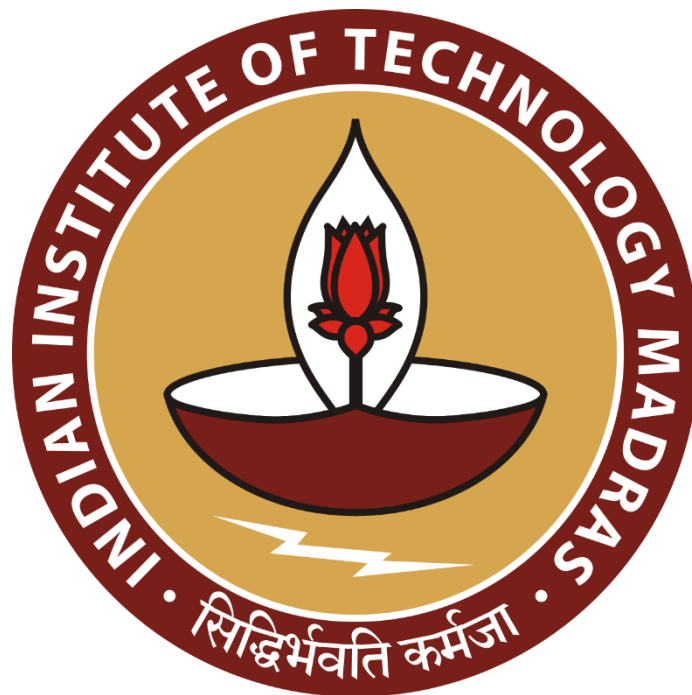# Improving Admissions Count and Student Performance through Data Analysis

Mid-Term Report for the BDM Capstone Project

Submitted by

Name: Krishna Varshney

Roll number: 22F3000605

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1   Executive Summary and Title

R G Institute of IT & Engg. is an education institution located in Shivaji Nagar Jhansi, running under SSSGM Society, with a vision to provide the best cheapest education to students of various communities, both urban and rural. The institute has been committed offering high-quality education and training to empower individuals. It is equipped with modern facilities, experienced faculties, and accreditation from various universities and government bodies like DOEACC/NIELIT, and has excelled in gaining a reputation not only in Jhansi but other cities too.

The institute has been facing some challenges from last year. It aims to increase the number of enrollments in its various courses and also improve the quality of grades of its students and develop methods to enhance the engagement of students in the courses. This project aims to address these challenges and provide effective solutions.

This report presents the initial analysis conducted. It includes detailed explanation of various processes including data collection, data cleaning and processing, along with the insights achieved from the students' enrollment and performance data. The initial analysis is mainly focused on understanding the enrollment patterns, various trends, and performance metrics across various courses offered. The analysis present in this report will help institute in getting some basic insights and also give the direction for further detailed analysis.

2   Proof of Originality

   a. **LINK TO LETTER FROM ORGANIZATION**

   b. **LINK TO IMAGES**

Figure 1: Visit to the Institute

Figure 2: Interaction with Founder

   c. **LINK TO INTERACTION VIDEO**

3 Metadata

The institute provides admissions to a wide variety of courses. These include projects from the government, courses which are provided by the government, and the institute's course. The details about which is given below:

Sr No	Program Name	Type	Category	Qualification	Duration	Fee
1	CCC	Certificate	Govt. Course	NIL	3 Months	4200/-
2	Tally ERP 9	Certificate	Institute	NIL	5 Months	5000/-
3	Desktop Publishing	Certificate	Govt. Project	10th Pass	5 Months	Free
4	PMKVY	Certificate	Govt. Project	10th Pass	4 Months	Free
5	Web Designing	Certificate	Institute	NIL	5 Months	6000/-
6	DCA	Diploma	Institute	12th Pass	1 Year	12000/-
7	O Level	Diploma	Govt. Course	12th Pass	1 Year	20000/-
8	BCA	UG	University	12th Pass	3 Year	25000/- per Year
9	MBA	PG	University	BBA Pass	2 Year	35000/- per Year
10	MCA	PG	University	BCA Pass	2 Year	35000/- per Year

In correspondence to this, I collected the students' data of different courses. The institute provided me with different Excel sheets containing course-wise student details. The data consists of the following fields:

   a. **Candidate Name:** Name of the student
   b. **Father/Guardian Name:** Father's name of the student
   c. **Gender:** Gender of the student
   d. **DOB:** Date of Birth of the student
   e. **Category:** Category of the student (Values- General, OBC, SC, ST)
   f. **City:** City the student belongs to
   g. **Course:** Course for which the student enrolled for

In addition to this, I also gathered performance data of students. It was also in the form of an Excel sheet. It contains the columns as mentioned above and also some other columns as below:

   h. **Roll No.:** Roll number of the student
   i. **Month:** Month in which the student gave the exam
   j. **Result:** Grade of the student for that course (Values- A, B, C, D, E, F, Absent)

LINK FOR THE ENTIRE DATA

4 Descriptive Statistics

The following descriptive statistics have been noted based on one-year data collected from the institute.

➢ The institute had a total of 1,709 enrollments across its various courses.

➢ In the past year, the institute has enrollments in 10 major courses, including 2 government courses, 2 government projects, 4 degree courses, and 2 institute courses.

➢ CCC and O Level courses have the most number of enrollments contributing to 1267 out of a total of 1709 enrollments.

➢ Students of age from 13 to 46 are enrolled in different programs.

➢ Gender distribution in the institute consists of 64% male and 36% female students.

➢ The institute has admission from local and 5 nearby cities.

➢ The current student performance rate is 76%.

➢ There is no monthly seasonality in number of students as enrollment for different programs occurs in different timelines. Therefore, the monthly average, variance, etc. is not significant.

5 Detailed Explanation of Analysis Process/Method

A. Visit to the Institute

For initials, I had a visit to the institute and met the director of the institute Mr. R. G. Gupta, who told me about the institute, their business model, and the problems they are facing. Then, he made me seek whatever help I needed from the institute's staff.

B. Data Collection

After going through the working process and problems, the first step was to collect relevant data. The institute had all the data about students and their performance in the form of Excel sheets made program-wise.

C. Data Cleaning and Processing

- After going through the data, I found a need to remove some columns that were of institute personal use and not significant for analysis. Also, I need to insert some extra fields to gain valuable insights.

- Certain columns contained the details of fee payment by the students, and some fields that the institute uses for the processing of admission. I removed those columns.

- **For student enrollments data:**

  I merged all the program-wise data into a single table and perform some **Pivot Table** operations on it. Further I realized that analyzing the age distribution of students can get me some valuable insights so I extracted age from the 'DOB' column using the following formula:

  =INT(YEARFRAC(F2,TODAY(), 1))

- **For students' performance data:**

  The institute provided me with the monthly results of students, however, in that sheet only two columns were there, **Roll Number**, and **Grade** of students. Therefore, I merged the student details table and result table into one, using the **VLOOKUP** operation on column Roll Number as follows:

  =VLOOKUP(A1, Sheet2!A:B, 2, FALSE)

  Also, in the table the grade was mentioned as A, B, C, D, F, and ABS, so to simplify I created another column named Status which had three values (Pass, Fail, and Absent) corresponding to the grade. This operation was done using the following formula:

  =IF(OR(L2="S", L2="A", L2="B", L2="C", L2="D"), "Pass", IF(L2="F", "Fail", IF(L2="ABS", "Absent", "Invalid Grade")))

  The data was then ready for analysis and gaining insights.

D. Data Analysis

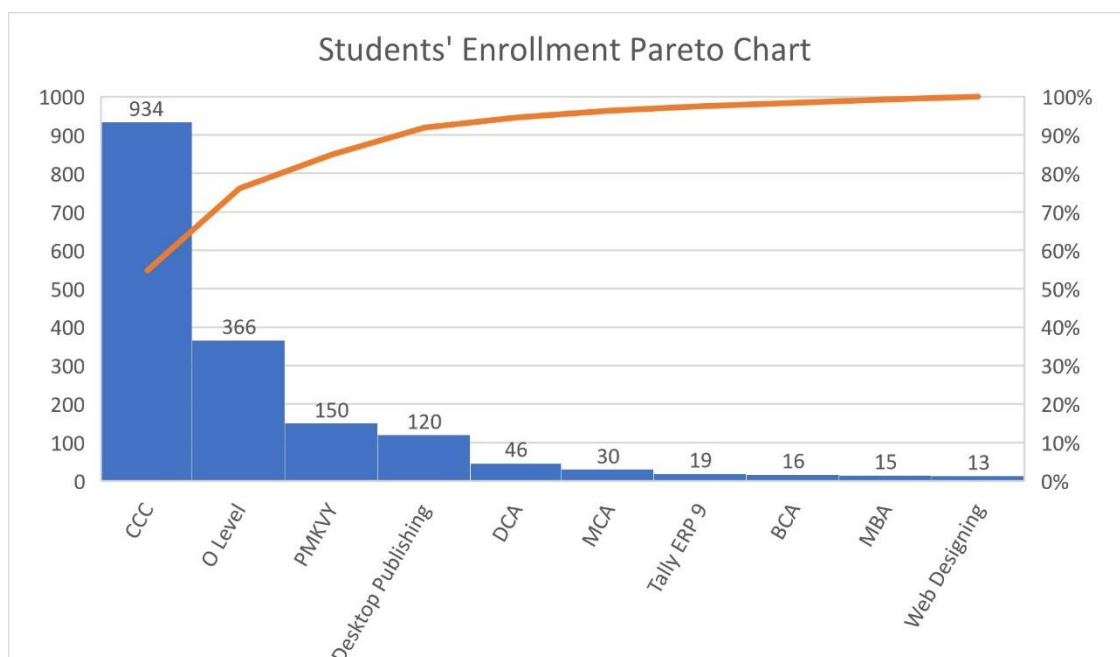Using Pivot Table and graphs I performed the following initial analysis on both the data I prepared:

Figure 3: Student Enrollments Pareto Chart
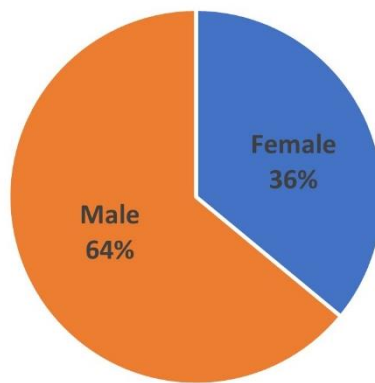
Gender Wise Distribution
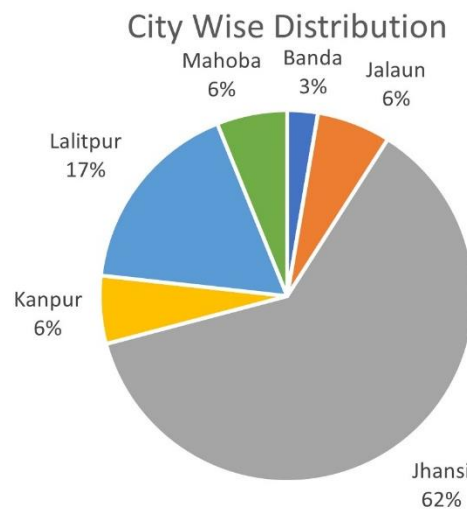

Figure 4: Gender-wise Enrollments Distribution

City Wise Distribution

Figure 5: City-wise Enrollments Distribution
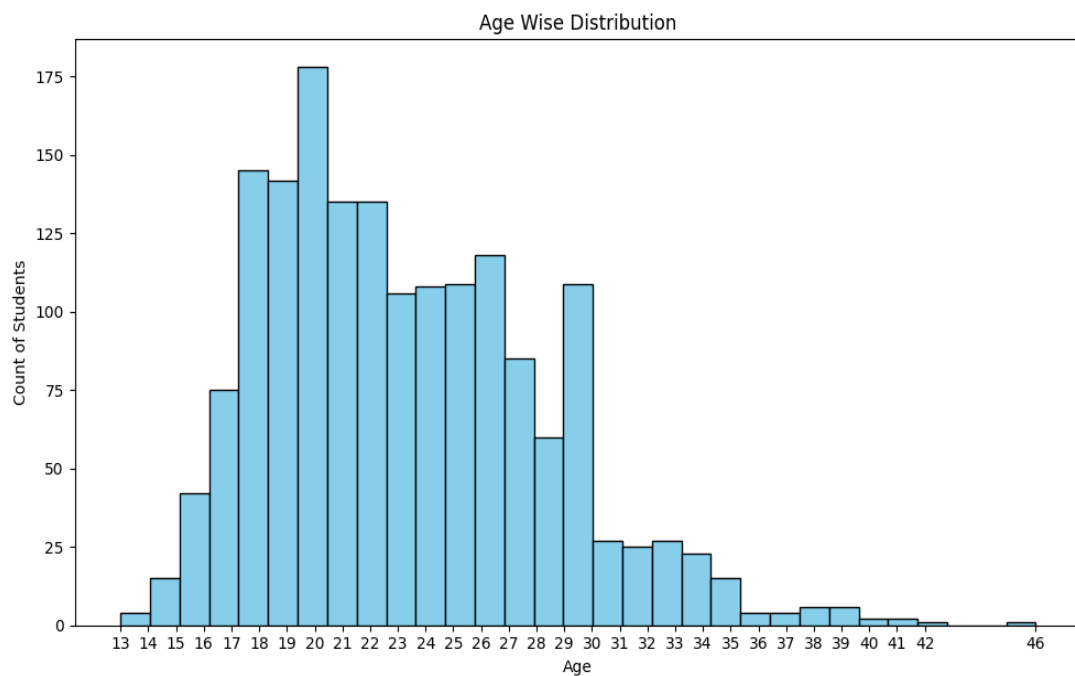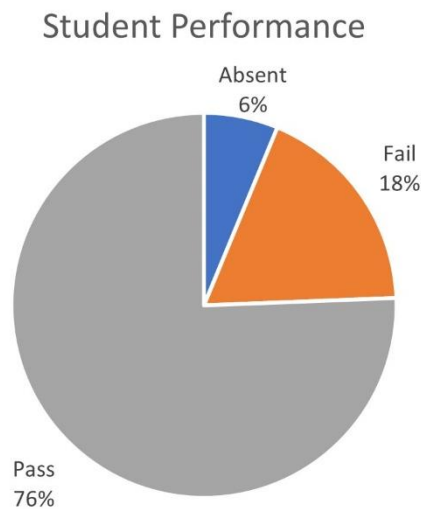
Age Wise Distribution

Figure 6: Age-wise Enrollments Distribution

Figure 7: Student Performance Chart

6 Results and Findings

➢ As shown in the Pivot Chart, CCC and O Level have the greatest number of enrollments contributing approximately 80% of the total. PMKVY and Desktop Publishing are government projects run by the institute and these courses contain a fixed number of seats. Also, the institute has a remarkably low rate of admission to other courses.

➢ As per the gender distribution, it's seen that female students constitute only 36% which should be a point to focus on by the institute.

➢ As the institute is located in Jhansi, it gets admissions from various cities of the Bundelkhand region which includes cities like Lalitpur, Kanpur, Mahoba, Jalaun, and Banda.

➢ Institute gets 62% admissions from local areas and 38% admissions from other cities. On showing the analysis to the staff responsible for enrollments, I was told that the percentage of admission from other cities has seen a significant rise in the past year which was only about 8-10% before that.

➢ From the age distribution graph, it is observed that the institute gets the most number of enrollments in the age group of 18-23, which are the students who completed their school or have completed their bachelor's degree and are seeking for master's or job.

➢ The institute could provide the result data only for CCC and O Level. The institute's current students pass record is 76%, while 18% of students fail the exams and 6% don't appear.

These results have been shown to the respective authorities, and further detailed analysis is to be done based on these results and feedback, which are going to be mentioned in the final report of the project along with the recommendations.