Mobile Phone Usage and Digital Wellbeing

Himanshu Jindal
MA21BTECH11007@iith.ac.in

Shubham Vishwakarma
MA21BTECH11020@iith.ac.in

Kaustubh Dandegaonkar
MA21BTECH11003@iith.ac.in

Roshan Kumar
MA21BTECH11013@iith.ac.in

Naman Chhibbar
MA21BTECH11011@iith.ac.in

Deepinder Singh
MA21BTECH11004@iith.ac.in

Devashish Chaudhari
MA21BTECH11005@iith.ac.in

Karthik Dhanavath
MA21BTECH11006@iith.ac.in

Abstract

Smartphones have become an indispensible part of our everyday lives, yet there is growing concern about the possible detrimental effects of excessive screen time. We conducted a survey as students to investigate the consequences of smartphone use on physical health, mental health, academic performance, and social behaviour. Our study concentrated on smartphone usage, which is currently the most popular gadget. We gathered information from students and analysed it statistically. Our findings suggest that excessive smartphone use has a negative impact on students' health, academic performance, and social connections. Students who spent more time on their smartphones had poorer sleep quality, higher stress and anxiety levels, poorer academic achievement, and less social engagement.

1 Introduction

The impact of smartphone screen time on students' physical health, mental health, academic performance, and social behaviour is investigated in our study. We performed a survey to gather information about students' smartphone screen time habits and perceived effects. The collected data is statistically analysed to determine the association between smartphone screen time and its impacts. Our findings will provide vital insights into the possible hazards of excessive smartphone screen time and the significance of keeping a good lifestyle balance with other pursuits.

The survey consisted of the following questions:

1. What is your average screen time in a day?
2. How many times do you check your phone in a day?
3. How many notifications do you receive in a day?
4. What percentage of your screen time is productive?
5. What is the average time you study daily (outside college hours)?
6. How much do you usually study in one sitting? (Hours)
7. Which hostel are you staying in?
8. Which degree are you pursuing?
9. Which year are you currently in?
10. Gender of the student.
11. Do you wear spectacles?
12. Do you use the phone in class?
13. Do you attend classes?
14. How do you rate your focus?
15. How do you rate your happiness or mental well-being?

Based on the responses obtained to the questions mentioned above, we pre-processed the data and the data visualisation is as follows:

2 Pre-processing and Visualisation of Data

The following steps were taken to pre-process the data:

1. We began by removing white-spaces and removing columns not required for further analysis.
2. Any existing `NaNs` were replaced with the modal values of the specific column, since we do not have any model to predict the unentered values.
3. Finally, to make the data more interpretable, we assumed that any person studying between n and $n+1$ hours would be studying $n+\frac{1}{2}$ hours on average.

After pre-processing, the collected data was visualized as simple bar plots:

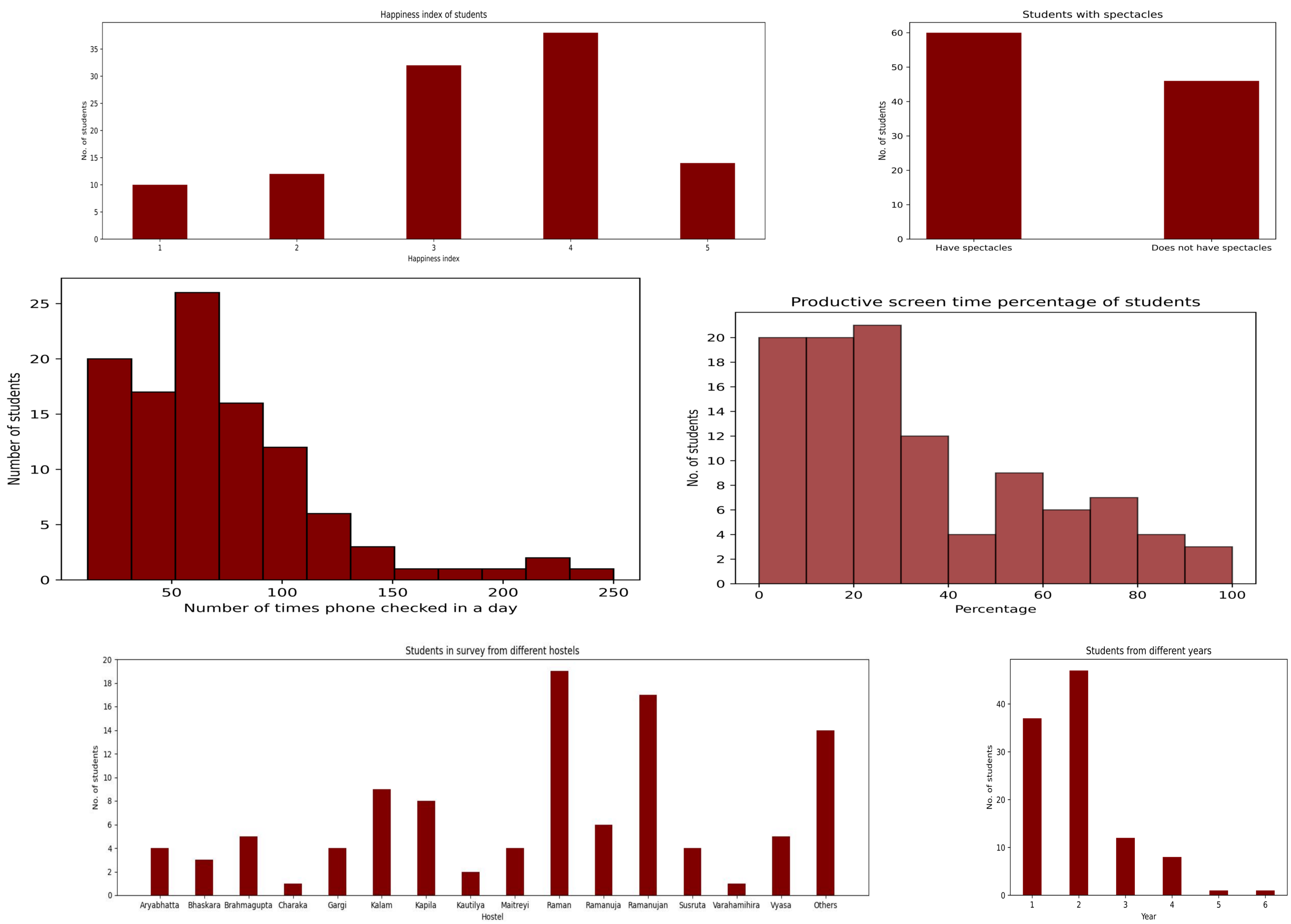

Figure 1. Visualizing the collected data

3 Analysis and Conclusions

After pre-processing the data, we had the data of 106 students from the survey, and we began to analyze the data. This involved the analysis of the one question involving a numerical variable:
"What is your average screen time in a day?"
The following results were obtained:

- Mean $= 5.96$ hours
- Median $= 5$ hours
- Mode $= 5$ hours

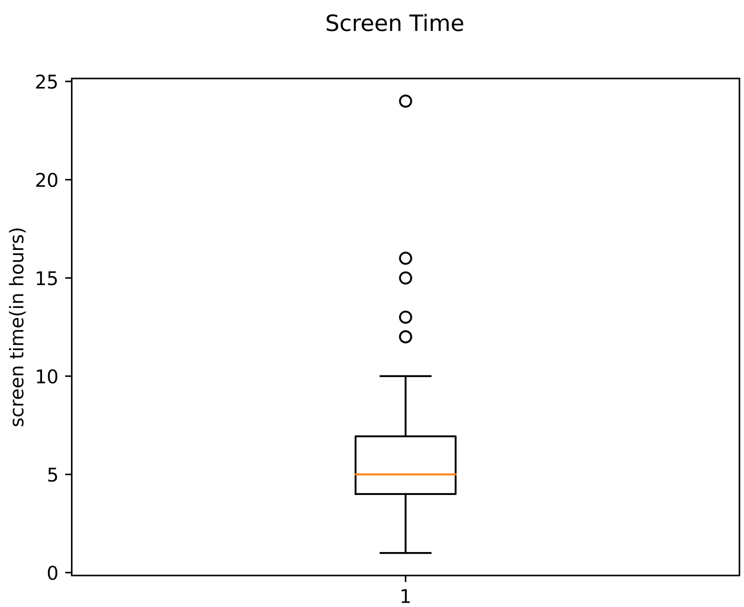

Figure 2. Screen Time of Students (Box Plot)

From the Box Plot: Figure 2, as well as the statistics, since the mean, median and mode are all approximately equal, we can say the data approximately follows a normal distribution.

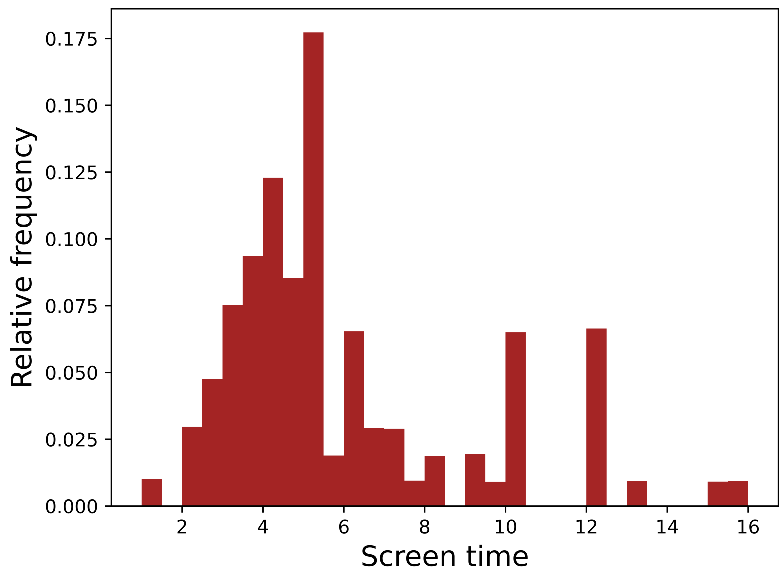

Figure 3. Screen Time of Students (distribution function) (Stem Plot)

After getting an idea of how the data stacks up, answers to some intriguing questions which involved the comparison of two questions were also given using segmented bar charts:

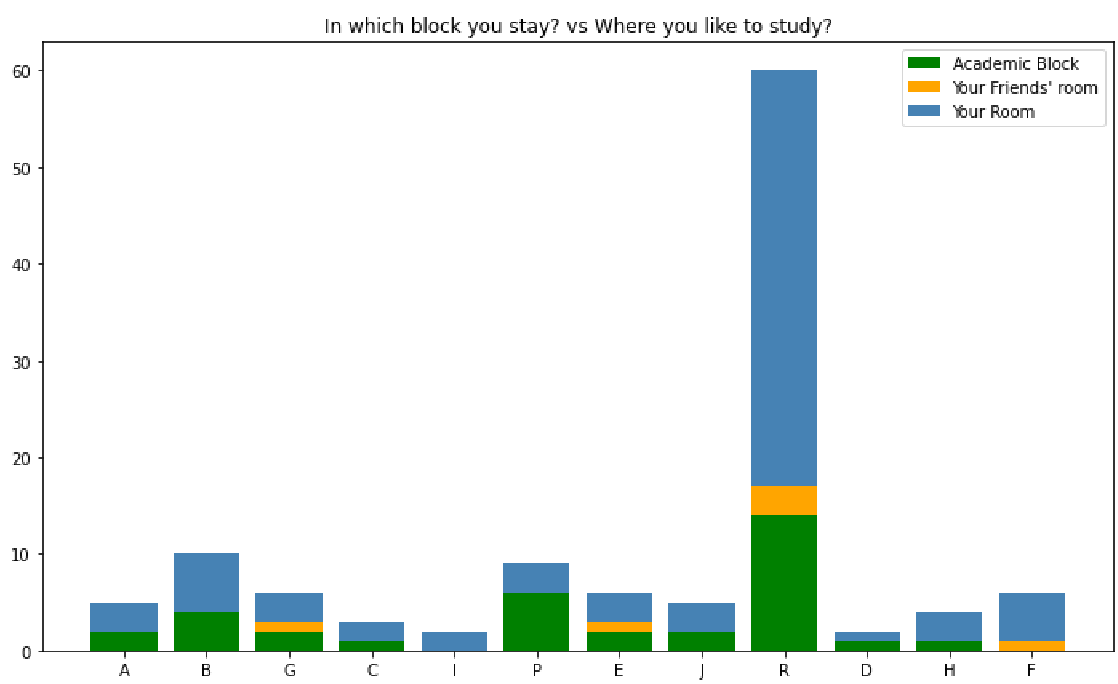

Figure 4. Residing Block vs Study room

Based on the Segmented Bar Plot: Figure 4, the following conclusions can be made:

1. A majority of students from the R block prefer to study in their rooms alone, whereas a majority of students from the P block (mainly involving Masters and Research students) prefer to go the academic blocks.
2. Very few prefer to study in their friends' room. More or less, from every block, people want to study in their room itself.

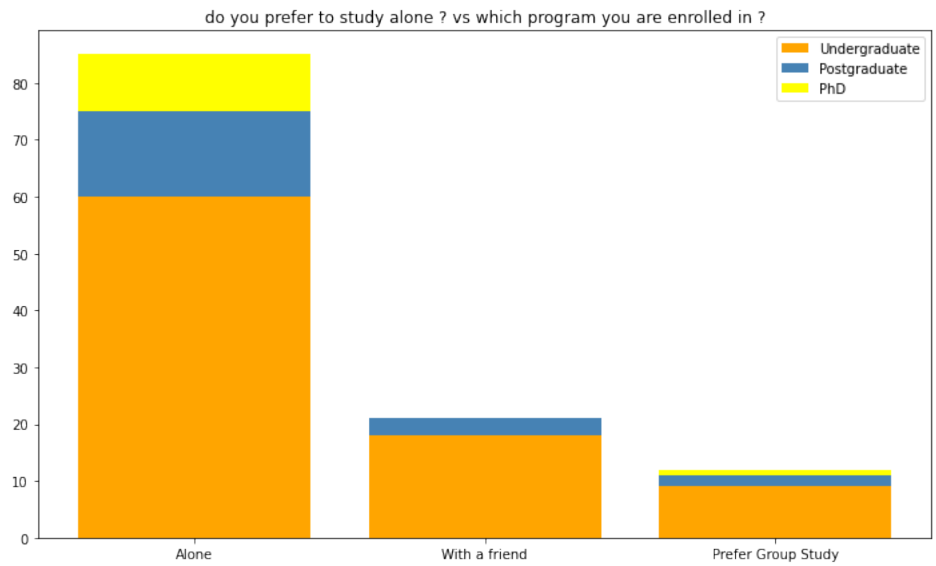

Figure 5. Preference towards studying alone vs Program

Based on the Segmented Bar Plot: Figure 5, the following conclusion can be made: Undergraduate students are open to studying alone, with a friend or even involving themselves in group studies. However, as we progress to Postgraduate and PhD students, they prefer studying alone in general.

Based on the Segmented Bar Plot: Figure 6, the following conclusion can be made: More or less, equal number of people prefer lecture recordings and/or slides. However, most people who prefer recordings study for the minimum time at one go, and as this time tends to increase, people prefer to use slides to study.

Based on the Segmented Bar Plot: Figure 7, the following conclusion can be made: A larger proportion of people studying for longer hours prefer having snacks with them as compared to people studying just for $1 - 2$ hours

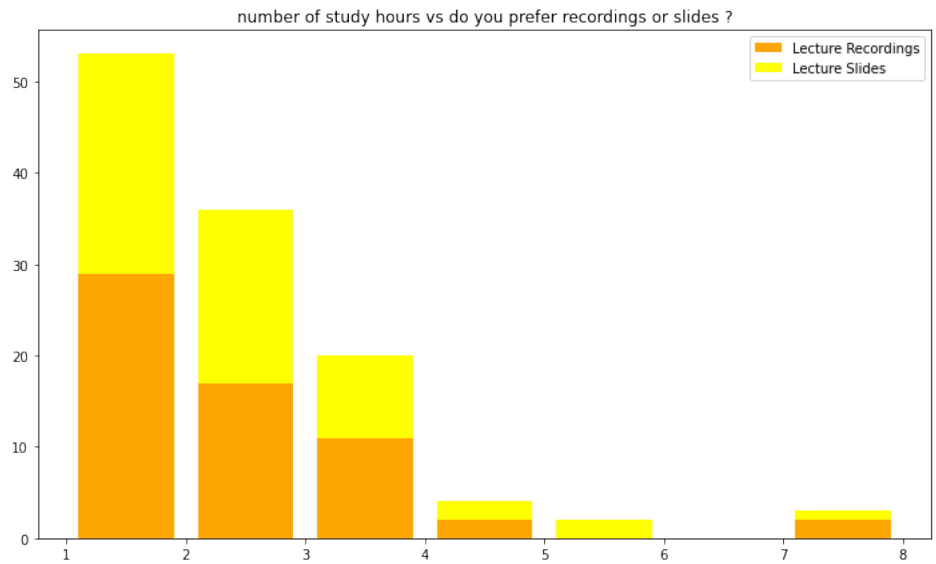

Figure 6. Hours studied vs Slides or Recordings

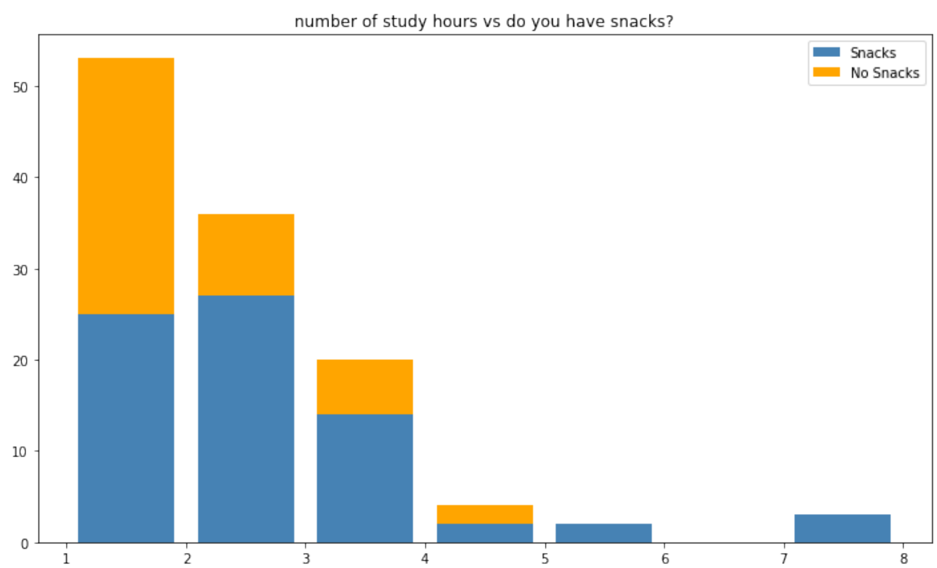

Figure 7. Study Hours vs Snacks Consumptions

4 Confidence Interval:

4.1 Confidence Interval for Mean Screen Time :

Let $\bar{x}$ = Sample Mean of target variable
Let S^2 = Sample Variance of target variable

let n= size of sample The Confidence interval is given by:

$$\bar{x} \pm E \tag{1}$$

also given by,

$$[\bar{x} - E, \bar{x} + E] \tag{2}$$

Where, E(Margin of Error) is given by,

$$E = t_{\alpha,n-1}(\frac{S}{\sqrt{n}}) \tag{3}$$

Width of the Confidence Interval is given by, $W = 2E$

4.1.1 Confidence interval for screen time of all students :

Let $\bar{x}$ = Sample Mean of screentime of students
Let S^2 = Sample Variance of screentime of students

Based on the sample selected, we have the following information-

$$\bar{x} = 5.96 \text{ hours} \tag{4}$$

$$S = 3.52 \tag{5}$$

$$n = 106 \tag{6}$$

We find the confidence interval of Screen time of students .
Confidence testing for the mean of the data gives us the following:

- The 95% confidence interval is given by: (5.28,6.63).
  The width of the interval is given by: 1.35
- The 99% confidence interval is given by: (5.06,6.85).
  The width of the interval is given by: 1.79

Increasing the confidence level, the width of the interval also increases.

4.1.2 Confidence interval for screen time of all male students :

Let $\bar{x}$ = Sample Mean of screentime of male students
Let S^2 = Sample Variance of screentime of male students

Based on the sample selected, we have the following information-

$$\bar{x} = 5.98 \text{ hours} \tag{7}$$

$$S = 3.6 \tag{8}$$

$$n = 86 \tag{9}$$

$$\tag{10}$$

We find the confidence interval of Screen time of students .
Confidence testing for the mean of the data gives us the following:

- The 95% confidence interval is given by: (5.21,6.79).
  The width of the interval is given by: 1.55
- The 99% confidence interval is given by: (4.95,7.01).
  The width of the interval is given by: 2.06

Increasing the confidence level, the width of the interval also increases.

4.1.3 Confidence interval for screen time of all female students :

Let $\bar{x}$ = Sample Mean of screentime of female students
Let S^2 = Sample Variance of screentime of female students

Based on the sample selected, we have the following information-

$$\bar{x} = 5.88 \text{ hours} \tag{11}$$
$$S = 3.21 \tag{12}$$
$$n = 20 \tag{13}$$
$$\tag{14}$$

We find the confidence interval of Screeen time of students .
Confidence testing for the mean of the data gives us the following:

- The 95% confidence interval is given by: (4.33,7.43).
  The width of the interval is given by: 3.1
- The 99% confidence interval is given by: (3.76,7.997).
  The width of the interval is given by: 4.237

Increasing the confidence level, the width of the interval also increases.

4.1.4 Confidence interval for screen time of all students who dont wear specs :

Let $\bar{x}$ = Sample Mean of screentime of students who dont wear specs
Let S^2 = Sample Variance of screentime of students who dont wear specs

Based on the sample selected, we have the following information-

$$\bar{x} = 5.03 \text{ hours} \tag{15}$$
$$S = 2.54 \tag{16}$$
$$n = 46 \tag{17}$$
$$\tag{18}$$

We find the confidence interval of Screeen time of students .
Confidence testing for the mean of the data gives us the following:

- The 95% confidence interval is given by: (4.27,5.78).
  The width of the interval is given by: 1.51
- The 99% confidence interval is given by: (4.02,6.03).
  The width of the interval is given by: 2.01

Increasing the confidence level, the width of the interval also increases.

4.1.5 Confidence interval for screen time of all students who wear specs :

Let $\bar{x}$ = Sample Mean of screentime of students who wear specs
Let S^2 = Sample Variance of screentime of students who wear specs

Based on the sample selected, we have the following information-

$$\bar{x} = 6.68 \text{ hours} \tag{19}$$
$$S = 3.99 \tag{20}$$
$$n = 60 \tag{21}$$

We find the confidence interval of Screen time of students .
Confidence testing for the mean of the data gives us the following:

- The 95% confidence interval is given by: (5.65,7.71).
  The width of the interval is given by: 2.06
- The 99% confidence interval is given by: (5.31,8.05).
  The width of the interval is given by: 2.74

Increasing the confidence level, the width of the interval also increases.

4.2 Confidence Interval for Variance :

Let $\bar{x}$ be the sample mean of the target variable, S^2 be the sample variance of the target variable, and n be the sample size.

If $X_1, X_2, ..., X_n$ are normally distributed and $a = \chi^2_{(1-\frac{\alpha}{2},n-1)}$ and $b = \chi^2_{(\frac{\alpha}{2},n-1)}$, then a $(1-\alpha)$

$$\left(\frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a}\right) \tag{22}$$

Similarly, a $(1-\alpha)$

$$\left(\frac{\sqrt{(n-1)}S}{\sqrt{b}}, \frac{\sqrt{(n-1)}S}{\sqrt{a}}\right) \tag{23}$$

Note that the confidence interval formulas above are only applicable if the sample data is normally distributed.

4.2.1 Confidence interval for Variance of screen time of all students :

Let S^2 = Sample Variance of screentime of students
Based on the sample selected, we have the following information-

$$S = 3.52 \tag{24}$$
$$n = 106 \tag{25}$$
$$a = 78.54 \tag{26}$$
$$b = 135.25 \tag{27}$$

We find the confidence interval of Variance of Screen time of students .
Confidence testing for the Variance of the data gives us the following:

- The 95% confidence interval is given by: (9.62,16.57).
  The width of the interval is given by: 6.95

4.2.2 Confidence interval for Variance of screen time of all male students :

Let S^2 = Sample Variance of screentime of male students
Based on the sample selected, we have the following information-

$$S = 3.6 \tag{28}$$
$$n = 86 \tag{29}$$
$$a = 61.39 \tag{30}$$
$$b = 112.39 \tag{31}$$

We find the confidence interval of Variance of Screen time of students .
Confidence testing for the Variance of the data gives us the following:

- The 95% confidence interval is given by: (9.80,17.94).
  The width of the interval is given by: 8.14

4.2.3 Confidence interval for Variance of screen time of all female students :

Let S^2 = Sample Variance of screentime of female students
Based on the sample selected, we have the following information-

$$S = 3.21 \tag{32}$$
$$n = 20 \tag{33}$$
$$a = 8.91 \tag{34}$$
$$b = 32.85 \tag{35}$$

We find the confidence interval of Variance of Screen time of students .
Confidence testing for the Variance of the data gives us the following:

- The 95% confidence interval is given by: (5.96,21.98).
  The width of the interval is given by: 16.02

4.2.4 Confidence interval for Variance of screen time of all students who dont wear specs :

Let S^2 = Sample Variance of screentime of students who dont wear specs
Based on the sample selected, we have the following information-

$$S = 2.54 \tag{36}$$
$$n = 46 \tag{37}$$
$$a = 28.37 \tag{38}$$
$$b = 65.41 \tag{39}$$

We find the confidence interval of Variance of Screen time of students .
Confidence testing for the Variance of the data gives us the following:

- The 95% confidence interval is given by: (4.44,10.23).
  The width of the interval is given by: 5.80

4.2.5 Confidence interval for Variance of screen time of all students who wear specs :

Let S^2 = Sample Variance of screentime of students who wear specs
Based on the sample selected, we have the following information-

$$S = 3.99 \tag{40}$$
$$n = 60 \tag{41}$$
$$a = 39.66 \tag{42}$$
$$b = 82.12 \tag{43}$$

We find the confidence interval of Variance of Screen time of students who wear specs .
Confidence testing for the Variance of the data gives us the following:

- The 95% confidence interval is given by: (11.44,23.68).
  The width of the interval is given by: 12.24

4.3 Confidence Interval for Proportions :

Let $\hat{p}$ = Sample Proportion
Let n= size of sample size

For large Random samples, a 100% CI for the Population proportion p is given by :

$$\left(\hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) \tag{44}$$

4.3.1 Confidence interval for Proportion of students who use phone between 4 to 6 hours :

$$\hat{p} = 0.47 \tag{45}$$

$$n = 106 \tag{46}$$

We find the confidence interval of Proportion of students who use phone between 4 to 6 hours.
Confidence testing for the Proportion of the data gives us the following:

- The 95% confidence interval is given by: (37.67,56.67).
  The width of the interval is given by: 20.00

5 Hypothesis Testing:

5.1 Case-1: Comparing the screen-time for Undergraduates and Postgraduates

We compare screen-time of Undergraduates and Postgraduates. We assume our null hypothesis to be that undergraduates have more screen-time than postgraduates. Let $\alpha = 0.05$

Let $\bar{x_1}$ = Sample Mean of screen-time of Undergraduates
Let $\bar{x_2}$ = Sample Mean of screen-time of Postgraduates
Let S_1^2 = Sample Variance of screen-time of Undergraduates
Let S_2^2 = Sample Variance of screen-time of Postgraduates

For Hypothesis Testing we make the following statements-

$$H_0 : \mu_1 - \mu_2 \geq 0 \text{ and } H_a : \mu_1 - \mu_2 < 0$$

We will now select a random sample of 18 entries for the Undergraduates and 18 entries for the Postgraduates from our data. Based on the sample selected, we have the following information-

$$\bar{x_1} = 5.479 \text{ hours} \tag{47}$$
$$\bar{x_2} = 7.038 \text{ hours} \tag{48}$$
$$S_1^2 = 7.388 \tag{49}$$
$$S_2^2 = 22.279 \tag{50}$$
$$n_1 = 73 \tag{51}$$
$$n_2 = 33 \tag{52}$$
$$\tag{53}$$

-

Since $\dfrac{S_1^2}{S_2^2} = 0.3316 > 0.25$, we can assume the population variances would be equal. Thus, we can say:

The degrees of freedom, $df = n_1 + n_2 - 2 = 104$

The pooled variance will be:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = 11.9716 \tag{54}$$

The test statistic is then given by:

$$t = \frac{\bar{x_1} - \bar{x_2} - 0}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} = -2.15 \tag{55}$$

Using the rejection region approach, we reject H_0 if $t \leq -t_{0.05,104}$, where $t_{0.05,104} = 1.659$.
Because the observed value of $t = -2.15$ is less than -1.659, we have enough statistical evidence to reject the null hypothesis, and thus, we can say, the postgraduates have more screentime in one go than the undergraduates on average.

5.2 Case-2: Comparing the screen time of people with specs and people with no specs.

We compare screen times of people who wear specs and people who do not . We assume our null hypothesis that people who wear specs have less screen time compared to people who wear specs.

Let $\bar{x_1}$ = Sample Mean of screen-time of people with no specs
Let $\bar{x_2}$ = Sample Mean of screen-time of people with specs
Let S_1^2 = Sample Variance of screen-time of people with no specs
Let S_2^2 = Sample Variance of screen-time of people with specs

For Hypothesis Testing we make the following statements-

$$H_0 : \mu_1 - \mu_2 \geq 0 \text{ and } H_a : \mu_1 - \mu_2 < 0$$

$$\bar{x_1} = 5.027 \text{ hours} \tag{56}$$
$$\bar{x_2} = 6.683 \text{ hours} \tag{57}$$
$$S_1^2 = 6.453 \tag{58}$$
$$S_2^2 = 15.934 \tag{59}$$
$$n_1 = 60 \tag{60}$$
$$n_2 = 46 \tag{61}$$

Since $\dfrac{S_1^2}{S_2^2} = 0.40 > 0.25$, we can assume the population variances would be equal. Thus, we can say:

The degrees of freedom, $df = n_1 + n_2 - 2 = 104$

The pooled variance will be:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = 10.5556 \tag{62}$$

The test statistic is then given by:

$$t = \frac{\bar{x_1} - \bar{x_2} - 0}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} = -2.601 \tag{63}$$

Using the rejection region approach, we reject H_0 if $t \leq -t_{0.05,116}$, where $t_{0.05,116} = 1.659$.
Because the observed value of $t = -2.601$ is less than -1.659, we have enough statistical evidence to reject the null hypothesis, and thus, we can say, those with specs have more screen time on average than those with no specs

5.3 Case-3: Comparing the productive screen time of Undergraduate students to that of Postgraduate students.

We compare productive screen time of Undergraduate students and Postgraduate students ! We assume our null hypothesis that Undergraduate students have more productive screen-time compared to Postgraduate students .

Let $\bar{x_1}$ = Sample Mean of screen time of UG students
Let $\bar{x_2}$ =Sample Mean of screen time of PG students
Let S_1^2 = Sample Variance of screen time of UG students
Let S_2^2 = Sample Variance of screen time of PG students

For Hypothesis Testing we make the following statements-

$$H_0 : \mu_2 - \mu_1 \leq 0 \text{ and } H_a : \mu_2 - \mu_1 > 0$$

$$\bar{x_1} = 27.89 \text{ hours} \tag{64}$$
$$\bar{x_2} = 35.80 \text{ hours} \tag{65}$$
$$S_1^2 = 522.002 \tag{66}$$
$$S_2^2 = 887.061 \tag{67}$$
$$n_1 = 73 \tag{68}$$
$$n_2 = 33 \tag{69}$$

Since $\frac{S_1^2}{S_2^2} = 0.588 > 0.25$, we can assume the population variances would be equal. Thus, we can say:

The degrees of freedom, $df = n_1 + n_2 - 2 = 104$

The pooled variance will be:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = 634.32 \tag{70}$$

The test statistic is then given by:

$$t = \frac{\bar{x_2} - \bar{x_1} - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 1.513 \tag{71}$$

Using the rejection region approach, we reject H_0 if $t \geq t_{0.05,104}$, where $t_{0.05,104} = 1.658$.

Because the observed value of $t = 1.513$ is less than 1.659, we fail to reject the null hypothesis, and thus, do not have enough evidence to say Undergraduate students have more productive screen-time compared to Postgraduate students .

5.4 Case-4: Comparing the screen time of people who attend all the classes and those who don't attend all the classes.

We compare screen time of people who attend all the classes and those who don't attend all the classes! We assume our null hypothesis that people who attend all the classes have higher screen time than those who don't.

Let $\bar{x_1}$ = Sample Mean of screen time of people who attend all the classes
Let $\bar{x_2}$ = Sample Mean of screen time of people who don't attend all the classes
Let S_1^2 = Sample Variance of screen time of people who attend all the classes
Let S_2^2 = Sample Variance of screen time of people who don't attend all the classes

For Hypothesis Testing we make the following statements-

$$H_0 : \mu_1 - \mu_2 \geq 0 \text{ and } H_a : \mu_1 - \mu_2 < 0$$

$$\bar{x_1} = 5.33 \text{ hours} \tag{72}$$
$$\bar{x_2} = 6.67 \text{ hours} \tag{73}$$
$$S_1^2 = 2.90 \tag{74}$$
$$S_2^2 = 3.95 \tag{75}$$
$$n_1 = 56 \tag{76}$$
$$n_2 = 50 \tag{77}$$

Since $\frac{S_1^2}{S_2^2} = 0.734 > 0.25$, we can assume the population variances would be equal. Thus, we can say:

The degrees of freedom, $df = n_1 + n_2 - 2 = 104$

The pooled variance will be:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = 11.83 \tag{78}$$

The test statistic is then given by:

$$t = \frac{\bar{x_1} - \bar{x_2} - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = -1.995 \tag{79}$$

Using the rejection region approach, we reject H_0 if $t \leq -t_{0.05,104}$, where $t_{0.05,104} = 1.658$.
Because the observed value of $t = -1.995$ is lesser than -1.658, we have enough statistical evidence to reject the null hypothesis, and thus, we can say, those who attend all the classes have lower screen time than those who don't attend all the classes.

5.5 Case-5: Comparing the screen time of people who are sad and those who are happy.

We compare screen time of people who are higher on the happiness metric and those who are lower on the happiness metric! We assume our null hypothesis that people who are higher on happiness metric (≥ 3) have higher screen time than those who are lower on happiness metric (≤ 2).

Let $\bar{x_1}$ = Sample Mean of screen time of people who have happiness (≥ 3)
Let $\bar{x_2}$ = Sample Mean of screen time of people who have happiness (≤ 2)
Let S_1^2 = Sample Variance of screen time of people who have happiness (≥ 3)
Let S_2^2 = Sample Mean of screen time of people who have happiness (≤ 2)

For Hypothesis Testing we make the following statements-

$$H_0 : \mu_1 - \mu_2 \geq 0 \text{ and } H_a : \mu_1 - \mu_2 < 0$$

$$\bar{x_1} = 5.457 \text{ hours} \tag{80}$$
$$\bar{x_2} = 7.897 \text{ hours} \tag{81}$$
$$S_1^2 = 3.2928 \tag{82}$$
$$S_2^2 = 3.6139 \tag{83}$$
$$n_1 = 84 \tag{84}$$
$$n_2 = 22 \tag{85}$$

Since $\frac{S_1^2}{S_2^2} = 0.911 > 0.25$, we can assume the population variances would be equal. Thus, we can say:

The degrees of freedom, $df = n_1 + n_2 - 2 = 104$

The pooled variance will be:

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = 3.357 \tag{86}$$

The test statistic is then given by:

$$t = \frac{\bar{x_1} - \bar{x_2} - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = -3.03 \tag{87}$$

Using the rejection region approach, we reject H_0 if $t \leq -t_{0.05,104}$, where $t_{0.05,104} = 1.658$.
Because the observed value of $t = -3.03$ is lesser than -1.658, we have enough statistical evidence to reject the null hypothesis, and thus, we can say, those who are higher on the happiness metric (≥ 3) have lower screen time compared to people at lower happiness (≤ 2).

5.6 Case-6: Comparing the screen time of students with phone checking habits vs other students

We compare screen time of people with frequent phone checking habits (Phone-check/day $\geq$ 150) with others (Phone-check/day $<$ 150)! We assume our null hypothesis that people who check phone ≥ 150 times have lower screen time than those who check phone < 150 times.

Let $\bar{x_1}$ = Sample Mean of screen time of people who check phone (≥ 150) times
Let $\bar{x_2}$ = Sample Mean of screen time of people who check phone (< 150) times
Let S_1^2 = Sample Variance of screen time of people who check phone (≥ 150) times
Let S_2^2 = Sample Mean of screen time of people who check phone (< 150) times

For Hypothesis Testing we make the following statements-

$$H_0 : \mu_1 - \mu_2 \leq 0 \text{ and } H_a : \mu_1 - \mu_2 > 0$$

$$\bar{x_1} = 5.457 \text{ hours} \tag{88}$$
$$\bar{x_2} = 7.897 \text{ hours} \tag{89}$$
$$S_1^2 = 3.2928 \tag{90}$$
$$S_2^2 = 3.6139 \tag{91}$$
$$n_1 = 84 \tag{92}$$
$$n_2 = 22 \tag{93}$$

Since $\frac{S_1^2}{S_2^2} = 0.911 > 0.25$, we can assume the population variances would be equal. Thus, we can say:

The degrees of freedom, $df = n_1 + n_2 - 2 = 104$

The pooled variance will be:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = 3.357 \tag{94}$$

The test statistic is then given by:

$$t = \frac{\bar{x_1} - \bar{x_2} - 0}{s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} = -3.03 \tag{95}$$

Using the rejection region approach, we reject H_0 if $t \leq -t_{0.05,104}$, where $t_{0.05,104} = 1.658$.
Because the observed value of $t = -3.03$ is lesser than -1.658, we have enough statistical evidence to reject the null hypothesis, and thus, we can say, those who are higher on the happiness metric (≥ 3) have lower screen time compared to people at lower happiness (≤ 2).

6 Contributions:

- **Devashish Chaudhari**
    *
- **Himanshu Jindal**
    *
- **Shubham Vishwakarma**
    *
- **Kaustubh Dandegaonkar**
    *
- **Roshan Kumar**
    *
- **Deepinder Singh**
    *
- **Naman Chhibbar**
    *
- **Karthik Dhanavath**