SIES (NERUL) COLLEGE OF ARTS, SCIENCE AND COMMERCE

NAAC ACCREDITED 'A' GRADE COLLEGE
(ISO 9001:2008 CERTIFIED INSTITUTION)
NERUL, NAVI MUMBAI – 400706

INTERNSHIP REPORT ON

Data Analyst Intern

SUBMITTED BY

Bosco Trevor Dsouza

SUBMITTED IN THE PARTIAL FULLFILMENT FOR THE DEGREE OF

MSc. COMPUTER SCIENCE
SEMESTER – IV, 2021 - 2022

SIES (NERUL) COLLEGE OF ARTS SCIENCE AND COMMERCE

NAAC ACCREDITED 'A' GRADE COLLEGE

(ISO 9001:2015 CERTIFIED INSTITUTION)

Certificate

NERUL, NAVI MUMBAI - 400706

THIS IS TO CERTIFY THAT

Bosco Trevor Dsouza

HAS UNDERTAKEN AN INTERNSHIP IN

EdIndia Foundation

Seat No: 04

In partial fulfilment of MSc - CS Degree Semester ___IV___ Examination in the academic year __2021-22__ . It is further certified that he/she has completed 600 hours of the Internship.

External Examiner

Head of Department   Principal

__Acknowledgement__

I am grateful to EdIndia Foundation for recruiting me as an intern. I express my professional gratitude and deep regards to **Mr. Varun Bhamre**, project manager of team Nirnay for his co-ordinal support, valuable information, inspiration which strengthened my confidence during my internship program and his precious guidance which were extremely valuable for my professional front as well as formy study.

I am quite grateful for this opportunity, throughout the internship, I learned about a lot of new data analysis tools and a whole new way of looking and analyzing data. I have learned how to effectively use data that data to generate a story out of it. I consider myself as a very lucky individual as I was provided with an opportunity to be a part of at EdIndia Foundation.

Subject: Regarding performance of Mr. Bosco during winter internship

Dear Sir/Madam,

This is to certify that Mr. **Bosco Trevor Dsouza** of <u>SIES College of Arts, Science and Commerce, Nerul</u> College/Institution worked as an intern as part of his MSc course in Computer Science of University of Mumbai. The particulars of internship are given below:

Internship starting date: 05/01/2022

Internship ending date: 19/05/2022

Actual number of days worked: 100 days

Tentative number of hours worked: 850 Hours

Broad area of work: End-to-end data analysis and visualization of educational, district administration specific as well as organization's internal data

A small description of work done by the intern during the period:
Bosco started his work as a support around data analysis and visualization to ongoing projects with government stakeholders. Eventually he worked on end-to-end data analysis of organization's internal program specific datasets, which include operational efficiency specific data, impact assessment data and digital media outreach analysis data. He has helped EdIndia in improving modes of visualization significantly in the process. He has also worked on creating wireframe as well as one of the modules for a WebApp of our data analytics project, Project Nirnay. Apart from that, he collaborated with the team in development of in-house version of nowadays trending game, 'Wordle'.

Thanking you,

Name: Varun Bhamre
Designation: Lead – Project Nirnay
Contact number: 8087044415
Email: varun.bhamre@edindia.org

Authorized Signatory

+91 7710903515
info@edindia.net

www.edindia.org

Maker Maxity, 5 North Avenue, Level 5Th,
Bandra Kurla Complex, Bandra East,
Mumbai, MH, (400051)

Name of intern: Bosco Trevor Dsouza

College/institution: SIES College of Arts, Science and Commerce, Nerul

[Note: Give a score in the 1-5 scale by putting √ in the respective cells]

Sr No	Particulars	Excellent	Very Good	Good	Moderate	Satisfactory
1	Attendance	√				
2	Punctuality	√				
3	Adaptability	√				
4	Ability to shoulder responsibility		√			
5	Ability to work in a team	√				
6	Written and oral communication skills		√			
7	Problem solving skills		√			
8	Ability to grasp new concepts	√				
9	Ability to complete task		√			
10	Quality of work done	√				

EdIndia™

Comments:

Bosco has been brilliant during his internship stint with EdIndia Foundation and has supported all our projects in every possible way. His attitude towards work is encouraging to see. He has always been humble and ready to shoulder any responsibility. He is a very good team player and always willing to help others in his spare time.

Signature:

Name: Varun Bhamre

Designation: Lead – Project Nirnay

Contact number: 8087044415

Email: varun.bhamre@edindia.org

📞 +91 7710903515
✉ info@edindia.net

🌐 **w w w . e d i n d i a . o r g**

Maker Maxity, 5 North Avenue, Level 5Th,
Bandra Kurla Complex, Bandra East,
Mumbai, MH, (400051)

Table Of Content

Title: *Data Analyst Intern*.
Description of the organization:

EdIndia is a section 8 Not for Profit Company that leverages technology and data analytics to create innovative solutions to impact the quality of education at scale. It has been provided incubation support by Sterlite Power Transmission Ltd. (SPTL) and its group companies. EdIndia is working with 6 states and providing analytical solutions to almost 53,000 schools and have trained almost 2,50,000 teachers.

There are 4 teams that function each having its own importance. Pragyan, Teachable, Digital Classroom and Nirnay.

Pragyan is a Multi-State In-Service teacher capacity building program that aims to improve the Teaching-Learning environment in the classrooms to capacitate the teachers of public education systems of the country, Pragyan focuses on Large-Scale interventions leveraging technology. It mainly deals with E-content creation and Teacher support.

Teachable is basically to support and provide webinars for the students who are pursuing or want to pursue their career in teaching.

Digital Classroom looks after the student's assessment, teacher training and conducting live sessions and webinars.

Nirnay All the analytical part is looked after by team Nirnay. All the data is sent to Nirnay team so that they can generate meaningful insights and present it in a dashboard format. Also, with dashboard we provide some statistical insights. Nirnay also works with administrative level of data that is government related data with respect to electricity connection, drinking water connection, type of roads, etc.…

So, I was a part of team Nirnay and I worked as a data analyst.

Recent project that we have acquired are:

1. MTGSY (Mukhyamantri Tripura Gram Sammridhi Yojana) project of South Tripura.

2. Vidya Jyoti 100 School Yojana

3. Hamara Vidhyalaya school level data with students' assessment (Namsai)

4. Sephaijala district Basic Survey project administrative level of data

Description about the activities done by the section where the intern has worked:

Google Data Studio:

A Business Intelligence tool or an analytical tool used to represent visualization on a dashboard and create an interactive dashboard. We can create are own dashboards and embed those dashboards in are website. We can also perform calculations and other operations. We can also implement map graphs to give a better location specific distribution or result to the user.

Advantage:
1. It is easy to use.
2. It has many connectors to connect directly with data source.
3. Consist of a tons of community visualization.
4. It is a 100% free source.

Google Spreadsheets:

The data from the stakeholders come in a spreadsheet format. All the operations and all the file cleaning are done in the spreadsheets.

Advantage:
1. It is easy to use.
2. It comes with a huge list of formulas that can be very helpful to perform operations and in manipulating the data as per the requirements.
3. It is a 100% free source.

Google BigQuery:

The data is also collected in real time and in such case spreadsheet will not be able to hold that amount of data in such a case we need a warehouse to store the data. So, Bigquery is one of the best options.

Advantage:
- It is easy to use similar to SQL.
- It allows you to use 10GB of free data.

Python:

I worked with python to clean data and also perform some statistical insights with the help of SciPy library and plotting those insights with the help of matplotlib and seaborn

Description of work allotted and actually done by the intern:

       I started working with support around data analysis projects and tried to involve myself in maximum number of projects to get maximum exposure. So, I have developed dashboards with respect to administrative data, student assessment data, and Internal purpose data which consist of data from various digital platforms.
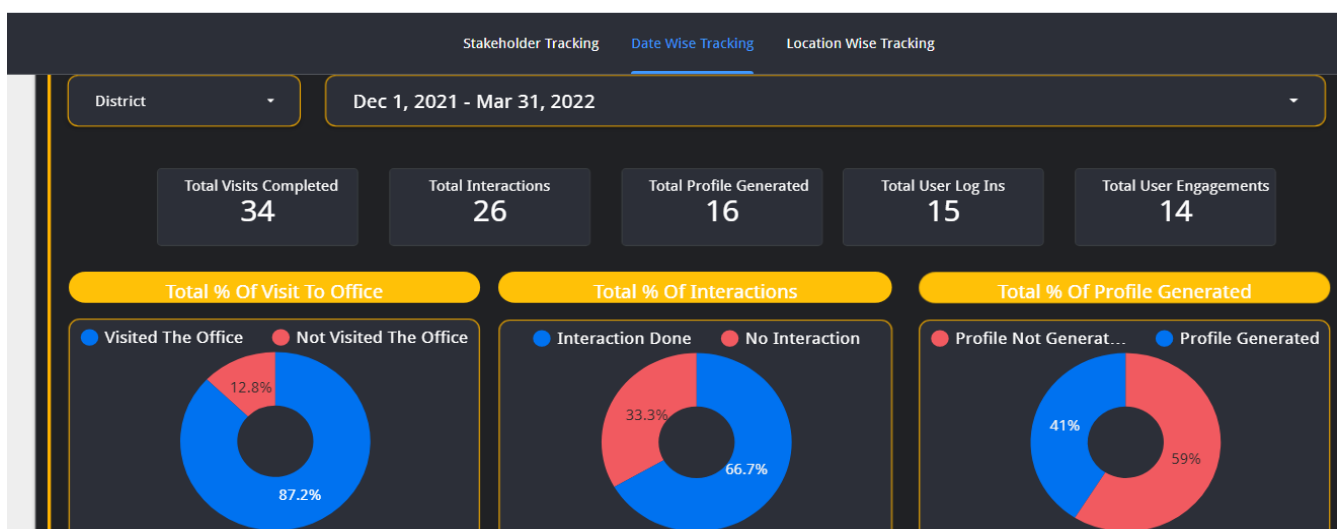
Work description:

❖ **Week 1**
- ✓ First week the first two days was like training they gave me some dashboards and ask me my views like third person review.
- ✓ After that my first task was to modify the Sephaijala district basic survey dashboard.
- ✓ After which my next task was to find out new visualizations to implement in the dashboards.
- ✓ The next task was to develop a automated data cleaning model which can highlight the errors and null values in an excel file.
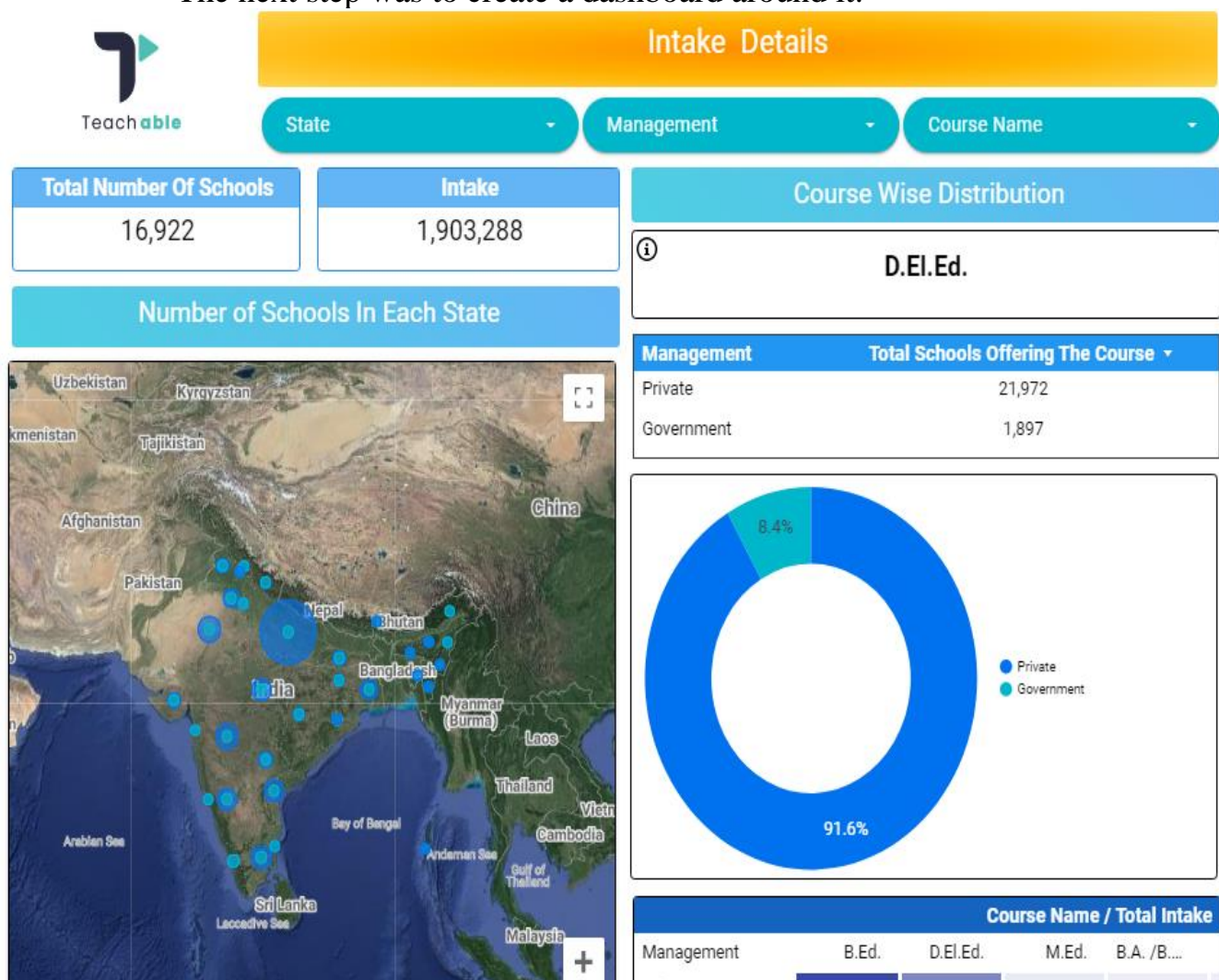
❖ **Week 2**
- ✓ In Second week of Internship, I was asked to do my very first dashboard that is to make a dashboard to track the progress of the internal field visit team of Nirnay.
- ✓ It took me around three days to create a dashboard and clean the data.
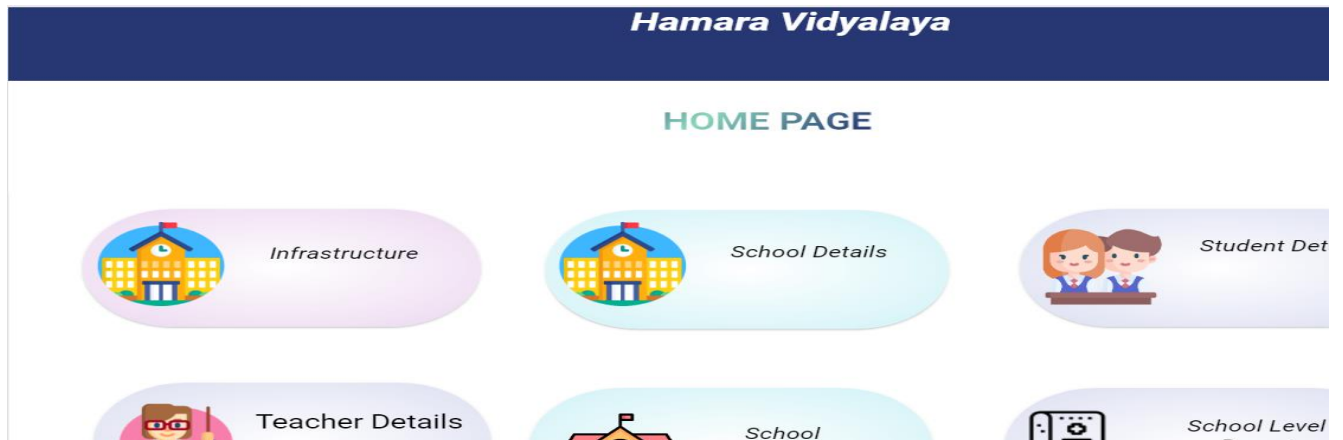- ✓ After that there were some issues in the date format of the dashboard which I modified.

❖ **Week 3-4**
- ✓ In This week I created a dashboard around the Teachable Students intake data.
- ✓ This involved collecting data from pdf format and then converting it into excel format.
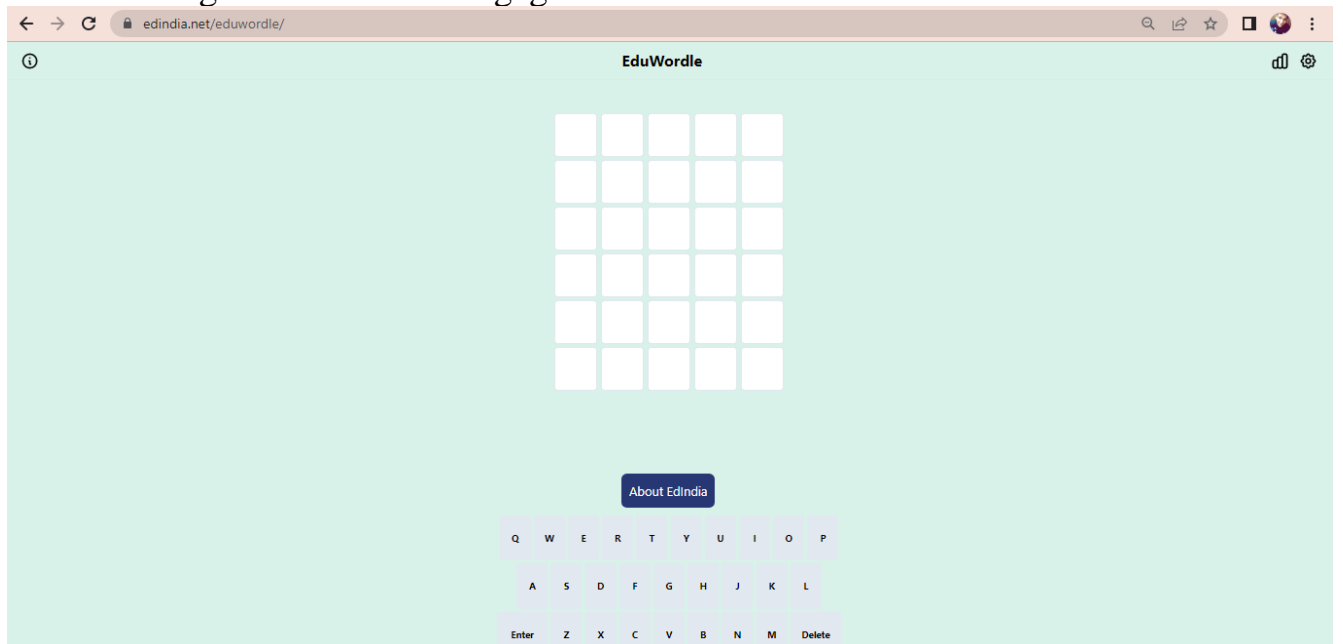- ✓ The next step was to create a dashboard around it.

❖ **Week 5-6-7**
- ✓ I worked with creating a dashboard template for Namsai Hamara Vidhyalaya dataset and then helping with the observation and dashboard creation.
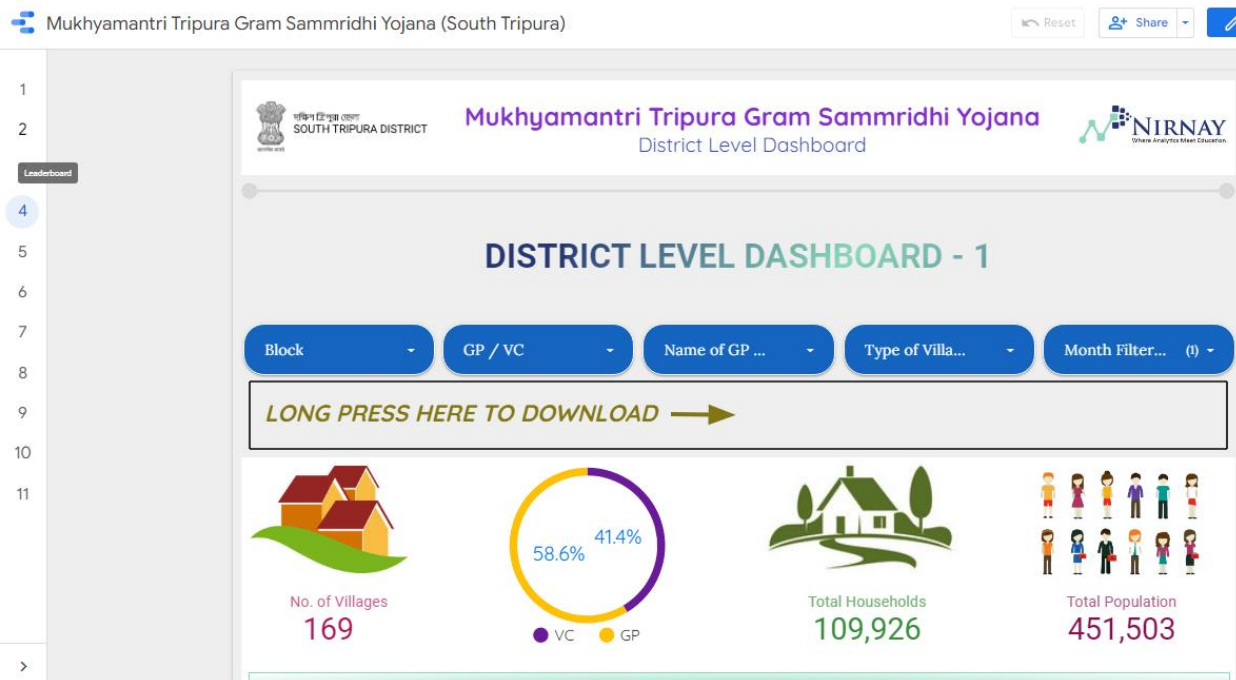- ✓ I also helped in modification of the new Sephaijala dashboard.

❖ Week 8-9-10-11

- ✓ I collaborated with the team in creation of an education related wordle game for teacher engagement and to boost their interest in education.
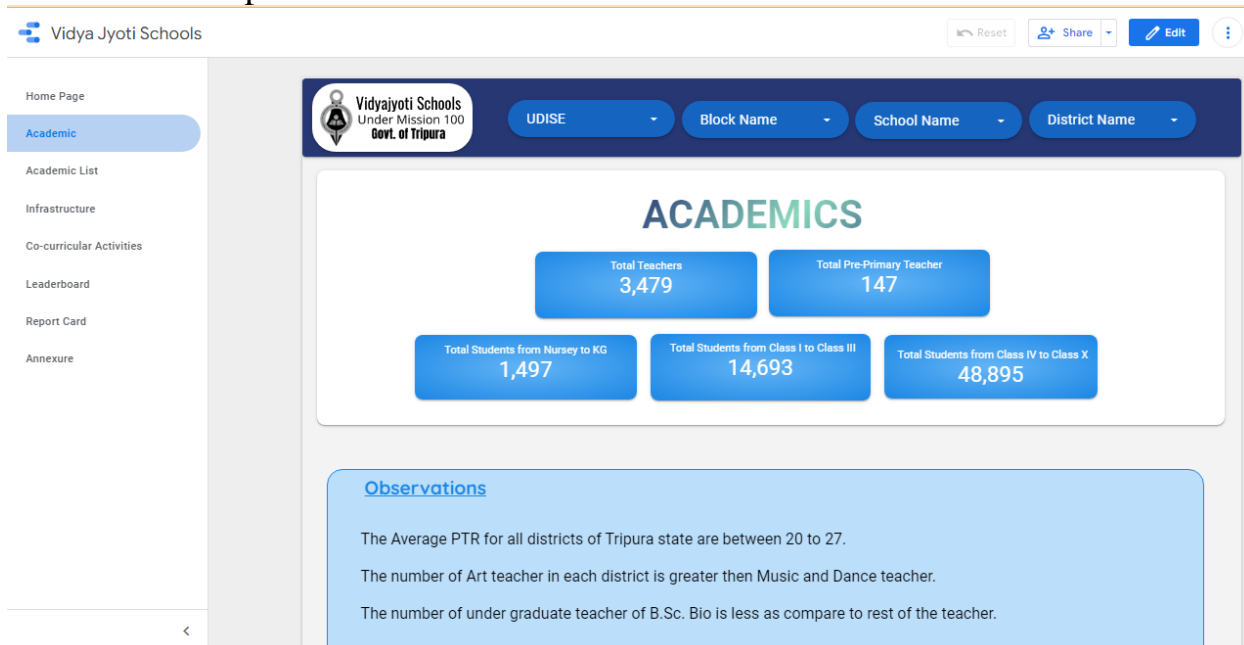

❖ Week 12-13

- ✓ Created a dashboard around the MTGSY dataset for administrative level Tripura data.
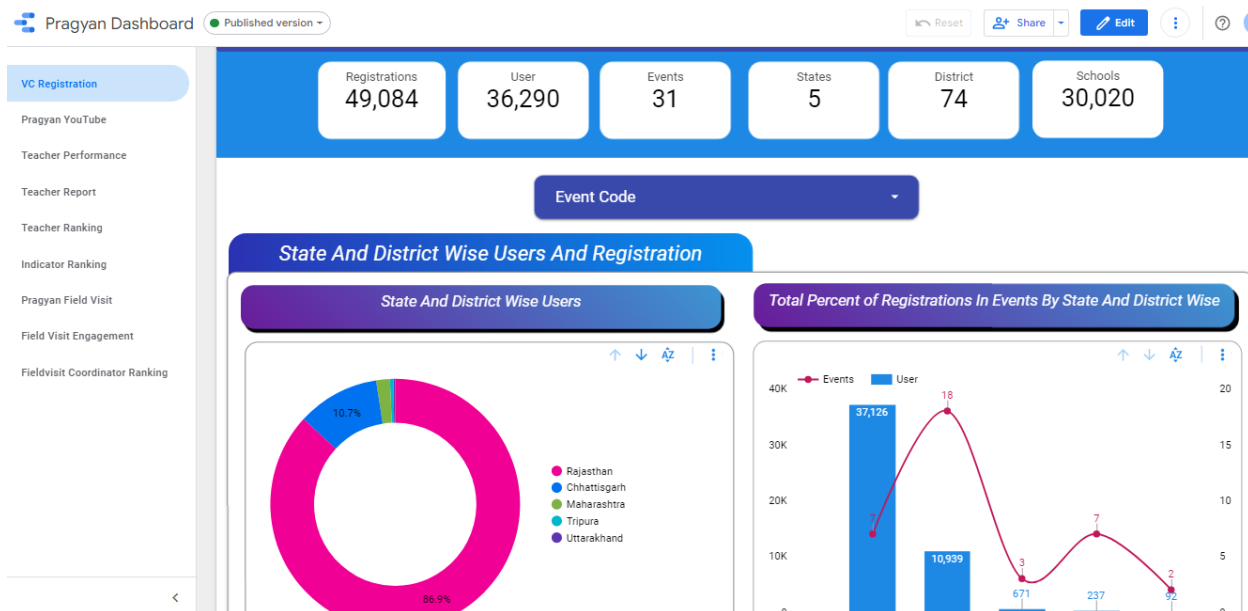- ✓ Created a wireframe for the Nirnay web application.

❖ **Week 14-15**

✓ Collaborating with my team I developed a Vidya Jyoti School CBSE board dashboard sample which helped us to impress the stakeholders and gaining that project.

✓ I also also created a dashboard around students' dropout rate from Tripura district.
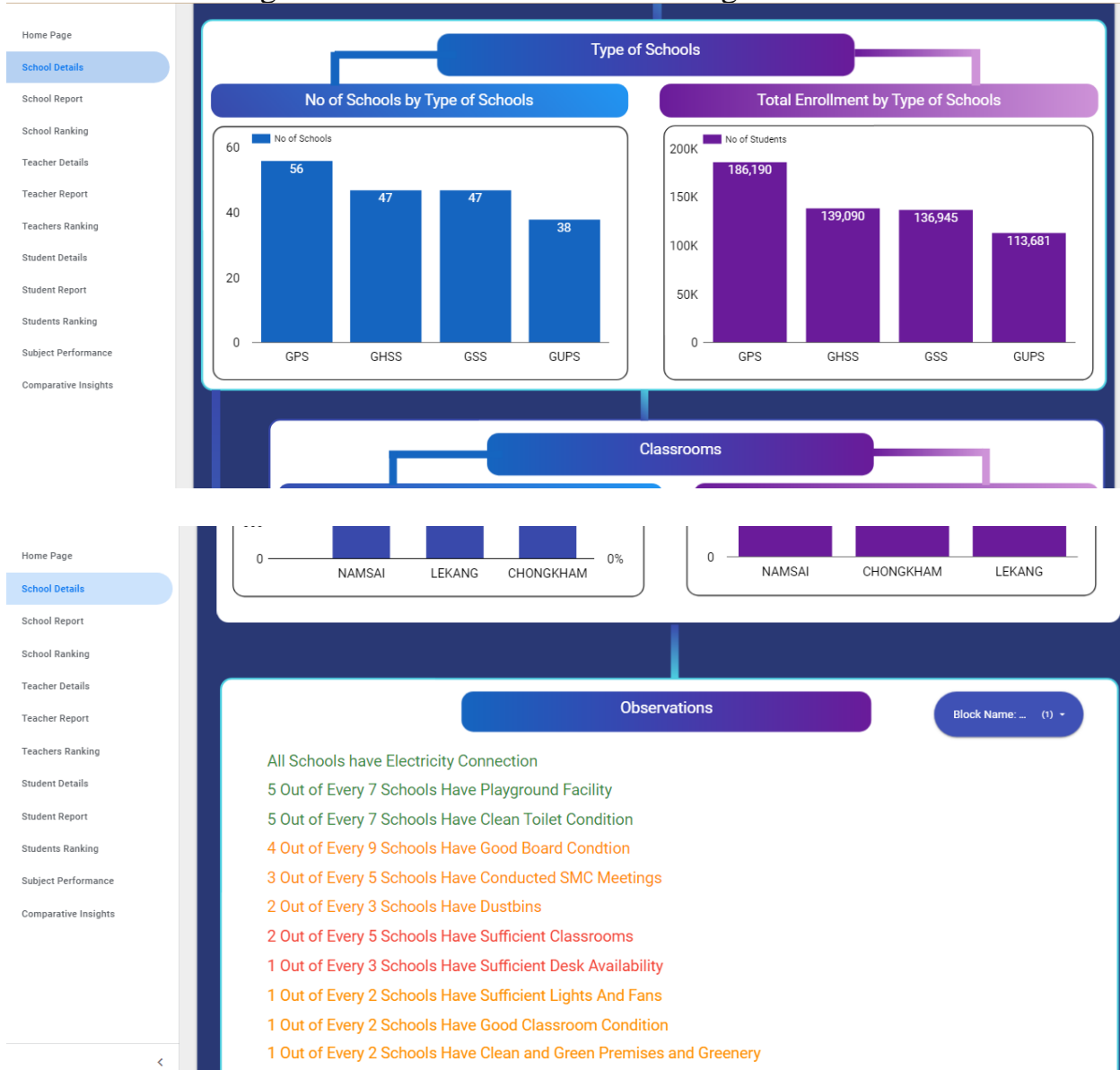
❖ Week 16-17-18

- ✓ Created dashboard for Tripura student baseline dataset which is internal assessment of students based on two subjects which is English and Math.
- ✓ Created a dashboard around Pragyan YouTube data, Pragyan Teacher assessment and Ranking data, Pragyan Mobile App Engagement data and Pragyan Field Visit Tracking data.

❖ Week 19-20(Extended Contract)

- ✓ In my extended contract period, I was given the task to create an input model front end for data collection.
- ✓ I was also given the task to develop a story like dashboard.
- ✓ I also gave observation based on the insights.

Self-assessment :

During this internship I learnt many things like react.js, python, Flask, Google spreadsheet, google Bigquery, postgres SQL and google data studio. I developed a soft skill to work with a team and how to effectively manage time. As a data analyst I think I developed one of the most important skills that I represent data in a story format.

Self-learning has always left a positive impact on me, as it gives me confidence to achieve something on my own. There are so many new things to learn and there are also lot of resources available, the key is to be patient and give some time to learn things. Industry experience made me realize that every person needs to go beyond his/her syllabus because syllabus always lack what industry is already working with. I got to work in a team which improved my communication skills.

I am grateful towards Mr. Varun Bhamre, at EdIndia Foundation and the entire team of Nirnay for providing guidance and also for helping me whenever I had queries. I achieved an experience which definitely made change in my life, work style and personality as well