NM2207 project diary entry

Wu Yutong

2023-11-03

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?

(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_SleepEnvirnmont, 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_acccept, 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 (Answer: What are the variables that you are going to plot? How will it answer your larger question?)

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_SleepEnvirnmont, 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_SleepEnvirnmont, ASHS_DaytimeSleep, ASHS_substances, ASHS_bedtimeRoutine, ASHS_sleepStability, ASHS_BedroomSharing): histogram

predisposition to insomnia - severity of insomnia: scatter plot

dysfunctional beliefs - severity of insomnia: scatter plot

cognitive instrusiveness - 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? (Answer: features of ggplot2/shiny/markdown do you plan to use to make the story 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? (Answer: Create a table with topics in one column and Weeks in the other to indicate which concept taught in which week is being used. Leave the entry of the Week column empty for self-learnt concepts)

Weeks	Topics
2	exploratory data analyses; intro to ggplot and shiny
3	working with variables and summary statistics
5	writing functions
7	visualising data using ggplot
8	visualising data using shiny
	wordcloud in shiny
	plotly
	gganimate
	transformr

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.