Assigning programming languages to cities based on GitHub users activities

Gelera, Rodney
University of Victoria

McCulloch, Kaileen
University of Victoria

Warwick, Nick
University of Victoria

March 18, 2017

Contents

Abstract

Data has never before been generated at such high speeds. With the ever increasing volume of data it has become more and more challenging to sift through and extract useful information. Data mining and data visualization are two tools that can help solve this problem. In this paper we are going to use classification to simplify the information and a visualizer for quick and easy evaluation of information.

1 Introduction

2 Related Work

3 Data Collections

3.1 Scraping

3.2 Mapping

4 Data Mining

5 Data Visualization

6 Conclusion

7   Appendix A

7.1   Code