Internship Report on Mayalogy Limited

by

Khondokar Sakib Hasan
16101183

An internship report submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
January 2022

Declaration

It is hereby declared that

1. The report submitted is my own original work while completing degree at BRAC University.
2. The report does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The report does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

Student's Full Name & Signature:

Khondokar Sakib Hasan
16101183

Approval

The Internship Report titled "Internship as a Data Engineer on Mayalogy Limited" submitted by

1. Khondokar Sakib Hasan (16101183)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science.

Examining Committee:

Supervisor:
(Member)

Dr. Muhammad Iqbal Hossain
Assistant Professor
Department of Computer Science and Engineering
Brac University

Co- Supervisor:
(Member)

Razib Hasan Shuvo
Senior Software Engineer
Mayalogy Limited

Head of Department:
(Chair)

Sadia Hamid Kazi
Chairperson, School of Data and Sciences
Department of Computer Science and Engineering
Brac University

Letter of Transmittal

Dr. Muhammad Iqbal Hossain
Assistant Professor
Department of Computer Science and Engineering
BRAC University

Subject: Submission of Internship Report on Mayalogy Limited

Dear Sir,
This is to inform you that I have successfully completed my four-month internship at Mayalogy Limited, which lasted from 14 March to 14 July 2021. This is my internship report for the work I completed at that company. It is an honor and a privilege for me to be able to submit my internship report at the conclusion of my internship. This report is a reflection of my internship experience, daily activities, and educational outcomes. The entire report is devoted to the experience I gained while working for the company for the last seven months, three of which were spent as a full-time employee. I've attempted to incorporate what I learned during my internship. I hope this internship report meets your expectations and satisfies you. Additionally, you'll be able to evaluate the caliber of my work and institutional education. I would be delighted to provide you with adequate rationale for any clarification or explanation regarding this report.

Thank you.

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis have been completed without any major interruption. Secondly, to my advisor Dr. Muhammad Iqbal Hossain sir for his kind support and advice in my work. He helped us whenever I needed help. And finally to my parents without their throughout sup-port it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

List of Figures

Chapter 1

Company Background

1.1 Background Information

Maya is an anonymous digital healthcare delivery service that links anyone seeking assistance with specialists who work on-demand. Maya began its journey in 2015. Maya Bangladesh, where there are numerous hurdles to obtaining credible advice (especially for women), Maya provides an easy-to-use service (dubbed "Maya") via which users can submit anonymous questions. Queries are then sent through a real-time, uber-like backend to a vetted network of professionals (doctors, therapists, etc.). Maya's long-term goal is to be a self-sustaining business, but they are extremely motivated by their purpose to help people. They are lucky to have a fantastic partner in BRAC, which provides financial assistance and invaluable counsel. Additionally, they feel that establishing a self-sustaining business will ultimately result in Maya having the greatest impact on society. While Maya's current focus is on scaling the service, a premium service was introduced to provide expert advise to organizations (e.g. textile manufacturers). Packages for business-to-consumer transactions were introduced in late 2018. BRAC, Multi Source and Multi-Mode, a private individual, ACI, and Osiris are among their investors. In underdeveloped nations, where many subjects are stigmatized or where infrastructural impediments to expert assistance exist, an anonymous service connecting customers to specialists is critical. For example, domestic violence has increased by 13% in Bangladesh, with 87 percent of women reporting some sort of abuse. Regrettably, only a third take action (source: UNDP, 2014).

1.2 Current Situation

A mother looking for guidance on a topic like this can go to Google, social media, family members, or an expert for help. Information found on search engines, websites, and blogs is often useful, but it's not always tailored to the individual. This is a widespread problem in Bangladesh, where people are afraid to open up to their mothers, sisters, and other family members about their personal problems because they don't feel comfortable doing so.

1.3 Value Proposition

Access to the expert, one-on-one, will be more easier and free (unless the user is on a subscription service) by Maya. A lack of information, distance, costs, and socio-cultural barriers can make a doctor or lawyer difficult to get. They're trying to offer value by giving their customers a simple, safe way to communicate with one another. Maya then frequently directs the user to a local doctor, lawyer, support group, etc. Users will never have to worry about being judged because Maya will always keep their identities private. They're also placing a lot of money on mobile devices. Despite their plans to create an SMS service, their research shows that the number of people who have access to the internet is increasing rapidly.

1.4 Key Principals

In some nations, it may be difficult to ask for help or instruction, especially in the developing world. Because of this, even if a person is willing to seek treatment, there are several barriers. Messaging apps can have a significant impact on businesses that overlook their potential. It must be a two-way conversation in order to succeed. Despite its tremendous scalability, sending SMS texts is not the answer. Although Google is an excellent source of information, it's not always the greatest option because individual data isn't validated. For Maya users, it's critical to know that someone is responding to their specific issue. In order to meet demand and maintain efficiencies, traditional techniques of hiring and managing specialists (e.g. in clinics and call centers) will not work. A more real-time demand-supply method, like Uber, is being used by Maya to grow more rapidly and efficiently.

1.5 Name background

When Ivy (CEO of Maya) was pregnant with Maya, she was inspired to create the company by her mother's battle with breast cancer and her own personal experience.

1.6 Who is their target audience?

1. Maya is designed to be a general-audience service, open to anyone, regardless of age, with a focus on urban youth (18-24 years old) and smartphone users.
2. Currently, they are concentrating their efforts in Bangladesh, with the goal of spreading Maya throughout South Asia.
3. Given their background and the fact that women are typically most in need, they conduct a lot of outreach to women, especially young women, but they are increasingly seeing males asking Maya questions about the women in their lives and about personal issues.

4. New mothers are a big target for Maya, looking to support women from pregnancy, delivery to child care.

5. People in Bangladesh are afraid to discuss their sexual health because of the stigma attached to it. As a result, the Maya app is an excellent resource for those seeking information about sexual health while maintaining their anonymity.

6. .Maya is also helping people who are suffering from mental health problems.

7. Maya App gives the following bundles to a patient: The Maya Plus 7-day free trial is available for all of the Maya Plus Silver, Gold, and Platinum editions of the software. One-time, Weekly, Monthly, Pregnancy, and Mini Packs are all available. Maya Counseling offers a FREE mental fitness assessment as part of her commitment to holistic health.

1.7 Vision of the Company

Maya Apa started in Dhaka a few years ago and day by day it expanded in whole Bangladesh. Every day we get thousands of paid questions from every point of Bangladesh. In 2018 Maya got some new investors such as ACI and Anchorless. Maya is already expanding in Pakistan to serve the people from Pakistan. We have also recruited some employees from Pakistan and currently working with them. By 2022 we are hoping to reach our break even where company will be able to run by its own revenue.

1.8 Organizational Chart and Employee details

Here in Maya there are 200 employees including the Doctors. We have separate leads for each and every team. Those leads report to their managers and Managers reports to the General Manager and Country head. After their assessment they send all the details to the CEO or to the investors.

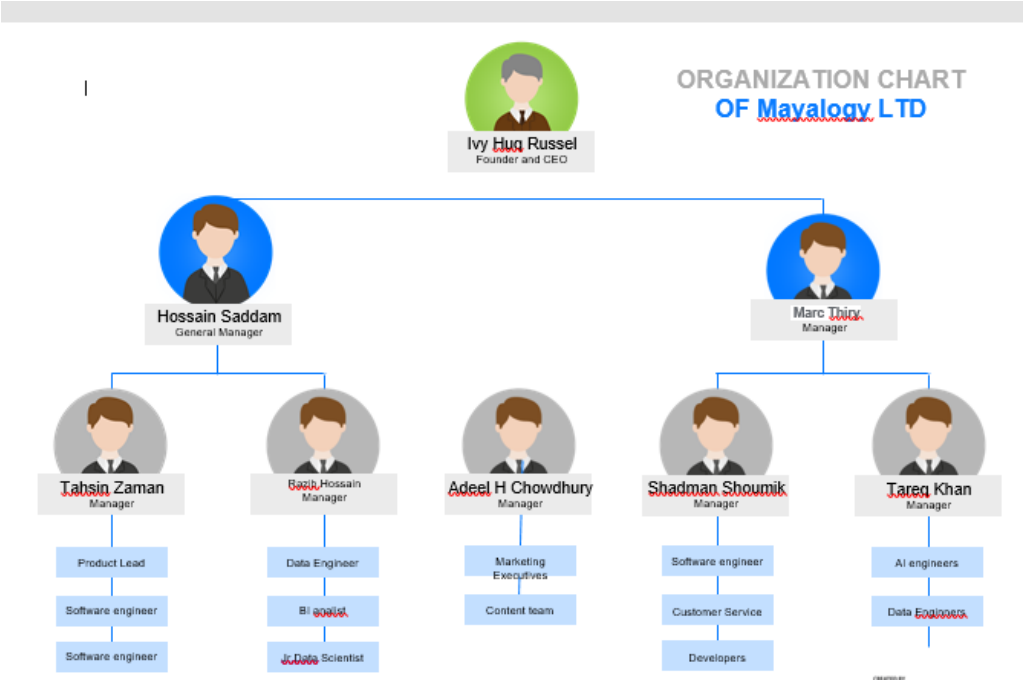

Figure 1.1: Organization Chart of Mayalogy

1.9 Departments and their work details

Just like other startups Maya has Finance, Tech, AI, Data, HR, Marketing, Sales, Business Development, customer experience and Dev ops. Other than that we have full time and part time doctors. Among them we have leads as well. Just like every startup we have Financial analysts to analyze and make financial models and also they need to research for the details that helps us to raise funding.Tech Team plays a major role whenever we are facing any difficulties or we need to design or implement any technological decisions. There are software engineers, UI/UX designers and developers in the tech team.AI team has AI engineers to support our system and make our AI more responsible. Data team has the most important role in the day to day task. Data team needs to response to the other team by providing necessary data's also people from data team needs to analyze every possible behavior of the users and build business models for the upcoming days and plan the OKR for coming days. There are business analyst, data scientists and data engineers in the data team. HR manages every organizational behavior as well as the look in to the appraisals and hiring part. Marketing and content creators work together for generating demands as well as making Maya a better place in the social medias. Business development and sales work together they plan how to expand and generate assets from the market.

Chapter 2

Internship activities and responsibilities

2.1 Job responsibilities

1. Design, implement, and automate deployment of our distributed system for collecting and processing log events from multiple sources.
2. Design data schema and operate internal data warehouses and SQL/NoSQL database systems.
3. Own the design, development, and maintenance of ongoing metrics, reports, analyses, and dashboards to drive key business decisions.
4. Monitor and troubleshoot operational or data issues in the data pipelines.
5. Work collaboratively with Business Analysts, Data Scientists, and other internal partners to identify opportunities/problems.
6. Provide assistance to the team with troubleshooting, researching the root cause, and thoroughly resolving defects in the event of a problem.

2.2 Till now my contribution for Maya Data team

- Database Documentation: The data team needs to work with 20-30 table each and every day. So whenever someone needs to write a query he needs to go through every table. So my first work was to understand the database and document the most frequently used tables.
- Query Documentation: There are some queries that needs to run every day or every week. So I write some of them and I collected some of them and documented in a google sheet.
- Finding anomaly values from database: There were many missing values and some illogical values in the database. I had to find them and report to my supervisor.
- Responding to data requests: Every day I had data request from other teams. For that I had to write queries in MySQL and provide them necessary data's.

- Extracting Data's for Churn Analysis: Churn analysis is very important to understand the user base. For that project we had to extract different data matrices.
- Making Effecting Dashboards
- User longitude and Latitude plotting on map: To find out in which area we have a large number of users I plotted longitude and latitude in the map by using folium library in python.
- Writing automated queries inside python: For avoiding same data requests everyday I wrote code in python so that the time range can be selected automatically and run queries inside python and convert the result in csv. After that it will email the CSV file from python code. I used MIME library for that.
- Used Python for ETL process: This helped to to insert query data's in a google sheet. I wrote code in python. Used Google sheet api and gspread and pandas.
- Used Mixpanel for Finding out the retention and some confidential user behavior data.
- Making funnels for business matrices: For Mayas business I have created some funnels such as install to registration funnel, Registration to ask question or registration to pay funnel.
- Retention analysis based on user behavior such as purchasing, app activity.
- Working on company due diligence questions.
- Making net promoter dashboard and implementing effectively.

2.3 Work Samples

2.3.1 Query Documentation Example

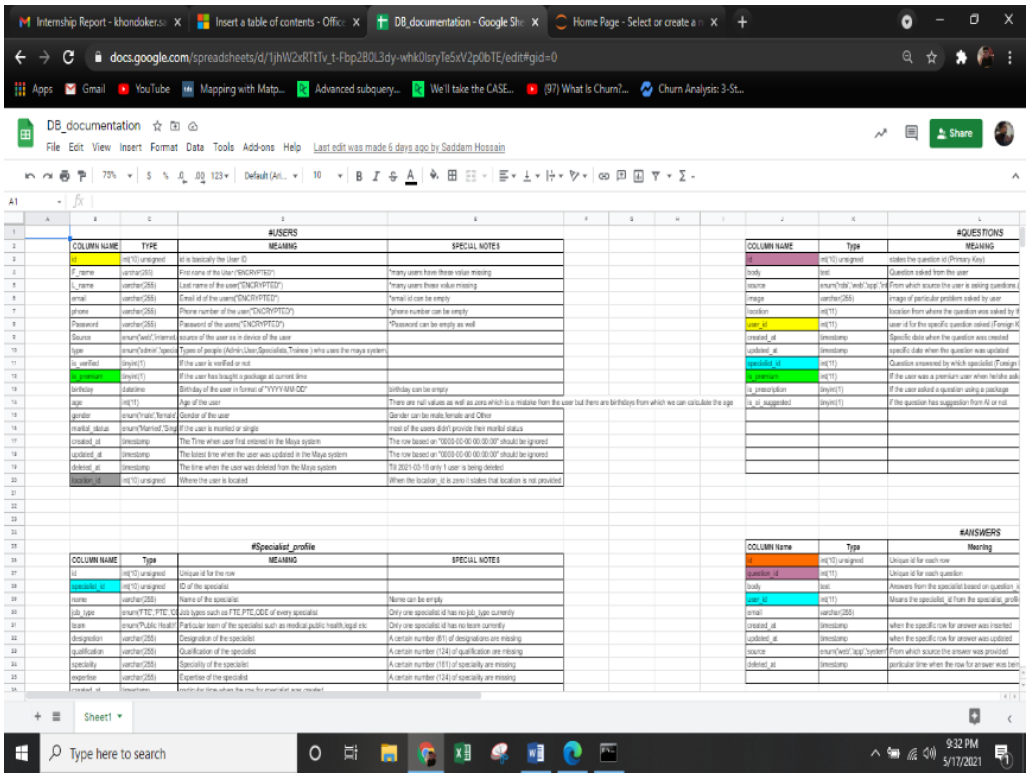

Figure 2.1: Database Documentation

This documentation shows the inner meaning of each and every individual column of every most frequent used tables for writing queries.

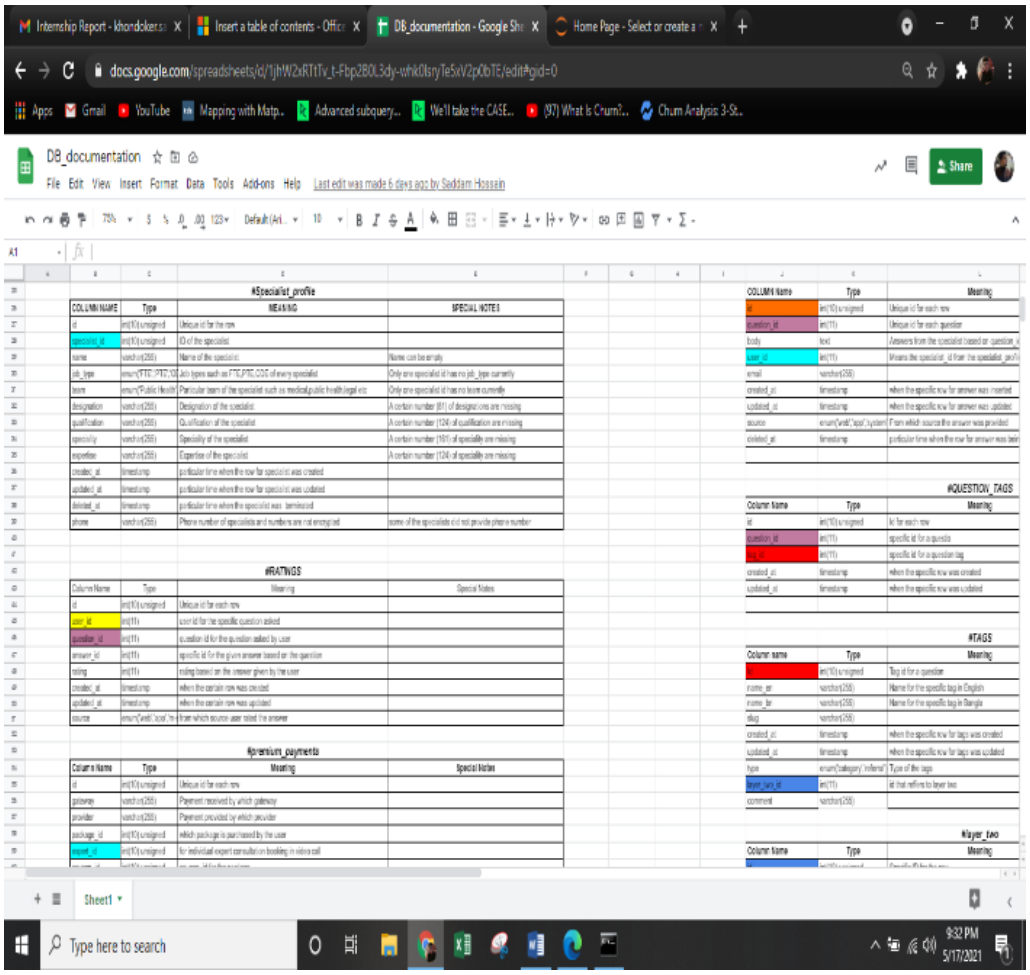

Figure 2.2: Database Documentation

2.4 Sample Queries

Link of Sample Queries

QUERY NAME	QUERY	WHAT IT DOES
Question count of all tags for any layer_one	select name, count(*) as Qs from (select q.* , t.name_en as name, l1.id as l12 from questions q, questions_tags qt, tags t, layer_two l2 , layer_on where q.id = qt.question_id and qt.tag_id = t.id and t.layer_two_id = l2.id and l2.layer_one_id = l1.id and l1.id = 2 and date(q.created_at) between '2020-07-01' and '2020-12-31') C inner join users ON C.user_id=users.id inner join locations ON C.user_id=locations.user_id group by 1 order by 2 desc	This query is giving the count of
How many questions were directly answered by AI and how	select DATE1 as month_name, APPLIED, DIRECT, (APPLIED/QU (select DATE_FORMAT(created_at, "%Y-%m") AS DATE ,count(*) date(created_at) >= '2019-01-01' group by DATE_FORMAT(created_at, "%Y-%m")) AS B, (select q.created_at, DATE_FORMAT(q.created_at, "%Y-%m") AS DATE1 from questions q, `ai_response_logs` aal where q.id = aal.`question_id` and q.`is_ai_suggested` = 1 and aal.status = "relevant" and q.status = 'answered' and date(q.created_at) >= '2019-01-01' group by DATE_FORMAT(q.created_at, "%Y-%m")) as A, (select DATE_FORMAT(created_at, "%Y-%m") AS DATE3, count(d date(created_at) >= '2019-01-01' group by DATE_FORMAT(created_at, "%Y-%m")) as C where A.DATE1 = C.DATE3 and A.DATE1 = B.DATE order by created_at	This query is giving the count of

Figure 2.3: Sample Queries

There are times when finding an answer through data is extremely important. For the time optimization I have made this query documentation so that anyone can optimally use this documentation.

2.5 Sample Work in Python

In [28]:
```
#For EMAIL
import smtplib
from email.mime.text import MIMEText
from email.mime.multipart import MIMEMultipart
from email.mime.base import MIMEBase
from email import encoders

email_user = 'sakib@maya.com.bd'
email_password = '[redacted]'
email_send = 'khondoker.sakib.hasan@gmail.com'

subject = 'DF in csv'

msg = MIMEMultipart()
msg['From'] = email_user
msg['To'] = email_send
msg['Subject'] = subject

body = 'Hi there, sending this email from Python!'
msg.attach(MIMEText(body,'plain'))

filename='practice.csv'
attachment = open(filename,'rb')

part = MIMEBase('application','octet-stream')
part.set_payload((attachment).read())
encoders.encode_base64(part)
part.add_header('Content-Disposition',"attachment; filename= "+filename)

msg.attach(part)
text = msg.as_string()
server = smtplib.SMTP('smtp.gmail.com',587)
server.starttls()
server.login(email_user,email_password)


server.sendmail(email_user,email_send,text)
server.quit()
```

Out[28]:
```
(221, b'2.0.0 closing connection k69sm3952521pgc.45 - gsmtp')
```

Figure 2.4: Automated Email Sending in Python

In [35]:
```
import pandas as pd
import requests
from xml.etree import ElementTree
import numpy as np
import folium
import mysql.connector
%matplotlib inline
```

In [36]:
```
maya_db = mysql.connector.connect(
  host="[redacted]",
  user="sakib",
  password="[redacted]",
    database = 'maya_db'
)
```

In [37]:
```
with maya_db.cursor() as cursor:
        sql = ("""SELECT DISTINCT lat,`long`,area
FROM locations
where (date(created_at) BETWEEN '2021-03-01' and '2021-03-10') and (lat!=0 and `long`!=0) and area is not null
limit 300""")
        cursor.execute(sql)
        cols = cursor.column_names
        result = cursor.fetchall()
```

In [38]:
```
df=pd.DataFrame(result)
df.columns = cols
#df.head()
df
```

Out[38]:

	lat	long	area
0	23.757193	90.3667057	Mohammadpur
1	23.8035401	90.3729659	Mirpur
2	23.4634017	91.1842282	Comilla Adarsha Sad
3	23.51449	90.2084941	Sreenagar
4	23.7277831	90.3811415	Azimpur
...	...	...	...
295	22.584258	91.6671113	Sitakunda
296	23.6068255	89.4487217	Sreepur
297	22.5018866	92.2110633	Kaptai

Figure 2.5: Loading Latitude and Longitude of a user by writing query inside Python

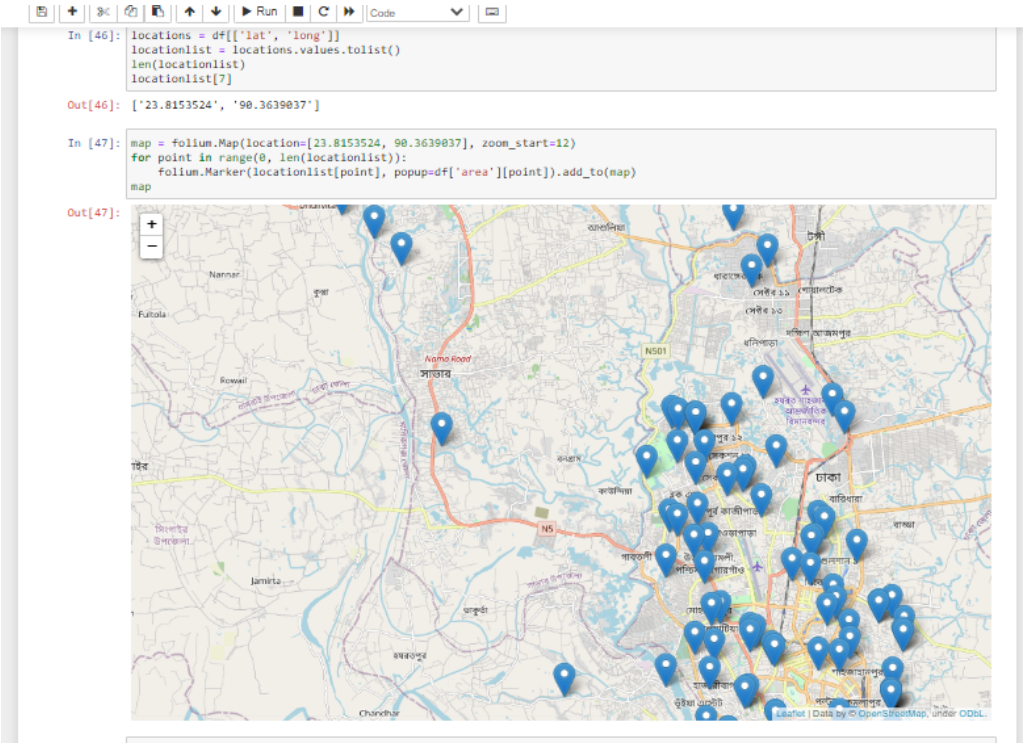

Figure 2.6: Latitude and Longitude of a user in Python

Link of Sample Python Code and Results

Figure 2.7: Net Promoter Score Calculation

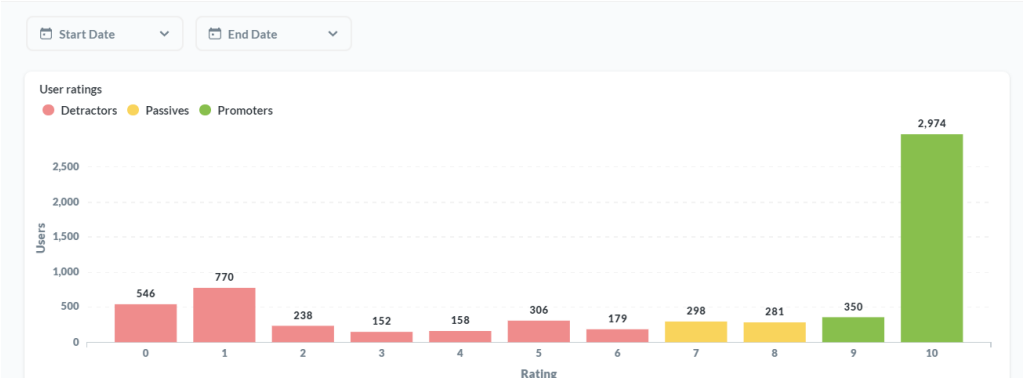

Figure 2.8: Net Promoter Score Visuals

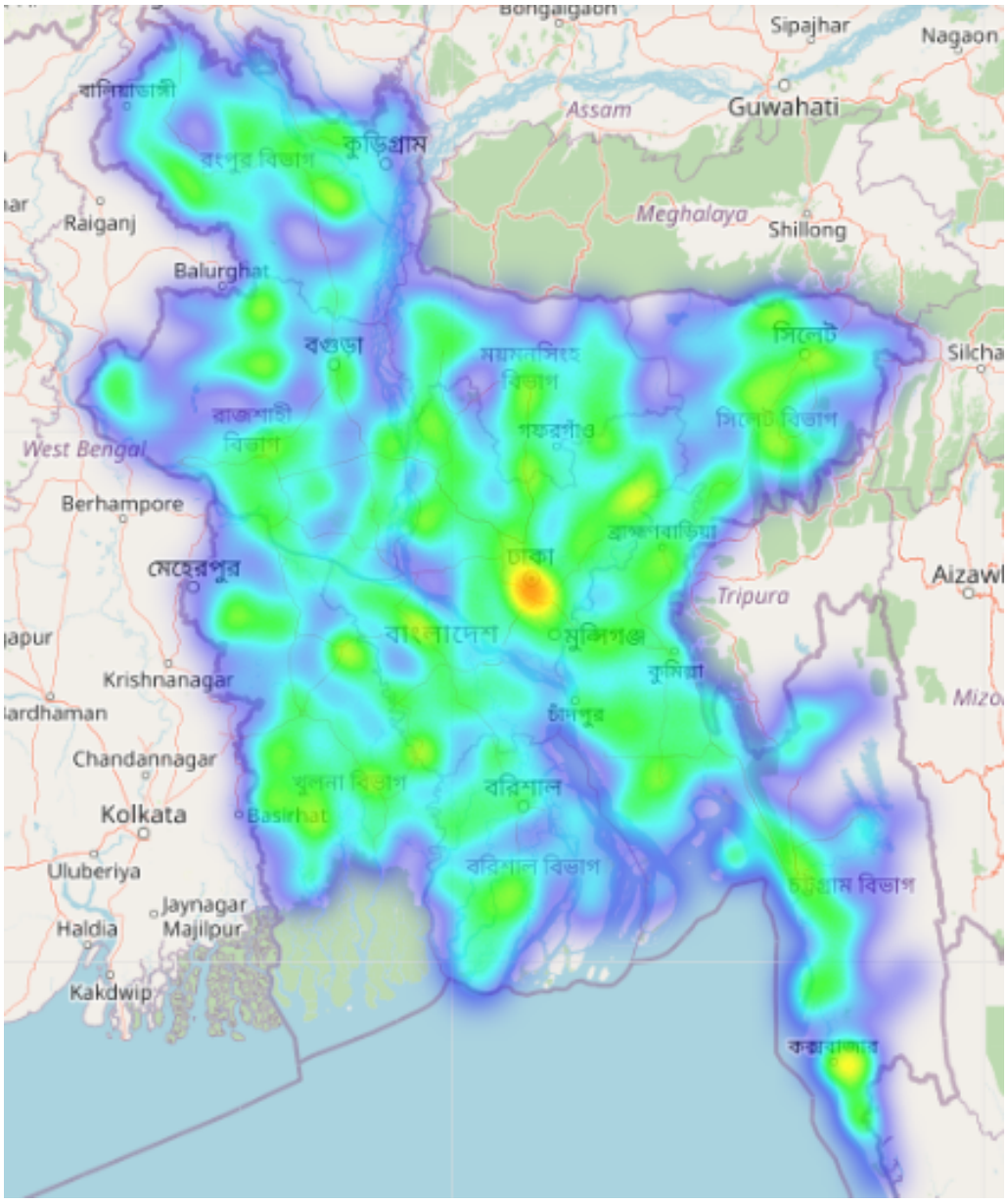

Figure 2.9: X users heatmap using folium library in Python

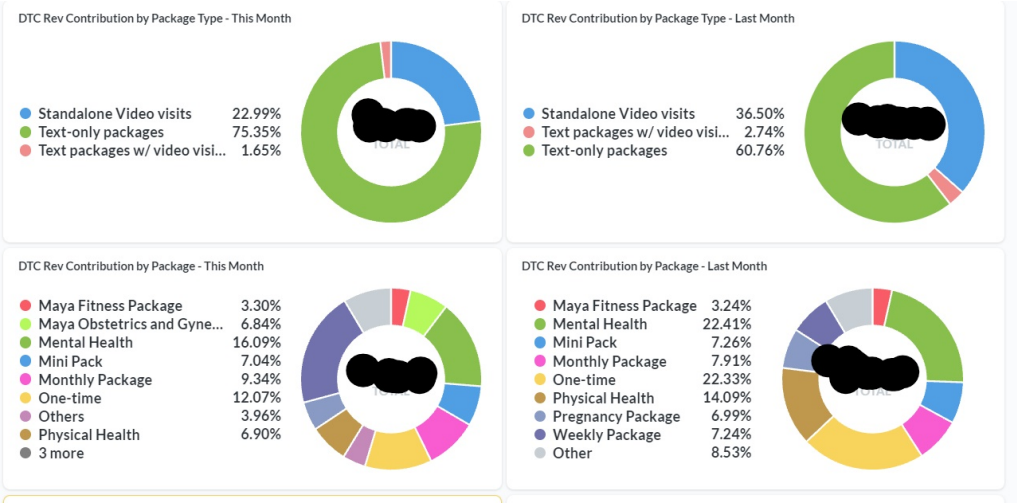

Figure 2.10: Sample board meeting dashboard

Chapter 3

Conclusion

3.1 Conclusion

In classes we learn theory, without knowing the actual implementations, learning theory makes us root strong but only that is not enough. In the last 7 months the as an intern and as a full-time employee it helped me to apply the knowledge which I learned till now, which was a really exciting opportunity for me. Learning while building something interesting is really fun. This internship opportunity really helped me to learn new things every day. While working, I had new experiences everyday which forced me to think critically and do things in deferent ways. This internship really helped me to increase my experience as a data guy. I gained real data analysis experience here. I was encouraged to be responsible and confident by the tasks on which I worked. The internship program assisted me in the sense that it enhanced my ability to function in a group setting. It is critical for a corporation to keep its functions as a unit and to value each teammate's creativity and contributions. I really want to conclude by thanking my supervisor Dr. Muhammad Iqbal Hossain sir as he guided me in the project in a great way.