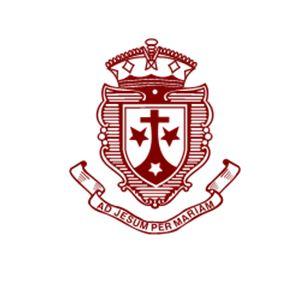
MOUNT CARMEL COLLEGE AUTOMONOUA

Department of Statistics and Analytics



Project Report done by:

1. Meghana Lokhande- MS193641[3 B.SC EMS]
2. Tanya Tibrewal-MS193666[3 B.SC EMS
3. Shreya Dey -MS193610 [3 B.SC EMS]
4. Samriddhi Roy-MS193608[3 B.SC EMS]
5. Vanshikha Agarwal-MS193652[3 B.SC EMS]

STATISTICS PROJECT REPORT

Effect of Social Media on youth

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Topic** | **Pg.no** |
| 1 | Acknowledgement | 02 |
| 2 | Introduction | 03 |
| 3 | Survey terminology | 04 |
| 4 | Survey methodology | 05 |
| 5 | Sampling technique | 06 |
| 6 | Objective | 07 |
| 7 | Target population | 08 |
| 8 | Pilot survey | 09 |
| 9 | Changes made in main questionnaire | 11 |
| 10 | Main questionnaire | 12 |
| 11 | Hypothesis | 15 |
| 12 | Representation through charts | 18 |
| 13 | Data analysis and interpretation | 25 |
| 14 | Limitations | 37 |
| 15 | Future use | 38 |
| 16 | Bibliography | 39 |

**ACKNOWLEDGEMENT**

We would like to express our deep sense of gratitude to Dr. Sr. Aparna, the principal of Mount Carmel College, and Dr. Lakshmi S.K., Head of the Department of Statistics and Analytics for allowing us to conduct our study “Effect of social media on youth”.

We would like to thank our guide, Miss Srishti Kumari, for her constant support and encouragement. We extend our gratitude to the entire department of Statistics and Analytics for allowing us to undertake this project.

Further, we would like to thank all the respondents for their contribution to our project.

**INTRODUCTION**

*The term “Social Media” refers to a computer-based technology that includes websites and applications that enable users to create and share content or to participate in social networking. Social media is a powerful tool to share, retrieve and exchange ideas and information on virtual networks and plays a pivotal role in our daily routine. Social media contains social networking sites, blogs and microblogs, wikis, discussion groups, videos and podcasts, online forums, mobile applications et cetera that has had an adaptive role on today’s Gen-Z. Thus, we, the students of Final Year, BSc. Statistics are going to do a sample study on the effect of social media on youth. The study is Bengaluru-based and solely for educational purposes.*

*Our main objectives are to analyse how much social media influencers affect the lives of youth by considering the ages 16-24 years, to interpret which gender is more influenced by social media influencers, to observe how social media life has an effect on the people’s career choices and to study the relationship between the time spent on social media and its impact on youth.*

*The main purpose of our study is to analyse the effect of social media and social media influencers on our lives using statistical tools like proportion tests, comparison tests, hypothesis tests (attributes), ratio tests, correlation, and regression concepts and these tools will also be used to compute the objectives of the study, thus enabling us to obtain aggregated data which will further help us to make an intelligent and accurate conclusion about a greater population.*

**SURVEY TERMINOLOGY**

**Social Media:** Socia*l media are interactive technologies that facilitate the* [*creation*](https://en.wikipedia.org/wiki/Content_creation) *and* [*sharing*](https://en.wikipedia.org/wiki/Information_sharing) *of information, ideas, interests, and other forms of expression through* [*virtual communities*](https://en.wikipedia.org/wiki/Virtual_communities) *and* [*networks*](https://en.wikipedia.org/wiki/Network_virtualization)*.*

***Social Networking:*** *The term social networking refers to the use of internet-based* [*social media*](https://www.investopedia.com/terms/s/social-media.asp) *sites to stay connected with friends, family, colleagues, customers, or clients. Social networking can have a social purpose, a business purpose, or both, through sites like Facebook, Twitter, LinkedIn, and Instagram.*

***Social media influencer:*** *Social media influencers are people who have influence and power in social media. The opinion of these social media influencers is valued by their followers because they are viewed as “experts” in their respective fields or niche.*

***Content Creator:*** *A content creator is someone who creates entertaining or educational material to be expressed through any medium or channel.*

**SURVEY METHODOLOGY**

***Data Source***

*Primary data was collected by conducting a sample survey using questionnaires through both offline and online modes. It was crucial to know how much impact does social media and social media influencers have on the life of the youth, how much time does an individual spend on social media, to interpret which gender is more influenced through social media and how much effect does social media have on people’s career choices.*

***SAMPLING  TECHNIQUE***

*The type of sampling method used in our survey is convenience sampling. It is a type of nonprobability sampling in which people are sampled because they are “Convenient” sources of data for researchers. Instagram polls, Google forms are some examples of convenience sampling. Data can also be obtained through convenience sampling by stopping people on the streets to fill out our survey form.*

*The questionnaire was divided into three parts.*

*1. To filter out respondents that reside in Bangalore and are between the age group of 16-24.*

*2. Personal information of the respondents which includes name, age, gender, and qualification.*

*3. Questions regarding how social media has affected their lives and the change in choices, tastes, and preferences made by them due to the impact of social media.*

***Data Compilation and Analysis:***

*After the collection of data, it was filtered, organized, and tabulated in MS Excel. The findings of this project were further used for analysis and to reach conclusions concerning our objectives.*

**OBJECTIVE**

1. *To analyse the effect of social media influencers on youth by considering the ages 16-24 years.*
2. *To interpret which gender is more influenced by social media influencers.*
3. *To observe how social media life affects people's career choices.*
4. *To study the relationship between the time spent on social media and its impact on youth.*

**TARGET POPULATION**

*Target population refers to the entire group of individuals on which researchers are interested in generalizing the conclusions. The target population usually has varying characteristics it is also known as the theoretical population.*

*Our survey is targeted on people aged 16- 24 years. We surveyed people from Bengaluru only.*

**PILOT SURVEY QUESTIONNAIRE**

**Survey on Effect of social media on Youth**

*1. NAME: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*

*2. AGE:\_\_\_\_\_\_*

*3. GENDER:\_\_\_\_\_\_\_\_\_\_\_\_\_\_*

*4. QUALIFICATION:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*

*5. Which social media channels are you most active on?*

* *Facebook*
* *Twitter*
* *Instagram*
* *YouTube*
* *Others\_\_\_\_\_\_\_\_\_\_\_\_\_\_*

*6. How much time do you spend on social media?*

* *Less than one hour*
* *1 to 2 hours*
* *2 to 3 hours*
* *More than 3 hours*

*7. Do you follow any influencers or content creators on social media?*

* *Yes*
* *No*

*8. If yes, do you think they affect your life?*

* *Yes*
* *No*
* *Kind of*
* *Depends*

*9. Why do you think you get persuaded by social media influencers?*

* *Content*
* *Lifestyle*
* *You relate to them*
* *Knowledge*

*10. Have you ever thought of becoming an influencer, if so what kind of influencer do you want to be?*

* *Food blogger/vlogger*
* *Fashion and Lifestyle*
* *Science and Technology*
* *Others\_\_\_\_\_\_*

*11. Is there any changes in your taste and preferences due to usage of social media?*

* *Yes*
* *No*

*12. Would you prefer to be a content creator than doing a job( 9 to 5 )?*

* *Yes*
* *No*
* *Both of them*

**CHANGES MADE IN PILOT SURVEY**

* *The question "Where do you live?" was added, as we realized that people outside Bangalore might also end up filling the form.*

* *In the Gender section, the options were included instead of a single dash.*

* *In the Qualification section, the different degrees were included instead of a single dash.*

**MAIN QUESTIONNAIRE**

1. Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

2. Age: \_\_\_\_\_\_

3. Gender

* Female
* Male

4. Where do you live?

* East Bangalore
* North Bangalore
* South Bangalore
* West Bangalore

5. Which social media channels are you most active on?

* Facebook
* Twitter
* Instagram
* YouTube
* Others\_\_\_\_\_\_\_\_\_\_\_\_\_\_

6.How much time do you spend on social media?

* Less than one hour
* 1 to 2 hours
* 2 to 3 hours
* More than 3 hours

7. Do you follow any influencers or content creators on social media?

* Yes
* No

8. If yes, do you think they affect your life?

* ·Yes
* No
* Kind of
* Depends

99.. Why do you think you get persuaded by social media influencers?

* Content
* ·Lifestyle
* You relate to them
* Knowledge

10. Have you ever thought of becoming an influencer, if so what kind of influencer do you want to be?

* Food blogger/vlogger
* Fashion and Lifestyle
* Science and Technology
* Others\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

11. Are there any changes in your taste and preferences due to the usage of social media?

* Yes
* No

12 . Would you prefer to be a content creator than doing a job (9 to 5)?

* Yes
* No
* Both of them

**HYPOTHESIS TO BE TESTED & STATISTICAL TOOLS TO BE USED**

**Objective 1: To analyse the most preferred social media network people are active on when compared with the other channels**

Test: Test for two proportions

H0: The proportion of users preferring Instagram is the same as the proportion of users preferring YouTube as social media handles

H1: The proportion of users preferring Instagram is not the same as the proportion of users preferring YouTube as social media handles

**Objective 2: To analyse the proportion of the male and female ratio using Instagram**

Test: Test for two proportions

H0: The proportion of females using Instagram is not significantly different as compared to the proportion of males using Instagram

H1: The proportion of females using Instagram is significantly different as compared to the proportion of males using Instagram

**Objective 3: To analyse whether following an influencer is independent of time spent on social media**

Test: Chi-square test by pooling method

H0: Following an influencer is independent of time spent on social media

H1: Following an influencer is not independent of time spent on social media

**Objective  4: Does geographical location affect the usage of social media by the youth?**

(Pie chart representation)

**Objective 5: To analyse the relationship between choosing to be a certain kind of influencer and usage of YouTube**

Test: Independence of Attribute

H0: Choosing to be a certain kind of influencer is independent of the usage of YouTube

H1: Choosing to be a certain kind of influencer is dependent on the usage of YouTube

**Objective 6: Are females more interested in food vlogging/blogging than males?**

(Pie chart representation)

**Objective 7: To analyse the change in taste and preferences due to usage of social media and content creators**

Test: Independence of Attribute

H0: Change in taste and preference due to social media is independent of following content creators

H1: Change in taste and preference due to social media is dependent on following content creators

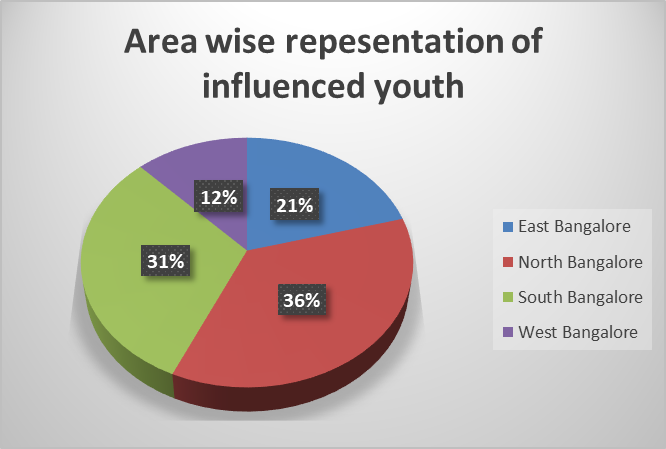
**Objective 8: What is the male-female proportion of willingness to 9 to 5 job and become a content creator?**

(Pie chart representation)

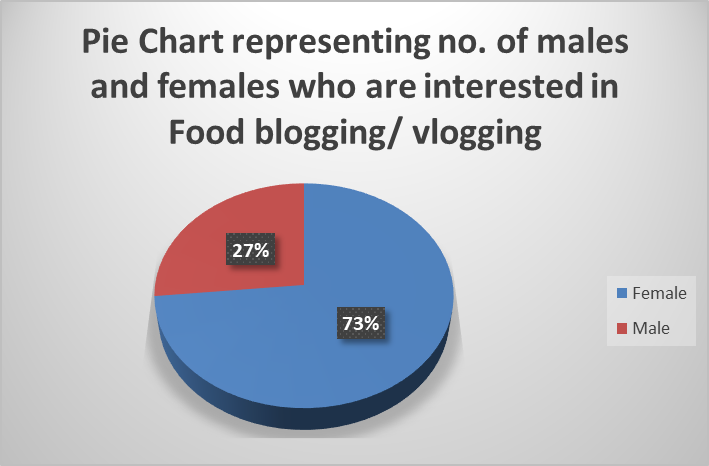
**REPRESENTATION THROUGH CHARTS**

***PIE CHARTS***

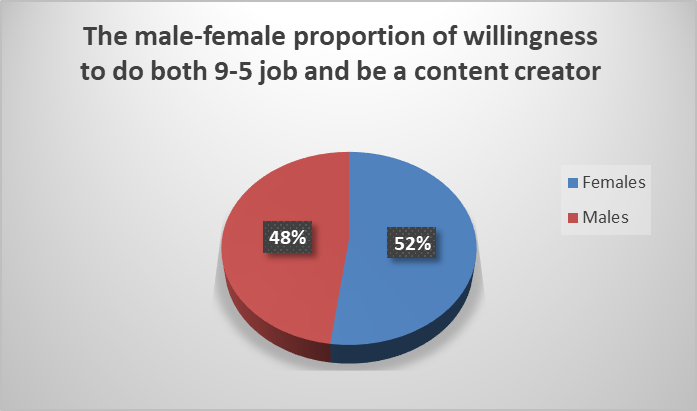
a) Does geographical location affect the usage of social media by the youth?



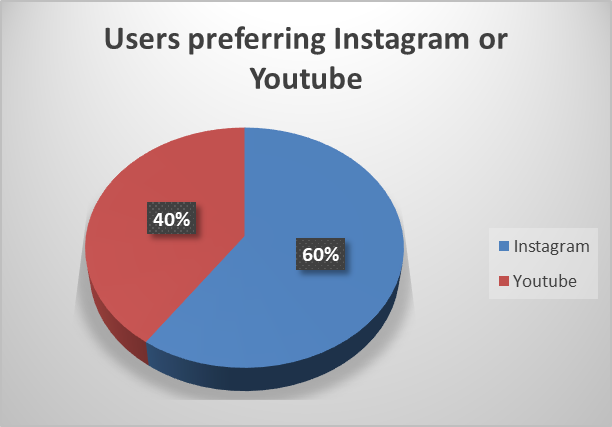
b) Are females more interested in food vlogging/blogging than males?



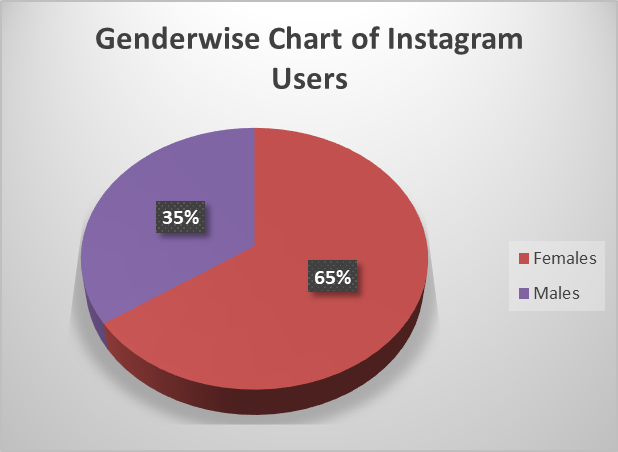
c) What is the male-female proportion of willingness to 9 to 5 job and become a content creator?



d) To analyse the most preferred social media network people are active on



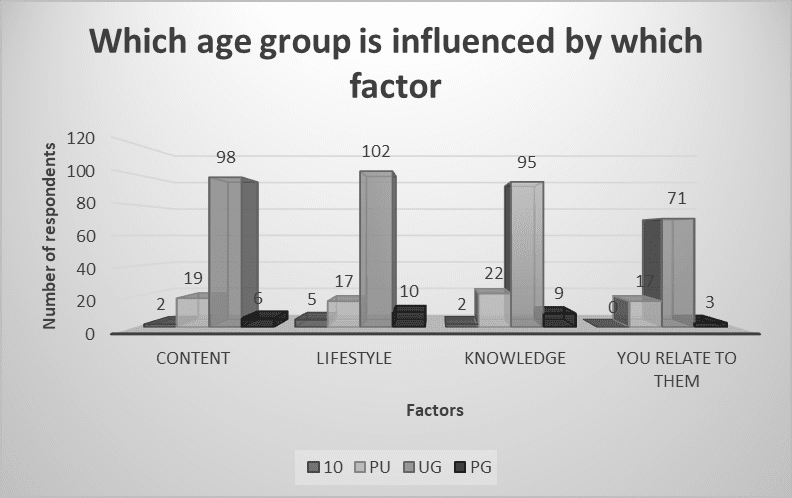
e) To analyse the usage of Instagram gender wise



***Additional Information: ( Analysis through charts)***

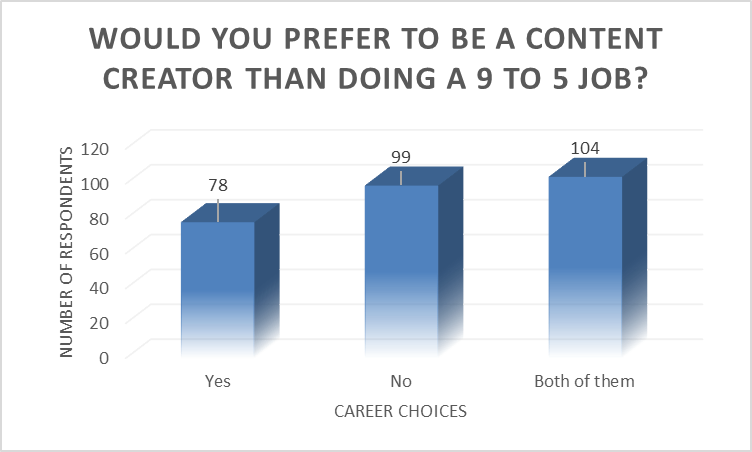
***MULTIPLE BAR CHARTS***

f) To analyse which factor influences people the most



***BAR GRAPH***

g) Do you want to do a 9 to 5 job or to be a content creator or both?



**DATA ANALYSIS AND INTERPRETATION**

***Analysis using “TEST FOR TWO PROPORTIONS”***

To analyse the most preferred social media channel by the people wrt the other channels

**Assumptions**

Let X1 be the number of people using Instagram.

Let X2 be the number of people using YouTube.

Let n1 & n2 be the total number of observations.

Here, n1=n2=289

X1~B (n1, P1)

X2~B (n2, P2)

For large values of n1 and n2, Binomial Distribution is approximated to Normal Distribution.1

P1~N (P1, P1Q1/n1) and

P2~N (P2, P2Q2/n2)

Here, α = 0.05

**Hypotheses**

H0: The proportion of users preferring Instagram is the same as the proportion of users preferring YouTube as social media handles. i.e., P1=P2

H1: The proportion of users preferring Instagram is not the same as the proportion of users preferring YouTube as social media handles. i.e., P1≠P2

**Test statistic**

Since n is large, the test statistic is given by:

Under the null hypothesis,

Here, p1= sample proportion of the number of people using Instagram = (X1/n1) =0.8408

p2= sample proportion of the number of people using YouTube = (X2/n2) =0.5674

n1 = n2 = 289

 Q = 1- P

**Computation**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Social Media Channels** | **No. of users** | **Total users** | **Prop-Value (P)** | **z-stat** | **p-value** |
| Instagram | 243 | 289 | 0.7041 | 7.1994 | 0 |
| YouTube | 164 | 289 |  |  |  |

**Conclusion**

We observe that the p-value is less than the α value. So, we reject the null hypothesis and conclude that the proportion of users preferring Instagram is not the same as the proportion of users preferring YouTube as a social media channel.

For any value of α, we accept the null hypothesis if p-value > α.

                             we reject the null hypothesis if p-value <= α.

***Analysis using “TEST FOR TWO PROPORTIONS”***

To analyse the gender wise usage of Instagram

**Assumptions**

Let X1 be the number of males using Instagram.

Let X2 be the number of females using YouTube.

Let n1 & n2 be the total number of observations.

Here, n1=n2=289

X1~B (n1, P1)

X2~B (n2, P2)

For large values of n1 and n2, Binomial Distribution is approximated to Normal Distribution.1

P1~N (P1, P1Q1/n1) and

P2~N (P2, P2Q2/n2)

Here, α = 0.05

**Hypotheses**

H0: The proportion of females using Instagram is not significantly different as compared to the proportion of males using Instagram. i.e., P1=P2.

H1: The proportion of females using Instagram is significantly different as compared to the proportion of males using Instagram. i.e., P1≠P2

**Test statistic**

Since n is large, the test statistic is given by:

Under the null hypothesis,

Here, p1= sample proportion of the number of males using Instagram = (X1/n1) =0.2906

P2= sample proportion of the number of females using Instagram = (X2/n2) =0.5501

n1 = n2 = 289

 Q = 1- P

**Computation:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gender** | **No. of users** | **Total Users** | **Prop- Value (P)** | **z-stat** | **p-value** |
| Males | 84 | 289 | 0.4204 | -3.8976 | 0 |
| Females | 159 | 289 |  |  |  |

**Conclusion:**

We observe that the p-value is less than the α value. So, we reject the null hypothesis and conclude that the proportion of females using Instagram is significantly different as compared to the proportion of males using Instagram.

For any value of α, we accept the null hypothesis if p-value > α.

                               we reject the null hypothesis if p-value <= α.

***Data Analysis using “TEST FOR INDEPENDENCE OF ATTRIBUTE”***

To analyse whether following an influencer is independent of time spent on social media

**Assumptions**

Attribute 1- Following an influencer

Attribute 2- Time spent on social media

**Hypothesis**

H0: Following an influencer is independent of time spent on social media

H1: Following an influencer is not independent of time spent on social media

**Test Statistic**

Under null Hypothesis, the test statistics is given by:  
 = ~   
 where, m=2, n=4

**Decision Rule**

For = 0.05, 3  
sign = 7.815  
If < sign, we accept the null hypothesis  
If > sign, we reject the null hypothesis

**Computation**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Row Labels** | **1 to 2 Hours** | **2 to 3 Hours** | **Less than 1 Hour** | **More than 3 Hours** | **Total** |
| No | 19 | 18 | 7 | 20 | 64 |
| Yes | 56 | 62 | 14 | 93 | 225 |
| Grand Total | 75 | 80 | 21 | 113 | 289 |

= 3.2682, sign = 7.815

**Conclusion**

At = 0.05, 3 we observe that < sign thus, we accept the null hypothesis, H0 and conclude that following an influencer is independent of time spent on social media.

***Data Analysis using “TEST FOR INDEPENDENCE OF ATTRIBUTE”***

To analyse the relationship between choosing to be a certain kind of influencer and usage of YouTube

**Assumptions**

Attribute 1- choosing to be a certain kind of influencer

Attribute 2- usage of YouTube

**Hypothesis**

H0: Choosing to be a certain kind of influencer is independent of usage of YouTube

H1: Choosing to be a certain kind of influencer is not independent of usage of YouTube

**Test Statistic**

Under the null hypothesis, the test statistics is given by

= ~

where, m=2, n=3

**Decision Rule**

For = 0.05, 2  
sign = 5.991  
If < sign, we accept null hypothesis  
If > sign, we reject null hypothesis

**Computation**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Observed Frequency (o)** | **Fashion and Lifestyle** | **Food blogger/vlogger** | **Science and Technology** | **Grand Total** |
| No | 41 | 32 | 20 | 93 |
| Yes | 31 | 47 | 32 | 110 |
| Grand Total | 72 | 79 | 52 | 203 |

= 5.6220, sign = 5.991

**Conclusion**

At = 0.05, 2 we observe that < sign thus, we accept the null hypothesis, H0 and conclude that choosing to be a certain kind of influencer is independent of usage of YouTube

***Data Analysis using “TEST FOR INDEPENDENCE OF ATTRIBUTE”***

To analyse the change in taste and preferences due to the usage of social media and content creators

**Assumptions**

Attribute 1- Change in taste and preferences

Attribute 2- following content creators

**Hypothesis**

H0: Change in taste and preferences due to usage of social media is independent of following content creators

H1: Change in taste and preferences due to usage of social media is dependent on following content creators

**Test Statistic**

= ~

where, m= 2, n= 2

**Decision Rule**

For = 0.05, 1  
sign = 3.841  
If < sign, we accept null hypothesis  
If > sign, we reject null hypothesis

**Computation**

|  |  |  |  |
| --- | --- | --- | --- |
| **Row Labels** | **No** | **Yes** | **Grand Total** |
| No | 28 | 53 | 81 |
| Yes | 31 | 170 | 201 |
| Grand Total | 59 | 223 | 282 |

= 12.7904, sign = 3.841

**Conclusion**

At = 0.05, 2 we observe that > sign thus, we reject the null hypothesis, H0 and conclude that change in taste and preferences due to usage of social media is dependent on following content creators

**LIMITATIONS**

* 1. The data collected in this study has a particular geographical area. This survey cannot be used to analyse the country-wise situation.
  2. Ideology and usage of people may vary by time, which can result in changing the conclusion of the by time.
  3. Respondents may give socially acceptable and convenient responses, which may skew the results.
  4. Respondents may be biased because of convenience sampling.
  5. Statistical measures like Anova, T-test are not used because of the large sample size, qualitative data.

**FUTURE USE**

1. Our survey can help students as a secondary source of data, further for their work or projects.

2. Our project can be used by telecommunication companies to understand which app youth are mostly in or spending much time on.

3. Our project can be used to enlighten people in society about the effect of social media on youth.

4. Our project can be used by telecommunication companies to improvise social media apps.

5. Our project can be used by social media influencers to know which field they can get into more and where to improvise the content.

6. Our project can be also used by the education department to know which age group and gender is getting most affected.

**.BIBLIOGRAPHY**

[*https://www.springwise.com/pros-cons-social-media-influencers*](https://www.springwise.com/pros-cons-social-media-influencers)

[*https://onlinedegrees.unr.edu/online-master-of-public-health/impact-of-social-media-on-youth-mental-health/*](https://onlinedegrees.unr.edu/online-master-of-public-health/impact-of-social-media-on-youth-mental-health/)

[*https://simple.wikipedia.org/wiki/Main\_Page*](https://simple.wikipedia.org/wiki/Main_Page)

[*https://www.statista.com/*](https://www.statista.com/)

[*https://www.groupdiscussionideas.com/social-media-influencers-impact-on-youth/#:~:text=Impact%20of%20social%20media%20influencers%20on%20youth%3A,-People%20trust%20social&text=Followers%20and%20especially%20youngsters%20feel,influencers%20interact%20with%20them%20frequently.&text=Some%20influencers%20do%20not%20mention,money%20and%20mislead%20the%20youngsters*](https://www.groupdiscussionideas.com/social-media-influencers-impact-on-youth/#:~:text=Impact%20of%20social%20media%20influencers%20on%20youth%3A,-People%20trust%20social&text=Followers%20and%20especially%20youngsters%20feel,influencers%20interact%20with%20them%20frequently.&text=Some%20influencers%20do%20not%20mention,money%20and%2)*.*

[*https://www.newportacademy.com/resources/well-being/effect-of-social-media-on-teenagers/*](https://www.newportacademy.com/resources/well-being/effect-of-social-media-on-teenagers/)