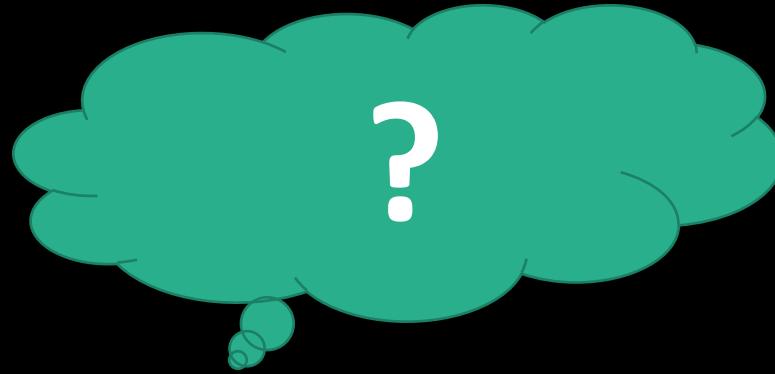


DATA ANALYTICS



DATA ANALYTICS

DATA ANALYTICS

Is the Science of analyzing **data in order to
make **conclusions** about that information**



same as Data Science?

DATA ANALYTICS

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

**14 Data Analysis Documents
12 Data Science Documents**

Converted word counts to word rates

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

Most frequent concept in both fields

(1) DATA

(2) MODEL

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

Same concept for both Data Analytics and Data Science

COMPUTATIONAL MATH

STATISTICS

DISCRETE MATH

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

PROBLEM SOLVING has higher word rates for Data Science

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

Data Analytics and Data Science is a blend of

STATISTICS

MATHEMATICS

COMPUTER SCIENCE

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

**Software development is lower on both
Data Analytics and Data Science**

**THEREFORE: they do not write the program,
they use it**

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

on both Data Analytics and Data Science

without databases would be ineffective

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

Closer look at **STATISTICS CONCEPTS**

**DATA
ANALYTICS**



**Exploratory concerns
(searching pattern in
sample data)**

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

Closer look at **STATISTICS CONCEPTS**

**DATA
SCIENCE**



Inferential activities use sample data to draw conclusions about the population

A Comparison of Key Concepts in Data Analytics and Data Science

Information Systems Education Journal (ISEDJ)
ISSN: 1545-679X

16 (1)
February 2018

Data Analytics vs Data Science

Closer look at **STATISTICS CONCEPTS**

**DATA
ANALYTICS**



Higher rate for big data

**DATA
SCIENCE**



**Traditional small sample
problems**

A Comparison of Key Concepts in Data Analytics and Data Science

Information & Communication Technology
ISSN: 1541-3226

16 (1)
January 2018

Goals,
Methods and
Applications
are
evolving

ANALYTICS

DATA
SCIENCE

small sample
problems

DATA ANALYTICS

Is the Science of analyzing **data in order to
make **conclusions** about that information**

DATA SCIENCE

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

DIAGNOSTIC ANALYTICS

PREDICTIVE ANALYTICS

PRESCRIPTIVE ANALYTICS

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

Describes what happened over
a period of time

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

Have the number of enrollees
gone up?

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

**Are enrollees for School of
Engineering better this year
than last year?**

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

DATA SET ON ENROLLMENT - Excel (Product Activation Failed)

	B	C	D	E	F	G	H	I	J	K	L
1	SY	YEAR	SEMESTER	GENDER	ENROLLEES						
2	2012-2013	FIRST YEAR	FIRST SEM	MALE	145						
3	2012-2013	SECOND YEAR	FIRST SEM	MALE	108						
4	2012-2013	THIRD YEAR	FIRST SEM	MALE	65						
5	2012-2013	FOURTH YEAR	FIRST SEM	MALE	33						
6	2012-2013	FIFTH YEAR	FIRST SEM	MALE	0						
7	2012-2013	FIRST YEAR	FIRST SEM	MALE	145						
8	2012-2013	SECOND YEAR	FIRST SEM	MALE	108						
9	2012-2013	THIRD YEAR	FIRST SEM	MALE	104						
10	2012-2013	FOURTH YEAR	FIRST SEM	MALE	34						
11	2012-2013	FIFTH YEAR	FIRST SEM	MALE	0						
12	2012-2013	FIRST YEAR	FIRST SEM	MALE	85						
13	2012-2013	SECOND YEAR	FIRST SEM	MALE	38						
14	2012-2013	THIRD YEAR	FIRST SEM	MALE	30						
15	2012-2013	FOURTH YEAR	FIRST SEM	MALE	11						
16	2012-2013	FIFTH YEAR	FIRST SEM	MALE	0						

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

DATA SET

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

TABLES

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

GRAPHS

SEAS SBMA STHM SE SN

TOTAL 1ST YEAR

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

**focuses more on why
something happened**

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

**This involves more diverse
data inputs and a bit of
hypothesizing.**

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

DATA SET FOR APPLICANTS 2017 2018 2019 COMBINED - Excel (Product Activation Failed)

	APPLICANT	STATUS	GPA	INCOME	GPA_A	INCOME_A	SCHOLAR	WORKING	SCHOOL/DEPT/COLLEGE START	SCHOOL/ORIGIN	SCHOOL/DEPT/COLLEGE CURRENT	
1												
2	1	ENROLLED	85-90	30K AND ABOVE	85	73746	1	1	3	PRIVATE	3	
3	2	ENROLLED	85-90	30K AND ABOVE	85	31633	0	0	1	PUBLIC	1	
4	3	ENROLLED	85-90	30K AND ABOVE	87	63525	0	0	2	PRIVATE	2	
5	4	ENROLLED	85-90	30K AND ABOVE	98	43081	1	1	2	PUBLIC	2	
6	5	ENROLLED	85-90	30K AND ABOVE	89	90054	1	1	3	PUBLIC	3	
7	6	ENROLLED	85-90	30K AND ABOVE	87	46836	0	1	1	PRIVATE	1	
8	7	ENROLLED	85-90	30K AND ABOVE	93	94500	1	1	2	PUBLIC	2	
9	8	ENROLLED	85-90	30K AND ABOVE	93	89981	0	0	1	PRIVATE	1	
10	9	ENROLLED	85-90	30K AND ABOVE	92	81838	0	0	1	PUBLIC	1	
11	10	ENROLLED	85-90	30K AND ABOVE	85	39466	0	0	4	PRIVATE	4	
12	11	ENROLLED	85-90	30K AND ABOVE	85	71155	1	1	1	PUBLIC	1	
13	12	ENROLLED	85-90	30K AND ABOVE	92	55779	0	0	4	PUBLIC	4	
14	13	ENROLLED	85-90	BELOW 30K	93	27871	1	1	1	PUBLIC	1	
15	14	ENROLLED	85-90	30K AND ABOVE	85	54183	0	0	3	PUBLIC	3	
16	15	ENROLLED	85-90	30K AND ABOVE	86	65164	0	1	1	PUBLIC	1	
17	16	ENROLLED	85-90	BELOW 30K	92	20990	1	0	2	PRIVATE	2	
18	17	ENROLLED	85-90	30K AND ABOVE	90	47848	1	0	2	PUBLIC	2	
19	18	ENROLLED	85-90	BELOW 30K	95	21332	0	1	5	PUBLIC	5	

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

Data set of applicants

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

**Income somehow affects the
Status of applicants
enrolling in the school?**

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

INCOME

**30k and above
below 30k**

STATUS

**Enrolled
Not Enrolled**

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

We want to check that the proportions who belong to
INCOME categories
and
STATUS categories
are different or not

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

We will use cross-tabulation of
INCOME categories

vs

STATUS categories

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

STEP 1: Highlight Raw Data

	YEAR OF ADMISSION	APPLICANT	STATUS	GPA	INCOME	GPA_A	INCOME_A	SCHOLAR	WORKING	SCHOOL/DEPT/COLLEGE START	SCHOOL/DEPT/	SCHOOL/DEP
2019	285	NOT ENROLLED	75-79	BELOW 30K	78	17354		0		PUBLIC		
2019	286	NOT ENROLLED	75-79	BELOW 30K	75	10946		1		PRIVATE		
2019	287	NOT ENROLLED	75-79	BELOW 30K	77	11667		1		PRIVATE		
2019	288	NOT ENROLLED	75-79	BELOW 30K	78	10739		1		PUBLIC		
2019	289	NOT ENROLLED	75-79	BELOW 30K	77	14366		1		PRIVATE		
2019	290	NOT ENROLLED	75-79	BELOW 30K	79	16267		0		PUBLIC		
2019	291	NOT ENROLLED	75-79	BELOW 30K	78	12985		0		PUBLIC		
2019	292	NOT ENROLLED	75-79	BELOW 30K	79	14668		1		PUBLIC		
2019	293	NOT ENROLLED	75-79	BELOW 30K	77	14615		1		PRIVATE		
2019	294	NOT ENROLLED	75-79	BELOW 30K	76	15994		1		PRIVATE		
2019	295	NOT ENROLLED	75-79	BELOW 30K	75	15802		1		PUBLIC		
2019	296	NOT ENROLLED	75-79	BELOW 30K	79	10925		0		PRIVATE		
2019	297	NOT ENROLLED	75-79	BELOW 30K	76	10204		1		PUBLIC		
2019	298	NOT ENROLLED	75-79	BELOW 30K	76	18308		0		PUBLIC		



2017

2018

2019

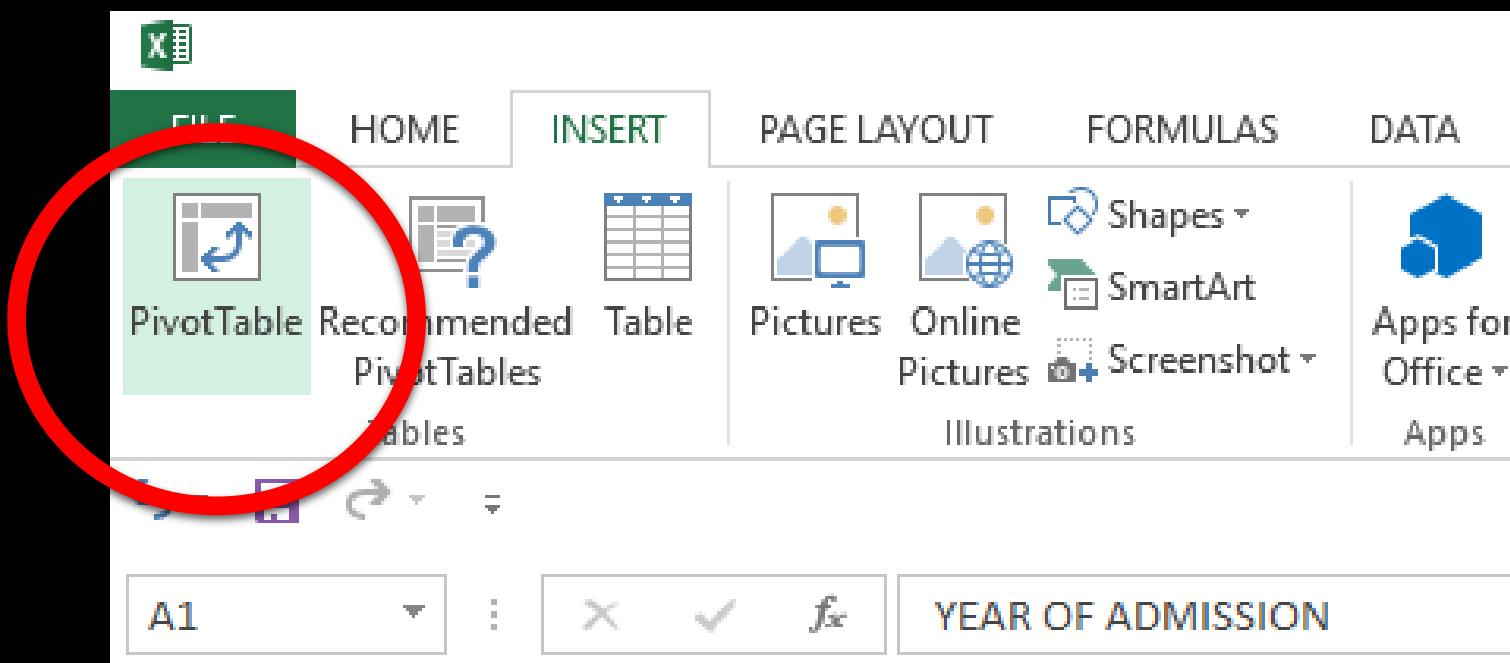
COMBINED



Types of Data Analytics:

DIAGNOSTIC ANALYTICS

STEP 2: Click PIVOT table



Types of Data Analytics:

DIAGNOSTIC ANALYTICS

STEP 3: Click OK

The image shows a Microsoft Excel spreadsheet titled "COMBINED" with data from 2017, 2018, and 2019. The data includes columns for Year of Admission, Applicant Status, GPA, Income, and Scholarships. A PivotTable creation dialog box is overlaid on the spreadsheet, prompting the user to select a table or range (COMBINED!\$A\$1:\$L\$1001), choose a new worksheet, and add the data to the Data Model.

	YEAR OF ADMISSION	APPLICANT	STATUS	GPA	INCOME	GPA_A	INCOME_A	SCHOLAR	WEIGHTAGE START	WEIGHTAGE END	COLLEGE
2019	287	NOT ENROLLED	75-79	BELOW 30K	77	11667		1			
2019	288	NOT ENROLLED	75-79	BELOW 30K	78	10739		1			
2019	289	NOT ENROLLED	75-79	BELOW 30K	77	14366		1			
2019	290	NOT ENROLLED	75-79	BELOW 30K	79	16267		0			
2019	291	NOT ENROLLED	75-79	BELOW 30K	78	12985		0			
2019	292	NOT ENROLLED	75-79	BELOW 30K	79	14668		1			
2019	293	NOT ENROLLED	75-79	BELOW 30K	77	14615		1			
2019	294	NOT ENROLLED	75-79	BELOW 30K	76	15994		1			
2019	295	NOT ENROLLED	75-79	BELOW 30K	75	15802		1			
2019	296	NOT ENROLLED	75-79	BELOW 30K	79	10925		0			
2019	297	NOT ENROLLED	75-79	BELOW 30K	76	10204		1			
2019	298	NOT ENROLLED	75-79	BELOW 30K	76	18308		0			PUBLIC
2019	299	NOT ENROLLED	75-79	BELOW 30K	76	11177		1			PUBLIC
2019	300	NOT ENROLLED	75-79	BELOW 30K	78	13079		0			PUBLIC

COMBINED

AVERAGE: 7064.832143 COUNT: 9612 SUM: 39563060

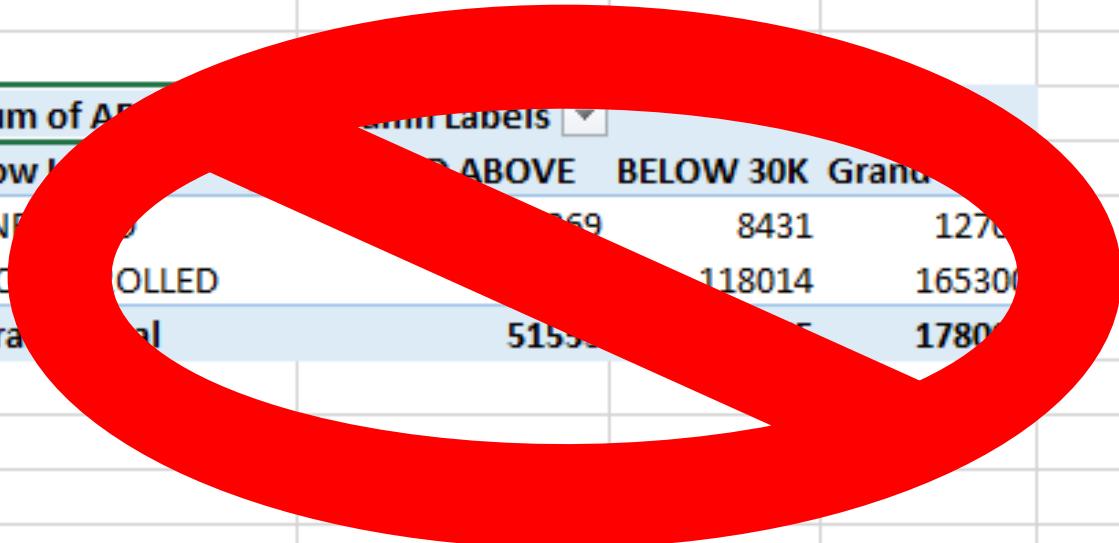
DIAGNOSTIC ANALYTICS

**STEP 4: Drag the following:
STATUS – ROWS
INCOME – COLUMN
APPLICANT – VALUES**

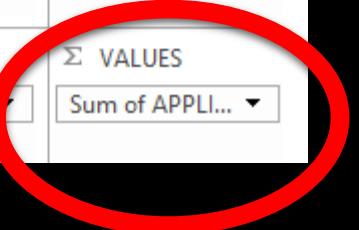
The screenshot shows the 'PivotTable Fields' ribbon interface. In the 'Choose fields to add to report:' section, several fields are listed with checkboxes: YEAR OF ADMISSION (unchecked), APPLICANT (checked), STATUS (checked), GPA (unchecked), INCOME (checked), GPA_A (unchecked), INCOME_A (unchecked), SCHOLAR (unchecked), and WORKING (unchecked). Below this, under 'Drag fields between areas below:', the 'FILTERS' section contains 'INCOME'. The 'ROWS' section contains 'STATUS'. The 'COLUMNS' section contains 'INCOME'. The 'VALUES' section contains 'Sum of APPLI...'. The entire interface is set against a dark background.

Types of Data Analytics:

DIAGNOSTIC ANALYTICS



	A	B	C	D	E	F
Sum of APPLICANT						
Row Labels						
EDUCATIONAL LEVEL	ABOVE 30K	BELOW 30K	Grand Total			
ENROLLED	1269	8431	12780			
NOT ENROLLED	118014	165300	283314			
Grand Total	51559	253621	178083			



PivotTable Fields

Choose fields to add to report:

- YEAR OF ADMISSION
- APPLICANT
- STATUS
- GPA
- INCOME
- GPA_A
- INCOME_A
- SCHOLAR
- WORKING

Drag fields between areas below:

FILTERS

COLUMNS

INCOME

ROWS

VALUES

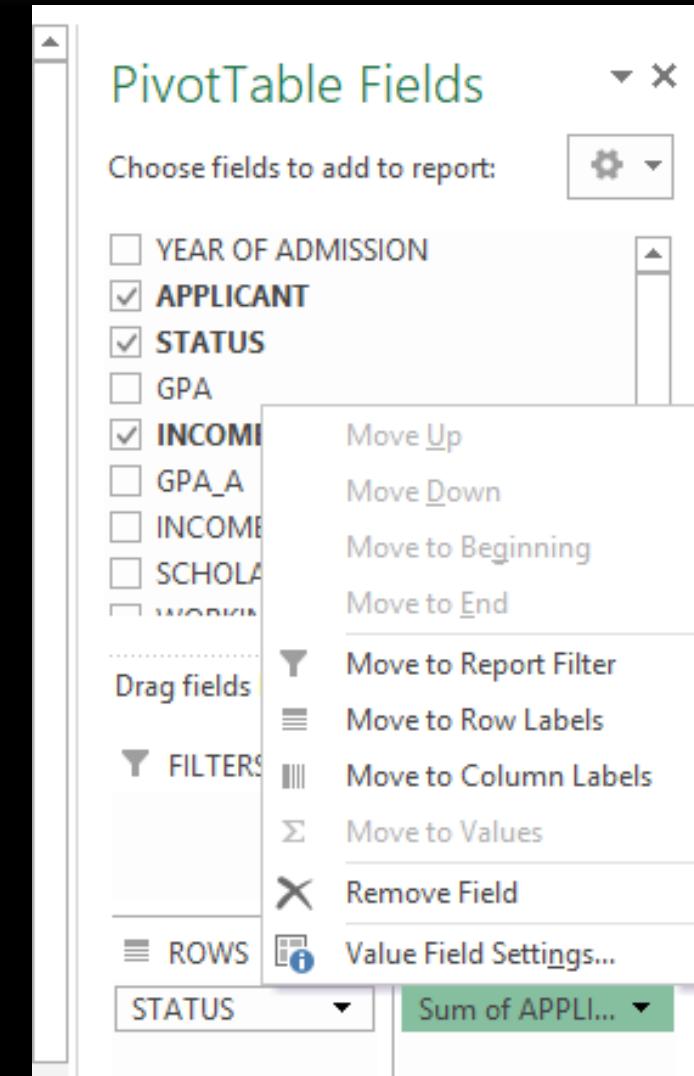
Sum of APPLICANT

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

STEP 5: Click (as shown, this will show the menu)

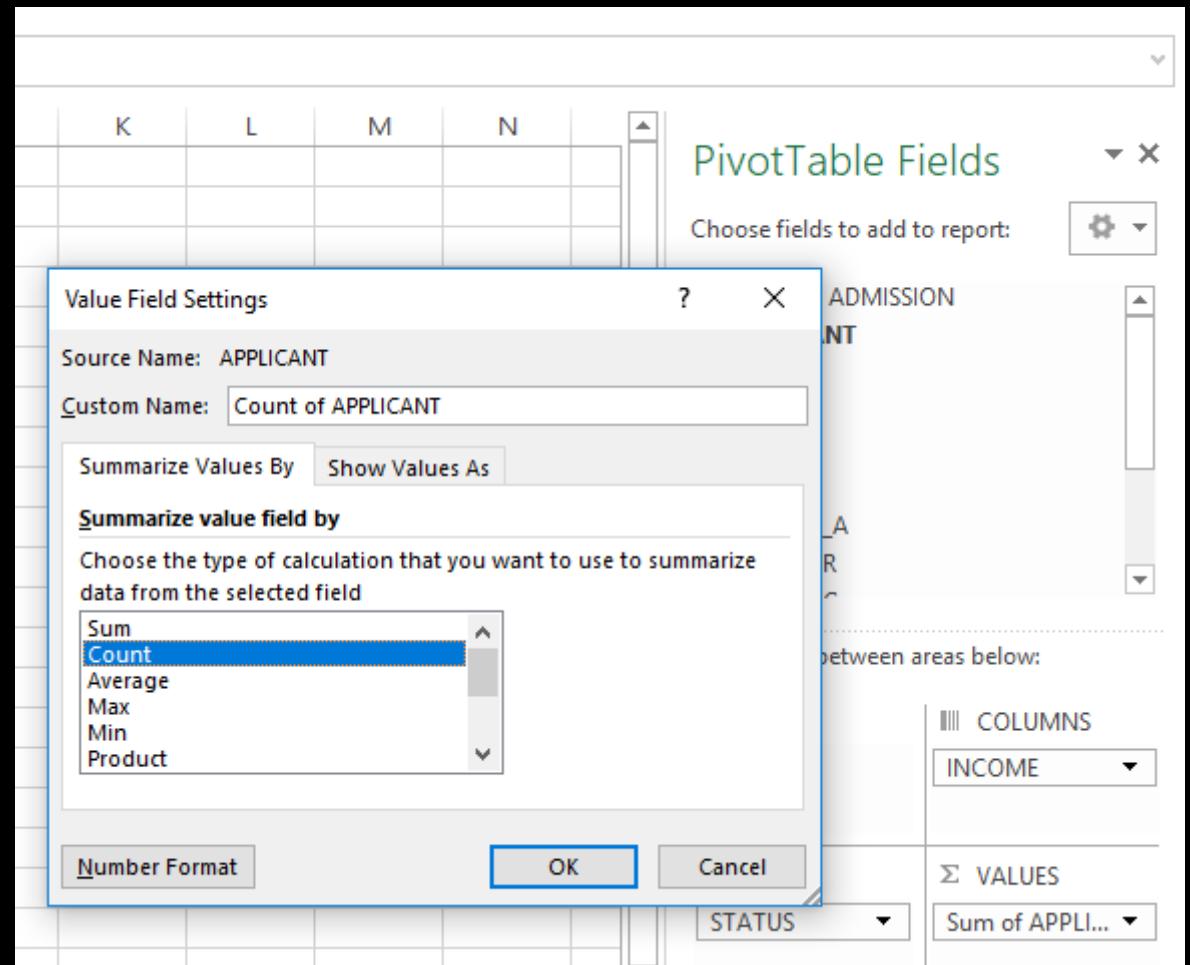
Click VALUE FIELD SETTINGS



Types of Data Analytics:

DIAGNOSTIC ANALYTICS

**STEP 6: Under the
SUMMARIZE VALUES
BY
Select
COUNT**



Types of Data Analytics:

DIAGNOSTIC ANALYTICS

A	B	C	D
Count of APPLICANT	Column Labels		
Row Labels	30K AND ABOVE	BELOW 30K	Grand Total
ENROLLED	85	115	200
NOT ENROLLED	346	454	800
Grand Total	431	569	1000

COUNT

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

**STEP 7: Click
SHOW VALUES AS
As well as the**



**then choose the
% needed in the report**

**e.g. % OF GRAND TOTAL
then click
OK**

The screenshot shows the 'Value Field Settings' dialog box in Microsoft Excel. The 'Source Name' is set to 'APPLICANT' and the 'Custom Name' is 'Count of APPLICANT'. The 'Summarize Values By' tab is selected. In the 'Show values as' dropdown, the option '% of Grand Total' is highlighted with a blue selection bar. Other options listed include 'No Calculation', '% of Grand Total', '% of Column Total', '% of Row Total', '% Of', '% of Parent Row Total', 'INCOME', and 'GPA_A'. At the bottom of the dialog are 'Number Format', 'OK', 'Cancel', and 'STATUS' buttons.

PivotTable Fields

Choose fields to add to report:

ADMISSION

IN

A

R

C

between areas below:

COLUMNS

INCOME

SUM VALUES

Sum of APPL...

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

A	B	C	D
Count of APPLICANT	Column Labels		
Row Labels	30K AND ABOVE	BELOW 30K	Grand Total
ENROLLED	85	115	200
NOT ENROLLED	346	454	800
Grand Total	431	569	1000

A	B	C	D
Count of APPLICANT	Column Labels		
Row Labels	30K AND ABOVE	BELOW 30K	Grand Total
ENROLLED	8.50%	11.50%	20.00%
NOT ENROLLED	34.60%	45.40%	80.00%
Grand Total	43.10%	56.90%	100.00%

% of Grand Total

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

A	B	C	D
Count of APPLICANT	Column Labels		
Row Labels	30K AND ABOVE	BELOW 30K	Grand Total
ENROLLED	85	115	200
NOT ENROLLED	346	454	800
Grand Total	431	569	1000

Count of APPLICANT	Column Labels		
Row Labels	30K AND ABOVE	BELOW 30K	Grand Total
ENROLLED	19.72%	20.21%	20.00%
NOT ENROLLED	80.28%	79.79%	80.00%
Grand Total	100.00%	100.00%	100.00%

% of Column Total

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

A	B	C	D
Count of APPLICANT	Column Labels		
Row Labels	30K AND ABOVE	BELOW 30K	Grand Total
ENROLLED	85	115	200
NOT ENROLLED	346	454	800
Grand Total	431	569	1000

Count of APPLICANT	Column Labels		
Row Labels	30K AND ABOVE	BELOW 30K	Grand Total
ENROLLED	42.50%	57.50%	100.00%
NOT ENROLLED	43.25%	56.75%	100.00%
Grand Total	43.10%	56.90%	100.00%

% of Row Total

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

**Other softwares can be used
that can somehow provide a
more complete output.**

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

COMBINED
Income does not affect Status applicants

STATUS * INCOME Crosstabulation

		INCOME		Total
		BELOW 30K		
STATUS	NOT ENROLLED	Count	454	346
		% within STATUS	56.8%	43.3%
		% within INCOME	79.8%	80.3%
		% of Total	45.4%	34.6%
	ENROLLED	Count	115	85
		% within STATUS	57.5%	42.5%
		% within INCOME	20.2%	19.7%
		% of Total	11.5%	8.5%
Total		Count	569	431
		% within STATUS	56.9%	43.1%
		% within INCOME	100.0%	100.0%
		% of Total	56.9%	43.1%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.037 ^a	1	.848		
Continuity Correction ^b	.012	1	.911		
Likelihood Ratio	.037	1	.848		
Fisher's Exact Test				.873	.456
Linear-by-Linear Association	.037	1	.848		
N of Valid Cases	1000				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 86.20.

b. Computed only for a 2x2 table

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

2017

Income does
not affect
the Status of
applicants

STATUS * INCOME Crosstabulation

STATUS	NOT ENROLLED		INCOME		Total
			BELOW 30K	30K AND ABOVE	
NOT ENROLLED		Count	19	131	150
		% within STATUS	12.7%	87.3%	100.0%
		% within INCOME	55.9%	60.6%	60.0%
		% of Total	7.6%	52.4%	60.0%
ENROLLED		Count	15	85	100
		% within STATUS	15.0%	85.0%	100.0%
		% within INCOME	44.1%	39.4%	40.0%
		% of Total	6.0%	34.0%	40.0%
Total		Count	34	216	250
		% within STATUS	13.6%	86.4%	100.0%
		% within INCOME	100.0%	100.0%	100.0%
		% of Total	13.6%	86.4%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.278 ^a	1	.598		
Continuity Correction ^b	.115	1	.735		
Likelihood Ratio	.276	1	.600		
Fisher's Exact Test				.707	.364
Linear-by-Linear Association	.277	1	.599		
N of Valid Cases	250				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.60.

b. Computed only for a 2x2 table

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

2018

Income does
affect the
Status of
applicants

STATUS * INCOME Crosstabulation

		INCOME		Total
		BELOW 30K		
STATUS	NOT ENROLLED	Count	226	184
		% within STATUS	55.1%	44.9%
		% within INCOME	85.0%	100.0%
		% of Total	50.2%	40.9%
ENROLLED	Count	40	0	40
		% within STATUS	100.0%	.0%
		% within INCOME	15.0%	.0%
		% of Total	8.9%	.0%
Total		Count	266	184
		% within STATUS	59.1%	40.9%
		% within INCOME	100.0%	100.0%
		% of Total	59.1%	40.9%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	30.369 ^a	1	.000		
Continuity Correction ^b	28.540	1	.000		
Likelihood Ratio	44.736	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	30.301	1	.000		
N of Valid Cases	450				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 16.36.

b. Computed only for a 2x2 table

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

2019

Income does
affect the
Status of
applicants

STATUS * INCOME Crosstabulation

STATUS	NOT ENROLLED		INCOME		Total
			BELOW 30K	30K AND ABOVE	
ENROLLED		Count	209	31	240
		% within STATUS	87.1%	12.9%	100.0%
		% within INCOME	77.7%	100.0%	80.0%
		% of Total	69.7%	10.3%	80.0%
Total		Count	60	0	60
		% within STATUS	100.0%	.0%	100.0%
		% within INCOME	22.3%	.0%	20.0%
		% of Total	20.0%	.0%	20.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.643 ^a	1	.003		
Continuity Correction ^b	7.305	1	.007		
Likelihood Ratio	14.704	1	.000		
Fisher's Exact Test				.001	.001
Linear-by-Linear Association	8.614	1	.003		
N of Valid Cases	300				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.20.

b. Computed only for a 2x2 table

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

Case Summaries

INCOME A

YEAR	STATUS	N	Mean	Std. Deviation
2017	NOT ENROLLED	150	60936.5400	23115.17748
	ENROLLED	100	61938.1200	24276.20059
	Total	250	61337.1720	23543.23122
2018	NOT ENROLLED	410	37270.2268	27725.78815
	ENROLLED	40	14374.8000	2765.18383
	Total	450	35235.0778	27266.26146
2019	NOT ENROLLED	240	21744.3958	19620.24299
	ENROLLED	60	14532.0833	2629.66889
	Total	300	20301.9333	17816.30745
Total	NOT ENROLLED	800	37049.9112	28032.03664
	ENROLLED	200	38203.6450	29375.05777
	Total	1000	37280.6580	28294.51772

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

Case Summaries

INCOME A

YEAR	STATUS	N	Mean	Std. Deviation
2017	NOT ENROLLED	150	60936.5400	23115.17748
	ENROLLED	100	61938.1200	24276.20059
	Total	250	61337.1720	23543.23122
2018	NOT ENROLLED	410	37270.2268	27725.78815
	ENROLLED	40	14374.8000	2765.18383
	Total	450	35235.0778	27266.26146
2019	NOT ENROLLED	240	21744.3958	19620.24299
	ENROLLED	60	14532.0833	2629.66889
	Total	300	20301.9333	17816.30745
Total	NOT ENROLLED	800	37049.9112	28032.03664
	ENROLLED	200	38203.6450	29375.05777
	Total	1000	37280.6580	28294.51772

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

Case Summaries

INCOME A

YEAR	STATUS	N	Mean	Std. Deviation
2017	NOT ENROLLED	150	60936.5400	23115.17748
	ENROLLED	100	61938.1200	24276.20059
	Total	250	61337.1720	23543.23122
2018	NOT ENROLLED	410	37270.2268	27725.78815
	ENROLLED	40	14374.8000	2765.18383
	Total	450	35235.0778	27266.26146
2019	NOT ENROLLED	240	21744.3958	19620.24299
	ENROLLED	60	14532.0833	2629.66889
	Total	300	20301.9333	17816.30745
Total	NOT ENROLLED	800	37049.9112	28032.03664
	ENROLLED	200	38203.6450	29375.05777
	Total	1000	37280.6580	28294.51772

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

Case Summaries

INCOME A

YEAR	STATUS	N	Mean	Std. Deviation
2017	NOT ENROLLED	150	60936.5400	23115.17748
	ENROLLED	100	61938.1200	24276.20059
	Total	250	61337.1720	23543.23122
2018	NOT ENROLLED	410	37270.2268	27725.78815
	ENROLLED	40	14374.8000	2765.18383
	Total	450	35235.0778	27266.26146
2019	NOT ENROLLED	240	21744.3958	19620.24299
	ENROLLED	60	14532.0833	2629.66889
	Total	300	20301.9333	17816.30745
Total	NOT ENROLLED	800	37049.9112	28032.03664
	ENROLLED	200	38203.6450	29375.05777
	Total	1000	37280.6580	28294.51772

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

**How many scholarships had
been granted per
School/Department/College per
School Year?**

Types of Data Analytics:

DIAGNOSTIC ANALYTICS

YEAR * SCHOLAR * SCHOOL/DEPT/COLLEGE START Crosstabulation

Count

SCHOOL/DEPT/COLLEGE START			SCHOLAR		Total
			NO	YES	
1.00	YEAR	2017	14	11	25
		2018	3	8	11
		2019	5	1	6
	Total		22	20	42
2.00	YEAR	2017	11	13	24
		2018	2	2	4
		2019	2	8	10
	Total		15	23	38
3.00	YEAR	2017	7	9	16
		2018	2	4	6
		2019	8	8	16
	Total		17	21	38
4.00	YEAR	2017	10	7	17
		2018	8	1	9
		2019	6	9	15
	Total		24	17	41
5.00	YEAR	2017	11	7	18
		2018	5	5	10
		2019	7	6	13
	Total		23	18	41

Types of Data Analytics:

PREDICTIVE ANALYTICS

**moves to what is likely going
to happen in the near term**

PREDICTIVE ANALYTICS

TREND ANALYSIS

A trend analysis is a method of analysis that allows researchers to **predict** what will happen in the future. Trend analysis is based on historical data about the performance given the overall trends of a data and particular indicators within the data.

Types of Data Analytics:

PREDICTIVE ANALYTICS

FIRST YEAR	2012-2013	2013-2014	2014-2015	2015-2016
FIRST SEM	1457	1492	1374	1265
SECOND SEM	1026	1042	945	924
SUMMER	100	112	65	154

Types of Data Analytics:

PREDICTIVE ANALYTICS

FIRST YEAR	2012-2013	2013-2014	2014-2015	2015-2016
FIRST SEM	1457	1492	1374	1265

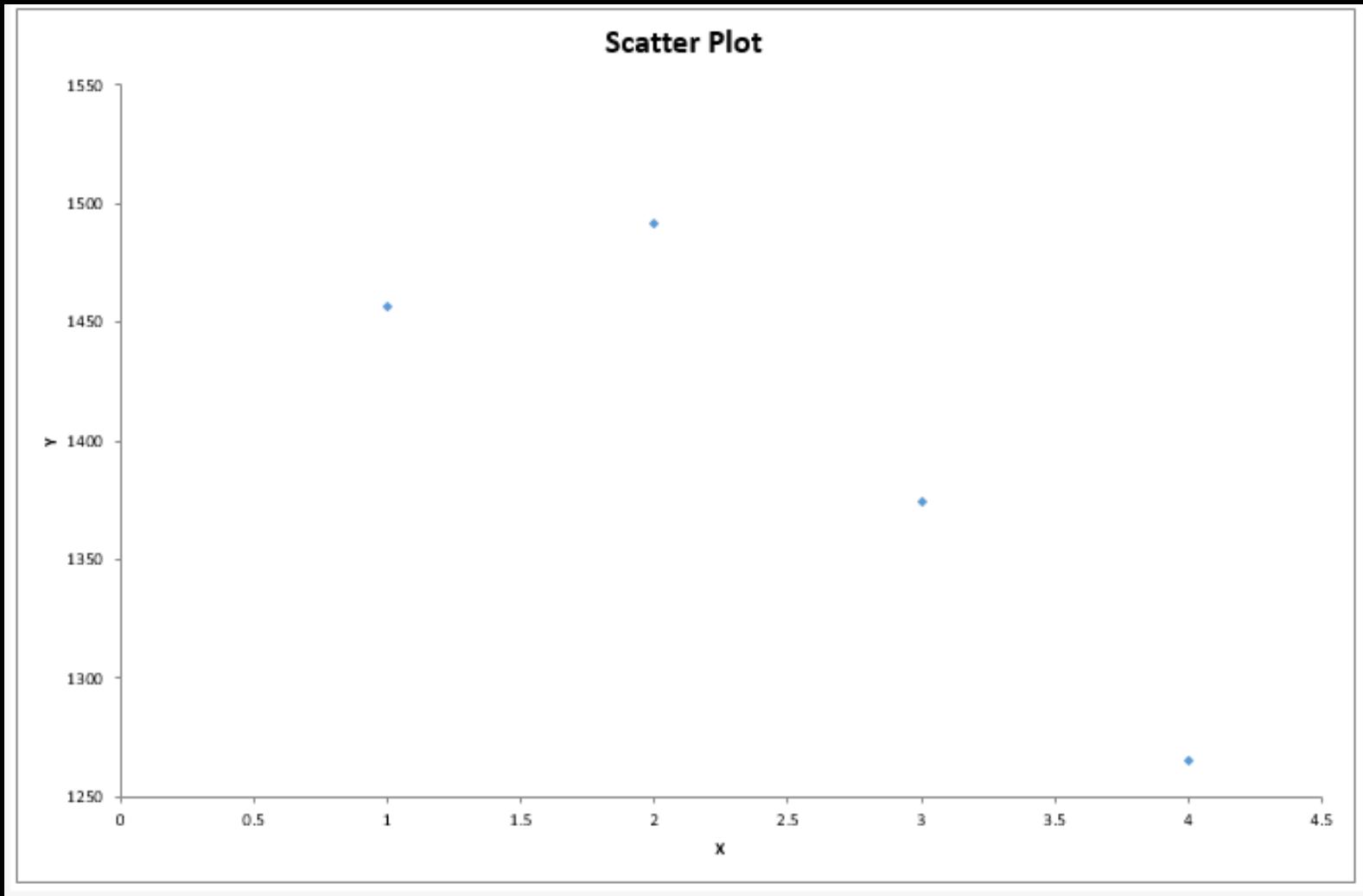
Types of Data Analytics:

PREDICTIVE ANALYTICS

X (School Year)	Y (enrolees)
1 (2012-2013)	1457
2 (2013-2014)	1492
3 (2014-2015)	1374
4 (2015-2016)	1265

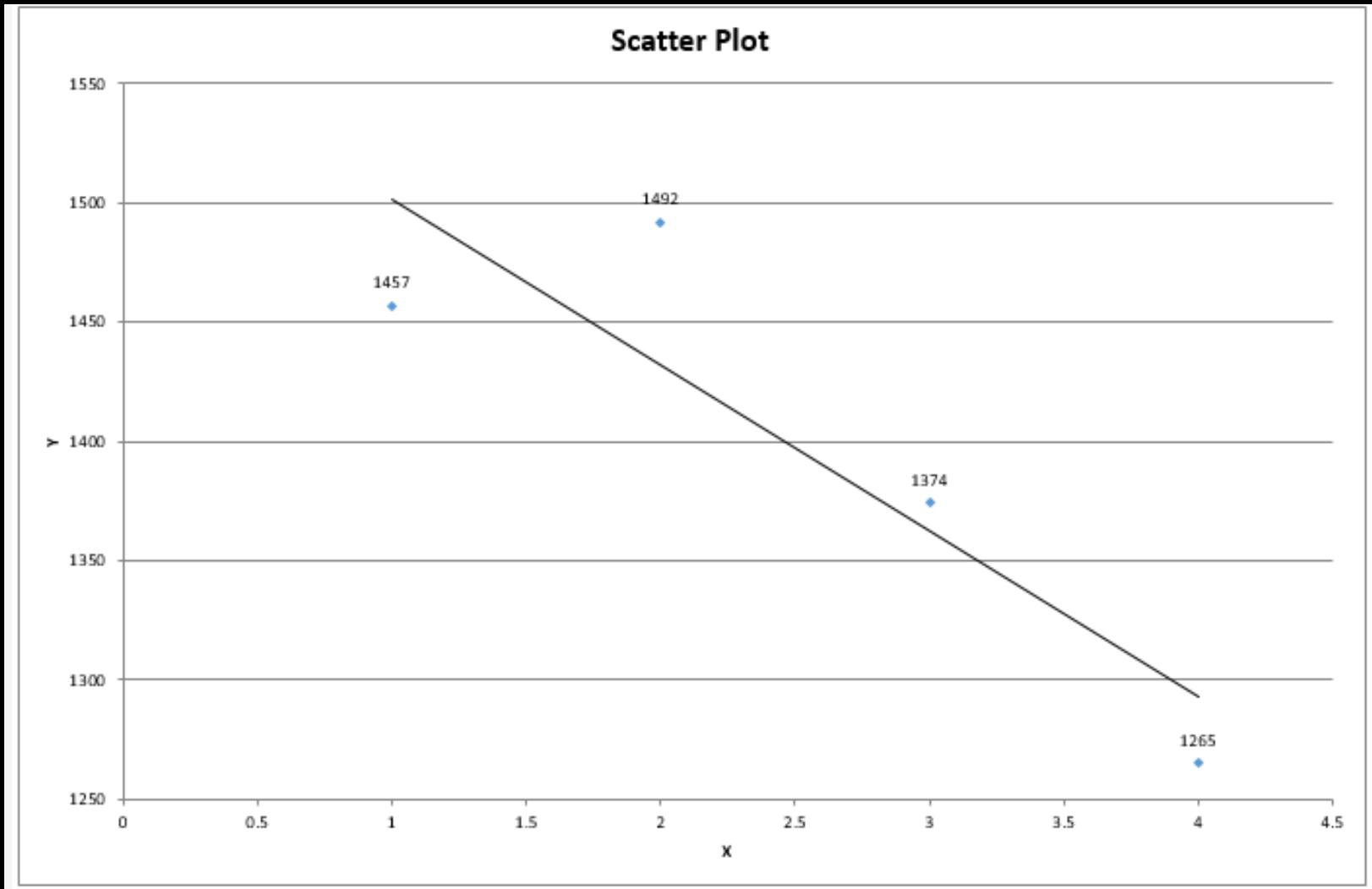
Types of Data Analytics:

PREDICTIVE ANALYTICS



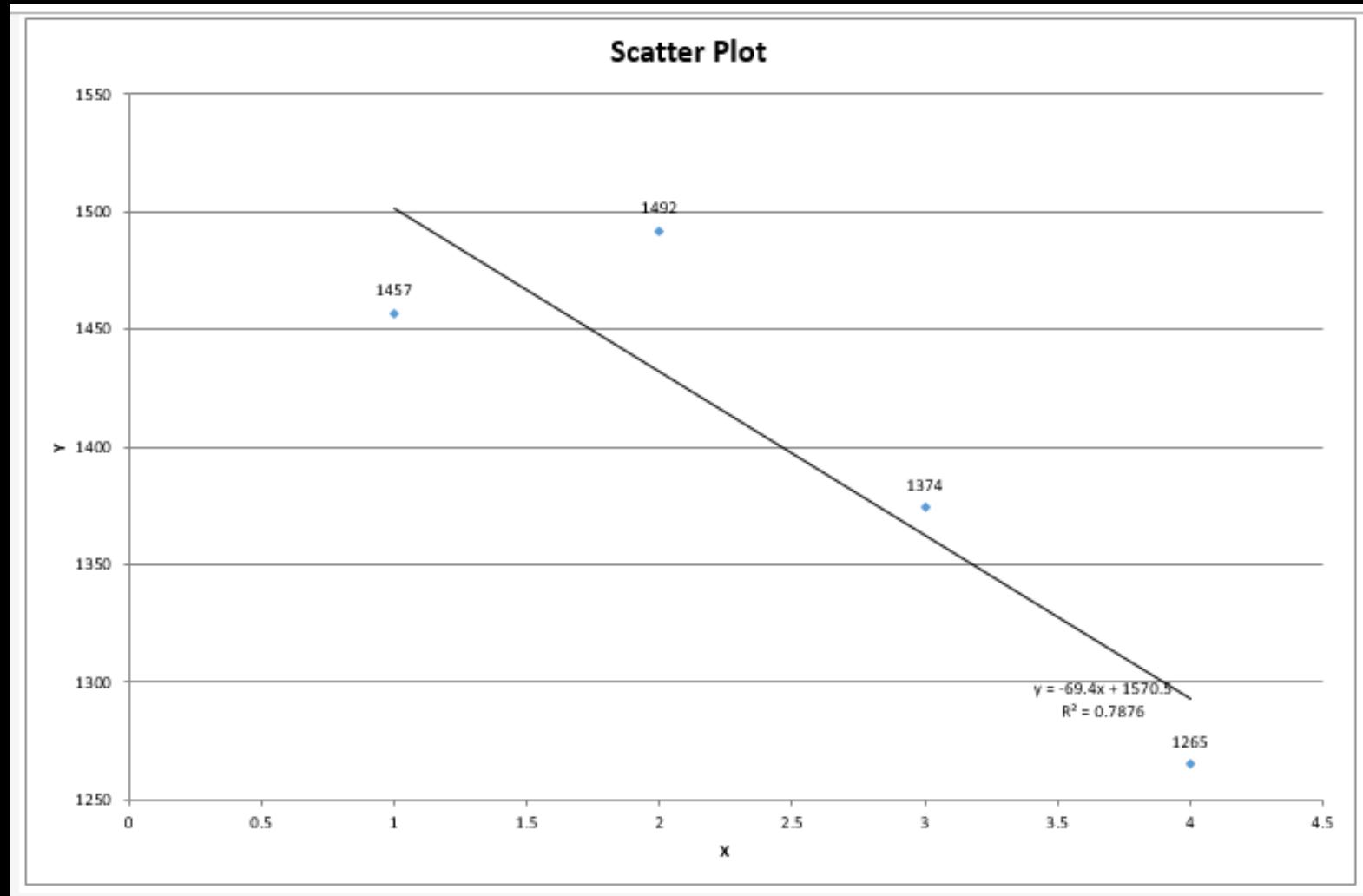
Types of Data Analytics:

PREDICTIVE ANALYTICS



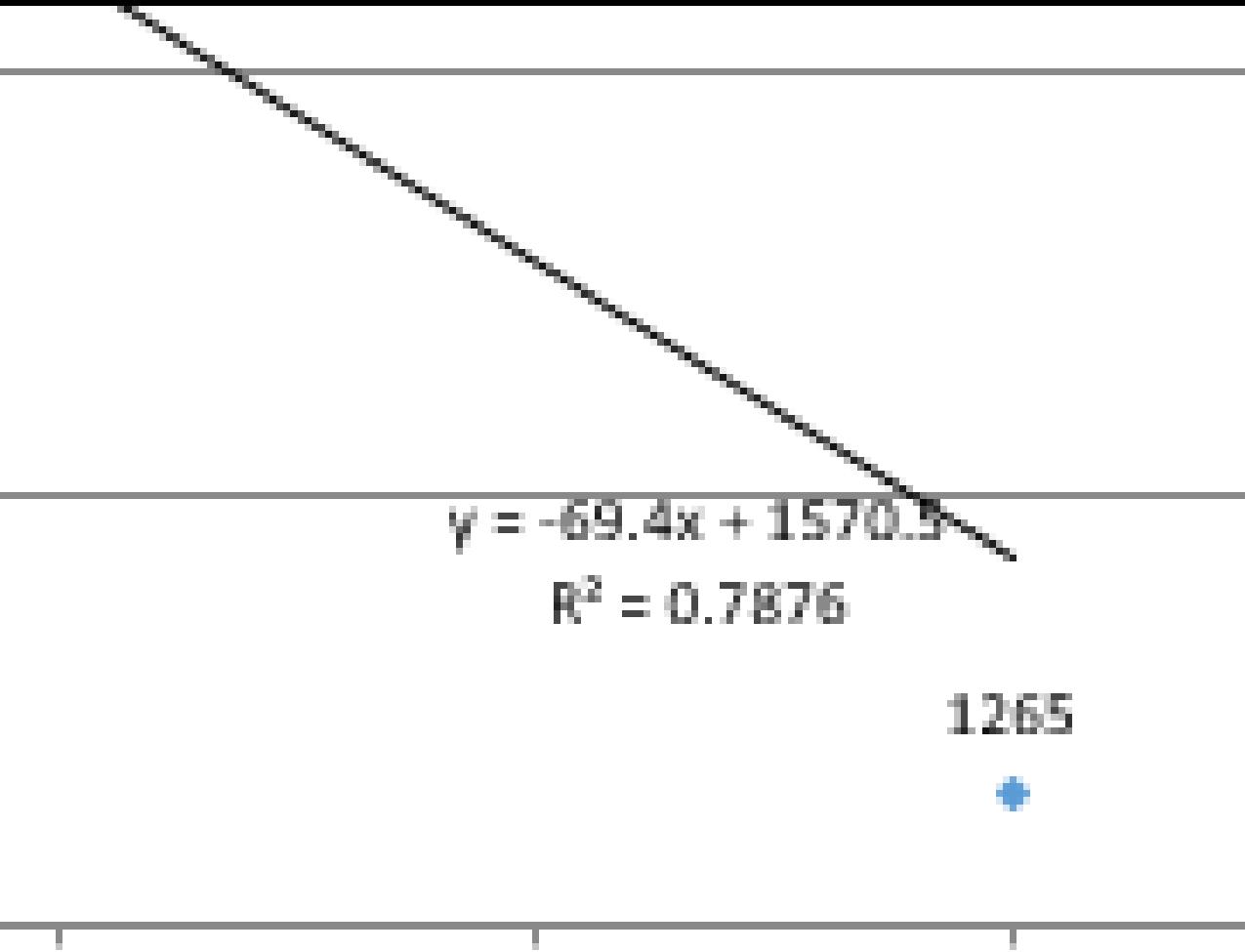
Types of Data Analytics:

PREDICTIVE ANALYTICS

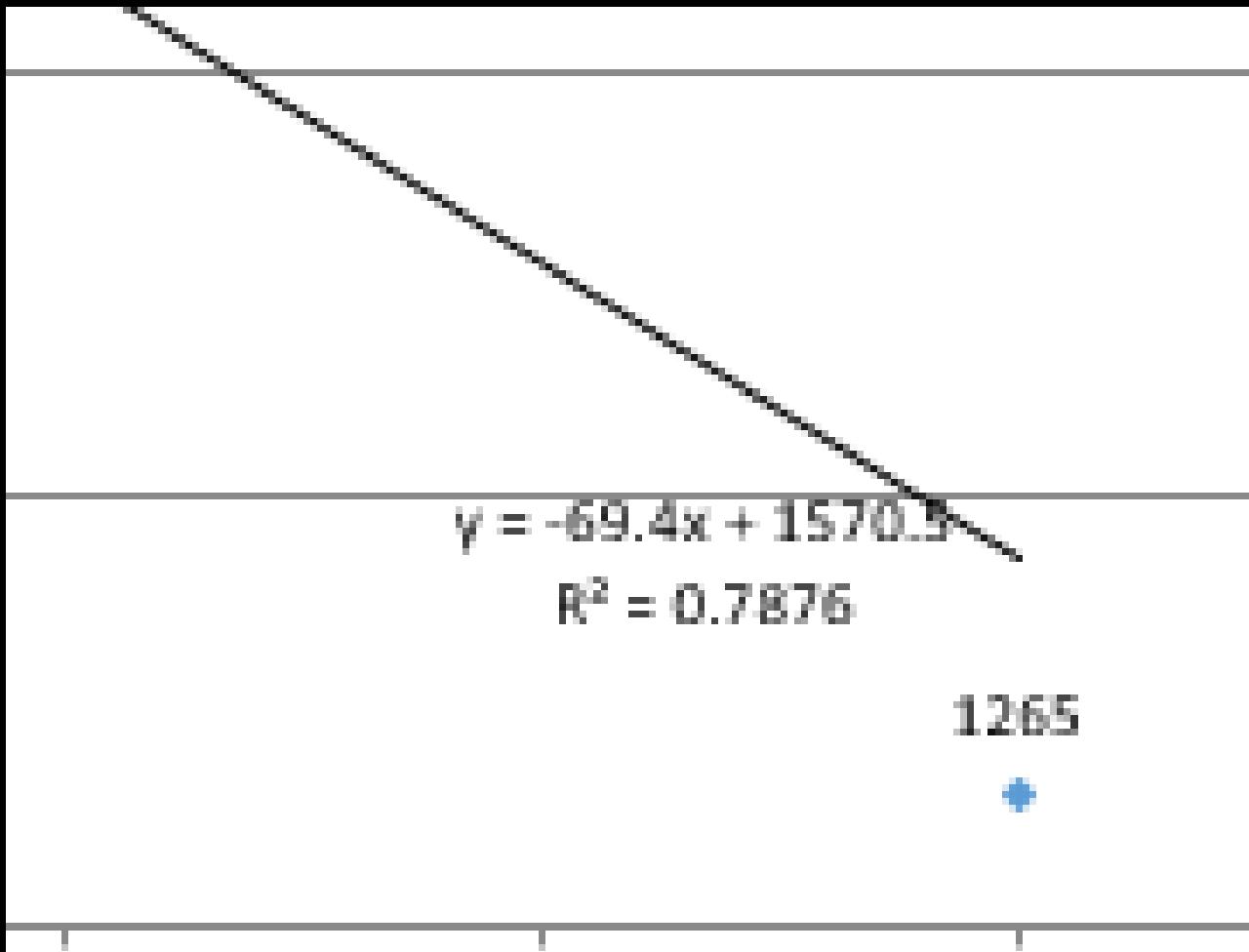


Types of Data Analytics:

PREDICTIVE ANALYTICS



PREDICTIVE ANALYTICS



$$\begin{aligned}y &= -69.4x + 1570.5 \\x &= 5 \quad (\text{the following year}) \\y &= -69.4(5) + 1570.5 \\y &= 1223.5 \quad (\text{predicted enrollment})\end{aligned}$$

Types of Data Analytics:

PREDICTIVE ANALYTICS

X (School Year)	Y (enrollees)
1 (2012-2013)	1457
2 (2013-2014)	1492
3 (2014-2015)	1374
4 (2015-2016)	1265
5 forecast	1224

Types of Data Analytics:

PREDICTIVE ANALYTICS

How do we distribute it by
department
college
school?

Types of Data Analytics:

PREDICTIVE ANALYTICS

FIRST YEAR	2012-2013	2013-2014	2014-2015	2015-2016	FORECAST
FIRST SEM	1457	1492	1374	1265	1224
SBMA	384	385	392	334	329
SCST	177	172	121	155	133
SE	434	460	397	409	375
SEAS	247	246	276	211	217
SN	18	11	11	8	9
STHM	197	218	177	148	160
SBMA	26%	26%	29%	26%	27%
SCST	12%	12%	9%	12%	11%
SE	30%	31%	29%	32%	31%
SEAS	17%	16%	20%	17%	18%
SN	1%	1%	1%	1%	1%
STHM	14%	15%	13%	12%	13%

Types of Data Analytics:

PREDICTIVE ANALYTICS

FIRST
Breakdown by
school/college/
department

FIRST YEAR	2012-2013	2013-2014	2014-2015	2015-2016	FORECAST
FIRST SEM	1457	1492	1374	1265	1224
SBMA	384	385	392	334	329
SCST	177	172	121	155	133
SE	434	460	397	409	375
SEAS	247	246	276	211	217
SN	18	11	11	8	9
STHM	197	218	177	148	160
SBMA	26%	26%	29%	26%	27%
SCST	12%	12%	9%	12%	11%
SE	30%	31%	29%	32%	31%
SEAS	17%	16%	20%	17%	18%
SN	1%	1%	1%	1%	1%
STHM	14%	15%	13%	12%	13%

SECOND
Percentage
distribution

FOURTH
Breakdown of
forecast

THIRD
3 year moving
average

Types of Data Analytics:

PREDICTIVE ANALYTICS

A	B	C	D	E	F	G	H	I	J	K	L	M
X	Y		SUMMARY OUTPUT									
1	1457											
2	1492		Regression Statistics									
3	1374		Multiple R	0.887442								
4	1265		R Square	0.787553								
FIRST SEM			Adjusted R Square	0.68133								
	1223.5		Standard Error	56.9921								
			Observations	4								
			ANOVA									
				df	SS	MS	F	Significance F				
			Regression	1	24081.8	24081.8	7.414119	0.112558				
			Residual	2	6496.2	3248.1						
			Total	3	30578							
				Coefficients	standard Err.	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%	
			Intercept	1570.5	69.80079	22.49975	0.00197	1270.171	1870.829	1270.171	1870.829	
			X	-69.4	25.48764	-2.72289	0.112558	-179.064	40.26448	-179.064	40.26448	
								1223.5				

Types of Data Analytics:

PREScriptive ANALYTICS

suggests a course of action

PRESCRIPTIVE ANALYTICS

What are the other factors that may affect enrollment?

- **Marketing expense?**
- **Number of school campaign visits?**

PRESCRIPTIVE ANALYTICS

CORRELATION ANALYSIS

A correlation is simply defined as a relationship between two variables. The whole purpose of using correlations in research is to figure out which variables are connected.

Types of Data Analytics:

PRESCRIPTIVE ANALYTICS

CORRELATION ANALYSIS

CORRELATION COEFFICIENT

r

has a value from -1 to 1

it measures the DEGREE of relationship

PRESCRIPTIVE ANALYTICS

**Marketing expense
MARKET (in millions)**

**Number of school
campaign visits
VISITS**

ENROLL	MARKET	VISITS
1457	0.82	34
1492	1.1	45
1374	0.72	38
1265	0.53	28
1245	0.67	34
1138	0.33	23
1465	0.56	40
2235	1.9	56

PRESCRIPTIVE ANALYTICS

Correlations

[DataSet1]

Correlations

		ENROLL	MARKETING	SCHOOL_VISITS
ENROLL	Pearson Correlation	1	.954**	.905**
	Sig. (2-tailed)		.000	.002
	N	8	8	8
MARKETING	Pearson Correlation	.954**	1	.911**
	Sig. (2-tailed)	.000		.002
	N	8	8	8
SCHOOL_VISITS	Pearson Correlation	.905**	.911**	1
	Sig. (2-tailed)	.002	.002	
	N	8	8	8

**. Correlation is significant at the 0.01 level (2-tailed).

PRESCRIPTIVE ANALYTICS

Nonparametric Correlations

[DataSet1]

Correlations

		ENROLL	MARKETING	SCHOOL_VISITS
Spearman's rho	ENROLL	Correlation Coefficient	.810*	.922**
		Sig. (2-tailed)	.015	.001
		N	8	8
	MARKETING	Correlation Coefficient	.810*	.814*
		Sig. (2-tailed)	.015	.014
		N	8	8
	SCHOOL_VISITS	Correlation Coefficient	.922**	.814*
		Sig. (2-tailed)	.001	.014
		N	8	8

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Types of Data Analytics:

PRESCRIPTIVE ANALYTICS

**Which among the math courses
in the curriculum contributes
in the subject Math and
Science of ME Board Exam?**

Types of Data Analytics:

PRESCRIPTIVE ANALYTICS

		college algebra	advanced algebra	spherical trigonometry	analytic geometry	differential calculus	integral calculus	solicitation	duration	differentiation	equation solving	probability and statistics
mathandscience	Pearson Correlation	.435 **	.478 **	.314	.274	.124	.433 **	.363 *	.367 *	.367 *	.106	
	Sig. (2-tailed)	.008	.003	.066	.091	.435	.004	.041	.015	.500		

Types of Data Analytics:

PRESCRIPTIVE ANALYTICS

**Which professional subjects in
the curriculum contributes
most in the ME Board Exam?**

Types of Data Analytics:

PREScriptive ANALYTICS

		overallt	machin	machin	machin	machin	machin	mechan	industri	industri	melaws	powerpl	powerpl					
		t	1	1lab	2	2lab	ab3	icalengi	alprante	alplante	ethicsco	antengi	antengi	casestu				
		estresul	edesign	edesign	edesign	edesign	neeringl	ngandd	ngineeri	ngineeri	desands	neering	nerringa	diproje				
overalltest	Pearson Correlation		1	-.078	-.081	.201	.190	.247	.160	.123	.258	.370 *	.258	.188	.042	.152	.317 *	.092
	Sig. (2-tailed)			.615	.603	.191	.218	.110	.298	.425	.090	.013	.090	.222	.786	.325	.036	.551

For Analysis:

		FRESHMEN				GRADUATES					
		CAS	COM	ENG	TOTAL	CAS	COM	ENG	TOTAL		
19	83	—	84	106	173	0	279	34	92	0	126
19	84	—	85	80	113	188	381	44	82	71	197
19	85	—	86	92	154	122	368	46	79	62	187
19	86	—	87	131	265	244	640	43	133	67	243
19	87	—	88	172	295	345	812	63	133	113	309
19	88	—	89	154	237	450	841	50	155	117	322
19	89	—	90	168	197	429	794	81	103	131	315
19	90	—	91	158	226	361	745	56	90	146	292
19	91	—	92	154	239	316	709	52	99	111	262
19	92	—	93	145	275	318	738	67	126	130	323
19	93	—	94	156	291	296	743	74	171	110	355
19	94	—	95	153	327	391	871	71	181	123	375
19	95	—	96	135	288	453	876	76	152	169	397
19	96	—	97	263	389	350	1002	153	214	121	488
19	97	—	98	255	397	481	1133	144	205	180	529
				2322	3866	4744	10932	1054	2015	1651	4720

For Analysis:

SBMA (1 ST SEM)	2012-2013	2013-2014	2014-2015	2015-2016	2012-2013
1 YEAR	384	385	392	334	1495
2 YEAR	279	360	223	257	1119
3 YEAR	207	201	283	296	987
4 YEAR	102	109	113	149	473
5 YEAR	7	8	26	1	42

Summary:

Types of Data Analytics:

DESCRIPTIVE ANALYTICS

DIAGNOSTIC ANALYTICS

PREDICTIVE ANALYTICS

PRESCRIPTIVE ANALYTICS

Summary:

DATA ANALYTICS

Is the Science of analyzing **data in order to
make **conclusions** about that information**

DATA ANALYTICS