## Data Statistics analysis for distinct attributes

Most well-liked and accepted value for each attribute for example:

Citizenship: Mostly students belong to which country during 2010-2011?

StudentMajor: Which was most studied course among students in 2010-2011?

Gender: which gender category mostly graduated in 2010-2011 either male or female?

College: Most graduated students belongs to which college from 7000 records in 2010-2011?

Diploma Description: What was the diploma description for most of the students in 2010-2011?

School: Which school students mostly graduated in 2010-2011?

School city: Which school city students mostly did graduation in 2010-2011?

## Correlation data analysis using Correlation matrix

**It would be done** on two attributes**. For example**

**School** and **School\_City** : it would do comparison between different schools of city on basis of graduated students. Like I would tell which school graduated student's ratio is higher and which school of same city has less graduated students than other one.

**Nation** and **Citizenship** here comparison would be done about citizenship of graduated students belongs to same nation.

Similarly I would see for other attributes.

## Histogram Analysis for visualization of data

It would be visualization of data by showing frequency of attribute-values.

## Scatter Plots for visualization of data

It would be visualization of data between any two attributes e.g. Gender and Citizenship, Diploma description and Gender and similarly for any other two-attributes can be compared with respect to graduated students.

## Frequent Itemset Mining using FPGrowth algorithm

It would tell itemset counts for various distinct and combined attributes.

ItemSet 1 if required. It would tell that from 7000 records, 97% of student major is sociology. 42.9% students are under age 30. Here I would state all ratios of distinct attributes.

ItemSet 2: [Diploma description, Gender]:- One result I'm writing here. 74.3% male students have deploma description as Bachelor.

ItemSet 3: [Nation,Age,DateOfBirth],[Age:37-43,Residency, DoB],[School City, Gender, Residency]-One result from last item-set is: 203 Graduated female students among 7000 belongs to Saudi-Arabia AlKhubr

ItemSet4: here would counting of 4 item-sets

ItemSet5, ItemSet6, ItemSet7, ItemSet8, ItemSet9, ItemSet10, ItemSet11