisfied with their work. We can show this in the PCP Plot.
- Which roles are the highest paid around the world?
 - From the developer type attribute and the salary we can deduce which roles are highest paid around the world. This can also be shown in ScatterPlot.
- What are the average salaries across the world?
 - From the attributes country and the salaries, we can deduce the average salaries from country to country. This can be shown in the Worldmap.
- How the trends are changing over the years?
 - From the datasets of the past 4 years, we will analyze the trend for major attributes for each and every year.

6 Rough outline of the dashboard

The following image roughly shows about our dashboard which we are planning present in the final project.

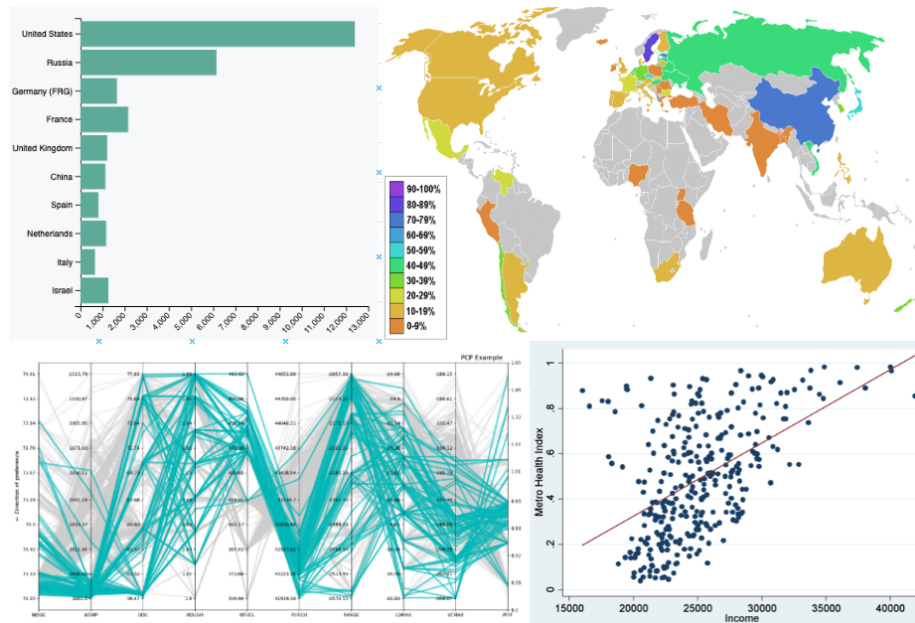


Figure 1: The Researcher and Supervisor