Correlational Analysis of Academic Parameters of First Year Engineering Students For Evaluation of Significance To Final Result

Minakshi S. Panchal[1], Dhanashree Munot[2], Vrushali Kamble[3], Dhiraj Wakharde[4]
[1, 2, 3, 4] Dept of Applied Science & Humanities
[1, 2, 3, 4] PimpriChinchwad College of Engineering,

Abstract- *We aim to study the association between observed variables and estimate the strength of the relationship. Research objectives is to quantitatively address by correlation analysis, which provides information about not only strength but also the direction of a relationship. In this paper we studied degree of correlationsbetween the SGPA with some factors affecting it most. The study will be helpful to improve the students SGPA by identifying the factors with strong positive relation.*

Keywords- Correlation, Pearson correlation coefficient, Kendall rank correlation, Spearman rank correlation, statistical analysis, dataset.

I. INTRODUCTION

The concept of correlation was first proposed by Sir Francis Galton in 1894, which was further mathematically described by Karl Pearson in 1896. Correlation, one of the most important statistical techniques which tests the relationship and association between variables is employed in scientific studies.

In statistical terms, correlation is a method of evaluating an anticipated two-way linear association between two quantifiable uninterrupted variables. The extent of "correlation" is measured by a statistic called the correlation coefficient, which represents the strength of the presumed linear association between the two selected variables. It is observed as to how closely two variables covary. It is an immeasurable quantity whose value can vary from −1 (Perfectly negative correlation) through 0 (no correlation) to +1 (perfect positive correlation). Direct relation is portrayed by a positive coefficient of correlation, which means that as the value of one variable increases, the value of the other variable also tends to increase.Any other configuration of relation between two incessant variables that is not linear is not considered as correlation.

The SGPA (Semester Grade Point Average) of students is an important factor for deciding their engineering results. This SGPA depends on various parameters including unit tests, attendance, academic marks, HSC exam result, CET results, etc. So, we are trying to find the relation between SGPA and these parameters, as to how strongly/weakly these parameters affect the SGPA using correlation technique.

II. LITERATURE REVIEW

1. Wouroud ELFARMAWI [2019] – This research was done to find the correlation between Customer Relationship Management (CRM) system, product innovation and satisfaction of customers showed that CRM is an effective aid for businesses to grow. The research revealed that CRM systems and customer satisfaction shows a strong correlation trend, so this will help companies to retain existing customers and attract new ones.

2. DR. T S R MURTHY and D. SIVA RAMA KRISHNA[2011] – The authors focussed on use of cell phones and tried to analyse its usage using correlation techniques. Studywas done to find out the correlation between different demographics of a population upon their cell phone usage viz. men and women, different age groups dependency on cell phone, usage of phones with respect to occupation of people and brand market. The research findings revealed that different usage patterns were observed with regards to different types of population.

3. Galadanci, B. S., Mukhtar, M.I. [2017] – This paper speaks about the correlational study conducted in Bayero University, Kano, Nigeria. The study was mainly focussed on understanding if there was any correlation between theoretical and practical examination scores of students. It was found that there existed a statistically significant positive correlation between scores of students in theoretical and practical examinations. It was also observed that correlation findings during

4. Irene L. Balmes [2017] – This was an interesting study done to find the correlation between a student's mathematical and programming ability, who opted for the BSCS program. This study was made to find the extent of relation between these two abilities rather than to check their dependencies on one another. One important conclusion was that the scores of students in maths can be used to predict whether the student can pass in programming courses under BSCS, which will further be a prominent factor for accepting students under BSCS program.

5. VikasMahalawat and Dr. Bharti Sharma [2018] – The authors tried to find out if there existed any relationship between business process management and organizational performance.. The conclusions were drawn based on a correlation matrix which showed that there is a positive correlation between BPRI and QI, BPROS and CS, BPROP and PPM and also between BPRTI and EB.

6. Eze, Felix John, Odigbo, Benedict Ejikeme and Ufot, Juliet Alfred [2015] – The authors jointly carried out a research to study the relationship between business location and customer patronage with an objective to find out its connotation on policies of business and their decision making process. The findings interpreted that even if a business launches a great product but still its location matters for consumers.

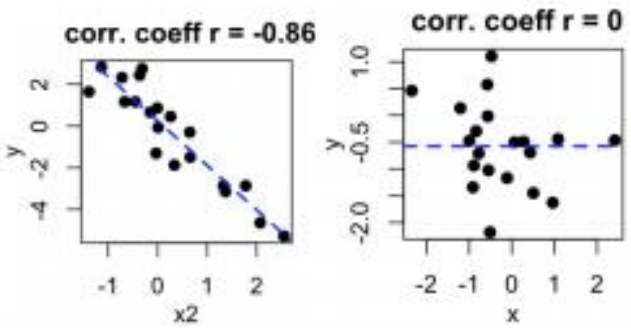
7. Suniti Yadav [20] – This paper talks about an experienced study in biological studies where correlation is used for understanding the relation between two variables and controlling the effects of covariates. In biological studies, the independent variables are predictor variables and outcomes to be seen are dependent variables. So to study this dependency between two variables i.e. risk factors and outcomes, correlation is considered to be the best measure, as it provides a quantitative way for measuring the strength of relation between selected variables. Further, the mathematical background of correlation was explained by the author and its use in biological studies was explained.

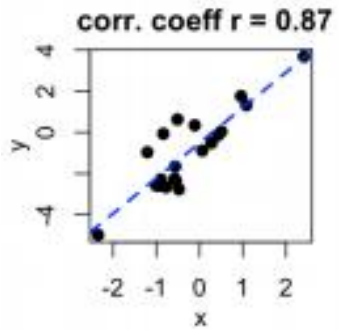
III. MATHEMATICAL BACKGROUND

Francis Galton's invention of correlation, a measure that demonstrates a linear relationship between two variables. The correlation coefficient can indicate any values from -1 to 1, where -1 indicating perfect negative correlation and 1 indicating perfect positive correlation.

Interpretation of correlation coefficient :

1. Negative Correlation
2. No Correlation
3. Positive Correlation





The **Pearson correlation coefficient** indicates the degree of the relationship between linearly related variables. It is known as the best method of measuring the association betweenvariables of interest because it is based on the method of covariance. It gives informationabout the magnitude of the association, or correlation, as well as the direction of therelationship. It is determined by using the formula:

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

1. r_{xy} – the correlation coefficient of the linear relationship between the variables x and y
2. x_i – the values of the x-variable in a sample
3. $\bar{x}$ - the mean of the values of the x-variable
4. y_i - the values of the y-variable in a sample
5. $\bar{y}$ - the mean of the values of the y-variable

Kendall rank correlation is a non-parametric test that measures the strength of dependence between two variables. If we consider two samples, a and b, where each sample size is *n*, we know that the total number of pairings with a b is *n*(*n*-

1)/2.The following formula is used to calculate the value of Kendall rank correlation:

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)}$$

Where, n_c – number of concordant (ordered in the same way)

n_d – number of discordant (ordered differently)

Spearman rank correlation is a non-parametric test that is used to measure the degree associated between two variables. This test does not carry any assumptions about the distribution of the data and is the appropriate correlation analysis when the variables are measured on a scale that is at least ordinal. The formula to calculate the Spearman rank correlation is :

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

Where, ρ = Spearman rank correlation,d_i = difference between the ranks of corresponding variables
n = number of observations

IV. PROPOSED WORK

The performance of a student in a semester is given by the Semester Grade Point Average(SGPA). It is the weighted average of the grade points obtained in all the courses, seminarsand projects registered by the student during the semester. SGPA is a tool used for calculatingresults of students.

This is a study research to find out the factors affecting this SGPA .We considered data from of randomly selected students of First year of PCCoE for one academic year.We considered the factors as entranceexams (CET, JEE), 12th board exams, internal exams , attendance and results of individual subjects in external exam.

Correlation matrix for parameters affecting SGPA

	Attendance	UT-1	UT-2	CET	HSC	CM-I	Phy	FPL	BEX	BCE	EG1	Total SGPA
Attendance	1											
UT-1	0.630	1.000										
UT-2	0.669	0.841	1.000									
CET	0.306	0.738	0.765	1.000								
HSC	0.456	0.755	0.756	0.770	1.000							
CM-I	0.413	0.820	0.779	0.824	0.811	1.000						
Phy	0.396	0.819	0.773	0.783	0.783	0.854	1.000					
FPL	0.423	0.696	0.691	0.627	0.641	0.692	0.665	1.000				
BEX	0.618	0.829	0.796	0.753	0.803	0.853	0.831	0.709	1.000			
BCE	0.441	0.778	0.884	0.817	0.881	0.866	0.886	0.866	0.789	1.000		
EG1	0.602	0.730	0.710	0.678	0.648	0.778	0.838	0.609	0.709	0.711	1.000	
Total SGPA	0.612	0.880	0.817	0.805	0.871	0.897	0.888	0.827	0.931	0.827	0.894	1.000

V. RESULTS AND ANALYSIS

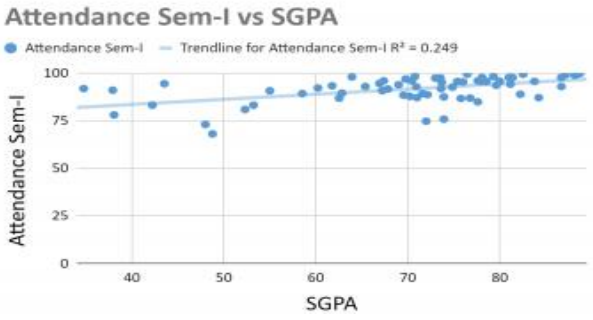
Correlation table:

Correlation value	Conclusion for relationship
0.80 to 1	Very strong relationship
0.60 to 0.79	Strong
0.40 to0.59	Moderate
0.20 to 0.39	Weak

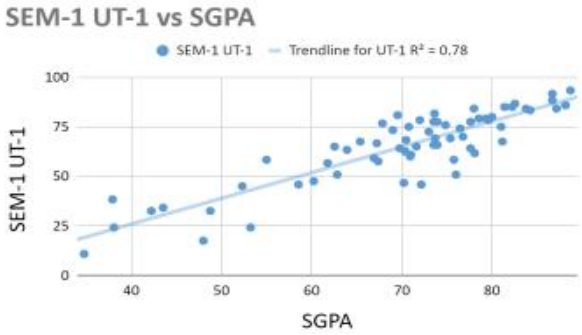
As seen from the scatter plots the points are scattered around the straight line which impliesthat the correlation is linear.

The positive values implies that there is positive correlation and the pair of x and y lie on thestraight line with a positive slope.

● The correlation between the attributes A1 and A22 is 0.4990 which implies that A1and A22 have a moderate correlation. That means Attendance Sem - I has moderateeffect on SGPA.



● The correlation between the attributes A3 and A22 is 0.8834 which indicates that Unittest - I results have a very strong correlation with SGPA.



VI. CONCLUSION

Correlation is an important concept that helped us to find the degree of relation betweenvariables. This is a study research related to factors affecting SGPA(Semester Grade PointAverage) . The tool used to serve this purpose is correlation.From the correlation table as seen in figure 1.2 it is observed that:

- HSC marks play an important role in determining SGPA.
- CET marks are also a deciding factor when it comes to SGPA results.
- Unit test results also share a strong relation with SGPA.
- Some of the subject results have moderate to very strong correlation with SGPA.
- Attendance of students has

REFERENCES

[1] Wouroud ELFARMAWI, 2019, “Correlation Between Customer RelationshipManagement System Usage, Product Innovation, And Customer Satisfaction”,Foundations of Management, Vol. 11, eISSN 2300-5661, pISSN 2080-7279, DOI:10.2478/fman-2019-0002

[2] DR. T S R MURTHY and D. SIVA RAMA KRISHNA, 2011, “Analysis Of CellPhone Usage Using Correlation Techniques”, International Journal of Wireless &Mobile Networks (IJWMN) Vol. 3, No. 2

[3] Galadanci, B. S., Mukhtar, M.I., 2017, “A Correlational Study Of Students'Theoretical And Practical Examinations Scores In Computer Applications Courses InBayero University Kano”, Science World Journal Vol 12(No 2), ISSN 1597-6343

[4] Irene L. Balmes, 2017, “Correlation of Mathematical Ability and ProgrammingAbility of the Computer Science Students”, Asia Pacific Journal of Education, Artsand Sciences, Vol. 4 No. 3, 85-88, P-ISSN 2362-8022, E-ISSN 2362-8030

[5] VikasMahalawat and Dr. Bharti Sharma, 2018, “Correlation between BusinessProcess Management and Organizational Performance: A Study of Bank X”,International Journal of Engineering and Management Research, Volume-8, Issue-2,Page Number: 209-213, ISSN (ONLINE): 2250-0758, ISSN (PRINT): 2394-6962

[6] Eze, Felix John, Odigbo, Benedict Ejikeme and Ufot, Juliet Alfred, 2015, “TheCorrelation between Business Location and Consumers Patronage: Implications forBusiness Policy Decisions”, British Journal of Economics, Management & Trade8(4): 294-304, Article no.BJEMT.2015.118, ISSN: 2278-098X, SCIENCEDOMAINinternational, DOI: 10.9734/BJEMT/2015/16998

[7] Suniti Yadav, “Correlation Analysis in Biological Studies”, Journal of the Practiceof Cardiovascular Sciences, Volume 4, Issue 2, May-August 2018