

**COLLEGE OF COMPUTING AND INFORMATION SCIENCES**

**DEPARTMENT OF NETWORKS**

**BACHELOR OF SCIENCE IN SOFTWARE ENGINEERING (YEAR 2) RECESS TERM 2 (BSE 2301)**

**CONCEPT PAPER FOR: STACK OVERFLOW DEVELOPER’S SURVEY RESULTS ANALYSIS PROJECT.**

**GROUP 7**

**PROJECT MEMBERS:**

|  |  |  |
| --- | --- | --- |
| **NAME** | **REG. NO** | **STD. NO** |
| NAMAGANDA MAGGIE | 16/U/9137/PS | 216014476 |
| KISEMBO ERINAH TUMUSIIME | 16/U/6168/PS | 216007296 |
| NALUMANSI FAITH | 16/U/861 | 216000271 |
| NYENDE MAHMOOD | 16/U/10636/PS | 216007311 |

**PROJECT LEADER**: KISEMBO ERINAH TUMUSIIME

**SUPERVISOR**: NOAH KANGE

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE SOFTWARE ENGINEERING RECESS PROJECT BSE 2301 20TH​JUNE, 2018.

Table of Contents

[1. INTRODUCTION 4](#_Toc517778263)

[1.1BACKGROUND 4](#_Toc517778264)

[2. PURPOSE 4](#_Toc517778265)

[3. PROJECT DESCRIPTION 4](#_Toc517778266)

[3.1GOAL 4](#_Toc517778267)

[3.2OBJECTIVES 4](#_Toc517778268)

[4. PROBLEM DESCRIPTION 5](#_Toc517778269)

[4.1.Techniques and tools: 5](#_Toc517778270)

[4.2.Benefits 5](#_Toc517778271)

[4.3.Project beneficiaries 5](#_Toc517778272)

[5. CONTACT INFORMATION 5](#_Toc517778273)

[6. References 6](#_Toc517778274)

# 1. INTRODUCTION

This document describes the Stack Overflow developers’ survey analysis project which is a software project for analysing through visualization of the 2018 stack overflow annual developers’ survey results.

## 1.1BACKGROUND

[1]Stack overflow is a popular question and answer site for programmers around the world. Every year a survey is carried out on the stack overflow developers and results are released. [1]In January 2018, a 30 minute survey was carried out and around 100,000 developers participated, they were asked [2]questions about everything from their favourite technologies to their job preferences.

## 2. PURPOSE

[1]The purpose of this project is to come up with various insights about the data release including the relationship between salaries, machine learning and diversity in technology using the survey results public and survey results schema tables.

[1]Over 100,000 developers took the 30 minute survey in January 2018, 98855 responses in this public data release considered qualified for analytical purposes based on completion and time spent on the survey and included at least one non PII question. Approximately 20,000 responses where started but not included in the data release because respondents did not answer enough questions or only answered questions with personally identifying information. Of the qualified responses 67441 completed the entire survey. The underrepresented groups in technology responded to the survey at even lower rates than expected from their participation the work force.

# 3. PROJECT DESCRIPTION

## 3.1GOAL

We purpose to critically analyse the survey results using various software tools to come up with insights on what kind of users are more or less likely to identify as part of the stack overflow community, participate or feel kinship with fellow developers.

3.2OBJECTIVES[3]

* To find out how long most developers have spent coding professionally
* To find out which programming language is most used by the developers
* To find out which operating system most developers work with
* To find out the education level of most developers
* To find out which age group of people participates most in the stack overflow activities

## 4. PROBLEM DESCRIPTION

The problem this project will address is that the relationship between the main survey result columns in the survey results public and survey results schema tables is not clearly laid out and thus insights about the data release cannot be obtained.

## 4.1. Techniques and tools:

We plan on using different data science techniques in R such as linear regression,Naïve Bayes classifier, KNN algorithm, decision trees, among others.

We also plan on using data tools in R such as data visualisation, random generation, advanced data visualisation.

## 4.2. Benefits

The following are some of the anticipated benefits at the end of this project;

* Obtaining insights about the relationship between salaries, machine learning and diversity in technology.
* Identifying the kind of users less likely to identify as part of the stack overflow community.
* Understanding the kind of users that are likely to participate in the stack overflow activities.
* Understanding those kinds of users that feel Kinship with fellow developers.

## 4.3. Project beneficiaries

The following are some off the beneficiaries of this project; Stack overflow developers, aspiring developers, applicants for jobs at stack overflow.

# 

# 5. CONTACT INFORMATION

For more information,please contact our project chief executive on:

Mobile +256 705 395 400

Mobile +256 791 504 377

Email: faithnalumansi@gmail.com.

# 6. References

|  |  |
| --- | --- |
| [1] | “ Developer Survey Results,” January 2018. [Online]. Available: https://stackoverflow.com/?ref=survey-2018&\_ga=2.133828686.1197689813.1529917966-1389724618.1521781189. [Accessed 19 June 2018]. |
| [2] | J. Silge, “Public Data Release of Stack Overflow’s 2018 Developer Survey,” Stack Overflow, 30 May 2018. [Online]. Available: https://stackoverflow.blog/2018/05/30/public-data-release-of-stack-overflows-2018-developer-survey/. [Accessed 26 June 2018]. |
| [3] | Q. Larson, Free Code Camp, 15 March 2018. [Online]. Available: file:///C:/Users/THE%20MAHMOOD/Downloads/Insights%20from%20Stack%20Overflow%E2%80%99s%202018%20survey%20of%20100,000%20developers.htm. [Accessed 19 June 2018]. |
| [4] | “Stack Overflow,” [Online]. Available: https://en.wikipedia.org/wiki/Stack\_Overflow. [Accessed 25 June 2018]. |