PROPOSAL

Qile (Cheela) Zhu, Jiaxin (Cecilia) Liu

1. Research Question

Are high school students' sleeping times negatively associated with their screen usage times?

In our daily life, we always complain that high school students on the international campus are not able to get enough sleep. However, we found that most students use a phone to communicate with others near bedtime and they might stay up later to wait for a reply. Thus, we want to find a correlation between their screen usage time and students' sleeping time.

2. Background Research

Numerous studies have found a correlation between using screen-based devices before bed and increased sleep latency, or the time it takes someone to fall asleep. For our ages, an estimated two out of every three teenagers sleep less than the recommended amount, and screen usage may be to blame for sleep deprivation and other issues.

Teenagers' use of screens is associated with a variety of insomnia symptoms. Screen usage prolongs the time it takes to fall asleep and results in less restful sleep overall by delaying the production of melatonin. A later bedtime often leads to disturbed sleep and increased next-day drowsiness because most teens have stringent school start times. The circadian rhythm is eventually disrupted by routinely staying up late throughout the week and sleeping in on the weekends. Long-time use of electronic devices might cause measurement errors (of both screen time exposure and sleep measures). Measures of screen time might evolve to reflect teenagers' contemporary digital landscape.

References:

Byrne, Rebecca et al. “Measurement of screen time among young children aged 0-6 years: A systematic review.” Obesity reviews : an official journal of the International Association for the Study of Obesity vol. 22,8 (2021): e13260. doi:10.1111/obr.13260

Lacy, Kathleen E., et al. "Screen time and physical activity behaviours are associated with health-related quality of life in Australian adolescents." Quality of life Research 21 (2012): 1085-1099.

https://www.researchgate.net/publication/351048905_Screen_time_and_its_impact_on_health

3. Sampling and Experimental Design

Variables

Explanatory Variable: High School Students’ Screen Usage Times

Response Variable: High School Students’ Sleeping Times

Type of Study

Only data is collected but no treatment is conducted on the individuals, so this inference is an observational study.

Data Collection

We will randomly select a sample of AP students in grade 10 and grade 11. First, a list of G10 AP students (sorted in alphabetical sequence by initials) will be obtained. Then, a random number generator will be used to randomly generate 25 distinct numbers within the total number of G10 AP students. The students with the corresponding rankings in the list will be chosen, and a form will be sent to them. The same procedure will be repeated for G11 AP students. The form will include the following questions:

1. What is your screen usage time in the past 7 days?
2. What is your sleeping time? (Multiple choices allowed)
    a. 3 – 4 hours
    b. 5 – 6 hours
    c. 7 – 8 hours

d. 9 – 10 hours
e. More than 10 hours

3. What is your current grade?
    a. G10
    b. G11

Data will be collected from their responses.

Scope of Inference

Because the data comes from an observational study, it cannot be used to establish causal links. The sample is randomly selected from solely all AP students in grade 10 and grade 11 at Beijing National Day School, hence there will be slight differences between different programs, different grades and different schools. However, in the rough, the conclusion from this inference can provide a reasonable estimation of the true association between sleeping time and screen usage time for all high school students.

4. Exploratory Data Analysis

Let the sleeping times of all high school students in the past 7 days be Y, the screen usage times of all high school students in the past 7 days be X.

Hence the sample sleeping times in the past 7 days is y, and the sample screen usage times in the past 7 days is x.

The relevant descriptive statistics and visualization of the data are:

- The average value ($\bar{y}$), standard error (s_y) and degree of freedom of y
- The average value ($\bar{x}$), standard error (s_x) and degree of freedom of x
- The sample size (n)
- A least-squares regression line of y vs. x will be drawn to visualize the data. The slope of the line (b) will be calculated.
- A residual plot of residuals (observed y – estimated y from least-squares regression line, $y_i - \hat{y}$) vs. x

Conditions that are checked or assumed for analysis:

- The sample is randomly collected. (Checked)
- The data points are independent. (Assumed)
- For a particular value of x, y values are normally distributed. (Checked from the least-squares regression line)
- The standard deviation of errors must be constant across values of x (Checked from the residual plot)
- The true relationship between x and y is linear (Assumed)

A one-sample t-test for the slope of the least-squares regression line for Y vs. X (β) will be conducted.

Null Hypothesis: $\beta = 0$ Alternative Hypothesis: $\beta < 0$

Test Statistics: $t_{sta} = \frac{b-\beta}{s_b}$ Degree of Freedom: $n - 2$

The p-value is calculated. If $p - value < \alpha = 0.05$, there is convincing evidence that Y is negatively associated with X. If $p - value > \alpha = 0.05$, there is no convincing evidence.

5. Group Task Assignments and Timeline

Timeline

5.16--5.21 collecting data and finishing the parts before data analysis.

5.21--5.22 Performing relevant descriptive statistics

5.23 Writing suggestions and conclusion.

5.24--5.25 preparing for presentation and revising the first draft.

Group Task Assignments

Cheela: parts A, B, C

Cecilia: parts D, E, F

6. Data

Average Screen Usage Time in the Past 7 Days (hours)	Average Sleeping Time in the Past 7 Days (hours)
5.50	7 - 8
4.00	7 - 8
7.14	6 - 7
2.57	6 - 7
8.57	7 - 8
4.29	7 - 8
8.14	7 - 8
3.71	6 - 7
6.00	5 - 6
2.86	6 - 7
3.50	5 - 6
10.00	5 - 6
6.00	7 - 8
10.57	3 - 4
4.00	7 - 8