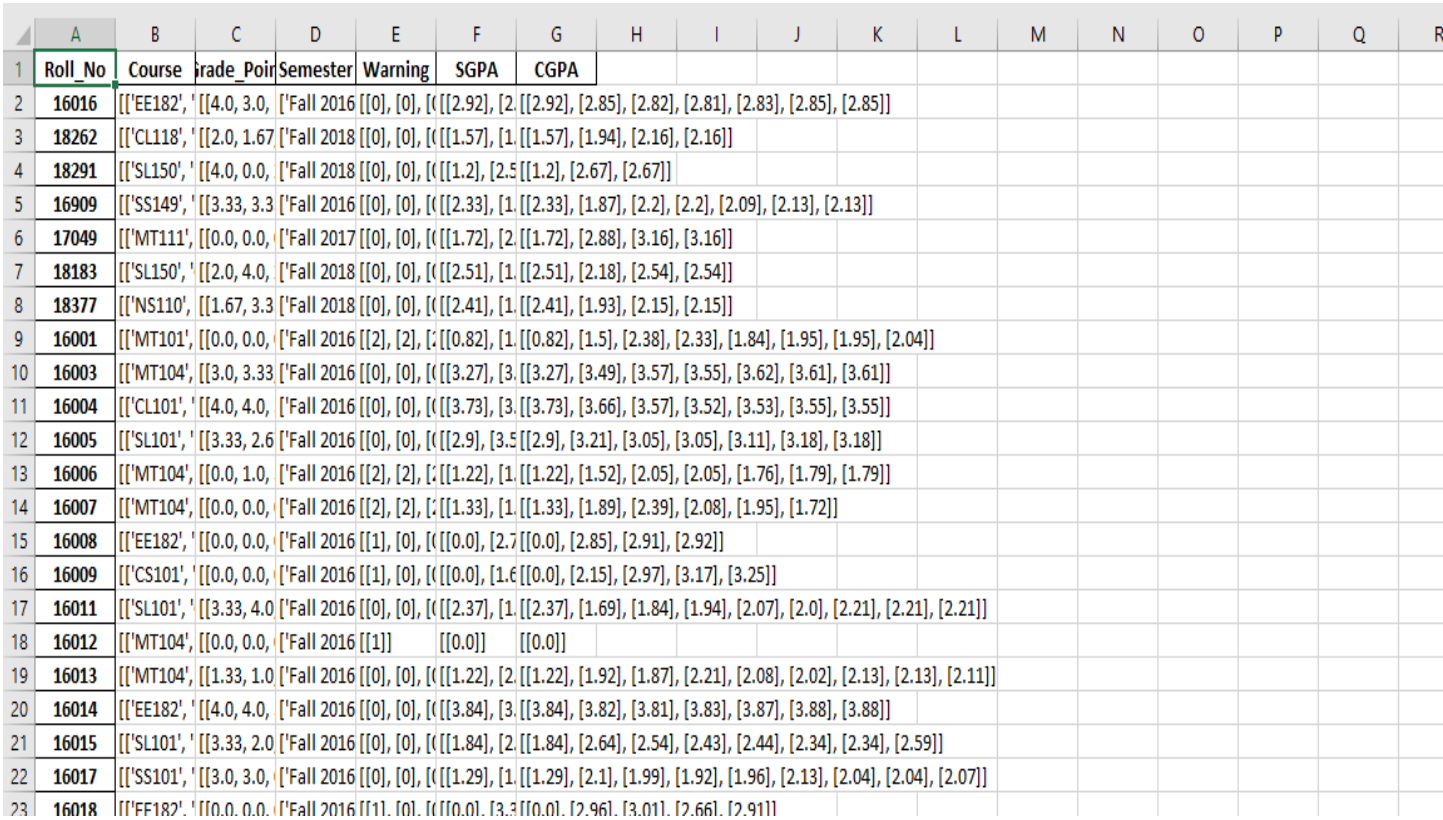
Course Advisory System

Data Loading & Preprocessing:

In data preprocessing, we combined individual entries of course taken to a 2D array of size 8. Each instance of array represents courses taken in a single semester. In this way, we also combined grade points of each course, warning counts, SGPA and CGPA. This data is then saved into excel file.



	A	B	C	D	E	F	G
1	Roll_No	Course	rade_Poir	Semester	Warning	SGPA	CGPA
2	16016	[['EE182', '	[[4.0, 3.0,	['Fall 2016	[[0], [0], [(	[[2.92], [2.	[[2.92], [2.85], [2.82], [2.81], [2.83], [2.85], [2.85]]
3	18262	[['CL118', '	[[2.0, 1.67	['Fall 2018	[[0], [0], [(	[[1.57], [1.	[[1.57], [1.94], [2.16], [2.16]]
4	18291	[['SL150', '	[[4.0, 0.0,	['Fall 2018	[[0], [0], [(	[[1.2], [2.5	[[1.2], [2.67], [2.67]]
5	16909	[['SS149', '	[[3.33, 3.3	['Fall 2016	[[0], [0], [(	[[2.33], [1.	[[2.33], [1.87], [2.2], [2.2], [2.09], [2.13], [2.13]]
6	17049	[['MT111',	[[0.0, 0.0,	['Fall 2017	[[0], [0], [(	[[1.72], [2.	[[1.72], [2.88], [3.16], [3.16]]
7	18183	[['SL150', '	[[2.0, 4.0,	['Fall 2018	[[0], [0], [(	[[2.51], [1.	[[2.51], [2.18], [2.54], [2.54]]
8	18377	[['NS110',	[[1.67, 3.3	['Fall 2018	[[0], [0], [(	[[2.41], [1.	[[2.41], [1.93], [2.15], [2.15]]
9	16001	[['MT101',	[[0.0, 0.0,	['Fall 2016	[[2], [2], [2	[[0.82], [1.	[[0.82], [1.5], [2.38], [2.33], [1.84], [1.95], [1.95], [2.04]]
10	16003	[['MT104',	[[3.0, 3.33	['Fall 2016	[[0], [0], [(	[[3.27], [3.	[[3.27], [3.49], [3.57], [3.55], [3.62], [3.61], [3.61]]
11	16004	[['CL101', '	[[4.0, 4.0,	['Fall 2016	[[0], [0], [(	[[3.73], [3.	[[3.73], [3.66], [3.57], [3.52], [3.53], [3.55], [3.55]]
12	16005	[['SL101', '	[[3.33, 2.6	['Fall 2016	[[0], [0], [(	[[2.9], [3.5	[[2.9], [3.21], [3.05], [3.05], [3.11], [3.18], [3.18]]
13	16006	[['MT104',	[[0.0, 1.0,	['Fall 2016	[[2], [2], [2	[[1.22], [1.	[[1.22], [1.52], [2.05], [2.05], [1.76], [1.79], [1.79]]
14	16007	[['MT104',	[[0.0, 0.0,	['Fall 2016	[[2], [2], [2	[[1.33], [1.	[[1.33], [1.89], [2.39], [2.08], [1.95], [1.72]]
15	16008	[['EE182', '	[[0.0, 0.0,	['Fall 2016	[[1], [0], [(	[[0.0], [2.7	[[0.0], [2.85], [2.91], [2.92]]
16	16009	[['CS101', '	[[0.0, 0.0,	['Fall 2016	[[1], [0], [(	[[0.0], [1.6	[[0.0], [2.15], [2.97], [3.17], [3.25]]
17	16011	[['SL101', '	[[3.33, 4.0	['Fall 2016	[[0], [0], [(	[[2.37], [1.	[[2.37], [1.69], [1.84], [1.94], [2.07], [2.0], [2.21], [2.21], [2.21]]
18	16012	[['MT104',	[[0.0, 0.0,	['Fall 2016	[[1]]	[[0.0]]	[[0.0]]
19	16013	[['MT104',	[[1.33, 1.0	['Fall 2016	[[0], [0], [(	[[1.22], [2.	[[1.22], [1.92], [1.87], [2.21], [2.08], [2.02], [2.13], [2.13], [2.11]]
20	16014	[['EE182', '	[[4.0, 4.0,	['Fall 2016	[[0], [0], [(	[[3.84], [3.	[[3.84], [3.82], [3.81], [3.83], [3.87], [3.88], [3.88]]
21	16015	[['SL101', '	[[3.33, 2.0	['Fall 2016	[[0], [0], [(	[[1.84], [2.	[[1.84], [2.64], [2.54], [2.43], [2.44], [2.34], [2.34], [2.59]]
22	16017	[['SS101', '	[[3.0, 3.0,	['Fall 2016	[[0], [0], [(	[[1.29], [1.	[[1.29], [2.1], [1.99], [1.92], [1.96], [2.13], [2.04], [2.04], [2.07]]
23	16018	[['EE182', '	[[0.0, 0.0,	['Fall 2016	[[1], [0], [(	[[0.0], [3.3	[[0.0], [2.96], [3.01], [2.66], [2.91]]

Data Scaling:

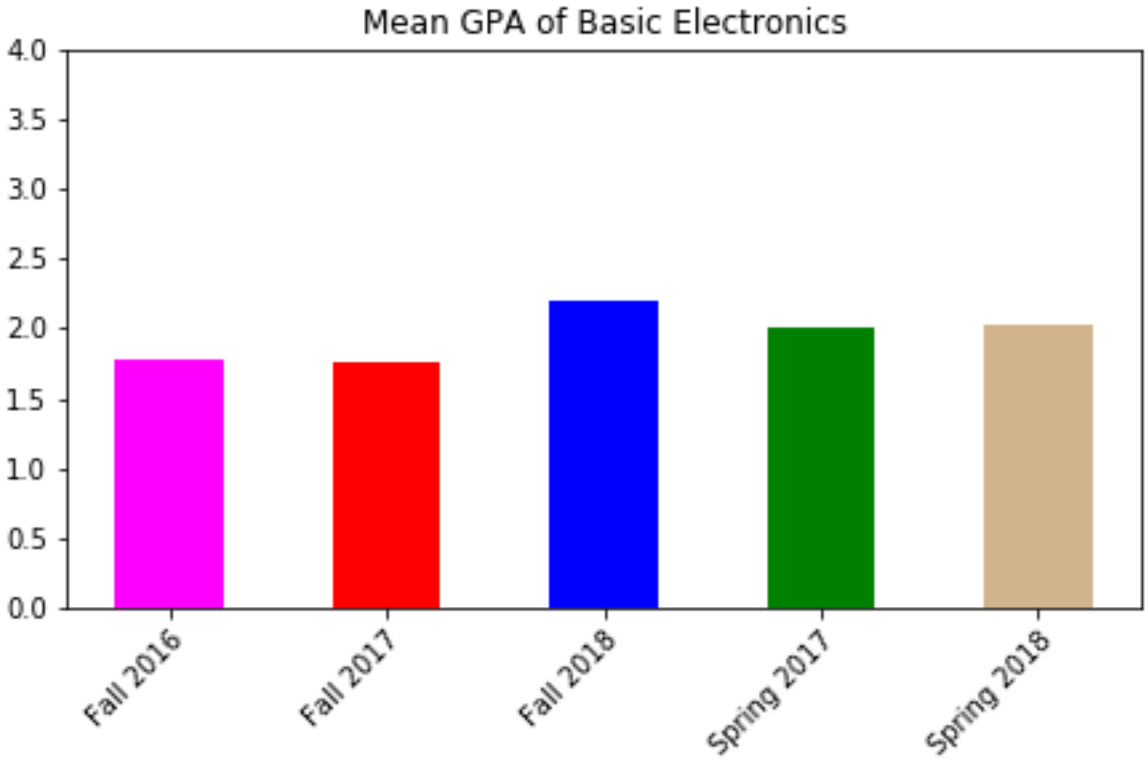
We scaled our data from 0 to 1 in normalize() using the formula taught in class.

$$\frac{C+\text{Min}\ (C)}{Max(C)-Min(C)}$$

Data Exploration/Visualization/Analysis:

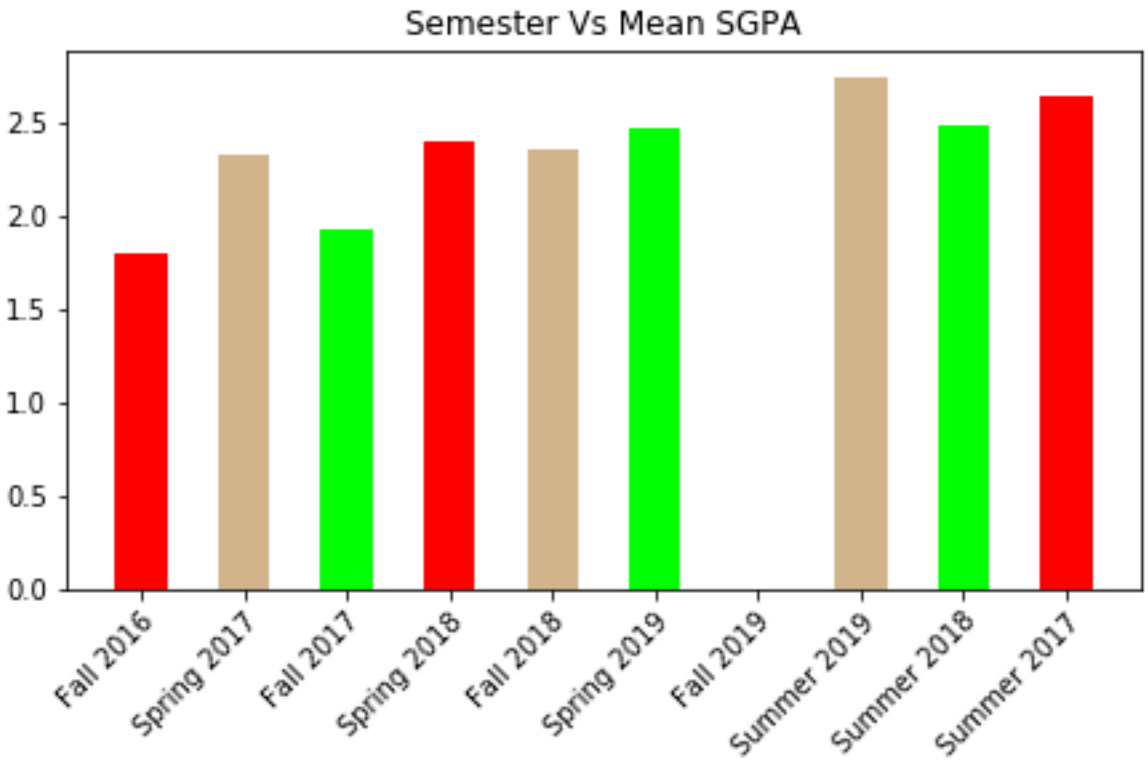
Mean GPA of a Course:

We give name of Course and it gives us a histogram. Each bar of histogram represents mean GPA in a semester. Figure shows the mean GPA of basic electronics. Fall 2017 has minimum mean GPA and Fall 2018 has highest mean GPA.



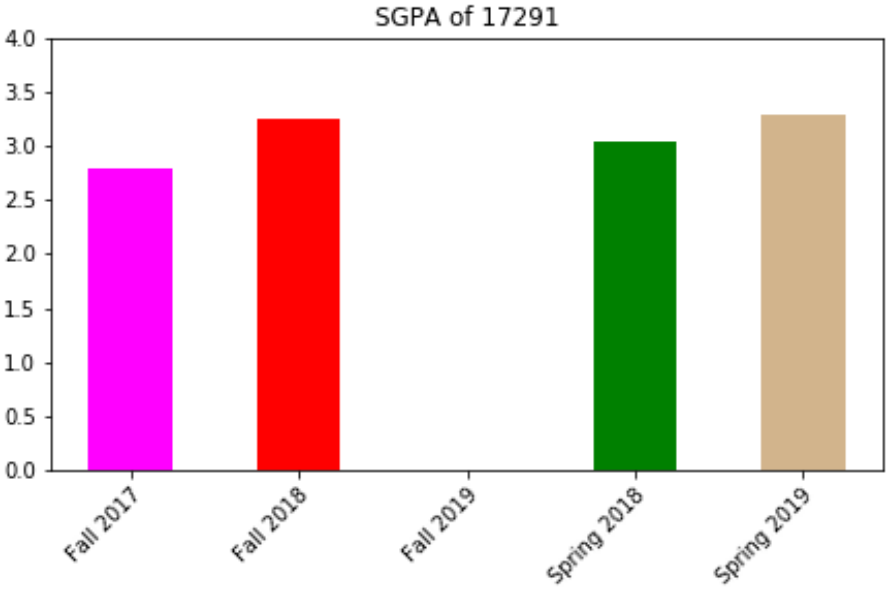
Semester Vs Mean SGPA

It gives histogram of mean SGPA of every semester. It shows that Summer 2019 has highest mean SGPA and Fall 2016 has minimum SGPA.



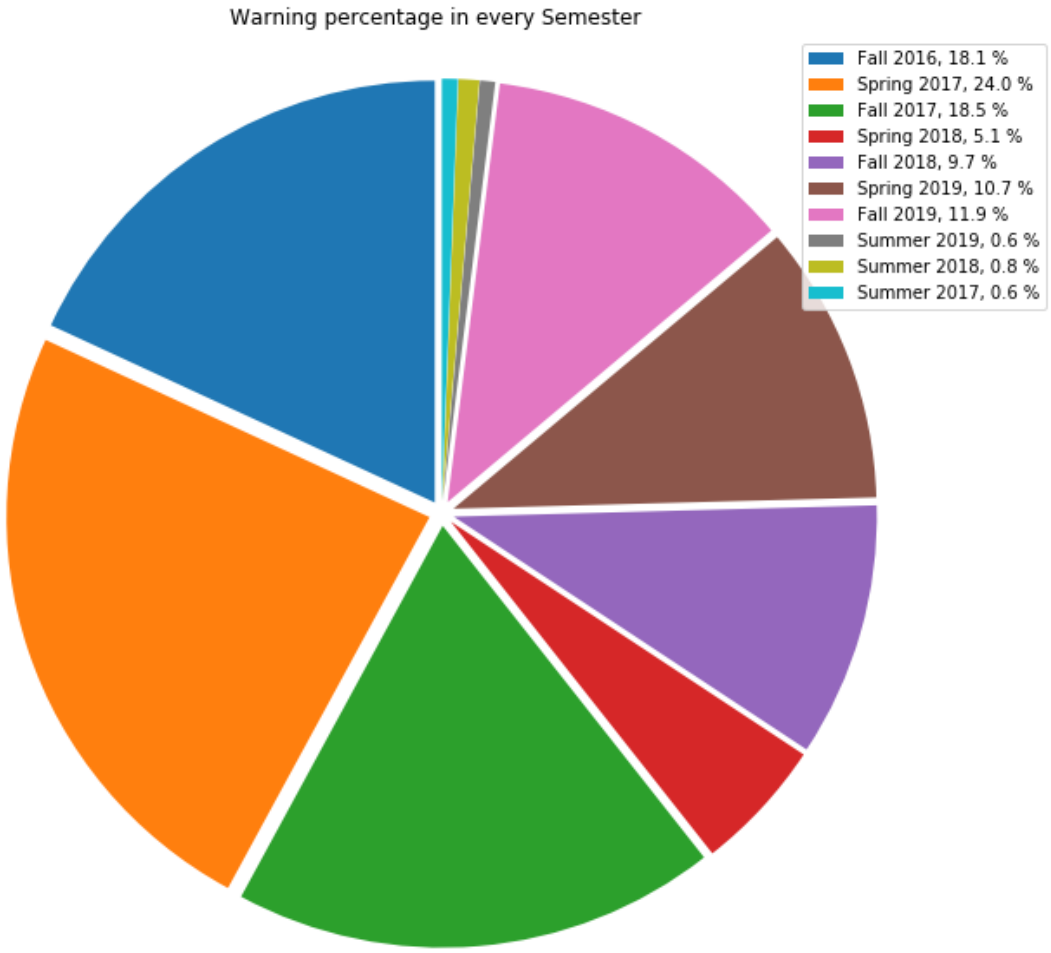
SGPA of a Student:

It gives histogram of SGPA of every semester of a particular student. Figure shows the SGPA of a student which has highest GPA in Fall 2018 and minimum GPA in Fall 2017.



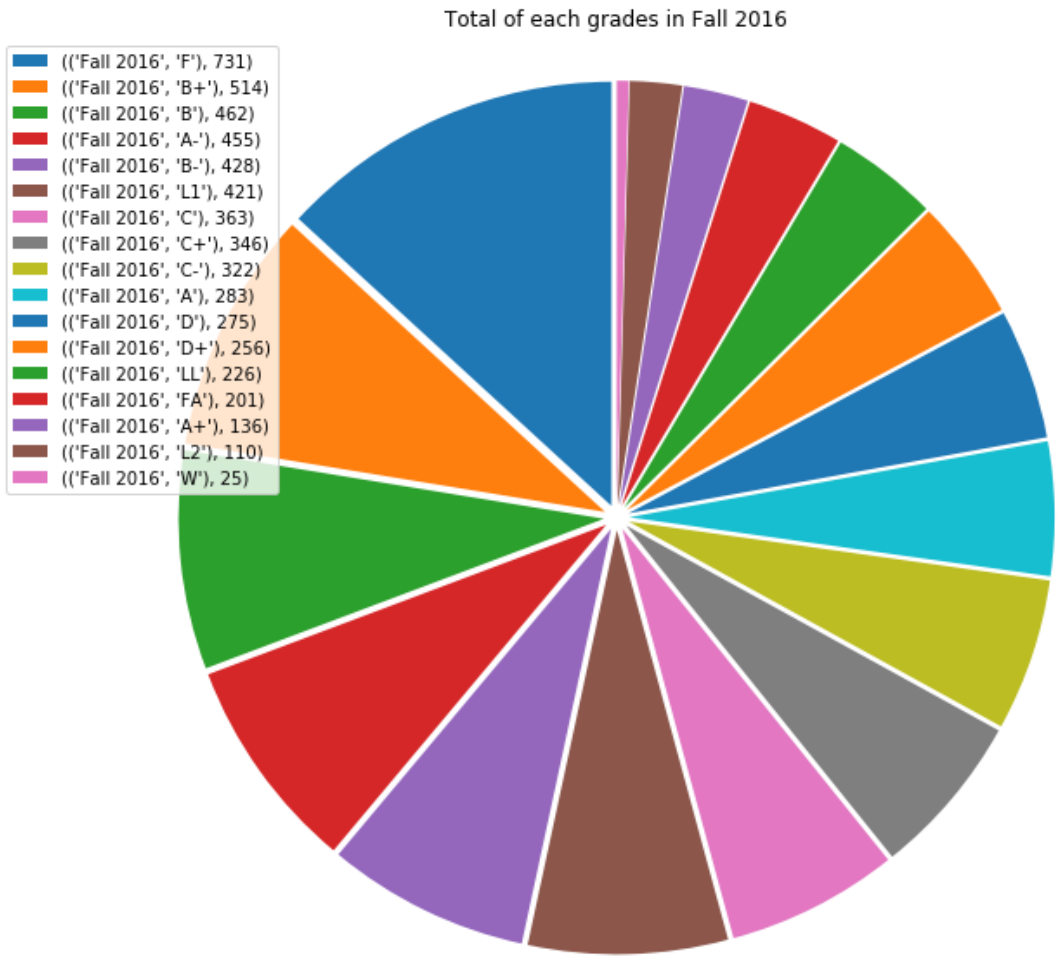
Warning Percentage:

It gives a pie chart of warning percentage in a semester. Percentage is computed based on total warning counts of the data given. Figure illustrates Summer 2019 and Summer 2018 has minimum warning percentage while spring 2017 has highest warning percentage.



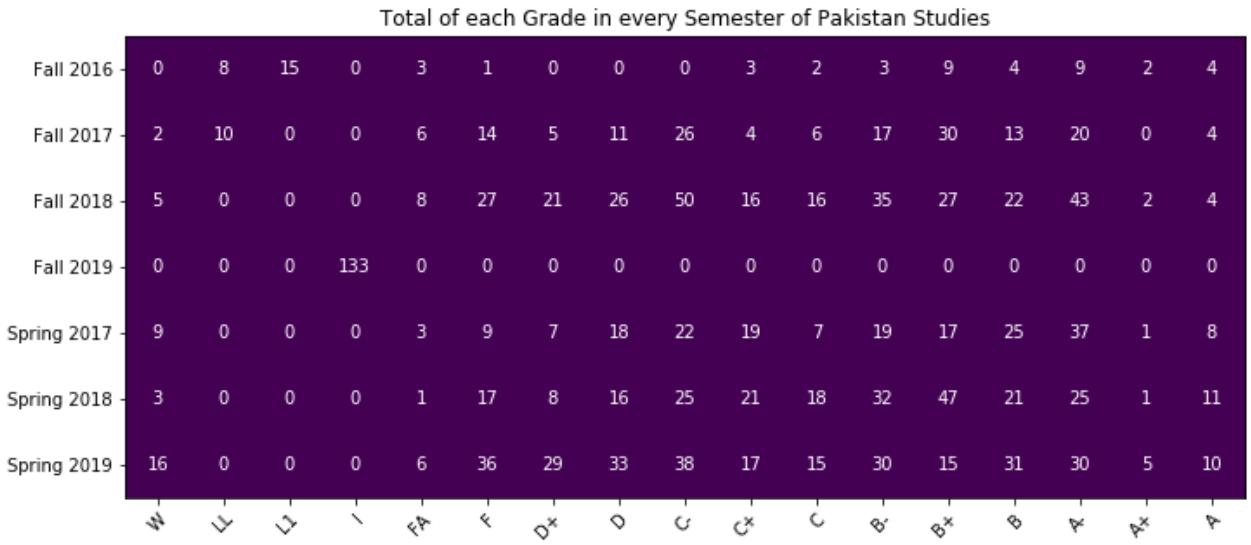
Total of each Grade in a Semester:

It gives a pie chart showing the total number of a grade in a specific semester. Figure is showing grades of Fall 2016. F is assigned to 731 students which is maximum. Only 25 students withdrew their courses which is minimum.



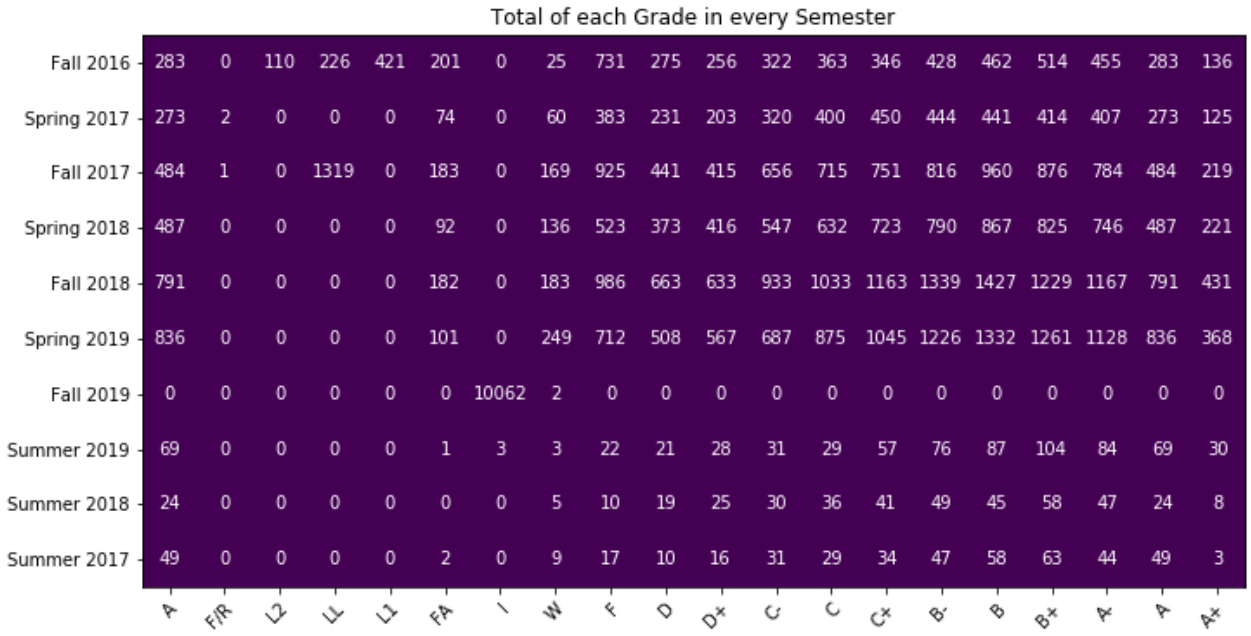
Total of each Grade in every Semester of a Course:

It gives a heat map of grades of a particular course assigned in all semesters. Below is the heat map of grades of Pakistan Studies.



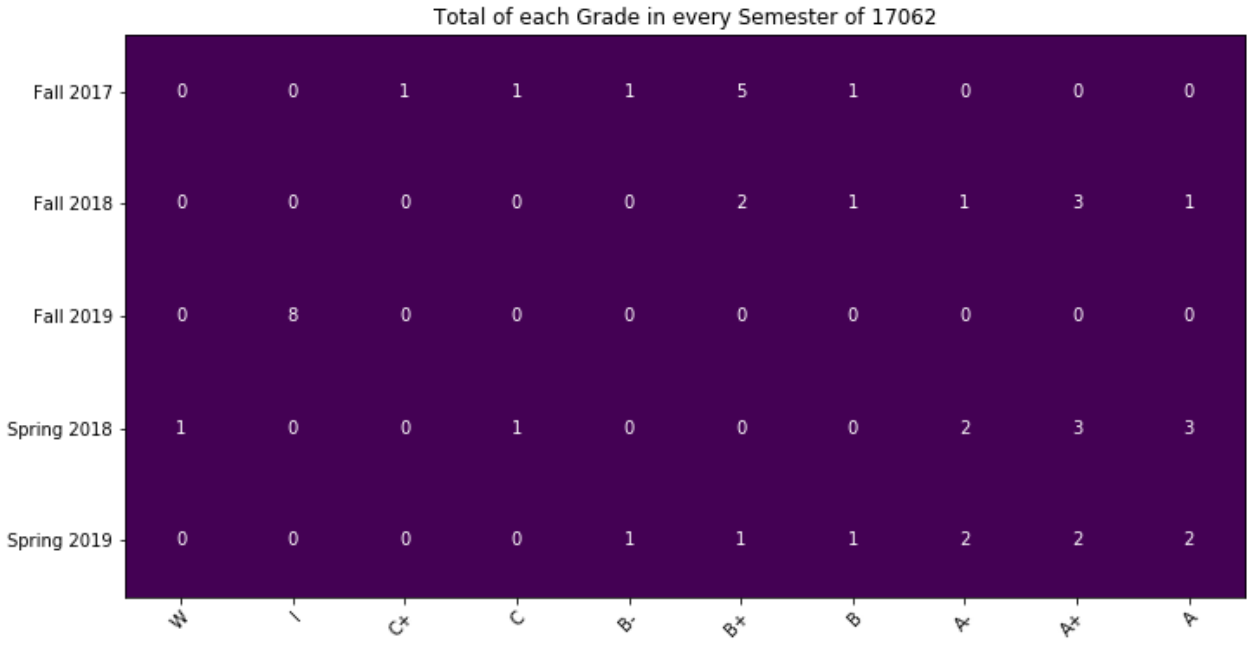
Total of each Grade in every Semester:

It gives a heat map of grades of a particular course assigned in all semesters.



Total of each Grade in every Semester of a Student:

It gives a heat map of grades of a particular student assigned in all semesters. Below is the heat map of grades of 170062.



Students advisory using kNN:

Take each semester as x and the next semester as y

Taken Courses(X1)

Courses Taken in a semester DIP, CNET, Automata

Grades (X2)

Grades in That Semester A, B, C

CGPA (X3)

CGPA till that semester 2.5

SGPA (X4)

SGPA in a semester 3.5

Warning (X5)

Warning Count in a semester 1

Recommended Courses (Y)

Courses Taken in next semester DM, Num. Meth, SE

Match with best x with given weights to find suitable y.

Problems in Data:

- In some columns of SGPA there is no value.
- There is some data of MS students which have identical roll numbers to Undergraduate students.
- Courses like Data Structures have more than one course codes that was why we were unable to check courses on the basis of course codes.
- Excel file is a mixture of Business, Computer Science and Electrical schools' data which makes hard to determine the recommended courses. Except one or two courses, courses are similar in first 4 semesters that's why some courses of other departments are recommended. Accuracy can be increased if the data is of one particular department.