Summary

Mobile phones have become an important part of our daily lives. People use them for communication, entertainment, studying, and even working. Because of this, many people spend a lot of time looking at their phone screens every day. This is called "screen time".

With the rise of smartphones in daily routines, we aim to understand how much time people spend on their smartphones and how this may relate to different personal conditions, such as weight, heart rate, age, gender, eyesight, and sleep duration. In addition, we explore which days of the week people tend to have the highest and lowest screen time. We also look into which application types are used the most, such as social media, entertainment, or productivity apps. This information can help us understand the impact of smartphone usage on health and daily habits.

Following our analysis, we found that age has a weak negative correlation with screen time, meaning that older people tend to spend slightly less time on their smartphones. In contrast, weight and sleep duration showed no correlation with screen time. Furthermore, the data indicated that respondents spent the most time on their mobile phones on Sundays, while Mondays recorded the lowest usage.

Group 6

6710545873	Vorapop Prasertkul
6710545521	Chaiyapat Kumtho
6710545741	Pasin Tongtip
6710545989	Amornrit Sirikham

Data Collection

We collected the data with a survey using Google Forms. Our target is people in different age groups. The survey is mostly in Thai and was accepting the results from April 12, 2025, until April 17, 2025, and 57 people responded to the survey.

The survey questions are :

1. What is your age?
2. What is your gender?
3. What is your weight (kg.) ?
4. What is your heart rate while answering the survey?
5. What is your approximate average sleeping duration (Hr.min) ?
6. How much screen duration did you spend on Sunday (Hr.min) ?
7. How much screen duration did you spend on Monday (Hr.min) ?
8. How much screen duration did you spend on Tuesday (Hr.min) ?
9. How much screen duration did you spend on Wednesday (Hr.min) ?
10. How much screen duration did you spend on Thursday (Hr.min) ?
11. How much screen duration did you spend on Friday (Hr.min) ?
12. How much screen duration did you spend on Saturday (Hr.min) ?
13. What is your most used application type?
14. How much notification did you receive on Sunday?
15. How much notification did you receive on Monday?
16. How much notification did you receive on Tuesday?
17. How much notification did you receive on Wednesday?
18. How much notification did you receive on Thursday?
19. How much notification did you receive on Friday?
20. How much notification did you receive on Saturday?
21. What is the current condition of your left eye?
    - 21.1 How is your eyesight (normal, near-sighted, far-sighted or Myopic Presbyopia) ?
        - 21.1.1 If you are near-sighted, how much?
        - 21.1.2 If you are far-sighted, how much?
        - 21.1.3 If you have compound vision, how much?
    - 21.2 Other abnormalities?
22. What is the current condition of your right eye (normal, near-sighted, far-sighted or compound vision) ?
    - 22.1 How is your eyesight (normal, near-sighted, far-sighted or Myopic Presbyopia) ?
        - 22.1.1 If you are near-sighted, how much?
        - 22.1.2 If you are far-sighted, how much?
        - 22.1.3 If you have compound vision, how much?
    - 22.2 Other abnormalities?

After we collected the data, we cleaned up the data. We found that there was data that had the same answers in every question(both surveys have the same answer and the answers came from the example provided in the form) except gender and age, so we did not use their answers. Another problem is the respondents did not follow the instructions that we provided, so the answers might not be accurate.

We found that most of the respondents were Male with 52.2%, followed by female with 47.8%. The majority of the respondents are 18 & 19 years old (45.65%), followed by older ages.

If we grouped the data by type of application used the most per day, 58.7% of the respondents spent the most time on entertainment applications, followed by social media at 34.8%, education at 4.3%, and information & reading at 2.2%.

If we grouped the data by day of the week that the respondents spent their time on the screen the most, 23.9% of the respondents spent time on their screen on Sunday the most, followed by Thursday & Tuesday, Wednesday & Friday & Saturday and Monday with 15.3%, 13.0% and 6.5% of the respondents respectively.

If we grouped the data by day of the week that the respondents spent their time on the screen the least, 28.3% of the respondents spent time on their screen on Monday the least, followed by Thursday & Saturday, Wednesday & Sunday, Tuesday and Friday with 19.6%, 10.9%, 8.7% and 2.2% of the respondents respectively.

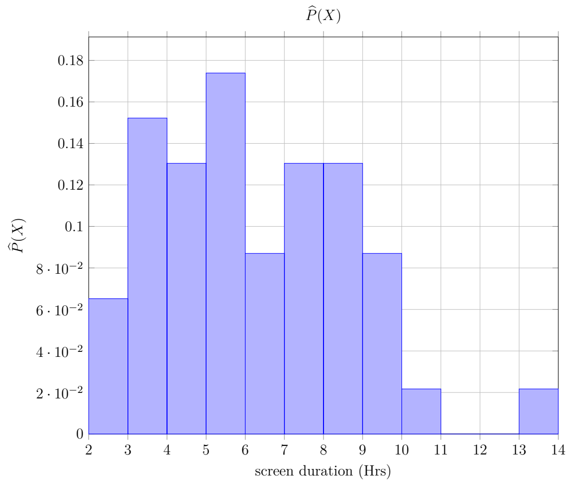
Methodology

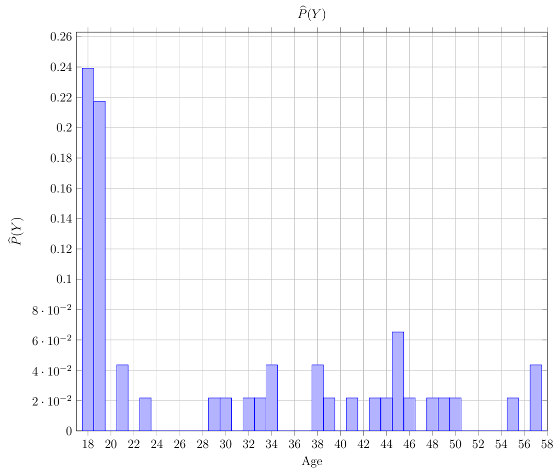
Probability

Screen time vs Age

y \ x	2	3	4	5	6	7	8	9	10	11	12	13	$\widehat{P}(Y)$	$y \cdot \widehat{P}(Y)$
	3	4	5	6	7	8	9	10	11	12	13	14		
18	0.0217	0.0652	0	0.0217	0	0.0217	0.0435	0.0435	0.0217	0	0	0	0.2391	4.3043
19	0	0	0.0217	0.0435	0.0435	0.0217	0.0217	0.0435	0	0	0	0.0217	0.2174	4.1304
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0.0217	0.0217	0	0	0	0	0	0.0435	0.913
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0.0217	0	0	0	0	0	0	0	0	0	0	0	0.0217	0.5
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0.0217	0	0	0	0	0	0	0	0	0	0	0.0217	0.6304
30	0	0	0	0	0.0217	0	0	0	0	0	0	0	0.0217	0.6522
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0.0217	0	0	0	0	0	0	0	0	0	0.0217	0.6957
33	0	0	0.0217	0	0	0	0	0	0	0	0	0	0.0217	0.7173913
34	0	0	0	0	0	0.0435	0	0	0	0	0	0	0.0435	1.4783
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0.0217	0.0217	0	0	0	0	0	0	0	0	0.0435	1.6522
39	0	0	0	0	0	0.0217	0	0	0	0	0	0	0.0217	0.8478
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0.0217	0	0	0	0	0	0	0	0	0.0217	0.8913
42	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	0	0.0217	0	0	0	0	0	0	0	0	0	0	0.0217	0.9348
44	0	0	0	0.0217	0	0	0	0	0	0	0	0	0.0217	0.9565217
45	0	0	0	0.0217	0	0	0.0435	0	0	0	0	0	0.0652	2.9348
46	0	0.0217	0	0	0	0	0	0	0	0	0	0	0.0217	1.
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0.0217	0	0	0	0	0	0	0	0	0	0	0	0.0217	1.0435
49	0	0	0.0217	0	0	0	0	0	0	0	0	0	0.0217	1.0652
50	0	0	0	0.0217	0	0	0	0	0	0	0	0	0.0217	1.087
51	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52	0	0	0	0	0	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0	0	0	0
54	0	0	0	0	0	0	0	0	0	0	0	0	0	0
55	0	0	0	0	0.0217	0	0	0	0	0	0	0	0.0217	1.19565217
56	0	0	0	0	0	0	0	0	0	0	0	0	0	0
57	0	0.0217	0.0217	0	0	0	0	0	0	0	0	0	0.0435	2.4783
$\widehat{P}(X)$	0.0652	0.1522	0.1304	0.1739	0.087	0.1304	0.1304	0.087	0.0217	0	0	0.0217	$\widehat{E}(X)$	6.23913
midpoint	2.5000	3.5000	4.5000	5.5000	6.5000	7.5000	8.5000	9.5000	10.5000	11.5000	12.5000	13.5000	$\widehat{E}(Y)$	30.10870
mid $\cdot \widehat{P}(X)$	0.163	0.5326	0.587	0.9565217	0.5652	0.9783	1.1087	0.8261	0.2283	0	0	0.2935		

The individual probability from the table comes from dividing frequency with the number of samples. $P[X]$ comes from summing the data for all columns. $P[Y]$ comes from summing the data for all rows. The expected value of X ($E[X]$) is 6.23913 which is calculated by summing all of $P[X]$ and the expected value of Y ($E[Y]$) is 30.10870 which is calculated by summing all of $P[Y]$.





The correlation coefficient is computed by using the formula

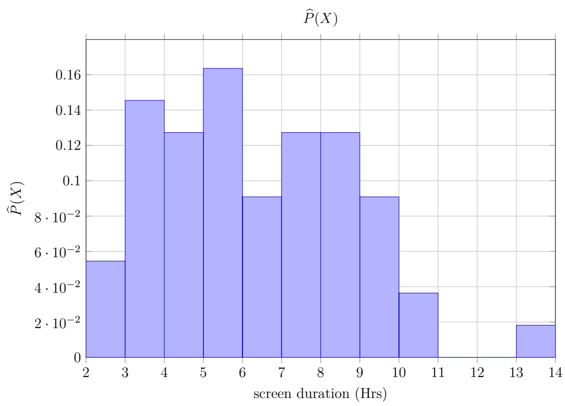
$$\widehat{\rho}_{X,Y} = \frac{\widehat{S}_{X,Y}}{s_X s_Y} \;\&\; \widehat{S}_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)\left(y_i - \overline{Y}\right)$$

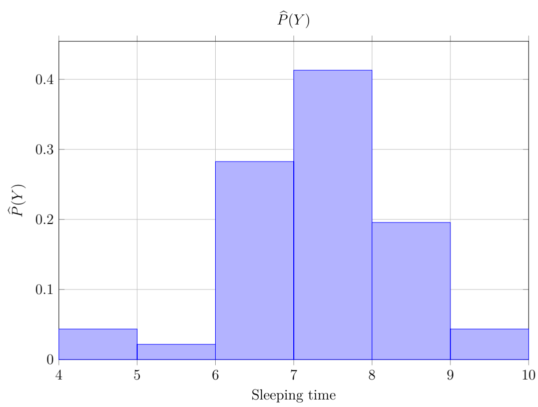
In this case, $\overline{X} = 6.271$, $\overline{Y} = 30.1087, n = 46$, $s_x = 2.4862$, $s_y = 13.2031$. Therefore, $\widehat{S}_{X,Y} = -10.9625$ and $\widehat{\rho}_{X,Y} = -0.3340$. Thus, the relation between X and Y has a weak negative relationship.

Screen time vs Sleeping time

y \ x		2	3	4	5	6	7	8	9	10	11	12	13	$\widehat{P}(Y)$	midpoint	mid $\cdot \widehat{P}(Y)$
		3	4	5	6	7	8	9	10	11	12	13	14			
4	5	0	0	0	0	0	0	0	0.0435	0	0	0	0	0.0435	4.5	0.1957
5	6	0	0	0	0.0217	0	0	0	0	0	0	0	0	0.0217	5.5	0.1195652
6	7	0	0.0217	0.0217	0.0435	0.0652	0.0435	0.0652	0	0.0217	0	0	0	0.2826	6.5	1.837
7	8	0.0217	0.0652	0.1087	0.087	0.0217	0.0652	0.0435	0	0	0	0	0	0.413	7.5	3.0978
8	9	0.0435	0.0435	0	0.0217	0	0.0217	0.0217	0.0435	0	0	0	0	0.1956522	8.5	1.6630435
9	10	0	0.0217	0	0	0	0	0	0	0	0	0	0.0217	0.0435	9.5	0.413
$\widehat{P}(X)$		0.0652	0.1522	0.1304	0.1739	0.087	0.1304	0.1304	0.087	0.0217	0	0	0.0217			
midpoint		2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5		$\widehat{E}(X)$	6.2391
mid $\cdot \widehat{P}(X)$		0.163	0.5326	0.587	0.9565217	0.5652	0.9783	1.1087	0.8261	0.2283	0	0	0.2935		$\widehat{E}(Y)$	7.3261

The individual probability from the table comes from dividing frequency with the number of samples. $P[X]$ comes from summing the data for all columns. $P[Y]$ comes from summing the data for all rows. The expected value of X ($E[X]$) is 6.23913 which is calculated by summing all of $P[X]$ and the expected value of Y ($E[Y]$) is 7.3261 which is calculated by summing all of $P[Y]$.





The correlation coefficient is computed by using the formula

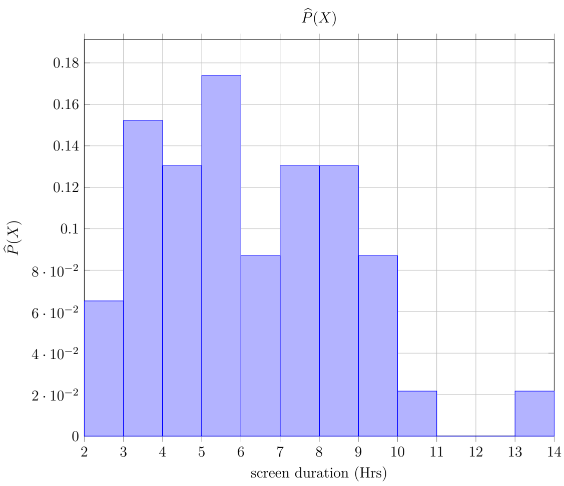
$$\widehat{\rho}_{X,Y} = \frac{\widehat{S}_{X,Y}}{s_X s_Y} \quad \& \quad \widehat{S}_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)\left(y_i - \overline{Y}\right)$$

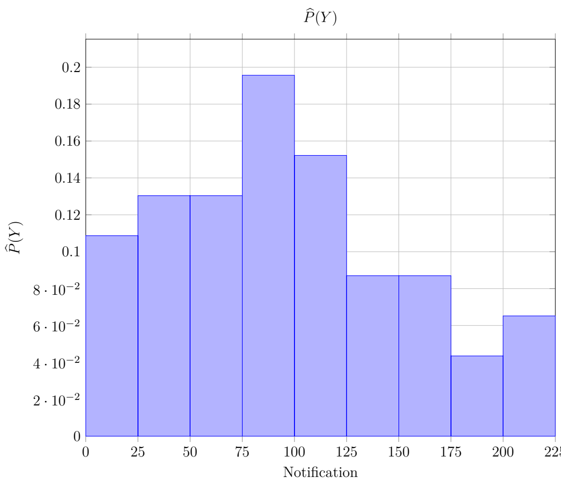
In this case, $\overline{X} = 6.271, \overline{Y} = 7.2178, n = 46, s_x = 2.4862, s_y = 1.0068$. Therefore, $\widehat{S}_{X,Y} = -0.3581$ and $\widehat{\rho}_{X,Y} = -0.1431$. Thus, the relation between X and Y has a weak negative relationship.

Screen time vs Notification

y \ x		2	3	4	5	6	7	8	9	10	11	12	13	$\widehat{P}(Y)$	midpoint	mid $\cdot \widehat{P}(Y)$
		3	4	5	6	7	8	9	10	11	12	13	14			
0	25	0.0217	0	0.0217	0.0652	0	0	0	0	0	0	0	0	0.1087	12.5	1.3587
25	50	0	0	0.0217	0	0.0217	0.0435	0.0217	0	0	0	0	0.0217	0.1304	37.5	4.8913
50	75	0.0217	0.0217	0	0.0435	0	0	0.0435	0	0	0	0	0	0.1304	62.5	8.1522
75	100	0.0217	0.0652	0.0217	0.0435	0.0435	0	0	0	0	0	0	0	0.1956522	87.5	17.1195652
100	125	0	0.0435	0.0217	0.0217	0.0217	0.0217	0	0.0217	0	0	0	0	0.1522	112.5	17.1196
125	150	0	0	0.0217	0	0	0	0.0217	0.0435	0	0	0	0	0.087	137.5	11.95652174
150	175	0	0	0	0	0	0.0652	0	0	0.0217	0	0	0	0.087	162.5	14.1304
175	200	0	0	0	0	0	0	0.0435	0	0	0	0	0	0.0435	187.5	8.1522
200	225	0	0.0217	0.0217	0	0	0	0	0.0217	0	0	0	0	0.0652	212.5	13.8587
$\widehat{P}(X)$		0.0652	0.1522	0.1304	0.1739	0.087	0.1304	0.1304	0.087	0.0217	0	0	0.0217			
midpoint		2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5		$\widehat{E}(Y)$	96.7391
mid $\cdot \widehat{P}(X)$		0.163	0.5326	0.587	0.9565	0.5652	0.9783	1.1087	0.8261	0.2283	0	0	0.2935		$\widehat{E}(X)$	6.2391

The individual probability from the table comes from dividing frequency with the number of samples. $P[X]$ comes from summing the data for all columns. $P[Y]$ comes from summing the data for all rows. The expected value of X ($E[X]$) is 6.23913 which is calculated by summing all of $P[X]$ and the expected value of Y ($E[Y]$) is 81.5217 which is calculated by summing all of $P[Y]$.





The correlation coefficient is computed by using the formula

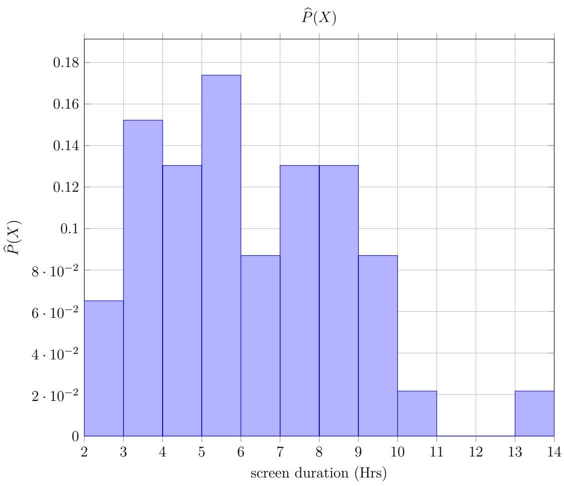
$$\widehat{\rho}_{X,Y} = \frac{\widehat{S}_{X,Y}}{s_X s_Y} \quad \& \quad \widehat{S}_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)\left(y_i - \overline{Y}\right)$$

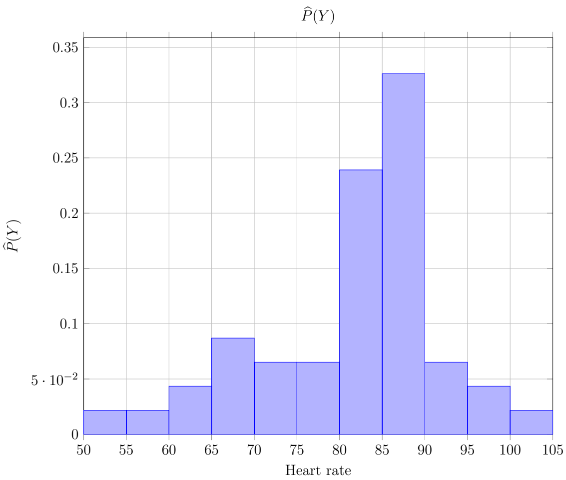
In this case, $\overline{X} = 6.271, \overline{Y} = 80.7174, n = 46, s_x = 2.4862, s_y = 10.8805$. Therefore, $\widehat{S}_{X,Y} = -1.3246$ and $\widehat{\rho}_{X,Y} = -0.0490$. Thus, the relation between X and Y has a weak negative relationship.

Screen time vs Heart rate

$y \backslash x$		2	3	4	5	6	7	8	9	10	11	12	13	$\widehat{P}(Y)$	midpoint	mid $\cdot\ \widehat{P}(Y)$
		3	4	5	6	7	8	9	10	11	12	13	14			
50	55	0	0	0	0	0	0	0.0217	0	0	0	0	0	0.0217	52.5	1.1413
55	60	0	0.0217	0	0	0	0	0	0	0	0	0	0	0.0217	57.5	1.25
60	65	0	0	0	0	0	0.0435	0	0	0	0	0	0	0.0435	62.5	2.7174
65	70	0	0	0	0	0.0217	0.0217	0.0217	0.0217	0	0	0	0	0.087	67.5	5.8696
70	75	0	0.0217	0	0.0217	0	0	0	0.0217	0	0	0	0	0.0652	72.5	4.7283
75	80	0	0.0435	0.0217	0	0	0	0	0	0	0	0	0	0.0652	77.5	5.0543
80	85	0.0217	0.0435	0.0435	0.0435	0.0217	0.0217	0.0217	0.0217	0	0	0	0	0.2391	82.5	19.72826087
85	90	0.0217	0.0217	0.0435	0.0652	0.0435	0.0435	0.0435	0.0217	0.0217	0	0	0	0.3261	87.5	28.5326
90	95	0	0	0	0.0435	0	0	0.0217	0	0	0	0	0	0.0652	92.5	6.0326
95	100	0.0217	0	0	0	0	0	0	0	0	0	0	0.0217	0.0435	97.5	4.2391
100	105	0	0	0.0217	0	0	0	0	0	0	0	0	0	0.0217	102.5	2.2283
$\widehat{P}(X)$		0.0652	0.1522	0.1304	0.1739	0.087	0.1304	0.1304	0.087	0.0217	0	0	0.0217			
midpoint		2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5		$\widehat{E}(X)$	6.2391
mid $\cdot\ \widehat{P}(X)$		0.163	0.5326	0.587	0.9565	0.5652	0.9783	1.1087	0.8261	0.2283	0	0	0.2935		$\widehat{E}(Y)$	81.5217

The individual probability from the table comes from dividing frequency with the number of samples. $P[X]$ comes from summing the data for all columns. $P[Y]$ comes from summing the data for all rows. The expected value of X ($E[X]$) is 6.23913 which is calculated by summing all of $P[X]$ and the expected value of Y ($E[Y]$) is 81.5217 which is calculated by summing all of $P[Y]$.





The correlation coefficient is computed by using the formula

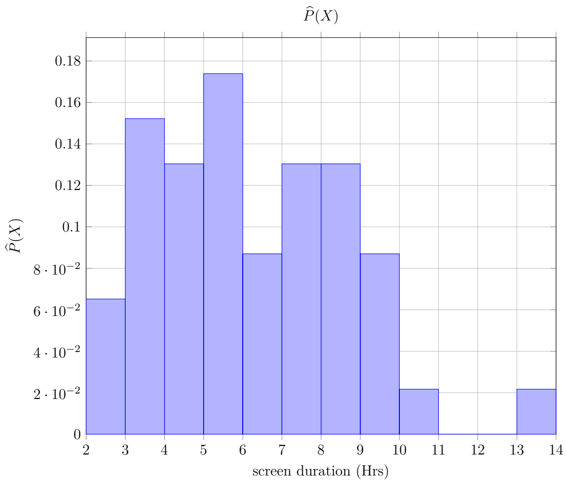
$$\widehat{\rho}_{X,Y} = \frac{\widehat{S}_{X,Y}}{s_X s_Y} \ \& \ \widehat{S}_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)\left(y_i - \overline{Y}\right)$$

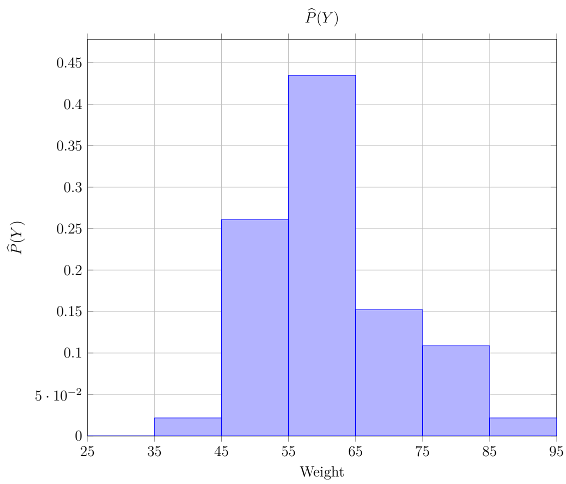
In this case, $\overline{X} = 6.271, \overline{Y} = 80.7174, n = 46, s_x = 2.4862, s_y = 10.8805$. Therefore, $\widehat{S}_{X,Y} = -1.3246$ and $\widehat{\rho}_{X,Y} = -0.0490$. Thus, the relation between X and Y has a weak negative relationship.

Screen time vs Weight

y \ x		2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12	12–13	13–14	$\widehat{P}(Y)$	midpoint	mid $\cdot\, \widehat{P}(Y)$
25	35	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0
35	45	0	0	0	0	0	0	0.0217	0	0	0	0	0	0.0217	40	0.8696
45	55	0	0.1087	0	0.0652	0	0.0435	0.0217	0.0217	0	0	0	0	0.2609	50	13.0435
55	65	0.0435	0.0217	0.0652	0.087	0	0.0435	0.0652	0.0652	0.0217	0	0	0.0217	0.4348	60	26.087
65	75	0.0217	0	0.0435	0.0217	0.0435	0.0217	0	0	0	0	0	0	0.1522	70	10.6522
75	85	0	0	0.0217	0	0.0435	0.0217	0.0217	0	0	0	0	0	0.1087	80	8.6957
85	95	0	0.0217	0	0	0	0	0	0	0	0	0	0	0.0217	90	1.9565
$\widehat{P}(X)$		0.0652	0.1522	0.1304	0.1739	0.087	0.1304	0.1304	0.087	0.0217	0	0	0.0217			
midpoint		2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5		$\widehat{E}(X)$	6.2391
mid $\cdot\, \widehat{P}(X)$		0.163	0.5326	0.587	0.9565	0.5652	0.9783	1.1087	0.8261	0.2283	0	0	0.2935		$\widehat{E}(Y)$	61.3043

The individual probability from the table comes from dividing frequency with the number of samples. $P[X]$ comes from summing the data for all columns. $P[Y]$ comes from summing the data for all rows. The expected value of X ($E[X]$) is 6.23913 which is calculated by summing all of $P[X]$ and the expected value of Y ($E[Y]$) is 61.3043 which is calculated by summing all of $P[Y]$.





The correlation coefficient is computed by using the formula

$$\widehat{\rho}_{X,Y} = \frac{\widehat{S}_{X,Y}}{s_X s_Y} \;\&\; \widehat{S}_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)\left(y_i - \overline{Y}\right)$$

In this case, $\overline{X} = 6.271, \overline{Y} = 61.0587, n = 46, s_x = 2.4862, s_y = 10.3099$. Therefore, $\widehat{S}_{X,Y} = -0.3518$ and $\widehat{\rho}_{X,Y} = -0.0137$. Thus, the relation between X and Y has a weak negative relationship.

Statistics

Descriptive Statistics

	Age	Average daily screen time (Hrs)	Average daily notification received	Weight (Kg)	Heart Rate	Sleeping time (Hrs)
Count	55	55	55	55	55	55
MEAN	29.1818	6.5771	103.7506	61.9709	80.8909	7.3098
median	21	6.2024	86	60.000	83	7.34
mode	19	6.2024	86	50.000	87	6.50
MIN	7	2.2357	6.4286	25.70	51	4.50
MAX	57	16.6714	357	111.1	111	10.5333
range	50	14.4357	350.5714	85.4	60	6.0333
variance	175.4478	7.9495	5585.4741	194.314	168.4323	1.4176
SD	13.2457	2.8195	74.736	13.9397	12.9781	1.1906
cv	0.4539	0.4287	0.7203	0.2249	0.1604	0.1629
MAD	11.7752	2.2193	56.0073	9.9298	9.9451	0.8687
quartile1 (Q1)	19	4.569	49.1429	52.5	73	6.5
quartile3 (Q3)	43	8.3167	146.7143	67	88	8
IQR	24	3.7476	97.5714	14.5	15	1.5
Q1-1.5IQR	-17	-1.0524	-97.2143	30.75	50.5	4.25
Q3+1.5IQR	79	13.9381	293.0714	88.75	110.5	10.25
Outliers (based on IQR) If no outlier, answer None.	None	16.671	294.857, 357	25.7, 93, 111.1	111	10.417, 10.533
MEAN-3SD	-10.5552	-1.8814	-120.4574	20.1519	41.9565	3.7379
MEAN+3SD	68.9188	15.0355	327.9587	103.7899	119.8254	10.8818
Outliers (based on SD) If no outlier, answer None.	None	16.6714	357	111.1	None	None
Mean after removing outliers based on IQR. If no outlier, type NA	NA	6.3901	95.3666	61.1269	80.3333	7.1904
SD after removing outliers based on IQR. If no outlier, type NA	NA	2.4782	61.5869	10.6309	12.4173	1.0355
Measure of Centrality	Median					
Reason	The data is right-skewed				The data is left-skewed	
Measure of Dispersion	IQR					
Reason	The data is skewed	The dataset has an outlier				

Goodness of Fit test

1. **Data set : Average screen duration**

   Type of distribution : Normal distribution

   Known parameter : 0

   Unknown parameter (m) : 2, which are μ, σ

   H_0 : Screen duration is normally distributed with $\mu = 6.2709$ and $\sigma = 2.4862$

   H_a : Screen duration is not normally distributed with $\mu = 6.2709$ and $\sigma = 2.4862$

   Number of cells with the expected number of samples (k) : 6

   Test static $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 2.0956$

   Significant level (α) : 0.05

   Degree of freedom 1 (ν_1) : $k - 1 - m = 6 - 1 - 2 \Rightarrow 3$

   Cutoff of non-rejection region : 7.8147

   Degree of freedom 2 (ν_2) : $k - 1 = 6 - 1 \Rightarrow 5$

   Cutoff of rejection region : 11.070

   Non-rejection regions : $\chi^2 < \chi^2_{0.05,3} = 7.8147$

   Rejection regions : $\chi^2 \geq \chi^2_{0.05,5} = 11.070$

   Rejection decision : Don't need to reject null hypothesis

   Conclusion : Screen duration is normally distributed with $\mu = 6.2709$ and $\sigma = 2.4862$

2. **Data set : Average Sleeping Duration**

Type of distribution : Normal distribution

Known parameter : 0

Unknown parameter (m) : 2, which are μ, σ

H_0 : Screen duration is normally distributed with $\mu = 7.2178$ and $\sigma = 1.0068$

H_a : Screen duration is not normally distributed with $\mu = 7.2178$ and $\sigma = 1.0068$

Number of cells with the expected number of samples (k) : 4

Test static $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 1.3428$

Significant level (α) : 0.05

Degree of freedom 1 (ν_1) : $k - 1 - m = 4 - 1 - 2 \Rightarrow 1$

Cutoff of non-rejection region : 3.8415

Degree of freedom 2 (ν_2) : $k - 1 = 4 - 1 \Rightarrow 3$

Cutoff of rejection region : 7.8147

Non-rejection regions : $\chi^2 < \chi^2_{0.05,1} = 3.8415$

Rejection regions : $\chi^2 \geq \chi^2_{0.05,3} = 7.8147$

Rejection decision : Don't need to reject null hypothesis

Conclusion : Screen duration is normally distributed with $\mu = 7.2178$ and $\sigma = 1.0068$

3. **Data set : Weight**

Type of distribution : Normal distribution

Known parameter : 0

Unknown parameter (m) : 2, which are μ, σ

H_0 : Screen duration is normally distributed with $\mu = 61.0587$ and $\sigma = 10.3099$

H_a : Screen duration is not normally distributed with $\mu = 61.0587$ and $\sigma = 10.3099$

Number of cells with the expected number of samples (k) : 6

Test static $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 3.1761$

Significant level (α) : 0.05

Degree of freedom 1 (ν_1) : $k - 1 - m = 6 - 1 - 2 \Rightarrow 3$

Cutoff of non-rejection region : 7.8147

Degree of freedom 2 (ν_2) : $k - 1 = 6 - 1 \Rightarrow 5$

Cutoff of rejection region : 11.070

Non-rejection regions : $\chi^2 < \chi^2_{0.05,3} = 7.8147$

Rejection regions : $\chi^2 \geq \chi^2_{0.05,5} = 11.070$

Rejection decision : Don't need to reject null hypothesis

Conclusion : Screen duration is normally distributed with $\mu = 61.0587$ and $\sigma = 10.3099$

4. **Data set : Average Notification**

Type of distribution : Normal distribution

Known parameter : 0

Unknown parameter (m) : 2, which are μ, σ

H_0 : Screen duration is normally distributed with $\mu = 93.7671$ and $\sigma = 56.1233$

H_a : Screen duration is not normally distributed with $\mu = 93.7671$ and $\sigma = 56.1233$

Number of cells with the expected number of samples (k) : 6

Test static $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 1.4815$

Significant level (α) : 0.05

Degree of freedom 1 (ν_1) : $k - 1 - m = 6 - 1 - 2 \Rightarrow 3$

Cutoff of non-rejection region : 7.8147

Degree of freedom 2 (ν_2) : $k - 1 = 6 - 1 \Rightarrow 5$

Cutoff of rejection region : 11.070

Non-rejection regions : $\chi^2 < \chi^2_{0.05,3} = 7.8147$

Rejection regions : $\chi^2 \geq \chi^2_{0.05,5} = 11.070$

Rejection decision : Don't need to reject null hypothesis

Conclusion : Screen duration is normally distributed with $\mu = 93.7671$ and $\sigma = 56.1233$

Hypothesis Test

1. **Data set : Average screen duration**

   Test hypothesis claims that *average screen duration* is greater than 7 Hrs.
   Collect data from 55 people. Sample mean sample standard deviation : 6.5771 & 2.8195 Hrs respectively.

   Case of Hypothesis test : Large sample size

   Upper-tailed, lower-tailed or two-tailed test : upper-tailed

   7 steps test

   1.1 Choose parameter of interest (μ)

   Parameter : μ

   1.2 Specify null value (μ_0) and null hypothesis (H_0)

$$\begin{aligned} \mu &: \text{average screen duration} \\ \mu_0 &: 7 \\ H_0 &: \mu = 7 \end{aligned} \tag{1}$$

   1.3 State alternative hypothesis (H_a)

$$H_a : \mu > 7 \tag{2}$$

   1.4 Compute test statistic (z)

$$z = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{6.2709 - 7}{\frac{2.4862}{\sqrt{46}}} \Rightarrow -1.989 \tag{3}$$

   1.5 Indicate significance level (α) and find rejection region (z_a)

$$\alpha : 0.05$$

$$1 - \alpha = 1 - 0.05 \quad = 0.95$$
$$z_a = z_{0.05} \quad = 1.6449 \tag{4}$$
$$z \geq 1.6449$$

1.6 Determine whether we reject null hypothesis or not

Test static does not fall inside the rejection region. Null hypothesis is not rejected.

1.7 Conclude the problem

We do not reject the null hypothesis. Average screen duration is 7 Hrs.

2. **Data set : Average notification**

Test hypothesis claims that *average notification* is greater than 100 times.
Collect data from 55 people. Sample mean sample standard deviation : 103.7506 & 74.7360 times respectively.

Case of Hypothesis test : Large sample size

Upper-tailed, lower-tailed or two-tailed test : upper-tailed

7 steps test

2.1 Choose parameter of interest (μ)

Parameter : μ

2.2 Specify null value (μ_0) and null hypothesis (H_0)

$$\mu : \text{average notification}$$
$$\mu_0 : 100 \tag{5}$$
$$H_0 : \mu = 100$$

2.3 State alternative hypothesis (H_a)

$$H_a : \mu > 100 \tag{6}$$

2.4 Compute test statistic (z)

$$z = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{93.7671 - 100}{\frac{56.1233}{\sqrt{46}}} \Rightarrow -0.7532 \tag{7}$$

2.5 Indicate significance level (α) and find rejection region (z_a)

$$\alpha : 0.01$$

$$1 - \alpha = 1 - 0.05 \quad = 0.99$$
$$z_a = z_{0.01} \quad = 2.3263 \tag{8}$$
$$z \geq 2.3263$$

2.6 Determine whether we reject null hypothesis or not

Test static does not fall inside the rejection region. Null hypothesis is not rejected.

2.7 Conclude the problem

We do not reject the null hypothesis. Average notification is 100 times.

Conclusion

From the question that we asked the respondents, we want to evaluate how demographic factors (age, weight) influence screen behavior and analyze the impact of screen duration on daily routines, including which days of the week have the highest and lowest usage.

We found that age has a weak negative relationship with screen duration, with a correlation coefficient of -0.3340. Weight does not have any relationship with screen duration, with a correlation coefficient of -0.0137. The median screen duration is 6.2024, the median weight is 60.00 and the median age is 21.

The day that most respondents have spent their screen duration is Sunday with 23.9% and the lowest screen duration is Monday with 28.3%. The median sleep duration is 7.34. However, we found that screen duration and sleep duration are not correlated with the correlation coefficient of -0.1431.

The flaw of this is that the majority of people who answer the survey are around 18 & 19 years old which might cause the result to be inaccurate.

Appendix

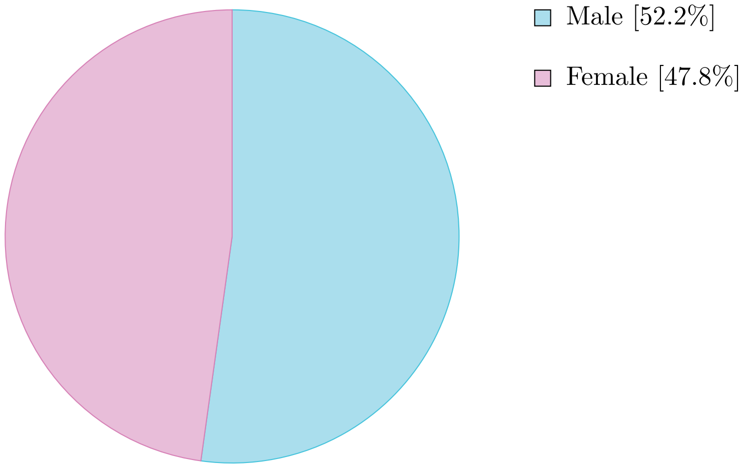


Figure 1: Pie Chart of Gender

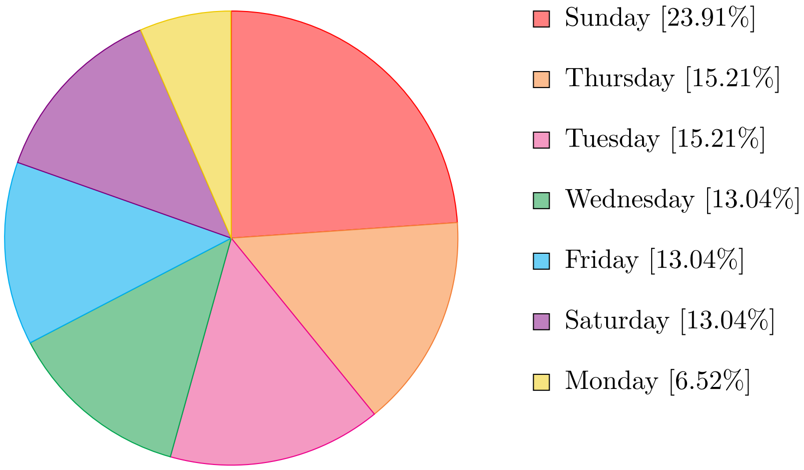


Figure 2: Pie chart with legend to the right (always outside).

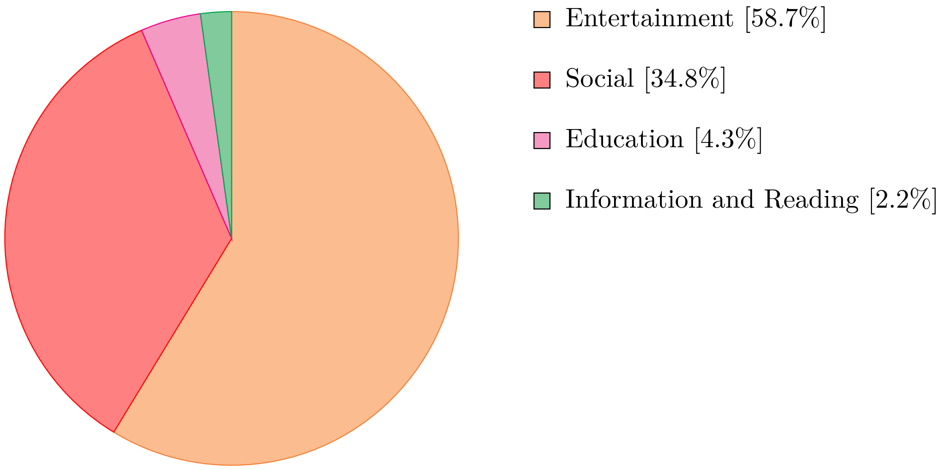


Figure 3: Pie chart with legend to the right (always outside).

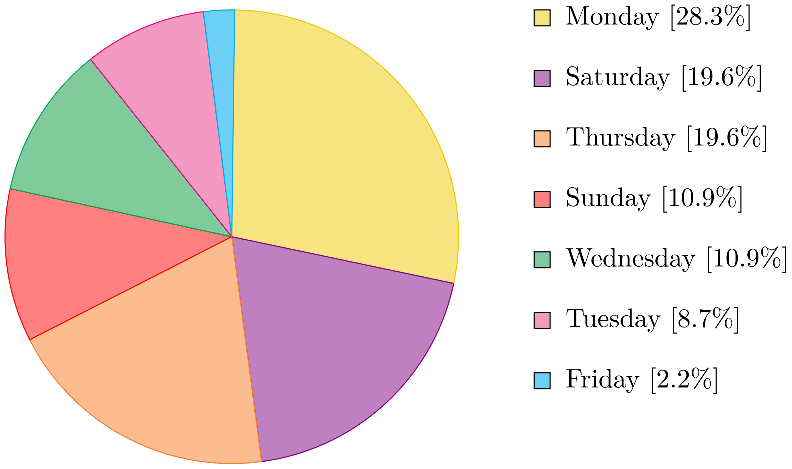


Figure 4: Pie chart with legend to the right (always outside).

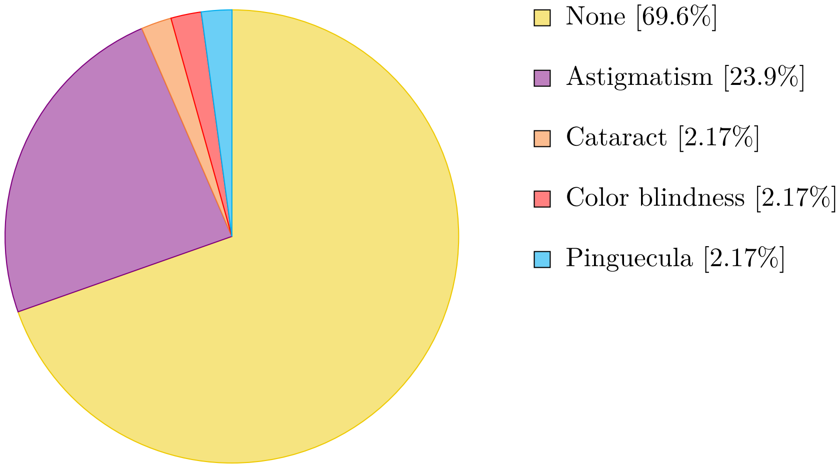


Figure 5: Pie chart with legend to the right (always outside).

y \ x		2	3	4	5	6	7	8	9	10	11	12	13	$\widehat{P}(Y)$	midpoint	mid $\cdot \widehat{P}(Y)$
		3	4	5	6	7	8	9	10	11	12	13	14			
-850	-750	0	0	0	0	0	0	0.0227	0	0	0	0	0	0.0227	-800	-18.18181818
-750	-650	0	0	0	0	0	0	0	0	0	0	0	0	0	-700	0
-650	-550	0	0	0	0	0.0227	0.0227	0	0	0	0	0	0	0.0455	-600	-27.27272727
-550	-450	0	0.0227	0	0	0	0	0	0	0	0	0	0	0.0227	-500	-11.36363636
-450	-350	0	0	0	0.0455	0	0	0.0227	0	0.0227	0	0	0	0.0909	-400	-36.36363636
-350	-250	0	0.0227	0.0227	0	0	0.0227	0	0.0227	0	0	0	0	0.0909	-300	-27.27272727
-250	-150	0.0227	0	0	0.0227	0.0227	0.0227	0	0	0	0	0	0	0.0909	-200	-18.18181818
-150	-50	0	0.0227	0.0227	0.0227	0	0.0227	0	0	0	0	0	0.0227	0.1136	-100	-11.36363636
-50	50	0.0455	0.0682	0.0455	0.0682	0.0455	0.0455	0.0682	0.0682	0	0	0	0	0.4545	0	0
50	150	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
150	250	0	0	0	0.0227	0	0	0.0227	0	0	0	0	0	0.0455	200	9.090909091
250	350	0	0	0.0227	0	0	0	0	0	0	0	0	0	0.0227	300	6.818181818
$\widehat{P}(X)$		0.0682	0.1364	0.1136	0.1818	0.0909	0.1364	0.1364	0.0909	0.0227	0	0	0.0227			
midpoint		2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5		$\widehat{E}(X)$	6.340909091
mid $\cdot \widehat{P}(X)$		0.1704545455	0.4772727273	0.5113636364	1	0.5909090909	1.022727273	1.159090909	0.8636363636	0.2386363636	0	0	0.3068181818		$\widehat{E}(Y)$	-134.0909091

Table 1: Left eye

Min X	Min Y
2.2357	-800
Max X	Max Y
13.8500	250
$\overline{X}$	$\overline{Y}$
6.3647	-142.5000
$\sum(X-\overline{X})(Y-\overline{Y})$	-3074.4167
$\widehat{\text{cov}}_{X,Y} = \dfrac{\sum(X-\overline{X})(Y-\overline{Y})}{n-1}$	-69.8731
s_X	2.4974
s_Y	225.1834
$\rho_{X,Y} = \frac{\widehat{\text{cov}}_{X,Y}}{s_X \cdot s_Y}$	-0.1242

Table 2: Left eye correlation

y \ x		2	3	4	5	6	7	8	9	10	11	12	13	$\widehat{P}(Y)$	midpoint	mid $\cdot \widehat{P}(Y)$
		3	4	5	6	7	8	9	10	11	12	13	14			
-950	-850	0	0	0	0	0	0	0.0233	0	0	0	0	0	0.0233	-900	-20.93023256
-850	-750	0	0	0	0	0	0	0	0	0	0	0	0	0	-800	0
-750	-650	0	0	0	0	0	0	0	0	0	0	0	0	0	-700	0
-650	-550	0	0	0	0	0.0233	0.0233	0	0.0233	0	0	0	0	0.0698	-600	-41.86046512
-550	-450	0	0.0233	0	0	0	0	0	0	0	0	0	0	0.0233	-500	-11.62790698
-450	-350	0.0233	0	0	0.0465	0	0.0233	0.0233	0	0	0	0	0	0.1163	-400	-46.51162791
-350	-250	0	0	0.0233	0	0	0	0	0	0	0	0	0	0.0233	-300	-6.976744186
-250	-150	0	0	0	0.0233	0.0233	0.0233	0	0	0.0233	0	0	0	0.093	-200	-18.60465116
-150	-50	0	0.0233	0.0233	0	0	0.0233	0.0233	0	0	0	0	0.0233	0.1163	-100	-11.62790698
-50	50	0.0465	0.0698	0.0465	0.093	0.0465	0.0465	0.0465	0.0698	0	0	0	0	0.4651	0	0
50	150	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
150	250	0	0	0	0.0233	0	0	0.0233	0	0	0	0	0	0.0465	200	9.302325581
250	350	0	0	0.0233	0	0	0	0	0	0	0	0	0	0.0233	300	6.976744186
$\widehat{P}(X)$		0.0698	0.1163	0.1163	0.186	0.093	0.1395	0.1395	0.093	0.0233	0	0	0.0233			
midpoint		2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5		$\widehat{E}(X)$	6.406976744
mid $\cdot \widehat{P}(X)$		0.1744186047	0.4069767442	0.523255814	1.023255814	0.6046511628	1.046511628	1.186046512	0.8837209302	0.2441860465	0	0	0.3139534884		$\widehat{E}(Y)$	-141.8604651

Table 3: Right eye

Min X	Min Y
2.2357	-875
Max X	Max Y
13.8500	250
$\overline{X}$	$\overline{Y}$
6.4204	-151.7442
$\sum(X-\overline{X})(Y-\overline{Y})$	-3016.2099
$\widehat{\text{cov}}_{X,Y} = \dfrac{\sum(X-\overline{X})(Y-\overline{Y})}{n-1}$	-70.1444
s_X	2.4992
s_Y	244.4865
$\rho_{X,Y} = \frac{\widehat{\text{cov}}_{X,Y}}{s_X \cdot s_Y}$	-0.1148

Table 4: Right eye correlation