

NM2207 project diary entry

Wu Yutong

2023-11-19

week 9

Topic

I will be conducting an analysis of insomnia data with the aim of offering a succinct overview on sleep difficulties. The goal is to provide users and viewers of the interactive data with a sense of connection at the demographic level and to inspire them with a range of coping techniques.

Data source

I will be examining the primary factors that lead to insomnia, exploring its effects on health, and evaluating the effectiveness of coping strategies, all while considering variations across different age groups and genders.

week 10

(1) What is the question that you are going to answer?

What is the gender and age disparity in insomnia (severity, causes and coping skills) in young adults? pre ## (2) Why is this an important question?

Sleep is a fundamental component of our lives, intricately linked to our mental well-being and daily functioning (NIH, 2022). Nevertheless, the issue of insomnia has risen to concerning levels, with a significant increase in sleep-related problems observed in young adults, going from 22.6% in 2010 to 30.5% in 2018 (Sivertsen, B. et al, 2019). In contrast to the commonly held belief that older adults experience more sleep difficulties than their younger counterparts, research indicates that insomnia affects over 1 in 4 individuals aged 18 to 24, with 29% experiencing it every night, making this age group the most vulnerable in the United States (Helsestart, 2022). Through an examination of the underlying causes of insomnia across different age groups and genders, we can unveil the distinctive challenges that this population faces today and explore effective approaches to address these concerns.

(3) Which rows and columns of the dataset will be used to answer this question?

All rows with valid data entry are used (86 rows after data cleaning).The columns used are categorised into 4 main groups: participants info, severity of insomnia, causes of insomnia and coping skills.

participants info: Group, SubGroup, Sex, Age

severity of insomnia: ISI_total

causes of insomnia:BDI_total, ASHS_total, ASHS_physiological, ASHS_cognitive,ASHS_emotional, ASHS_SleepEnvironmnt, ASHS_DaytimeSleep, ASHS_substances, ASHS_bedtimeRoutine, ASHS_sleepStability, ASHS_BedroomSharing, DBAS_total, FIRST_total, GCTI_total, ACE_tot, asq_home, asq_school, asq_attendance, asq_romantic, asq_peer, asq_teacher, asq_future, asq_leisure, asq_finance, asq_responsibility

coping skills: cope_disengage_su, cope_growth, cope_disengage_mental, cope_emotions, cope_socialsupp_instr, cope_active, cope_denial, cope_religion, cope_humor, cope_disengage_emo, cope_restraint, cope_socialsupp_emo, cope_accept, cope_suppression, cope_planning

Include the challenges and errors that you faced and how you overcame them.

challenge 1: unable to publicise the app (when all workings on data cleaning and graph plotting is done on a R markdown file, app.R can still show on local host because information required is found in the workspace. However, when trying to publicise the app, it will not work)

solution: transfer all codes, including library(tidyverse), read.csv(), and data cleaning to app.R file so that when running the app, data files can be accessed.

challenge 2: for my eventual data visualisation, I need to use a lot of variables from the data set and a lot of them have specific qualities which make it very messy to manipulate them

solution: I tried to regroup and recategorise them by assigning 4 big titles: participants, sleep quality, casues, and coping skills; subsequently creating 4 different datasets. Within each group, there are also smaller groups like ASHS_total (Adolescent Sleep Hygiene Scale) and distinct categories of ASHS (eg. cognitive domain, emotional domain). I will further organise them by creating subgroups of data.

week 11

(1) List the visualizations that you are going to use in your project

First group of visualisation: I am going to plot age (Age) and gender (Sex) against severity of insomnia (ISI_total). Age-severity of insomnia: line plot Gender-severity of insomnia: bar plot

Second group of visualisation: analyse causes of insomnia (BDI_total, ASHS_total, ASHS_physiological, ASHS_cognitive,ASHS_emotional, ASHS_SleepEnvironmnt, ASHS_DaytimeSleep, ASHS_substances, ASHS_bedtimeRoutine, ASHS_sleepStability, ASHS_BedroomSharing, DBAS_total, FIRST_total, GCTI_total, ACE_tot, asq_home, asq_school, asq_attendance, asq_romantic, asq_peer, asq_teacher, asq_future, asq_leisure, asq_finance, asq_responsibility)

depression (BDI_total) -severity of insomnia: scatter plot

sleep hygiene (ASHS_total) -severity of insomnia: scatter plot

rank all sleep hygiene factors (ASHS_physiological, ASHS_cognitive,ASHS_emotional, ASHS_SleepEnvironmnt, ASHS_DaytimeSleep, ASHS_substances, ASHS_bedtimeRoutine, ASHS_sleepStability, ASHS_BedroomSharing): histogram

predisposition to insomnia (FIRST_total)- severity of insomnia: scatter plot

dysfunctional beliefs (DBAS_total) - severity of insomnia: scatter plot

cognitive intrusiveness (GCTI_total) - severity of insomnia: scatter plot

rank all stress factors (asq_home, asq_school, asq_attendance, asq_romantic, asq_peer, asq_teacher, asq_future, asq_leisure, asq_finance, asq_responsibility): histogram

word cloud of all the causes of insomnia

Third group of visualisation: compare effectiveness of coping skills - pie chart

(2) How do you plan to make it interactive?

ggplot2: to plot the graphs (scatter plot, histogram, bar plot, etc.)

shiny: 1. create drop down bar for the viewer to select the specific factor he/she wants to focus at; 2. wordcloud to give a visual representation of the causes of insomnia

plotly, gganimate, transformr: to make the graph mroe interactive portray a clearer story (eg. highlight the highest point on graph)

(3) What concepts incorporated in your project were taught in the course and which ones were self-learnt?

Weeks	Topics
2	exploratory data analyses; intro to ggplot and shiny
3	working with variables and summary statistics
4	tidy data
5	writing functions
6	loop
7	visualising data using ggplot
8	visualising data using shiny
	plotly
	dslib
	dplyr

the challenges and errors that you faced and how I overcame them

challenge 3 Initially, I wanted to show the ranking of causes of insomnia using histogram but when after data manipulation, I realised that the causes of insomnia are measured not on a consistent scale. Therefore, comparing all causes in a single graph is not valid.

solution I decided to separate the data sets and regroup the causes into depression, sleep hygiene, predisposition, dysfunctional beliefs, cognitive intrusiveness. My focus is now changed to display their association with insomnia.

week 12

the challenges and errors that you faced and how I overcame them

challenge 4 When trying to make the plot more asthetic, I faced difficulty trying to customise the x-axis label as as there are 5 plots in the function and I don't know how to change x-axis label for each plot.

solution instead of modifying the function of the app, I renamed the variable to the desired x-axis label. This is easier way to achieve my desired outcome.

challenge 5 The initial plan to use word cloud and animation is not actualised as when trying out the visualisation, the display is not clear and defeats the purpose of making the message clear to the viewers.

solution Instead of word cloud and animation on the plots, I chose to make the ideas concise by using simple widgets (eg. radial button).

Final submission

(1) What is the theme of your data story?

The focus of my data story centres around the theme of insomnia.

(2) Why is it important to address this question?

Sleep is a fundamental aspect of our lives intricately linked to mental well-being and daily functioning (HHS, 2022) that faces a growing concern with the prevalence of insomnia. Notably, sleep-related issues in young adults have surged from 22.6% in 2010 to 30.5% in 2018 (Sivertsen et al., 2019). Contrary to the common belief that older adults experience more sleep difficulties, research reveals that insomnia affects over 25% of individuals aged 18 to 24, with 29% experiencing it every night, making this age group the most vulnerable in the United States (Helsestart, 2022). By examining the distribution of insomnia across various age groups and genders and exploring its underlying causes, we can uncover the distinct challenges faced by this population and identify effective approaches to address these concerns.

(3) Why do you think the data sources that you have curated can help you answer the question?

The selected data source (Singh, 2023) comprises a comprehensive set of responses from a sample of young adults potentially experiencing sleep problems. It includes information on gender and age, ratings for insomnia severity, and causes of insomnia.

(4) What are the insights from the data and how are they depicted in plots?

The data visualization provides a clearer representation of the distribution of insomnia among young adults, with those between 17 and 17.5 years old and females being the most severely affected. Scatter plots and best fit lines offer insights into how various causes of insomnia correlate with its severity. For instance, depression, predisposition to insomnia, dysfunctional beliefs, and cognitive intrusiveness positively correlate with insomnia severity, while sleep hygiene shows a negative correlation.

(5) How did you implement this entire project? Were there any new concepts that you learnt to implement some aspects of it?

The entire project commenced with data cleaning and sense-making of the available data. Initially, there was ambition regarding the potential manipulation of numerous variables for the final data visualization outcome. However, a more in-depth understanding of the data revealed challenges in interpreting certain data sets (e.g., coping skills with only numerical inputs and no detailed explanation of the rating scale or assessment used during data collection). Consequently, it was deemed irrational to proceed with visualizing these ambiguous data sets. The finalized data sets were then used to experiment with the Shiny App.

Two final apps were developed: (1) Demographics of Insomnia and (2) Causes of Insomnia. The interaction element for App (1) includes a drop-down bar for users to select demographic factors to view specific plots (e.g., Age as a line graph or Gender as a pie chart). Noteworthy concepts learned during the construction of App (1) include highlighting the maximum point on a line graph, customizing the app theme using the dslib library, and embedding two different types of plots (line graph and pie chart) into one app. App (2) incorporates a radial button for users to choose specific causes correlated with insomnia (e.g., depression, sleep hygiene, predisposition to insomnia, dysfunctional beliefs, or cognitive intrusiveness). Key concepts learned during the construction of App (2) include creating best-fit lines for scatter plots and developing a function that allows for the selection of different graphs without duplicating the code for each plot.

Beyond the apps, additional concepts were learned building the website, including creating a fade-in effect for images, dividing the webpage into columns for information compartmentalization, using collapse blocks for additional information, incorporating callout blocks for highlighting important information, utilizing flip cards, incorporating text boxes, adding background images, and customizing the website font.

References

Bruce, E. S., Lunt, L., McDonagh, J. E. (2017). Sleep in adolescents and young adults. Clin Med (Lond), 17(5): 424–428. <https://doi.org/10.7861/clinmedicine.17-5-424>.

Helsestart. (n.d.). Global Insomnia Statistics in 2022 & 2023. Global insomnia statistics in 2022 & 2023. <https://www.helsestart.no/news/global-insomnia-statistics>.

Roth, T. (2007). Insomnia: Definition, prevalence, etiology, and consequences. Journal of Clinical Sleep Medicine, 3(5 suppl). <https://doi.org/10.5664/jcsm.26929>.

Singh, U. (2023, July 25). Adolescent insomnia study data. Kaggle. <https://www.kaggle.com/datasets/utkarshx27/insomnia-symptomatology-in-adolescence>.

Sivertsen, B., Vedaa, Ø., Harvey, A. G., Glozier, N., Pallesen, S., Aaro, L. E., Lonning, K. J., & Hysing, M. (2018). Sleep patterns and insomnia in young adults: A national survey of norwegian university students. Journal of Sleep Research, 28(2). <https://doi.org/10.1111/jsr.12790>.

U.S. Department of Health and Human Services (HHS). (n.d.). Why is sleep important?. National Heart Lung and Blood Institute. <https://www.nhlbi.nih.gov/health/sleep/why-sleep-important>.