

**Statistical Analysis of**

**Programme wise**

**Foreign Students Enrollment in Year 2011-2012**

R Project Report in

Mathematical Foundation of Computer Science

**Submitted By: Submitted To :**

Akanksha Goel (02) Prof. Vasudha Bhatnagar

Ashish (11) Deepti (18)

**Department of Computer Science**

# University of Delhi

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **INDEX** | **Page** |
| 1. | **Dataset Description** | 2 |
| 2. | **Exploratory Data Analysis (EDA) and**  **Data Preprocessing** | 3-7 |
| 3. | **Problem Set and Motivation** | 8 |
| 4. | **Data Analysis and Visualization** | 9-16 |
| 5. | **Conclusion** | 17 |
| 6. | **References** | 18 |

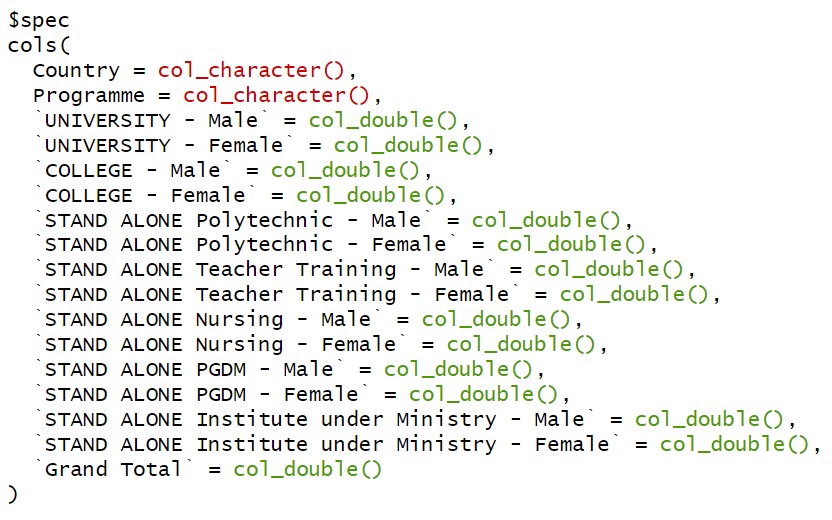
# Dataset Description

The analysis is done on **Foreign students Enrollment – 2011-2012** . The data is collected under the *Study Programme & Discipline according to AISCHE Survey from data.gov.in .* The data shows the number of Foreign fe/male students enrolled in Indian Universities and Colleges in different Programme . Dataset covers entire Higher Education Institutions in the country. Dataset is categorised in 3 broad Categories; University, College and Stand‐Alone Institutions.

## Data Source : [data.gov](https://data.gov.in/resource/programme-wise-foreign-students-enrolment-2011-2012)

**Dimensions :** Dataset consists of **1905** tuples and **17** attributes.

**Attributes specification :**



**Exploratory Data Analysis(EDA)**

**Statistical Summary of each column -** using R summary function, each column’s measure of location and centrality is computed.

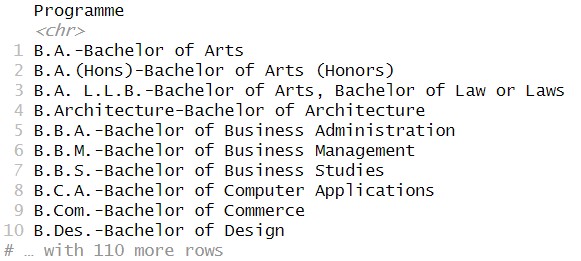
**Attributes:**

CATEGORICAL

1. "COUNTRY" :- There are 155 unique Countries.

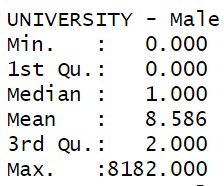
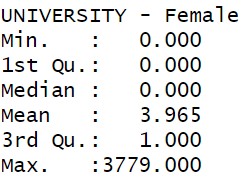


1. "PROGRAMME" :- There are 120 unique Programmes.

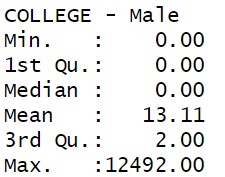
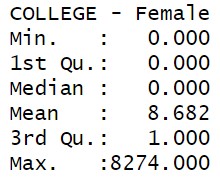


### NUMERICAL

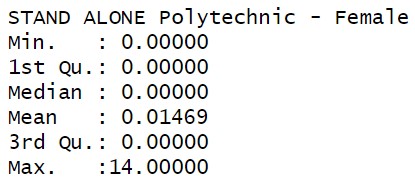
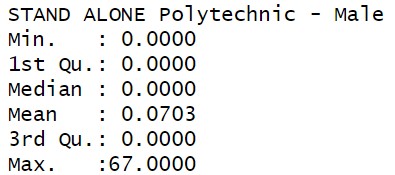
[3]"UNIVERSITY - Male" [4]"UNIVERSITY - Female"

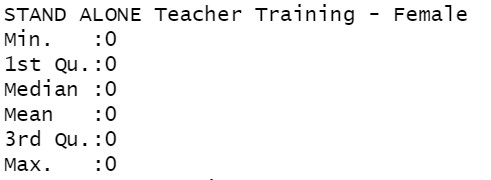
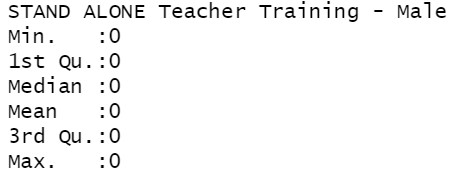
[5]"COLLEGE - Male" [6]"COLLEGE - Female"

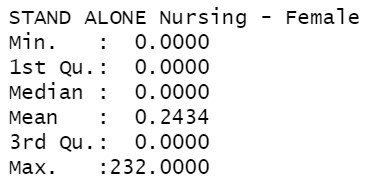
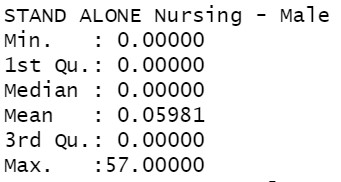
[7]"STAND ALONE Polytechnic - Male" [8]"STAND ALONE Polytechnic - Female"



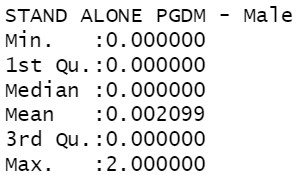
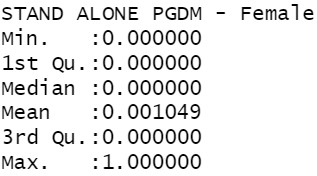
[9]"STAND ALONE Teacher Training-Male" [10]"STAND ALONE Teacher Training-Female



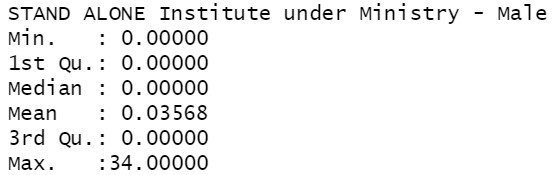
[11]"STAND ALONE Nursing - Male" [12]"STAND ALONE Nursing - Female"



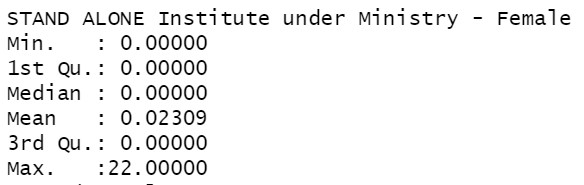
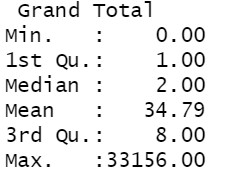
[13]"STAND ALONE PGDM - Male" [14]"STAND ALONE PGDM - Female"

[15]"STAND ALONE Institute under Ministry -Male"



[16]"STAND ALONE Institute under Ministry - Female" [17]"Grand Total"

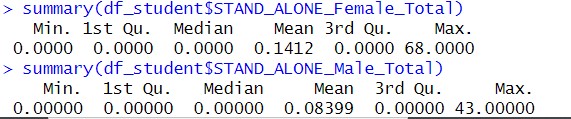
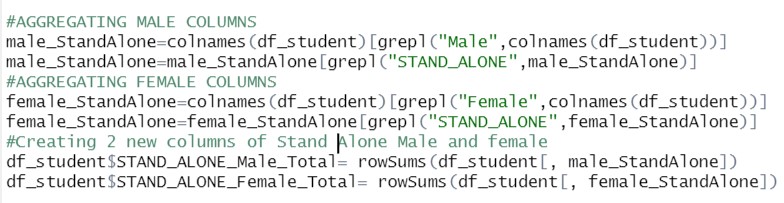
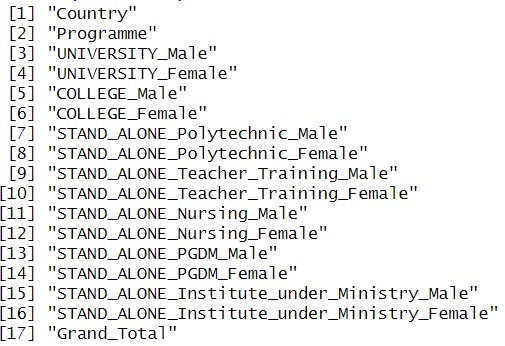
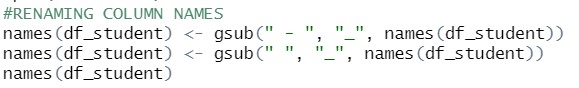
 

**NOTE:**

**Preprocessing need to be done Before Visualizations :**

* Removing the last redundant tuple.
* Renaming Column Names – Extra spaces should be removed.
* Stand Alone Columns have a lot of null values so these columns need to be integrated.

# Data Preprocessing



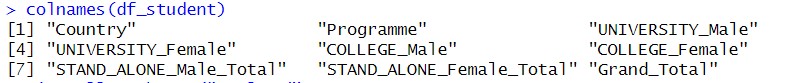
# Problem Set and Motivation

**Problem statement:**- The enrollment of foreign students in universities and colleges is decreasing, posing a challenge for institutions that rely on international students for revenue and diversity. Additionally, changes in immigration policies and visa regulations in some countries have made it harder for foreign students to obtain the necessary documents to study abroad. This decrease in foreign student enrollment threatens the financial stability of many institutions and limits the cultural exchange opportunities for domestic students.

Analyzing the trends and factors affecting the enrollment of foreign students in universities/colleges for the academic year 2011-12 in different programmes in different countries. The study aims to identify the most populated foreign country ; most studied Courses in University/Colleges/Stand Alone .

The research will also examine the distribution of students among these countries and within different programmes. The findings of the study can provide insights into selecting and improving the recruitment and retention of foreign students and enhancing the internationalization of the higher education system.

**DATA SET:**



**DIMENSIONS : *1905*** tuples and ***9*** columns

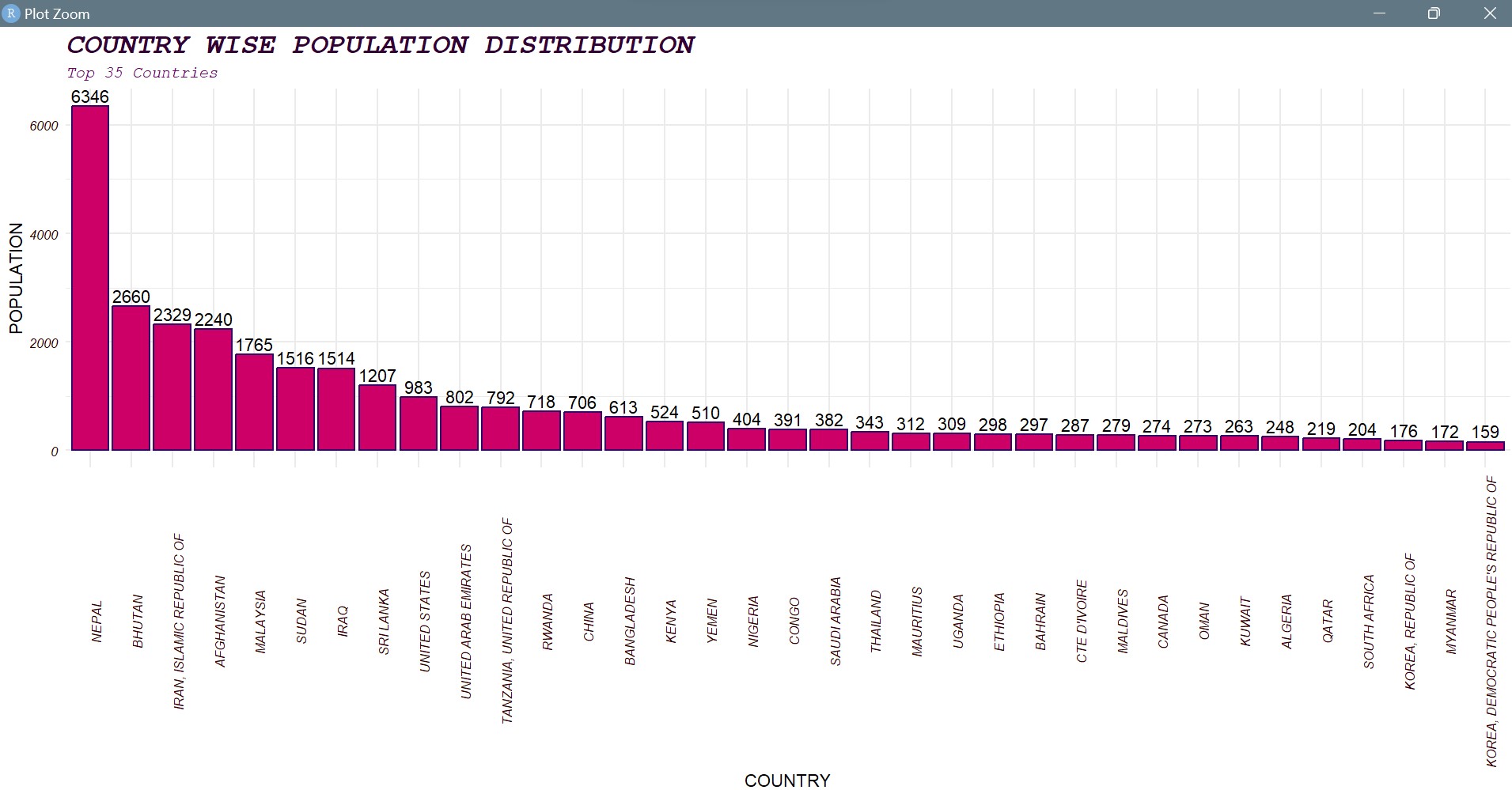
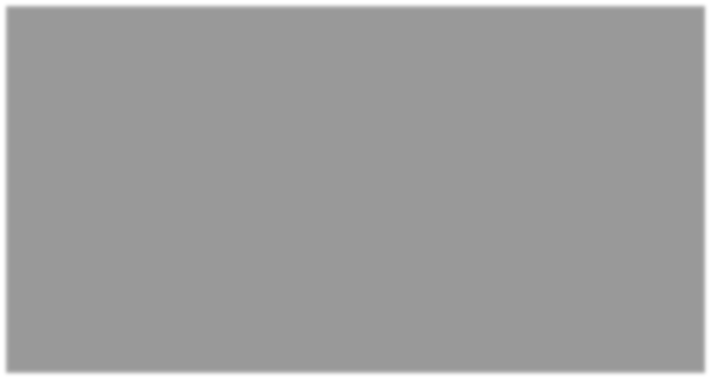
# Data Analysis and Visualization

## 1. COUNTRY WISE :-

**Plot**

**:**

**-**



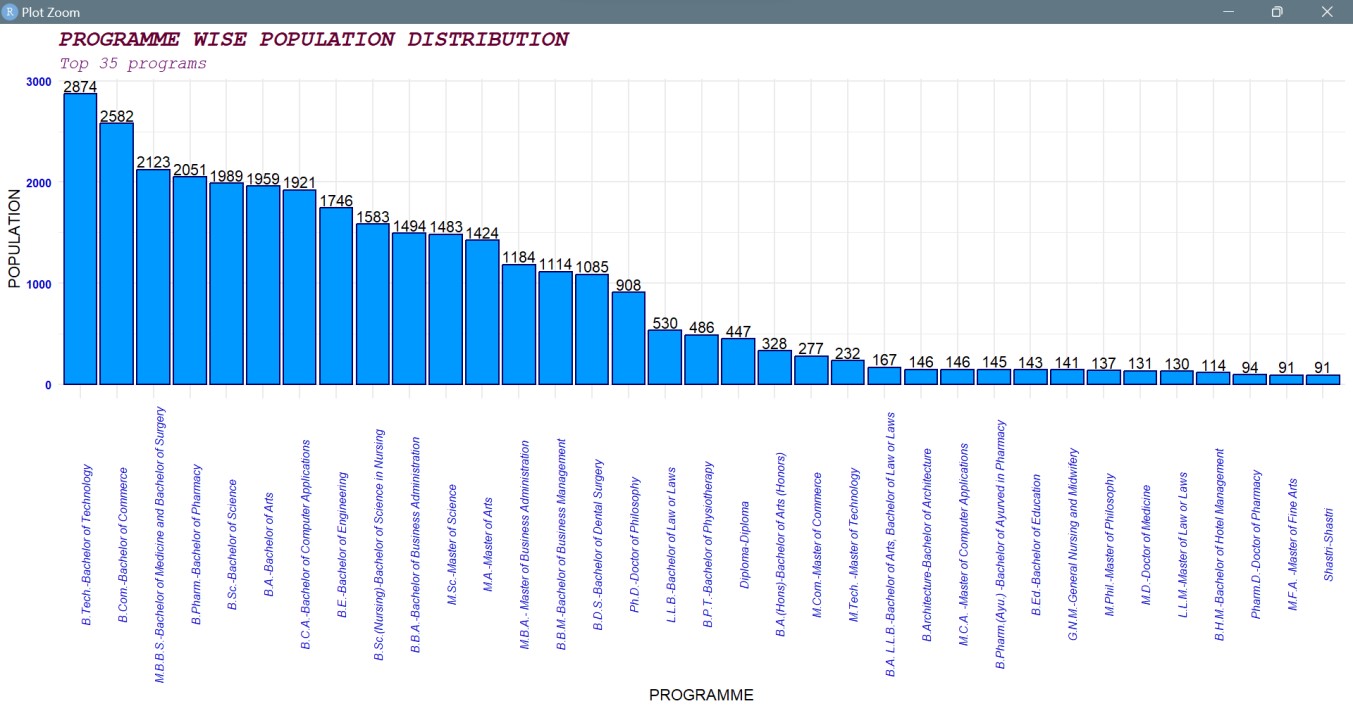
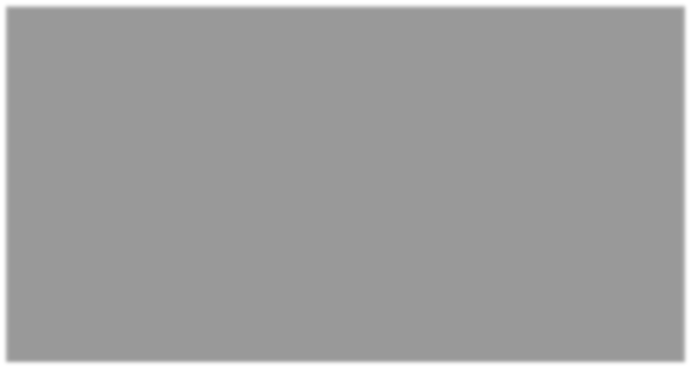
## Observation :-

• Visualized Top 35 countries , and observed Nepal had highest population of 6346 followed by Bhutan with 2260 … and least populated country is Korea, Democratic People’s Republic Of with 159 population.

## 2. PROGRAMME WISE :-

**Plot :**

**-**



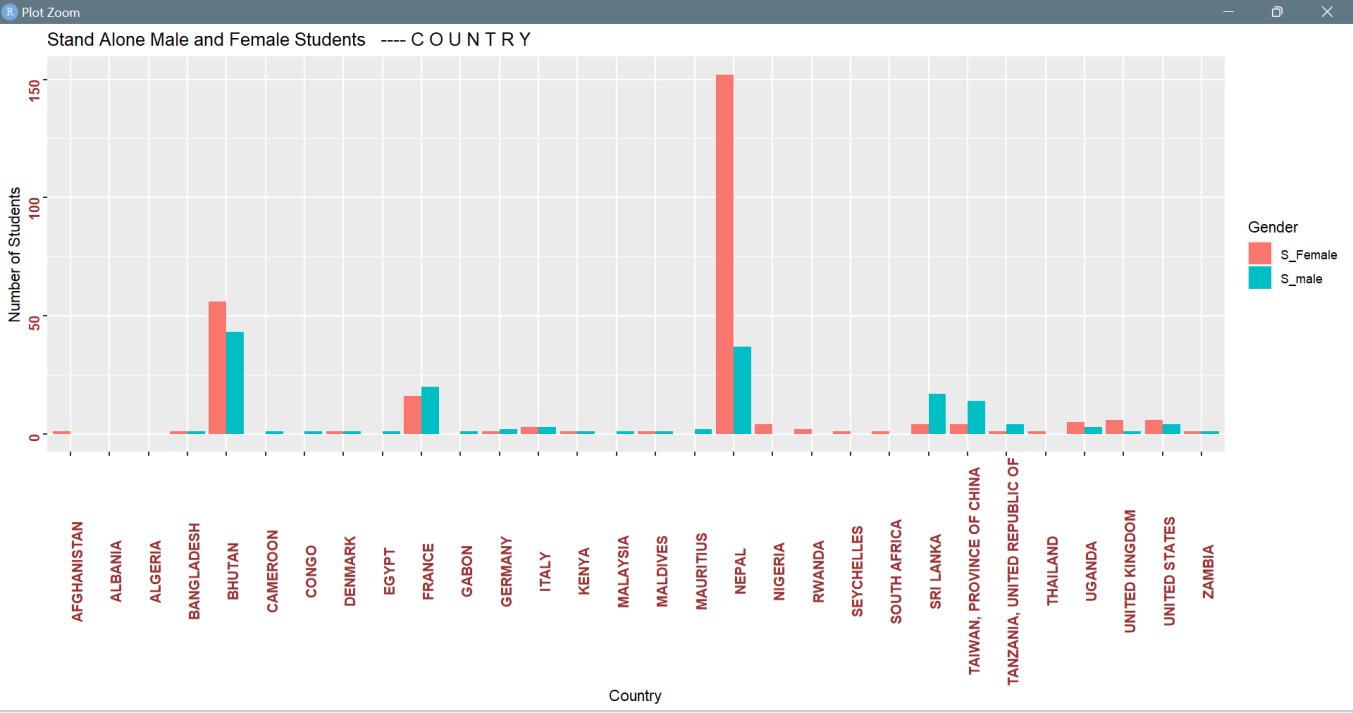
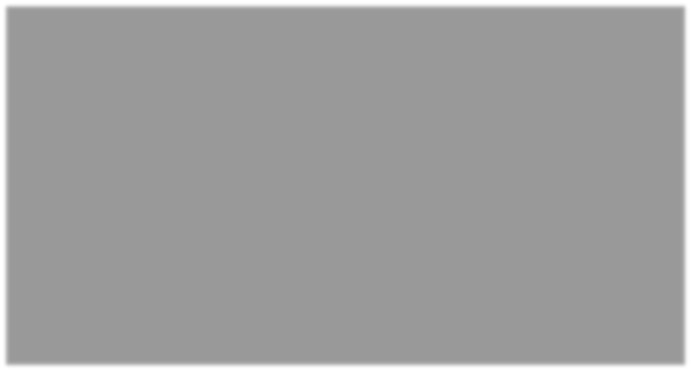
## Observation :-

• Visualized Top 35 programmes , and observed B.Tech.-Bachelor of Technology had highest population of 2874 followed by B.Com.Bachelor of Commerce with 2260 … and least studied programme is Shastri-Shastri and M.F.Sc. -Master of Fishery Science with 91 students .

## 3. COUNTRY WISE (STAND ALONE ) :-

**Plot:**

**-**



## Observation :-

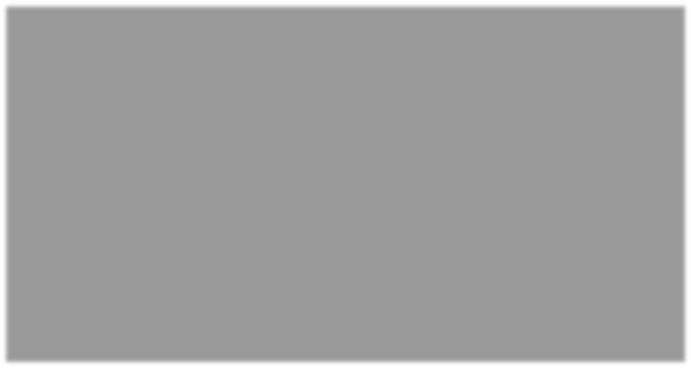
• Visualized distribution of students among different countries and observed that Nepal had more than 150 Female students (maximum), and Bhutan had more than 50 Male students. Further Afghanistan,

Nigeria, Rwanda, Seychelles , South Africa had only Female students , Also Cameroon, Congo, Egypt, Gabon, Mauritius had only Male students.

## 4. PROGRAMME WISE (STAND ALONE ) :-

**Plot:**

**-**



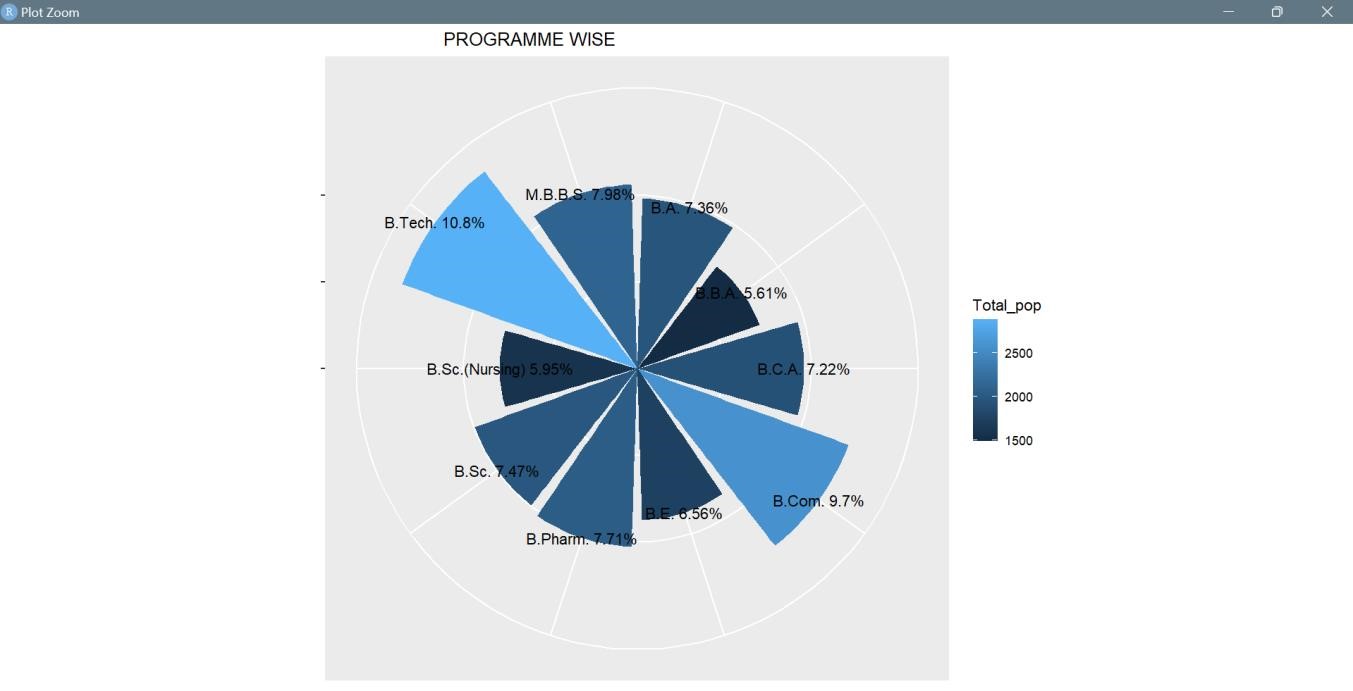
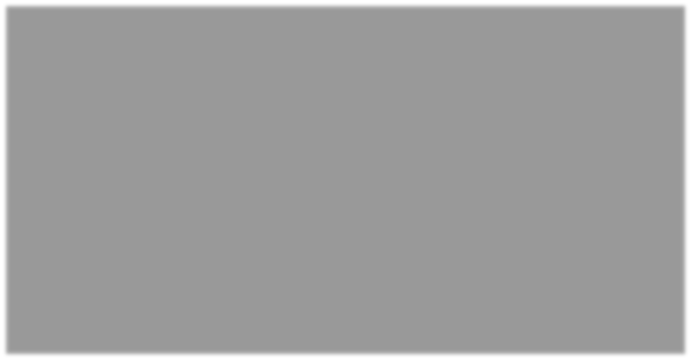
### Observation :-

* Visualized distribution of students among different Programmes and observed that *Diploma-Diploma* had more than 120 Male students (maximum), and G.N.M.-General Nursing and Midwifery had more than 120 Female students(maximum). Further B.Sc(Hons), B.Sc (post Basic)-B.Sc (post Basic) had only Female students and Also, P.G.P.PostGraduate Diploma in Management had only Male students.
* Only 10 programmes among 120 programmes are studied.

### 5. PROGRAMME WISE

**Plot:**

**-**



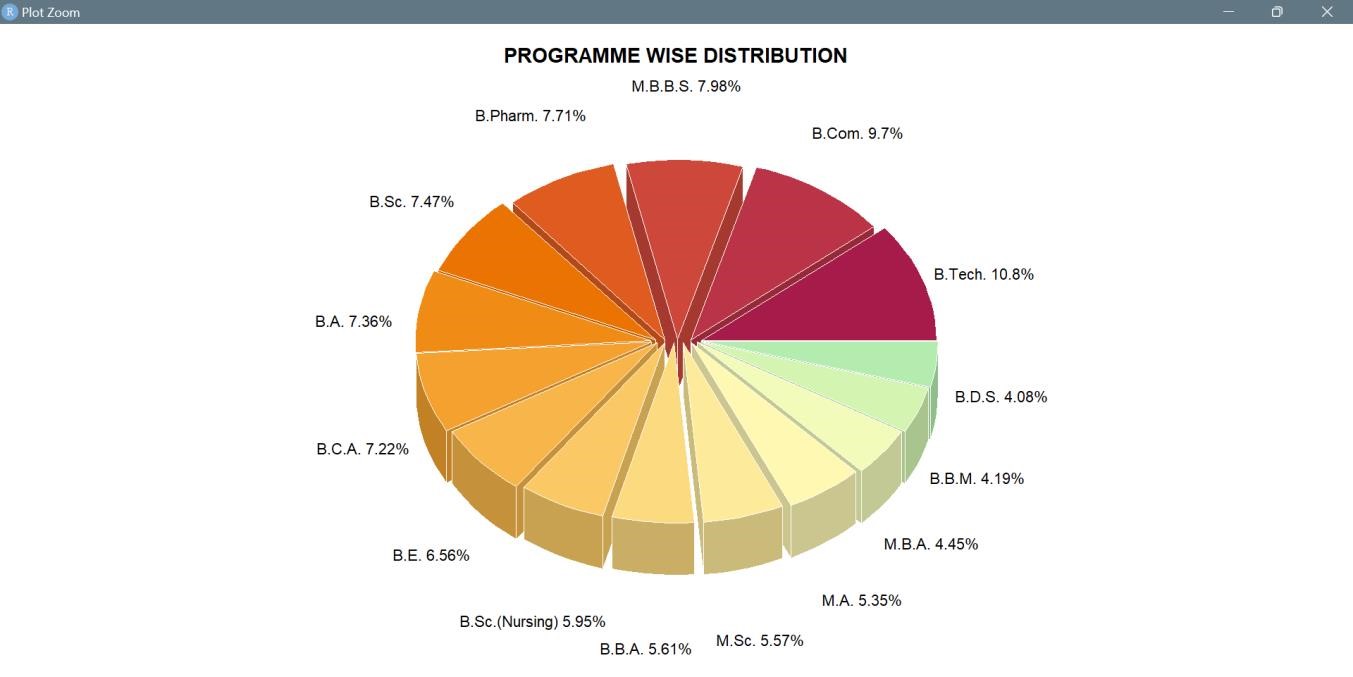
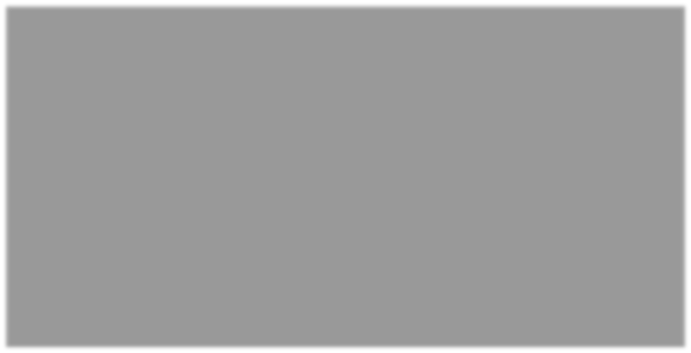
### Observation :-

• Visualized distribution of students among different Programmes and observed that *B. Tech* had been studied by approx. 10.8% students, and *B.B.A.* had been studied by only 5.61% students among top 10 programmes

### 6. PROGRAMME WISE DISTRIBUTION

**Plot:**

**-**



### Observation :-

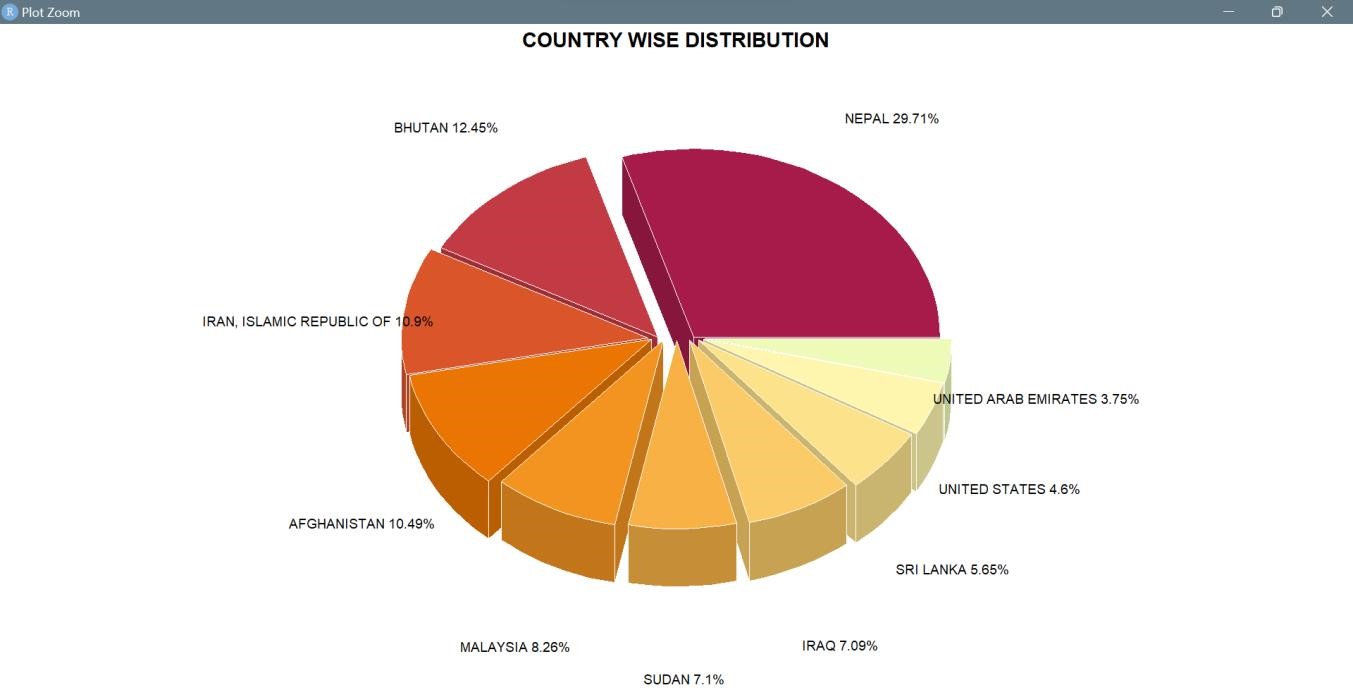
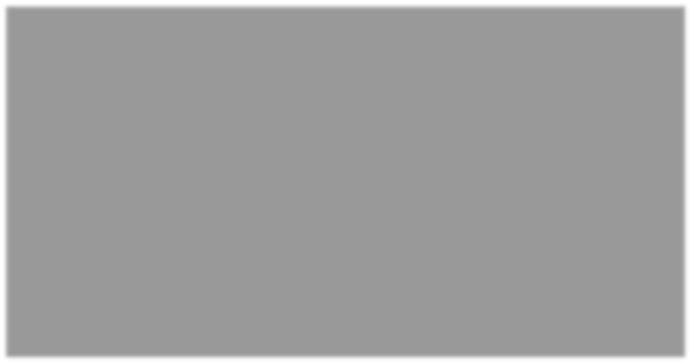
• Visualized distribution of students among different Programmes and observed that *B. Tech* had been studied by approx. 10.8% students, and *B.D.S.* had been studied by only 4.08% students among top 15 programmes

**7. COUNTRY WISE DISTRIBUTION**

### Plot:-

**Observation :**

**-**



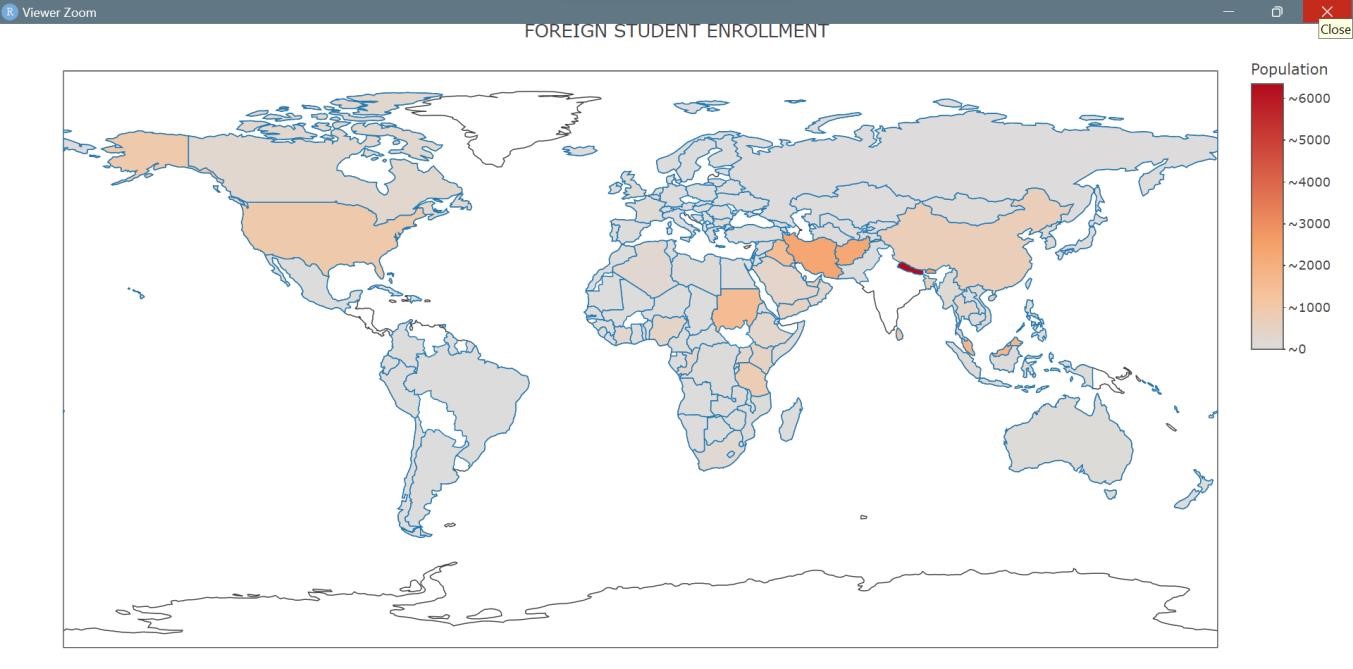
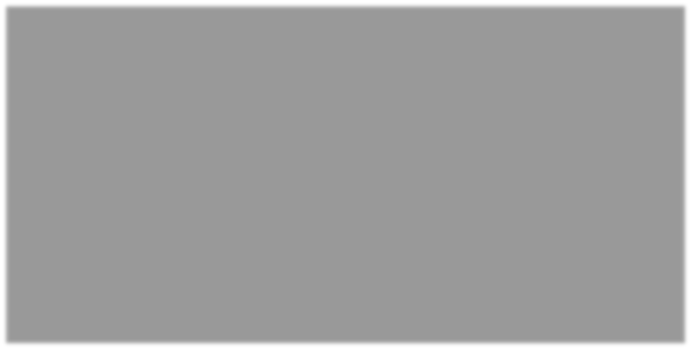
Highest share of foreign students come from the neighboring countries of which Nepal is 29.71% of the total, followed by Bhutan (12.45%), Iran (19.9%), Afghanistan (10.49%), Malaysia (8.28%).

Sudan and Iraq constitutes 7% each and Sri Lanka, United States and UAE, constitutes 5.65%, 4.6% and 3.75% of the foreign students respectively

### 8. CHOROPLETH GRAPH

**Plot:**

**-**



### Observation :-

Highest share of foreign students come from the neighboring countries of which Nepal is 29.71% of the total, followed by Bhutan (12.45%), Iran (19.9%), Afghanistan (10.49%), Malaysia (8.28%).

Sudan and Iraq constitutes 7% each and Sri Lanka, United States and UAE, constitutes 5.65%, 4.6% and 3.75% of the foreign students respectively

# Conclusion

* The total number of foreign students enrolled in higher education is 33,156.
* The foreign students come from 155 different countries from all across the globe. The top 10 countries constitute 64% of the total foreign students enrolled.
* Highest share of foreign students come from the neighboring countries of which Nepal is 29.71% of the total, followed by Bhutan (12.45%), Iran (19.9%), Afghanistan (10.49%), Malaysia (8.28%). Sudan and Iraq constitutes 7% each and Sri Lanka, United States and UAE, constitutes

5.65%, 4.6% and 3.75% of the foreign students respectively

**References**

1. [data.gov.in](http://data.gov.in/)
2. [towardsdatascience.com](https://towardsdatascience.com/)
3. [Exploratory Data Analysis in R](https://www.geeksforgeeks.org/exploratory-data-analysis-in-r-programming/)
4. [Shiny Tutorials - RStudio](https://shiny.rstudio.com/tutorial/written-tutorial/lesson1/)
5. [Data Visualization with R](https://rkabacoff.github.io/datavis/)

➔