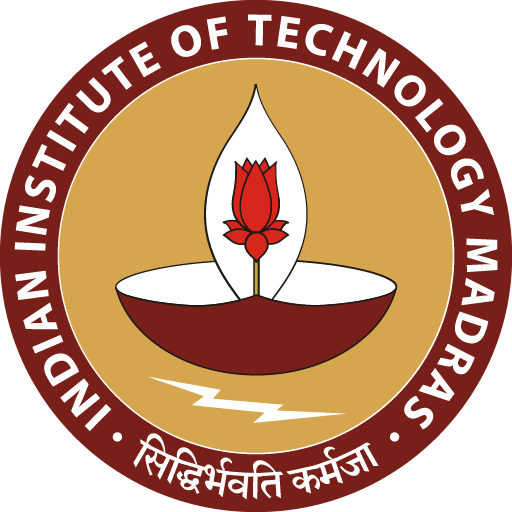
Pioneer Kochi, a case study

BDM Capstone Project Final Submission

Pravin Anavaratham M

[22ds2000111@ds.study.iitm.ac.in](mailto:22ds2000111@ds.study.iitm.ac.in)

MAY 2025 TERM



INDIAN INSTITUTE OF TECHNOLOGY, MADRAS, CHENNAI TAMIL NADU, INDIA, 600036

(BS) DEGREE IN DATA SCIENCE AND APPLICATIONS

|  |  |
| --- | --- |
| **INDEX** | |
| **Topic** | **Page No.** |
| Executive Summary | 2 |
| Detailed Explanation of Analysis processes and methods | 3 |
| Results, Findings, and Interpretation of Results | 9 |
| Recommendations for the Problem | 17 |
| Proof of Originality of Data | 20 |

Executive Summary

This report explores the challenges faced by **Pioneer Kochi**, a coaching center in Kerala that prepares students for government exams and engineering college admissions. After the COVID-19 pandemic, the number of student enquiries dropped, and competition from other online education platforms increased. At the same time, converting enquiries into actual admissions has been difficult—only 2 out of every 10 enquiries result in a student joining. This project focuses on finding the reasons behind the drop in enquiries and suggests low-cost ways to reach more students. It also aims to improve the admission rate from 20% to 40% by using data to better understand and manage student leads.

The project began by collecting data from various sources including student inquiries, course details, payments etc. Due to the absence of a structured database system, a considerable amount of effort was devoted to cleaning, validating, and organizing the raw data. Enquiry records were cross-verified with fee payment data, and a robust data model was created to analyse performance across key variables like courses, batches, marketing efforts, and student demographics.

Power BI was used to create dashboards that provided valuable insights into business performance. Some notable findings include:

* Only 20% of inquiries resulted in course enrolment, indicating a gap in lead conversion.
* A large portion of inquiries were left unaddressed, suggesting a lack of consistent follow-up.
* Most student leads were concentrated around **Kochi**, signalling where marketing efforts should be intensified.

Based on these insights, several practical suggestions were made:

1. Deploy a Customer Relationship Management (CRM) system to streamline lead tracking and improve follow-ups.
2. Enhance staff training by standardizing communication protocols and ensuring quality engagement with prospective students.
3. Boost the institute’s presence on social media to connect with a younger, more digitally engaged audience.
4. Launch a dedicated website and leverage existing online learning platforms to offer hybrid courses.
5. Explore offering short-term, customized courses tailored to the local market demand.

A SWOT analysis revealed that Pioneer Kochi’s key strengths lie in its faculty and reputation, while weaknesses include limited online visibility and manual data management. Opportunities exist in hybrid learning and outreach to local institutions, though threats from online competitors and changing student preferences remain significant.

Overall, this project highlights the importance of using data to inform strategic decisions. The insights and recommendations outlined can help Pioneer Kochi optimize its operations, increase student enrolment, and remain competitive in a fast-changing educational landscape.

# Detailed Explanation of Analysis processes and methods

**Data Analysis journey of Pioneer Kochi**

The data analysis journey at Pioneer Kochi came with several challenges. It began with understanding the business and identifying the key issues it was facing. Once the problems were clear, relevant data was collected, cleaned, and transformed for analysis. This analysis helped in identifying important performance indicators and creating metrics that could guide strategic planning and execution.

**Data Collection**

Data collection was the most difficult part of my analysis. Pioneer Kochi is a B2C business where income was generated by giving services to students who are aspiring for several Central and State Government Jobs in areas of specialized technical and non-technical domains, also focuses on students who want to enter into Engineering colleges in Kerala after completing diploma in a technical field. I was able to convince the owner in sharing the details. The owner was a friend of mine and at that point of time firm was running through some crisis and was in rebuilding phase after COVID-19.

He had some limitations as he was operating a franchise of Pioneer Kerala (a popular coaching institute). Some of the operations were controlled by franchise such as digital marketing.

Understanding business was the major challenge. I visited the center multiple times and tried to understand the course offerings. Later I got the enquiry sheets for various years. The data was not structured well. Enquiry and student information represented in a single sheet. My earlier analysis was baselined only on the enquiry sheet that I gathered from the owners. The enquiry sheet was maintained by the operator which was a simple excel file without additional layer of data quality checks. Attached link which I received earlier. [Link](https://drive.google.com/drive/u/0/folders/1tUi3I-_BUHexsQURrRmmTX02Un7VcooH)

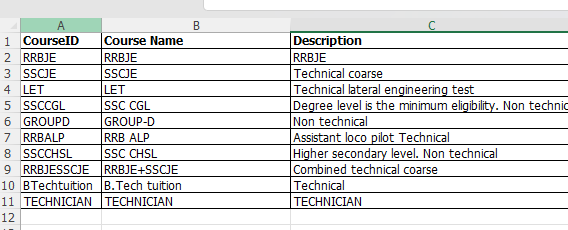
I had no control over the past data that were captured also many of the students who were in the sheet had already moved on.

To address the above issues, I worked with business to reconcile enquiry data received for the year 2024. I received fees payment information for 2024 this data was reconciled with enquiry sheet to ensure data quality.

[Enquiry](https://docs.google.com/spreadsheets/d/1N7LtBXyYpdRps6muj9VNYQKIUT5i6X38/edit?usp=drive_link&ouid=103614183911011697164&rtpof=true&sd=true)

[Fees](https://docs.google.com/spreadsheets/d/1ayH6a9ix3ycaoOiveZarBrKoFZU47z0P/edit?usp=drive_link&ouid=103614183911011697164&rtpof=true&sd=true)

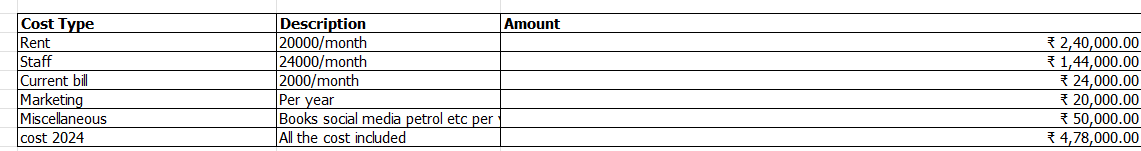
Gathering coarse information had its own challenges. Each course has a mixture of technical and non-technical modules. Every time I spoke with owner I had lot of information, once I go back and see the enquiry data I had several other questions. In order to handle that situation, I started sending voice notes through WhatsApp. The questions were specific and the answers were precise on to point. A particular course was operated in 1 or more batches in a year. Each batch had a schedule of 3 to 6 months. Due to limitation of space a particular batch has weekly 2 to 3 classes based on availability of lectures. Following are the list of courses offered in the year 2024.



Sharing the coarse details in the link for reference

[Course](https://docs.google.com/spreadsheets/d/1yWZ3a4_5Bucgj7BAjMGjhDcFOqB8618d/edit?usp=drive_link&ouid=103614183911011697164&rtpof=true&sd=true)

Understanding the expenses of the firm is an important step in my analysis. I had several interactions with the owner and came up with following fixed expense details. This does not include lecture expenses which varies for course to course



**Data Cleaning and transformation**

Data cleaning was a tedious task with all the data shared files were manually inputted and had no data quality checks. Some of the data cleaning tasks performed are correcting the date formatting issues, correcting categorical information in the file.

**Enquiry**

Enquiry sheet is all about the leads. Leads data is mostly generated in following ways

1. Person directly contacts Pioneer Kochi by seeing local advertisement, posters, notices and newspaper advertisement
2. Contact information shared by Pioneer headquarters. People visiting the website manually entering the information
3. Pioneer marketing staffs directly reaching out to students and institutions to market course offerings after a notification comes.
4. Existing student referring his friend

Data obtained is entered into the enquiry sheet by the telephone operator who connects with potential student and updates the information. The method is quite manual and is prone to high error. I have manually checked the data with Fee data to verify data entered are correct

In the Enquiry sheet, I have removed inaccurate college information to retain only the data I am fully confident about. Additionally, I have introduced a Comments Category column to systematically classify the feedback. This categorization will help in understanding the reasons why an aspirant did not proceed with enrollment.

|  |  |
| --- | --- |
| **Column** | **Description** |
| EnquiryID | Random Id generated for a particular enquiry info |
| StudentName | Name of the student |
| EnquiryDate | Date when student was contacted first |
| Contact | Contact number |
| Location | Location of the student |
| Qualification | Qualification |
| Branch | Additional information on the speciality of Education. Eg: Mechanical Engineering |
| College | College where the aspirant completed education |
| CourseID | Coarse he is interested |
| BatchType | Regular or Week end batch |
| KnowAboutyou | How do they know about Pioneer. |
| Comments | What is their thought on joining. A quick note on the conversation |
| Comments Category | Based on response comments can be categorized. Added field for analysis purpose |

Please find the link for enquiry sheet

**Course**

Course data was prepared with continuous interaction with business owner. Understanding the daily operations was the ultimate challenge. The cycle starts when a notification comes for a particular job say RRBJE. Coarse offerings are marketed through local newspaper, notices, websites etc. Lectures are temporary staffs who are contacted based on demand. Fees planned for the course is based on number of hours required to complete the course. Number of hours is further divided into hours which needs technical expertise and non-technical expertise. Technical expertise means lectures needs to be subject matter experts. Upon discussion, I understood that lectures were scheduled up to 3 days per week. Each day consists of 5-hour lectures for a course. Lectures had to be paid on an average Rs 400/- for non-technical course and Rs 700/- for technical course. With this information I was able to calculate the lecture cost of a batch. The formula for lecture cost is 400\* (No of non-technical hrs) + 700\* (No of technical hours)

|  |  |
| --- | --- |
| **Column** | **Description** |
| CourseID | Identifier for Course |
| Course Name | Name of the course |
| Description | Brief description of the course |
| Coursefee | Course fees planned |
| HrsRequired | Number of hours required to complete the course |
| No\_Modules | Number of modules in a course |
| No\_Tech\_Lectures | Number of technical lectures required |
| No\_NonTech\_Lectures | Number of non-technical lectures |
| No of Days | Number of days required to complete the course. Estimated based on 5 hours per day. |
| No of weeks | Number of weeks required to complete the course. Calculated field based on 3 days per week schedule, 5 hours per day |
| No of Months | Number of weeks required to complete the course. Calculated field based on 3 days per week schedule, 5 hours per day |
| No of technical hours | Number of technical hours for the course |
| No of non technical hrs | Number of non-technical hours |

**Student**

Student table is designed to include people who have at least made 1 payment towards a course. The purpose of the table is to establish relationship among different dimensions of the data. This table has references to enquiry, batch, course etc.

|  |  |
| --- | --- |
| **Column** | **Description** |
| StudentName | Name of student |
| Contact | Personal contact number |
| BatchID | Batch number |
| CourseID | Course ID |
| Location | Location student hails from |
| Qualification | Educational qualification |
| JoinDate | Course joining date |
| CompleteDate | Course completion date. |
| FirstFeeDate | Date when the first fees payment was made |
| Discount | Any discounts provided |
| Status | Course completion status |
| EnquiryID | Reference to the enquiry |

**Batch**

Courses are mapped to batches. Some of the courses were conducted more than once in a year. To track this for reporting purpose I have mapped courses to batches. Batch has a start date and end date. Batch information is derived out of the payments data using student information. This was again reconciled with business owner to set batch start date and end date.

|  |  |
| --- | --- |
| **Column** | **Description** |
| Batch ID | Random Id for batch |
| Batch Code | Batch code |
| Batch Name | Name of the batch |
| CourseID | Course id. Foreign key to course |
| BatchStartDate | Start date of batch |
| BatchEndDate | End Date of batch |

**Payment**

Payment data was the cleanest data in the files that were shared with me. It had major details required for analysis. Some of the interesting observations include fees not complete. Upon checking I realized that they were either given discount or stopped the course in middle or students who did not pay fees after completion of the course. Payment data has voucher no to uniquely identify a transaction. Each student makes payment one or more times during the course life time. Batchid info is used to refer the batch where the student is. StudentId is used to set up relationship with student table

|  |  |
| --- | --- |
| **Column** | **Description** |
| PaymentID | Payment identifier |
| StudentID | Id corresponding to student |
| BatchID | Batch id to refer Batch info |
| VNO | Unique alpha numeric to record payment |
| InstallmentNo | The payment is associated with an instalment made by a student. |
| JoinDate | Course joining date |
| PaymentDate | Date when payment was made |
| PaymentAmt | Amount correponding to a voucher number |

**Location**

Location information is taken out of the enquiry sheet. Distinct location name is used to find latitude, longitude data

|  |  |
| --- | --- |
| **Column** | **Description** |
| Location | Location Name |
| Latitude (N) | Latitude |
| Longitude (E) | Longitude Info |

A data model was created for analysis purpose. The cleaned file is available in the shared folder.[**Model**](https://docs.google.com/spreadsheets/d/1C6yB2xbXz1sIPN9b0-bTQ6xJ0l6Pv4As/edit?usp=drive_link&ouid=103614183911011697164&rtpof=true&sd=true)

**Tools**

Power BI is used for data analysis. Refined clean file is uploaded to Power BI to create data model required to produce data visualizations. Pbix file is available in the location [Folder](https://drive.google.com/drive/u/0/folders/1CN-eug1bkWNd_YeIIn_swkCYx9TkUoMW) with name Study\_V4.pbix

Results, Findings, and Interpretation of Results

**Enquiry Data Analysis**

Enquiry data is the lead data that the business got from multiple mediums. Important channels include Internet, website, newspaper and peer group. The target group for marketing are high school to graduate level educated people who are in the age group of 18 to 32. Majority of competitive exam eligibility fall in this range.

Summarized enquiry data is given below.

|  |  |
| --- | --- |
| **Lead Summary** | **Value** |
| Total number of enquiries | 1005 |
| Total Joined Students | 200 |
| Total did not join | 805 |
| Conversion Rate | 20% |
| Number of courses Planned | 10 |
| Number of Courses Executed | 6 |

## The total enquiries include all prospective students who may or may not go on to join the course. This represents the target group. Pioneer group has marketing channels which makes use of digital platforms such as Sulekha for lead gathering. Some of these leads may not be accurate. The conversion rate gets affected due to this.

## Analyzing the enquiries a little bit further gives valuable insights.

## Enquiry summary

## The summary dashboard is intended for a quick view of lead generation and conversion. I have added filters so that business owner can fetch data based on dates, course, batch, lecture etc. One of the important objectives of the project is to increase lead generation. Business needs to set short term and long-term strategic goals to continue running in this fast-paced world. The dashboard helps owner see how many have joined between certain dates, which are the courses generating maximum enquiries what percentage of enquiries have been converted to enrollment. Which are the locations that are leading in lead generation

## Fig1-Enquiry Report

## 

## Course wise Enquiry distribution

## Figure 2 (Course-Wise Enquiry) illustrates the distribution of enquiries across all courses offered in 2024. RRB ALP received the highest number of enquiries, totaling 406. In contrast, BTech Tuition, a course that was planned but eventually discontinued due to zero enrollment, sits at the lower end of the graph. SSC CHSL also saw very low enrollment, leading to financial loss for the company.

## It’s important to note that every course incurs a minimum cost in terms of faculty and resource allocation, which varies based on the balance of technical and non-technical instructional hours. To be able to kickstart a course a batch needs minimum students to get enrolled. It is easy to scrap a course if nobody enrolls but if very few enrolls it becomes difficult to scrap the course as the company needs to meet the obligation to the student. On a different note, RRB ALP, was offered in multiple batches last year.

## Fig2 Course-wise Enquiry

## Enquiry vs. Join Trends by Month Year

## The graph is intended to check any particular pattern in enquiries and course enrollment. It is observed beginning of the year has a definite spike in enquiry and joining

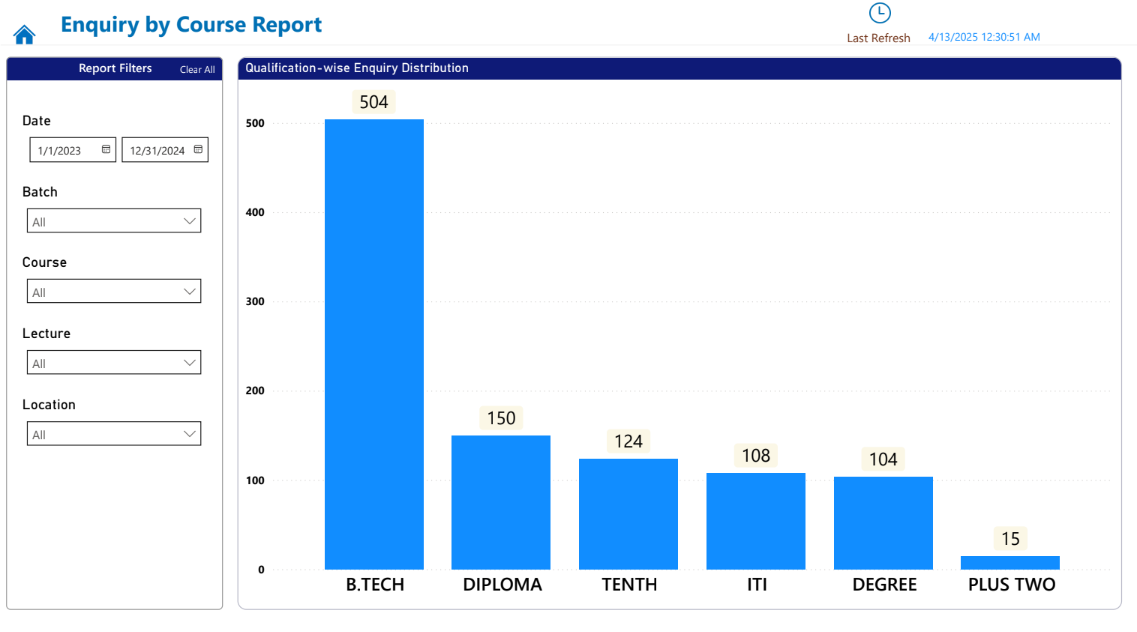
## Fig 3 Enquiries and Joins detail Report

## 

## Qualification-wise Enquiry Distribution

The objective of the analysis was to understand the qualification levels of the aspirants. B.Tech-qualified engineers form the core target group for Pioneer, followed by diploma holders, although their numbers are comparatively much smaller. The analysis is quite reflection of educational institutions in Kerala. Majority of students opt for BTECH.

**Fig-4 Qualification- wise Enquiry Distribution**

****

## Marketing Channel Insight

## The analysis is used to understand which are the best channels by which aspirants can be reached. It is very clear that Internet is the best medium to reach out to aspirants.

## Next is Thozhil Vartha, a weekly publication extensively read throughout Kerala by individuals seeking job opportunities. Interestingly, despite being a platform that offers leads for potential students, Sulekha is not commonly used by course aspirants. Direct messages between aspirants and the firm are another source of leads. Broadcast messages like notice, posters also played role in lead generation. References through friend (internal reference) had a high potential for conversion. This channel needs to be used better. Join data based on marketing channel

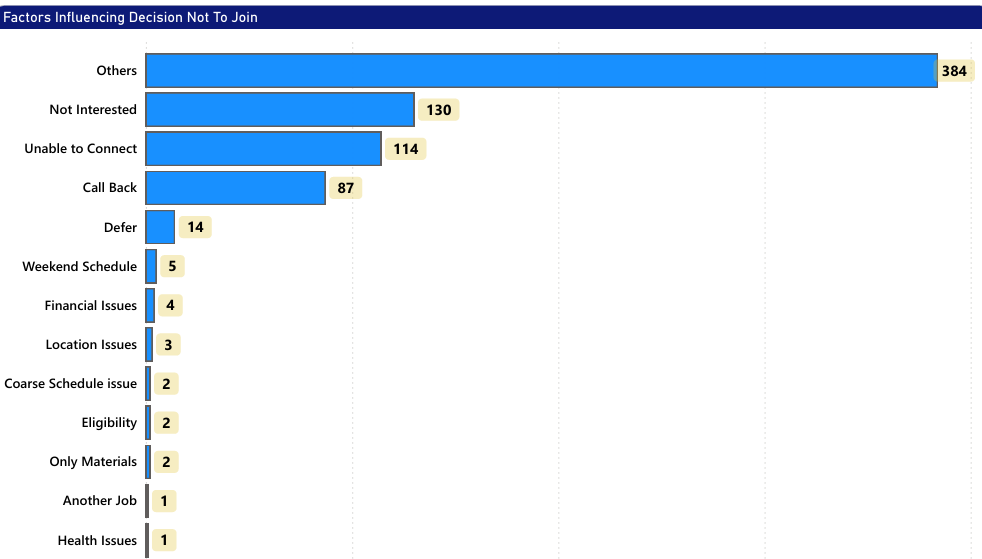
## Fig-5 Marketing Channel Insight Enquiry and Join

## 

## Factors Influencing decision not to join

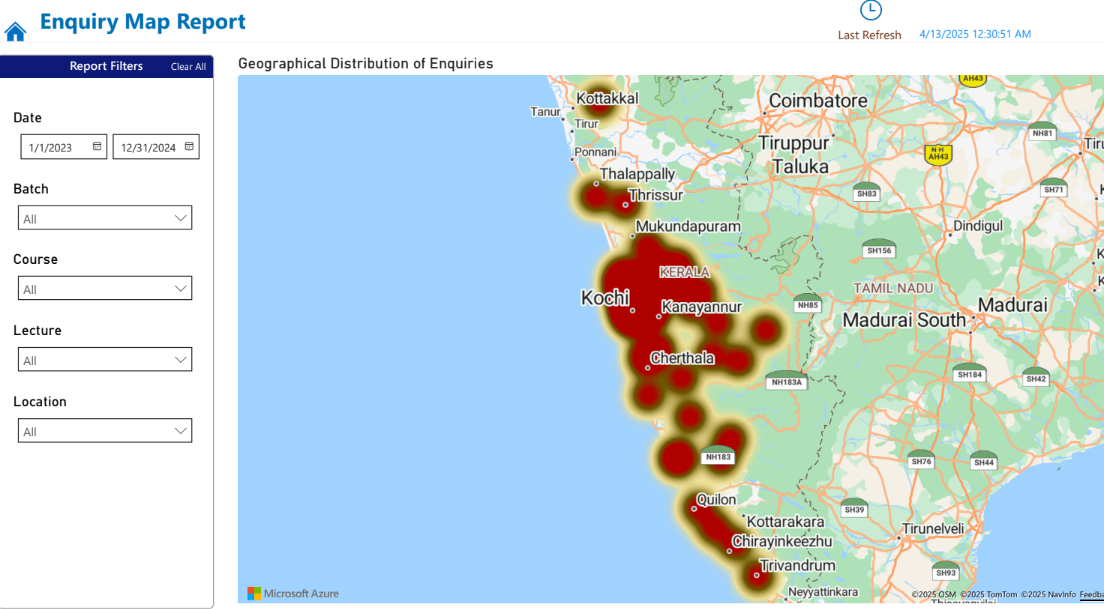
I wanted to see what all factors are making aspirants to drop the idea of joining the institution. The majority of individuals did not disclose their reason for not joining. While some explicitly stated a lack of interest, others could not be reached. This points to a potential issue in the lead generation process. Additionally, 87 individuals mentioned they would call back—indicating possible interest and representing a missed opportunity for enrollment.

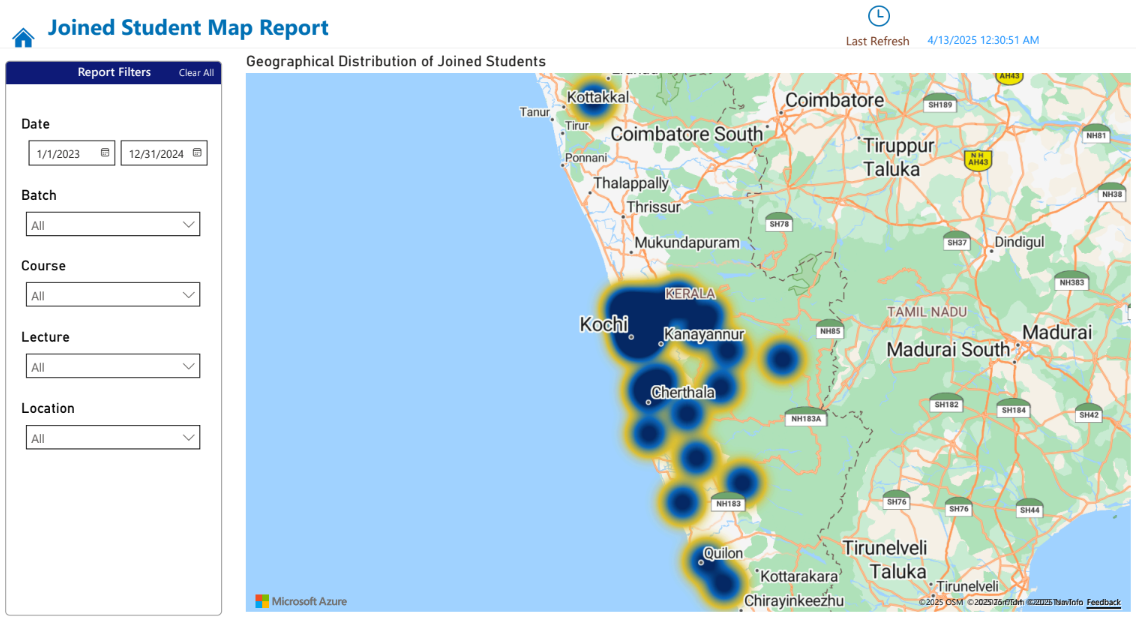
**Fig- 6 Factors influencing Decision not to join**

****

**Location Analysis**

The objective was to identify the regions from which the next batch of students is likely to emerge. The analysis clearly indicated that marketing efforts should be concentrated in and around Kochi. The geographic distribution of both enquiries and enrolled aspirants has been visualized through a geo-map for better insight. Direct marketing should be carried out in these areas to tap in more leads in future. Kochi is the most important city in Kerala. Location of the institution holds a strategic advantage. Within 5 KM radius there are railway station, bus stand, and metro. These are the common transportation options for common man

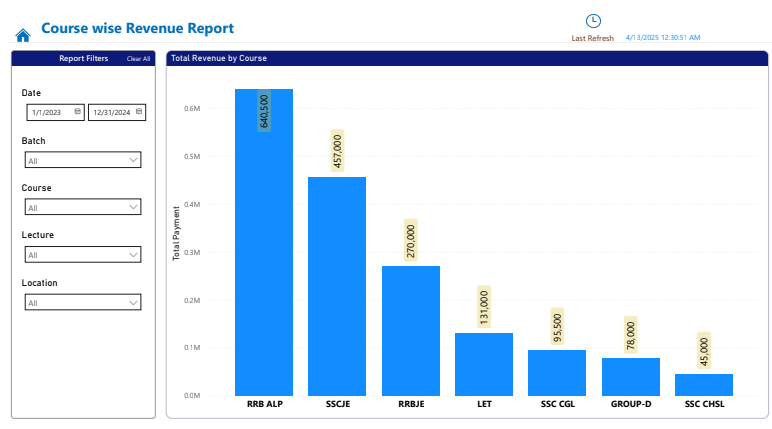
**Fig -7 Geographical distribution of Enquiries**

**Fig -8 Geographical distribution of Enquiries**

**Revenue Analysis**

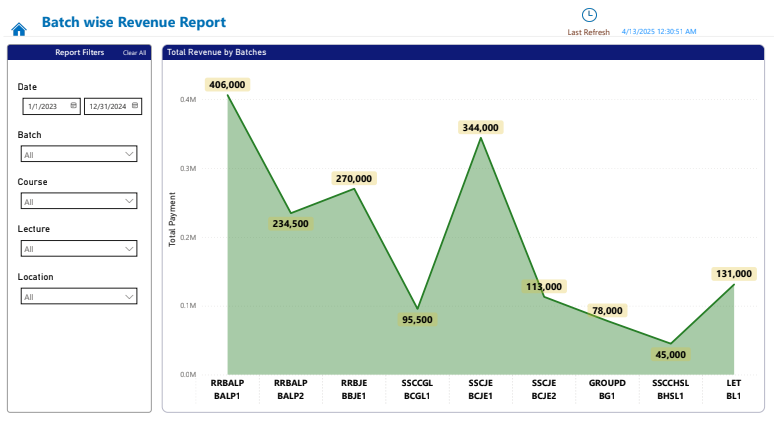
Fee payment by students is the ultimate source of revenue of Pioneer Kochi. I tried to analyze revenue in different dimensions. The chart shows which are the revenue generating courses in terms of value.

**Fig-9 Course wise Revenue**



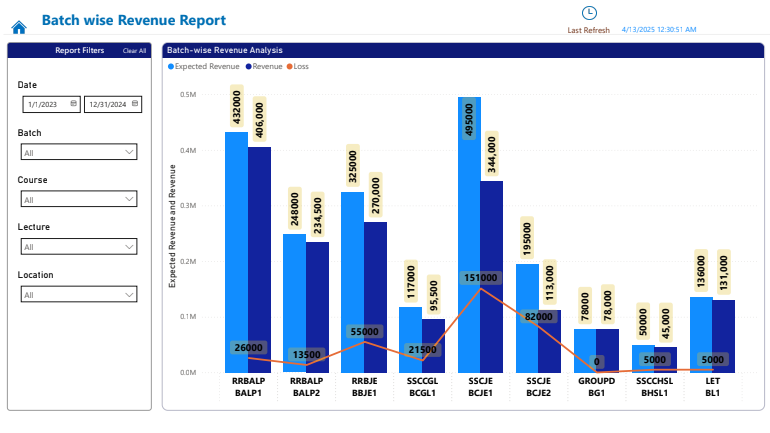
Following shows the revenue by batch. It is interesting to know that RRB ALP had 2 batches and first batch was way more successful than the second similarly for SSCJE.

**Fig- 10 Batch wise Revenue Report**

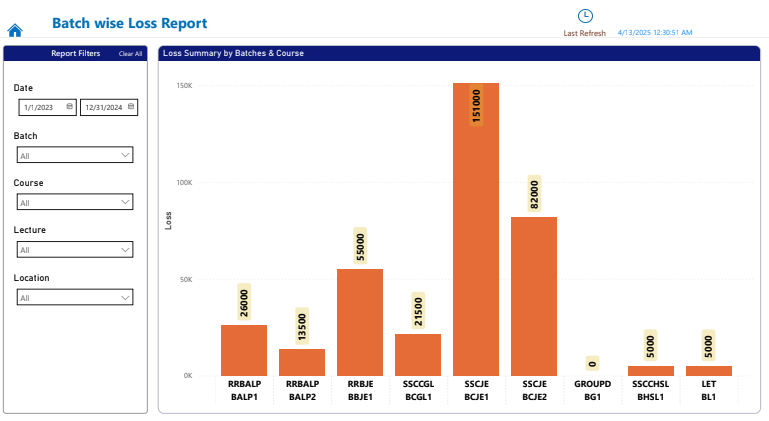
****

**Fig - 11 Forecast Revenue, Actual Revenue and Loss**

Following graph shows the revenue loss due to nonpayment of fees by students. Expected Revenue shows amount expected, Revenue is the amount paid and Loss is the difference. Most of the courses had nonpayment of fees by student issue. This is a severe issue for the growth of company



The figure shows losses per batch. Group D BG1 incurred 0 loss where as SSCJE BCJE1 incurred heavy loss of Rs 151000. This is mainly due to students dropping the course ahead of the schedule which needs to be investigated

**Fig 12 Batch wise loss report** 

Recommendation for the Problems

**Data-Driven Conclusions:**

The insights suggest that there is room for improvement in the firm’s operational efficiency. At present, a common strategy is used to manage a variety of situations. Typically, the process begins when a competitive exam is announced—leading to marketing efforts, course planning, and the formation of batches with assigned students and faculty. This approach tends to work well in scenarios where demand is high and supply is limited. However, with the changing landscape and increasing competition, it may be beneficial to revisit and adapt this strategy.

**Leads Management**

Leads represent a critical source of potential income and must be managed effectively. While Excel offers a convenient way to record information, it is not ideal as a long-term data repository. Without proper access controls and validation checks, data in spreadsheets is vulnerable to errors and tampering. Unlike structured entry systems used in websites or company portals, spreadsheets rely heavily on manual data entry, which can lead to inconsistencies. To address this, there is a growing need for a dedicated software solution—one that can automatically initiate a campaign upon receiving a new enquiry and send periodic reminders to both the operator and manager. This ensures that each lead is tracked throughout its lifecycle, either until enrollment or until it is marked inactive after a defined period of no response. The cleaned data enables concise and focused reporting, as demonstrated in the report. As part of the process, I have created an excel template which can be used for temporary reporting until we have a system in place. Please find the excel file used for Reporting. [Link](https://docs.google.com/spreadsheets/d/1C6yB2xbXz1sIPN9b0-bTQ6xJ0l6Pv4As/edit?usp=drive_link&ouid=103614183911011697164&rtpof=true&sd=true)

**Staff Education**

Educating staff and regularly reviewing their work is essential, as they represent the company to prospective students. Providing them with a structured handbook—including a checklist of key questions, along with clear dos and don’ts—can be an effective way to ensure consistency and professionalism in communication.

**Precision Marketing**

As observed in the analysis, the majority of leads originate from the parent company’s website. However, there is minimal marketing activity on social media platforms such as Facebook, Instagram, and others. Considering that potential leads for Pioneer who are in age group of 18 to 32 are active users of modern digital platforms, establishing a strong social media presence is essential. In today’s competitive landscape, having a solid digital footprint is not just beneficial—it’s crucial for business growth. Following is the digital presence of the institution. It can be observed that they have very fewer social media presence.

**Instagram**

[Instagram](https://www.instagram.com/pioneerkochi?igsh=MTUybjNyMzZqc3d0Mg==)

**Website**

[Pioneer Kerala](https://www.pioneerkerala.com/)

Facebook

[Facebook](https://www.facebook.com/share/1Bn5DNrpJF/)



As observed in the analysis, most high-quality leads have come through the company’s central website. However, Pioneer Kochi currently does not have a dedicated website of its own currently, primarily due to challenges in maintenance.

Parent company has a dedicated online learning portal called [Tuteetute](https://tuteetute.com/) which offers limited online learning. Pioneer Kochi needs to use the platform to market the course offerings outside geo locations where they have market access. Using the platform, they should offer students a hybrid learning opportunity.

It has been observed that a significant number of enquiries originate from Kochi and its surrounding areas. To capitalize on this opportunity, Pioneer Kochi should actively engage with engineering colleges and polytechnic institutions to directly promote its course offerings. Introducing customized learning plans and early bird offers can further enhance lead generation and attract more prospective students. Reaching out to school children is also a good option. Educational sessions with them on the prospective competitive advantages of starting early can motivate students to join the institution. Depending on the change in demands from the aspirants’ pioneer has to offer new short-term courses

**SWOT Analysis**

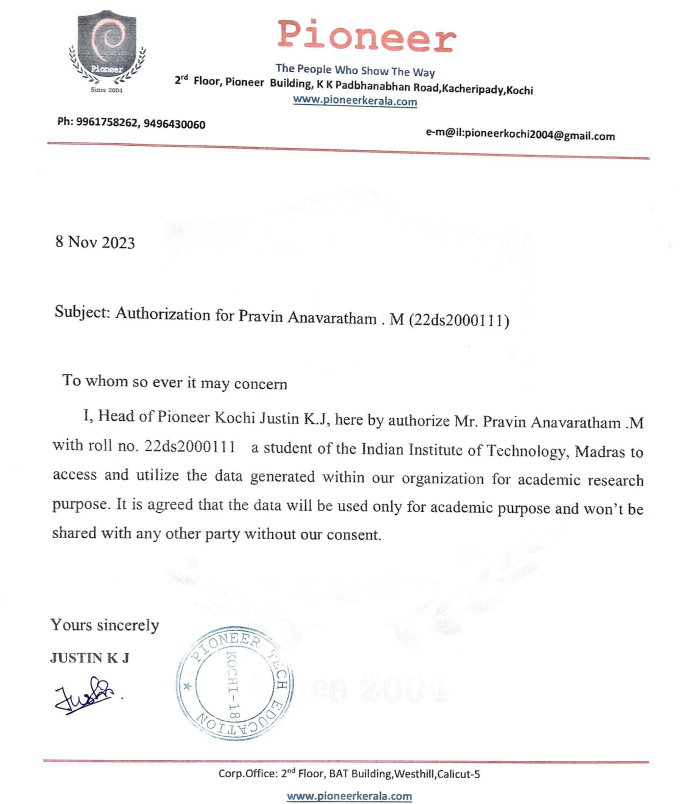
A SWOT analysis was performed to analyze the company in 4 quadrants of Strength, Weakness, Opportunities and Threat. The results are shown below.

|  |  |
| --- | --- |
| **Strength** | **Weakness** |
| - Experienced and qualified faculty - Strong results in competitive exams - Positive word-of-mouth and local reputation - Affordable fee structure for key courses  - Strategic location in the middle of city  - Support from parent company | - Nonpayment of full fees by enrolled students  - Manual lead tracking via Excel with no CRM system  - No dedicated website for Pioneer Kochi - Limited use of social media marketing - Low enrollment in some courses resulting in operational loss |
| **Opportunities** | **Threat** |
| - Growing interest in government jobs and competitive exams - Scope to introduce online and hybrid learning modules - Expanding marketing to areas around Kochi - Strategic partnerships with schools and local institutions | - Rising competition from online edtech platforms - Changing patterns and unpredictability in competitive exams - Student preference shifting towards digital learning resources - Lead loss due to inconsistent follow-up mechanisms |

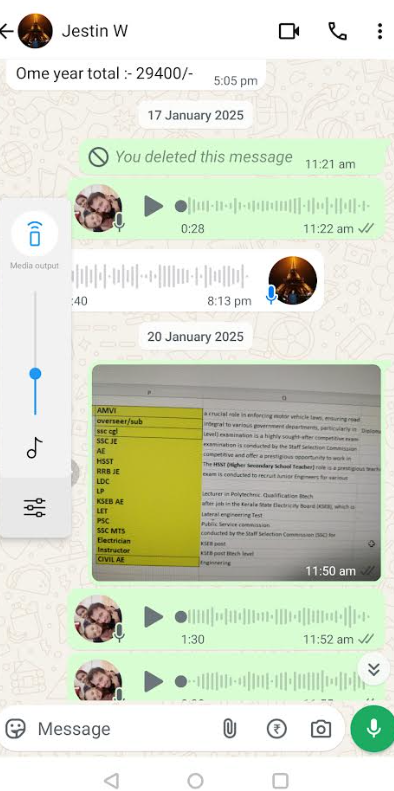
Following shared folder contains raw file, reports, field notes, model file used for model building and model screen shots

[Drive Link](https://drive.google.com/drive/folders/1CN-eug1bkWNd_YeIIn_swkCYx9TkUoMW?usp=drive_link)

Proof of Originality of data



**Company pictures**



**Discussion with business owner**