CHAPTER 1
INTRODUCTION

1.1 Introduction

The manual work of result analysing takes a tedious turn when a large number of records are to be analysed upon. So in order to reduce the manual process of documentation and analysing this project has been developed. We aimed at creating an analysing method which erases the need of removing conclusions manually.

We have considered university result pdf file for the academic year 2017-2018(SE COMPUTER) for the analysis. We analysed various parameters from result analysis of the university result pdf file (which is in the form of excel sheet).This project categorises the result distinction wise and also finds fail result percentage.

1.2 Problem Statement

To analyse various contents from university result, in order to see that the manual process of deriving conclusions is reduced using automated programs. This analysis is being done with the help of 2 languages and that are **Python** and **R**.

Data which is analysed helps to:

1. Compare online and theory result
2. Visualization of grades
3. Comparison of results division wise
4. Maximum and minimum result
5. Failed result percentage
6. Comparison of result semester wise

1.3 Design

The purpose of Result Analyzing System is to automate the existing manual system by this application project.

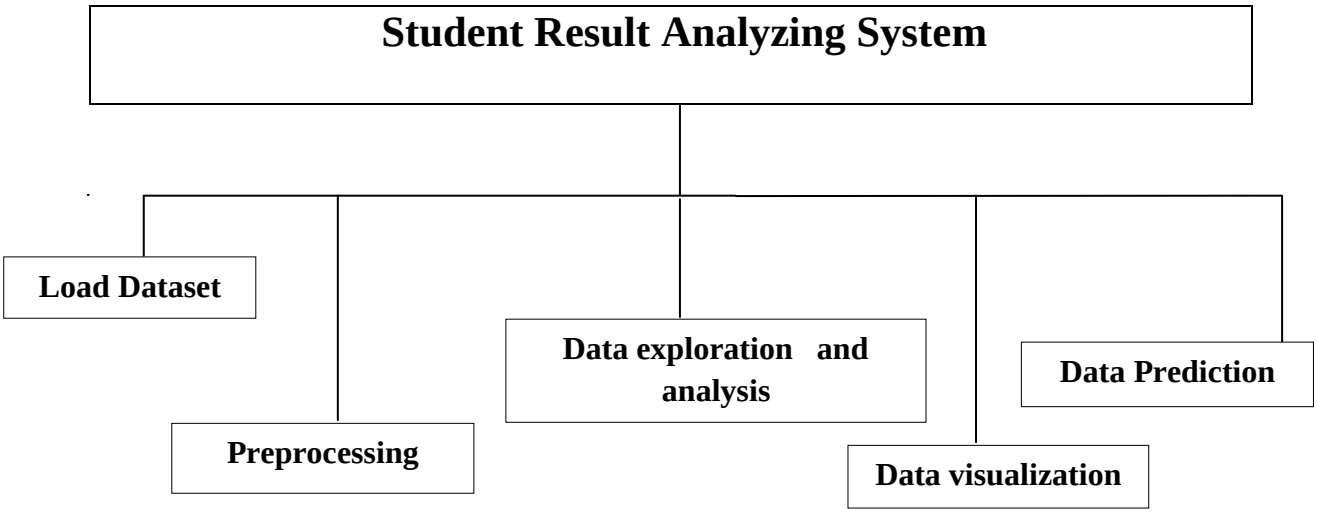


Fig. 1.3.1 Breakdown Structure

1. **Load Dataset**

It Include loading the data in the application. Here data is loaded in form of excel sheet.

2. **Data pre-processing**

It includes preprocessing of data where the unwanted data like special symbols, marks are eliminated or removed from data.

3. **Data Exploration and analysis**

The loaded data is explored and analyzed in ways such that to attain our objectives.

4. **Data Visualization**

It includes visualization of data in different mathematical forms like pi-chart, histogram, comparison graphs etc.

5. **Data Prediction**

Data prediction includes prediction of final data which will be helpful in concluding the data analysis application.

CHAPTER 2
SCOPE AND OBJECTIVE

2.1 Scope

The data is from university result pdf (SPPU) 2017-2018..

There are factors which we are considering for analysing are online marks, theory marks, total marks, pointers, grades of result.

The analysis will be mainly focused on university result pdf file 2017-2018 (SE-COMPUTER).

The analysis will show data of above mentioned topics in the form of Graphs, Bar charts, Pie-Charts, Histograms.

The analysis will show graphs of

1. Online marks vs theory marks
2. Visualization of grades
3. Percentage
4. Pointers
5. Overall Percentage
6. Fail students result

2.2 Objectives

1. The main objective of the project is to build an application program that provides a platform to reduce the manual work for analysis the result of students.
2. To conclude what is the performance of students in a subject.
3. To conclude which subject needs to be given more attention.
4. To show year result conclusion with pi-chart graph, bar chart, histogram, etc.

CHAPTER 3
DATA AND REQUIREMENTS

3.1 Data

The data is taken from university result pdf file .The data is in the form of excel sheet which is later converted into **CSV**(Comma Separated Value) files.

3.2 Hardware Requirements

1. 2GB RAM
2. Dual core Intel or AMD processor

3.3 Software Requirements

1. Python Programming Language
2. R Programming Language
3. Windows or Linux operating system

3.4 Tools and Techniques

There are some tools and libraries that are required while doing the analysis of the data and they are as follows:

1. matplotlib
2. numpy
3. csv
4. pandas

There are some techniques that have been used for the visualization of data and that are as follows:

1. Graphs
2. Bar Charts
3. Pie-charts
4. Histogram
6. Linear Regression

CHAPTER 4
DATA PRE-PROCESSING

The data which we have extracted from pdf may contain some missing values, some garbage values, extra characters, blank spaces, symbols which are not required for analysis.
So in order to make the data readable and executable(able to be analysed) we are pre-processing the data.

4.1 Data Pre-processing

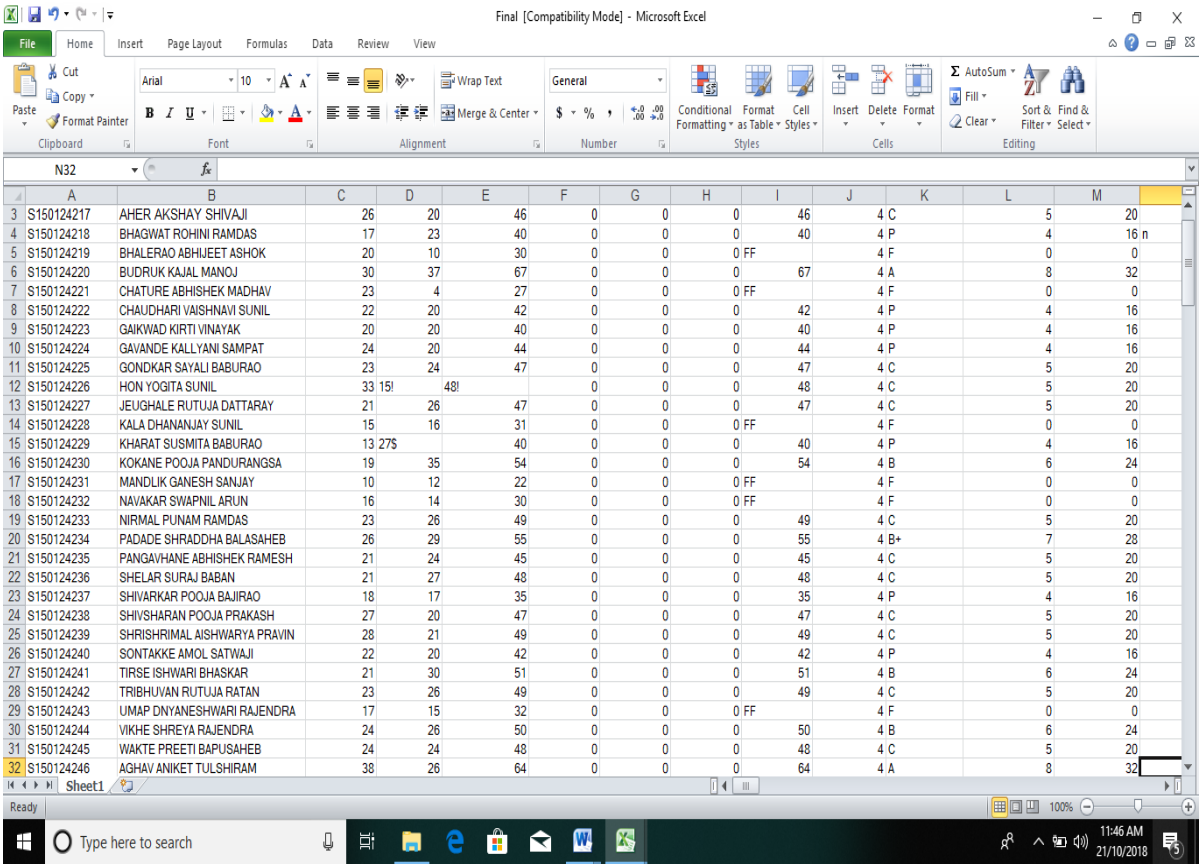
Identify missing values, blank spaces and special symbols in the given data, then put <NA> (Not Applicable) in the records where there are missing values and remove special symbols as well as blank spaces.

Code (in python):-

```
for row in data:
      str=row[3]
      if row[3]=="":
            row[3]="NA";
            serial_no.append(row[3])
      elif " " in str or "%" in str or "$" in str:
            str=str.replace(" ","")
            str=str.replace("$","")
            str=str.replace("%","")
```

Remove records consisting of <NA> and type caste the remaining values to integer or float respectively, so that the records are totally readable and can be **easily** analysed.

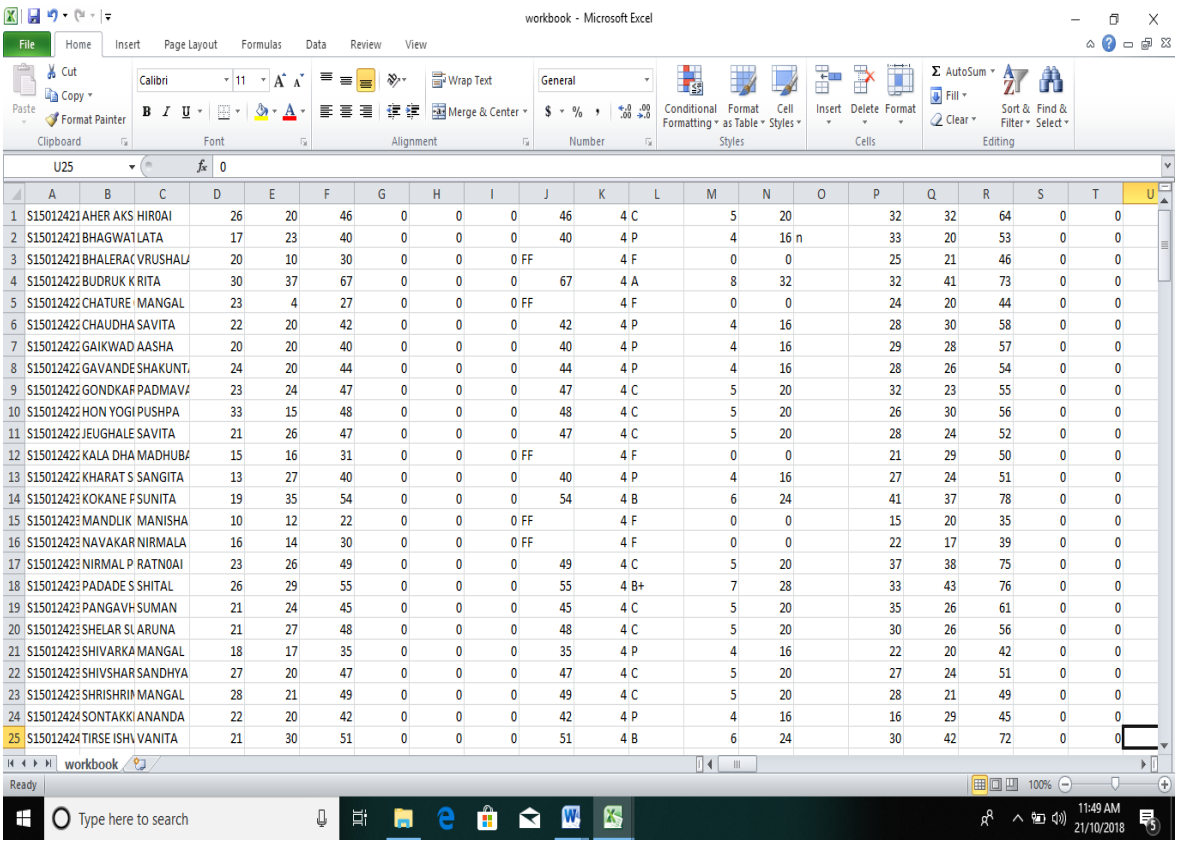
Below picture shows the data before pre-processing.



	A	B	C	D	E	F	G	H	I	J	K	L	M
3	S150124217	AHER AKSHAY SHIVAJI	26	20	46	0	0	0	46	4	C	5	20
4	S150124218	BHAGWAT ROHINI RAMDAS	17	23	40	0	0	0	40	4	P	4	16 n
5	S150124219	BHALERAO ABHIJEET ASHOK	20	10	30	0	0	0	FF	4	F	0	0
6	S150124220	BUDRUK KAJAL MANOJ	30	37	67	0	0	0	67	4	A	8	32
7	S150124221	CHATURE ABHISHEK MADHAV	23	4	27	0	0	0	FF	4	F	0	0
8	S150124222	CHAUDHARI VAISHNAVI SUNIL	22	20	42	0	0	0	42	4	P	4	16
9	S150124223	GAIKWAD KIRTI VINAYAK	20	20	40	0	0	0	40	4	P	4	16
10	S150124224	GAVANDE KALLYANI SAMPAT	24	20	44	0	0	0	44	4	P	4	16
11	S150124225	GONDKAR SAYALI BABURAO	23	24	47	0	0	0	47	4	C	5	20
12	S150124226	HON YOGITA SUNIL	33	15!	48!	0	0	0	48	4	C	5	20
13	S150124227	JEUGHALE RUTUJA DATTARAY	21	26	47	0	0	0	47	4	C	5	20
14	S150124228	KALA DHANANJAY SUNIL	15	16	31	0	0	0	FF	4	F	0	0
15	S150124229	KHARAT SUSMITA BABURAO	13	27$	40	0	0	0	40	4	P	4	16
16	S150124230	KOKANE POOJA PANDURANGSA	19	35	54	0	0	0	54	4	B	6	24
17	S150124231	MANDLIK GANESH SANJAY	10	12	22	0	0	0	FF	4	F	0	0
18	S150124232	NAVAKAR SWAPNIL ARUN	16	14	30	0	0	0	FF	4	F	0	0
19	S150124233	NIRMAL PUNAM RAMDAS	23	26	49	0	0	0	49	4	C	5	20
20	S150124234	PADADE SHRADDHA BALASAHEB	26	29	55	0	0	0	55	4	B+	7	28
21	S150124235	PANGAVHANE ABHISHEK RAMESH	21	24	45	0	0	0	45	4	C	5	20
22	S150124236	SHELAR SURAJ BABAN	21	27	48	0	0	0	48	4	C	5	20
23	S150124237	SHIVARKAR POOJA BAJIRAO	18	17	35	0	0	0	35	4	P	4	16
24	S150124238	SHIVSHARAN POOJA PRAKASH	27	20	47	0	0	0	47	4	C	5	20
25	S150124239	SHRISHRIMAL AISHWARYA PRAVIN	28	21	49	0	0	0	49	4	C	5	20
26	S150124240	SONTAKKE AMOL SATWAJI	22	20	42	0	0	0	42	4	P	4	16
27	S150124241	TIRSE ISHWARI BHASKAR	21	30	51	0	0	0	51	4	B	6	24
28	S150124242	TRIBHUVAN RUTUJA RATAN	23	26	49	0	0	0	49	4	C	5	20
29	S150124243	UMAP DNYANESHWARI RAJENDRA	17	15	32	0	0	0	FF	4	F	0	0
30	S150124244	VIKHE SHREYA RAJENDRA	24	26	50	0	0	0	50	4	B	6	24
31	S150124245	WAKTE PREETI BAPUSAHEB	24	24	48	0	0	0	48	4	C	5	20
32	S150124246	AGHAV ANIKET TULSHIRAM	38	26	64	0	0	0	64	4	A	8	32

Fig.4.1.1 Data before pre-processing

Now this picture shows the data after pre-processing which is in the cleaned format.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	S15012421	AHER AKS	HIR0AI	26	20	46	0	0	0	46	4	C	5	20		32	32	64	0	0	
2	S15012421	BHAGWAT	LATA	17	23	40	0	0	0	40	4	P	4	16	n	33	20	53	0	0	
3	S15012421	BHALERAO	VRUSHALA	20	10	30	0	0	0	FF	4	F	0	0		25	21	46	0	0	
4	S15012422	BUDRUK K	RITA	30	37	67	0	0	0	67	4	A	8	32		32	41	73	0	0	
5	S15012422	CHATURE	MANGAL	23	4	27	0	0	0	FF	4	F	0	0		24	20	44	0	0	
6	S15012422	CHAUDHA	SAVITA	22	20	42	0	0	0	42	4	P	4	16		28	30	58	0	0	
7	S15012422	GAIKWAD	AASHA	20	20	40	0	0	0	40	4	P	4	16		29	28	57	0	0	
8	S15012422	GAVANDE	SHAKUNT/	24	20	44	0	0	0	44	4	P	4	16		28	26	54	0	0	
9	S15012422	GONDKAR	PADMAVA	23	24	47	0	0	0	47	4	C	5	20		32	23	55	0	0	
10	S15012422	HON YOGI	PUSHPA	33	15	48	0	0	0	48	4	C	5	20		26	30	56	0	0	
11	S15012422	JEUGHALE	SAVITA	21	26	47	0	0	0	47	4	C	5	20		28	24	52	0	0	
12	S15012422	KALA DHA	MADHUBA	15	16	31	0	0	0	FF	4	F	0	0		21	29	50	0	0	
13	S15012422	KHARAT S	SANGITA	13	27	40	0	0	0	40	4	P	4	16		27	24	51	0	0	
14	S15012423	KOKANE P	SUNITA	19	35	54	0	0	0	54	4	B	6	24		41	37	78	0	0	
15	S15012423	MANDLIK	MANISHA	10	12	22	0	0	0	FF	4	F	0	0		15	20	35	0	0	
16	S15012423	NAVAKAR	NIRMALA	16	14	30	0	0	0	FF	4	F	0	0		22	17	39	0	0	
17	S15012423	NIRMAL P	RATN0AI	23	26	49	0	0	0	49	4	C	5	20		37	38	75	0	0	
18	S15012423	PADADE S	SHITAL	26	29	55	0	0	0	55	4	B+	7	28		33	43	76	0	0	
19	S15012423	PANGAVH	SUMAN	21	24	45	0	0	0	45	4	C	5	20		35	26	61	0	0	
20	S15012423	SHELAR SU	ARUNA	21	27	48	0	0	0	48	4	C	5	20		30	26	56	0	0	
21	S15012423	SHIVARKA	MANGAL	18	17	35	0	0	0	35	4	P	4	16		22	20	42	0	0	
22	S15012423	SHIVSHAR	SANDHYA	27	20	47	0	0	0	47	4	C	5	20		27	24	51	0	0	
23	S15012423	SHRISHRIM	MANGAL	28	21	49	0	0	0	49	4	C	5	20		28	21	49	0	0	
24	S15012424	SONTAKKI	ANANDA	22	20	42	0	0	0	42	4	P	4	16		16	29	45	0	0	
25	S15012424	TIRSE ISH	VANITA	21	30	51	0	0	0	51	4	B	6	24		30	42	72	0	0	

Fig.4.1.2 Data after pre-processing

CHAPTER 5
EXPLORATORY DATA ANALYSIS

5.1 Exploratory Analysis

The data can be explored with the help of various funtions like dim(), str(), summary(). We can see the structure of tha data frame that we have created and we also can see the dimensions as well as the summary of the data frame. We can see the whole data frame using print() in R programming language

.**Code (in R Language):-**

```
df=read.csv("workbook1.csv",header=TRUE)
print(df)
dim(df)
str(df)
summary(df)
```

1. The function **read.csv**() will read the data in mentioned csv file and will create a data frame.

2. The function **dim()** will give the dimensions of the data frame.

3. The function **str()** will give the structure of data frame and will show whether the attributes in it are factor or numeric data types, it will identify if factor is applicable or not and if yes then it will show the factors those are recognized and will also give the count as well as it will assign numbers to the factors identified by alphabetical precedences.

4. The function **summary()** will give the minimum, maximum, median value of each column(attribute).

CHAPTER 6
VISUALIZATION

This part includes the actual representation (visualization) of data. The data is visualized using both R and Python programming languages.
The libraries that are used while displaying the data are as follows:

1. numpy
2. matplotlib.pyplot

The functions that are used while displaying data are as follows:

1. plt.xticks()
2. plt.text()
3. plt.plot()
4. plt.legend()
5. plt.pie()
6. plt.hist()
7. plt.bar()
8. plt.xlabel()
9. plt.ylabel()
10. plt.show()

Codes (in Python) that include below functions:-

6.1 Analyzed Graphs (Bar chart) of each subject

Now we are going to see the grades secured by students in graphical manner. The graphs are plotted using mathplotlib and numpy.

```
barsticks=plt.bar(list3,list2,align='center',alpha=1,color='g''b''
y''m''c''r''k''b')
for rect in barsticks:
     height=rect.get_height()
     plt.text(rect.get_x()+rect.get_width()/2,height,'%d'%
int(height),ha='center')
plt.legend(barsticks[:8],grades)
plt.xticks(list3,grades)
plt.tight_layout()
plt.title('Visualization Of Grades Obtain in DM')
plt.xlabel('GRADES')
plt.ylabel('NO OF GRADES')
plt.show()
```

1. **np.arrange()** which is a function of numpy library will arrange the districts and will store it in x_pos variable.
2. **plt.plot()** which is a function of matplotlib.pyplot is used to plot the graph. The parameters that are passed must be in sequence like X-co-ordinates, then Y-co-ordinates and the ‘ro’ is point style and linestyle=’- -‘.
3. Then there is **plt.xlabel()** and **plt.ylabel()** to give labels to each X and Y axis; **plt.title()** is used to give title to the graph and **plt.show()** is used at last for displaying the graph.

6.2 Online Vs Theory comparison (Histogram) for each subject:

We have plotted histogram for comparison between online and theory marks of each subject.

Code:

```
for row in plot:
        ol.append(int(row[3]))
        th.append(int(row[4]))
legends=['Online','Theory']
ranges=['0-10','10-20','20-30','30-40','40-50']
plt.hist([ol,th],x,rwidth=0.7,alpha=1)
plt.title('Online vs Theory marks of DM [rangewise]')
plt.xlabel("RANGE OF MARKS")
plt.ylabel("NUMBER OF STUDENTS")
plt.legend(legends)
plt.show()
```

1. **plt.hist()** is used to draw histogram, which is used to show frequency of a particular thing in a particular range(category).
2. The parameters that are passed to the above function are list of things whose frequency we want to see and list categories in which we want to observe the frequency of things.
3. **rwidth=0.8** can be given as a parameter to **plt.hist()** in order to have spacing between each category of histogram.

6.1.1 Discrete Mathematics (DM):

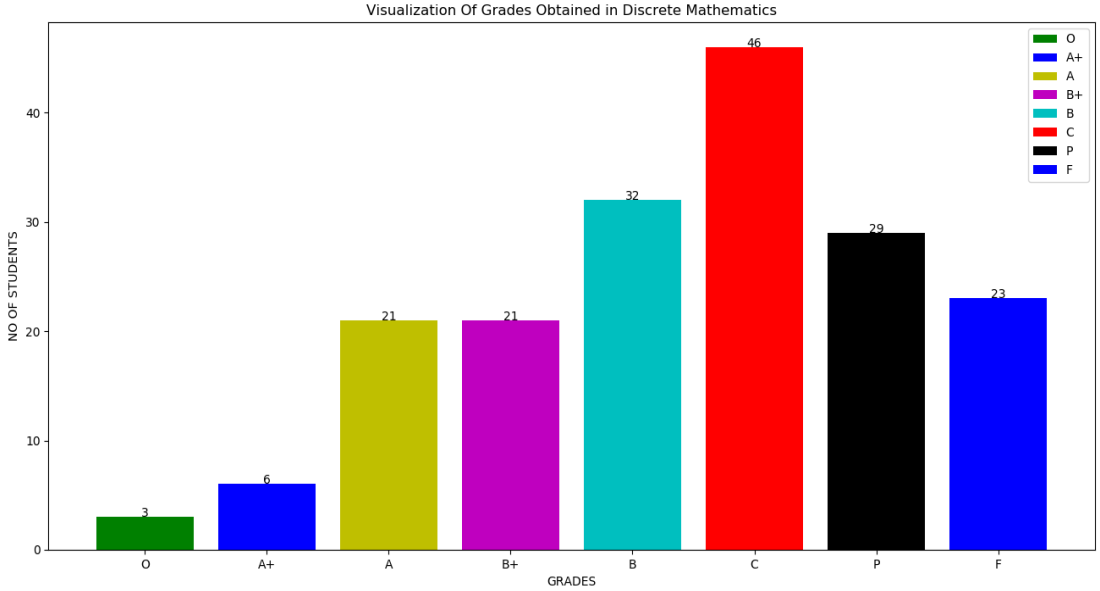


Fig. 6.1.1 Visualization of grades for the subject DM

From the above graph we can clearly see that out off 181 students

1. The number of students getting C grade (i.e.in the range of 60-69) is the most compared to others.
2. Students securing B grade are 32 which is the second highest.
3. Only 3 students have secured O grade (i.e. in range 90-99).
4. 23 students couldn't clear this subject in the second attempt.
5. Number of students securing A and B+ are equal.

Online Vs Theory marks comparison

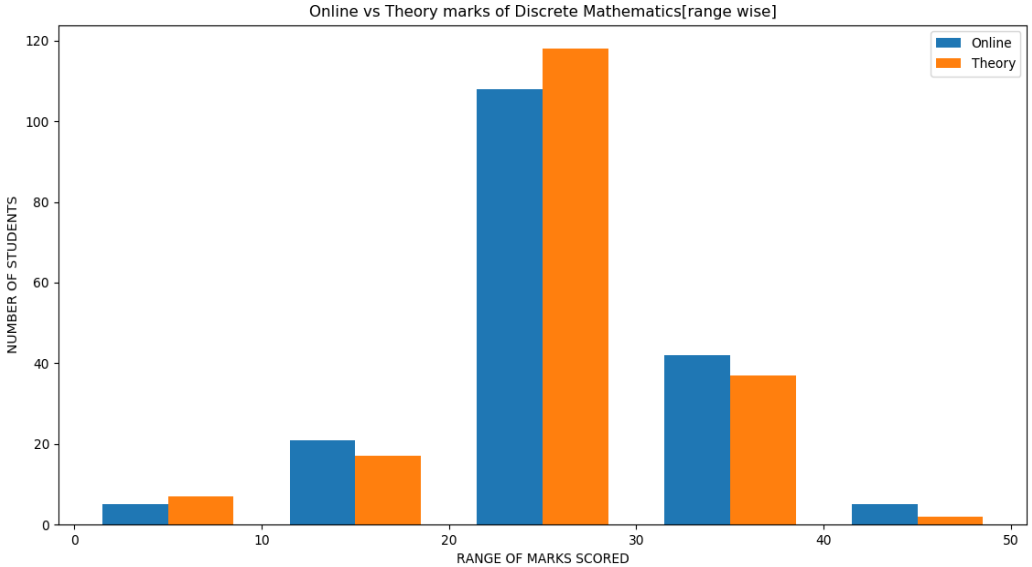


Fig. 6.2.1.Online Vs Theory marks for DM

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. In most of ranges the online marks beats the theory marks in case of scores, thus concluding that students get more in online compared to theory. Students securing marks in the range of 40-50 are the least for both online as well as theory.

6.1.2 Digital Electronics and Logic Design (DELD):

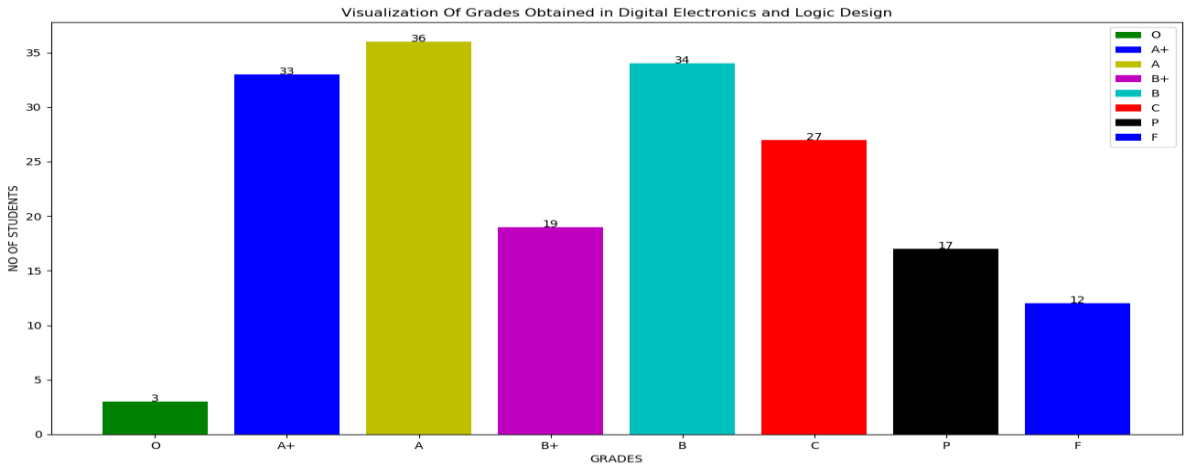


Fig. 6.1.2 Visualization of grades for the subject DELD

In the above given graph we can clearly see out off 181 students;

1. The number of students securing A+, A, B grades are nearly equal(i.e. 33, 36, 34 respectively)
2. Students securing O grade is the least.
3. Only 3 students have secured O grade (i.e. in range 90-99).
4. 12 students couldn't clear this subject in the second attempt.
5. The students securing B+ and C grades are almost equal.

Online Vs Theory marks comparison

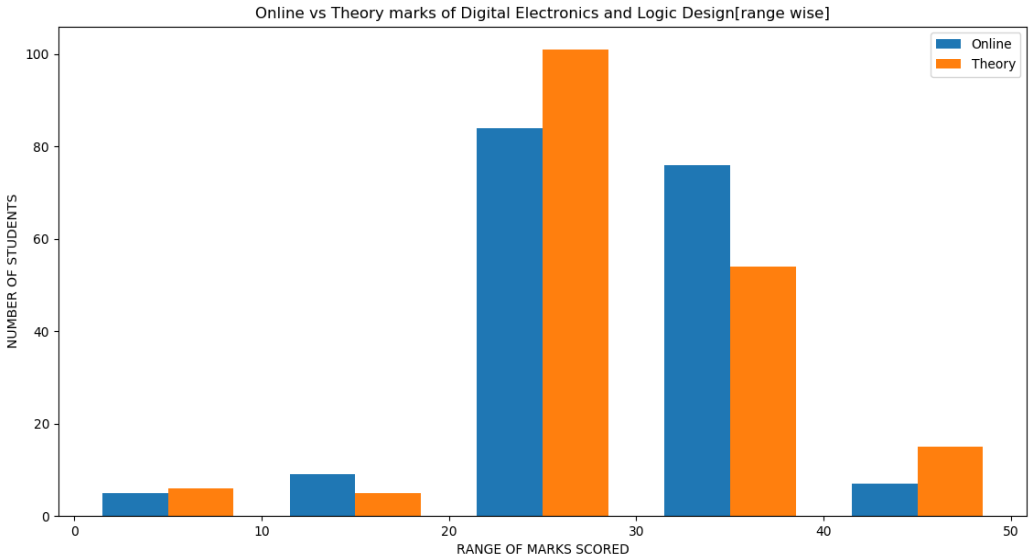


Fig. 6.2.2. Online Vs Theory marks for DELD

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Here, in some cases ranges the online marks beats the theory marks as well as vice versa, thus concluding that students ratio for marks scored in online as well as theory are equal. For the range 20-30 theory marks scored are more whereas for the range 30-40 online marks scored exceeds to that scored in theory.

6.1.3 Computer Organization and Architecture (COA):

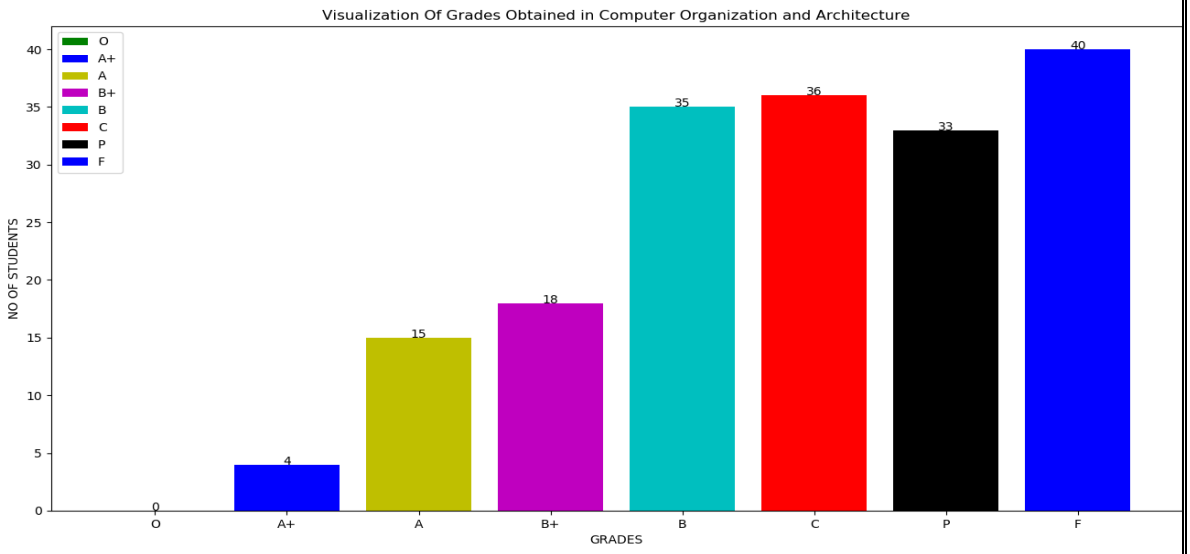


Fig. 6.1.3 Visualization of grades for the subject COA

In the above given graph we can clearly see

1. The number of students getting B and C grade are nearly equal.
2. The ratio of latter grades are lot higher(the scoring in this subject is less).
3. The students are clearing this subject on average.
4. No one could secure an O grade(i.e. in range 90-99).
5. 40 students couldn't clear this subject in the second attempt.
6. The rate of failure is high for COA(i.e. students have a hard time understanding theory subjects).
7. The students getting an A+ grade are 4.

Online Vs Theory marks comparison

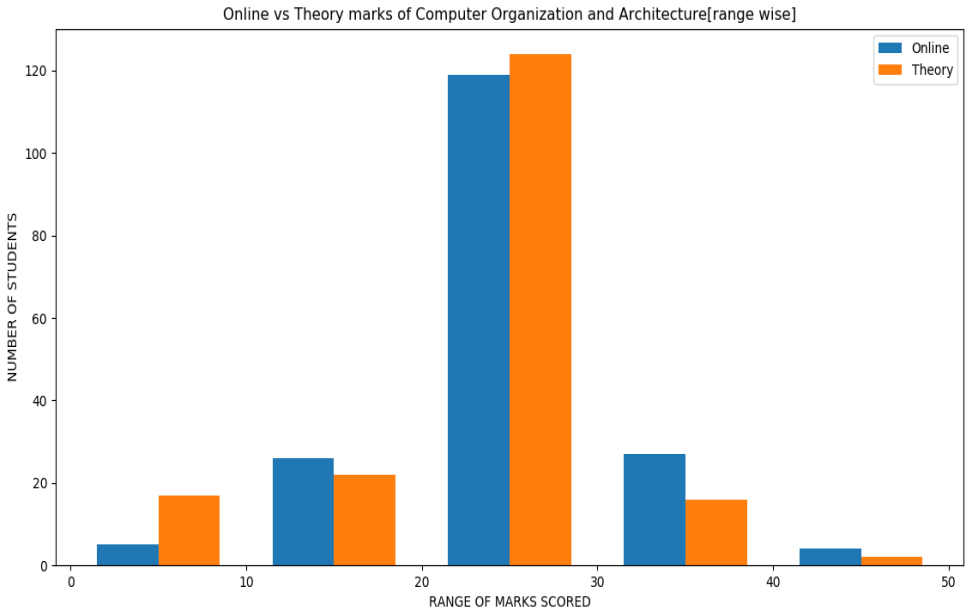


Fig. 6.2.3 Online Vs Theory marks for COA

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are the least for both online and theory.

6.1.4 Data Structure and Algorithms (DSA):

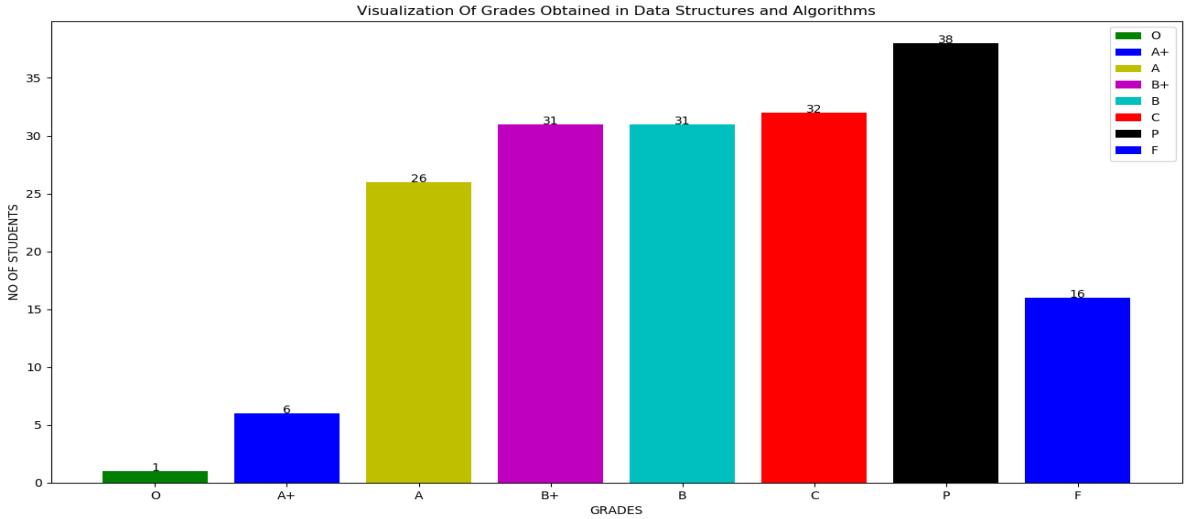


Fig. 6.1.4 Visualization of grades for the subject DSA

In the above given graph we can clearly see

1. The number of students getting P grade is the most compared to others.
2. The number of students securing A, B+, B, C grades are nearly equal(i.e. 26,31,31,32 respectively)
3. Students securing B+ and B grades are same
4. 6 students have secured an A+ grade
5. Only 1 student have secured O grade(i.e. above 95).
6. 26 students couldn't clear this subject in the second attempt.

Online Vs Theory marks comparison

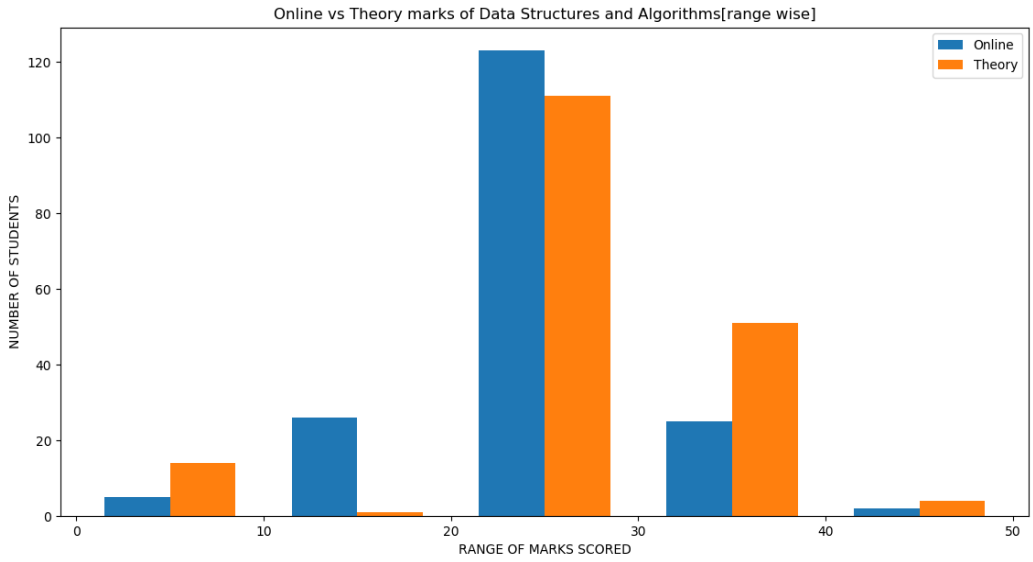


Fig. 6.2.4 Online Vs Theory marks for DSA

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are the least for online whereas for the theory the range 20-30 has the least score. Here, in some cases ranges the online marks beats the theory marks as well as vice versa, thus concluding that students ratio for marks scored in online as well as theory are equal.

6.1.5 Object Oriented Programming (OOP):

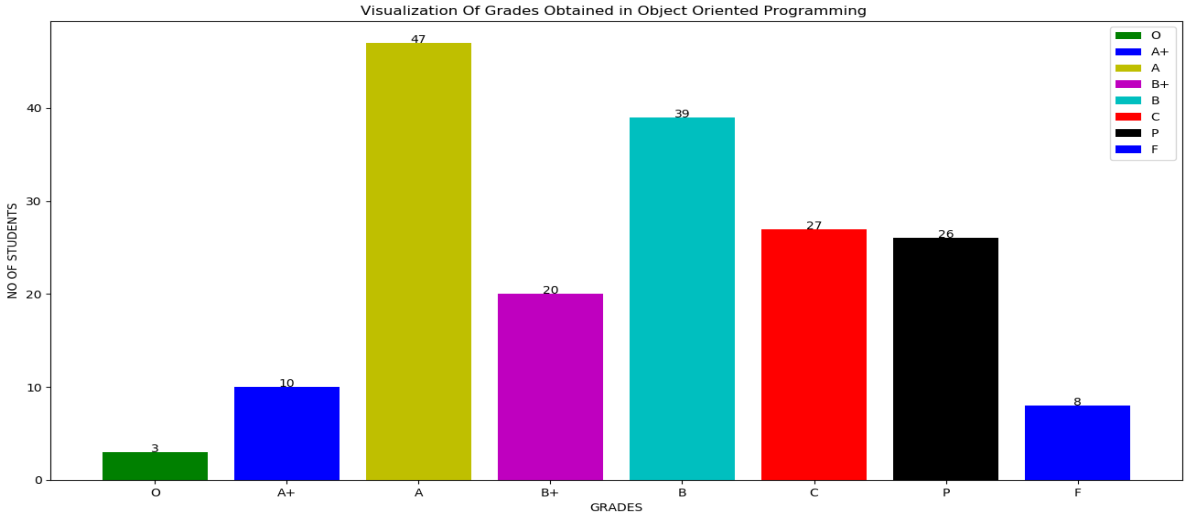


Fig. 6.1.5 Visualization of grades for the subject OOP

In the above given graph we can clearly see

1. The number of students getting A grade (i.e.in the range of 80-89) is the most compared to others.
2. The number of students securing C and P grades are nearly equal(i.e. 27, 26 respectively)
3. Students securing B grade are 32 which is the second highest.
4. 20 students have secured an A+ grade(i.e. above 90)
5. Only 3 students have secured O grade(i.e. in between 90-99).
6. 8 students couldn't clear this subject in the second attempt.

Online Vs Theory marks comparison

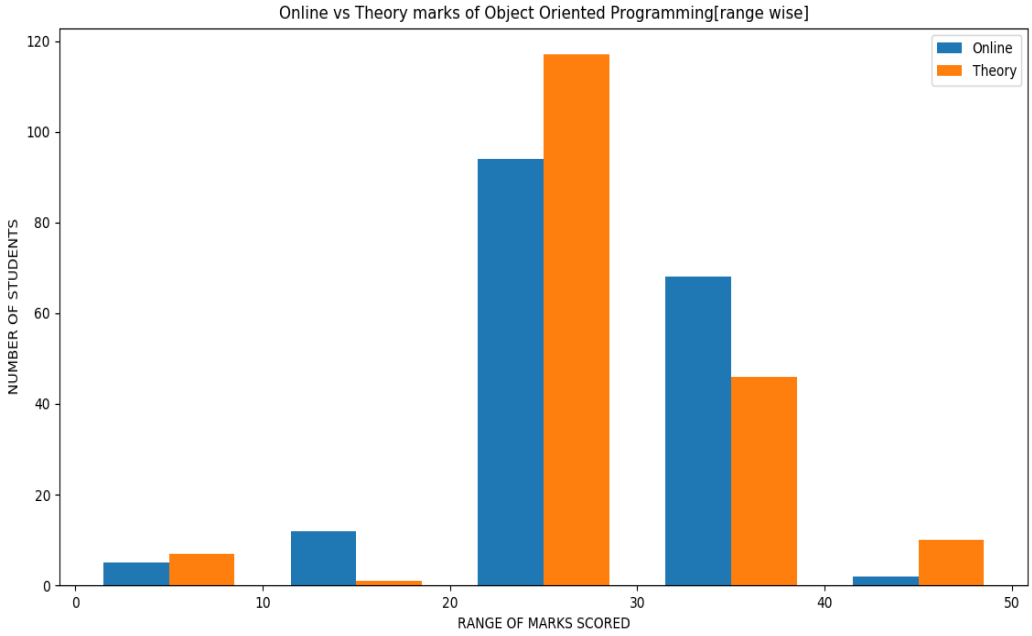


Fig. 6.2.5 Online Vs Theory marks for OOP

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are the least for online whereas for the theory the range 20-30 has the least score.

6.1.6 Engineering Mathematics III (EM-III):

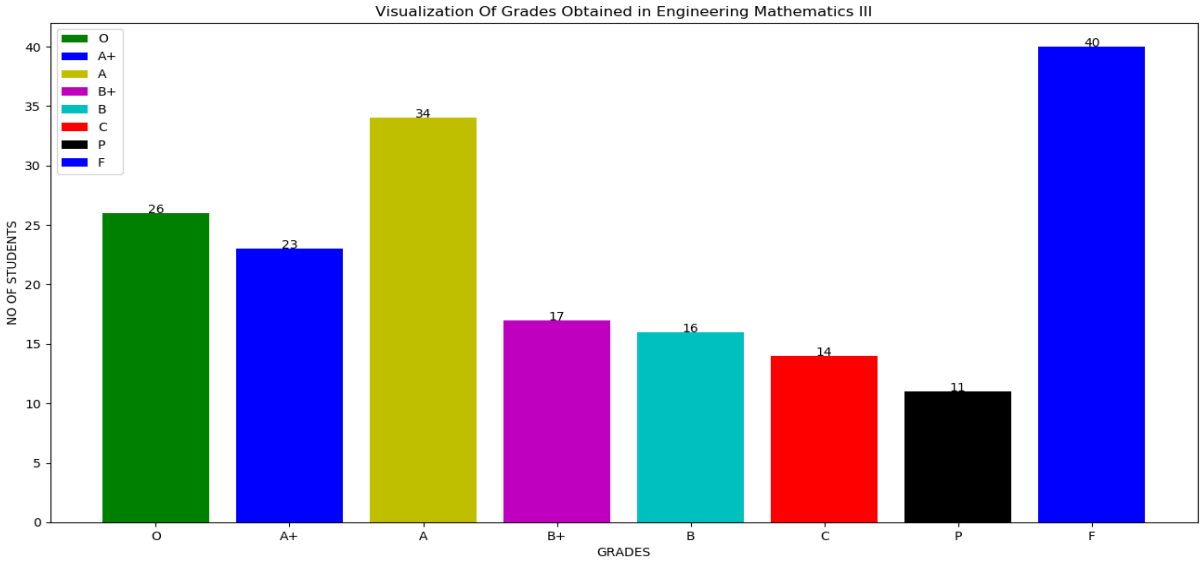


Fig. 6.1.6 Visualization of grades for the subject EM3

In the above given graph we can clearly see

1. The failure percentage is the most (i.e. 40 students couldn't clear this subject)
2. Students securing A grade are 34 which is the second highest.
3. 23 students have secured an A+ grade(i.e. above 95)
4. The students getting an O grade is the most compared to other subjects (i.e. Scoring is easy in mathematical subject.)
5. The result to contradictory to the result of COA.
6. The number of students securing B and B+ grades are nearly equal(i.e. 17,16 respectively).

Online Vs Theory marks comparison

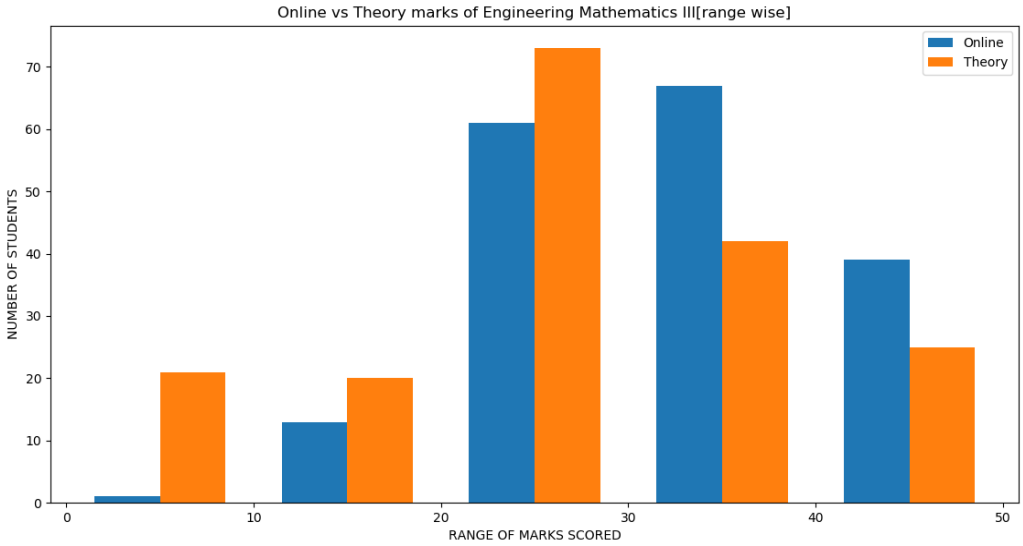


Fig. 6.2.6 Online Vs Theory marks for M3

In the above graph, for the range 20-30 and 30-40 firstly the theory marks exceeds the online marks whereas it is the opposite for the next range concluding that both the ranges equally balance each other. Students securing marks in the range of 0-10 are the least for both online as well as theory.

6.1.7 Computer Graphics (CG):

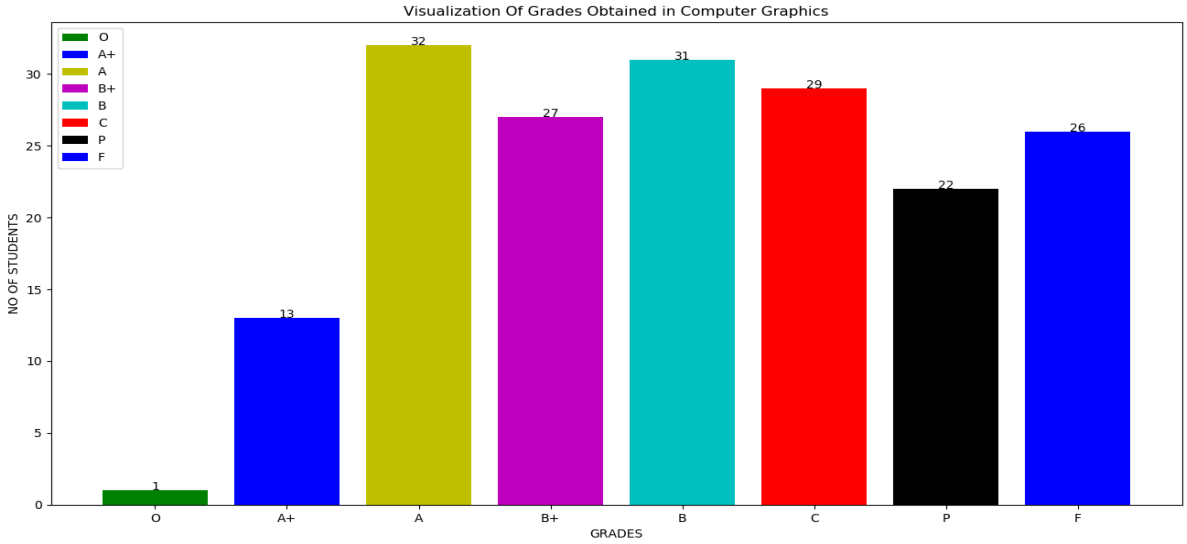


Fig. 6.1.7 Visualization of grades for the subject CG

In the above given graph we can clearly see

1. 26 students couldn't clear this subject
2. Students securing A grade are 32 which is the highest.
3. 13 students have secured an A+ grade(i.e. above 90)
4. Only 1 student could secured an O grade
5. The number of students securing B and A grades are nearly equal
6. The number of students securing B+ and C grades are nearly equal

Online Vs Theory marks comparison

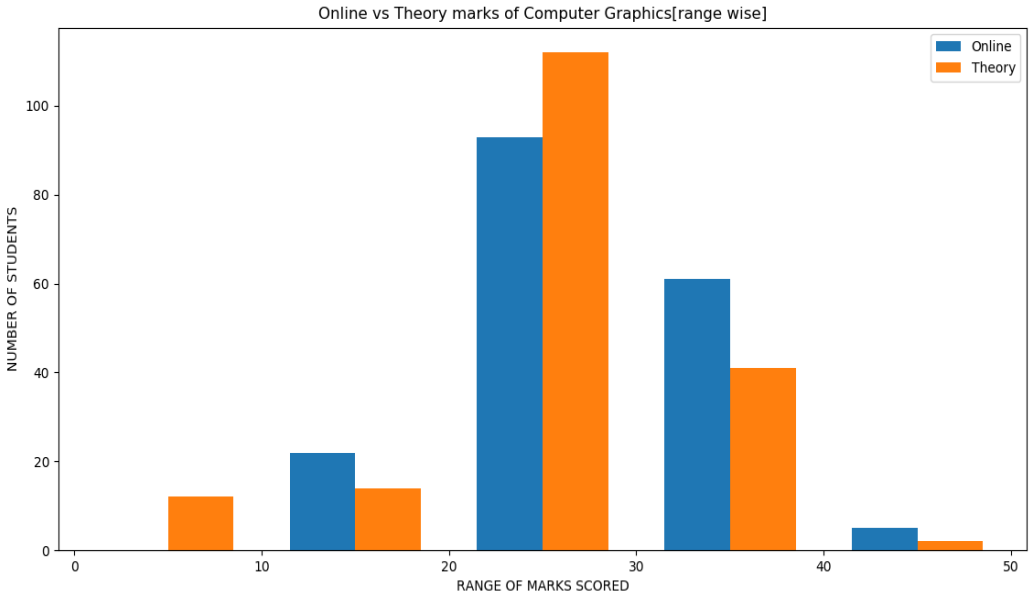


Fig. 6.2.7 Online Vs Theory marks for CG

In the above graph, for the range 20-30 and 30-40 firstly the theory marks exceeds the online marks whereas it is the opposite for the next range concluding that both the ranges equally balance each other. Students securing marks in the range of 0-10 are the least for online whereas least in the range of 40-50 for theory.

6.1.8 Advanced Data Structure (ADS):

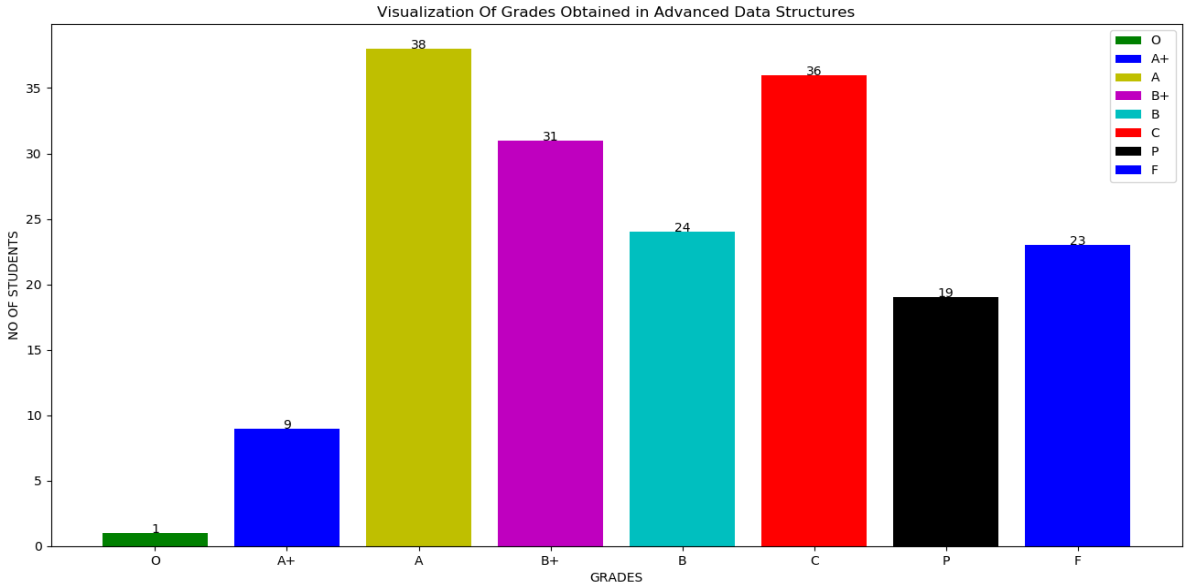


Fig. 6.1.8 Visualization of grades for the subject ADS

In the above given graph we can clearly see

1. 23 students couldn't clear this subject
2. Students securing A grade are 38 which is the highest.
3. Students securing C grade are 36 which is the second highest.
4. Only 9 students have secured an A+ grade(i.e. above 90)
5. Only 1 student could secured an O grade.

Online Vs Theory marks comparison

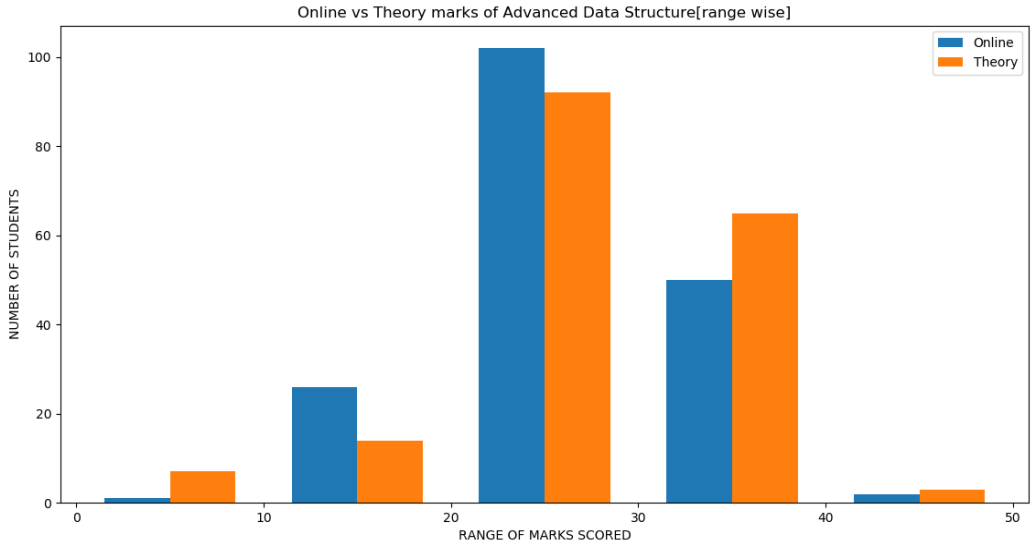


Fig. 6.2.8 Online Vs Theory marks for ADS

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are the least for online whereas for the theory the range 20-30 has the least score. Here, in some cases ranges the online marks beats the theory marks as well as vice versa, thus concluding that students ratio for marks scored in online as well as theory are somewhat equal.

6.1.9 Microprocessor (MP):

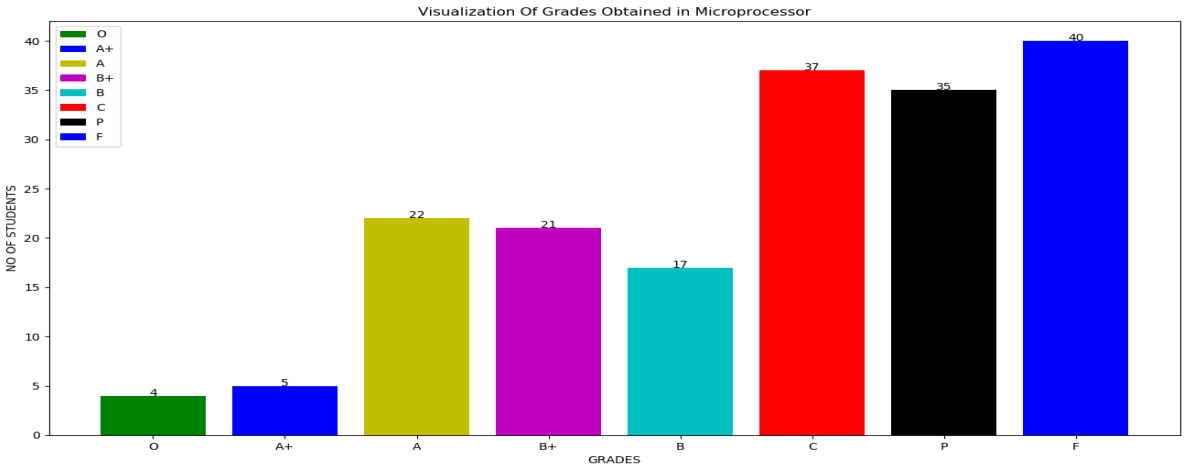


Fig. 6.1.9 Visualization of grades for the subject MP

In the above given graph we can clearly see

1. The failure result is the most (i.e.40 students couldn't clear this subject)
2. The failure result is same like M3
3. Students securing C grade are 37 which is the second highest.
4. Only 5 students have secured an A+ grade(i.e. above 90)
5. The students getting an O grade is the most compared to other subjects(i.e. Scoring is easy in mathematical subject.)
6. 4 students have secured an O grade
7. The number of students securing A and B+ grades are nearly equal (i.e. 22,21 respectively).

Online Vs Theory marks comparison

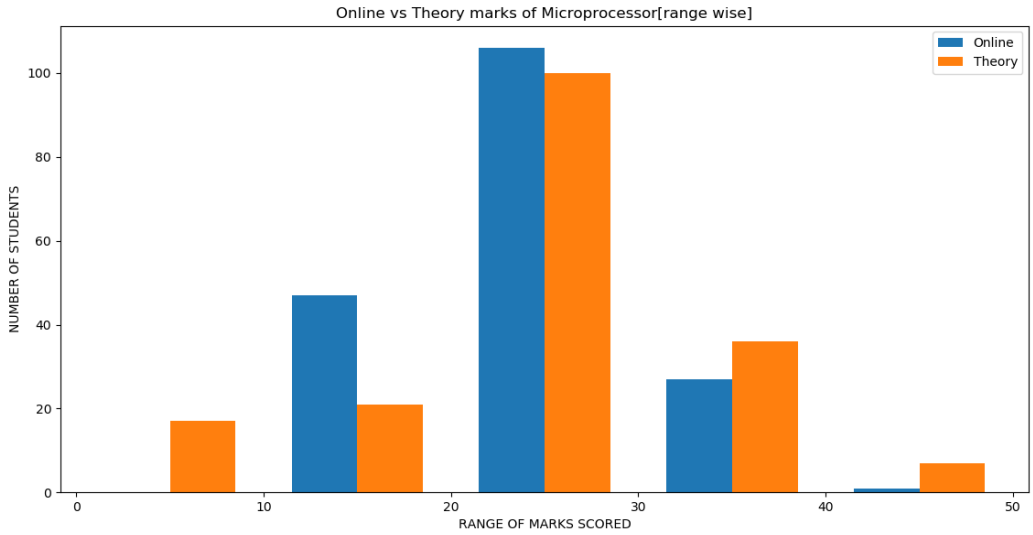


Fig. 6.2.9 Online Vs Theory marks for MP

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are the least for theory whereas for online the range 0-10 is applicable as there are no figures of marks .Here, in some cases ranges the online marks beats the theory marks as well as vice versa, thus concluding that students ratio for marks scored in online as well as theory are somewhat equal.

6.1.10 Principles of Programming Languages (PPL):

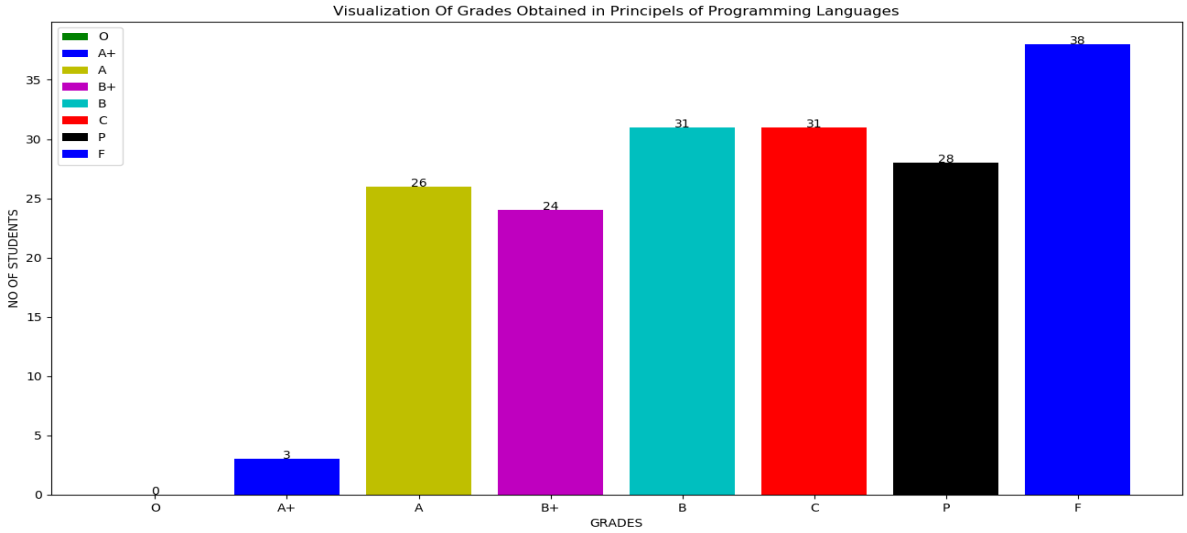


Fig. 6.1.10 Visualization of grades for the subject PPL

In the above given graph we can clearly see

1. The failure percentage is the most (i.e.38 students couldn't clear this subject)
2. Students securing A grade are 34 which is the second highest.
3. Only 3 students have secured an A+ grade(i.e. above 90)
4. No one could secure an O grade
5. The number of students securing B and C grades are nearly equal(i.e. 31,32 respectively)
6. The number of students securing A and B+ grades are nearly equal(i.e.26,24 respectively)

Online Vs Theory marks comparison

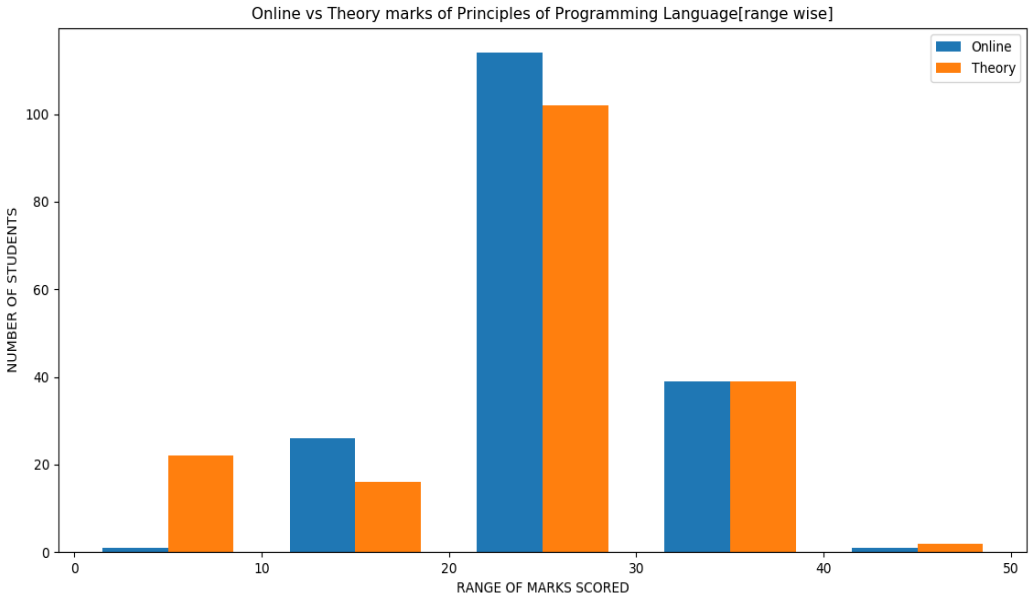


Fig. 6.2.10 Online Vs Theory marks for PPL

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are the least for both online and theory. For the range 30-40 online as well as theory marks are equal.

6.4 Pie Chart

Now we are going to print a pie-chart for semester wise comparison of performance

1. **plt.pie()** is used to draw pie-chart, which shows how much amount of (percentage of) a particular thing is capturing amongst all others.

2. The parameters that are passed to it are list of values of stake holders of the pie-chart(i.e. similar kind of things those who are to be compared against one another), **labels=** list of labels to be passed to the pie-chart.

3. Another parameters are **startangle =** angle with which you want to start drawing first pie, **colors =** list of colors, **autopct = '%2.2%%'** is used to calculate percentage of area captured of by a particular thing amongst all up to 2 floating digits.

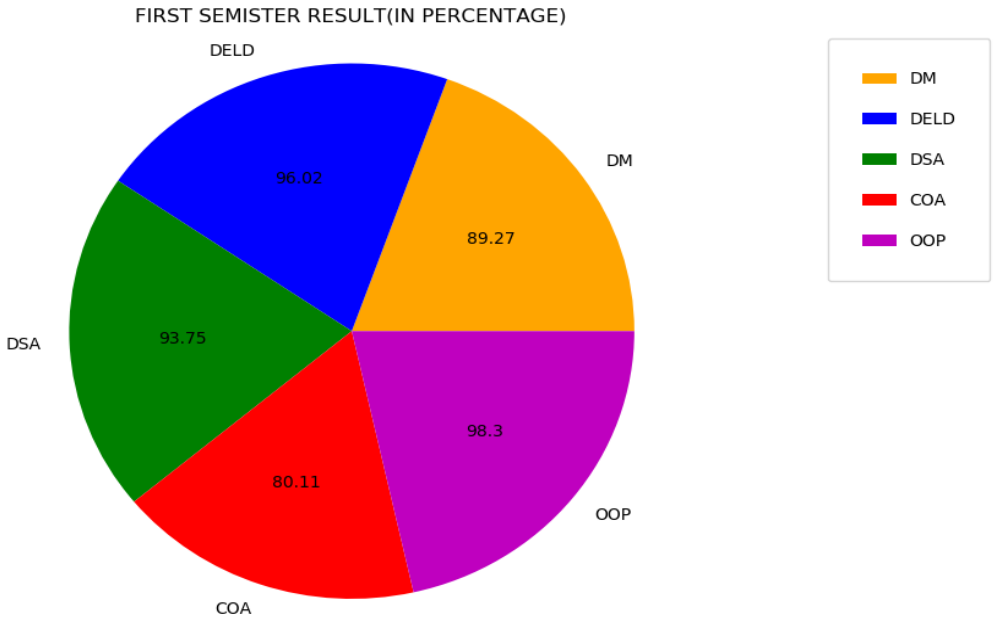
4. **explode=(0,0.1,0,0.2)** is used to highlight one particular pie by exploding it out of the pie-chart as per given factor.

Note:
All the lists and parameters passed to plt.pie() must contains same number of records in it, then only the pie-chart will be drawn.

Code:

```
# Data to plot
labels = 'DM', 'DELD', 'DSA', 'COA', 'OOP'
sizes = np.array(output)
colors = ['orange', 'blue', 'green','r', 'm']
#explode = (0, 0, 0, 0.1,0)  # explode 1st slice
total=sum(sizes)
# Plot
plt.pie(sizes,labels=labels,
colors=colors,autopct=absolute_value2)
```

6.4.1 First Semester Result



0Fig. 6.4.1 First Semester Result

From the above pie chart we can see that DM having 89.27% result, DELD having 96.02% result, DSA having 93.75% result, COA having 80.11% result and OOP having 98.30% result. OOP has the highest result than all other subjects and COA has the lowest result.

6.4.2 Second Semester Result

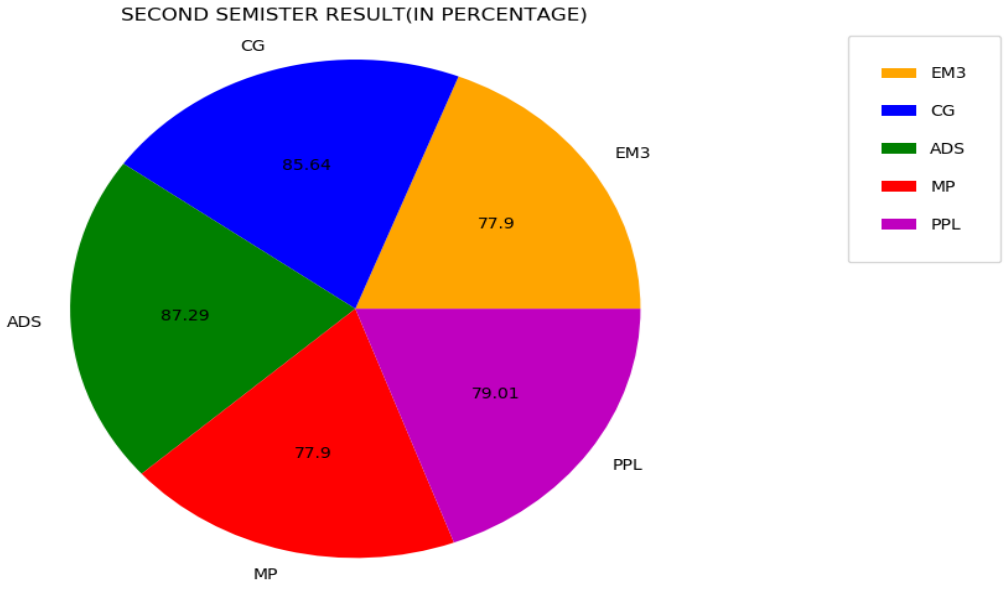


Fig. 6.4.2 Second Semester Result

From the above pie chart we can see that EM-III having 77.9% result, CG having 85.64% result, ADS having 87.29% result, MP having 77.9% result and PPL having 79.01% result. ADS has the highest result than all other subjects and EM3 and MP both having the same result i.e. the lowest result.

6.3 MIN, MAX, FAILED

This code is developed for calculating maximum marks, minimum marks scored and the failed result percentage for each subject semester wise.

Code:

```
for rect in barsticks:
      height=rect.get_height()
      plt.text(rect.get_x()+rect.get_width()/2,heigh,'%d'%
int(height),ha='center')
barsticks=plt.bar(list3+0.27,min1,align='center',alpha=1,widt
h=0.27)
for rect in barsticks:
      height=rect.get_height()
      plt.text(rect.get_x()+rect.get_width()/2,heigh,'%d'%
int(height),ha='center')
barsticks=plt.bar(list3+0.54,fail,align='center',alpha=1,widt
h=0.27)
for rect in barsticks:
      height=rect.get_height()
      plt.text(rect.get_x()+rect.get_width()/2,height,'%d'%
int(height),ha='center')

plt.legend(legends)
plt.xticks(list3,subjects,ha='left')
plt.tight_layout()
plt.title('Max marks, Min marks and Failed Percentage of
students(Div-B)')
plt.xlabel('Subjects')
plt.ylabel('Marks in percentage(APPROX)')
plt.show()
```

6.3.1 First semester:

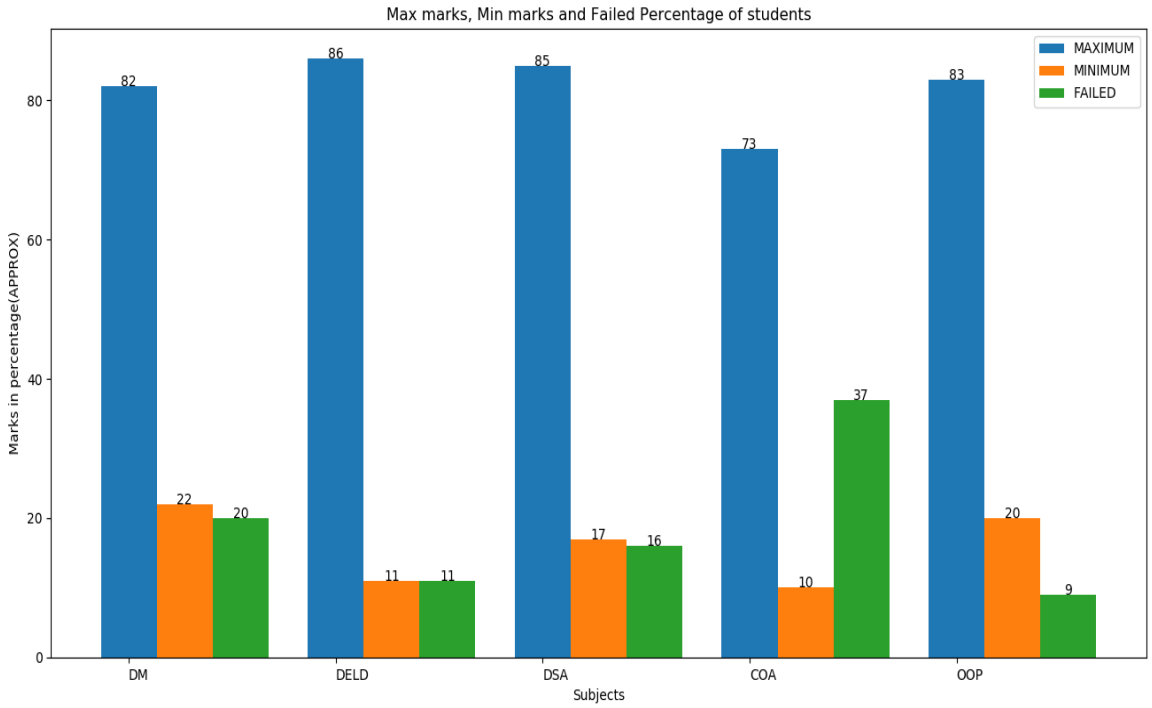


Fig. 6.3.1 Max, Min, Failed Percentage of students

From the above graph we come to know that,

1] Maximum and minimum marks scored in DM are 82 and 22 respectively while 20 students could not clear this subject.
2] Maximum and minimum marks scored in DELD are 86 and 12 respectively while 11 students could not clear this subject.
3] Maximum and minimum marks scored in DSA are 85 and 17 respectively while 11 students could not clear this subject.
4] Maximum and minimum marks scored in COA are 73 and 10 respectively while 17 students could not clear this subject.
5] Maximum and minimum marks scored in OOP are 83 and 20 respectively while 9 students could not clear this subject.

6.3.2Second semester:

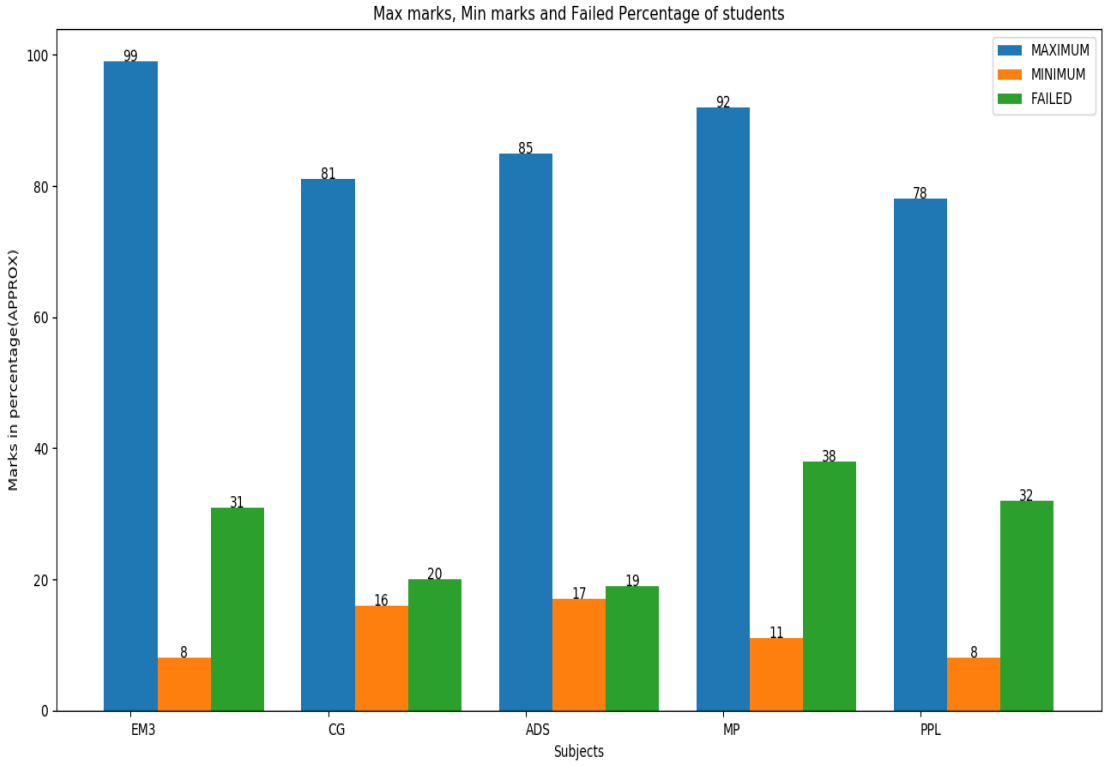


Fig. 6.3.2 Max, Min, Failed Percentage of students

From the above graph we come to know that,

1] Maximum and minimum marks scored in EM3 are 99 and 8 respectively while 31 students could not clear this subject.
2] Maximum and minimum marks scored in CG are 81 and 16 respectively while 20 students could not clear this subject.
3] Maximum and minimum marks scored in ADS are 85 and 17 respectively while 29 students could not clear this subject.
4] Maximum and minimum marks scored in MP are 90 and 11 respectively while 38 students could not clear this subject.
5] Maximum and minimum marks scored in PPL are 78 and 8 respectively while 32 students could not clear this subject.

DIVISION-WISE RESULT
(DIV A)

6.1.1a Discrete Mathematics (DM):

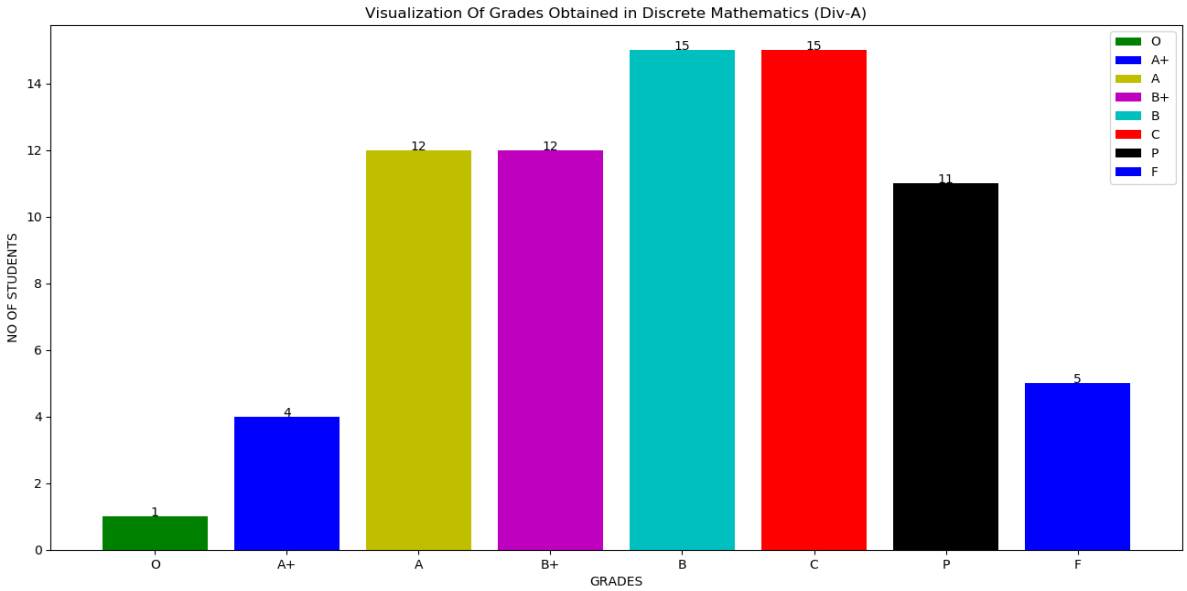


Fig. 6.1.1a Visualization of grades for the subject DM (div A)

From the graph we can clearly see that only 2 students could secure an O grade, whereas 4 students have secured A+ grade. The number of students securing A and B+ grades are equal(i.e.12 each),whereas the number of students securing B and C grade are equal(i.e.15 each). 5 students couldn't clear this subject in the second attempt from div A.

Online Vs Theory marks comparison

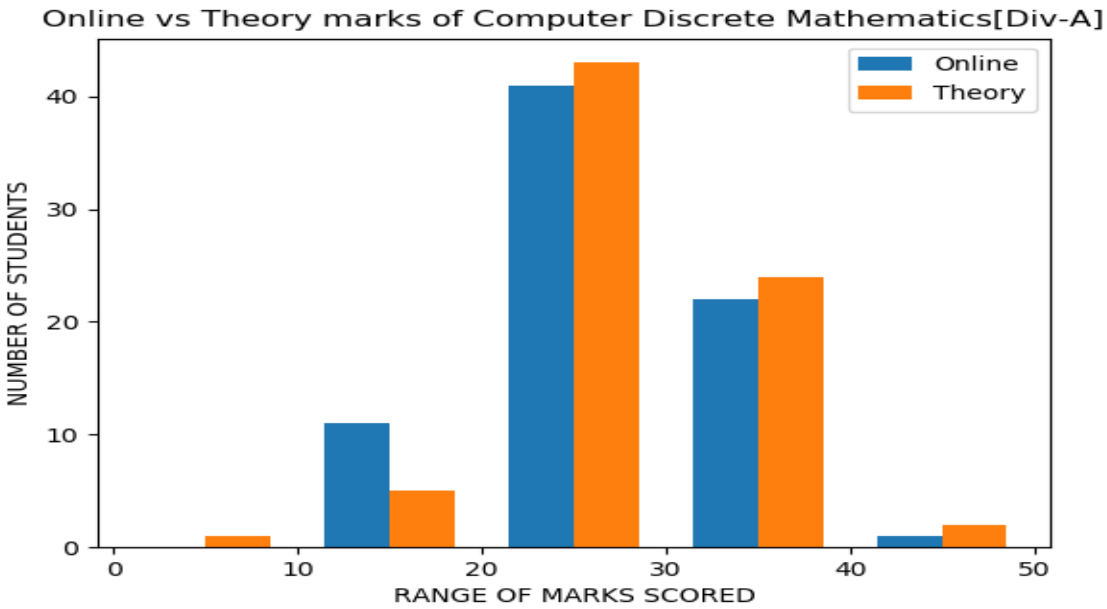


Fig. 6.2.1a Online Vs Theory marks for DM

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. No one has got less than 10 marks in online . Students securing marks in the range of 0-10 are the least for theory.

6.1.1b Digital Electronics and Logic Design (DELD):

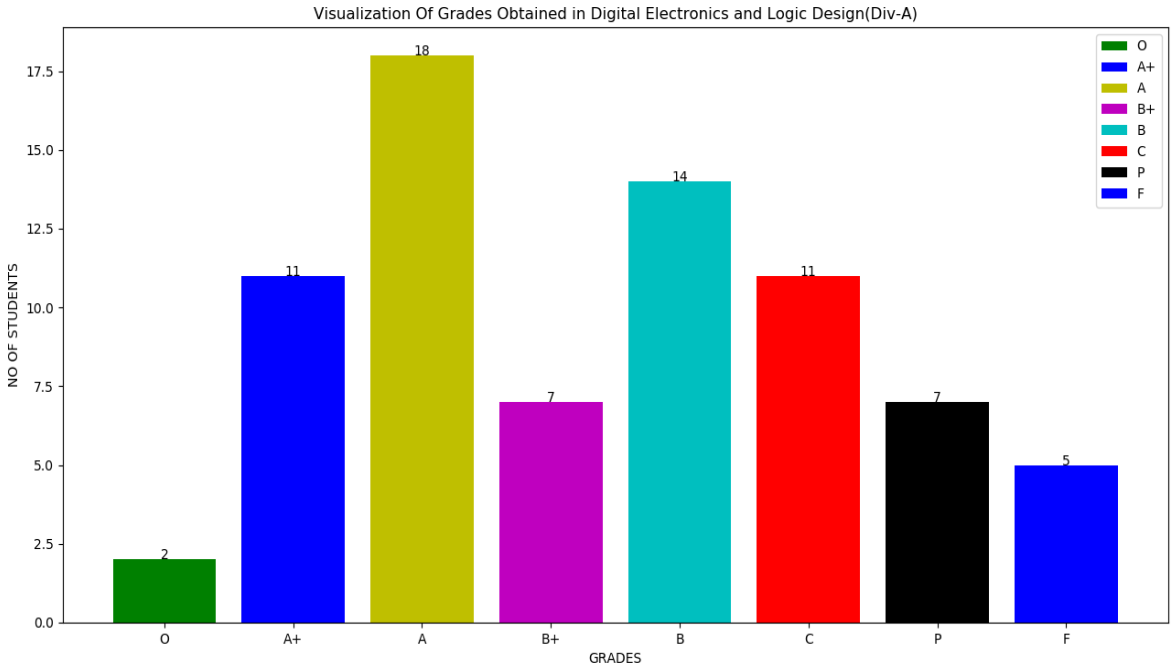


Fig. 6.1.1b Visualization of grades for the subject DELD (div A)

From the graph we can clearly see that, only 2 students could secure an O grade, whereas 11 students have secured A+ grade..5 students couldn't clear this subject in the second attempt from div A. Students securing A grade are 18 which is the highest. The number of students securing B+ and P grade are equal (i.e.7 each).

Online Vs Theory marks comparison

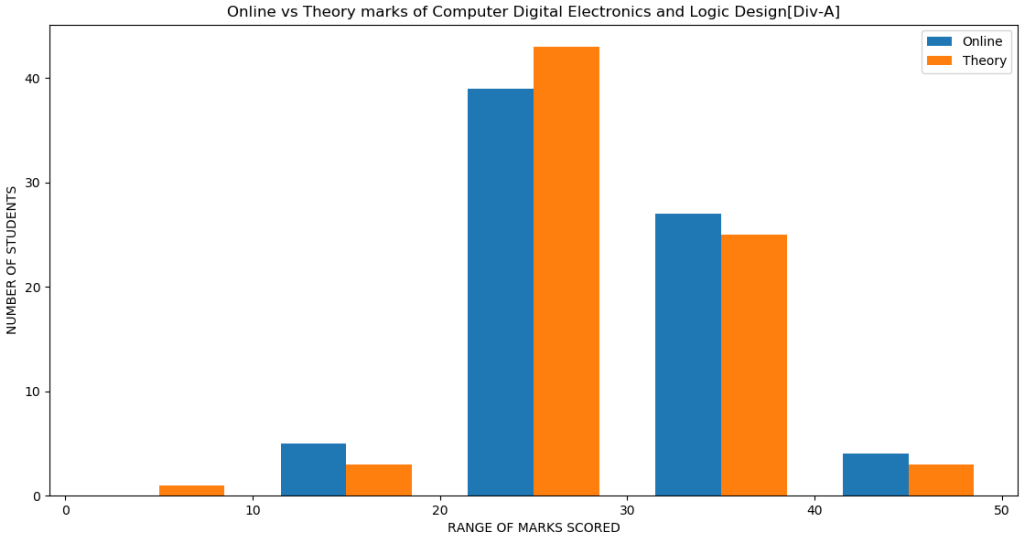


Fig. 6.2.1b Online Vs Theory marks for DELD

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Here, in some cases ranges the online marks beats the theory marks as well as vice versa. For the range 20-30 theory marks scored are more whereas for the range 30-40 online marks scored exceeds to that scored in theory. No one has got less than 10 marks in online. The percentage of students in the range 10-20 and 40-50 are nearly equal.

6.1.1c Data Structure and Algorithms (DSA):

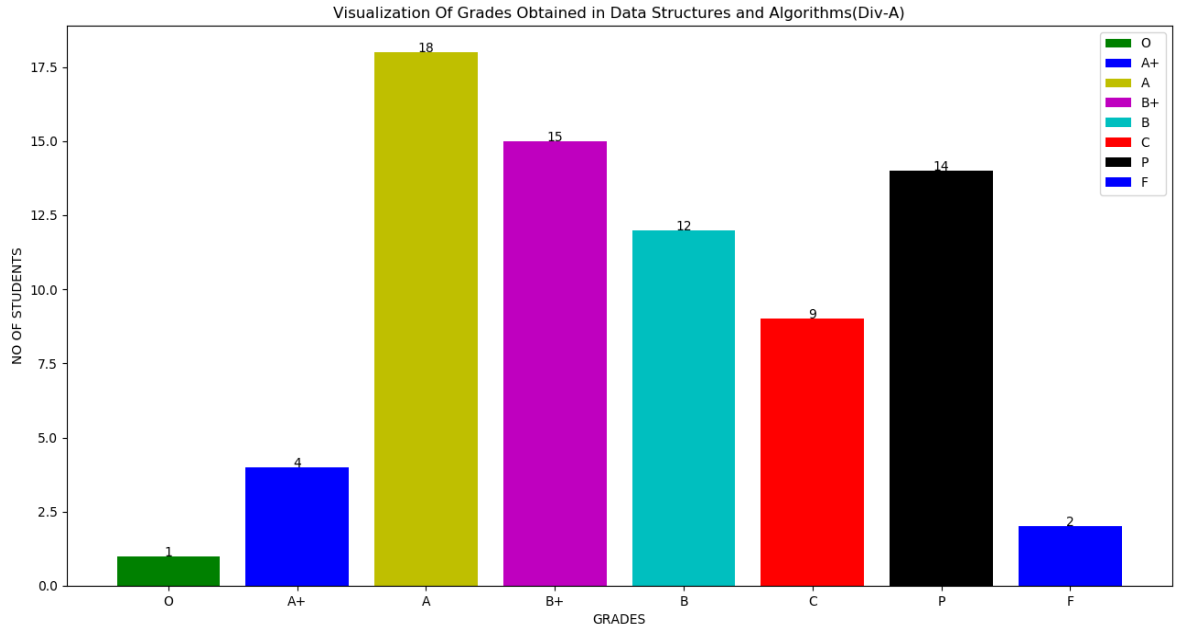


Fig. 6.1.1c Visualization of grades for the subject DSA (div A)

From the graph we can clearly see that, only 2 students could secure an O grade, whereas 4 students have secured A+ grade.2 students couldn't clear this subject in the second attempt from div A. Students securing A grade are 18 which is the highest. The number of students securing B+ and P grade are nearly equal.

Online Vs Theory marks comparison

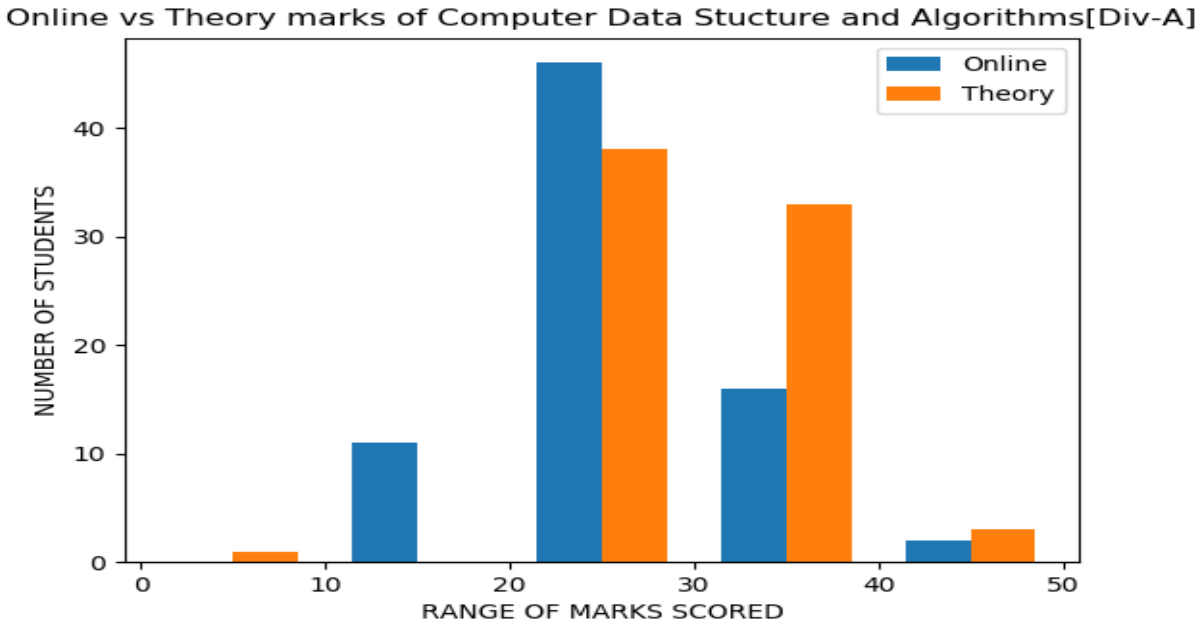


Fig. 6.2.1c Online Vs Theory marks for DSA

In the above graph we can clearly see that, maximum number of students have secured marks in the range of 20-30 for both online as well as theory. No student has got marks in the range of 0-10 for online and 10-20 in theory. Also 0-10 has the least number of students scoring theory marks and 40-50 has the least number of students scoring online marks.

6.1.1d Computer Organization and Architecture (COA):

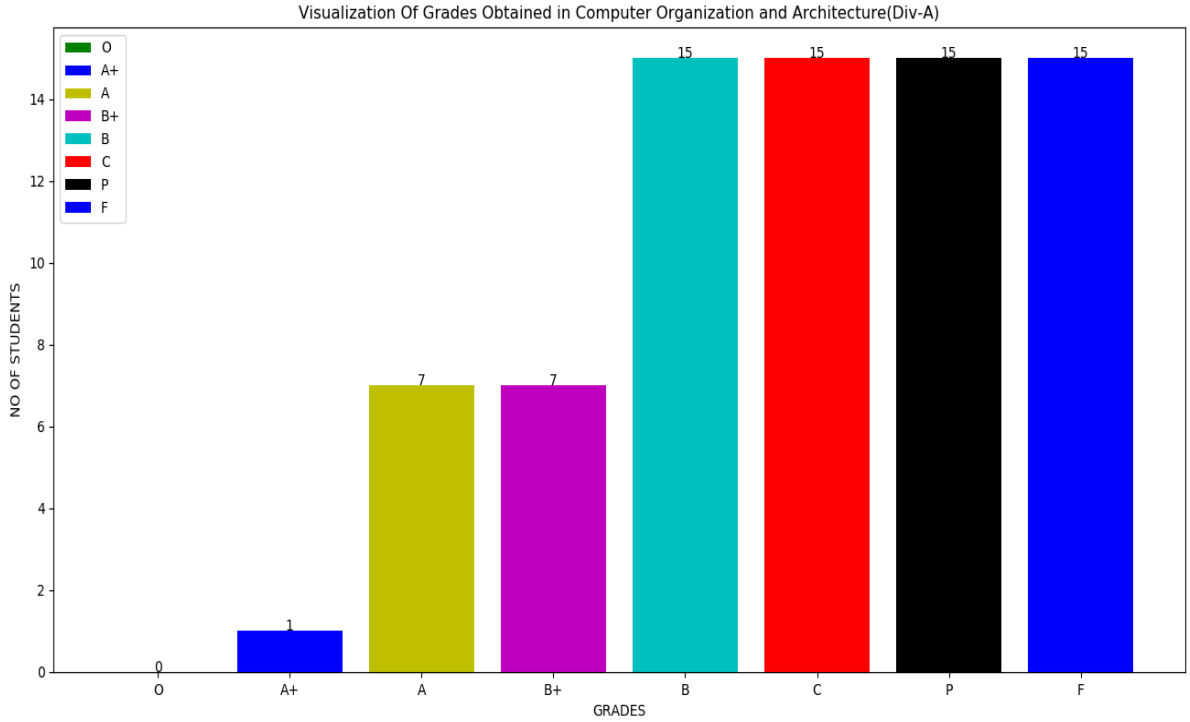


Fig. 6.1.1d Visualization of grades for the subject COA (div A)

From the graph we can clearly see that, no one could secure an O grade, whereas only 1 student have secured A+ grade. The number of students securing B+ and A grade are equal(i.e.7 each), whereas remaining grades each have same number of students securing them(i.e.15 each).

Online Vs Theory marks comparison

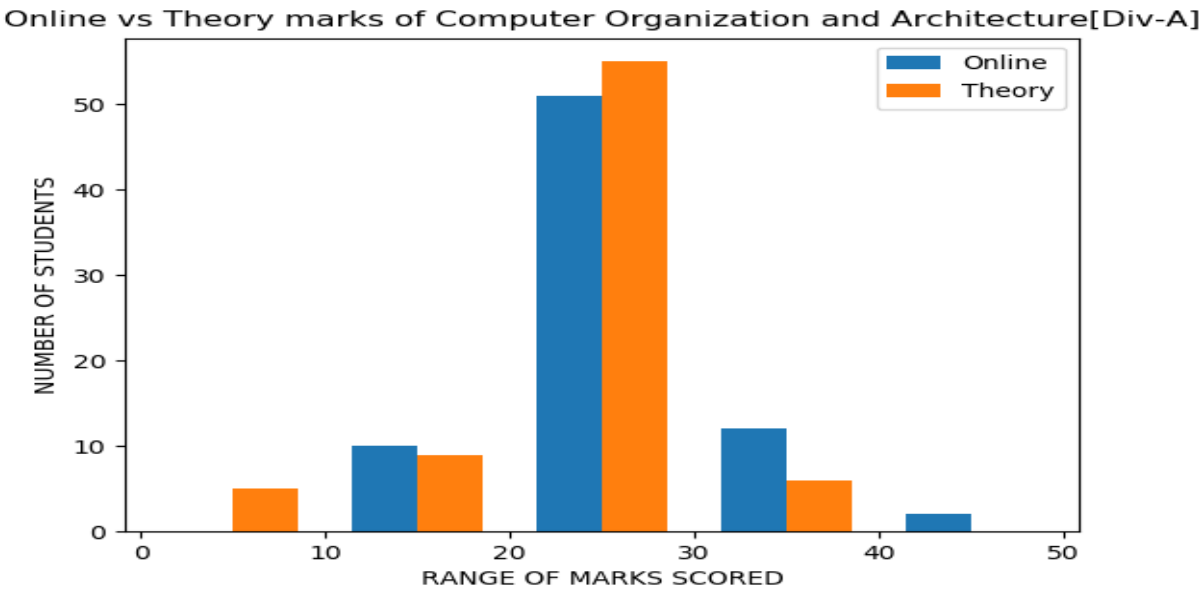


Fig. 6.2.1d Online Vs Theory marks for COA

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 and 0-10 are none for theory and online respectively. Here, in some cases ranges the online marks beats the theory marks as well as vice versa.

6.1.1e Object Oriented Programming (OOP):

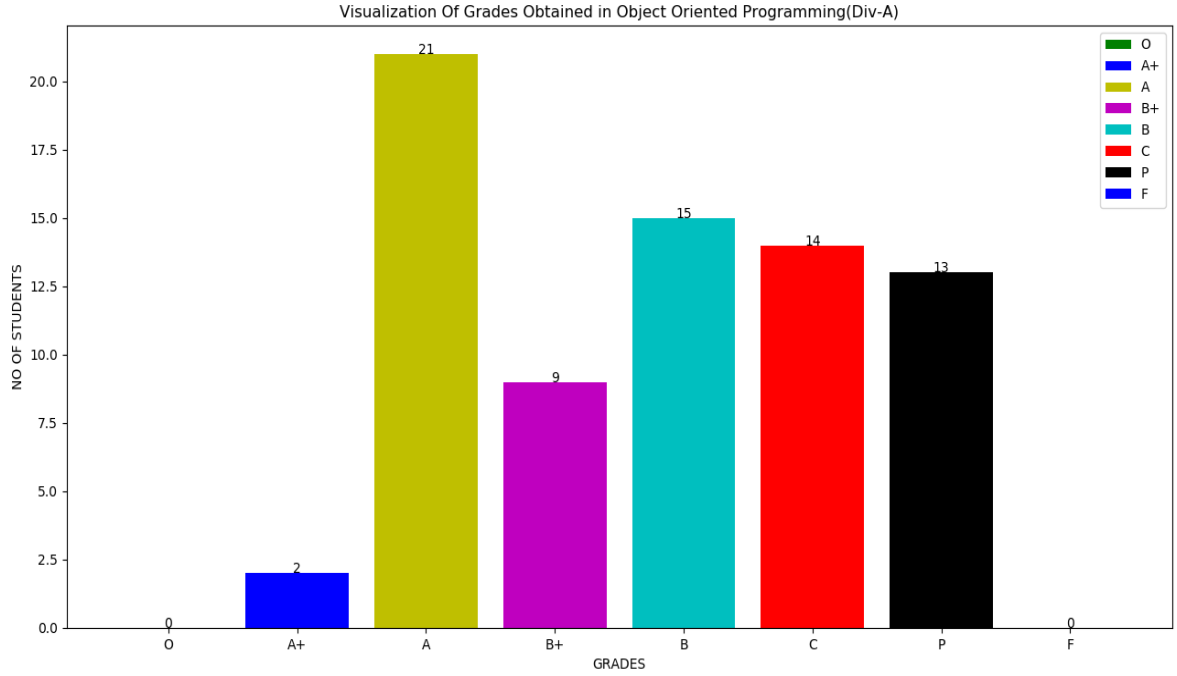


Fig. 6.1.1e Visualization of grades for the subject OOP(div A)

From the graph we can clearly see that, no one could secure an O grade, whereas only 2 students have secured A+ grade. Students securing A grade are 18 which is the highest. The best part is all students from div A have cleared this subject . Students securing A grade are 21 which is the highest.

Online Vs Theory marks comparison

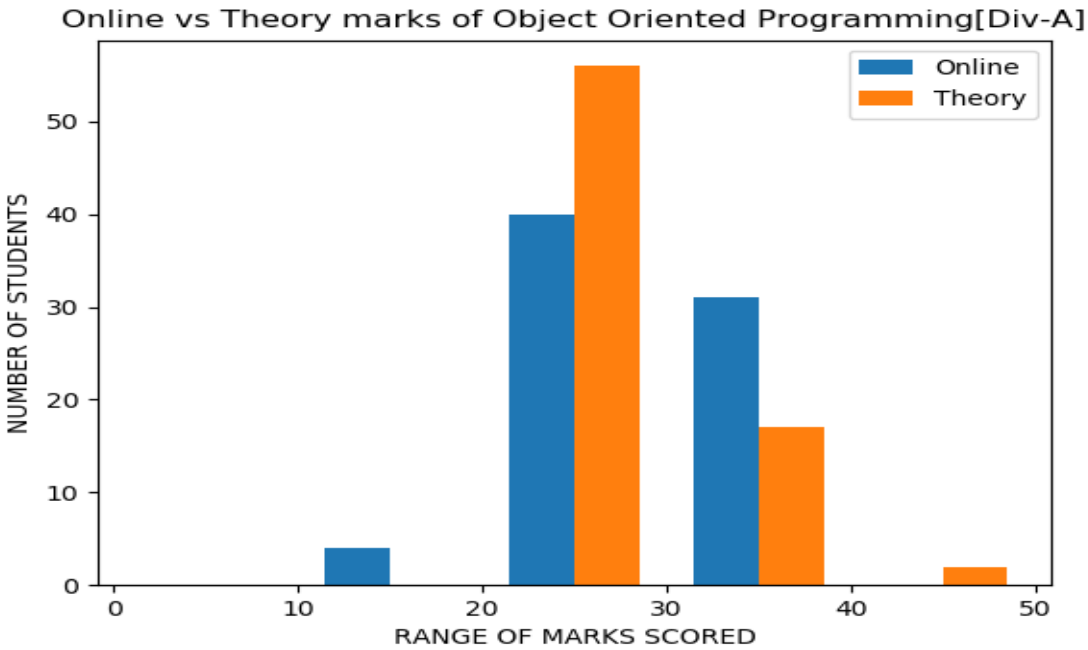


Fig. 6.2.1e Online Vs Theory marks for OOP

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. No student has got marks in the range of 40-15 for online and 10-20 in theory. To add to it no student has scored marks in the range of 0-10.

6.1.1f Engineering Mathematics III (EM-III):

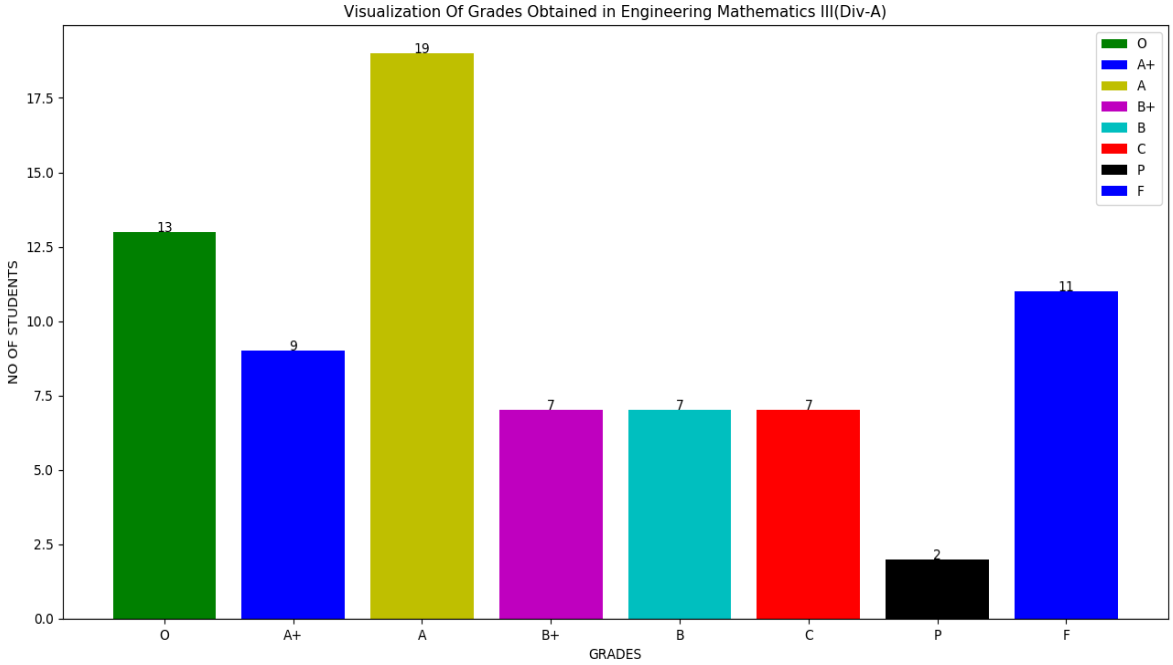


Fig. 6.1.1f Visualization of grades for the subject EM3 (div A)

From the graph we can clearly see that this subject has the highest passing rate, whereas only 2 students have secured P grade. Students securing A grade are 19 which is the highest.12 students couldn't clear this subject . Students securing B+,B, C grades are same.

Online Vs Theory marks comparison

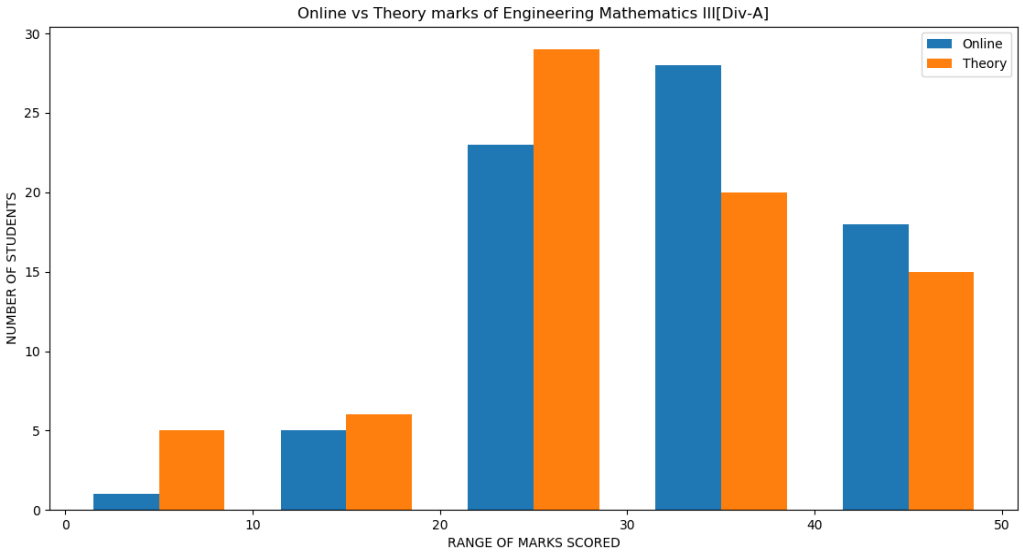


Fig. 6.2.1f Online Vs Theory marks for M3

In the above graph, for the range 20-30 and 30-40 firstly the theory marks exceeds the online marks whereas it is the opposite for the next range concluding that both the ranges equally balance each other. Students securing marks in the range of 0-10 are the least for both online as well as theory. Here ranges of 20-30, 30-40, 40-50 have greater ratio of students amongst them compared to first two ranges.

6.1.1g Computer Graphics (CG):

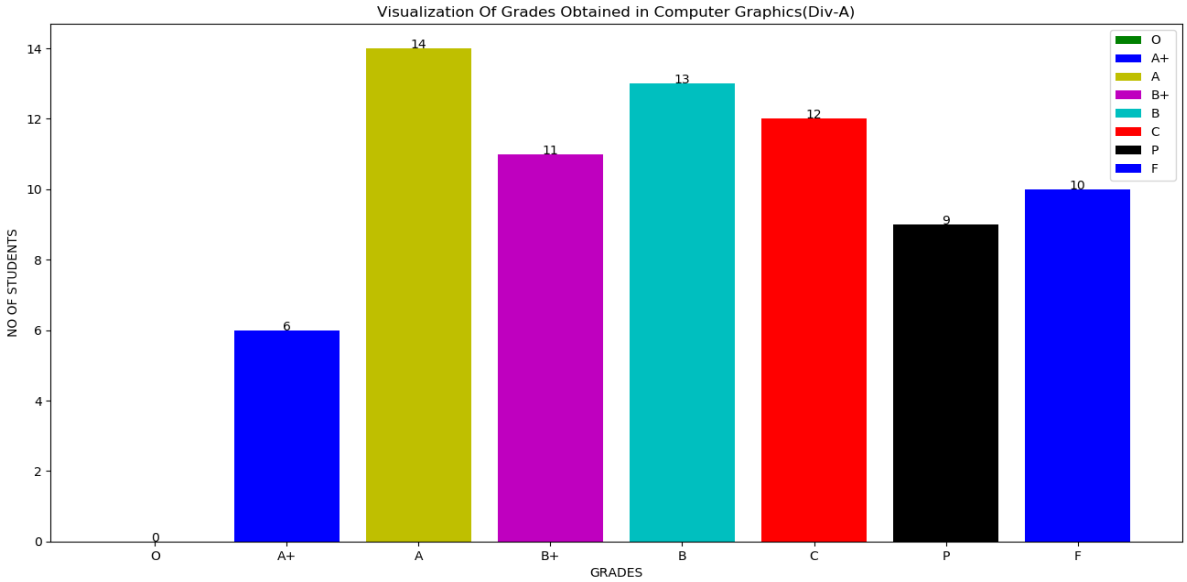


Fig. 6.1.1g Visualization of grades for the subject CG (div A)

From the graph we can clearly see that, no one could secure an O grade, whereas 6 students have secured A+ grade. 10 students couldn't clear this subject. Students securing A grade are 14 which is the highest. secure an O grade, whereas 4 students have secured A+ grade.2 students couldn't clear this subject in the second attempt from div A. Students securing A grade are 18 which is the highest. The number of students securing A,B+ ,B and C grade are in same range(14,11,13,12 respectively).

Online Vs Theory marks comparison

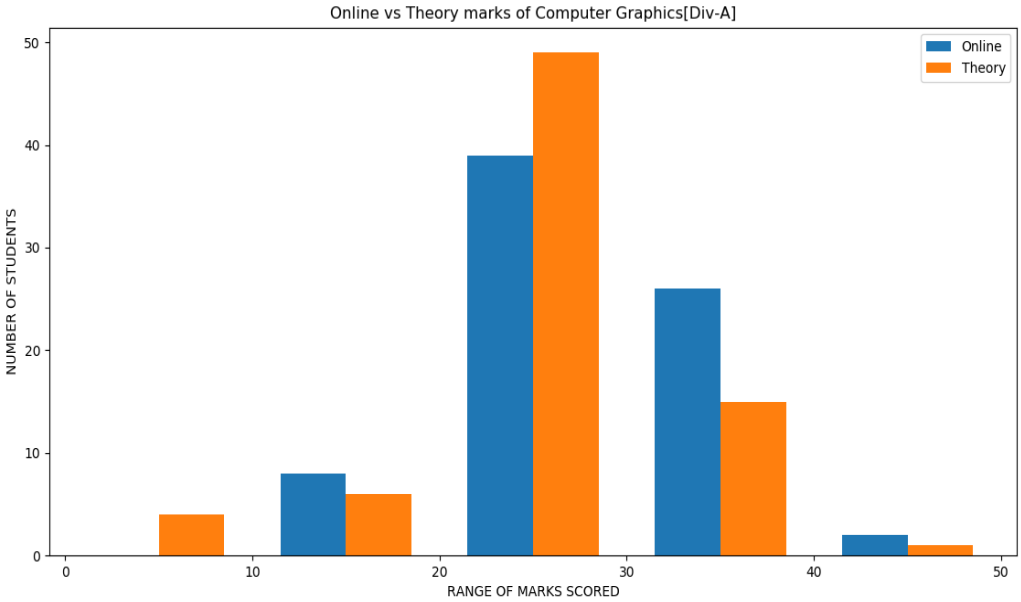


Fig. 6.2.1g Online Vs Theory marks for CG

In the above graph, for the range 20-30 and 30-40 firstly the theory marks exceeds the online marks whereas it is the opposite for the next range concluding that both the ranges equally balance each other. Students securing marks in the range of 0-10 are the least for online whereas least in the range of 40-50 for theory.

6.1.1h Advanced Data Structure (ADS):

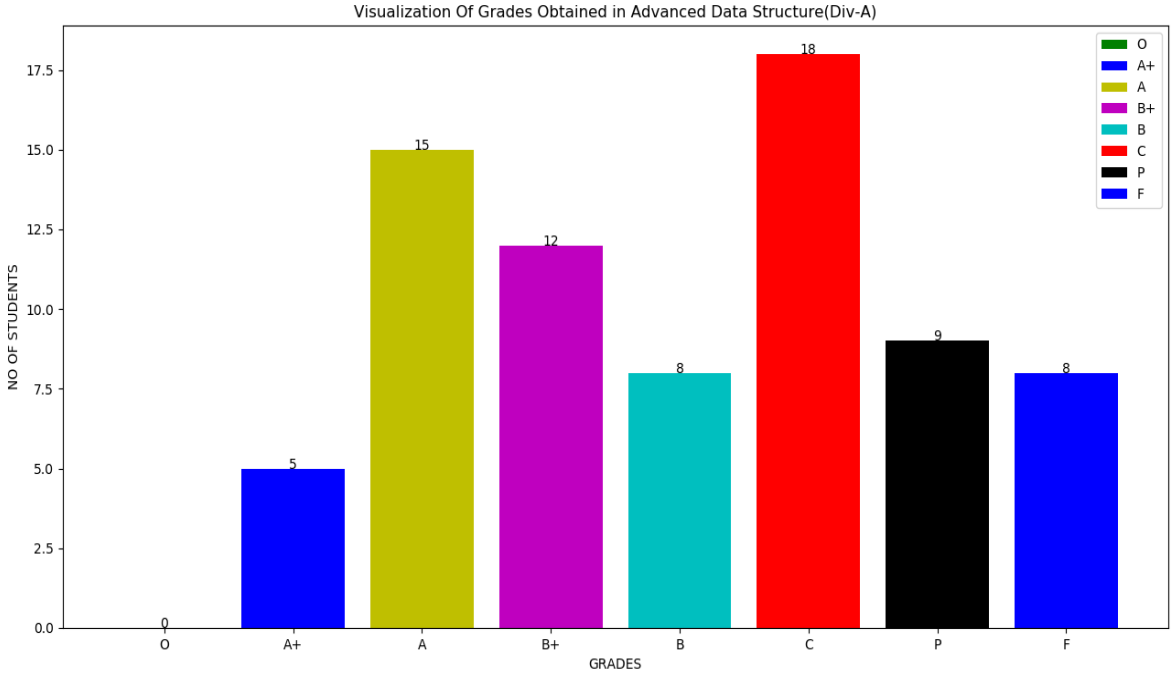


Fig. 6.1.1h Visualization of grades for the subject ADS (div A)

From the graph we can clearly see that, no one could secure an O grade, whereas 5 students have secured A+ grade. . Students securing C grade are 18 which is the highest. Students securing A grade are 15 which is the second highest.8 students couldn't clear this subject in the from div A.

Online Vs Theory marks comparison

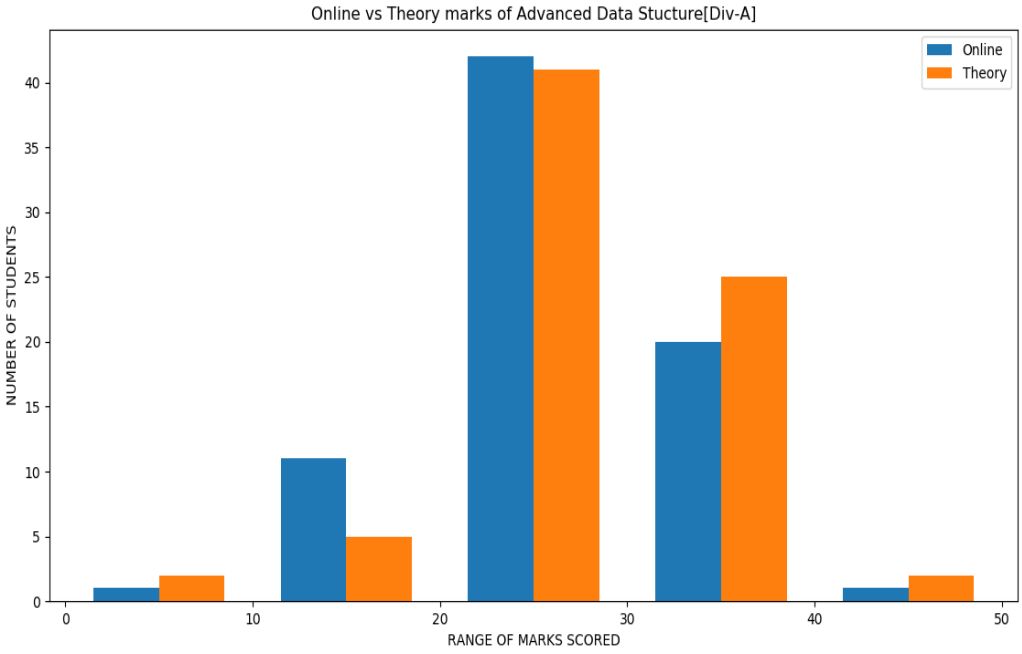


Fig. 6.2.1h Online Vs Theory marks for ADS

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 and 0-10 are the least for online and theory(i.e. both are same).

6.1.1i Microprocessor (MP):

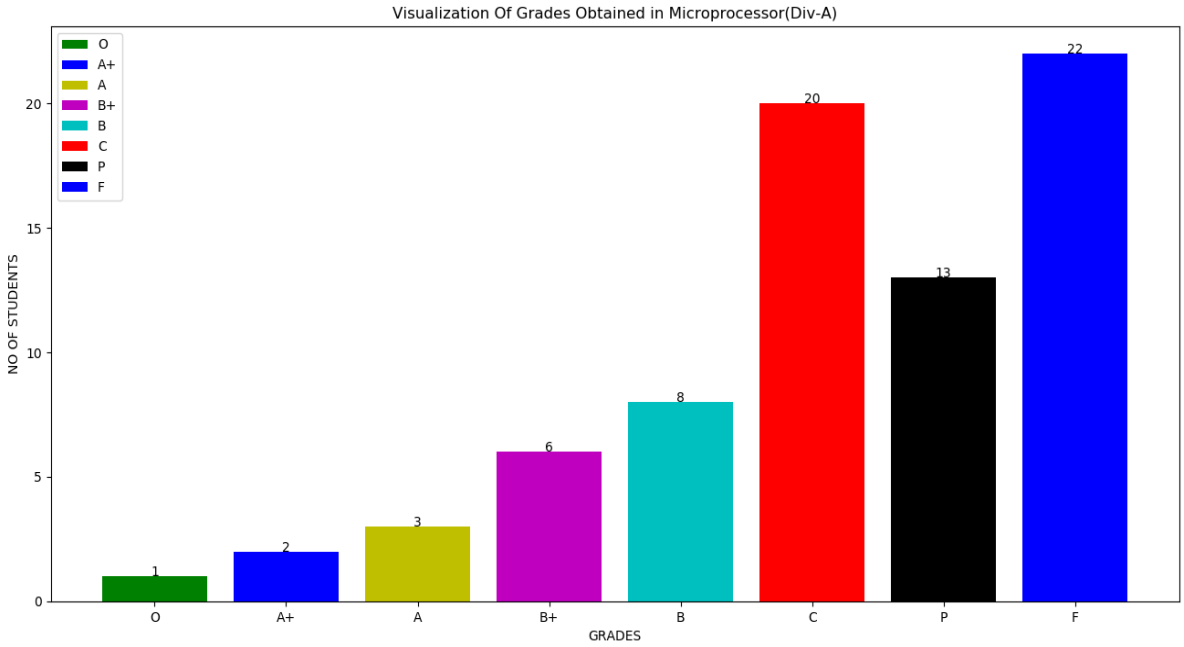


Fig. 6.1.1i Visualization of grades for the subject MP (div A)

From the graph we can clearly see that, only 1student could secure an O grade, whereas 2 students have secured A+ grade . Students securing C grade are 18 which is the second highest. Students securing A, B+ and B grade are 3 ,6 and 8 respectively.22 students couldn't clear this subject in the from div A.

Online Vs Theory marks comparison

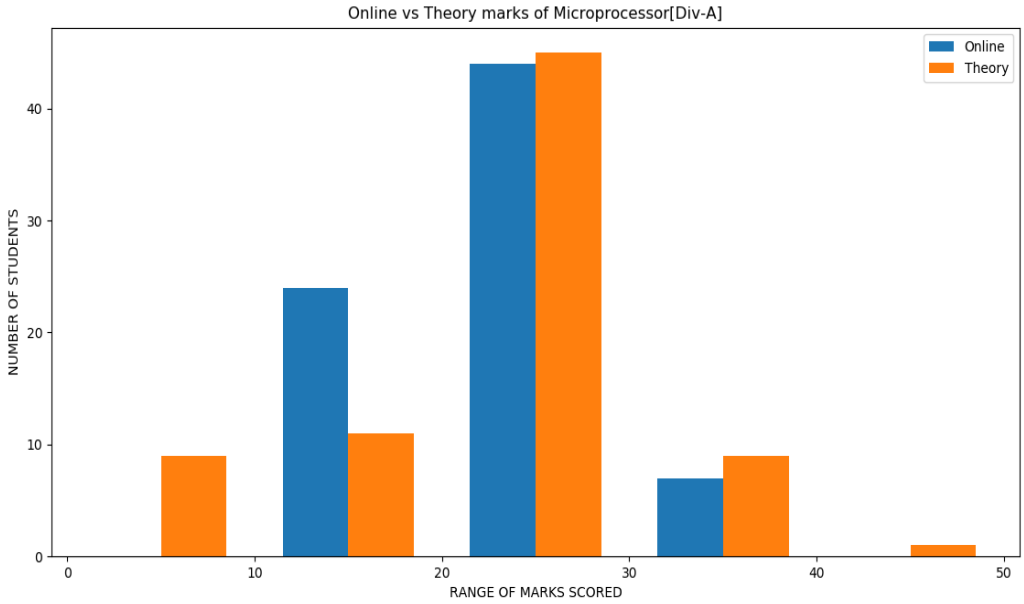


Fig. 6.2.1i Online Vs Theory marks for MP

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are the least for theory whereas for online the range 0-10 and 40-50 is applicable as there are no figures of marks .The ratio of online marks for 10-20 is a lot greater than its theory.

6.1.1j Principles of Programming Languages (PPL):

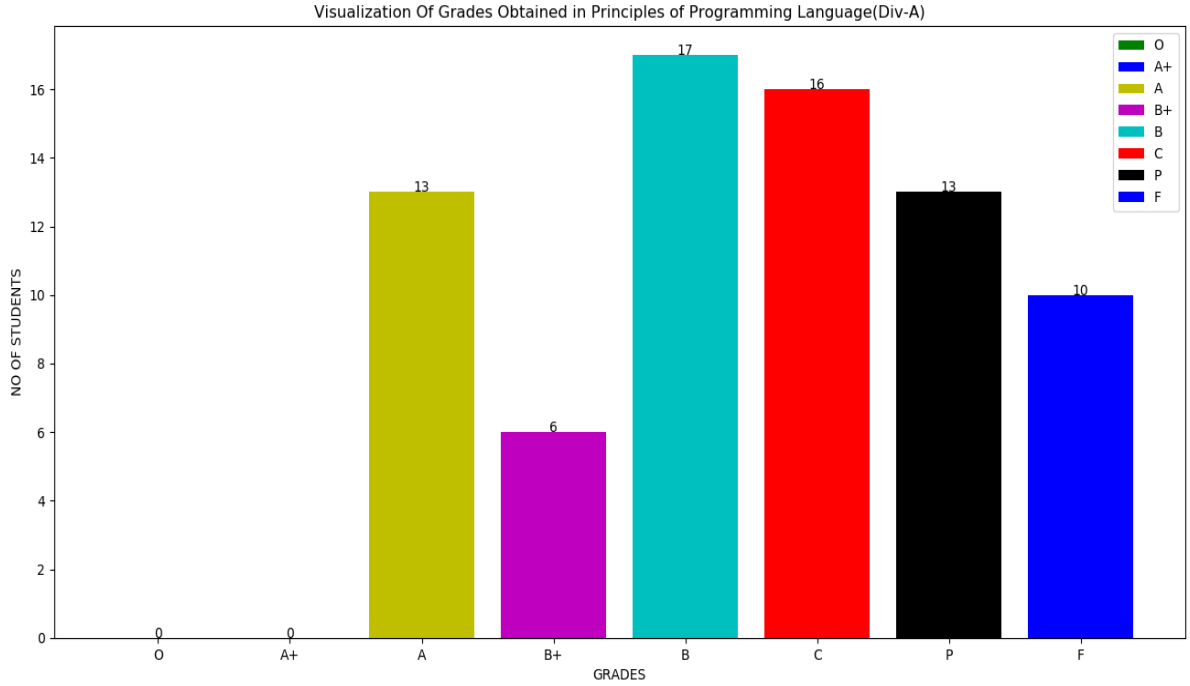


Fig. 6.1.1j Visualization of grades for the subject PPL (div A)

From the graph we can clearly see that, no one could secure an O as well as A+ grade, whereas only 6 students have secured B+ grade. The number of students securing B and C grade are nearly equal(i.e.17 and 16 respectively),10 students could not clear this subject.

Online Vs Theory marks comparison

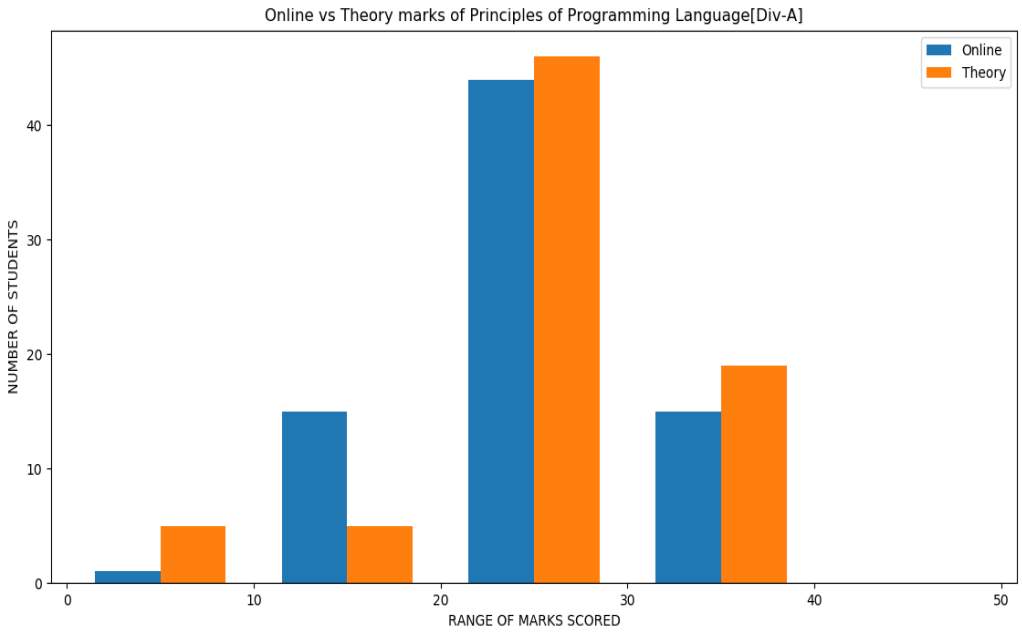


Fig. 6.2.1j Online Vs Theory marks for PPL

In the above graph we can clearly see that, maximum number of students have secured marks in the range of 20-30 for both online as well as theory. No student has secured marks in the range of 40-50 for online and theory. The ratio of marks scored in the range 0-10 is the least for both online and theory.

DIV B

6.1.2a Discrete Mathematics (DM):

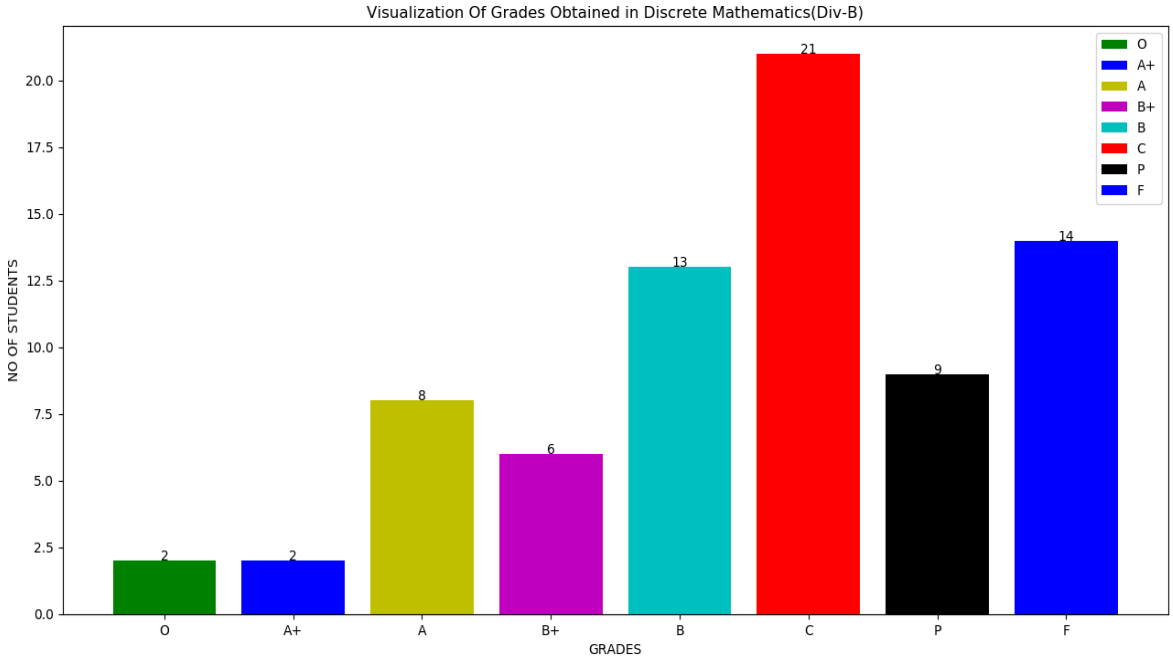


Fig. 6.1.2a Visualization of grades for the subject DM(div B)

From the graph we can clearly see that, only 2 students could secure an O as well as A+ grade, whereas only 8 students have secured A grade.21 students have got C grade which is the highest.14 students could not clear this subject.

Online Vs Theory marks comparison

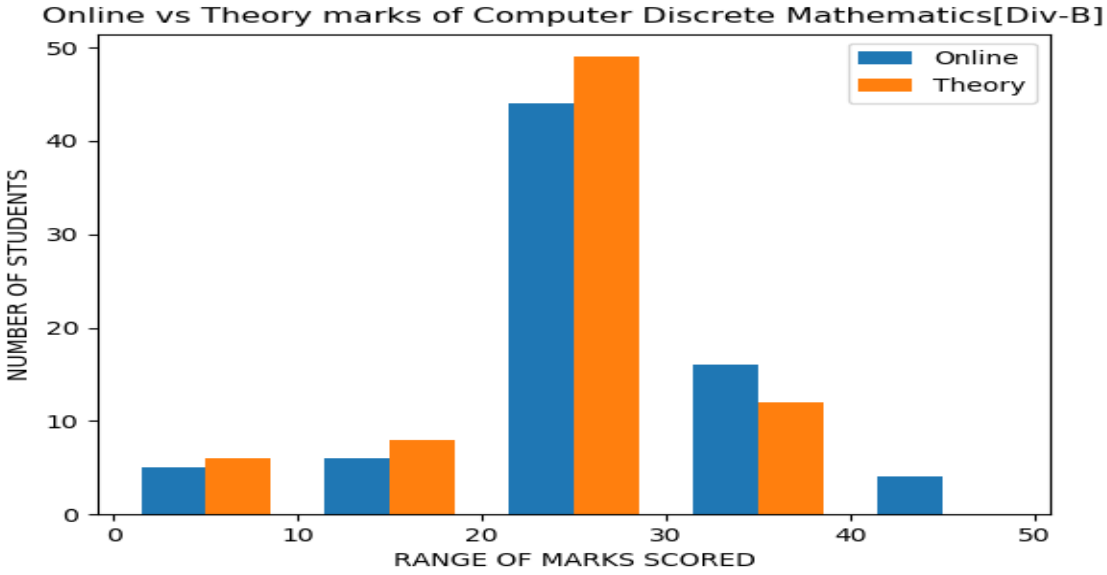


Fig. 6.2.2a Online Vs Theory marks for DM

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. This is contrasting to the graph of total DM marks where most of ranges the online marks beats the theory marks in case of scores. Students securing marks in the range of 40-50 are the least for both online as well as theory.

6.1.2b Digital Electronics and Logic Design (DELD):

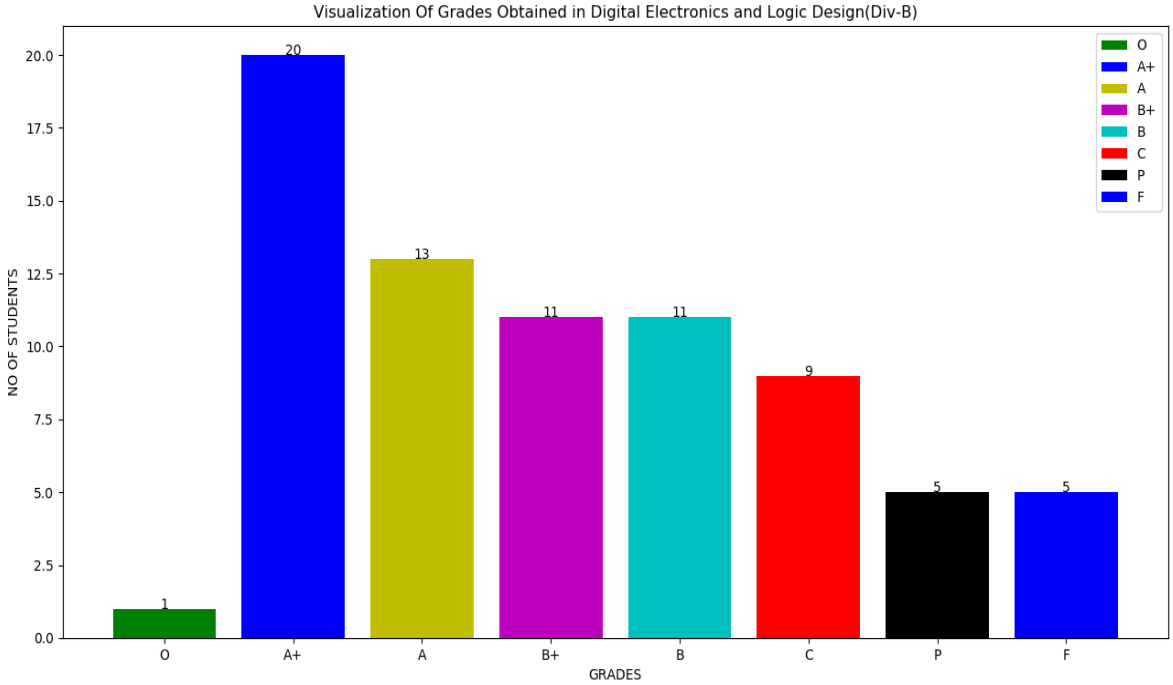


Fig. 6.1.2b Visualization of grades for the subject DELD(div B)

From the graph we can clearly see that, only 1 student could secure an O grade, whereas 20 students have secured A+ grade which is the highest.13 students have got A grade.5 students could not clear this subject in the second attempt. The number of students securing B and C grade are equal(i.e. 11 each).

Online Vs Theory marks comparison

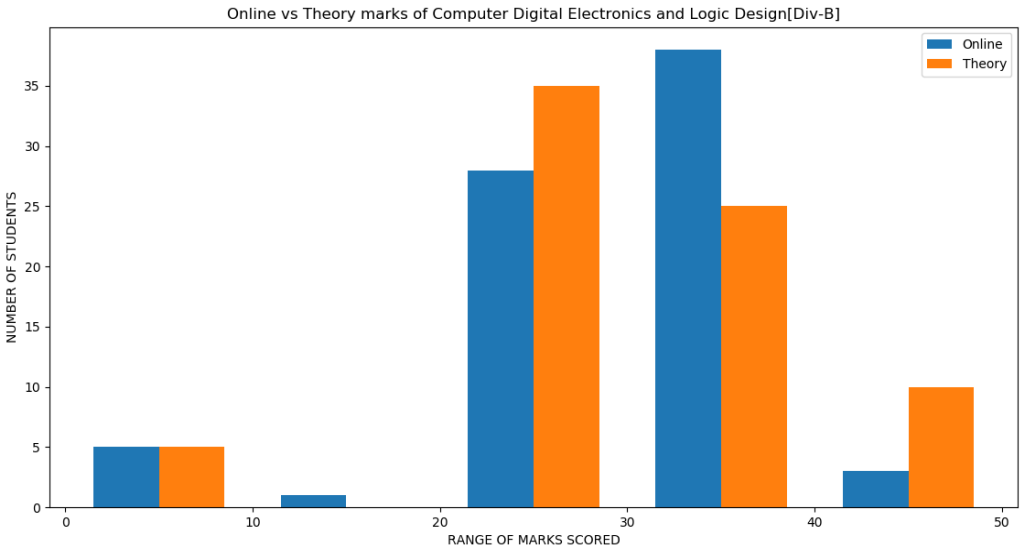


Fig. 6.2.2b Online Vs Theory marks for DELD

In the above graph we can clearly see that, most of students have secured marks in the range of 30-40 for both online as well as theory. This subject result of div B breaks the norm that has continuing until now (i.e. the most highest range marks is not in the range of 20-30 but in 30-40 for online). Maximum number of students are in range of 20-30 and 30-40. For the range 10-20 there are no entries of theory marks. Overall performance of this subject is better compared to others.

6.1.2c Data Structure and Algorithms (DSA):

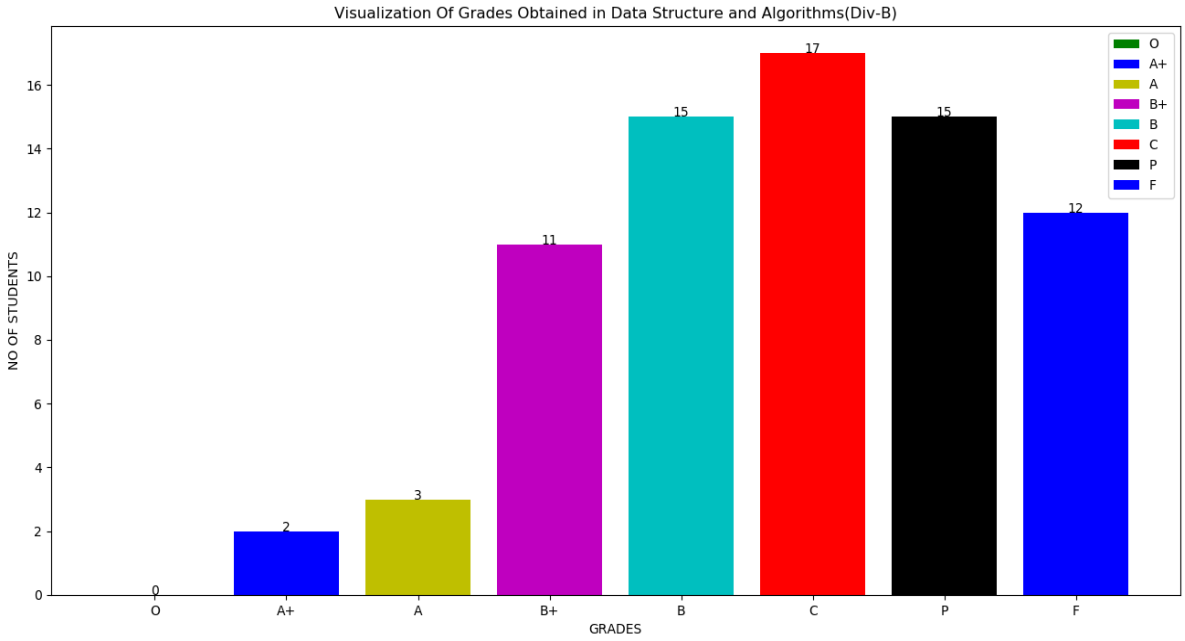


Fig. 6.1.2c Visualization of grades for the subject DSA (div B)

From the graph we can clearly see that, no one could secure an O grade, whereas 2, 3 students have secured A+ and A grade resp.11 students have got B+ grade.12 students could not clear this subject in the second attempt. The number of students securing B and P grade are equal(i.e.15 each). 17 students have got C grade which is the highest.

Online Vs Theory marks comparison

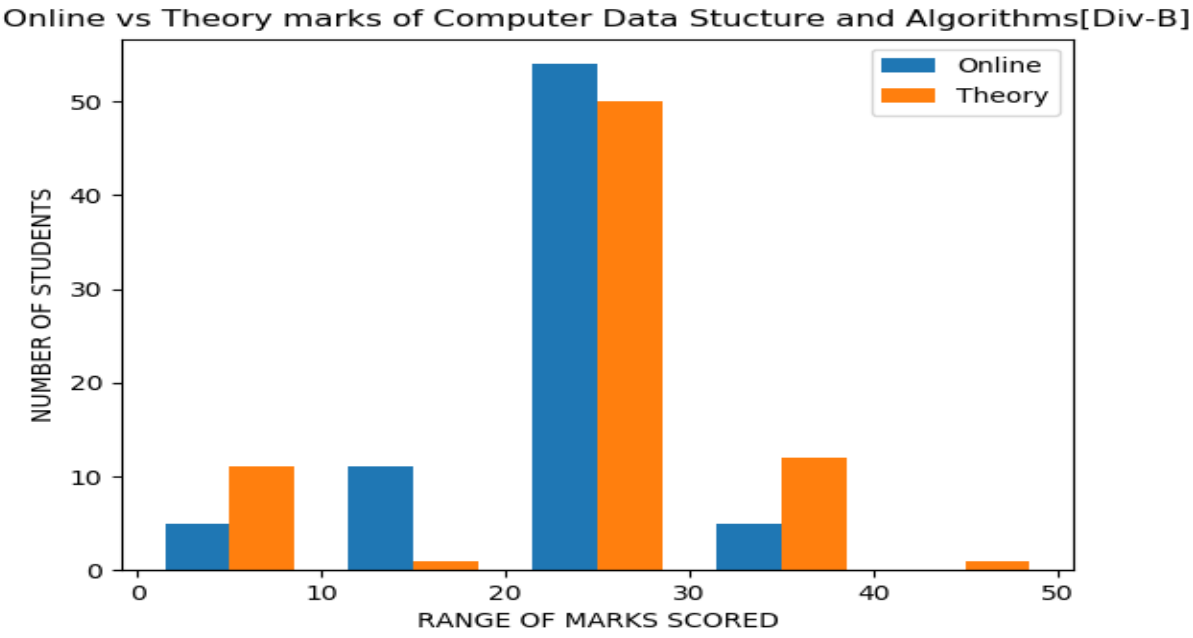


Fig. 6.2.2c Online Vs Theory marks for DSA

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are the least for online whereas for the theory the range 20-30 has the least score. Here more than half number of students are in the range of 20-30.

6.1.2d Computer Organization and Architecture (COA):

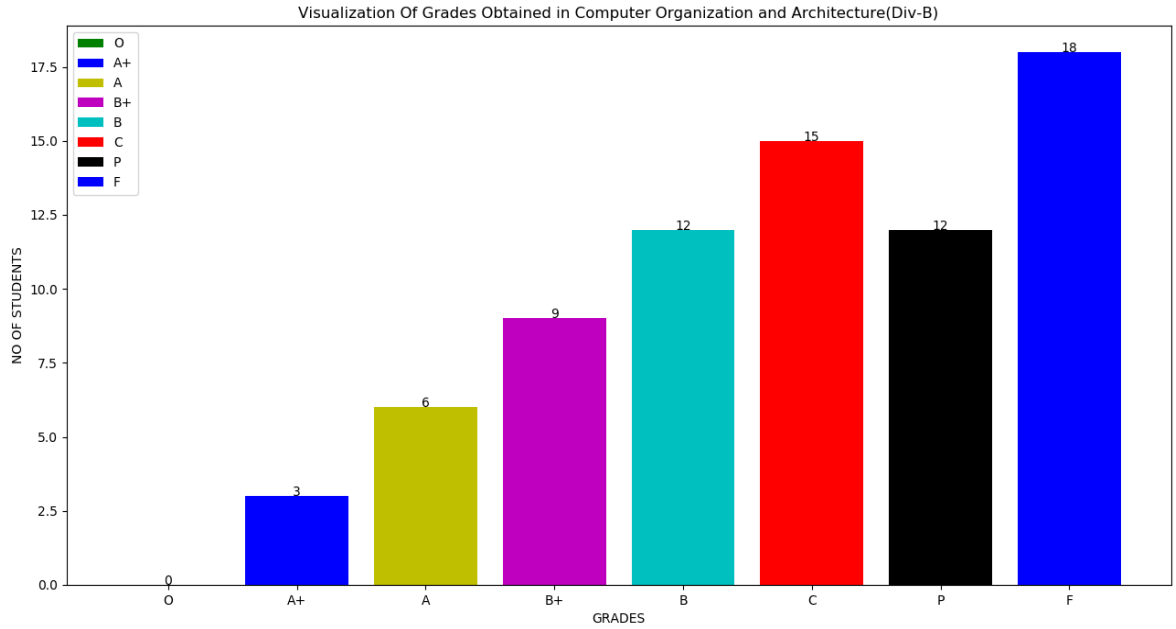


Fig. 6.1.2d Visualization of grades for the subject COA(div B)

From the graph we can clearly see that, no one could secure an O grade, whereas 3 students have secured A+ grade.9 students have got B+ grade.18 students could not clear this subject in the second attempt which is the highest.15 students have got C grade which is the second highest.

Online Vs Theory marks comparison

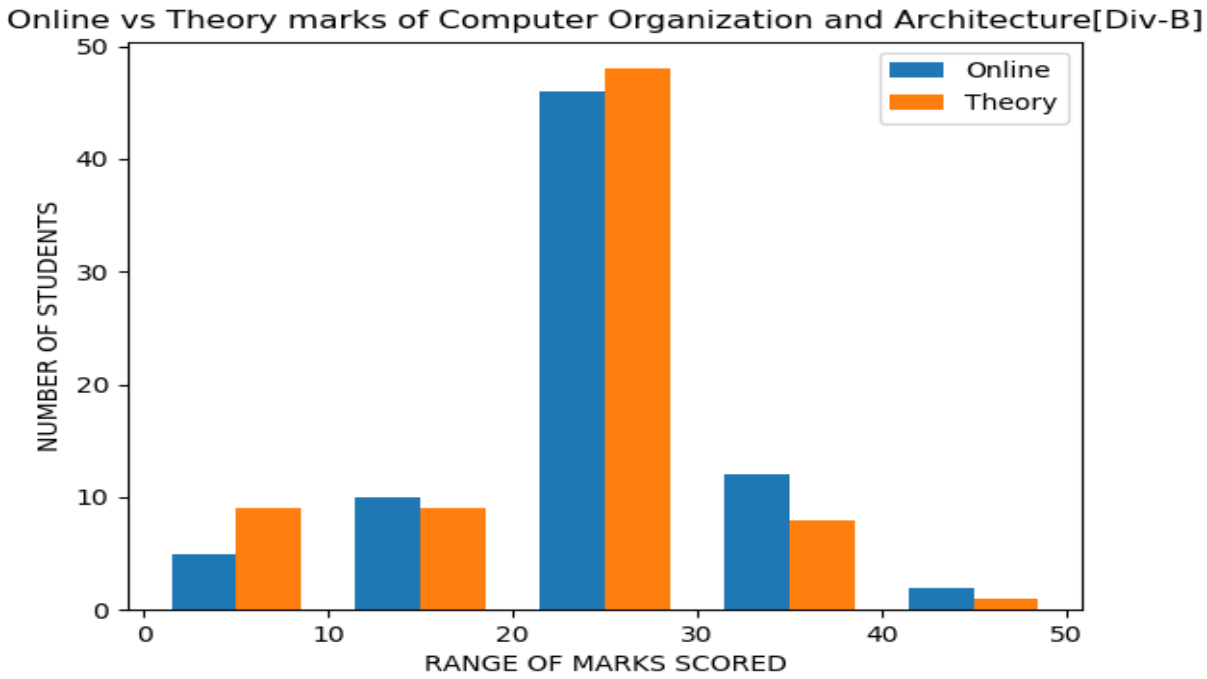


Fig. 6.2.2d Online Vs Theory marks for COA

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are least for online as well as theory. Here, in some cases ranges the online marks beats the theory marks as well as vice versa.

6.1.2e Object Oriented Programming (OOP):

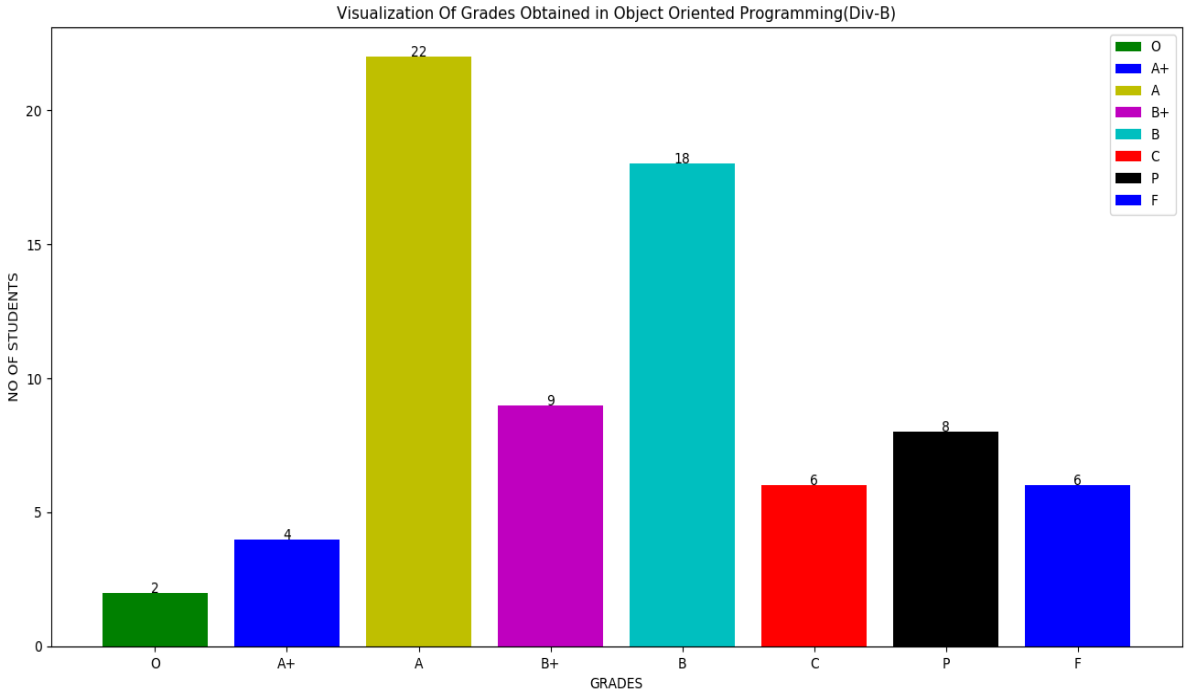


Fig. 6.1.2e Visualization of grades for the subject OOP(div B)

From the graph we can clearly see that, only 2,4 students could secure an O and A+ grade respectively, whereas 22 students have secured A grade which is highest.9 students have got B+ grade. 6 students could not clear this subject in the second attempt.18 students have got B grade which is the second highest.

Online Vs Theory marks comparison

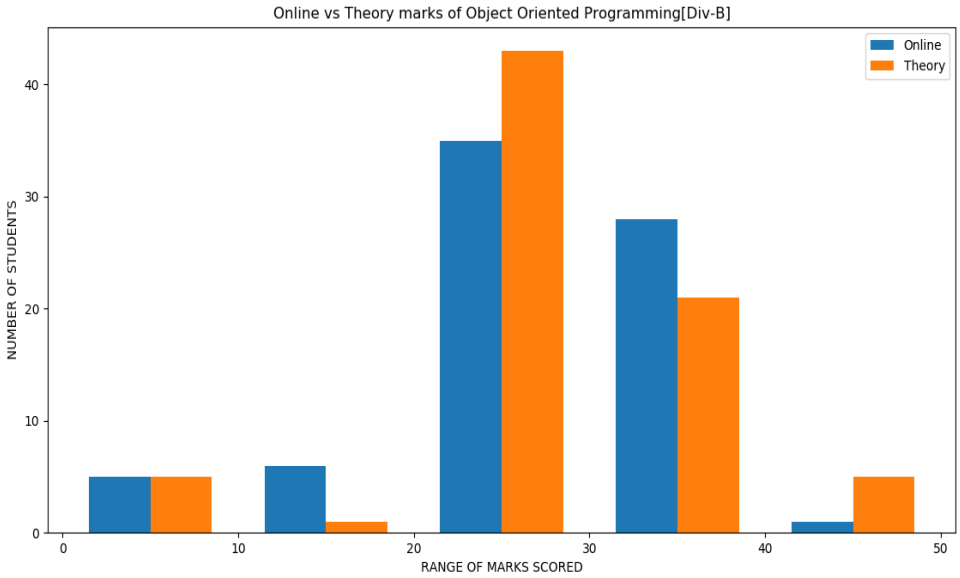


Fig. 6.2.2e Online Vs Theory marks for OOP

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are least for online as well as theory. Here, in some cases ranges the online marks beats the theory marks as well as vice versa.

6.1.2f Engineering Mathematics III (EM-III):

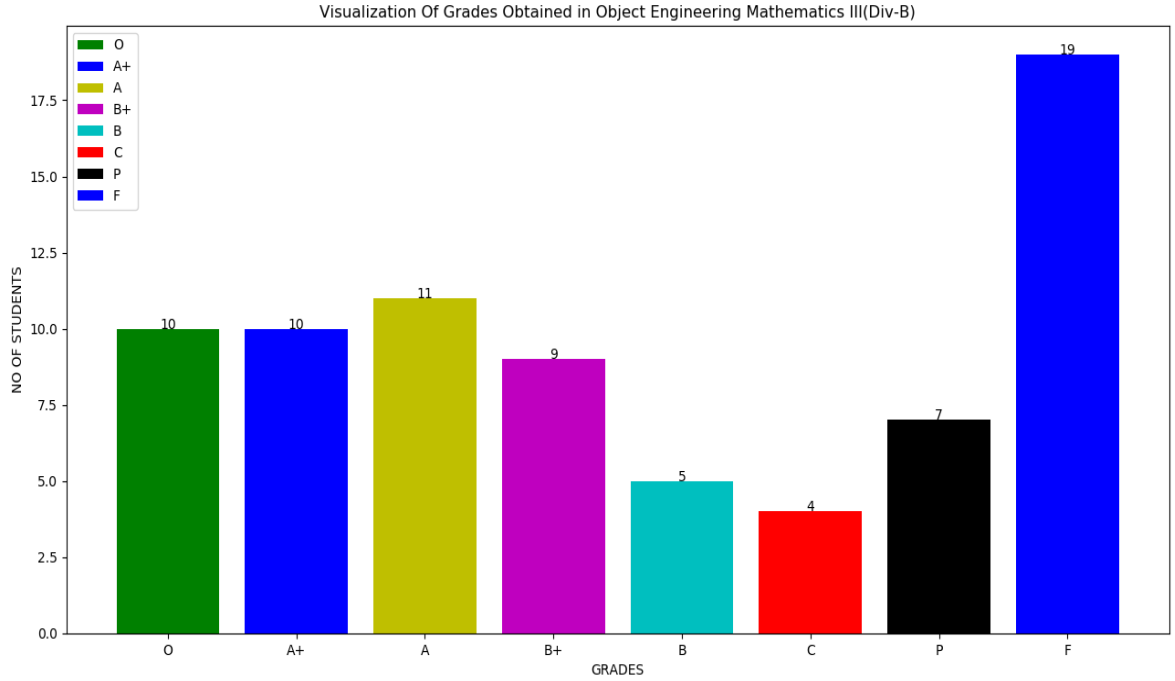


Fig. 6.1.2f Visualization of grades for the subject M3(div B)

From the graph we can clearly see that,10 students have secure an O and A+ grade each, whereas 11 students have secured A grade which is highest.9 students have got B+ grade. 19 students could not clear this subject. Scoring percentage is high here.

Online Vs Theory marks comparison

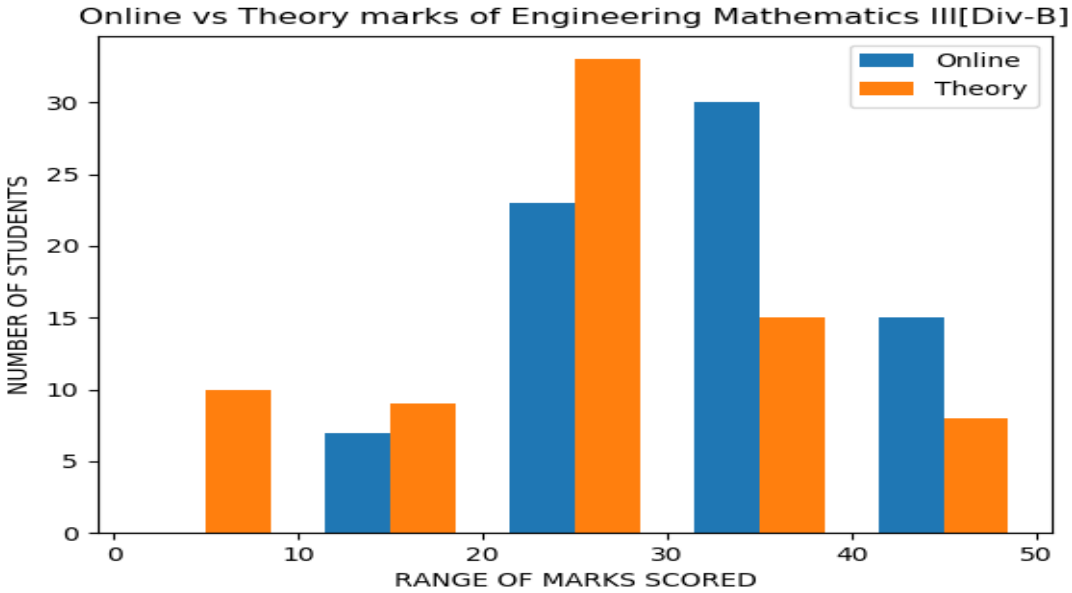


Fig. 6.2.2f Online Vs Theory marks for M3

In the above graph, for the range 20-30 and 30-40 firstly the theory marks exceeds the online marks whereas it is the opposite for the next range concluding that both the ranges equally balance each other. Students securing marks in the range of 0-10 are the least for both online as well as theory. Here ranges of 20-30,30-40,40-50 have greater ratio of students amongst them compared to first two ranges. No student from B div has got marks less than 10 in online.

6.1.2g Computer Graphics (CG):

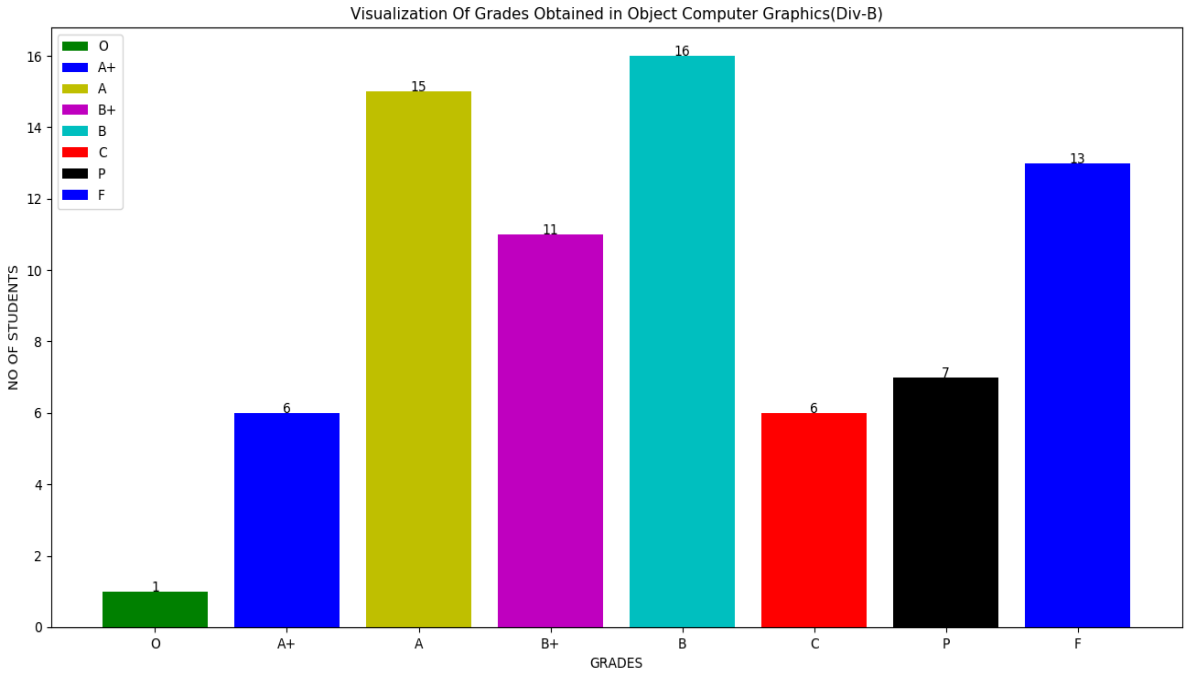


Fig. 6.1.2g Visualization of grades for the subject CG(div B)

From the graph we can clearly see that, only 1 student could secure an O grade, whereas 6 students have secured A+ grade.15 students have got A grade.13 students could not clear this subject.18 students have got B grade which is the highest.

Online Vs Theory marks comparison

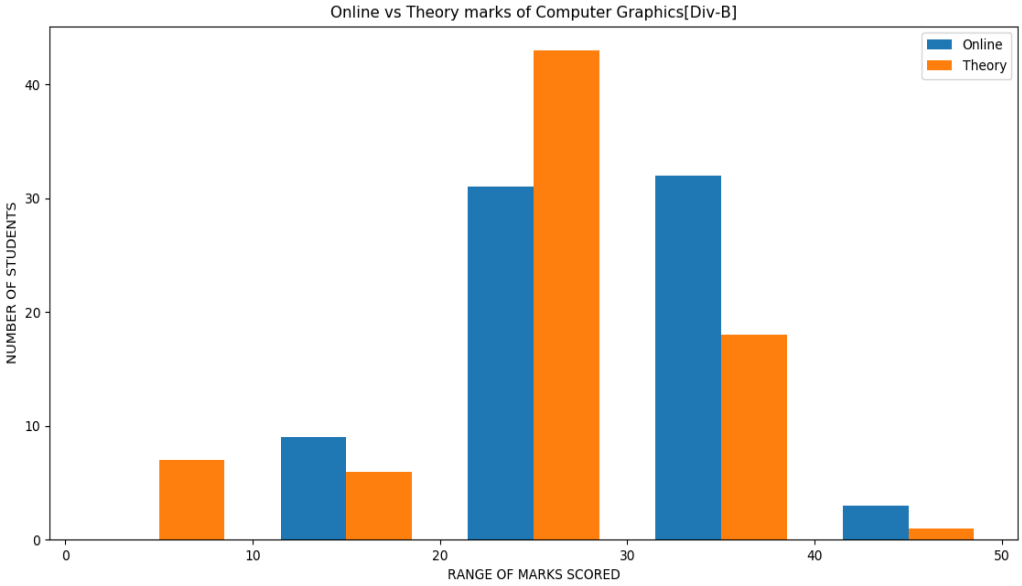


Fig. 6.2.2g Online Vs Theory marks for CG

In the above graph, for the range 20-30 and 30-40 firstly the theory marks exceeds the online marks whereas it is the opposite for the next range concluding that both the ranges equally balance each other. Students securing marks in the range of 0-10 are the least for online whereas least in the range of 40-50 for theory. No student from B div has got marks less than 10 in online.

6.1.2h Advanced Data Structure (ADS):

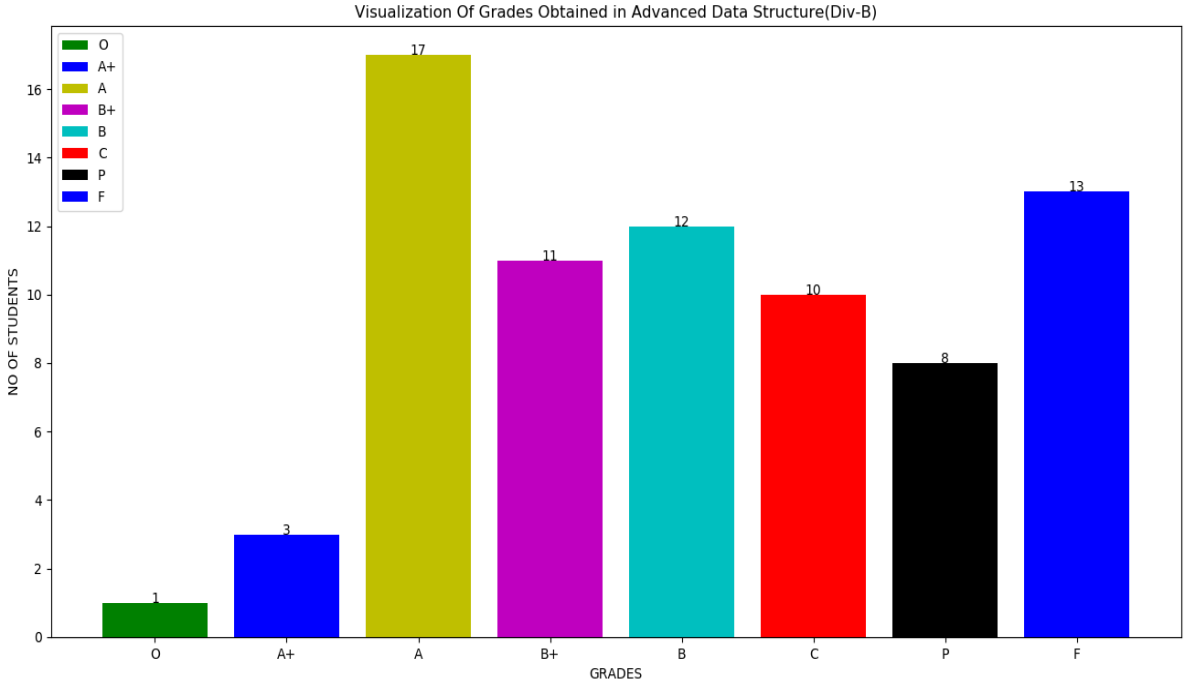


Fig. 6.1.2h Visualization of grades for the subject ADS(div B)

From the graph we can clearly see that, only 1 student could secure an O grade, whereas 3 students have secured A+ grade. 17 students have got A grade which is the highest.13 students could not clear this subject. Students getting B+,B, and C grade are nearly equal(i.e.11,12,10 respectively).

Online Vs Theory marks comparison

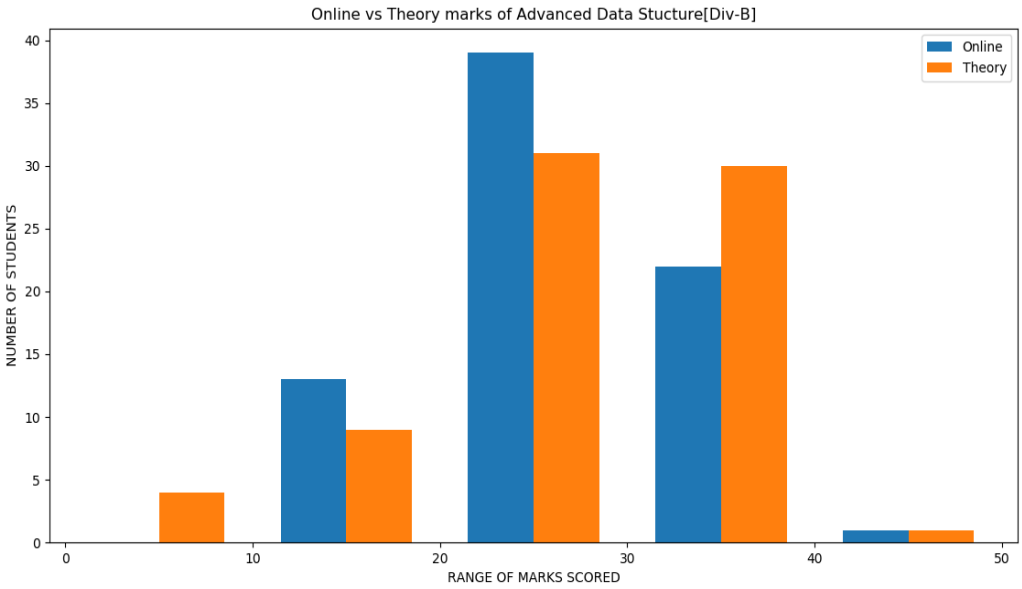


Fig. 6.2.2h Online Vs Theory marks for ADS

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 40-50 are same for theory and online. Second highest number of marks scored by students are in the range of 30-40. No student from B div has got marks less than 10 in online.

6.1.2i Microprocessor (MP):

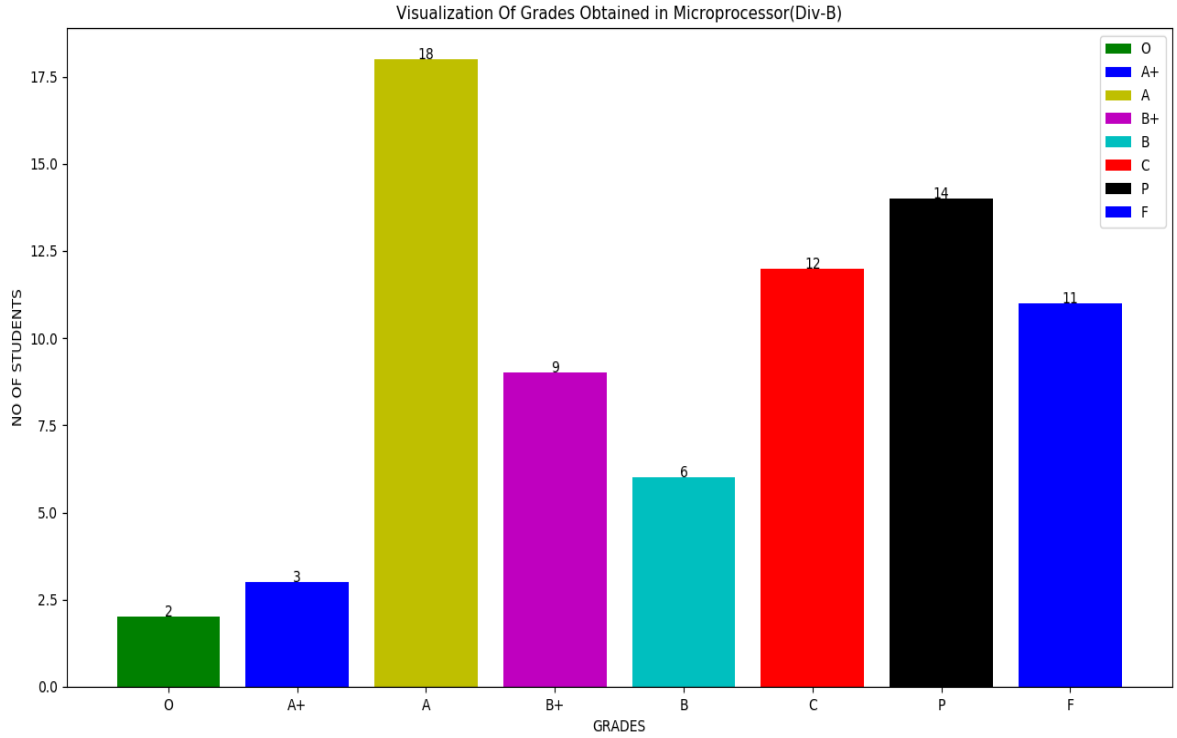


Fig. 6.1.2i Visualization of grades for the subject MP(div B)

From the graph we can clearly see that, only 2 students could secure an O grade, whereas 3 students have secured A+ grade.18 students have got A grade which is the highest.11 students could not clear this subject.

Online Vs Theory marks comparison

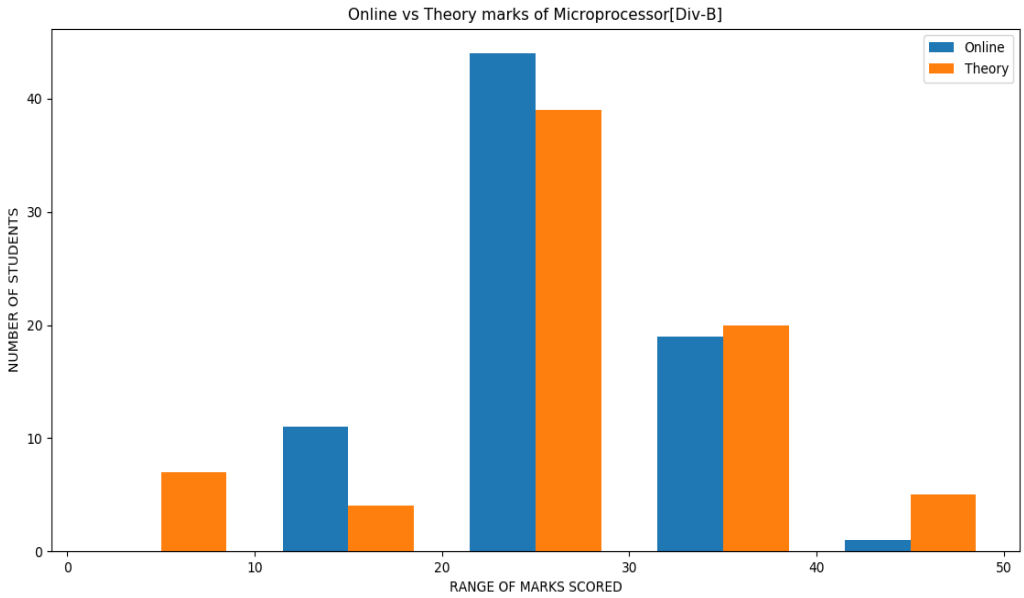


Fig. 6.2.2i Online Vs Theory marks for MP

In the above graph we can clearly see that, most of students have secured marks in the range of 20-30 for both online as well as theory. Students securing marks in the range of 10-20 are the least for theory whereas for online the range 0-10 is applicable as there are no figures of marks .The ratio of of online marks for 10-20 is a lot greater than its theory.

6.1.2j Principles of Programming Languages (PPL):

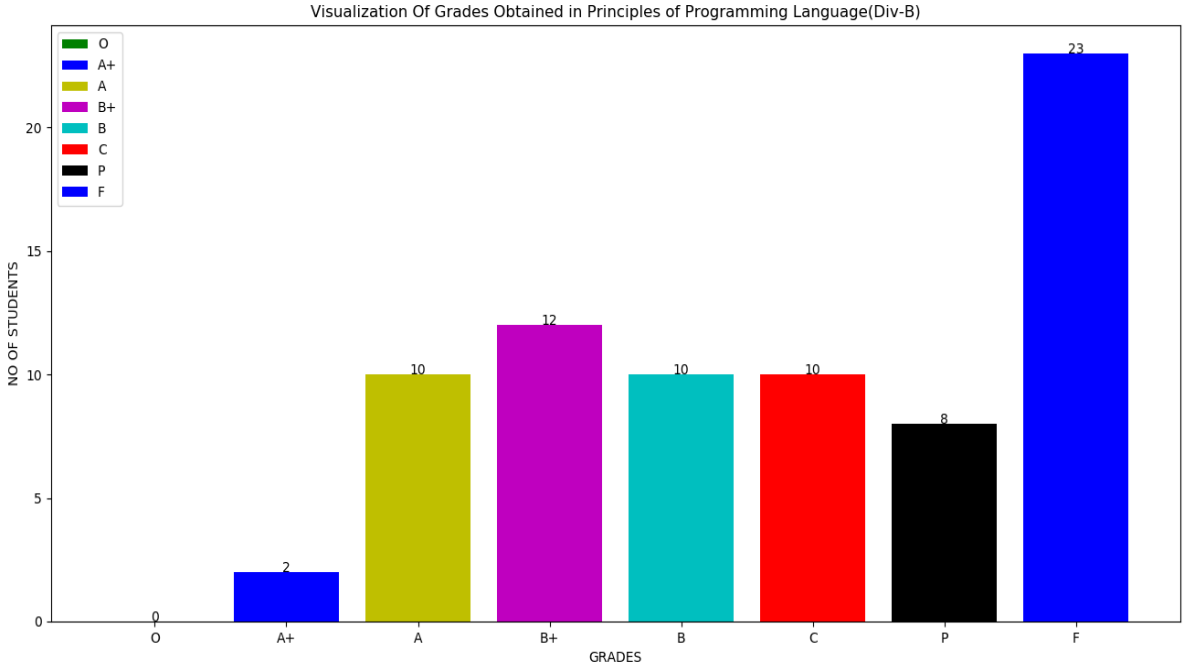


Fig. 6.1.2j Visualization of grades for the subject PPL(div B)

From the graph we can clearly see that, no one could secure an O grade, whereas only 2 students have secured A+ grade.10 students have got A grade. 23 students could not clear this subject. 12 students have got B+ grade

Online Vs Theory marks comparison

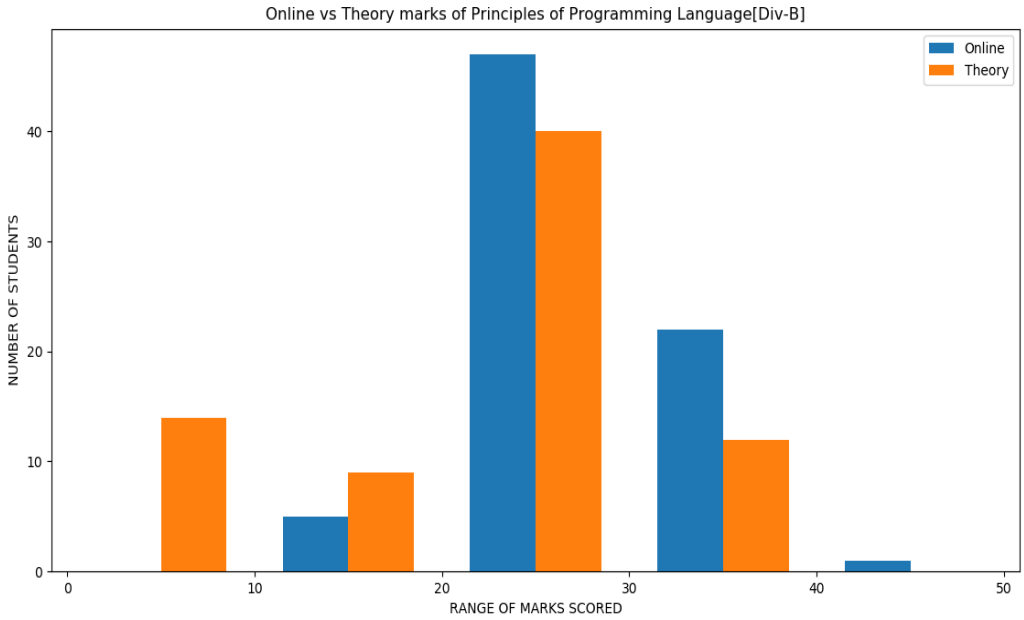


Fig. 6.2.2j Online Vs Theory marks for PPL

In the above graph we can clearly see that, maximum number of students have secured marks in the range of 20-30 for both online as well as theory. No student has secured marks in the range of 40-50 for theory and 0-10 for online. Second highest number of marks scored by students are in the range of 30-40.

6.5 Normal Linear Regression Graph for Prediction

Linear regression is the process of fitting a linear equation to set of sample data in order to predict the output. Here we generated the linear regression for predicting the credit points from the total (i.e. online and theory) marks. We used PYTHON for plotting the regression graph.

Code:

```
import matplotlib.pyplot as plt
import seaborn as sb
import pandas as pd
d = pd.read_csv('workbook.csv')
sb.regplot(x="Total Marks",y="Credit Points",data=d)
plt.show()
```

seaborn library :

seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

sb.regplot() :

This function in seaborn is used to visualize a linear relationship as determined through regression.

Let us see the regression for each subject.

6.5.1 Discrete Mathematics (DM)

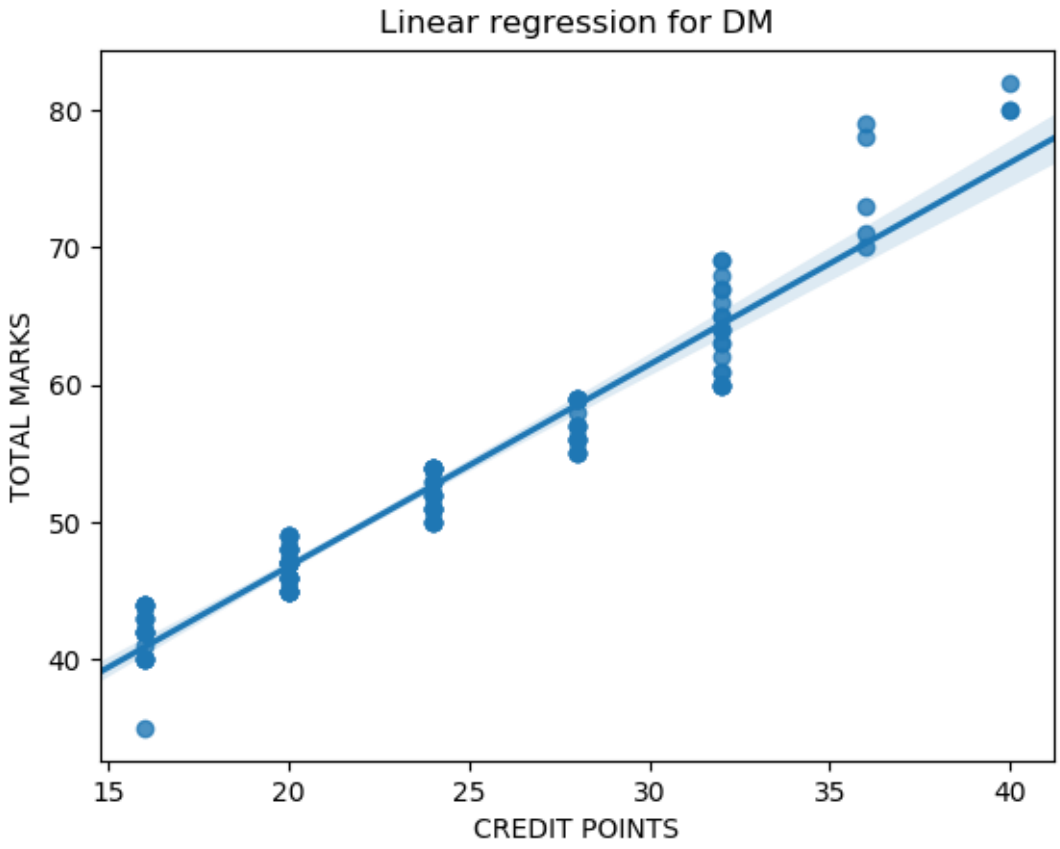


Fig. 6.5.1 Linear regression for DM

6.5.2 Digital Electronics and Logic Design

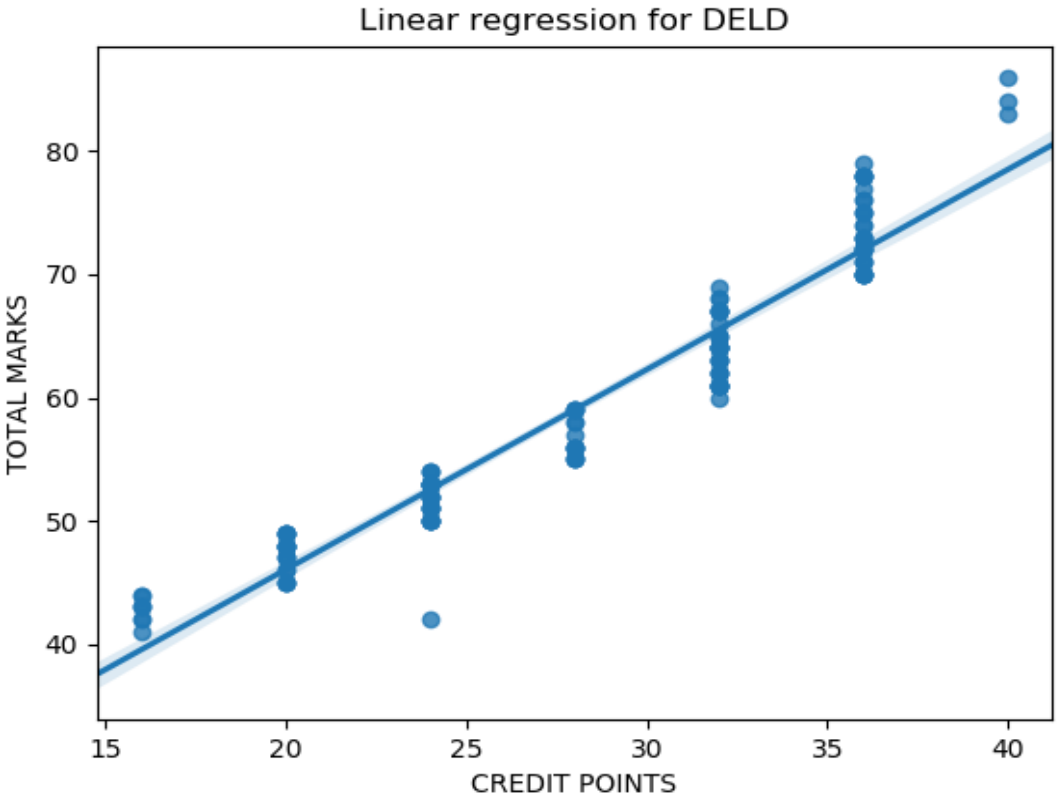


Fig. 6.5.2 Linear regression for DELD

6.5.3 Data Structure and Algorithm

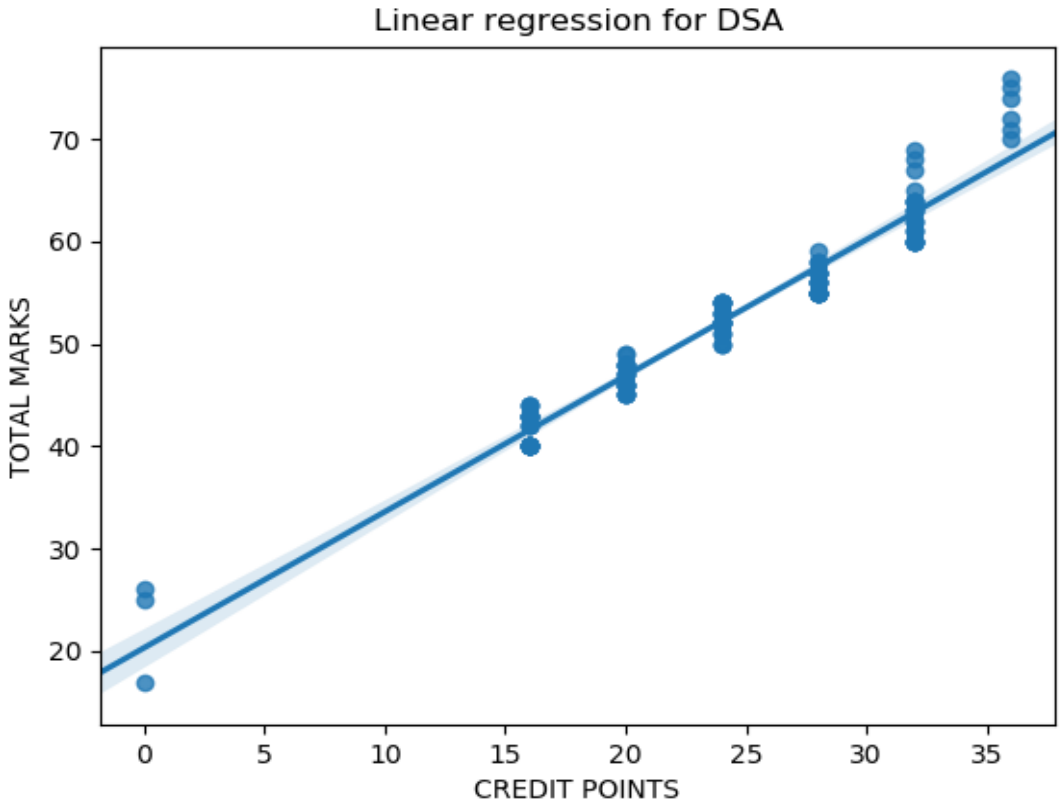


Fig. 6.5.3 Linear regression for DSA

6.5.4 Computer Organization and Architecture

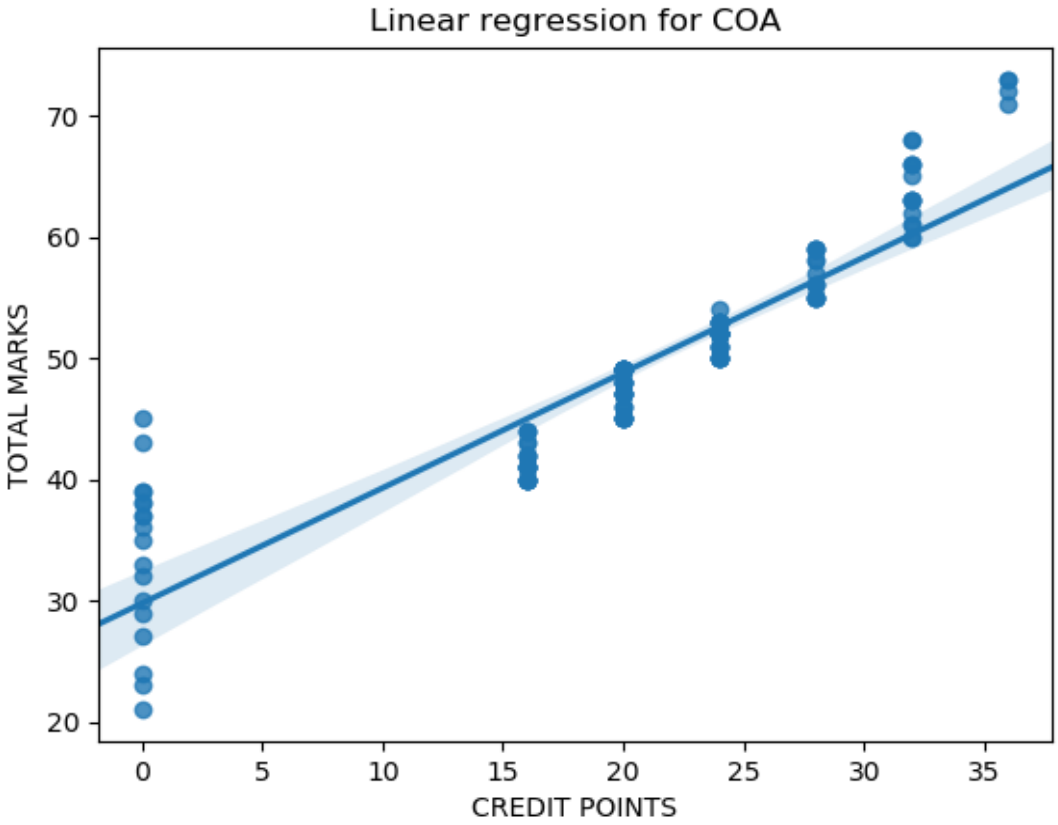


Fig. 6.5.4 Linear regression for COA

6.5.5 Object Oriented Programming

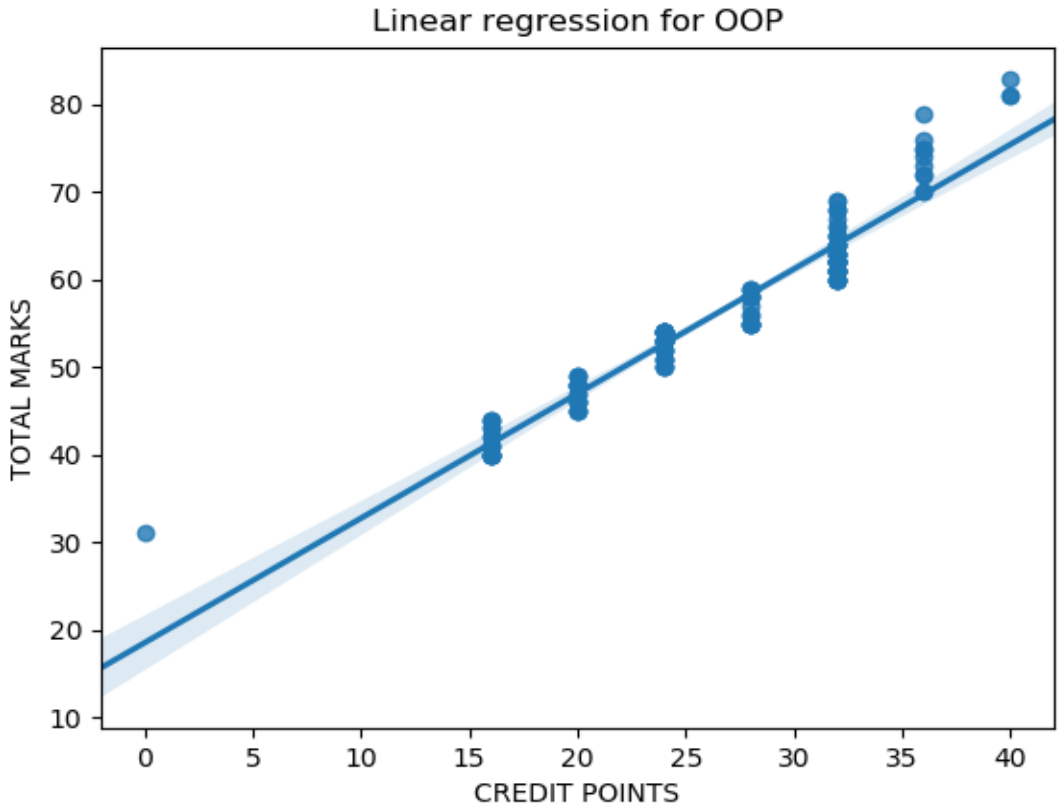


Fig. 6.5.5 Linear regression for OOP

6.5.6 Engineering Mathematics III

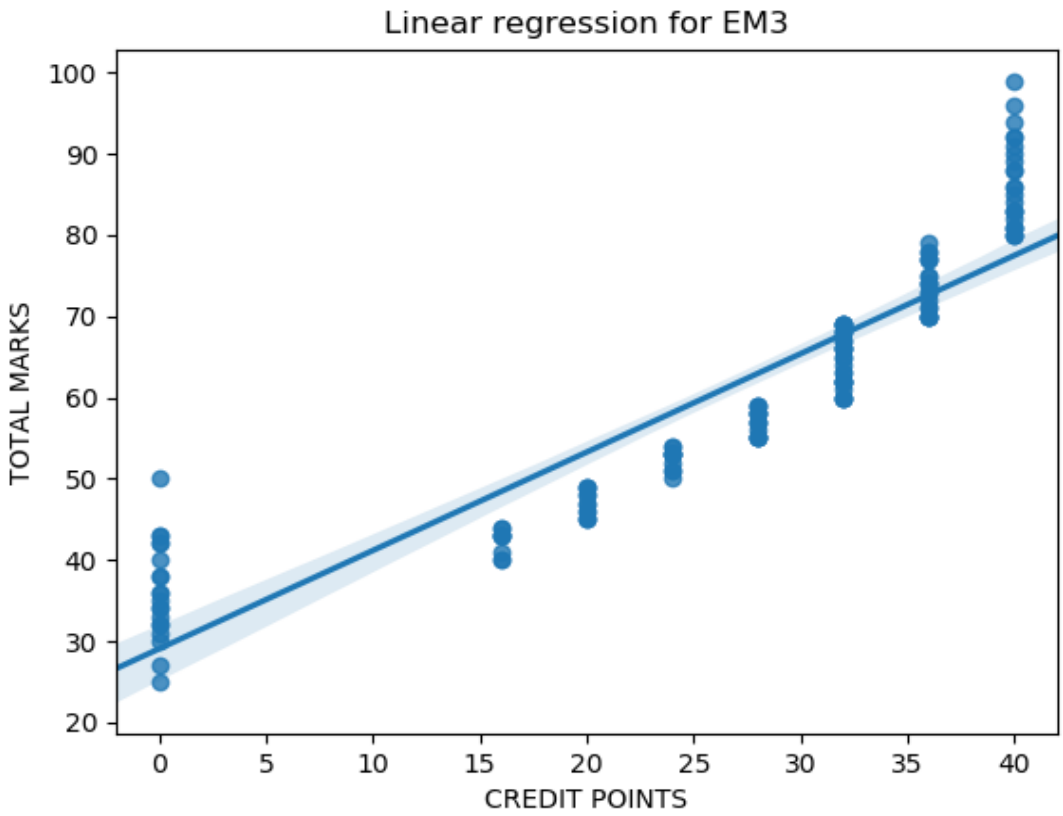


Fig. 6.5.6 Linear regression for EM III

6.5.7 Computer Graphics

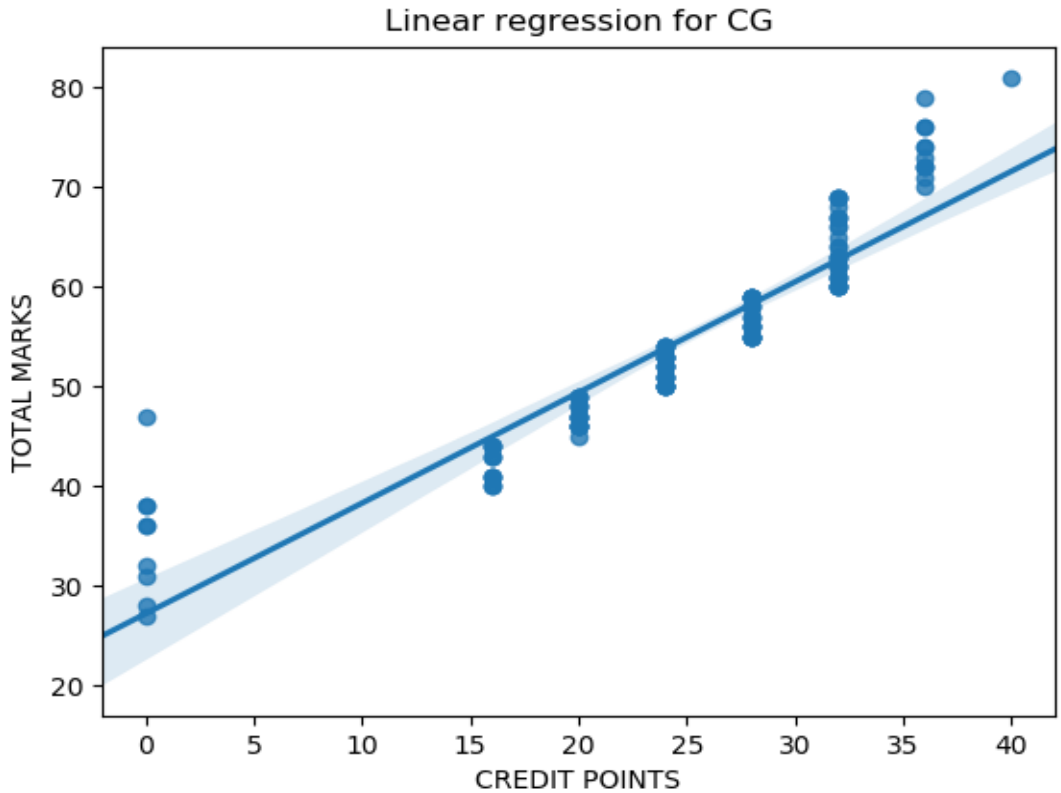


Fig. 6.5.7 Linear regression for CG

6.5.8 Advanced Data Structure

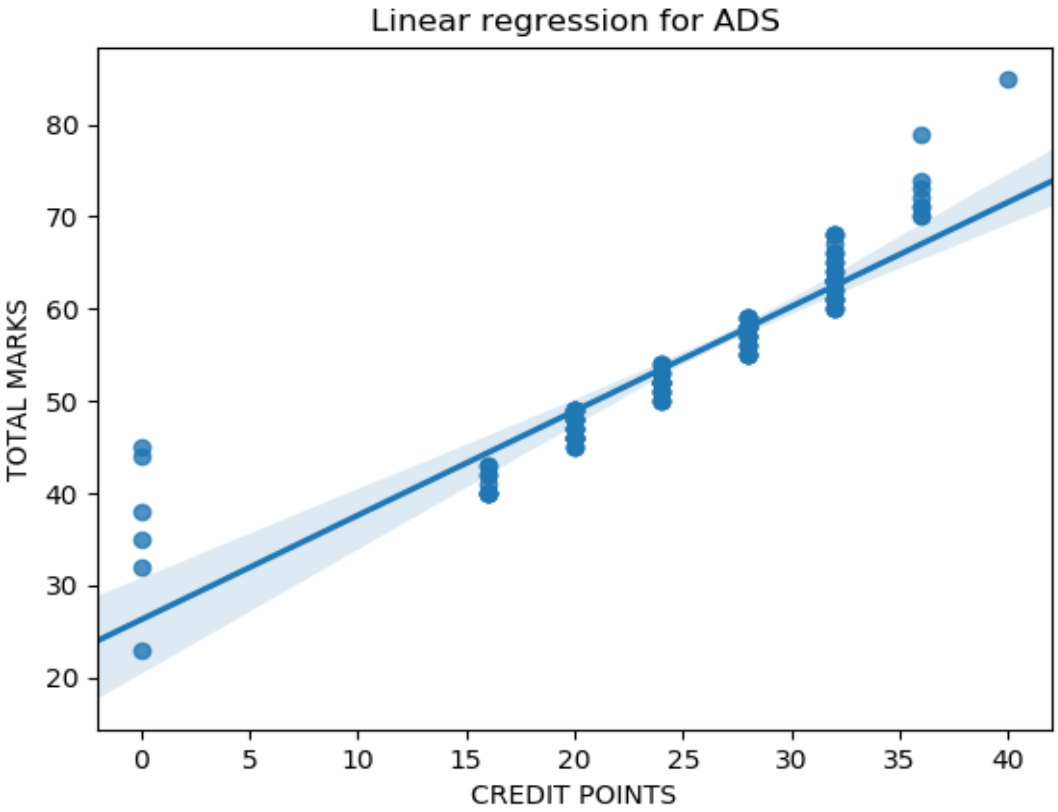


Fig. 6.5.8 Linear regression for ADS

6.5.9 Microprocessor

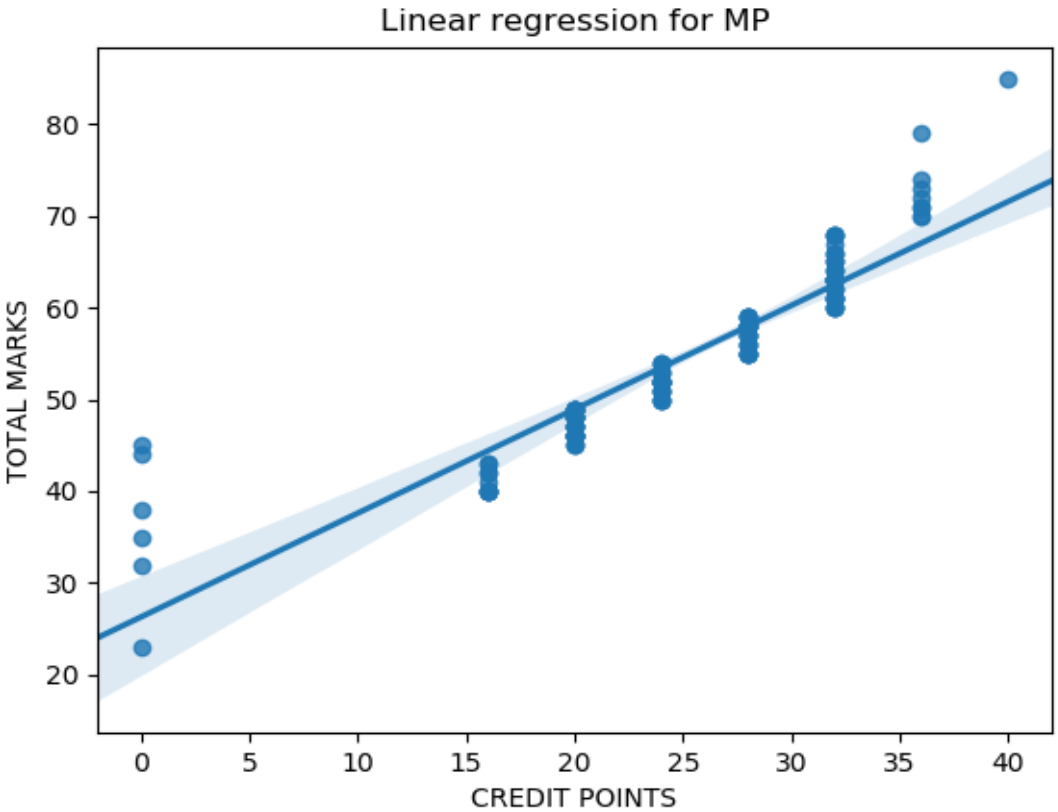


Fig. 6.5.9 Linear regression for MP

6.5.10 Principles of Programming Language

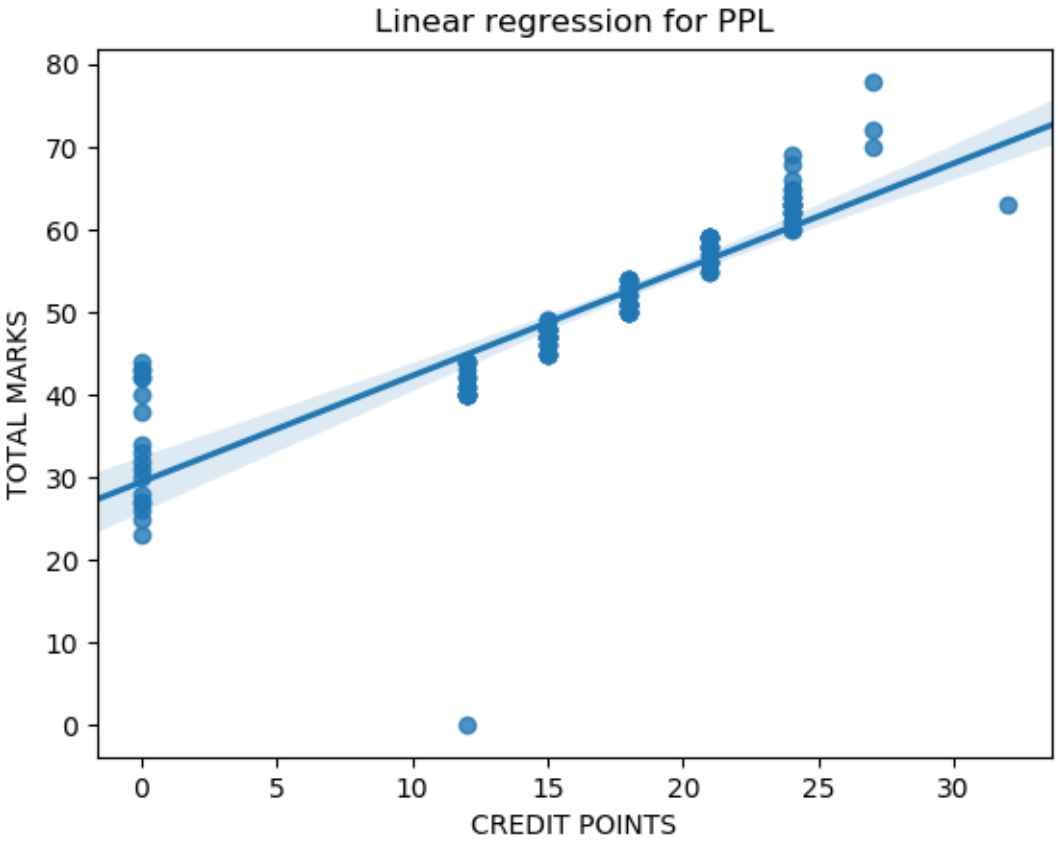


Fig. 6.5.10 Linear regression for PPL

CHAPTER 7
CONCLUSION

From this mini-project we came to know about the result of each and every subject of the second year (2017-18) [Computer department]. From the overall analysis we have concluded the performance of students in each subject. By analysing the result, we came to the conclusion that 2017-18 result is better than the previous years.

The result we got is as follows:

Discrete Mathematics - 89.27%
Digital Electronics and Logic Design - 96.02%
Data Structure and Algorithms – 93.75%
Computer Organization and Architecture – 80.11%
Object Oriented Programming – 98.30%
Engineering Mathematics III – 77.9%
Computer Graphics – 85.64%
Advanced Data Structure – 87.29%
Microprocessor - 77.9%
Principles of Programming Language – 79.01%

In this year, students scored more marks in OOP subject i.e. 98.30%. EM III has the lowest result as compared to all other subjects. All the remaining subjects are having average result.

By comparing the online and theory marks of students, we came to know that online is easy to score. All students scored more marks in online as compared to theory. Because of this the overall result (pointer) of each student is increased.

According to our analysis, we can say that most of students scored marks in the range of 20-30 for both online and theory.

In DM the number of students getting C grade (i.e.in the range of 60-69) is the most as compared to other grades.

In DELD the number of students securing A+, A, B grades are nearly equal (i.e. 33, 36, 34 students respectively).

In DSA the number of students getting P grade is the most compared to others. The number of students securing A, B+, B, C grades are nearly equal (i.e. 26, 31, 31, 32 respectively).

In COA the no. of students getting B and C grade are nearly equal. 40 students gets failed in this subject.

In OOP the number of students getting A grade (i.e.in the range of 80-89) is the most compared to others.

In EM III the failure percentage is the most i.e. 40 students couldn't cleared this subject. 26 students got O grade in EM III.

In CG, 26 students couldn't clear this subject. 32 students got A grade.

In ADS 38 students secured A grade.

In this year the highest pointer obtained is 9.7 and the lowest is 4.54.

From this result we can say that students have to take more efforts to increase the overall result of computer department.

Both A and B division has the average performance in the last year.

REFERENCES

1. http://www.w3schools.org
2. Fundamentals of Python
   - By B.L.JUNEJA
3. http://www.pythonprogramminglanguage.com
4. http://www.pythonspot.com
5. http://geeksforgeeks.com