Business Data Management (BDM) Capstone Project

Case Study of Studify Consultants

Mid-Term Submission

Submitted by

Name: Shahbaaz Singh

Roll number: 23f3001356

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

Declaration Statement

I am working on a Project titled "**Case Study of Studify Consultants**". I extend my appreciation to **Studify Consultants**, for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered from primary sources and carefully analyzed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the principles of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I understand that all recommendations made in this project report are within the context of the academic project taken up towards course fulfillment in the BS Degree Program offered by IIT Madras. The institution does not endorse any of the claims or comments.

Signature of Candidate: Shahbaaz Singh

Name: Shahbaaz Singh

Date: 24th July, 2024

1. Executive Summary:

The project focuses on a medium-sized Private Education company 'Studify Consultants' located at SCO – 64, 2nd Floor, Sector 47 C, Chandigarh. The business is B2C and deals in the segment of Overseas Education and Travel.

In the Project Proposal I stated the major business issues that the organization is facing are related to lesser leads because of Big Fishes of the same Industry. This issue can be addressed once the target audience, countries that are preferred by the audience and some other things are clear. For example, analysis on age can be done to find which age group is most likely to go abroad for study or which age group prefer travelling outside country. Another example is that if we can analyze the places which most of the students decide to study, it can be used in the process of advertisements.

Another issue that I stated is resource credibility as in this industry employees sometimes leak the client data outside the company to get extra benefit. This issue can be addressed when the leads increase as when leads increases the profit of the company increases because then the employes would have more trust on the company, which will increase the retention and credibility of the employes.

I also stated employes prefer company which have advance technology, it meaning is that some companies have their own software for customer's data entry, this software provide the company about the deadlines of applications, fees, etc. This was the other issue I stated that whenever the employes have to find any data about a particular person they have to go through physical forms, various excel files in various folders which is very time consuming. For this I am trying to make a software, which will make the entry of the data easy and mitigate the problem of physical forms and excel files, hence increasing the productivity of the employees and decreasing the load on them which will increase their morale and trust, and this will result in more retention and credibility of the employes. This software will not be very advanced or fancy, it will be created with my knowledge of python, Tkinter, etc.

The data is provided by the company but it was missing some values and have unnecessary fields because of which the data was cleaned. The data that was provided to me was about 8 months and is of year 2023. The tools that will be used for the analysis of the data will be Excel and its inbuilt functions, features like graphing and pivot tables.

2. Proof of Originality of the Data:

Details:
Name – Studify Consultants
Founders/Owners – Mr. Gurmukh Singh and Mr. Mandeep Hooda
Directors - Mr. Gurmukh Singh and Mr. Mandeep Hooda
Address – SCO – 64, 2nd Floor, Sector 47 C, Chandigarh

About:
The company that I am working with is Studify Consultants; which is a for-profit private corporation and overseas education, residency and travel based in Chandigarh, Punjab. It

was founded by two partners Mr. Gurmukh Singh and Mr. Mandeep. The company was founded in 2005. In present Mr. Gurmukh Singh and Mr. Mandeep Hooda are acting as the Director of the Company. The company helps students getting Student Visa for education in abroad and also helps in providing tourist visa and Permanent Residency of other countries.
The company operates with a dedicated team of approximately 23 members within a 1,800-square-foot space, generating an annual revenue of ₹1.2 crore. The team is led by two directors and includes a general manager, four counselors, four tele callers, four members in the admissions cell, three individuals handling filing, and three professionals in HR (Human Resources) department.
Company Website Link - https://studifyconsultants.in/

Proof of Originality:
I have captured images, video clips and a short interview with one of the directors of the company Mr. Mandeep Hooda, as proof of originality. In addition, I also asked the company to provide me a letter on their letterhead with stamp and sign proving that they participated with me in this project. All the proofs have been organized into a folder in Google Drive, accessible through the link provided here:
Google Drive Link - https://drive.google.com/drive/folders/1yyrmTX4oxbgXOsvKTlrPdNxOvqHRZCxS?usp=drive_link

3. Metadata

The data I collected can be divided into two types, first type is Student VISA data which provide the information about the "Students" who applied for education VISA through "Studify Consultants" and the second type is Tourist VISA data which provide the information about the "People" who applied for tourist VISA through "Studify Consultants". Both the data sets were not containing the ages or date of birth of the applicant, so I asked the company about their ages so they provided me the scanned passport PDFs of the individual and I had to meticulously open every file and added the *Date of Birth*, and by using the YEARFRAC function from Date of Birth to *Date of Lodgment* I calculated the *Age*. Let's discuss about the Metadata of both the data sets in detail:

3.1 Student VISA Metadata**:**

There are 12 columns in this Data set, First column is *Serial No.*, Second column is *Student Name*, third column is *Gender* (Male/Female), fourth column is *University/College Name*, fifth column is *Canadian Province*, sixth column is *Course Taken*, seventh column is *Date of Lodgment* in the format DD-MM-YYYY, eight column is *Date of Birth* in the format DD-MM-YYYY, ninth column is *Age* in the format XX.XX because of the precision, tenth column is *Intake* in the format MMM-YY providing information which academic session the student took, eleventh column is *File Status* (APPROVED/REJECTED) and the twelfth column is *Attempt* providing the information about the VISA attempt through the company.

I have shared a portion of the data as an image:

Serial No.	Student Name	Gender	University/College Name	Canadian Province	Course Taken	Date of Lodgement	Date of Birth	Age	Intake	File Status	Attempt
1	GURKIRAT SINGH DHILLON	MALE	UNIVERSITY OF LETHBRIDGE	ALBERTA	BACHELOR OF MANAGEMENT	02-01-2023	06-06-2004	18.57	Sep-23	REFUSED	FIRST
2	JASPAL SINGH	MALE	CAPE BRETON UNIVERSITY	NOVA SCOTIA	BACHELOR OF ARTS	06-01-2023	25-02-2001	21.86	May-23	REFUSED	FIRST
3	SIMRANJEET SINGH	MALE	UNIVERSITY OF LETHBRIDGE	ALBERTA	BACHELOR OF ARTS	07-01-2023	03-10-2001	21.26	Sep-23	REFUSED	FIRST
4	MONIKA SHARMA	FEMALE	GEORGIAN COLLEGE	ONTARIO	PROJECT MANAGEMENT	09-01-2023	11-01-1996	27.00	May-23	APPROVED	FIRST
5	ROHIT KUMAR	MALE	SHERIDAN COLLEGE	ONTARIO	ADVANCE DIPLOMA MECHANICAL ENG TECH	11-01-2023	03-10-2001	21.28	Aug-23	REFUSED	FIRST
6	VIKASH RAJ	MALE	UNIVERSITY OF LETHBRIDGE	ALBERTA	BACHELOR OF SCIENCE	11-01-2023	27-09-2002	20.29	Sep-23	REFUSED	FIRST
7	ADITYA BHARDWAJ	MALE	UNIVERSITY OF LETHBRIDGE	ALBERTA	BACHELOR OF SCIENCE	12-01-2023	26-01-2004	18.96	Sep-23	APPROVED	FIRST
8	JASJEET SINGH	MALE	CAPILANO UNIVERSITY	BRITISH COLUMBIA	ASSOCIATES OF SCIENCE	16-01-2023	02-02-2005	17.95	May-23	APPROVED	FIRST
9	NAVJEET KAUR	FEMALE	CAPILANO UNIVERSITY	BRITISH COLUMBIA	ASSOCIATES OF SCIENCE	16-01-2023	06-03-2004	18.86	May-23	APPROVED	FIRST
10	HIMANSHI	FEMALE	UNIVERSITY OF LETHBRIDGE	ALBERTA	BACHELOR OF SCIENCE	17-01-2023	29-10-2002	20.22	Sep-23	APPROVED	FIRST
11	AMRINDER SINGH	MALE	UNIVERSITY OF LETHBRIDGE	ALBERTA	BACHELOR OF ARTS	17-01-2023	23-09-2003	19.32	Sep-23	APPROVED	FIRST

Figure 3.1. Student VISA Data (Sample)

3.2 Tourist VISA Metadata:

There are 10 columns in this Data set, First column is *Serial No.*, Second column is *Candidate Name*, third column is *Gender* (Male/Female), fourth column is *Country* for which they want VISA, fifth column is *Purpose of Visit* in the country, sixth column is *Date of Lodgment* in the format DD-MM-YYYY, seventh column is *Date of Birth* in the format DD-MM-YYYY, eighth column is *Age* in the format XX.XX because of the precision, ninth column is *File Status* (APPROVED/REJECTED) and the tenth column is *Attempt* providing the information about the VISA attempt through the company.

I have shared a portion of the data as an image:

Serial Number	Candidate Name	Gender	Country	Purpose of Visit	Date of Birth	Date of Lodgement	Age	File Status	Attempt
1	AMRIK SINGH	MALE	CANADA	TO MEET SON & DAUGHTER IN LAW	03-01-1939	02-01-2023	84.00	APPROVED	FIRST
2	KULDIP KAUR	FEMALE	CANADA	TO MEET SON & DAUGHTER IN LAW	18-02-1942	02-01-2023	80.87	APPROVED	FIRST
3	TARANJIT KAUR	FEMALE	CANADA	TO MEET SISTER	04-10-1990	02-01-2023	32.25	APPROVED	FIRST
4	SATINDER KAUR	FEMALE	AUSTRALIA	TO ATTEND BROTHER MARRIAGE	01-11-1992	02-01-2023	30.17	APPROVED	FIRST
5	DEEPAK	MALE	AUSTRALIA	TO MEET BROTHER	10-08-1991	03-01-2023	31.40	APPROVED	FIRST
6	TARANBIR SINGH	MALE	CANADA	SOCIAL VISIT	02-10-1979	04-01-2023	43.26	REFUSED	FIRST
7	SURJIT SINGH	MALE	AUSTRALIA	TO CELEBRATE SON BIRTHDAY	24-09-1956	04-01-2023	66.28	APPROVED	FIRST
8	TARANJIT KAUR SAINI	FEMALE	AUSTRALIA	TO CELEBRATE GRANDCHILD BIRTHDAY	08-12-1961	04-01-2023	61.08	APPROVED	FIRST
9	SARBJIT KAUR	FEMALE	CANADA	TO MEET SON & DAUGHTER IN LAW	15-10-1975	13-01-2023	47.25	REFUSED	FIRST
10	JAGVIR SINGH	MALE	AUSTRALIA	TO ATTEND BROTHER CONVOCATION	14-02-1998	13-01-2023	24.91	REFUSED	FIRST
11	VIAN THAKUR	MALE	CANADA	VISITOR	21-03-2021	16-01-2023	1.82	APPROVED	FIRST
12	KULWINDER KAUR DHARNI	FEMALE	AUSTRALIA	TO ATTEND BROTHER MARRIAGE	20-01-1988	16-01-2023	34.99	APPROVED	FIRST

Figure 3.2. Tourist VISA Data (Sample)

4. Descriptive Statistics:

Now I will provide the descriptive statistics based on the data I collected.

4.1 Student VISA

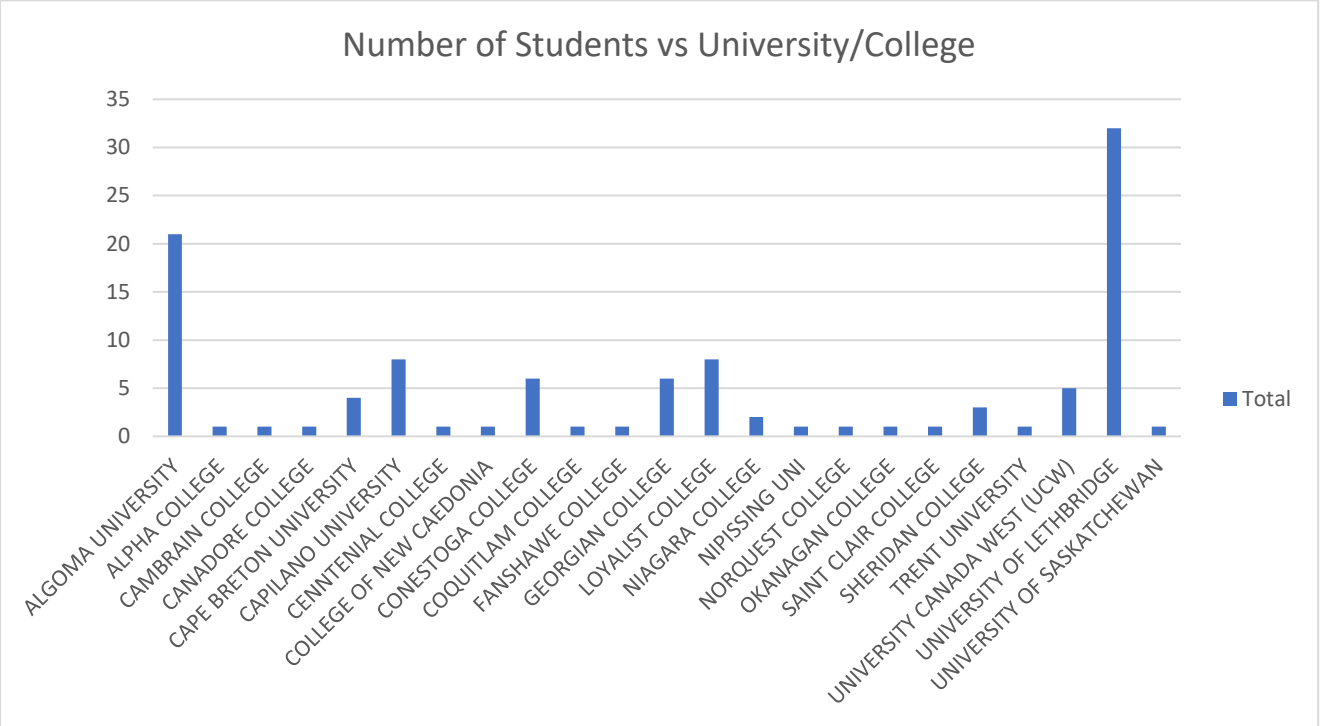

Figure 4.1. Number of Students vs University/College

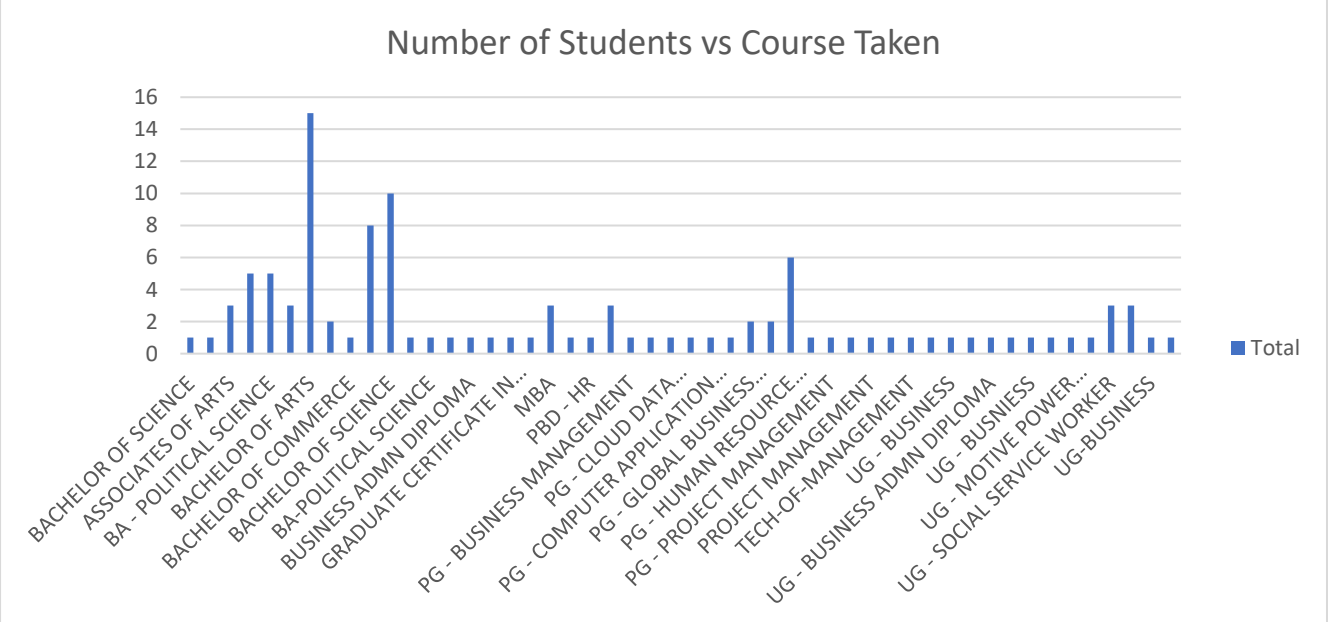

Figure 4.2. Number of Students vs Course Taken

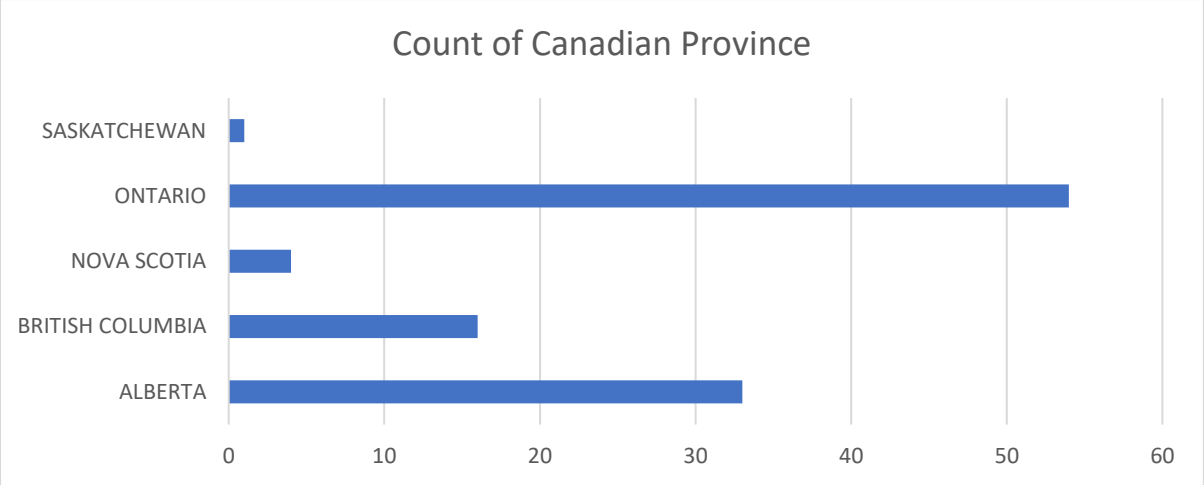

Figure 4.3. Canadian Province vs Number of Students

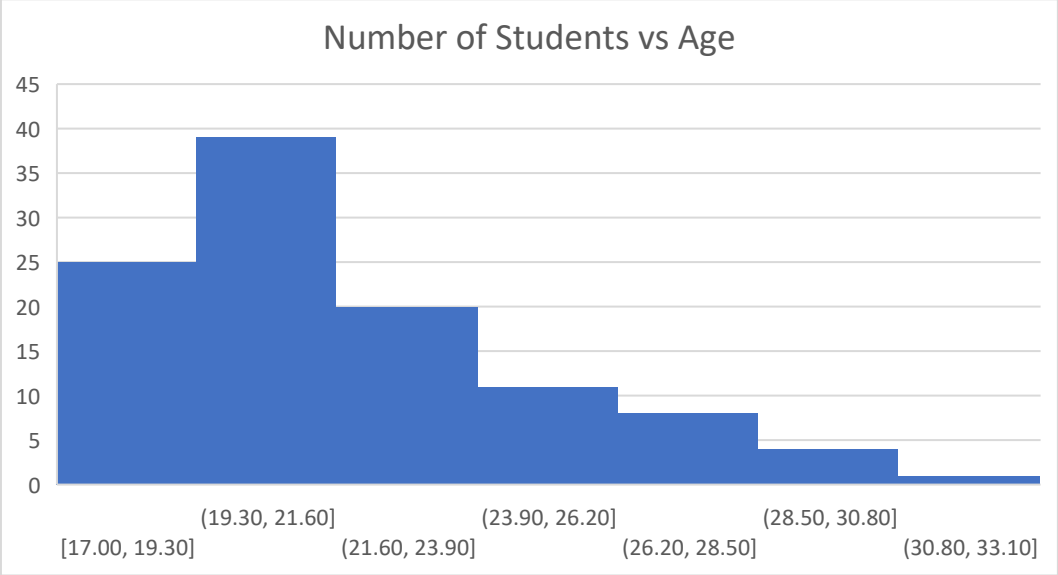

Figure 4.4. Number of Students vs Age

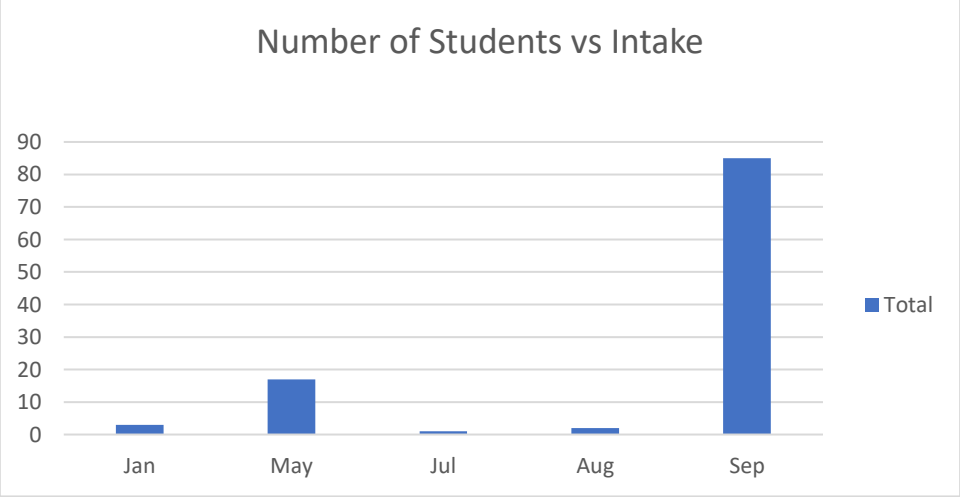

Figure 4.5. Number of Students vs Intake

4.2 Tourist VISA

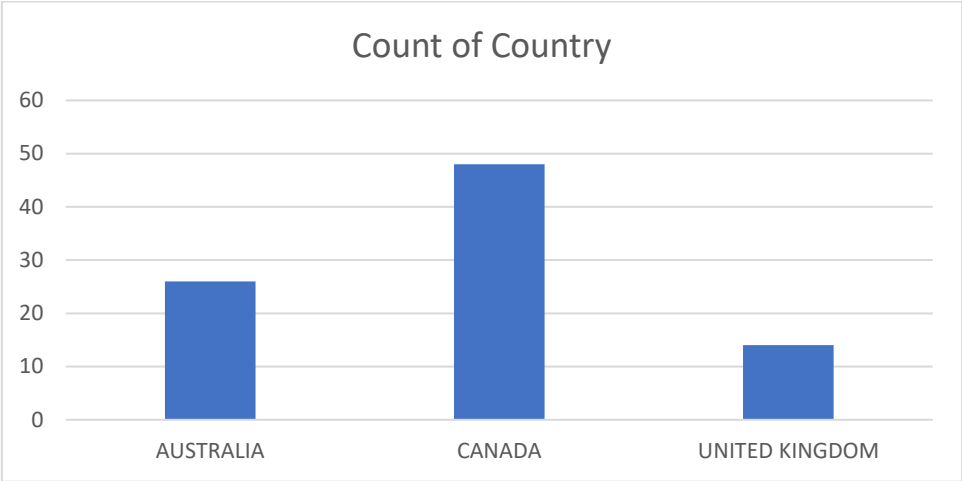

Figure 4.6. Number of People vs Country

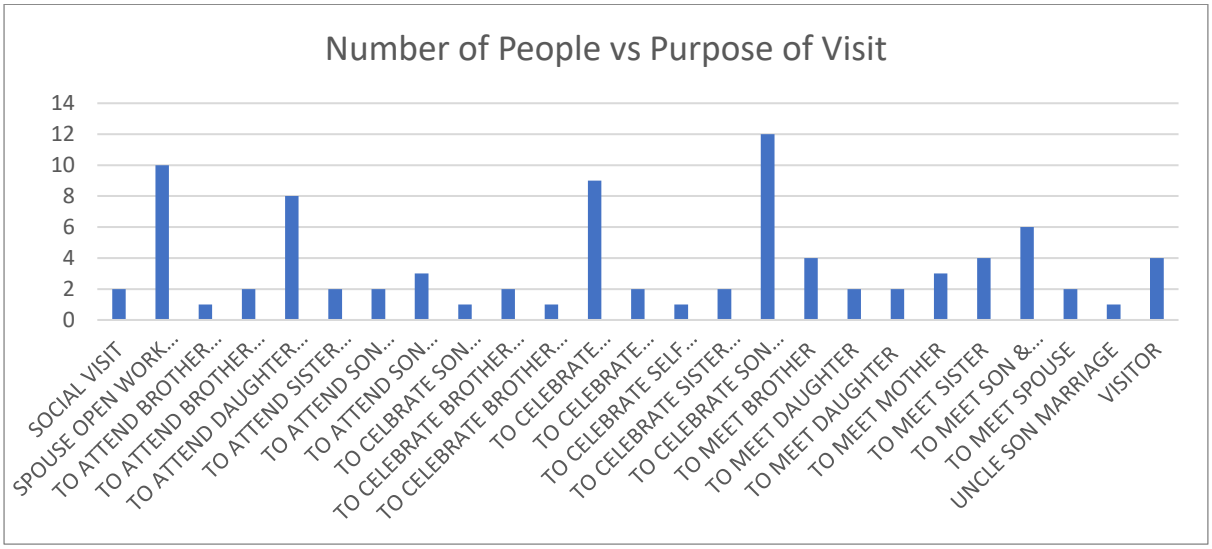

Figure 4.7. Number of People vs Purpose of Visit

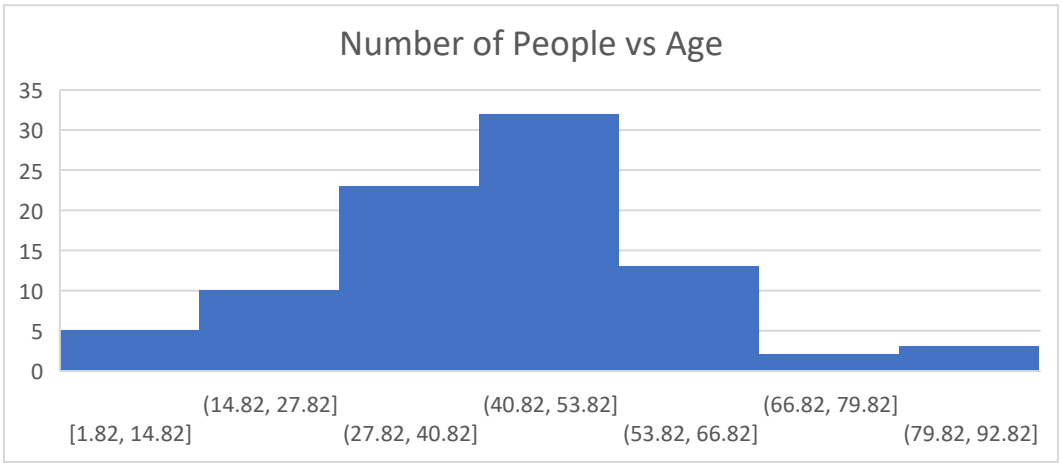

Figure 4.8. Number of Students vs Intake

5. Detailed Explanation of Analysis Process/Method:

First I collected 8 months of data of year 2023, the data was not clean, it consists of some unnecessary fields like *Application No.* and *Biometric Status*, this information was not needed by me for my analysis so I removed them. The data also consisted a sensitive information of *Passport No.*, the company directors asked me to not share this information so I removed it, but one problem that arise with it was that now I was not able to distinguish those who attempted third or second time for the VISA through the company so to mitigate the problem I created an extra column named *Attempt* which provided the information about the attempt. Also, I had added the *Date of Birth* and *Age* fields meticulously through the Passport PDFs of consumers.

After the data was cleaned, I first analyzed the Student VISA data, the column *University/College Name* have discrete values to I decided to plot a Bar Chart to see which University/College is chosen by the greatest number of students. The column *Course Taken* also had discrete values and Bar Chart was also plotted to see the most popular courses among the students. Both these fields can't be plotted directly through the insert chart function because of the number of distinct entries in the column, so for the plotting I used the *PivotChart & PivotTable* function by specifying the *Values* as *Count of Serial No.* and the field which to be analyzed was specified was *Axis.*
Then the *Canadian Province* field was analyzed to find the which provinces of Canada are more in demand among the students, for this a Horizontal Bar graph was plotted by the use of Pivot Chart and Table function.
After that the most important column *Age* was analyzed, this column have continuous values so I plotted a histogram to see which age group go abroad for education the most and which the least.
And in last I analyzed the *Intake* field, by plotting the bar graph through *PivotChart & PivotTable* function to find the which *Intake* is common among students.

Secondly I analyzed the Tourist VISA data, the country column is having discrete values so a bar chart was plotted with the *PivotChart & PivotTable* function to see which countries among *Canada, Australia* and *United Kingdom* have who much contribution in the Tourist VISA applications.
Then the *Purpose of Visit* Column was analyzed with the use of Pivot Table because it consisted of many distinct discrete values, so a bar graph was plotted to see frequent reasons of people for visiting other countries as tourists.
Lastly same as in Student VISA data, *Age* was analyzed with the use of a histogram because of the continuous nature of the data entries of this column. The graph provides information about the age group of people who visit other countries as tourists.

One important information about the analysis is that I used the *Attempt* field as filter to only include the rows which have the *Attempt* field as FIRST attempt because if I included all the rows then there would have been repetition of some students or people.

There were many numerical figures that I calculated using Pivot Table, those figures are helpful in visualizing the trends, I have included them in *Results and Findings* for better understanding.

More advance analysis will be done which will be included in the Final-Term Submission, that analysis would be more oriented towards numerical aspects of the data.

6. Results and Findings:

6.1 Student VISA

- *University of Lethbridge* (about 29.63%) followed by *Algoma University* (about 19.44%) are the two universities of top choices among the students.
- *Bachelor of Arts* is the most popular course among the students which is about 13.89% followed by *Bachelor of Science* about 9.26%.
- Most of the students chooses University/College which is located in *Ontario* (about 50%) followed by *Alberta* (about 30.56%).
- The age group which goes out for Overseas Education the most is 19.20 to 21.60 years of age; this age group contributes about 36.11% of the students.
- September Intake is the most common intake about 78.70% of the students take it.

6.2 Tourist VISA

- Most of the people go to *Canada* through the tourist VISA, about 54.54% of the people.
- The most common purposes for the tourist visits are *TO CELEBRATE SON BIRTHDAY, SPOUSE OPEN WORK PERMIT, TO CELEBRATE DAUGHTER BRITHDAY* and *TO ATTEND DAUGHTER CONVOCATION*, these purposes contribute about 44.32% of the total purposes.
- The most common age group who apply for tourist VISA is 40.82-53.82 years, which contribute to about 36.36% of the people.

In the Final-Term Submission, I would discuss about the numerical aspect of the data and find the solutions to the problem statements by the interpretation of the results that we will get by some more analysis of the data.