How Do I Spend my Time?

Stat231: Google Calendar Assignment

Due Friday, September 25 by 5:00 PM EST

Overview

This assignment will be undertaken individually, though I encourage discussion of general ideas with your peers. I also encourage you to use other resources (including course materials, the internet, me, ...).

How do I spend my time?

One challenge of remote learning that many students voiced last spring is the difficulty of structuring our time in the absence of the usual structures in place when on campus. Even for students *on* campus this semester, the increase of online and hyflex classes, and re-imagined extracurricular activities, may make it difficult to manage time efficiently.

How *are* you spending your time? In this assignment, you'll track how you're spending your time using Google Calendar, then import the calendar data into R to analyze and provide insights into how you're spending your time. Note that you can use another electronic calendar (macOS, Microsoft Outlook), so long as you can download a `.ics` file from it.

The assignment aims to give you the opportunity to "play the whole [Data Science] game", including proposing a question of interest that can be addressed with data and partaking in *data collection*, as well as wrangling, visualizing, analyzing, and reflecting on the information gathered. I hope an ancillary benefit is that it provides you with insights into how you're spending your time, and ideas for how better to manage your time for the rest of the semester. (One strategy: create a "weekly plan" at the beginning of each week, and document any changes as the week goes on.)

Feel free to extend upon this basic question "how do I spend my time?" or explore a variation of it. For instance, some other ideas include:

- Document "intended time" doing things (e.g. studying, sleeping) versus "actual time" doing those things, and compare results
- Document time spent on each course, and/or time spent on different parts of a course (e.g. synchronous session, discussion forum, homework, etc.)
- Document time spent on school vs. work vs. leisure vs. rest, etc.
- If you already use Google Calendar as a way to keep track of your schedule, you could compare how your time was spent last fall (on campus, pre-pandemic) to how your time is being spent this semester (mid-pandemic)

Learning Objectives

The learning goals of this assignment relate to:

- Data Collection

  - Experience creating measurable data observations (e.g. how is "one day" measured? what defines "studying"?)

  - Address data collection constraints due to limits in technological capacity and human behavior
- Data Wrangling
  - Use known data verbs in new application
  - Identify new data verbs that are appropriate to implement in this context
- Data Ethics
  - Practice the ethical and legal responsibilities of those collecting, storing, and analyzing data
  - Decide limits for personal privacy
  - Deliberate on the trade-offs between research results and privacy
- "Playing the whole game"
  - Tie together data collection, analysis, ethics and communication components
  - Iterate between and within the components of the "whole game"

Details

What to do

1. Identify 2-4 primary questions of interest to you about how you spend your time. We will be sharing results with peers, so be sure that you identify questions *you feel comfortable sharing with your peers and instructor*.
2. Track your time in your calendar application for 14 days or more. (If you already use a calendar app, I suggest you create a new calendar within the app dedicted to this activity. That way, your pre-existing calendars will be kept private.) Fill in blocks of time and mark an entry with the activity you were performing: sleeping, studying, eating, exercising, socializing, etc. How you fill in and categorize your blocks of time should depend on what your questions of interest are.
3. Export your calendar data to `.ics` file format (a universal calendar format used by several email and calendar programs). This should take less than 5 minutes: https://support.google.com/calendar/answer/37111?hl=en.
4. Import the `.ics` file into R as a dataframe using the `ical` package (code to get you started is given below).
5. Create two visualizations and one table, wrangling the data as necessary along the way. (Note that we will use class time on Thursday, Sept. 10 to work on wrangling our calendar data; to take full advantage of the class time, it would be helpful if you attempted some wrangling prior to that class.)
6. What insights can you glean from how you're spending your time? Write a report introducing your questions of interest, explaining what you found, and reflecting on the answers to your questions posed. Your report should follow the template provided.
7. Finally, write a one-page reflection on this experience. Particular questions to reflect on are included in the template provided.

You will inevitably encounter different types of problems during data collection (Step 2) and data analysis (Step 5). I recommend you iterate through these steps in their entirety **early and often**. That way, you can identify and address any problems early on.

Tip: The color coding of events is lost when exporting the data into the `.ics` file, so don't rely on color-coding your calendar to give you information for this assignment.

Getting started . . .

The code below can be used to import the `.ics` file into R as a dataframe. Be sure to include the relevant path (where your `.ics` file is located) and update to your file name (with your email address).

```
library(tidyverse)
library(lubridate)
library(ical)

my_calendar <- ical_parse_df(file = "INSERT_PATH_HERE/your_user_name@amherst.edu.ics") %>%
  mutate(start_datetime = with_tz(start, tzone = "America/New_York")
    , end_datetime = with_tz(end, tzone = "America/New_York")
    , length_sec = end_datetime - start_datetime
    , date = floor_date(start_datetime, unit = "day"))
```

References

Thank you to Albert Kim (Smith College) and Johanna Hardin (Pomona College) for the Google Calendar assignment. They credit Roger Peng's and Hilary Parker's **Not So Standard Deviations** podcast titled "Compromised Shoe Situation" (http://nssdeviations.com/size/5/?search=shoe), in which they discuss a data science design challenge on getting to work on time, for the inspiration.