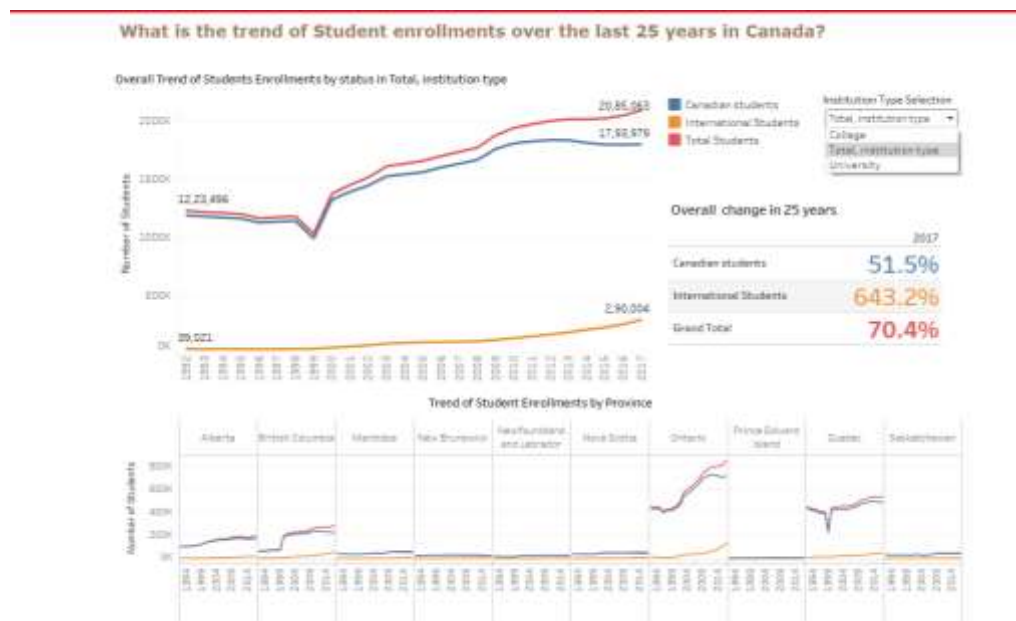## Analysis

In this project, we are analysing the patterns and trends of student enrolments in Canada, from different perspectives.

1. **What is the trend of Student enrolments over the last 25 years in Canada?**
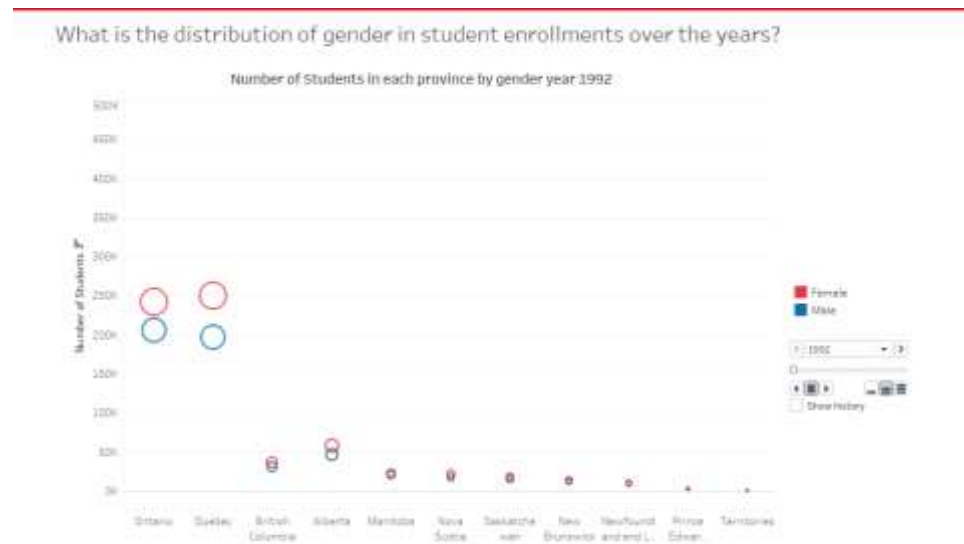


The overall trend of student enrollment shows a very steady increase over the period of 1992 to 2017 in Canada. Initially most of the enrollments were Canadians, while from 2003, there is an influx of international student enrollments. Recent years depict that the Canadian student enrollment has slightly decreased but the considerable increase of international students have maintained the upward trend in overall enrollment. We also noticed that there is a decline during the year 1999, in the Canadian student enrollment. When the details are drilled down to the province level, we see that majority of the enrollments are in Ontario and Quebec provinces followed by moderate enrollments in Alberta and British Columbia in comparison to other provinces. In addition, the decline in the 1999 enrollment seem to be a direct reflection from Quebec and Ontario provinces.

Universities remain as major post-secondary education institution type compared to colleges. Overall, a 70% increase over the period (from 1992 to 2017) in student
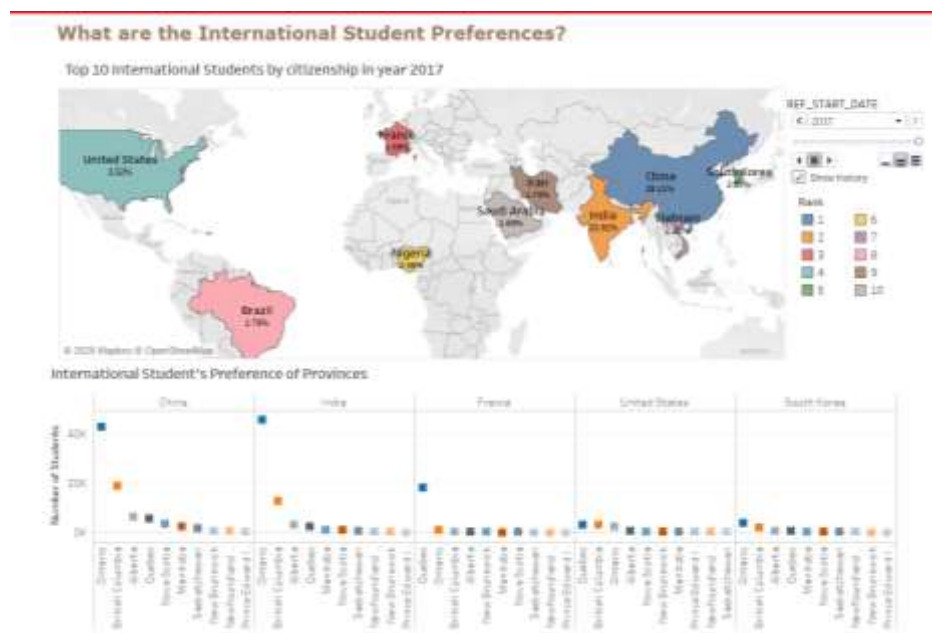
enrollments was noticed, whereas a staggering increase of 643% in International student enrollments.

2. **What is the distribution of gender in student enrollment over the year?**



This is an interactive visualization that reflects the gender distribution in student enrollment over the provinces during the years, 1992 to 2017. It is identified that, in all the provinces, female student enrolments are more than that of male. In the early 1990s, Quebec was the leading province for post-secondary education followed by Ontario but by the beginning of new century, Ontario started leading the position till now. Another noticeable fact is that, the gender Gap has been increasing consistently over the period. Also, there is a very significant growth in British Columbia, overtaking Alberta to become the 3rd province with more post-secondary student enrollment. Other provinces have a much lesser contribution to the Canadian post-secondary education with respect to the number of student enrollments.

3. **What are the International Student trends and location preferences?**



This graph shows the top 10 countries from where students come to study in Canada. It is observed that more students are from East or South East Asia. From 2000 onwards, China is the top country followed by India and France, that sends students to Canada. Also, in recent years, India and China together contributes about 50% of international students.

We were also interested in analyzing if the international students have any particular location preferences. The data and visualization show that majority of the international students goes to Ontario and then British Columbia. An interesting observation from this graph is that, students from United States and France prefer Quebec over Ontario.

Metadata and Clean-up steps

**Before clean up**

Total no: of rows: 1,646,382

Total no: of columns: 18

**Clean up:**

**Step 1: Removed the Columns:**

DGUID

UOM

UOM_ID

SCALAR_FACTOR

SCALAR_ID

VECTOR

COORDINATE

STATUS

SYMBOL

TERMINATED

DECIMALS

**Step 2: Check for duplicates**

No duplicates

**Step 3: Rearrange the columns**

Split the column - Country of citizenship into 2 to separate the country/region code and name

Split the column - REF_DATE split into start and end year. Remove original columns

Filter the data for which the country/region code is between 10000 and 99999 representing the countries alone. Remove the original and code only code.

Rename the 'GEO' column to 'INSTITUTION_LOCATION'

There are 548647 (41%) NULL VALUES. Replace null Values in 'Value' column to 0.

The Value field is renamed to NUMBER OF STUDENTS

All Column names renamed for Readability

Duplicates are checked for, none present.

After Cleaning there is a total of 1336463 Rows and 8 fields.

| Column Name | Column Description | Data Type | Possible Values |
|---|---|---|---|
| REF_START_DATE | The starting year of reference | Number for year | 1992-2017 |
| REF_END_DATE | The ending year of reference | Number for year | 1993-2018 |
| INSTITUTION_LOCATION | Different Provinces in Canada where the institution is located | String | Alberta |
| | | | British Columbia |
| | | | Canada |
| | | | Manitoba |
| | | | New Brunswick |
| | | | Newfoundland and Labrador |
| | | | Nova Scotia |
| | | | Ontario |
| | | | Prince Edward Island |
| | | | Quebec |
| | | | Saskatchewan |
| | | | Territories |
| INSTITUTION_TYPE | Type of the institution in which student is enrolled | String | College |
| | | | University |
| | | | Total, Institution Type |
| STATUS_OF_STUDENT | Describes the status of student in canada | String | Canadian Student |
| | | | International Student |
| | | | Not Reported |
| | | | Total,stat us of student in Canada |
| GENDER | Student's Gender | String | Female |
| | | | Male |
| | | | Gender Unknown |
| | | | Total gender |

| NUMBER_OF_ST UDENTS | Indicates the number of students in each category | Numerical Values | 0,1,2 etc… |
|---|---|---|---|
| COUNTRY_OF_CI TIZENSHIP | Country of Citizenship indicates student's country of citizenships | String | Name of the country |