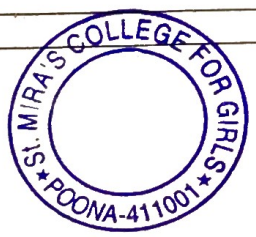
St. Mira's College for Girls, Pune

(Autonomous-Affiliated to Savitribai Phule Pune University)

1.1.3 - List of Activities for Courses having focus on employability/ entrepreneurship/ skill development offered by the institution during the last five years

Sr. No.	Name of the Course	Course Code
	MSC	
1	Cloud computing	MSE21905
2	Practical paper based on Software Architecture and Design Pattern, Machine Learning & Web Frameworks	MS32004
	MCOM	
3	Introduction to Behavioural Finance	MCM32004
4	Capital Markets and Financial Services	MCM42001
5	Project Work (Business Administration)	MCM42004
6	Project Work (Accountancy)	MCM42006
	MA ECONOMICS	
7	Research Project	MEC42003
	FYBCOM	
8	Compulsory English	AC#12001
9	Compulsory English	AC#22001
10	Business Mathematics and Statistics	AC12001
11	Business Mathematics and Statistics	AC22001
12	Consumer Protection & Business Ethics	C12006
13	Consumer Protection & Business Ethics	C22006
14	Business Administration	AC12008
15	Business Administration	AC22008
	SYBCOM	
16	Business Administration – I	C31608
17	Business Administration – I	C41608
	TYBCOM	
18	Business Administration - II	C51708
19	Business Administration - II	C61708
20	Marketing - III	C51713
21	Marketing - III	C61713
	FYBA	
22	Comp. English EM	A12001
23	Optional English	A12005
24	Optional English	A22005
	SYBA	
25	Comp English-EM	A31601
26	Comp English-EM	A41602
27	Comp English-MM	A31602
28	Comp English-MM	A41602
29	English-1	A31614
30	English-2	A41614
	TYBA	
31	Comp English EM	A51701
32	Comp English EM	A61701
33	Comp English MM	A51702
34	Comp English MM	A61702
35	Sociology Special-3	A51712
36	English-3	A51715
37	English-4	A61715
38	English -4	A51716
39	English -5	A61716
40	Psychology Special -3	A51718



St. Mira's College for Girls, Pune

(Autonomous-Affiliated to Savitribai Phule Pune University)

1.1.3 - List of Activities for Courses having focus on employability/ entrepreneurship/ skill development offered by the institution during the last five years

Sr. No.	Name of the Course	Course Code
	FYBSC	
41	Discrete Mathematics	BS12003
42	Principles of Digital Electronics	BS12008
43	Electronics Practical	BSP12012
44	Graph Theory	BS22003
	SYBSC	
45	Mathematics - Numerical Analysis	BS31604
46	Mathematics - Operations Research	BS41604
47	Digital System Design	BS31605
48	PIC Microcontroller Architecture, Interfacing & Programming	BS41605
	TYBSC	
49	Operating Systems	BS61701
50	Object oriented analysis and design	BS51706
51	Advanced database management system	BS61706
52	Lab Course III	BSP61709
	FYBBA	
53	Principles of Management	BB22001
	TYBBA	
54	Entrepreneurship Development	BB51702
55	Specialisation - Finance	BB61706A
56	Specialisation - HR	BB61706B
57	Specialisation - Marketing	BB61706C
	FYBBA(CA)	
58	Business Communication Skills	BC12001
59	Principles of Management	BC12002
60	C Language	BC12003
	SYBBA(CA)	
61	Software Engineering	BC31605
62	Opertaing System	BC31603
	TYBBA(CA)	
63	Web technology	BC51702
64	Project	BC51706
65	Advanced Web Technology	BC61701
66	Project	BC61706

IQAC Co-Ordinator
St. Mira's College for Girls, Pune

ST. MIRA'S COLLEGE FOR GIRLS ★ POONA-411001

G. H. Gidwani
Principal
St. Mira's College for Girls

BUSINESS MATHEMATICS.
PROJECT.

Name :- Dolly Suresh Rathod.

Roll No :- 3467

Class :- F.Y.B.Com.

Div :- C.

Semester :- II.

Year :- 2020-2021.

Title :- Correlation and Regression.

INDEX.

* Defination and Notation of Correlation and Regression :-

Defination :-

Correlation :- Correlation is a statistical measure that determine the association or correlationship between two variables.

Regression :- Regression describes how to numerically relate an independent variable to the dependent variable.

In dependent and Independent variables the correlation has no difference and in regression both variables are different. Correlation coefficient indicates the extent to which two variables move together. Regression indicates the impact of a change of unit on the estimated variable (y) in the known variable (x). The objective of correlation is to find a numerical value expressing the relationship between variables and the objective of Regression is to estimate values of random variables on the basis of the values of a fixed variables. The X variable can be fixed with correlation, but confidence intervals and statistical tests are no longer appropriate.

Correlation quantifies the direction and strength of the relationships between two numeric variables, x and y, and always lies between -1.0 and 1.0. Simple linear regression relates x to y through an equation of the form $Y = a + bx$.

★ Key similarties :-

- Both quantify the direction and strength of the relationship between two numeric variables.
- When the correlation (r) is negative, the regression slope (b) will be negative.
- When the correlation is positive, the regression slope will be positive.
- The correlation squared (r^2 or R^2) has special meaning in simple linear regression. It represents the proportion of variation in Y explained by X.

★ Key differences of correlation and Regression :-

★ Regression attempts to establish how x causes y to change and the results of the analysis will change

- if x and y are swapped. with correlation the x and y variables are interchangable

* Regression assumes x is fixed with no error, such as a dose amount or temperature setting. With correlation, x and y are typically both random variables, such as height and weight or blood pressure and heart rate.

* Correlation is a single statistic, whereas regression produces an entire equation.

---

* **Formula of correlation :-**
let x and y be the two random variables
The population correlation coefficient for x and y is given by the formulae.

$$\rho_{x,y} = corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

where,
ρ_{xy} = Population correlation coefficient between x and y
μ_x = Mean of the variable x
μ_y = Mean of the variable y.
σ_x = Standard deviation of x
σ_y = standaud deviation of y.
E = Expected value operator.
cov = covriance.

The above formula can also be written as :-

$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2}}$$

★ Regression Equation :-

As we know, linear regression is used to model the relationship between two variables. Thus, a simple linear regression equation can be written as :-

$$Y = a + bx$$

where,

Y = Dependent variable
X = Independent variable

$$a = [(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)] - [n(\Sigma x^2) - (\Sigma x)^2]$$

$$b = [n(\Sigma xy) - (\Sigma x)(\Sigma y)] / [n(\Sigma x^2) - (\Sigma x)^2]$$

Correlation Analysis :-

Correlation analysis is applied in quantifying the association between two continous variables, for example, and dependend and independent variable or among two independent variables.

Regression Analysis :-

Regression analysis refers to assessing the relationship between the outcome variable and one or more variables. The outcome variable is known as the dependent or response variable and the risk elements and cofounders are known as predictors or independent variables. The dependent variable is shown by "y" and independent variables are shown by "x" in regression analysis.

The sample of a correlation coefficient is estimated in the correlation analysis. It ranges between -1 and +1, denoted by r and quantifies the strength and direction of the linear association among two variables. The correlation among two variables can either be positive, i.e. a higher level of one variable is related to a higher level of another or negative, i.e. a higher level of one variable is related to a lower level of the other.

The sign of the coefficient of correlation shows the direction of the association. The magnitude of the coefficient shows the strength of the association.

For example, a correlation of $r = 0.8$ indicates a positive and strong association among two variables while a correlation of $r = -0.3$ shows a negative and weak association. A correlation near to zero shows the non-existence of linear association among two continuous variables.

* Key advantage of correlation :-

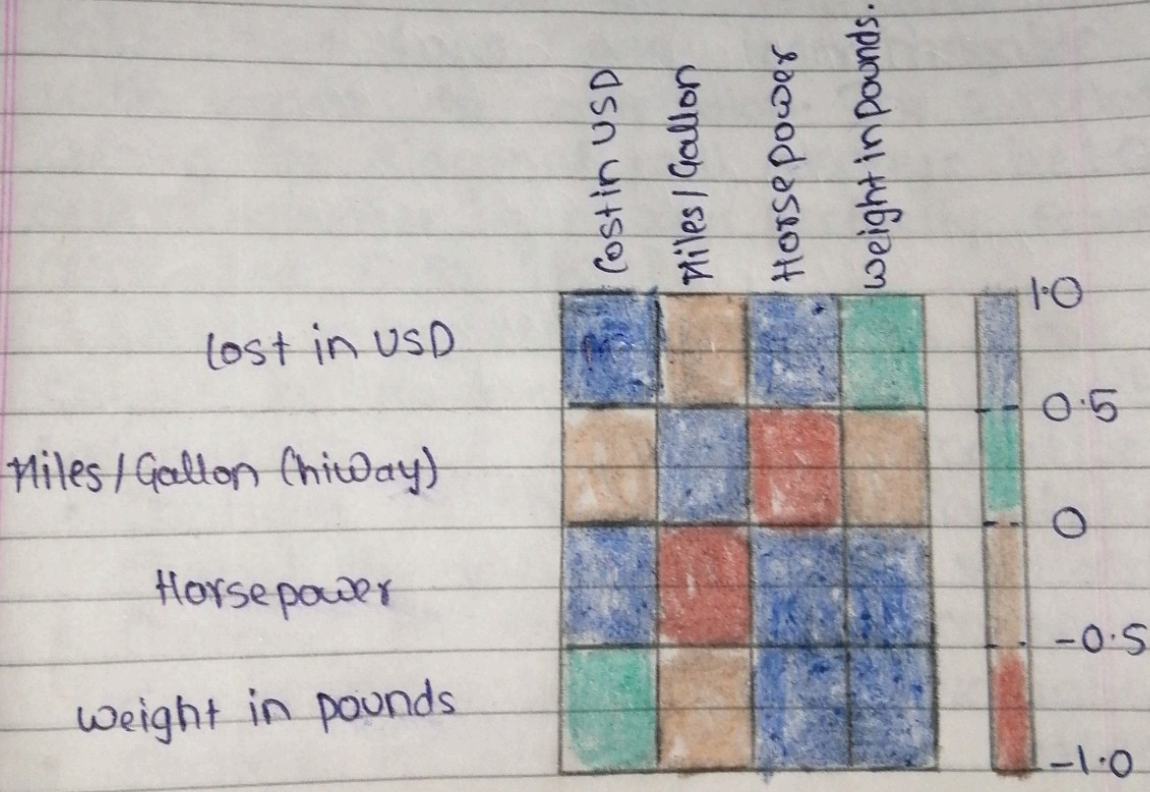
* Correlation is a more concise (single value) summary of the relationship between two variables than regression. In result, many pairwise correlations can be viewed together at the same time in one table.

* Key advantage of Regression :-

* Regression provides a more detailed analysis which includes an equation which can be used for prediction and/or optimization.

✱ Correlation example

As an example, let's go through the Prism tutorial on correlation matrix which contains an automotive dataset with cost in USD, MPG, Horsepower, and weight in Pounds as the variables. Instead of just looking at the correlation between one X and one Y, we can generate all pairwise correlations using Prism's correlation matrix.



The prism correlation matrix displays all the pairwise correlations for this set of variables.

- The red boxes represent variables that have a negative relationship
- The blue boxes represent variables that have a positive relationship.

Note that the matrix is symmetrix. For example, the correlation between "weights in pounds" and "cost in USD" in the lower left corner (0.52) is the same as the correlation between "cost in USD" and "weight in pounds" in the upper right corner (0.52). This reinforces the fact that X and Y are interchangable with regard to correlation. The correlation along the diagonal will always be 1.00 and a variable is always perfectly ~~correlated~~ correlated with itself.

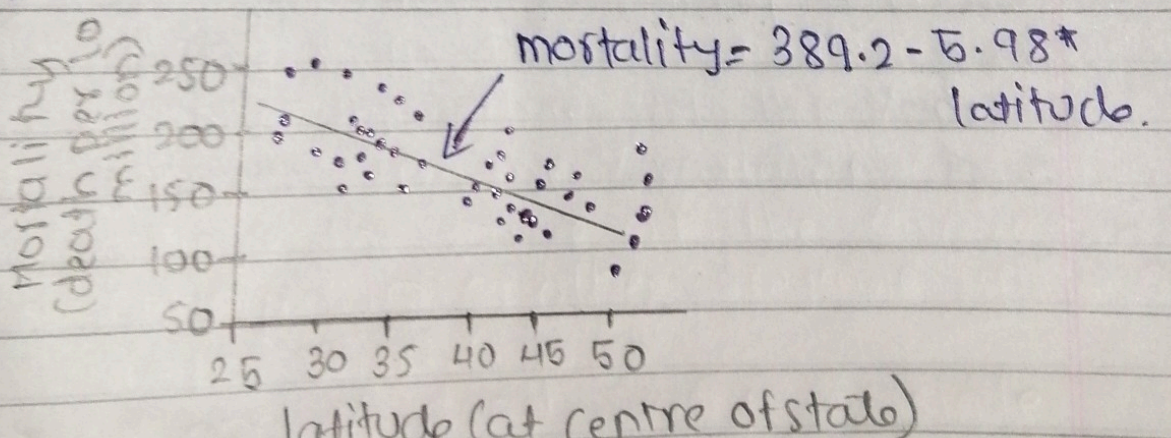
When interpreting correlations you should be aware of the four possible explanations for a strong correlation:-

- Changes in the X variable causes a change the value of the Y variable.
- Changes in the Y variable causes a change the value of the X variable.
- Changes in another variable influencies both X and Y.
- X and Y. don't really correlate at all, and you just happened to observe such a strong correlation by chance. The P value quantifies the likelihood that this could occur.

Regression example.

The strength UV rays varies by latitude. The higher the latitude, the less exposure to the sun, which corresponds to a lower skin cancer risk. So where you live can have an impact on your skin. Two variables, cancer mortality rate and latitude, were entered into Prism's XY table. The Prism graph (right) shows the relationship between skin cancer mortality rate (Y) and latitude at the centre of a state(X). It makes sense to compute the correlation between these variables, but taking it a step futher, let's perform a regression analysis and get a predictive equation.

Skin Cancer Mortality vs State latitude.



The relationship between X and Y is summarized by the fitted regression line on the graph with equation:
mortality rate = 389.2 - 5.98* latitude.

Based on the slope of -5.98 each 1 degree increase in latitude decreases deaths due to skin cancer by approximately 6 per 10 millions people.

Since regression analysis produces an equation, unlike correlation, it can be used for prediction. For example, a city at latitude 40 would be expected to have 389.2 - 5.98*40 = 150 death per 10 million due to skin cancer each year. Regression also follows for the interpretation of the model coefficients.

- Slope :- every one degree increase in latitude decreases mortality by 5.98 deaths per 10 millions

- Intercept :- at 0 degree latitude (equator), the model predicts 389.2 deaths per 10 million. Although, since there are no data at the intercept, this prediction relies heavily on the relationship maintaining its linear from to 0

Conclusion :-

Correlation examines the strength of the relation between two variables. neither of which is necessarily considered the target variable. Regression examines the strength of the relation between one or more predictor variables and a target variable. Regression can be very useful in formulating predictive models such as the risk of myocardial infaraction in patients presenting with chest pain, the risk of cardic events in patients undergoing non-cardic surgery, whether you are considering a correlation between variables or a regression analysis, you should consider not only the statistical significance of the relation but also its magntitude or strength, in terms of the proportion of variation explained by the model or the extent to which groups with very different risks can be specified.

* Bibliography :-

1] www. investopedia. com
2] en. m. wikipedia .org
3] https://www. toppr. com
4] graphpad. com
5] https://www. collectionscanda. gc.
6] www. byju's. com.