COMP90024 Team9 Project2

The liveability of Melbourne

Gengchang XU   1214774
Zhaoxiang NING   1261076
Xinyi ZHOU   1281911
Haochu WANG   1281962
Shiming ZHENG   1149897

Introduction

The purpose of this study is to investigate liveability of Melbourne via building a cloud based system to harvest and assess the sentiment class of the Twitter data about these three scenarios from Melbourne on the UniMelb Research Cloud (MRC), analyze the Twitter data with the data from the Australian Urban Research Infrastructure Network (AURIN) platform in CouchDB database, then, visualizing the results in a front-end web application.
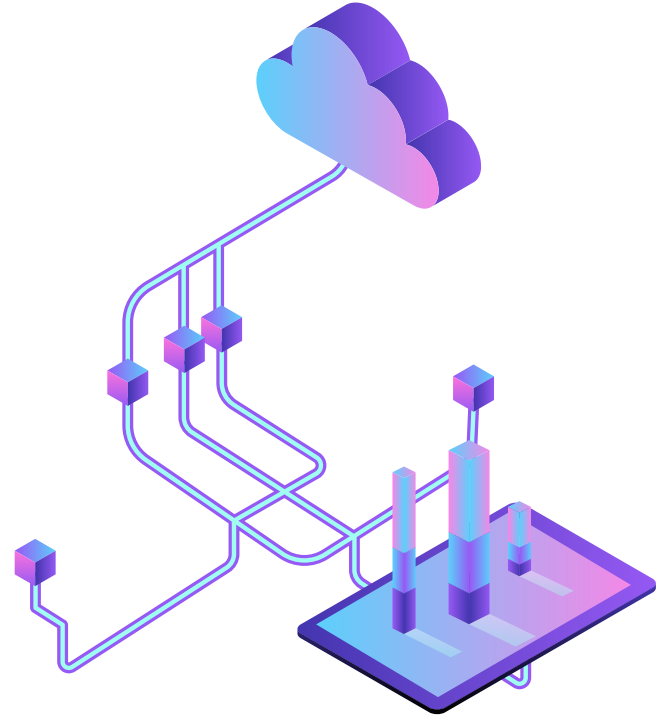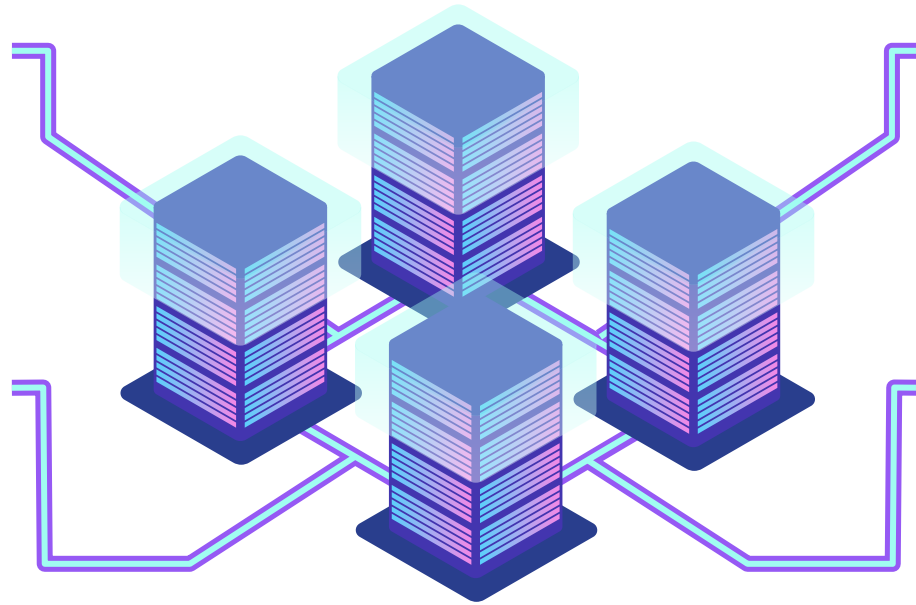
Content

1. System Architecture

2. System Design

3. Scenario Discussion

4. Data Visualization & Analysis
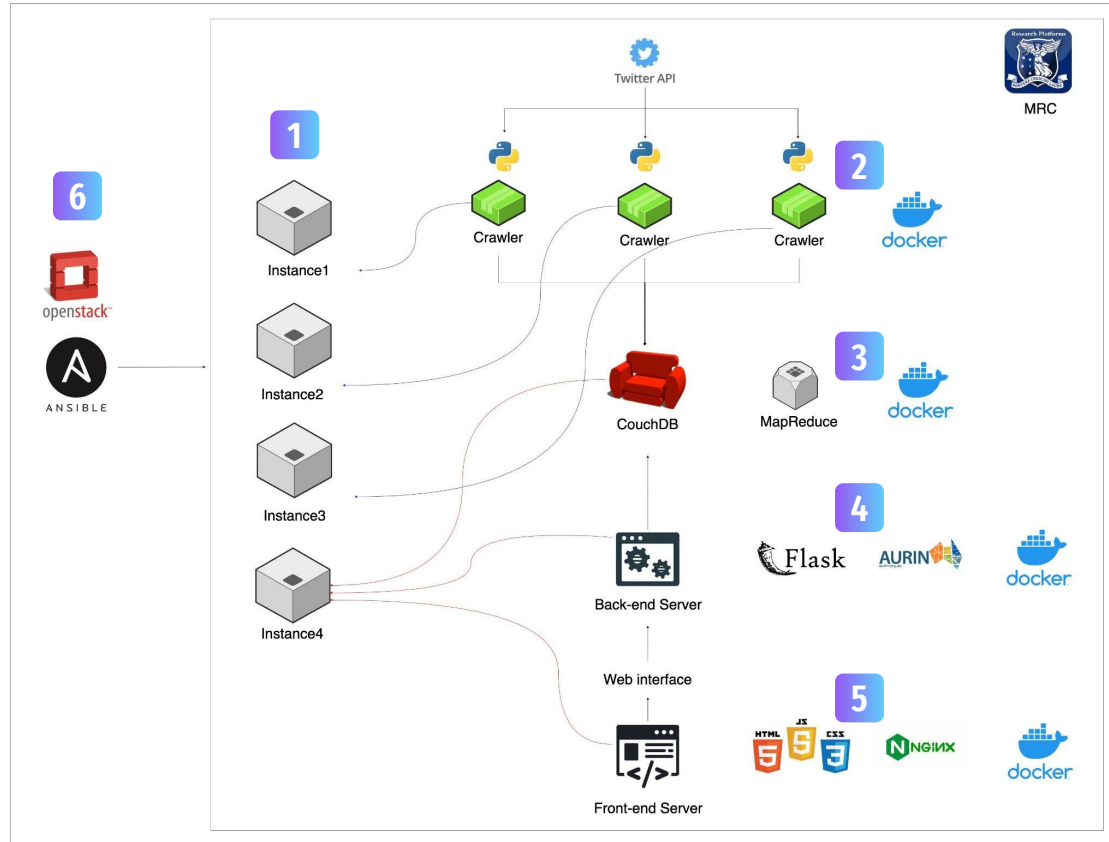
1.System Architecture

1

Resource Allocation

2

Crawlers

3

CouchDB

4

Back-end Server

5

Front-end Server

6

Ansible Automation

Twitter API

1

Instance1

Instance2

Instance3

Instance4

Crawler

Crawler

Crawler

2

docker

CouchDB

3

MapReduce

docker

MRC

4

Flask

AURIN

docker

Back-end Server

Web interface

5

HTML5

JS5

CSS3

NGINX

docker

Front-end Server

6

openstack

ANSIBLE

MRC Resource allocation

Instance	Flavor	Functionality	Volume
Instance 1	uom.mse.2c9g	Crawler(Data harvesting)	50 GB
Instance 2	uom.mse.2c9g	Crawler(Data harvesting)	50 GB
Instance 3	uom.mse.2c9g	Crawler(Data harvesting)	50 GB
Instance 4	uom.mse.2c9g	Docker and CouchDB deploying, Front-end and Back-end	200 GB

2. System Design Detail

2.How to retrieve data to the front-end?

1.How to remove duplicated tweets?
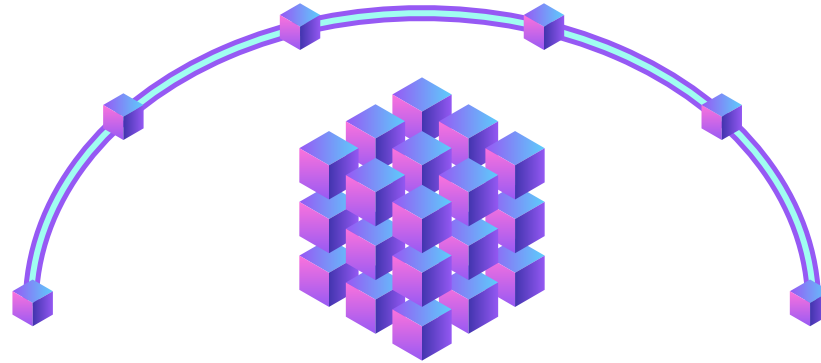
3.Which API we used in this system?

3. Scenario Discussion

Scenario 1

Whether the number of employed people in Melbourne on AURIN is positively correlated with the number of tweets that talk about recruitment towards a positive attitude.

Scenario 2

The correlation between the number of graduate students and the number of tweets with positive attitudes towards education.

Scenario 3

The relationship between the number of languages used in each suburb and the number of people whose salary reached median employment income.
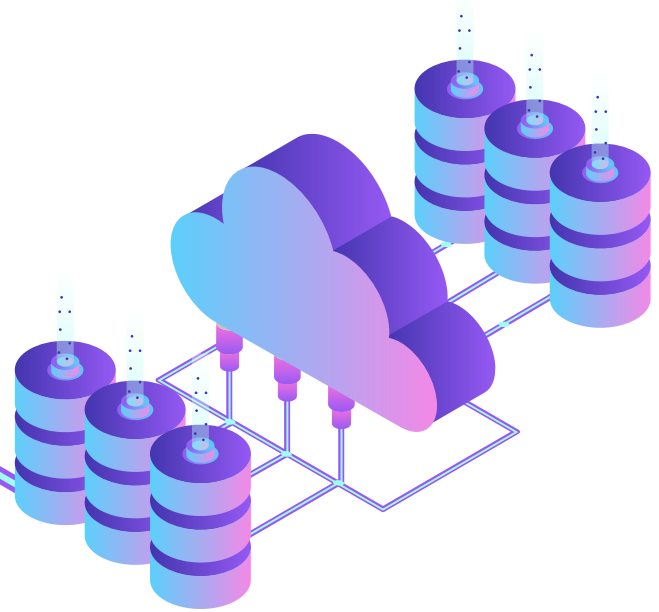
4. Data Visualization & Analysis

1 **Interactive maps**

To show the population distribution of each suburb.

2 **Histograms**

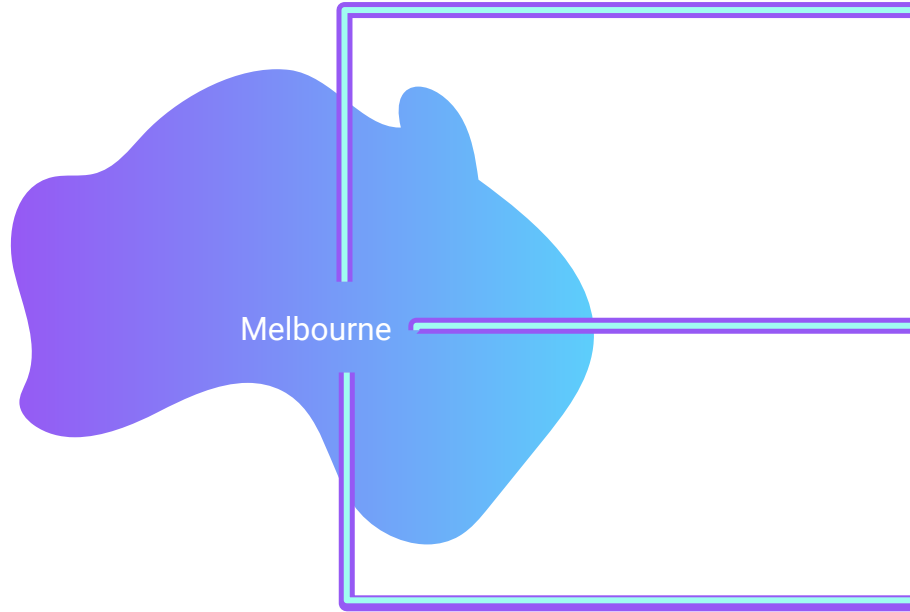To reveal the data from tweets analysis.

3 **Pie charts**

To display the data from AURIN we selected.

4 **Linear Regression**

To verify the hypothetical correlation for each suburb.

Web App Link: http://172.26.132.76:3000/

Conclusion

For scenario 1

The number of recruitment tweets towards a positive attitude is positively correlated with the number of employed people from AURIN.

For scenario 2

The correlation between the number of tweets with positive attitudes towards education is positively correlated with the number of graduate students from AURIN.

For scenario 3

The correlation between the number of languages used is negatively correlated with the number of people whose salary reached median employment income from AURIN for the given suburbs in Melbourne.

Melbourne

Thank you for watching!

Team 9

Gengchang XU	1214774
Zhaoxiang NING	1261076
Xinyi ZHOU	1281911
Haochu WANG	1281962
Shiming ZHENG	1149897