William Pembleton

<https://www.kaggle.com/stackoverflow/stack-overflow-2018-developer-survey#survey_results_public.csv>

Abstract

Background

The data I decided to use for this project is from the website Stack Overflow. Stack Overflow is a online community for developers to learn, share their knowledge, and build their careers. It’s essentially the Q&A website for developers similar to something like Yahoo Answers. Stack Overflow also does surveys asking their users to answer questions about their job. They asked questions like “How satisfied are you with your current Job?”, “On a typical day, how much time do you spend outside?” and “What is your current gross salary (before taxes and deductions)”.

Description of ID3 and how it works

Methodological

I cut some thing out of the dataset because Apriori couldn’t handle it. They asked some questions of email and like the like and asked participants to rank their feelings of the questions on a Likert scale (1-5 scale).

I also separated out lists in my dataset into their own individual column (for instance in the question “Which development environment(s) do you use regularly?” a participant would check all that apply. The way they represented this in the dataset was IDE1;IDE2. I separated each of these values out into their own column.

I ran the algorithm on the data

Graphs

Connections of the nodes

Conclusion

It appears that