Project Work For B.Sc. III In Year 2020-21.

Sr. No.	Name of the Students	Title of the Project	Name of Mentor
1	Shinde Gurudas Ananda Yamgar Bhavesh Bajirao Kulkarni Abhishekh Shrikant Kulkarni Sushilkumar Jagdish	To study the challenges of online education over traditional education	Dr. M. K. Patil
2	Bhapkar Amruta Shivaji. Mujawar Swaleha Yunus Patil Ankita Chavagounda. Pawar Swati Pradip Shaikh Simran Shoukat	To study the relation between age, weight and blood pressure	Mr. V. V. Koshti
3	Kadam Komal Sukhdev Kadam Gauri Suresh Bhosale Seema Ajay More Aishwarya Arvind	Roles of fertilizers on grapes production	Mr. A. P. Patil
4	Babar Aishwarya Sanjay Patil Bhagyashri Harishchandra Chavan Sneha Shrikant Patil Vaishnavi Vilas	Statistical Analysis of effect of fertilizers and soil type on sugarcane	Miss. S.S. Bhosale
5	Shinde Jyoti Subhash Mujawar Simaran Nisar Mali Pooja Sahadev Mali Sonali Vithoba Kore Pratiksha Jayant	To study the time spent by students on social media	Miss. S. S. Patil

A
Project Report
On

"TO STUDY THE CHALLENGES OF ONLINE EDUCATION OVER TRADITONAL EDUCATION"

Submitted to

Padmabhushan Vasantraodada Patil Mahavidyalaya,
KavatheMahankal.
DEPARTMENT OF STATISTICS

By,
Mr. Shinde Gurudas Ananda
Mr. Yamgar Bhavesh Bajirao
Mr. Kulkarni Abhishekh Shrikant
Mr. Kulkarni Sushilkumar Jagdish

Under the Guidance of,

Prof. Dr. Patil M. K.
Prof. Koshti V. V.

Prof. Patil A. P.
Prof. Bhosale S. S.
Prof. Patil S. S.

Through,

The Principal,
Padmabhushan VasantraodadaPatil
Mahavidyalaya, Kavathe Mahankal.

For the academic year 2020-2021

Padmabhushan Vasantraodada Patil Mahavidyalaya, Kavathe Mahankal

DEPARTMENT OF STATISTICS

CERTIFICATE

This is to certify that the following students are working under my guidance. They have satisfactorily carried out and completed the project work for the B. Sc. Part III. They have presently submitted their project entitled "TO STUDY THE CHALLENGES OF ONLINE EDUCATION OVER TRADITONAL EDUCATION"

This work is genuine, original and the project or the part of it has not been submitted for the award of any other degree.

1. Mr. Shinde Gurudas Ananda.
2. Mr. Yamgar Bhavesh Bajirao.
3. Mr. Kulkarni Abhishekh Shrikant.
4. Mr. Kulkarni Sushilkumar Jagdish.

Dr. M. K. Patil	Dr. M. K. Patil	Dr. M. K. Patil
Teacher in charge	**Examiner**	**Head of Department**

DECLARATION

We hereby declare that the project entitled "TO STUDY THE CHALLENGES OF ONLINE EDUCATION OVER TRADITIONAL EDUCATION".

It is the original work carried out by us at the Department of Statistics.

This project or part of it has not been submitted previously for any other degree of any other purpose.

Sr. No.	Roll No.	Name	PRN
1.	1	Shinde G.A.	2018010208
2.	2	Yamgar B.B	2018014653
3.	3	Kulkarni A.S	2017045567
4.	4	Kulkarni S.J	2018014505

Place : Kavathe Mahankal

Date :

ACKNOWLEDGEMENT

We express our hearty thanks and deepest gratitude to Dr. Patil M. K. who gave backing and his valuable guidance for this project work. We further thank to Prof. Koshti V.V. Prof. Bhosale S.S. Prof. Patil A.P. and Prof. Patil S.S. Who gave us their encouragement, kind support, and constant inspiration during our project, work further we are thankful to all of our college students who gave us necessary information to carry out this project.

Finally we give sincerest thanks to all our family members for keeping our spirit up and their immerse support during every stage of the project work.

CONTENTS

INTRODUCTION

As the first wave of COVID-19 entered India in early 2020, schools were closed in mid-March, just weeks before the end of the 2019-20 academic years.

The vast majority of the country's 26 crore school children have not set foot in a school since then, depending instead on various forms of distance education.

The availability of digital education – whether via live, synchronous teaching on platforms like Zoom, or through recorded lectures, emails, WhatsApp or educational apps - was largely dependent on whether schools, teachers and parents had access to the necessary infrastructure. In many States, teachers came to school and taught in their own empty classrooms, using their blackboards and lab facilities, while facing a computer screen that communicated the lessons to their students at home.

Our aim is to bring out real effect of COVID-19 on educational sector and whether do we have necessary infrastructure to enhance online education in future? And what students/ learners think about online education, what major issues they are facing to educate themselves through online method. We sampled students from various colleges, universities and high-schools, from very beginning to, that is, Primary, to last phase of one's education, that is, Masters, Because there seems to be big difference between how college graduates, who have been already learning from various online sources and the school noob's who had never used such method for education till now.

We also discussed **Online v/s Offline education** with some school teachers and professors, about what challenges they are facing to teach online, because none of them have been through this kind of time .During sampling we found that, it's really hard for teachers to teach student below sixth or seventh grade and thereafter collaborative curve increases exponentially.

The intention behind this project is to analyze the grave flexibility of educational system and its infrastructure to tackle unbearable situation like this, because we are spending hefty amounts to get on Mars but one unprecedented 80 nm sized virus could shut the earth down, and make us question on the reliability.

This project will provide all the aspect of students who are being educated by online scheme, what are the ups and downs of online education, comparison of online learning with traditional learning through various options. Conclusions can be drawn by statistical graphs and various statistical tests.

OBJECTIVES OF STUDY

1. To get the basic idea about the current online infrastructure through various statistical graphs.

2. To compare the effectiveness of online and offline learning.

3. To study the dependence of method of learning and its mode of education.

4. To study the various educational issues and their dependence on mode of education.

5. To study the preference of junior students (below age 18) that of the senior students (above age 18) regarding to online and offline education.

Source of Data & Sampling Method

The main objective of this project is, “To compare the online education with traditional education”. We have sampled students from different educational phases, to ensure that conclusions apply to various situations. Also this project is intended to know about the current methods that are students and teachers are using for online teaching-learning.

We designed a questionnaire to fulfill the objectives behind this project accurately. The data collection procedure was done using Google forms and by one-to-one approach. The perspective of teachers towards online education has affected crucially to this project.

We sampled subjects using simple random sampling without replacement (SRSWOR) method. The collection of sampling has been taken through online questionnaire using Google forms and direct interviewing with students.

STATISTICAL TOOLS

Following are some statistical tools us towards the appropriate results and finding.

- ***Graphical Representation***

To represent our data graphically, we use the multiple bar diagram, pie chart method etc. To construct the appropriable divided bar diagram of given data we used Microsoft Excel.

i. ***Pie Chart:*** A pie chart displays data, information and statistics in an easy-to-read 'pie-slice' format with varying slice sizes telling you how much of on data element exists. The bigger the slice the more of that particular data was gathered.

ii. ***Bar Chart:*** A bar chart or bar graph is a chart that presents grouped data with rectangular bars with lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. A vertical bar is sometimes called a column bar chart.

A bar graph is a chart that uses either horizontal or vertical bars to show comparisons among categories being compared, and the other axis represents a discrete value.

- ***Test for independence of attributes***

To test the independence of attributes, here data are in the form of cross tables. Where, A_i = Educational Methods/Problems

B_j = Mode of education

Hypotheses:

H_0: Attributes A_i and B_j are independent of each other.

vs

H_1: Attributes A_i and B_j are dependent of each other.

Under H_0 the test statistics is

$$\chi^2 = \sum_{i=1}^{R}\sum_{j=1}^{C} \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

Statistic χ^2 has (R-1) (C-1) d .f.

The value of χ^2 is obtained with the help of formula. To make decision about H_0, Calculated value of χ^2 is compared with table value of χ^2 for (R-1) (C-1) d. f and α % level of significance (l.o.s)

Conclusion:

If cal. $\chi^2 \geq$ tab. χ^2, then we reject H_0 at α % l. o. s. Otherwise we accept H_0.

- ***Test for population proportion:***

To test the population proportion of two samples.

Hypothesis:

H_0: Two population proportions are same. ($p_1 = p_2$)

Vs

H_1: Two population proportions are not same. ($p_1 \neq p_2$)

Under H_0, the test statistic is given by,

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Let α be the level of significance and Z_α being the critical value.

Conclusion:

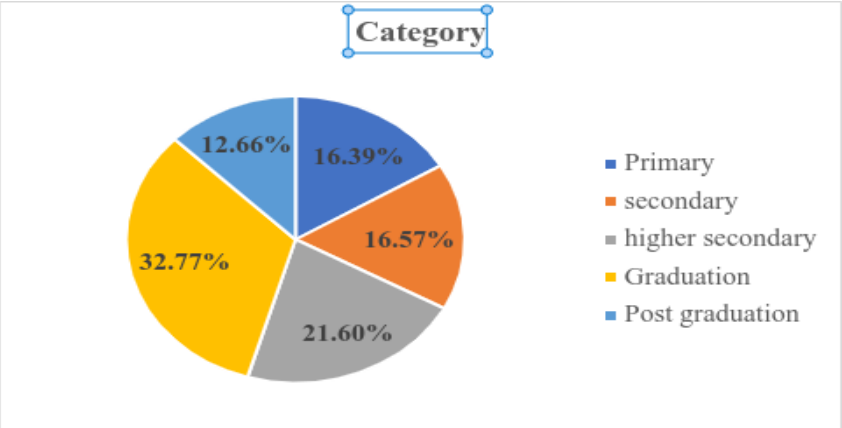
If calculated |Z| > tabulated Z, we reject H_0 at α % l.o.s otherwise we accept it.

STATISTICAL ANALYSIS

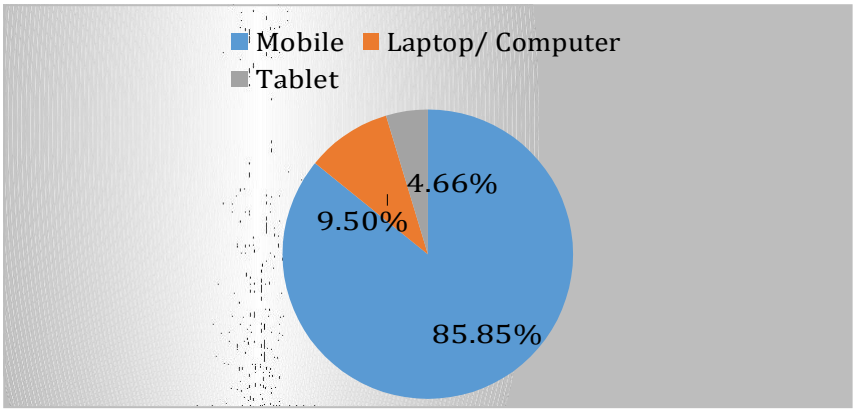
1. To get the basic idea about the current online infra-structure through various statistical graphs.

i) To which category do you belong?

This data are merely for sample and have done proportion test in further analysis.



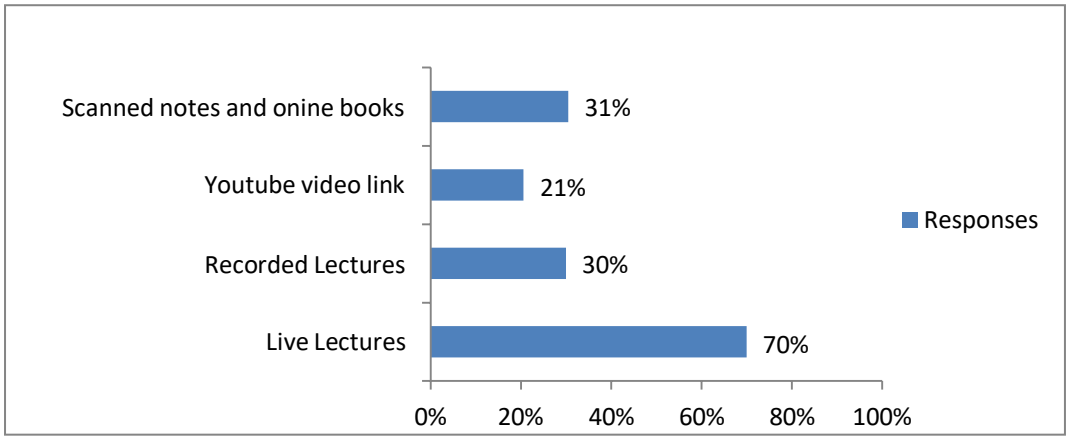
ii) Which device do you use for online learning?



Conclusion:

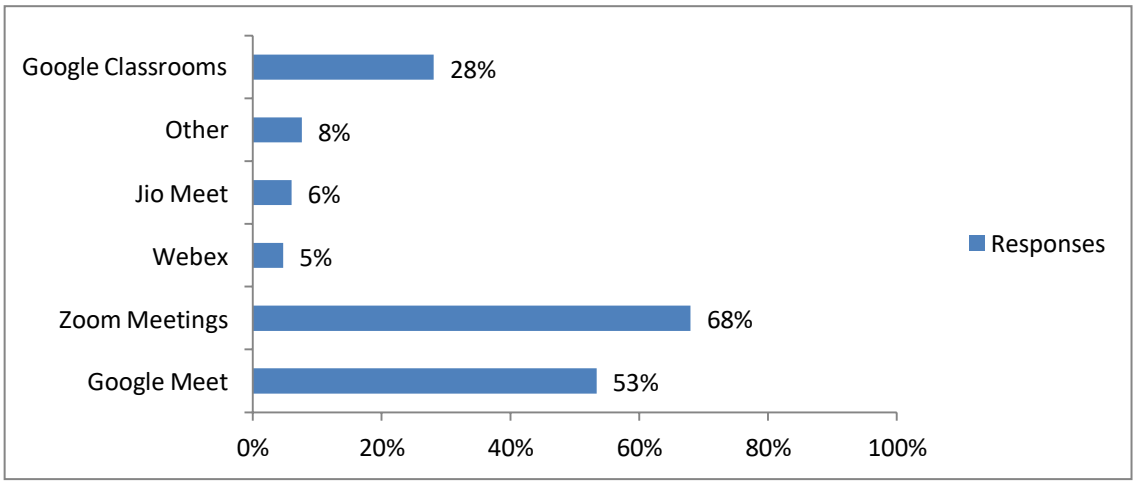
From the above graph we conclude that, most of the students are using Mobile as their primary learning device.

iii) Which method do you use for online learning?



Conclusion: From the above graph we can conclude that, **Live lectures** are popular among students and teachers.

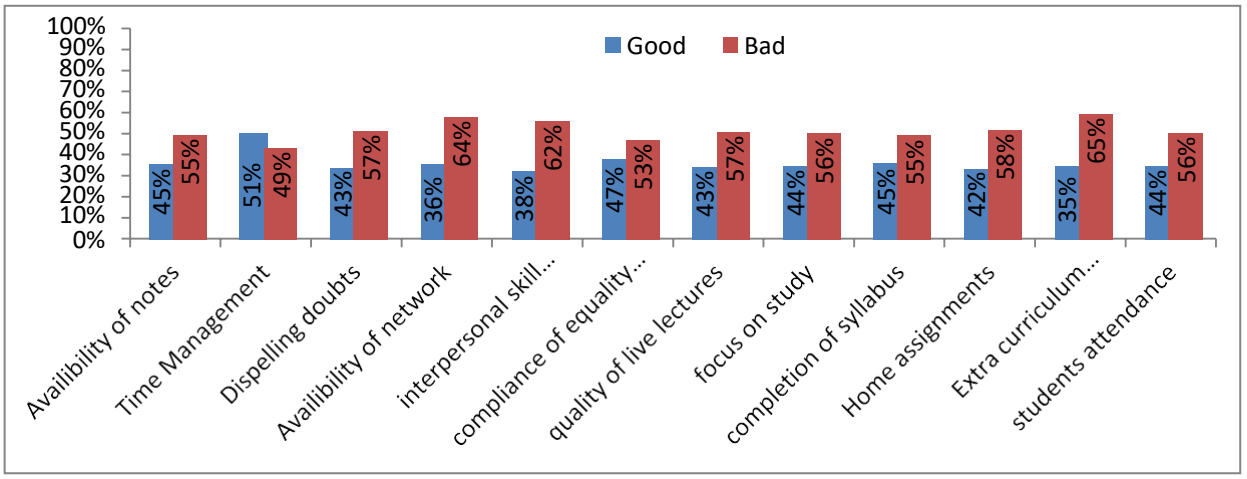
iv) Which application do you use for online learning?



Conclusion: Zoom meetings and **Google meet** are popular online learning platforms among all.

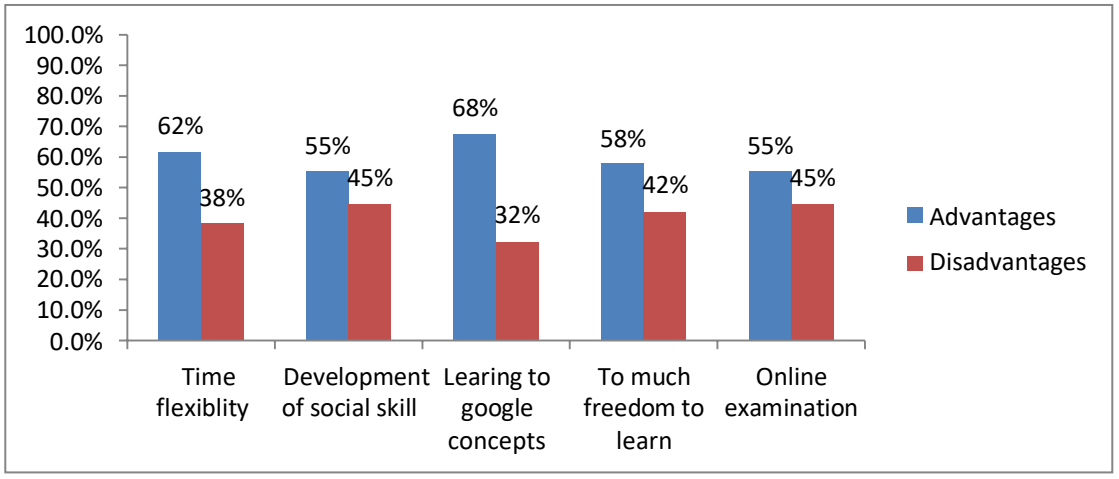
2. To compare the effect of online and offline learning.

i) Please answer by circling one number against each statement in case of 'Online Learning'.



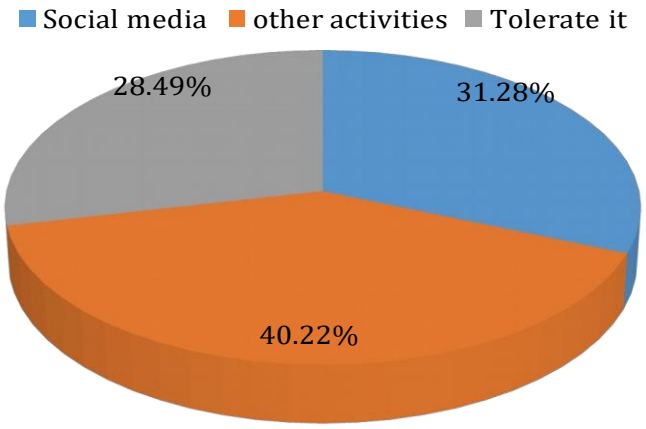
Conclusion: According to the above graph, if we were to summed up, during online lectures, technical issues and being deprived from social activities are inevitable.

ii) If online education implemented properly, its advantages or disadvantages



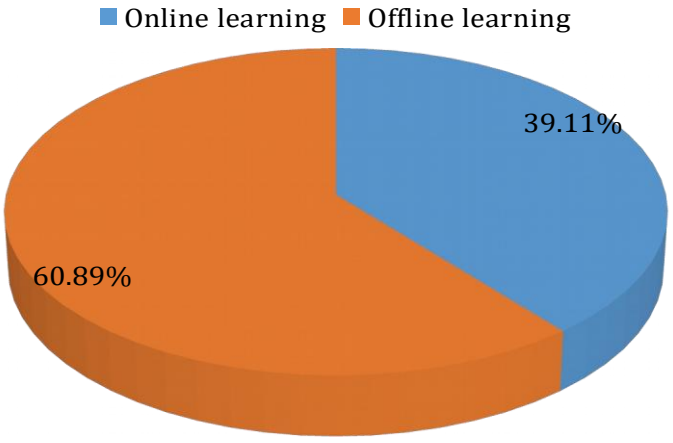
Conclusion: Above graph suggests that, students think Time management, Freedom and Online Examination are perks over traditional education system, if online education implemented properly.

iii) What do you do during weary online lectures?



Conclusions: Above graph suggests that, Students tend to **other activities** and **Social media** when online lectures are boring.

iv) What would you prefer for long learn?



Conclusion: From the above graph we conclude that, still most of the students prefer offline as their main method of education.

3. To study the dependence of method of learning and its mode of education.

Here we use Chi Square test for independence of attributes to test the hypothesis. To check whether different methods of learning and mode of education are independent or not.

H_0: Different methods of learning and mode of education are independent.

vs

H_1:. Different methods of learning and mode of education are dependent.

Observation Table:

	Online	Offline	Expected online	Expected offline
Project	211	326	208.5	328.5
Teaching/ Learning	211	326	208.5	328.5
Getting Knowledge	238	299	208.5	328.5
Dispelling Doubts	208	329	208.5	328.5
Social Interaction	194	343	208.5	328.5
Practicals	189	348	208.5	328.5

Under H_0, The test statistics is

$$\chi^2 = \sum_{i=1}^{R}\sum_{j=1}^{C}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$$

= 11.5526

Tabulated value of χ^2_5 at 5% level of significance = 11.0705

Result: Here, Calculated value of χ^2 is 11.5 526 > tabulated value of χ^2_5 is 11.0705,Hence we reject H_0..

Conclusion: Method of learning and mode of education are dependent on each other. And also concludes students find current online education method less effective.

4. To study the various educational issues and their dependence on mode of education.

Here we check whether current online infrastructure is more problematic or not.

H_0: Educational problems are independent of mode of education.

vs

H_1: Educational problems are dependent of mode of education.

Observation table:

Problems	Yes or it is	No or its not	Expected Yes's	Expected No's
Bad effect on mental health	267	270	291.8333	245.1667
Is digital divide an issue	265	272	291.8333	245.1667
Decrease in expenditure	310	227	291.8333	245.1667
Increase in procrastination	298	239	291.8333	245.1667
Bought new device for lectures	301	236	291.8333	245.1667
Participation in online competition	310	227	291.8333	245.1667

Under H_0 Test statistic is,

$$\chi^2 = \sum_{i=1}^{R} \sum_{j=1}^{C} \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

=15.9028

Tabulated value of χ^2_5 at 5% level of significance = 11.0705

Result: Here, Calculated value of χ^2 is 15.9028 > tabulated value of χ^2_5 is 11.0705 ,Hence we reject H_0..

Conclusion: Hence we conclude that, Educational problems are dependent on mode of education. Issue of digital divide and increase in procrastination or distractions are the biggest problems of online education.

5. To study the preference of junior students that ofthe senior students regarding to online and offline education.

Here we want to check, what proportion of students and from which phase prefer which education method. This is age wise comparison, junior students (<18) and senior students (>18) according to there preference.

H_0: There is no significant difference between preference of junior students and senior students regarding to online and offline education.

vs

H_1: There significant difference between preference of junior students and senior students regarding to online and offline education.

Observation Table:

	Online	Offline
Junior Students	100	144
Senior Students	83	161

Under H_0 Test statistic is,

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

= 1.589584

Tabulated value of Z at 5% level of significance = 1.64485

Result: Here, Calculated value of Z is 1.589584 < tabulated value of Z is 1.6448. Hence we accept H_0.

Conclusion: There is no significant difference between preference of junior students and senior students regarding to online and offline education. All students find offline should be the mode of the education irrespective of their age.

SUMMARY

Overall conclusion can be drawn from this project is, Students have strong bias towards traditional education.

Reasons are as follows:

i) Not having proper online infrastructure.

ii) Freedom of time and everything which tend to increase in procrastination.

iii) No social interfere and lack of interpersonal skill development.

iv) More distractions.

v) Online held Examinations and Competitions can't be fair.

vi) Absentee of students to online held lectures.

Whilst, online education could be game changer if implemented properly, as now it is in its early stages. Teacher dominant environment could solve most of its flaws, and issues related to social interactions might be solved by creating online communities.

Sampled data suggest that, above conclusions are strongly valid to the population. However, the absentee of students leads to increase in dropout rate, also during sampling we found some students don't have device to attend online lectures, which is one of the reason for absentee of the students. The conclusions made from this project further confirm the UDISE+ data released in July 2021.

BIBLIOGRAPHY

1. Basic Statistics, B. L. Agarwal, *New Age International (2006)*

2. Mathematical Statistics, Gupta & Kapoor, *Sultan Chand & Sons (2017)*

3. www.pib.gov.in.

4. www.udiseplus.gov.in

5. www.mospi.nic.in

QUESTIONNAIRE

Questionnaire for Statistical Study of online learning

A Project Report On

" TO STUDY THE RELATION BETWEEN AGE,WEIGHT AND BLOOD PRESSURE "

Submitted to.

(Department of Statistics)

Padmabhushan Vasantraodada Patil Mahavidyalaya, Kavathe Mahankal.

By

1) Bhapkar Amruta Shivaji.

2) Mujawar Swaleha Yunus.

3) Patil Ankita Chavagounda.

4) Pawar Swati Pradip.

5) Shaikh Simran Shoukat.

Under the Guidance of

Prof.Dr.M.K.Patil. (Head of the department.)

Prof.V.V.Koshti.

Mr.A.P.Patil.

Miss.S.S.Bhosale.

Miss.S.S.Patil.

Through

Prof.Dr.M.K.Patil

2020-2021

Shikshan Prasarak Sanstha's

Padmabhushan Vasantraodada Patil Mahavidyalaya

Kavathe Mahankal.

CERTIFICATE

This is to certify that the following students are working under my guidance.They have satisfactorily carried out and completed the project work for the B.Sc. Part III.They are presently submitted their project entitled.This work is genuine, original and the project or the part of it has not been submitted for the award of any other degree.

"TO STUDY THE RELATION BETWEEN AGE,WEIGHT AND BLOOD PRESSURE".

1. Miss. Bhapkar Amruta Shivaji.
2. Miss.Mujawar Swaleha Yunus.
3. Miss. Patil Ankita Chavagounda.
4. Miss. Pawar Swati Pradip.
5. Miss.Shaikh Simran Shoukat.

Teacher in charge. Examiner. Head •Department of Statistics•

Declaration

We hereby declare that the project entitled "TO STUDY THE RELATION BETWEEN AGE, WEIGHT AND BLOOD PRESSURE ". It is the original work carried out by us at the Department of Statistics.This project or part of it has not been submitted previously for any other degree or any other purpose.

Place: Kavathe Mahankal.

Date:

Sr.no	Name
1.	Bhapkar Amruta shivaji
2.	Mujawar Swaleha Yunus
3.	Patil Ankita Chavagounda
4.	Pawar Swati Pradip
5.	Shaikh Simran Shoukat

Acknowledgment

We express our hearty thanks and deepest gratitude goes to Shri. Koshti Sir, who made valuable guidance and co-operation during our project work. We are rather infused by the guidance of Mr.Patil sir, Miss. Bhosale madam, Miss. Patil madam, Who put us their encouragement, kind support, and constant inspiration during our project work further we are thankful to all of our college students who gave us the necessary information to carry this project.

Finally, we give sincerest thanks to all our family members For keeping our spirit up and their immense support during every stage of the project work.

Index

Introduction

Blood pressure (BP) is the pressure of circulating blood against the walls of blood vessels. Most of this pressure results from the heart pumping blood through the circulatory system. When used without qualification, the term "blood pressure" refers to the pressure in the large arteries. Blood pressure is usually expressed in terms of the systolic pressure (maximum pressure during one heartbeat) over diastolic pressure (minimum pressure between two heartbeats) in the cardiac cycle. It is measured in millimeters of mercury (mmHg) above the surrounding atmospheric pressure.

Blood pressure is one of the vital signs together with respiratory rate, heart rate, oxygen saturation, and body temperature that healthcare professionals use in evaluating a patient's health. Normal resting blood pressure, in an adult is approximately millimeters of mercury (16 kPa) systolic over80 millimeters of mercury (11 kPa) diastolic, denoted as "120/80 mmHg". Globally, the average blood pressure, age standardized, has remained about the same from 1975 to the present, at approx. 127/79 mmHg in men and 122/77 mmHg in women, although these average data mask quite large divergent regional trends. Blood pressure is influenced by cardiac output and systemic vascular resistance and arterial stiffness and varies depending on the situation.

OBJECTIVES OF STUDY

1) To study the information about blood pressure patients.

2) To study the symptoms of blood pressure.

3) To study at what age blood pressure is more likely to occur.

4) To study who has high blood pressure according to gender.

5) To study whether blood pressure is more dependent on weight and age.

6) To study whether blood pressure can be controlled without medicine.

7) To study the side effect of blood pressure.

Source of Data & Sampling Method

As the main aim of our project is “To study the relation between age,weight and blood pressure”. We consider different blood pressure are side effects, symptoms, more likely to age, more depend on age and weight, diet, high blood pressure to gender, etc. So we decide to collect all primary data from the villages Dashing, Hingangaon, Shindevadi, Vithurayachi Vadi, and Kavathe Mahankal, which will provide us the necessary and sufficient data within less time and more accuracy.

So we divide the selected area and from each village, we selected a sample of size 250 by simple random sampling method of the data is collected by filling questionnaires from each selected blood pressure patients by interview method.

Statistical Tools

The following are some statistical tools that lead us towards the appropriate results and findings.

1. Graphical representation:

To represent our data graphically, we use the multiple bar diagram, pie chart method, etc. To construct the appropriable divided bar diagram of given data we used Microsoft Excel.

i) Bar Chart: A bar chart or bar graph is a chart that presents grouped data with rectangular bars with lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. A vertical bar is sometimes called a column bar chart.

A bar graph is a chart that uses either horizontal or vertical bars to show comparisons among categories being compared, and the other axis represents a discrete value.

ii) Pie chart: A pie chart displays data, information, and statistics in an easy-to-read ‘pie-slice' format with varying slice size telling you how much of one data element exists. The bigger the slice the more of that particular data was gathered.

2.Multiple correlation coefficient:

Multiple linear regression means, " a numerical index of the degree of relationship between a particular variable and two or more other variables".

The multiple correlation coefficient of X1 on X2 and X3 is denoted by Rx1.x2x3 and is given by,

R1.23^2 = (r12^2+r13^2-2r12.r13.r23) /(1-r23^2)

Similarly, R2.13^2 =(r12^2+r23^2-2r12.r13.r23) /(1-r13^2)

And R3.12^2 =(r13^2+r23^2-2r12.r13.r23) /(1-r12^2)

The sample correlation coefficient is,

r12 = Cov(x1,x2)/σ1σ2

r13 =Cov(x1,x3)/σ1σ3

r23 =Cov(x2,x3)/σ2σ3

And the partial correlation coefficient is,

r12.3= (r13-r23*r12)/sqrt((1-r23^2)(1-r12^2))

r13.2= (r12-r12*r23)/sqrt((1-r12^2)(1-r23^2))

r23.1=(r23-r12*r13)/sqrt((1-r12^2)(1-r13^2))

Where,

Cov(x1,x2)=1/n* ∑x1.x2- x1bar*x2bar

Cov(x1,x3)=1/n*∑x1.x3-x1bar*x3bar

Cov(x2,x3)=1/n*∑x2.x3-x2bar*x3bar

$\sigma 1^2 = 1/n \sum(x1-x1bar)^2$

$\sigma 2^2 = 1/n \sum(x1-x1bar)^2$

$\sigma 3^2 = 1/n \sum(x1-x1bar)^2$ and

X1 bar= $\sum x1/n$

X2 bar=$\sum x2/n$ and X3 bar=$\sum x3/n$

Now,the test for sample correlation coefficient is,

H0: r12= 0 (The blood pressure is depend on age)

V/S

H1:r12≠0(The blood pressure is not depend on age)

The test statistic is,

t=r/sqrt(1-r^2)*sqrt(n-1)~tn-2

Let,albha be the 5% level of significance.

Conclusion: If cal t>tab t,we reject Ho at 5% level of significance otherwise accept Ho.

The test for partial correlation coefficient is,

Ho:r12.3=0 (The blood pressure is depend on age and weight.)

V/S

H1:r12.3≠0(The blood pressure is not depend on age and weight.)

Under Ho,the test statistic is,

$t = r/\sqrt{(1-r^2)} * \sqrt{(n-k-2)} \sim t_{n-k-2}$

Let,albha be the 5% level of significance.

Conclusion: If cal t > tab t ,we reject Ho at 5% level of significant otherwise accept Ho.

The test for multiple correlation coefficient is,

Ho:R1.23=0 (The blood pressure is depend on age and weight)

V/S

H1:R1.23≠0(The blood pressure is not depend on age and Weight.)

Under Ho,the test statistic is,

$F = R^2/(1-R^2) * (n-k-1)/k \sim F(k, n-k-1)$

Let,albha be the 5% level of significance.

Conclusion: If cal F > tab F,we reject Ho at 5% level of significance otherwise accept Ho.

3.Test for one sample population proportion:

we have to test,

Ho : P=Po=0.5 (The population proportion is equal to 0.5)

V/S

H1 : P > Po =0.5(The population proportion is greater than 0.5)

Under Ho, the test statistic is,

$$Z = (p-Po) / \sqrt{(Po.Qo / n)}$$

Where,

Po = Population proportion under Ho=0.5

P = Sample proportion

Qo = 1-Po=1-0.5=0.5

n = Sample size

Let,alpha be the 5% level of significance.

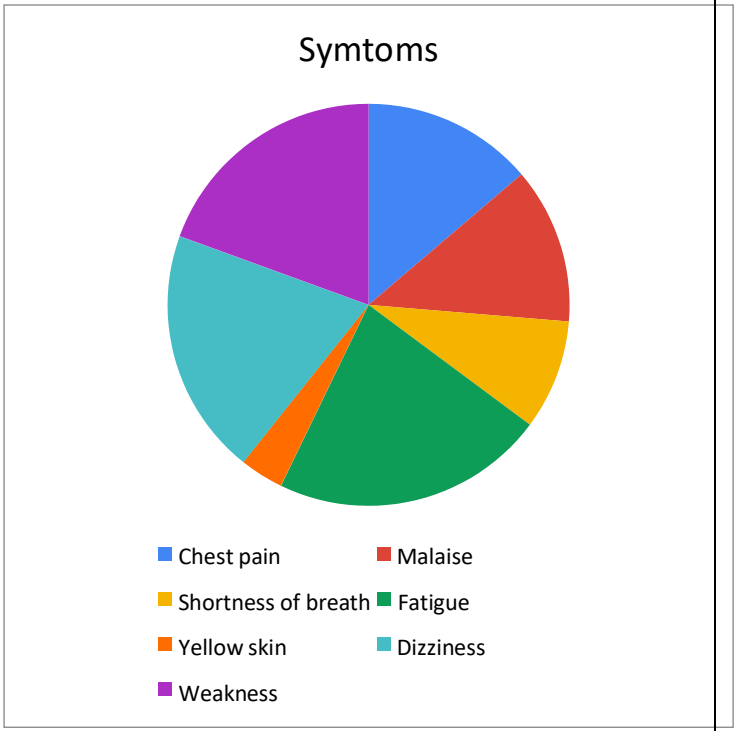
Conclusion:

If cal. $Z \geq 1.645$, then we reject H_0 otherwise accept Ho at 5% level of significance .

Statistical analysis

1) To study the symptoms of blood pressure.

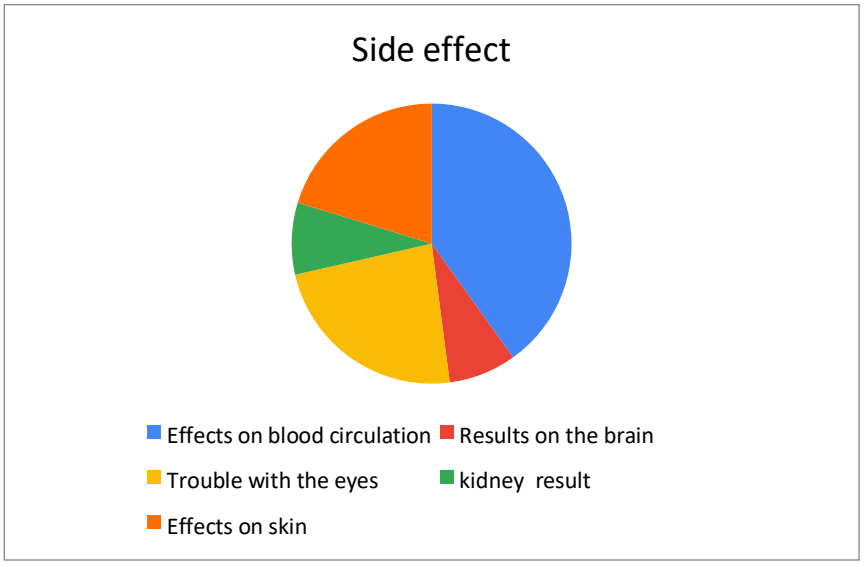
Symptoms	Total
Chest pain	23
Malaise	20
Shortness of breath	9
Fatigue	99
Yellow skin	6
Dizziness	50
Weakness	43



Comment: From the above diagram we conclude that , the most common symptom of blood pressure is fatigue.

2)To study the side effect of blood pressure.

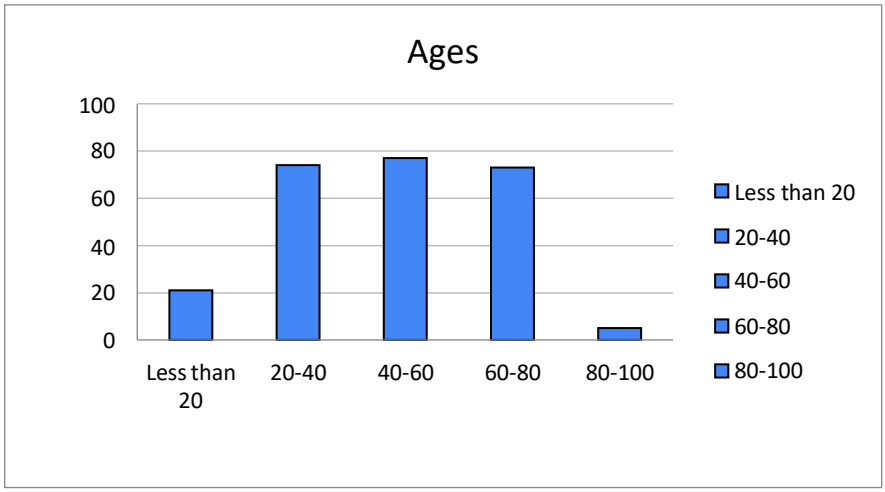
Side effect	Total
Effects on blood circulation	89
Results on brain	56
Trouble with the eyes	60
Kidney result	20
Effects on skin	25



Comment: From the above diagram we conclude that , the most common side effect of blood pressure is blood circulation.

3)To study at what age blood pressure is more likely occur.

Ages	Total
Less than 20	21
20-40	74
40-60	77
60-80	73
80-100	5



Comment: From the above diagram, The highest BP is in the age group of 40-60.

4)To study whether bp is more dependent on weight and age.

We have to test,

Ho:R1.23=0(The blood pressure is depend on age and weight.)

V/S

H1:R1.23≠0(The blood pressure is not depend on age and weight.)

Let,albha be the 5% level of significance.

Observation Table:

X1	X2	X3	X1^2	X2^2	X3^2	X1X2	X1X3	X2X3
140	40	75	19600	1600	5625	5600	10500	3000
140	49	70	19600	2401	4900	6860	9800	3430
142	45	66	20164	2025	4356	6390	9372	2970
145	60	45	21025	3600	2025	8700	6525	2700
157	50	75	24649	2500	5625	7850	11775	3750
135	54	73	18225	2916	5329	7290	9855	3942
139	59	67	19321	3481	4489	8201	9313	3953
137	43	70	18769	1849	4900	5891	9590	3010
147	62	75	21609	3844	5625	9114	11025	4650
142	69	65	20164	4761	4225	9798	9230	4485
140	68	80	19600	4624	6400	9520	11200	5440
128	60	62	16384	3600	3844	7680	7936	3720

135	75	45	18225	5625	2025	10125	6075	3375
140	56	45	19600	3136	2025	7840	6300	2520
155	72	75	24025	5184	5625	11160	11625	5400
135	64	65	18225	4096	4225	8640	8775	4160
120	59	65	14400	3481	4225	7080	7800	3835
76	48	69	5776	2304	4761	3648	5244	3312
190	85	50	36100	7225	2500	16150	9500	4250
75	46	73	5625	2116	5329	3450	5475	3358
170	65	52	28900	4225	2704	11050	8840	3380
160	75	87	25600	5625	7569	12000	13920	6525
136	19	50	18496	361	2500	2584	6800	950
124	36	56	15376	1296	3136	4464	6944	2016
147	40	54	21609	1600	2916	5880	7938	2160

Totals	∑X1=3455	∑X2=1399	∑X3=1609	∑X1^2=491067

∑X1X2=196965	∑X1X3=221357	∑X2X3=90291

Calculation: Here,

X1 bar=138.2 , X2 bar=55.96 and X3 bar=64.36

σ1 = 23.31179959 , σ2=14.40411052 and σ3=11.53734805

n1=25. and n2=25

Cov(X1,X2)= 144.928

Cov(X1,X3)= -40.272 and

Cov(X2,X3)=10.0544

Sample correlation coefficient is,

r12=0.431608564

r13=-0.14973434

r23=0.060501163

Multiple correlation coefficient between X1 on X2 and X3 is

R1.23=0.461373932

Similarly,

R2.13=0.444910847 and

R3.12=0.188920573

Partial correlation coefficient is,

r12.3=0.44487583

r13.2=-0.19458221 and

r23.1=0.137149403

Testing of the values:

Sample correlation coefficient is,

t cal=2.2946

t cal= 0.7261

t cal=0.2906

Partial correlation coefficient is,

t cal=2.2758

t cal= 0.9086

t cal=0.6342

Multiple correlation coefficient is,

F cal=2.9735

F cal=2.7146

F cal= 15.28

The tabulated value as,

Sample correlation=t tab=2.06865761

Partial correlation=t tab=2.079613845

Multiple correlation=F tab=3.443356779

Result: Here,α be the 5% level of significance.

Sample correlation coefficient

1) If cal t > tab t,we reject Ho
2) If cal t < tab t,we accept Ho
3) If cal t < tab t,we accept Ho

Partial correlation coefficient

1) If cal t > tab t,we reject Ho
2) If cal t < tab t,we accept Ho
3) If cal t < tab t,we accept Ho

Multiple correlation coefficient

1) If cal F< tab F,we accept Ho
2) If cal F <tab F,we accept Ho
3) If cal F > tab F,we reject Ho

Conclusion: We conclude that,

For the sample correlation coefficient,

1)The blood pressure is depend on age.

2) The blood pressure is depend on weight.

For the partial correlation coefficient,

1)The blood pressure is depend on age.

2) The blood pressure is depend on weight.

For the multiple correlation coefficient,

1)The age is depend on blood pressure.

2)The weight is does not depend on blood pressure.

5)To study who has high BP according to Gender.

we have to test,

Ho: P=Po =0.5 (The male and female are equally affected by blood pressure.)

V/S

H1: P<Po (P<0.5) (The females patients are less affected by blood pressure as compared to male blood pressure patients.)

Let, alpha=5% be the level of significance .

Calculation: Here,

n =Total number of samples = 250

X= Number of female=121

P = Sample proportion =X/n=121/250=0.48

Under Ho, the test statistics is,

$Z = (0.484-0.5)/\sqrt{(0.25/250)}$

Z = -0.5059

Cal Z = -0.5059

Let, alpha be the 5% level of significance.

Tab.Z = - 1.645

Result: Here, the calculated value of Z is greater than the tabulated value of Z.Hence, we reject Ho at a 5% level of significance.

Conclusion: We conclude that, the female blood pressure patients are less affected by blood pressure as compared to male blood pressure patients.

Overall all Conclusion

1) The most common symptom of blood pressure is fatigue.
2) The most common side effect of blood pressure is blood circulation.
3) The highest age group is the in blood pressure is 40-60.
4) i) The blood pressure is depend on age.
   ii) The blood pressure is depend on weight.
   iii) The blood pressure is depend on age.
   iv) The blood pressure is depend on weight.
   v) The age is depend on blood pressure.
   vi) The weight is does not depend on blood pressure.
5) The female blood pressure patients are less affected by blood pressure as compared to male blood pressure patients.

LIMITATIONS

1) We have to limit the scope of this research according to the result we obtain from the questionnaires.
2) We consider the whole population is normal but in reality, it may be or may not be normal.
3) The time constraint was one of the major limitations.

SCOPE

1) A well-planned Study can be taken including some characteristics.
2) It can be extended to a greater region for better estimates.

BIBLIOGRAPHY

This bibliography is intended as a resource for the design and implementation of blood pressure programs, and is useful for blood pressure treatment protocol development and adaptation. The aim is to provide a brief overview of the scientific basis of key aspects of blood pressure control for a comprehensive approach to blood pressure management. This list is not representative of all aspects of blood pressure management and is largely drawn from the scientific literature published in English.

The bibliography is organized in the following manner: first, a review of the burden of to underscore its importance as a non-communicable disease; second, a review of the evidence for protocol-based treatment; third, a review of the evidence on task sharing; fourth, a summary of the literature on medical supplies and patient-centered services; and lastly, a synopsis on information systems.

REFERRED BOOKS:

Sr. No	Book name	Author
1.	Control the blood pressure	Robert Rowan M.D
2.	Blood pressure solutions	Jessica Robbins
3.	The complete guide to lowering blood pressure naturally	Deborah Mitchell

Questionnaires

1) तुमचे वय किती आहे?

a) 20 पेक्षा कमी

b) 20-40

c) 40-60

d) 60-80

e) 80-100

2)तुमचे वजन किती आहे ?

a) 30-40

b) 40-50

c) 50-60

d) 60-70

e) 70-80

f) 80-90

g) 90-100

3) तुम्हाला कोणत्या प्रकारचे बीपी आहे?

a) जास्त

b) कमी

4) तुमच्या बीपी ची रेंज काय आहे?

a) 60-70

b) 70-80

c) 120-140

d) 140-160

e) 160-180

f) 180-200

g) 200 पेक्षा अधिक

5) तुम्ही कोणत्या प्रकारचे आहार घेता?

a) फळे

b) पालेभाज्या

c) मांसाहार

d) गोड पदार्थ

6) तुम्ही कोणत्या प्रकारचे आहार घेत नाही?

a) मीठ

b) आबट पदार्थ

c) दुग्धजन्य पदार्थ

d) चहा कॉफी

7) तुम्ही व्यायाम करता का?

a) हो **b)** नाही

8) जर व्यायाम करत असतील तर कोणते ?

a) चालणे

b) पळणे

c) सायिल चालवणे

d) प्राणायाम

e) योगासने

f) नतृ्य करिणे

9) तुम्ही कोणत्या प्रकारचे औषधि घेता ?

a) ऍलोपॅथिक

b) होमओपॅथिक

c) इतर

आलीत?

10) तुम्हाला कोणत्या प्रकारची लक्षणे आढळ

a) छातीत दखु णे

b) अस्वस्थता

c) श्वसनास त्रास

d) र्‍िवा

e) कपवळी त्वचा

f) चक्कर मारणे

g) अशक्तपणा

11) वजन वाढल्यास अर्वा कमी झाल्यास तुमच्या बीपी वर कोणता पररणाम होतो का?

a) हो

b) नाही

12) बीपी कण्ट्रोल मध्ये आणण्यासाठी तुम्ही कोणते प्रयत्न करता?

a) कनरोगी वजन ठेवणे

b) अकि ि व्यायाम िरणे

c) कमठाचा वापर िमी ि◌रणे

d) फळे- पालेभाज्या आहारात वाढवणे

13) अचानि◌ बीपी वाढल्यास तुम्ही ि◌ाय ि◌रता?

a) घरगुती उपाय

b) दवाखान्यात जाणे

14) घरगुती उपाय ि◌ोणते?

a) ग्लबू पाणी

b) तणाव िमी ि◌रणे

C) नारळ पाणी

d) इतर

15) बीपी मुळे तुमच्या शरीरावर पररणाम होतात ि◌ा?

a) हो

b) नाही

16) बीपी मुळे तुमच्या शरीरावर ि◌ोणते पररणाम होतात?

a) रक्ताकभसरणवर पररणाम

b) र पररणाम
मे◌ंदव

c) डोळ्याचा त्रास

d) किडनी वर पररणाम

e) त्वचेवर पररणाम

17) ब्पीमुळे मानसि◌ पररणाम होतात ि◌ा?

a) भीती वाटणे

b) कचडकचड होणे

c) िळजी वाटणे

18) िोणत्या िारणामुळे तुम्हाला कबपीचा त्रास सुरू झाला?

a) तणाव

b) अिि वजन

C) मद्यपान

d) अिि वय

e) हृदय िवि ार

f) किडनी प्रॉब्लेम

g) प्रेग्नेंसी

h) रक्ताची िमतरता

i) जीवनसत्वाचे िमतरता

Thank You

A

Project Report

On

"TO STUDY THE ROLES OF FERTILIZERS ON GRAPES PRODUCTION"

Submitted to

(Department of Statistics)

Padmabhushan Vasantraodada Patil Mahavidyalay Kavathe Mahankal

By

Miss. Kadam Komal Sukhdev

Miss.Kadam Gauri Suresh

Miss. Bhosale Seema Ajay

Miss.More Aishwarya Arvind

Under the Guidance of

Prof.Koshti Sir

Prof.Patil sir

Prof. Bhosale madam

Prof.Patil madam

Through

The Principal,

Dr.Patil.M.K

Padmabhushan Vasantraodada Patil Mahavidyalaya, Kavathe Mahankal

For the academic year 2020-2021

DEPARTMENT OF STATISTIC

Padmabhushan Vasantraodada Patil Mahavidyalaya, Kavathe Mahankal

CERTIFICATE

This is to certify that the following students are working under my guidance. They have satisfactorily carried out and completed the project work for the B. Sc. Part III. They is presently submitted their project entitled.

This work is genuine, original and the project or the part of it has not been submitted for the award of any other degree.

"TO STUDY THE ROLE OF FERTILIZERS ON GRAPES PRODUCTION".

1. Miss. Kadam Komal Sukhdev.
2. Miss.Kadam Gauri Suresh
3. Miss. Bhosale Seema Ajay
4. Miss. More Aishwarya Arvind

Teacher in charge **Examiner** **Head**

•Department of Statistic•

DECLARATION

We hereby declare that the project entitled "TO STUDY THE ROLE OF FERTILIZERS ON GRAPES PRODUCTION". It is the original work carried out by us at the Department of Statistics.

This project or part of it has not been submitted previously for any other degree of any other purpose.

Place: Kavathe Mahankal

Date:

Sr. No.	**Roll No .**	**Name**
1	1	Kadam.K.S
2	2	Kadam.G.S
3	3	Bhosale.S.A
4	4	More.A. A

AKNOWLEDGMENT

We express our hearty thanks and deepest gratitude goes to Prof.Koshti Sir, who made valuable guidance and co-operation during our project work. We are rather infused by the guidance of Prof.Patil sir, Prof. Bhosale madam , Prof. Patil madam, Who put us their encouragement, kind support, and constant inspiration during our project work further we are thankful to all of our college students who gave us necessary information to carry this project.

Finally we give sincerest thanks to all our family members For keeping our spirit up and their immense support during every stage of the project work.

INDEX

INTRODUCTION

The title of the project is “To study the role of fertilizer on Grapes production. “By using this data and some statistical tools for analysing that data. In this project we reported here some results related to the objectives of the survey.The main purpose of the project is to know the role of most important factor for production of grapes and are Fertilizers . India is agriculture country . 13.7% economy is depend on agriculture .In Maharashtra state Sangli and Nashik district are famous for grapes production .In Sangli district most production of grapes in Kavathe Mahankal Tehsil. In Kavathe Mahankal Tehsil 60% crops of grapes and occupation of most farmers are grapes .In grapes production the most affected factor are Fertilizers. Therefore we decided To Study The Role Of Fertilizers On Grapes Production. We have collected data in Agran Dhulgaon, landgewadi, Shelkevadi, Kokale in Kavathe Mahankal Tehsil. i.e. information of grapes production in three different villages .The information is age, education, grapes type,soil type, Fertilizers type, income per acre and many.We collected data by direct personal investigation and Google form.

Today Fertilizers has become essential to modern agriculture to feed the growing population . Chemical fertilizers are used extensively in modern agriculture , in order to improve crop yield. Urea is most popular and widely used dry N Fertilizers . Increased crop production largely relies on the type of Fertilizers used to supplement essential nutrients for plants. Now a day, there is increased emphasis on the impact on soil environment due to continuous use of chemical fertilizers.

OBJECTIVES

- To know the role of organic and inorganic Fertilizers in grapes production.
- To compare the proportion of farmers which use organic and inorganic fertilizers.
- To study independence among soil type and variety of grapes.
- To study difference between production of two years.
- To find percentage of farmers are satisfied with the rate for grapes.
- To study which variety give more income by using Fertilizers.
- To study the correlation between expenses on Fertilizers and income of grapes.

SOURCE OF DATA & SAMPLING METHOD

Sampling is part of our day to day life which we use advertently or inadvertently. For instance housewife takes one or two grains of rice from the cooking pan and decides whether it is cooked or not, a pathologist takes few drops of blood and test for any change in blood of the whole body. In all these situations, sampling is inevitable and gives satisfactory results.

As the main aim of our project " To Study the role of Fertilizers in grapes production." We consider different varieties of grapes ,types of soil,types of Fertilizers etc.So we have selected some villages Agran Dhulgaon, landgewadi,Shelkewadi, Kokale in Kavathe Mahankal Tehsil by convinient sampling method, which provides us necessary and sufficient data within less time and more accuracy. Thus the data is collected by filling questionnaire from the farmers by the interview method.

STATISTICAL TOOLS

The following are some statistical tools leads us towards the appropriate results and finding.

1) Graphical representation:

a) *Histogram*: A histogram is a graphical representation of the distribution of numerical data . It is an estimate of the probability distribution of a continuous variables (quantitative variable) and was first introduced by Karl Pearson.

b) *Pie Diagram:* A pie Diagram or pie chart is a circular statistical graphic, which is divided into slices to illustrate numerical proportion. In a pie chart the arc length of each slice is proportional to the quantity is represent.

c) *Multiple Bar Diagram:* In multiple bar diagram two or more sets of inter-related data are represented .The technique of making simple bar chart is used to draw this diagram but the difference is that we use different shades ,colors,or dots to distinguish between different phenomenon.

2) Chi-Square test for independence of attributes:

The hypothesis that, the one factor independent of other or not, can be tested by the Chi-Square test. The hypothesis that,

H_0: Two attributes are independent of each other.

H_1: Two attributes are dependent.

Under H_0 , the test statistic is,

$X^2 = \sum \sum$(Oij ^2/Eij)-N~χ^2(m-1) (n-1)d.f

Statistic X^2 has (m-1)(n-1) d.f in formula (a), Eij is the expected frequency corresponding to (i,j)th cell of observed frequency Oij.

Under the null hypothesis,the expected frequency is,

Eij=(i th row total * j th column total)/ sample size

= (Ri*Cj)/n

Once the expected frequency are calculated , the value of Chi-Square is obtained with the help of formula (a) . To make a design about H_0, calculated value of Chi-Square is compared with the table attributes are not independent. Otherwise accept H_0.

3) Paired t-Test for difference of two population means:

Let us consider a particular situation where,

a)The sample size are equal i.e. n1= n2=n (say)

b) The sample observation (x1,x2,…..,xn) and (y1, y2,…, yn) are not completely independent but they are dependent in pairs.

Let di= xi – yi , i=1,2,3,….n denote the difference in the observations for the i th unit.

Here, we want to test ,

H_0 : μd =0 i.e. the difference between means is insignificant.

V/S

H_1 : μd≠0 i.e. the difference between means is significant.

Under H_0 , the test statistic is ,

$t = (\bar{d})/(s/\sqrt{n}) \sim t_{n-1}$

Where , d= Yi – Xi , here , Yi >Xi , $\bar{d} = \sum d_i/n$ &

$S^2 = 1/(n-1)\sum (d_i - \bar{d})^2$

Conclusion:

If calculated value of $|t| > t_{(n-1,\ alpha/2)}$,we reject the null hypothesis at alpha% l.o.s and conclude that the difference between means is significant , otherwise we accept H_0 at alpha% l.o.s.

4) Estimating parameter values by using confidence interval:

The assumption is justifiable only if the size of sample n is large and neither p nor q is very small. If n is large the value of either p or q is not very small, then the difference between them is not much and the distribution would be very mush like a normal one. We have the value of pq is can never exceed 1/2 *1/2 or 1/4 . As such, if we do not know the value of p in the universe and if the value of p in the sample is very small or very large (so that the value of q is very large or very small) then instead taking the sample values of p and q. The value of pq may be taken as 1/4 which is maximum value.

Then the standard error is,

S.E= $\sqrt{(pq/n)}$

And the confidence interval at alpha% l.o.s is given by p +-q * Z alpha.

5) Correlation:

Correlation means , " The study of existence and the magnitude and direction of variation between two or more variable".

Karl Pearson's coefficient of correlation: The most commonly used Method which gives a mathematical expression for correlation is the one suggested by Karl Pearson.

Here , sigma^2x = 1/N $\sum$(xi – x bar)^2 gives a measure of variation in x.

& Sigma^2y =1/N $\sum$(yi – y bar)^2 gives a measure of variation in y.

Also, 1/N $\sum$(xi – x bar)$\sum$(yi – y bar) gives a measure of simultaneous variation in x &y. Then the correlation coefficient is,

r= ($\sum$(xi – x bar)(yi – y bar))/($\sqrt{\sum}$(xi – x bar)^2*$\sum$(yi – y bar)^2)

The probable error and coefficient of correlation:

The reliability of the coefficient of correlation is determined by calculating the probable error. If n is the size of the sample the probable error is given by,

Probable error= 0.6745* (1- r^2)/n

Where , r is correlation coefficient.

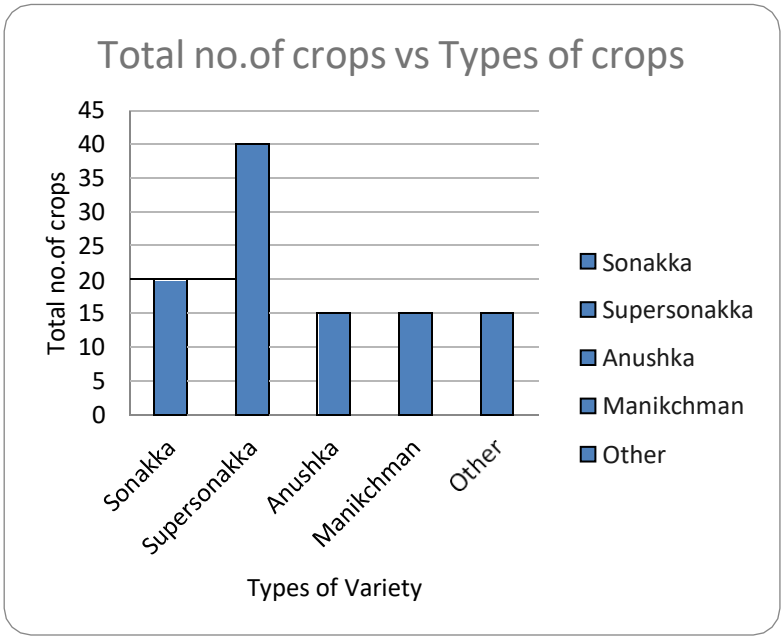
Conclusion:

If the value of r is less than probable error then the correlation is not significant. If the value of r is greater than six times the probable error then the correlation is significant.

STATISTICAL ANALYSIS

1) To study the distribution of variety of grapes:

Variety	Total
Sonakka	20
Supersonaka	41
Anushka	15
Manikchman	16
Other	14

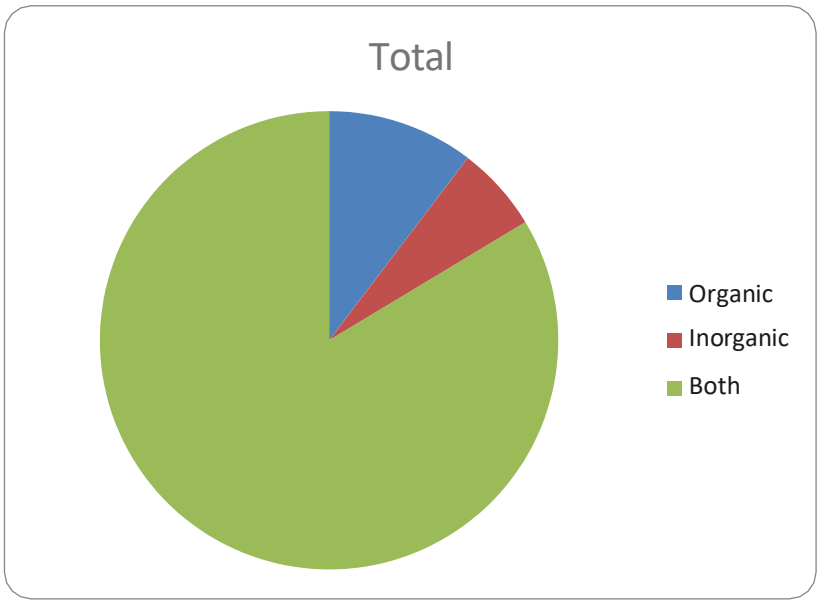


Conclusion:

From the above graph diagram we conclude that most farmers use Supersonaka variety of grapes as compare with other varieties.

2) To Study Distribution Of Fertilizers :

Fertilizers	Total
Organic	12
Inorganic	7
Both	97



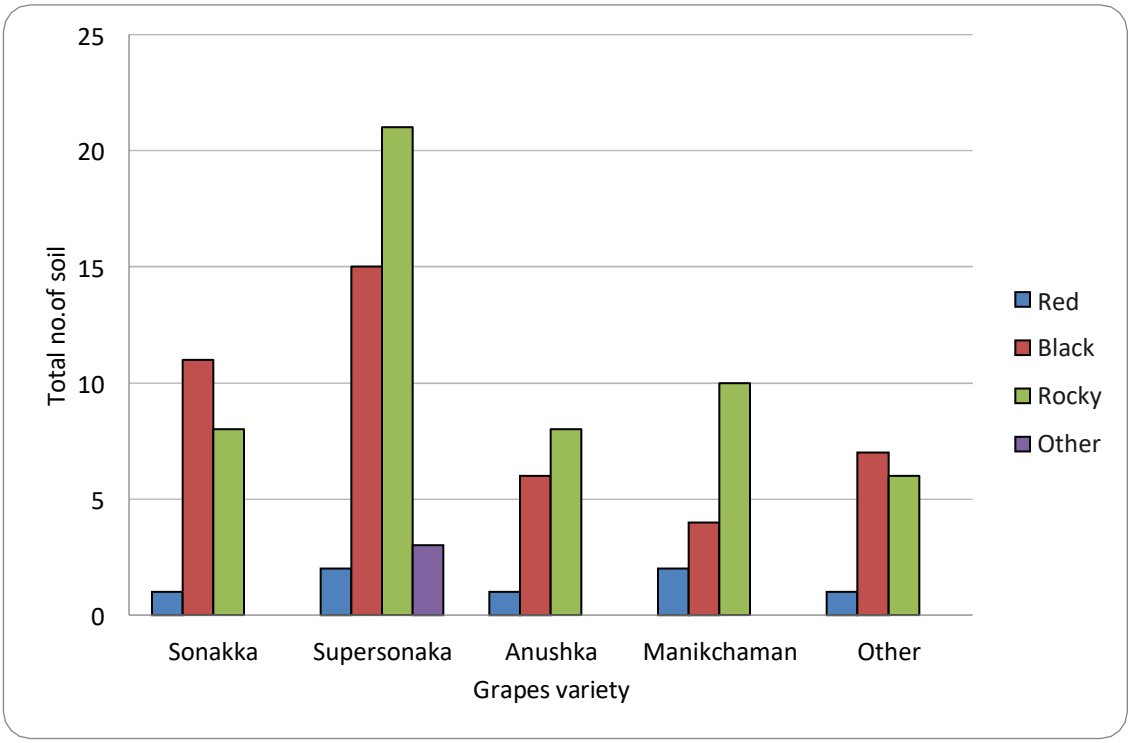
Conclusion:

From the above Pie Diagram , it is clear that most farmers uses the both fertilizers for more production of grapes. i.e. Organic as well as inorganic fertilizers play important role in more production of grapes.

Most of the farmers used both the fertilizers but use of organic fertilizers is greater than the inorganic fertilizers.

3) To study Distribution of grapes variety with soil type :

	Red	Black	Rocky	Other
Sonakka	1	11	8	0
Supersonaka	2	15	21	3
Anushka	1	6	8	0
Manikchaman	2	4	10	0
Other	1	7	6	0

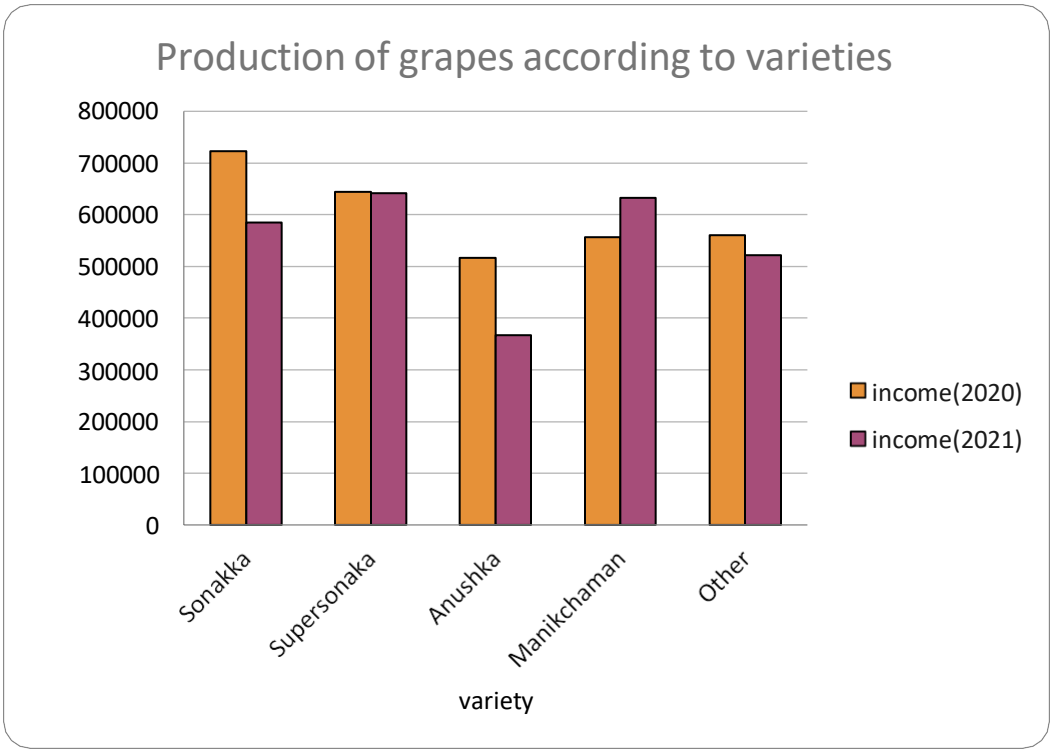


Conclusion:

From the above graph we conclude that cultivation of Supersonaka is greater in Rocky Soil compare with other soil.

4) To Study which variety of grapes give more income:

variety	Average income(2020)	Average income(2021)
Sonakka	723500	584750
Supersonaka	644268.2927	641268.2927
Anushka	516666.6667	366666.6667
Manikchaman	556250	632812.5
Other	560714.2857	521428.5714



Conclusion:

From the graph , we see that variety of Sonakka as well as Supersonaka give more income to the farmers. But Supersonaka gave most income in both 2020 & 2021 years.

5)To study independence between soil type and variety of grapes:

Here we check the soil type for the different varieties of grapes are Independent or not.

Attributes:A-variety if grapes &B-soil type

Here, A1-Sonakks, A2-Super sonakka, A3-Anushka, A4-Manikchaman

A5-Other

Here, we want to test the hypothesis that,

H0:Soil type and variety of grapes are independent

H1:Soil type and variety of grapes are not independent

Observation Table 1:

	Red	Black	Rocky	Other	Total
Sonakka	0	12	8	0	20
Super sonakka	2	15	21	3	41
Anushka	4	5	6	0	15
Manikchaman	2	4	10	0	16
Other	0	7	6	0	13
Total	8	43	51	3	105

Observation Table 2:

Sr.No	Oij	Eij	Oij^2/Eij
1	0	1.523809524	0
2	12	8.19047619	17.58139535
3	8	9.714285714	6.588235294
4	0	0.5714285714	0
5	2	3.123809524	1.280487805
6	15	16.79047619	13.40045377
7	21	19.91428571	22.14490674
8	3	1.171428571	7.682926829
9	4	1.142857143	14
10	5	6.142857143	4.069767442
11	6	7.285714286	4.941176471
12	0	0.4285714286	0
13	2	1.219047619	3.28125
14	4	6.552380952	2.441860465
15	10	7.771428571	12.86764706
16	0	0.4571428571	0
17	0	0.9904761905	0
18	7	5.323809524	9.203935599
19	6	6.314285714	5.701357466
20	0	0.3714285714	0
		Total	125.1854003

Now, Calculated X^2 is given by,

$X^2 = \sum \sum (Oij\text{^}2/Eij) - N =$ 20.1854003

The table value of X^2 with (m-1) (n-1) d.f. at 5% l.o.s. is,

$X^2(12, 0.05) = 21.026$

P – value =0.10 >alpha= 0.05

Result :

Here,

Tabulated value of chi- square is 21.026 > Calculated value chi-squre

Hence we Accept Ho .

Conclusion :

Hence we conclude that soil type and variety of grapes are independent.

6) To study the difference between production of two years:

In order to check the difference between production of two years, we use Paired t-test.

Let X1, X2,… ..... Xn be a random sample of size 105 from production Of 2020 of farmers.

Let Y1, Y2,… ..... Yn be a random sample of size 105 from production Of 2021 of farmers.

Here, we want to test the hypothesis that,

Ho:There is no significant difference between production of two years of Farmers .i.e. $\mu 1=\mu 2$

V/s

H1: The production of farmers in 2020 is greater than production of 2021

i.e. $\mu 1>\mu 2$

Under Ho. The test statistic is,

$T=\bar{d}/(s/\sqrt{n}) \sim t_{n-1}$

Where, $\bar{d}=\sum di/n$ & $S^2 =1/n-1(\sum (di-\bar{d})^2)$

Here, n=105, $S^2 = 7964039179$, $S =282206.2901$

& $\bar{d} =21361.90476$, Then,

$T(cal) =0.775653818$

$T(tab) =1.664$

Result: Here, $t(cal)<t(tab)$,

Hence, we accept Ho.

We conclude that, production of farmers in 2020 is greater than production Of 2021.

7) To compare the proportion of farmers using organic and inorganic fertilizers:

In a data, 12 farmers out of 105 farmers are found to be users of organic fertilizers.

Now, we have to discuss whether this information supports the view that majority

Of farmers are users inorganic fertilizers.

P=0.1142857143

Q=1-p=0.8857142857

S.E = $\sqrt{pq/n}$ =0.03104903799

Hence, the confidence limits for p, the proportion of farmers used organic fertilizers

At 95% confidence level is,

0.1142857143+-0.0310493799*1.96=0.0534289297 to 0.1751454989

Conclusion:

The view that majority of farmers used inorganic fertilizers is equivalent to the View that majority of farmers used organic fertilizers which amounts to situation That the proportion p of farmers used organic fertilizers should always be less than 0.5 since, the confidence limits for p include values more than 0.5 in this case , the The given information does not support the view that majority of farmers used Inorganic fertilizers at 95% confidence level.

8) To study the correlation between expense of fertilizers and income of grapes:

We have to study the correlation between expense of fertilizers and income of Grapes. For this we use Karl Pearson's coefficient of correlation.

$$r= (\sum(xi – x\ bar)(yi – y\ bar))/(\sqrt{\sum(xi – x\ bar)^2 * \sum(yi – y\ bar)^2})$$

Here, we have

x bar	112742.8571
y bar	1168447.619
∑(xi-xbar)	-112743
∑ (yi-ybar)	-1168447. 614
∑(xi-xbar)^2	260521008970
∑(yi-ybar) ^2	26865395800252
∑ (xi-xbar) (yi-ybar)	1.317342894*10^11

Hence, r=0.04979443803

Probable error:

reliability of coefficient of correlation is determined by calculating the Probable error & is given by,

Probable error = 0.6745*(1-r^2) /n

Here we have, r=0.04978443803

Hence,

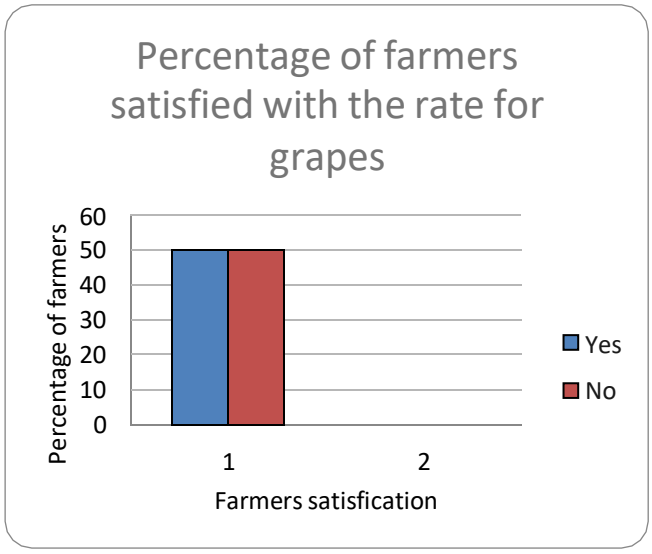
P. E= 0.006407881778

Conclusion:

Coefficient of correlation is greater than probable error hence the Correlation is significant.

9) To find percentage of farmers are satisfied with the rate for grapes:

Farmers satisfied	Yes	No	Total
percentage	48.54	51.46	100



Conclusion:

From the above graph we see that 48.54% farmers are satisfied with the rate getting for grapes and 51.46% farmers are not satisfied with the rate getting for grapes.

Overall Conclusions

1. Both the Organic as well as inorganic fertilizers play important role in production of grapes as most of the farmers used both the Fertilizers.

2. Types of soil and variety of grapes are independent.

3. Production of farmers in 2020 is greater than the production of 2021.

4. Supersonaka is the variety of grapes which give more income for the farmers.

5. There is a positive correlation between expenses on Fertilizers and income of grapes & it is significant.

6. Most of the farmers are not satisfied with the rate getting for grapes and 48% farmers are satisfied with the rate getting for grapes.

7. The most of the farmers used variety of grapes is Super Sonakka.

LIMITATIONS

1) We have to limit the scope of this research according to the result we obtain from the questionnaires.

2) This survey is limited to P.V.P Mahavidyalaya Kavathe Mahankal.

3) The time Constraints was one of the major limitations.

SCOPE

1) A well planned Study can be taken including some relevant characteristics.

2) It can be extended for greater region for better estimates.

BIBLIOGRAPHY:

REFERRED BOOKS:

Sr. No.	Book Name	Authors
1	Basic Statistics	B. L. Agarwal
2.	Fundamentals of statistics.	A.C.Gupta & V.K.Kapoor

Website and research engines:

1. Wikipedia
2. Google

QUESTIONNAIRES

1) द्राक्ष लागवडी खालील एकू ण कती क्षेत्र आहे.

2) तुम्ही कोणत्या पद्धतीने शेती करता?

a) पारंपररक

b) आधुननक

3) तुमच्या बागेसाठी लागणारी जमीन कोणत्या प्रकारची आहे?

a) काळी b) खडकाळ डी d)इतर
C) ताब

4) बागेला पाणीपुरवठा कसा करता?

a) ठठबक ससचन b) भुपाठ

5) तुम्ही कोणत्या जातीचे पीक घेता?

a) सोनाक्का b) सुपरसोनक्क

c) अनुष्का D) मननकचमन e) इतर

6) बागेवरती कधी रोगाचा प्रभाव झाला आहे का?

a) हो b)नाही

7) असेल तर कोणत्या रोगाचा प्रभाव झाला आहे?

a)थrips b)बुरशी c)डाऊनी d)इतर

8) बागेसाठी कोणत्या प्रकारची खते वापरता?

a) सेंद्रीय खत b) रासायननक खत c) दोन्ही

9) तुम्ही कोणत्या पद्धतीचे सेंठद्रय खत वापरता?

a) गांढुळखत b) शेणखत c) कोंबडीखत

d) स्लरी e) इतर

10)तुम्ही कोणत्या पद्धतीचे रासायनीक खत वापरता?

a)SOP b)NPK c)Urea

d)Ammonium sulphate e) इतर

11) खतांच्या वापरासाठी कोणाचं सल्ला घेता?

a) अनुभवी शेतकरी b) कृषितज्ञ

c) इंटरनेट/सोशलमीडया d)इतर

12) बागेवरती कोणत्या खतांचा चांगला पररणाम

जाणवतो?

a) सेंठद्रय b)रासाययननक c)दोन्ही

13) द्राक्ष बागेवरती हवामानाचा पररणाम जाणवतो का?

a) होय b) नाही

14) सेंठद्रय खतांचा वर्षिक सरासरी खर्च ककती येतो?

15) रासायनीक खतांचा वर्षीक सरासरी खर्च ककती येतो?

16) प्रतीएकर ककती उत्पन्न

समळतेे? a)2020-

b)2021-

17) खर्चाच्या तुलनेत अपेक्षेप्रमाणे उत्पन्न समळतेे का?

a) होय b) नाही

18) जर वीजी असणारा द्राक्षेचा दर तुम्हाला योग्य वाटतो का?

a) होय b) नाही

19) जर वीजी असणारा द्राक्षेच्या दराबद्दल शासनाला काय सांगू इच्छित का?

a) होय b) नाही

20) द्राक्ष शेती फायदेशीर वाटते का?

a) होय b) नाही

A Project Report on

"STATISTICAL ANALYSIS OF EFFECT OF FERTILIZERS AND SOIL TYPE ON SUGARCANE"

Submitted to

Padmabhushan Vasantraodada Patil Mahavidyalaya, Kavathe mahankal.

Department of Statistics

By

Miss. Babar Aishwarya Sanjay

Miss. Patil Bhagyashri Harishchandra

Miss. Chavan Sneha Shrikant

Miss. Patil Vaishnavi Vilas

Under the Guidance of

Prof.(Dr.) M.K.Patil (Head of department)

Assoc.Prof.V.V.Koshti

Mr.A.P.Patil

Miss.S.S.Bhosale

Miss.S.S.Patil

Through

The Principal,

Dr.M.K.Patil

Padmabhushan Vasantraodada Patil Mahavidyalaya, Kavathe mahankal

For The Academic Year 2020-2021

DEPARTMENT OF STATISTICS

PADMABHUSHAN VASANTRAODADA PATIL MAHAVIDYALAYA, KAVATHE MAHANKAL

CERTIFICATE

This is to certify that the following students are working under

My guidance. They have satisfactorily carried out and completed the project work

For the B.Sc. Part III. They have presently submitted their project entitled.

This work is genuine and the project or the part of it has not been

Submitted for the award of any other degree.

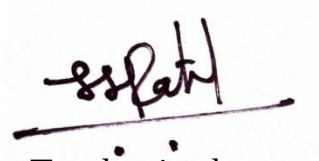
"STATISTICAL ANALYSIS OF EFFECT OF FERTILIZERS AND SOIL TYPE ON SUGARCANE"

1. Miss. Babar Aishwarya Sanjay

2. Miss. Patil Bhagyashri Harishchandra

3. Miss. Chavan Sneha Shrikant

4. Miss. Patil Vaishnavi Vilas



Teacher in charge

Examiner

Head of Department of statistics

DECLARATION

We hereby declare that the project entitled "STATISTICAL ANALYSIS OF EFFECT OF FERTILIZERS AND SOIL TYPE ON SUGARCANE" is the original work carried out by us at the Department of statistics.This project or part of it has not been submitted previously for any other degree of any other purpose.

Place: Kavathe mahankal

Date:

Sr. No.	Roll No.	Names
1		Babar Aishwarya Sanjay
2		Patil Bhagyashri Harishchandra
3		Chavan Sneha Shrikant
4		Patil Vaishnavi Vilas

ACKNOWLEDGEMENT

We express our hearty thanks and deepest gratitude goes to Dr. M.K.Patil Sir and Mr.V.V.Koshti Sir, who provided valuable guidance and co-operation during our project work. We are rather infused by the king's guidance of Mr.A.P.Patil , Miss.S.S.Bhosale, Miss. S.S.Patil, Who put us through their encouragement, kind support, and constant inspiration during our project work. Further we are thankful to all our college students and farmers who gave us necessary information to carry this project.

Finally we give sincerest thanks to all our family members for keeping our spirit up and their immense support during every stage of the project work.

INDEX

INTRODUCTION :

The title of the project is "Statistical analysis of the effect of fertilizers and soil type on sugarcane yield". By this data and using some statistical tools for analyzing that data. In this project we reported here some results related to the objectives of the survey. The main purpose of the project is to know the role of the most important factor for production of sugarcane yield and fertilizers. India is an agricultural country. 16.1% of the economy is dependent on agriculture(by the reference of the internet). In Maharashtra and Karnataka state Sangli and Belgaum district are famous for sugarcane production. In Sangli district most production of sugarcane is in Kavathe Mahankal Tehsil. In Kavathe Mahankal Tehsil 60% yield of Sugarcane and occupation of most farmers are sugarcane. In Sugarcane production the most affected factor are fertilizers. Therefore we decided to study the role of fertilizers in the yield of Sugarcane production. We have collected data in Jayagavhan,Kuchi, Dhulagaon ,Irali in Kavathe Mahankal Tehsil and Khilegaon-Pandegaon in Athani Tahsil. (We choose those villages because most of the people in that village grow sugarcane) In these four villages data was collected i.e. information of yield of Sugarcane production in four different villages. The information is age, education ,types of Sugarcane, soil type, fertilizers types, income per acre, loss per acre and many others. We collected data by direct personal investigation or interview/(Google forms). We proceed to study how the observations are classified and a frequency distribution is formed.

Nowadays fertilizer has become essential to modern agriculture to feed the growing population. Chemical fertilizers are used extensively in modern agriculture, in order to improve crop yield. Urea is the most popular and widely used dry N fertilizer. Increased crop production largely relies on the type of fertilizers used to supplement essential nutrients for plants. Fertilizer application is required to replace crop land nutrients that have been consumed by previous plant growth with the ultimate goal of maximizing productivity and economic returns. Nowadays, there is increased emphasis on the impact on the soil environment due to continuous use of chemical fertilizers.

OBJECTIVES

- To know the best season for sugarcane planting.
- To know the role of different types of fertilizers for Sugarcane production.
- To find average quantity of which fertilizer used per acre by the farmers for effective production of Sugarcane.
- To know which diseases are most prevalent in sugarcane crops.
- To compare the proportion of farmers which use organic and inorganic fertilizers on Sugarcane production.
- To study independence among soil types and varieties used for cultivation.
- To study the correlation between expense on fertilizers and income of sugarcane.

SOURCE OF DATA COLLECTION AND SAMPLING METHOD

Sampling is a part of our day to day life which we use advertently or inadvertently. For instance, a housewife takes one or two grains of rice from the cooking pan and decides whether it is cooked or not, a pathologist takes a few drops of blood and tests for any change in blood of the whole body. In all these situations, sampling is inevitable and gives satisfactory results.

As the main aim of our project "STATISTICAL ANALYSIS OF EFFECT OF FERTILIZERS AND SOIL TYPE ON SUGARCANE" we consider different varieties of Sugarcanes, types of soil, types of fertilizers etc. So we have selected some villages as Jaygavhan, Kuchi, Dhulagaon,Irali in Kavathe Mahankal Tehsil and Khilegaon-Pandegaon in Athani Tahsil, by convenient sampling method, which provide us necessary and sufficient data within less time and more accuracy. Thus the data is collected by filling questionnaires from the farmers by the interview method / Google forms (because of these Corona situation)

Statistical Tools

1] GRAPHICAL REPRESENTATION:

a) HISTOGRAM: A histogram is a graphical representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable (quantitative variable) and was first introduced by Karl Pearson.

b) PIE DIAGRAM: A pie diagram or pie chart is a circular statistical graphic, which is divided into slices to illustrate numerical proportion. In a pie chart the arc length of each slice is proportional to the quantity it represents.

c) MULTIPLE BAR DIAGRAM: In multiple bar diagrams two or more sets of interrelated data are represented. The technique of making a simple bar chart is used to draw this diagram but the difference is that we use different shades, colors, or dots to distinguish between different phenomena.

d) LINE CHART:A line graph is a graphical display of information that changes continuously over time. Within a line graph, there are various data points connected together by straight lines that represent values represented by the data points. A line chart is used to represent data over a continuous time span. It is generally used to show the trend of a measure (or a variable)over a time. Data values are plotted as points that are connected using line segments.

e)BAR-CHART : A bar chart or bar graph is a chart or a graph that present categorical data with rectangular bars with heights or length of proportional to the values that it represents. The bars can be plotted vertically or horizontally. A vertical bar chart is sometimes called column chart

2] ESTIMATING PARAMETER VALUES BY USING CONFIDENCE INTERVAL:

The assumption is justifiable only if the size of sample n is large and neither p nor q is very small. If n is large the value of either p or q is not very small, then the difference between them is not much and the distribution would be very much like a normal one. We have the value of pq is can never exceed $\frac{1}{2} * \frac{1}{2}$ or $\frac{1}{4}$. As such, if we do not know the value of p in the universe and if the value of p in the sample is very small or very large (so that the value of q is very large or very small), then, instead taking the sample values of p and q, the values of pq may be taken as $\frac{1}{4}$ which is the maximum value.

Then the standard error is,

$$S.E. = \sqrt{\frac{pq}{n}}$$

And the confidence interval at α% l.o.s. is given as, $p \pm q * Z_{\alpha}$

3] CHI-SQUARE TEST FOR INDEPENDENCE OF ATTRIBUTES:

The hypothesis that one factor independent of another or not, can be tested by the chi-square test. The hypothesis that,

$H_{0:}$ Two attributes are independent of each other.

H_1: Two attributes are dependent.

Under H_0, the test statistics is,

$$X^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{Oij}{Eij} - N \sim X^2_{(m-1)(n-1)}$$

Statistic X^2 has (m-1) (n-1) d.f. in formula (a), E_{ij} is the expected frequency corresponding to (i, j) th cell of observed frequency $O_{ij.}$

Under the null hypothesis, the expected frequency is,

$$E_{ij} = \frac{ith\ row\ total * ith\ column\ total}{sample\ size} = \frac{Ri*Cj}{n}$$

Once the expected frequencies are calculated, the value of chi – square is obtained with the help of formula (a). To make a design about H_0, the calculated value of chi – square compared with the table attributes are not independent. Otherwise accept H_0.

4] Correlation:

Correlation means, "The study of existence and the magnitude and direction of variation between two or more variables".

Karl Pearson's coefficient of correlation:

The most commonly used method which gives a mathematical expression for correlation is the one suggested by Karl Pearson.

Here, $\sigma^2_x = \frac{1}{N}\sum(xi - \bar{x})^2$ gives a measure of variation in x

& $\sigma^2_y = \frac{1}{N}\sum(yi - \bar{y})^2$ gives a measure of variation in y

Also, $\frac{1}{N}\sum(xi - \bar{x})\sum(yi - \bar{y})$ gives a measure of simultaneous variation in x & y. Then the correlation coefficient is,

$$r = \frac{\sum(xi-\bar{x})(yi-\bar{y})}{\sqrt{\sum(xi-\bar{x})^2\sum(yi-\bar{y})^2}}$$

The probable error and the coefficient of correlation:

The reliability of the coefficient of correlation is determined by calculating the probable error. If n is the size of the sample the probable error is given by,

Probable error= $0.6745 * \frac{(1-r^2)}{n}$

Where, r is the correlation coefficient.

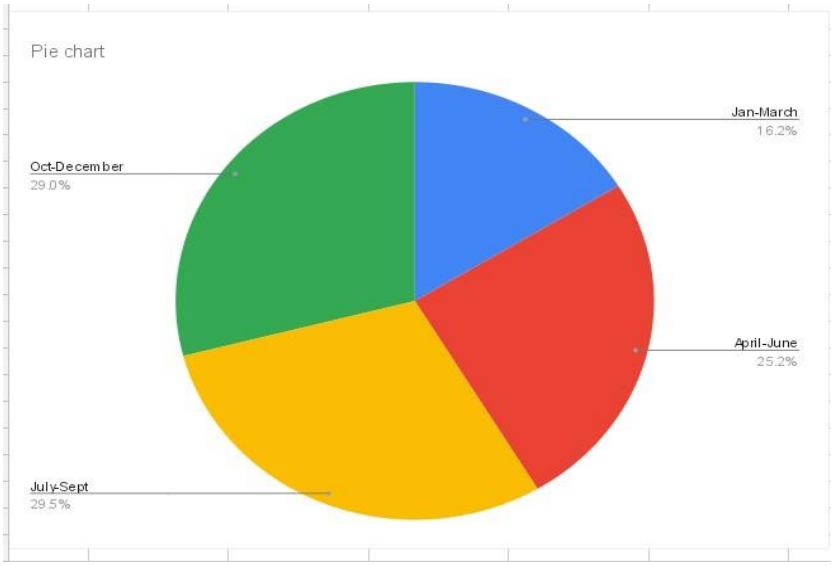
Result If the value of r is less than probable error then the correlation is not significant. If the value of r is greater than six times the probable error then the correlation is significant.

STATISTICAL ANALYSIS

1] To know the best season for sugarcane planting :(Depending on the season in which most farmers harvest)

Season	Toral(Farmers)
Jan-March	34
April-June	53
July-Sept	62
Oct-Decembe	61

Percentage of best season for planting of Sugarcane

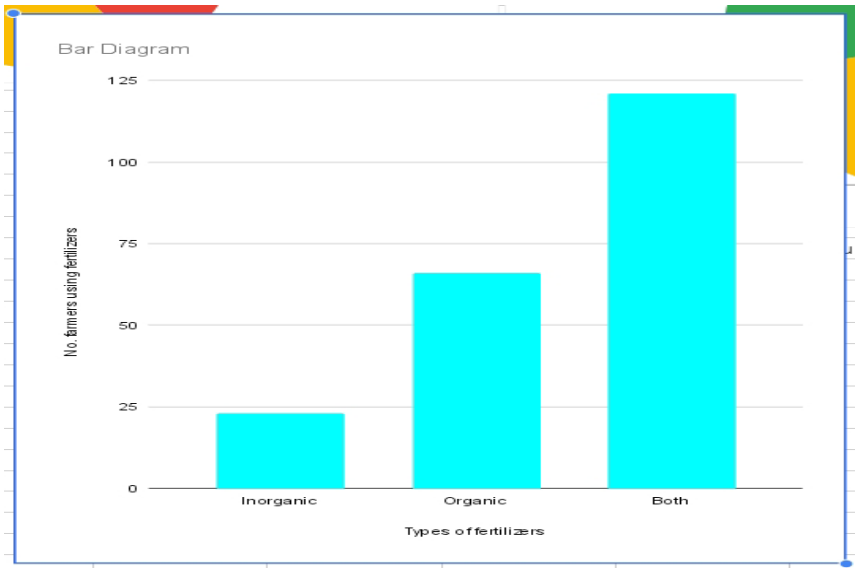


Conclusion:From the above Pie-diagram we observed that , there is higher percentage in July-sept season i.e.29.5%(depending on the season in which most farmers harvest)

2] To know the role of fertilizers in Sugarcane production:

Fertilizers Types	Total(fertilizer type used by farmers)
Inorganic	23
Organic	66
Both	121

Average quantity of fertilizer type used by farmers



Conclusion:

From the above Bar-diagram it is clear that, about 121 farmers use both i.e. organic & inorganic fertilizers for more production of Sugarcane yield .

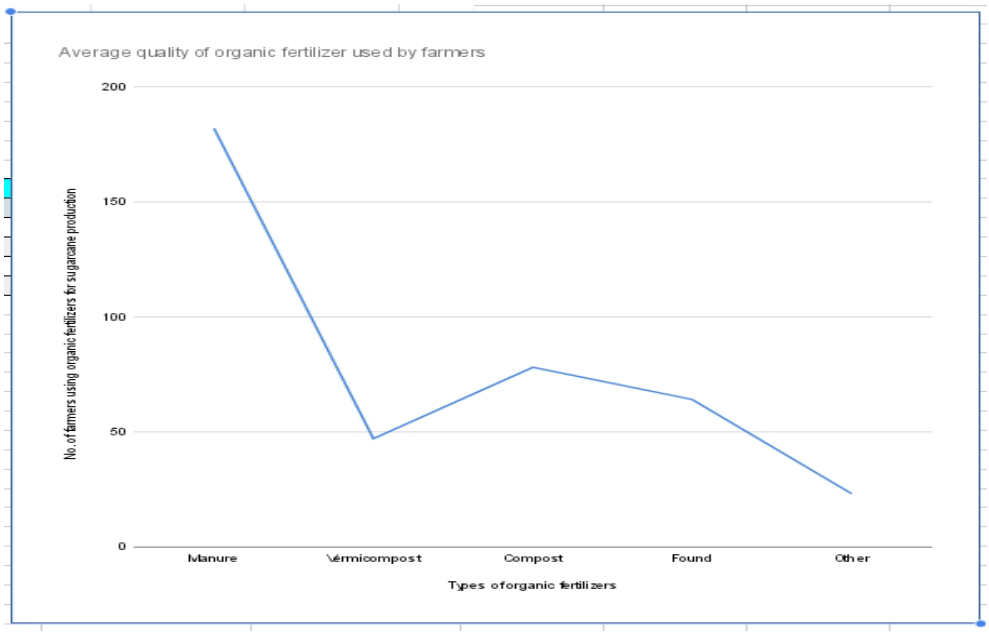
Most of the farmers used both fertilizers but use of organic fertilizers is greater than inorganic fertilizers.

3] To find the average quantity of fertilizers used per acre by the farmers for effective production of Sugarcanes yield:(ton per acre)

i) For organic fertilizers:

Organic fertilizer	No.of farmers
Manure	182
Vermicompost	47
Compost	78
Found	64
Other	23

Average quantity of organic fertilizers used by farmers(ton per acre

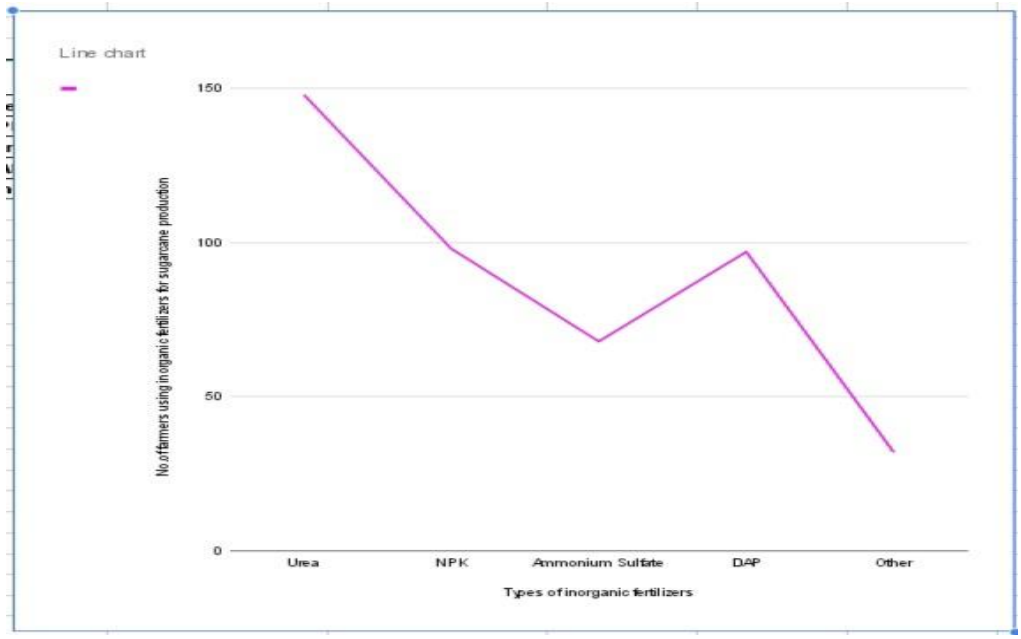


Conclusion:From the above line-chart, we see that above the 150 farmers used manure fertilizer, above 60 farmers used compost and found , less than 60 farmers used vermicompost and other.

ii) For inorganic fertilizer:(Kg per acre)

Inorganic fertilizer	No.of farmers
Urea	148
NPK	98
Ammonium sulfate	68
DAP	97
Other	32

Average quantity of inorganic fertilizer used by farmers (kg per acre)



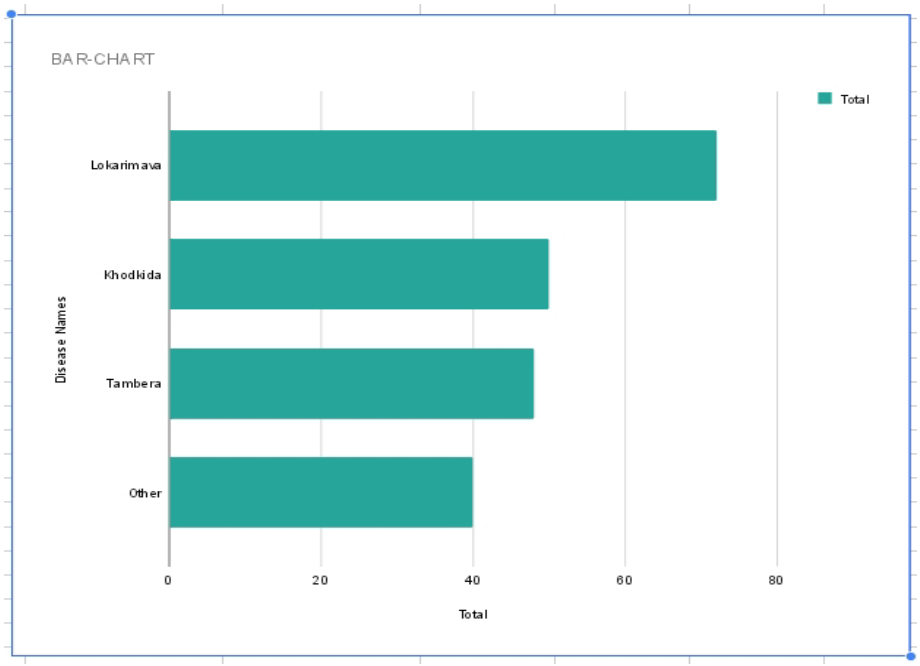
Conclusion:

From the above line chart , we see that above 100 farmers are using urea, above 50 and below 100 farmers are using NPK, Ammonium sulfate, DAP and below 50 farmers used other inorganic fertilizer.

4]To know which diseases are most prevalent in sugarcane crops:

Disease Names	Total
Lokarimava	72
Khodkida.	50
Tambera	48
Other	40

For finding most prevalent disease in Sugarcane crop



Conclusion: From the above Bar- chart we conclude that disease of Lokari Mava is more Prevalent in Sugarcane Crop.

5] To compare the proportion of farmers using organic and inorganic fertilizers:

In the data, 66 farmers out of 210 farmers are found to be users of organic fertilizers. Now, we have to discuss whether this information supports the view that the majority of farmers are users of inorganic fertilizers.

p= 0.3142

q= 1-0.3142 =0.6858

S.E. = $\sqrt{\frac{pq}{n}}$ =0.03203

Hence, the confidence limits for p, the proportion of farmers used organic fertilizers at 95% confidence level is,

0.31428571 $\pm$ 0.032031*1.96 = 0.25150 to 0.3770

Result:

The view that majority of farmers used inorganic fertilizers is equivalent to the view that majority of farmers used organic fertilizers which amounts to situation that the proportion p of farmers used organic fertilizers should always be less than 0.5 since, the confidence limits for p include values more than 0.5 in this case, the given information does not support the view that majority of farmers used inorganic fertilizers at 95% confidence level.

Conclusion:

There is a majority in use of organic fertilizers as compared with inorganic fertilizer.

6] To study independence between soil type & variety of Sugarcane yield :

Here we check the type of soil for the different varieties of Sugarcanes are independent or not.

Attributes: - A-variety of Sugarcanes & B- soil type

Here, A1-Nira86032, A2-265, A3-671, A4-8014

B1-Rocky, B2-Red clay, B3-Alluvial, B4-Black,B5-other

Here, we want to test the hypothesis that,

H_0: Soil type and variety of sugarcanes are independent.

H_1: Soil type and variety of Sugarcanes are not independent.

Observation Table 1]:

	Rocky	Red clay	Alluvial	Black	Other	Total
Nira86032	15	16	6	77	0	114
265	9	11	2	18	5	45
671	7	5	4	16	1	33
8014	4	1	5	8	0	18
Total	35	33	17	119	6	210

Observation Table 2]:

Sr.no.	Oij	Eij	Oij2/Eij
1	15	19	11.8421
2	16	17.9142	14.2903
3	6	9.2285	3.9009
4	77	64.6	91.7801
5	0	3.2571	0
6	9	7.5	10.8
7	11	7.0714	17.1111
8	2	3.6428	1.0980
9	18	25.5	12.7058
10	5	1.2857	19.4446
11	7	5.5	8.9090

12	5	5.1857	4.8209
13	4	2.6714	5.9893
14	16	18.7	13.6898
15	1	0.9428	1.06067
16	4	3	5.3333
17	1	2.8285	0.3535
18	5	1.4571	17.1573
19	8	10.2	6.2745
20	0	0.5142	0
			=276.7406

Now, calculated X^2 is given by,

Here, m=5 & n=4

$$X^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{Oij}{Eij} - N = 276.7406-210=\underline{66.7406}$$

The table value of X^2 with (m-1) (n-1) d.f. at 5% l.o.s. is,

$$X^2_{(12,\ 0.05)} = 21.026$$

Result:

Here, calculated X^2 >table X^2.

Therefore we reject H_0, we conclude that type of soil and the variety of Sugarcanes crops are not independent.

Conclusion: There is a relationship between soil type and varieties of Sugarcanes for the production of Sugarcanes yield.

7] To study the correlation between expense on fertilizers and income of Sugarcane:

We have to study the correlation between expense on fertilizers and income of Sugarcane . For this we use Karl Pearson's coefficient of correlation,

$$r = \frac{\sum(xi-\bar{x})(yi-\bar{y})}{\sqrt{\sum(xi-\bar{x})^2 \sum(yi-\bar{y})^2}}$$

Here, we have,

$\bar{X}$=24525.97613, $\bar{Y}$=10094.33160, $\sum$ (Xi- $\bar{X}$) =5150454.738

$\sum(Yi-\bar{Y})$=2119809.524 $\sum(Xi-\bar{X})^2$ =2.6527× 10^{13}, $\sum(Yi-\bar{Y})^2$ =4.4936×10^{12}

$\sum$ (Xi-$\bar{X}$)(Yi-$\bar{Y}$) = 1.0918×10^{13}

Hence, r = 0.97846

Probable error:

The reliability of coefficient of correlation is determined by calculating the probable error & it is given by,

$$\text{Probable error} = 0.6745 * \frac{(1-r^2)}{n}$$

Here we have, r=0.0690018

Hence,
P.E. = 0.0001368

Result:

Here, we see that there is a positive correlation between expense on fertilizers and income of Sugarcane yield.There is a linear relationship between expense and income of Sugarcanes yield.

Also, the probable error is greater than coefficient of correlation hence the correlation is significant

Conclusion: As we use more fertilizers, we will have more income on Sugarcanes yield.

OVERALL CONCLUSIONS

1] Both the Organic as well as inorganic fertilizers play important role in production of Sugarcane as most of the farmers used both the fertilizers.

2] The given data supports the view that the majority of farmers used organic fertilizers.

3] Different types of soil and varieties of Sugarcane crops are dependent.

4] From this analysis we conclude that the best season for planting Sugarcane crop is in July-Sept and Oct-December (according to farmers)

5] The Lokari Mava disease is more prevalent in Sugarcane.

6] Most of the farmers used manure fertilizer in organic fertilizer and Urea in inorganic fertilizer.

7] There is a positive correlation between expense on fertilizers and income of grapes & and it is significant.

Limitations

1. We have limits on the scope of this research according to the result we obtain from the questionnaires.

2. This survey is limited to P.V.P.college kavathe mahankal.

3. The time constraints was one of the major limitations.

4.Our population is not too large , therefore there is no restrictions over resources in terms of money or time

Scope

1. A well planned study can be taken up including some relevant characteristics.

2. It can be extended to a greater region for better estimates.

Bibliography

Referred books:

Sr. No.	Book Name	Authors
1.	Fundamentals of applied statistics	A. C. Gupta & V. K. Kapoor
2.	Basic Statistics	B. L. Agarwal
3.	Bsc.part II STATISTICS (paper VII)	Dr.K.G.Potdar,Dr.M.K.Patil, Dr.A.A.Kalagonda,Dr.A.V.Dorugade

From 1.Google

2.Wikipedia

(**a**) Tiller red rot

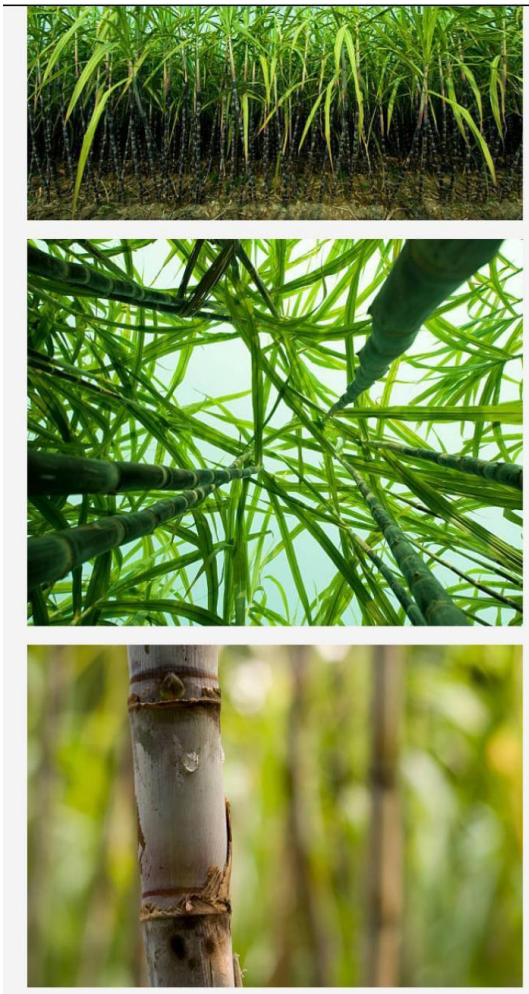
(**b**) Lamina red rot

(**c**) Midrib red rot

(**d**) Stem red rot

Ring Spot of Sugarcane | Pests ...
plantix.net

Eyespot of Sugarcane | Pests &...
plantix.net



Minor disease and ring spot disease

Thank you…….

A

Project Report

On

"TO STUDY THE TIME SPENT BY STUDENTS ON SOCIAL MEDIA"

Submitted to

(Department of Statistics)

Shikshan Prasarak Sanstha's

PADMABHUSHAN VASANTRAODADA PATIL MAHAVIDYALAYA, KAVATHE MAHANKAL.

By

Miss. Shinde Jyoti Subhash.

Miss. Mujawar Simaran Nisar.

Miss. Mali Pooja Sahadev.

Miss. Mali Sonali Vithoba.

Miss. Kore Pratiksha Jayant.

Under the Guidance of

Prof. Dr. Patil M.K. (Head of the dept.)

Asso. Prof. Koshti V.V.

Miss. Bhosale S.S.

Miss. Patil S.S.

Mr. Patil A.P.

Through

The principal

Dr.Shri.Patil M.K.

Padmabhushan Vasantraodada Patil Mahavidyalaya Kavathe Mahankal.

For the academic year 2020-2021

Shikshan Prasarak Sanstha's

PADMABHUSHAN VASANTRAODADA PATIL MAHAVIDYALAYA KAVATHE MAHANKAL.

(Department of Statistics)

CERTIFICATE

This is to certify that the following students are working under my Guidance. They have satisfactorily carried out and completed the project Work for the B. Sc. Part III. They are presently submitted their project entitled.

This work is genuine, original, and the project or the part of it has not has submitted for the award of any other degree.

"TO STUDY THE TIME SPENT BY STUDENTS ON SOCIAL MEDIA".

Miss. Shinde Jyoti Subhash.

Miss. Mujawar Simaran Nisar.

Miss. Mali Pooja Sahadev.

Miss. Mali Sonali Vithoba.

Miss. Kore Pratiksha Jayant.

Teacher in Charge	***Examiner***	***Head Department of statistic***

DECLARATION

We hereby declare that the project entitled **"TO STUDY THE TIME SPENT BY STUDENT ON SOCIAL MEDIA".**

It is the original work carried out by us at the department of statistics.

This project or part of it has not been submitted previously for any other degree or any other purpose.

Sr. No.	Roll No.	Name
1.		Shinde J. S
2.		Mujawar S. N
3.		Mali P. S
4.		Mali S. V
5.		Kore P. J

Place: Kavathe Mahankal.

Date:

ACKNOWLEDGEMENT

We express our heartly thanks and deepest gratitude goes to Prof Koshti V. V Sir, who made valuable guidance and co-operation during our project work. We are rather infused by the king guidance of Prof. Mr.Koshti V. V, Prof. Mr. Patil A.P, Prof. Miss Bhosale S. S, Prof. Miss Patil S.S.

Who put us their encouragement, kind support, and constant inspiration during our project work further we are thankful to all of our college students who gave us the necessary information to carry this project.

Finally, we give sincerest thanks to all our family members for keeping our spirit up and their immense support during every stage of the project work.

INDEX

INTRODUCTION

Social media is most important in human life and also helps to study. We study how much can social media is important in student's life and we also study the good and bad effects of social media.

According to statistics, it is estimated that, in 2020, around 3.6 billion people are using social media worldwide; up from 3.4 billion in 2019. This number is expected to increase to 4.41 billion in 2025.

We study how many students are using social media and how much time is spent, students on social media, and also check which app is more used by students. Students use social media to share information and make connections. On a personal level, social media allows due to communicating with friends and family learn new things. social media is neither inherently bad nor good it all depends on how it is being used.

The power of social media is the ability to connect and share information with anyone on earth or with many people simultaneously.

Social media is organized as a way to interact with family and friends. Social media is an interesting based form of communication. There are many forms of social media. Some of the most popular social media websites with over a hundred million registered users include whatsapp, facebook, telegram, instagram, youtube, hotstar.

METHOD OF DATA COLLECTION

We surveyed Kavathe Mahankal & Jath area and collect the data. We select the convenient sample by the random sampling method. We collect the data through the online questionnaire. 399 samples are the optimum sample size. We collect data from the online method. The data of our project is primary because we collect the data by taking online methods from each student male and female.

OBJECTIVES

Social media are being used in large quantities in the world today, and so we plan to project on this topic.

1. To find out the amount of time spent by students on social media.
2. To study which App more used by students.
3. To check most active times on social media by students.
4. To study age-wise use of social media by students.
5. To study the effects of social networking sites on the education of students.
6. To check whether social networking sites are popular in the male and female gender.

STATISTICAL TOOLS

The following are some statistical tools leads us towards the appropriate results and finding.

A) GRAPHICAL REPRESENTATION :

To represent our data graphically, we use the multiple bar diagram, pie chart method, etc. To construct the appropriable divided bar diagram of given data we used Microsoft Excel.

1. **Bar Chart:** A bar chart or bar graph is a chart that presents grouped data with rectangular bars with a lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. A vertical bar is sometimes called a column bar chart.

A bar graph is a chart that uses either horizontal or vertical bars to show comparisons among categories being compared, and the other axis represents a discrete value.

2. **Pie Chart:** A pie chart displays data, information, and statistics in an easy-to-read pie-slice format with varying slice sizes telling you how much of one data element exists. The bigger the slice the more of that particular data was gathered.

3. **Line Chart:** A line chart or line plot or line graph or curve chart is a type of chart that displays information as a series of data points called 'markers' connected by straight line segments. It is a basic type of chart common in many fields. It is similar to a scatter plot except that the measurement points are ordered typically by their x-axis value) and joined with straight line segments. A line chart is often used to visualize a trend in data over intervals of time a time series thus the line is often drawn chronologically. In these cases, they are known as run charts.

B)TEST OF INDEPENDENCE OF ATTRIBUTES:

In 2*2 Contingency Table

The hypothesis, that one factor is independent of each other or not, can be tested by the Chi-square area. The hypothesis,

Ho: Two attributes are independent of each other.

V/s

H1: Two attributes are dependent,

Attribute	B1	B2	Total
A1	a	b	a+b
A2	c	d	c+d
Total	a+c	b+d	N

The above hypothesis can be tested by the statistic,

$$\chi^2 = \frac{N[ad-bc]^2}{((a+b)*(a+c)*(c+d)*(b+d))} = \chi^2_{1\ d.f}$$

Conclusion :

If $\chi^2_{(cal)} > \chi^2_{(tab)}$ with 1 d.f , we reject Ho .

This means that the two attributes are not independent. Two attributes are independent of each other.

C)TEST FOR EQUALITY OF TWO POPULATION PROPORTIONS (P1=P2) :

Let P1 and P2 be the proportions of the unit of possing the given attribute in population. Let a random sample of n1 units be drawn from a population with the population proportion P1, in which the number of units possessing the given attribute is X1 and a random sample of n2 units be drawn from a population P2, in which the number of units possessing the given attribute is X2 . the corresponding standard normal variate is,

$$Z = \frac{(p1-p2)-(P1-P2)}{\sqrt{(\frac{P1Q1}{n1} + \frac{P2Q2}{n2})}} \sim N(0,1)$$

Here we want to test the null hypothesis, Ho: $P1 = P2$

i.e., there is no significant difference between population proportions.

V/s. $H1: P1 \neq P2$

Under the null hypothesis,

$$Z = \frac{(p11 - p21)}{\sqrt{PQ\,(\frac{1}{n1} + \frac{1}{n2})}}$$

Let α be the level of significance and Zα being the critical value.

Conclusion :

I. Suppose $H1: P1 \neq P2$ if compute$|Z| > \text{tab.}\, Z\alpha/2$ we reject Ho at $\alpha\%$ l.o.s and that there is a significant difference between population proportions.

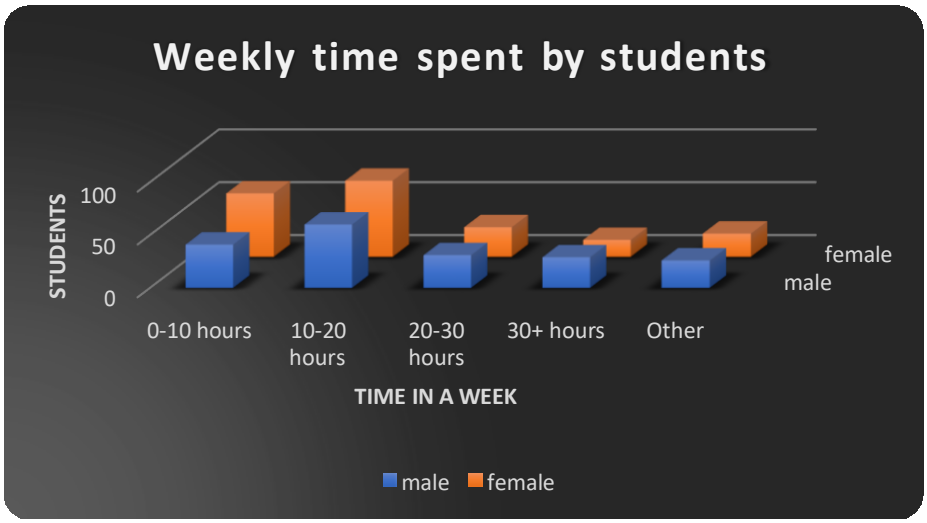
II. Suppose $H1: P1 > P2$ if computed $Z > \text{tab.}\, Z\alpha$ we reject Ho at α% l.o.s. otherwise, accept.

III. Suppose $H1: P1 < P2$, if computed $Z < \text{tab.} -Z\alpha$ we reject Ho at α% l.o.s. otherwise, accept.

STATISTICAL ANALYSIS

1. To find out the amount of time spent by students on social media.

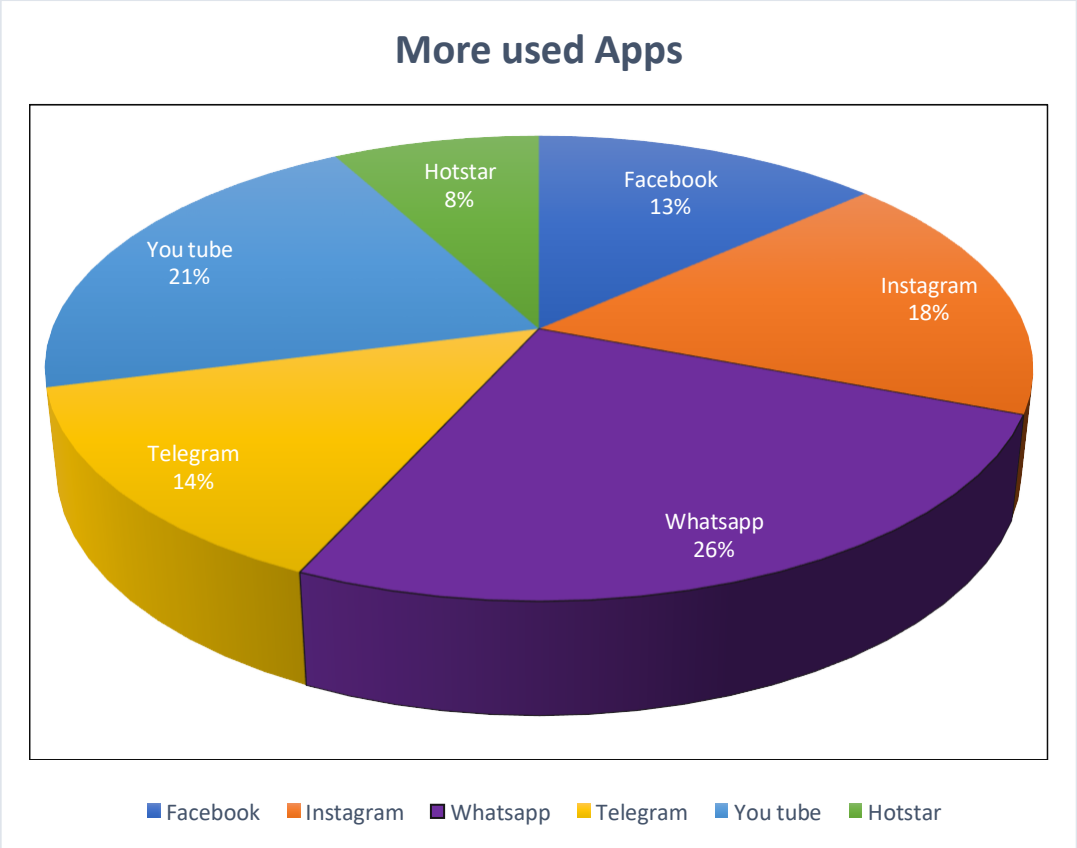
Hours in a week	Male	Female
0-10 hours	41	60
10-20 hours	60	72
20-30 hours	31	28
30+ hours	29	16
Other	26	22



Conclusion: From the above graph, we observe that the female & male students same time spent 20-30 hours a week on social media.

2. To study which App more used by students.

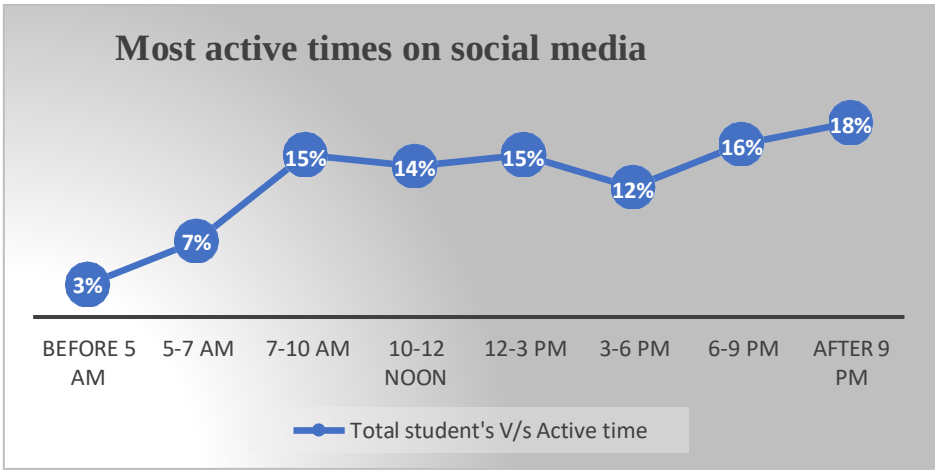
Social networking sites	Total no of students	Total no. of students in %
Facebook	183	13%
Instagram	238	18%
Whatsapp	355	26%
Telegram	194	14%
Youtube	287	21%
Hotstar	109	8%



Conclusion: From the above graph, we observed that most of the 26.0% of students use Whatsapp.

3.To check most active times on social media by students.

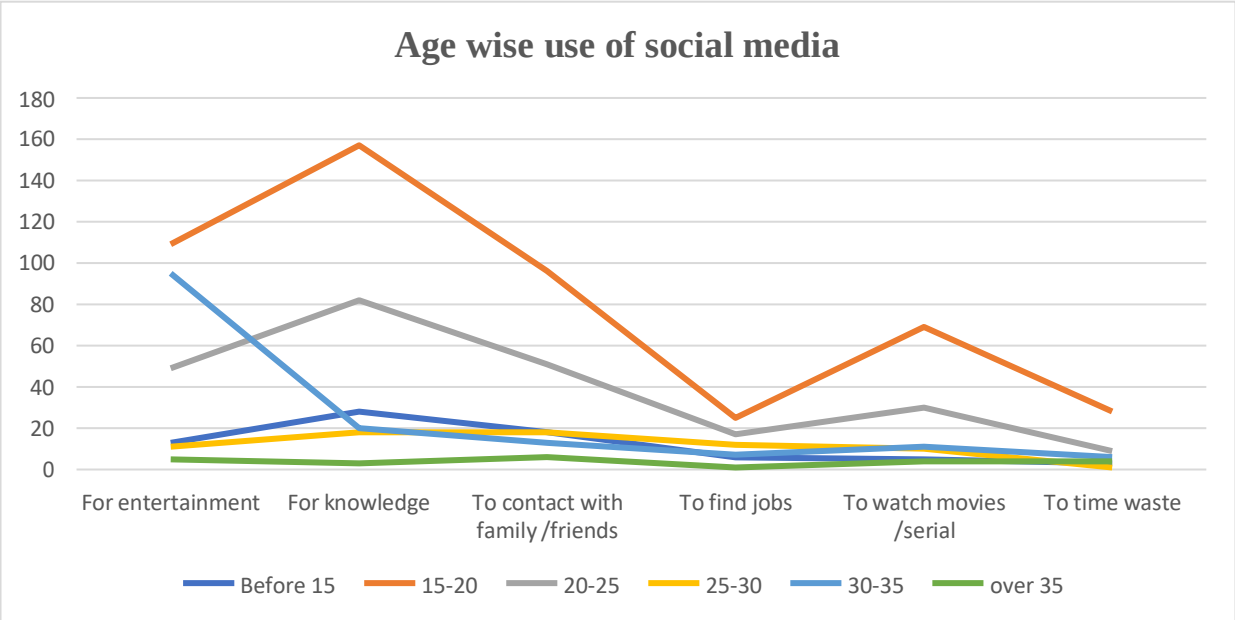
Active time	Total students	Total students in %
Before 5 am	23	3%
5-7 am	51	7%
7-10 am	101	15%
10-12 am	98	14%
12-3 am	101	15%
3-6 am	80	12%
6-9 am	110	16%
After 9 pm	126	18%



Conclusion: From the above graph, we observe that 18.3% of students are active on social media after 9 pm.

4.To study age-wise use of social media by students.

Use of social media	For entertainment	For knowledge	To contact with family and friends	To find jobs	To watch movies/ serial	To time waste
Before 15	13	28	18	6	5	3
15-20	109	157	96	25	69	28
20-25	49	82	51	17	30	9
25-30	11	18	18	12	10	1
30-35	95	20	13	7	11	6
Over 35	5	3	6	1	4	4



Conclusion: From the above graph, we observed that the 15-20 age group students most use social media.

5. To study the effects of social networking sites on the education of students.

Attribute :

A: Effect of social networking sites B: Student

Here we denoted, A : Yes α : No
B: Male β: Female
Ho: A and B are independent of each other.
H1: A and B are dependent on each other.

Statistical Table :

Attribute	A	α	Total
B	95	99	194
β	99	106	205
Total	194	205	399

The test statistic: $\chi^2 = \frac{N[ad-bc]^2}{((a+b)*(a+c)*(c+d)*(b+d))} = \chi^2_{1\ d.f}$

$\chi^2_{(cal)} = 0.0182$ $\chi^2_{(tab)} = 3.8414$

$\chi^2_{(cal)} = (0.0182) < \chi^2_{(tab)} = (3.8414)$

Result: Here, Cal χ^2 < tab $\chi^2_{(1,0.05)}$ hence, accept H0 i.e., reject H1.

i.e., A & B are independent to each other.

Conclusion: we concluded that the social networking sites have an effect on students' education.

6.To check whether social networking sites are popular in the male and female gender.

1) Z test for Facebook :

Hypothesis :

Ho: There is no significant difference between Facebook use in males and Facebook use in females. i.e. Ho: $P11 = P21$ V/s
H1: There is a significant difference between Facebook use in males and Facebook use in females. i.e. H1: $P11 \neq P21$

P11: Proportion of Facebook use in females.
P21: Proportion of Facebook use in males.

Observation Table :

n1	n2	X11(users) in n1	X21(users) in n2	p11	p21
205	194	61	126	0.2975	0.6494

Calculation:

$$P = \frac{n1p11+n2p21}{n1+n2} = 0.4685$$

Q = 1-P = 0.5314

Then calculated Z,

$$\text{Cal } |Z1| = \frac{(p11-p21)}{\sqrt{PQ(\frac{1}{n1}+\frac{1}{n2})}} = 7.0404$$

The table value of Z at 5% l.o.s., Tab |Z|= 1.96

Result:

Z1= Cal |*Z*| > Tab |Z|**,** hence we reject Ho at 5% l.o.s.

Conclusion: We conclude that the proportion of Facebook uses in males and females is not same.

2) Z test for Instagram:

Hypothesis :

Ho: There is no significant difference between Instagram use in males and Instagram use in females.i.e. Ho: $P12 = P22$

V/s

H1: There is a significant difference between Instagram use in males and Instagram uses in females. i.e. H1: $P12 \neq P22$

P12: Proportion of Instagram use in females.
P22: Proportion of Instagram use in males.

Observation Table :

n1	n2	X12(users) in n1	X22(users)in n2	p 12	p 22
205	194	121	112	0.5902	0.5773

Calculation:

$$P = \frac{n1p12+n2p22}{n1+n2} = 0.5839$$

Q = 1-P = 0.4160
Then calculated Z,

$$\text{Cal } |Z2| = \frac{(p12-p22)}{\sqrt{PQ(\frac{1}{n1}+\frac{1}{n2})}} = 0.2612$$

The table value of Z at 5% l.o.s., Tab |Z| = 1.96

Result:

Z2= Cal |Z| < Tab |Z| , here we accept Ho at 5% l.o.s.

Conclusion: We conclude that the proportion of Instagram uses in males and females is same.

3) Z test for Whatsapp :

Hypothesis :

Ho: There is no significant difference between Whatsapp use in males and Whatsapp use in females.i.e. Ho: $P13 = P23$
V/s
H1: There is a significant difference between Whatsapp use in males and Whatsapp use in females i.e. H1: $P13 \neq P23$

P13: Proportion of WhatsApp use in females.
P23: Proportion of WhatsApp use in males.

Observation Table :

n 1	n 2	X 13(users) in n1	X 23(users) in n2	p 13	p 23
205	194	182	177	0.8878	0.9123

Calculation:

$$P = \frac{n1p13+n2p23}{n1+n2} = 0.8997$$

Q = 1-P =0.1002

Then calculated Z,

$$\text{Cal } |Z3| = \frac{(p13-p23)}{\sqrt{PQ(\frac{1}{n1}+\frac{1}{n2})}} = 0.8142$$

The table value of Z at 5% l.o.s., Tab |Z| = 1.96

Result:

Z3= Cal |Z| < Tab |Z| , here we accept Ho at 5% l.o.s.

Conclusion: We conclude that the proportion of Whatsapp uses in males and females is same.

4) Z test for Telegram :

Hypothesis :

Ho: There is no significant difference between Telegram use in males and Telegram use in females.i.e. Ho: $P14 = P24$

V/s

H1: There is a significant difference between Telegram use in males and Telegram use in females. i.e. H1: $P14 \neq P24$

P14: Proportion of telegram use in females.

P24: Proportion of telegram use in males.

Observation Table :

n1	n2	X14(users) in n1	X24(users) in n2	p 14	p 24
205	194	94	99	0.4585	0.5103

Calculation :

$$P = \frac{n1p14+n2p24}{n1+n2} = 0.0125$$

Q = 1-P = 0.9874

Then calculated Z,

$$\text{Cal } |Z4| = \frac{(p14-p24)}{\sqrt{PQ(\frac{1}{n1}+\frac{1}{n2})}} = 1.4693$$

The table value of Z at 5% l.o.s., Tab |Z| = 1.96

Result: Z4= Cal |Z| < Tab |Z| , here we accept Ho at 5% l.o.s.

Conclusion: We conclude that the proportion of Telegram uses in males and females is same.

5) Z test for YouTube :

Hypothesis :

Ho: There is no significant difference between YouTube use in males and YouTube use in females.i.e. Ho: $P15 = P25$

V/s

H1: There is a significant difference between YouTube use in males and YouTube use in females. i.e. H1: $P15 \neq P25$

P15: Proportion of youtube use in females.

P25: Proportion of youtube use in males.

Observation Table :

n1	n2	X15(users) in n1	X25(users) in n2	p 15	p 25
205	194	156	138	0.7609	0.7113

Calculation :

$P = \frac{n1p15+n2p25}{n1+n2} = 0.7367$

Q = 1-P = 0.2632

Then calculated Z,

$$\text{Cal } |Z5| = \frac{(p15-p25)}{\sqrt{PQ(\frac{1}{n1}+\frac{1}{n2})}} = 1.1243$$

The table value of Z at 5% l.o.s., Tab |Z| = 1.96

Result:

Z5 = Cal |Z| < Tab |Z| , here we accept Ho at 5% l.o.s.

Conclusion:

We conclude that the proportion of Youtube use in males and females is same.

6) Z test for Hotstar :

Hypothesis :

Ho: There is no significant difference between Hotstar use in males and Hotstar use in females.i.e. Ho: $P16 = P26$
V/s
H1: There is a significant difference between Hotstar use in males and Hotstar use in females. i.e. H1: $P16 \neq P26$

P16: Proportion of Hotstar use in females.
P26: Proportion of Hotstar use in males.

Observation Table :

n1	n2	X16(users) in n1	X26(users) in n2	p 16	p 26
205	194	48	52	0.2341	0.2680

Calculation :

$$P = \frac{n1p15+n2p25}{n1+n2} = 0.2505$$

Q = 1-P = 0.7494

Then calculated Z,

$$\text{Cal } |Z6| = \frac{(p12-p22)}{\sqrt{PQ(\frac{1}{n1}+\frac{1}{n2})}} = 0.7809$$

The table value of Z at 5% l.o.s., Tab |Z| = 1.96

Result:

Z5 = Cal |Z| < Tab |Z| , here we accept Ho at 5% l.o.s.

Conclusion:

We conclude that the proportion of Hotstar uses in males & females is same.

OVERALL CONCLUSION

- Both male & female students same time spent 20-30 hours a week on social media.
- Most of the 26% of students use Whatsapp.
- 18.3% of students are active on social media after 9 pm.
- The 15-20 age group students most use social media.
- The social networking sites have an effect on students' education.
- Female students use Facebook more than male students.
- The proportion of Instagram use in female & male students is equal.
- WhatsApp use of female & male students is the same.
- The proportion of Telegram use in female & male students is equal.
- YouTube use of female & male students is the same.
- The proportion of Hotstar use in female & male students is equal.

LIMITATIONS

1) Sample size is very small to represent the whole city.

2) We have to limit the scope of this research according to the result we obtain from the questionnaires.

3) This survey is limited to Kavthe Mahankal and Jath in Maharashtra.

4) We consider the whole population is normal but in real reality is or may not be normal.

5) There is a time limit for collecting data which is not sufficient.

SCOPE

1) This project can be extended to the international level.

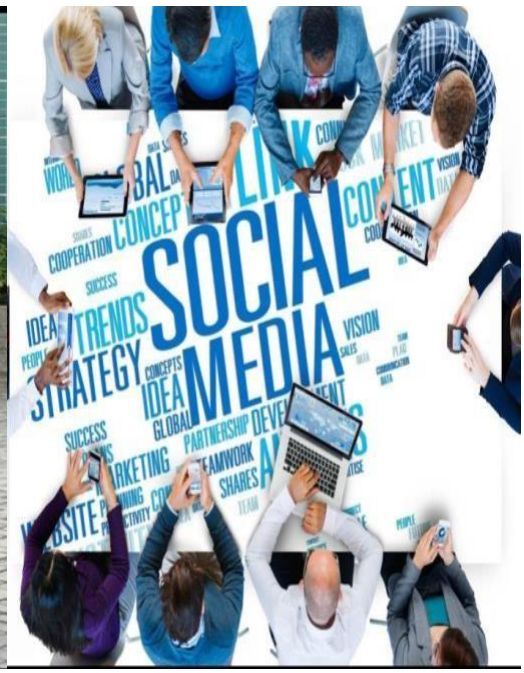
2) We will-planned Study can be taken including some relevant characteristics.

3) It can be extended to the greater region for better estimates

BIBLIOGRAPHY

Basic Statistics. B. L. Agarwal
Mathematical Statistics. Kapoor& Gupta

Websites and search engines:-
1. https://www.google.com
2. https://www.wikipedia.com



QUESTIONNAIRE

1) Gender
Female
Male

2) Age
Before 15
16-20
21-25
26-30
31-35
Over 35

3) How many types of social networking accounts do you maintain?
1
2
3
4
5
6

4) What social networking sites do you have accounts to maintain?
Facebook
Instagram
Whatsapp
Telegram
YouTube
Hotstar

5) How many times a day do you see social networking sites?
Once a day
2-5 times a day
5-10 times a day
10+ times a day

6) How many hours do you spend on social networking sites in a day?
Minimum 2 hour
2-5 hours
5-10 hours
10+ hours
Other

7) How many hours do you spend on social networking sites in a week?
0-10 hours
10-20 hours
20-30 hours
30+ hours
Other

8)What time are you most active on social media sites?
Before 5 am
5-7 am
7-10 am
10-12 noon
12-3 pm
3-6 pm
6-9 pm
After 9 pm

9) When do you use social networking sites?
When during free time
While at school/college work
Before/after mealtime
Before/after sleeping time

10) Do social networking social affect your studies?
Yes
No

11) What are the side effects of using social networking sites?
Feeling of tension
Incidence of headache
Difficulty in sleeping
Difficulty in concentration
May lead to other mental health issues

12) What are the benefits of using social networking sites?
To Help with studies
To learn new skills
Digital representations
Entertainment
Other

13) Why do you use social media?
For entertainment
For knowledge
To contact family /friends
To watch movies /serial
To time waste
Other