

# Diary of a Data Science Student: Cabin Fever During COVID-19

Stat231: Google Calendar Project

YOUR NAME HERE

Due Tuesday, May 5 by 5:00 PM

## Overview

This project will be undertaken individually. While you can discuss general ideas with your peers, please do not share code directly with each other. You can use other resources (including course materials, the internet, me, ...).

## How do I spend my time?

One challenge of remote learning that many students have voiced is the difficulty of structuring our time in the absence of the usual structures in place when on campus.

How *are* we spending our time now? In this project, we'll track how we're spending our time using Google Calendar, then import the calendar data into R to analyze and provide insights into how we're spending our time.

The project 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. (One tip: create a “weekly plan” at the beginning of each week, and document any changes as the week goes on. See the “Video: Google Calendar Tips” on our Moodle course page.)

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
- If you already use Google Calendar as a way to keep track of your schedule, you could compare how your time was spent at the beginning of the semester (on campus) to how your time is being spent at the end of the semester (since remote learning/COVID-19 isolation)
- Feeling artistic? Feel free to go the *data art* route, and create visualizations using other mediums outside of R. For examples: Dear Data (<http://www.dear-data.com/theproject>) and Environmental Graphiti (<http://www.environmentalgraphiti.org/series>).

## Details

### What to do

1. Identify 2-4 primary questions of interest to you about how you spend your time.
2. Track your time in your Google Calendar application for approximately 7-10 days (a longer amount of time would be ideal, but we're pressed for time here!). 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 relevant summaries and visualizations, wrangling the data as necessary along the way. (Note: MUCH, MUCH LESS wrangling is expected here than in the Clearinghouse Challenge. The `ical` package imports the calendar data into a nicely formatted dataframe, with variables for summary of the event, time start, time end, and description.) Wrangling should *not* take a substantial amount of time.
6. What insights can you glean from how you're spending your time? Write a 3-5 paragraph report introducing your questions of interest, explaining what you found, and reflecting on the answers to your questions posed.
7. Finally, write a short reflection (1-2 paragraphs) on the process. As someone who provides data: What expectations do you have when you give your data (e.g. to Facebook, Google, MapMyRun, etc.)? As someone who analyzes others' data: What legal and ethical responsibilities do you have?

## Tips

- Avoid creating calendar entries with a “repeat” schedule. Unfortunately, calendar entries based on “repeat” schedules only appear once (the first occurrence) in the R dataframe. (Or, if you do use repeat events – you would need to write an algorithm to code the recurring events.)
- The color coding of events is also lost when exporting the data into the `.ics` file, so don't rely on color-coding your calendar to give you information.
- There is a `googlecalendar` package under development that connects with Google's Calendar API that you may come across while searching online. Although you are free to explore that API if you want, it is not necessary nor expected that you utilize it for this project. The steps for importing your Google Calendar data into R are given below. (I have not explored the package or API myself; it could be a dead end.)

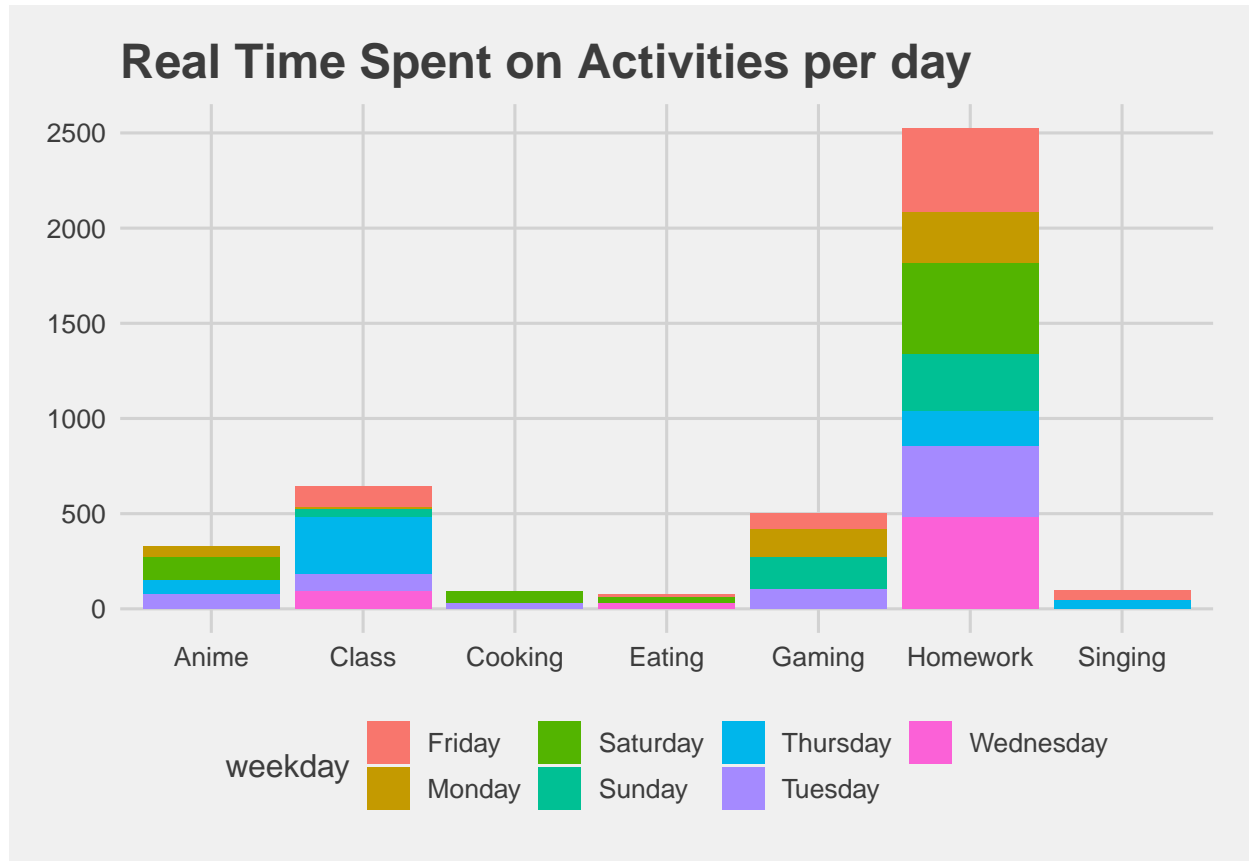
## Getting started . . .

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).

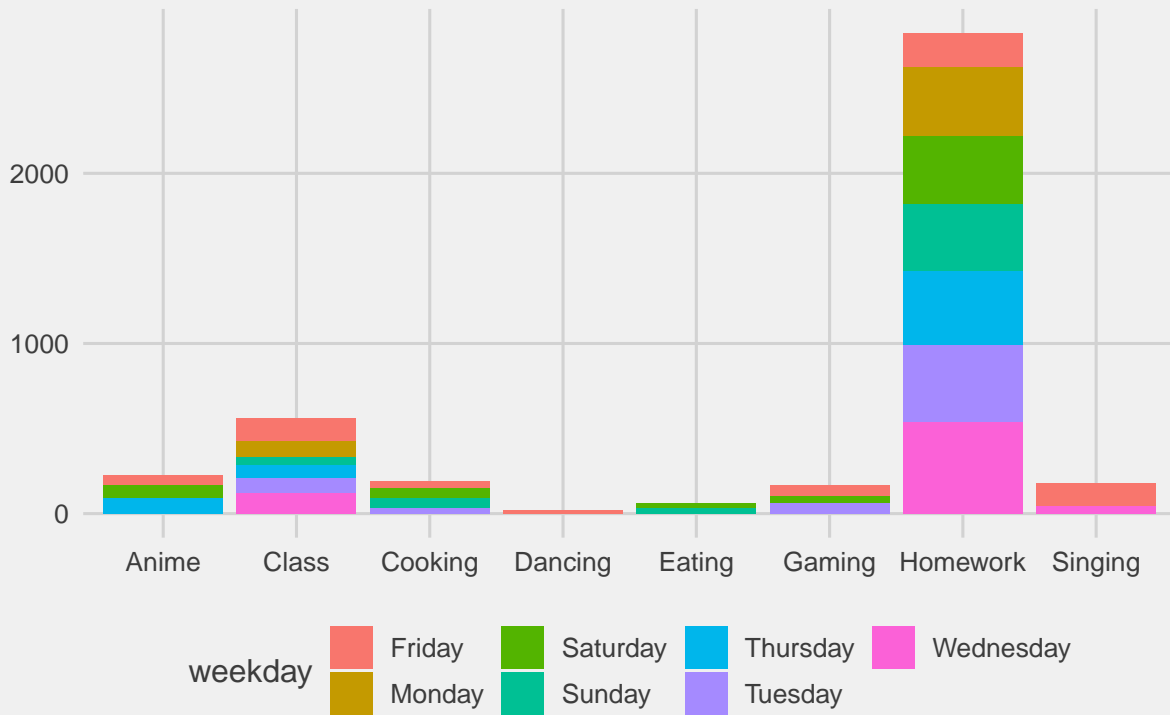
## Diary of a Data Science Student: Cabin Fever During COVID-19

For your final submission, please delete the above instructions, and include all of your (reproducible) code and written responses below.

Then, intersperse your written report and any summary tables or visualizations here.



## Intended Time Spent on Activities per day



# Homework

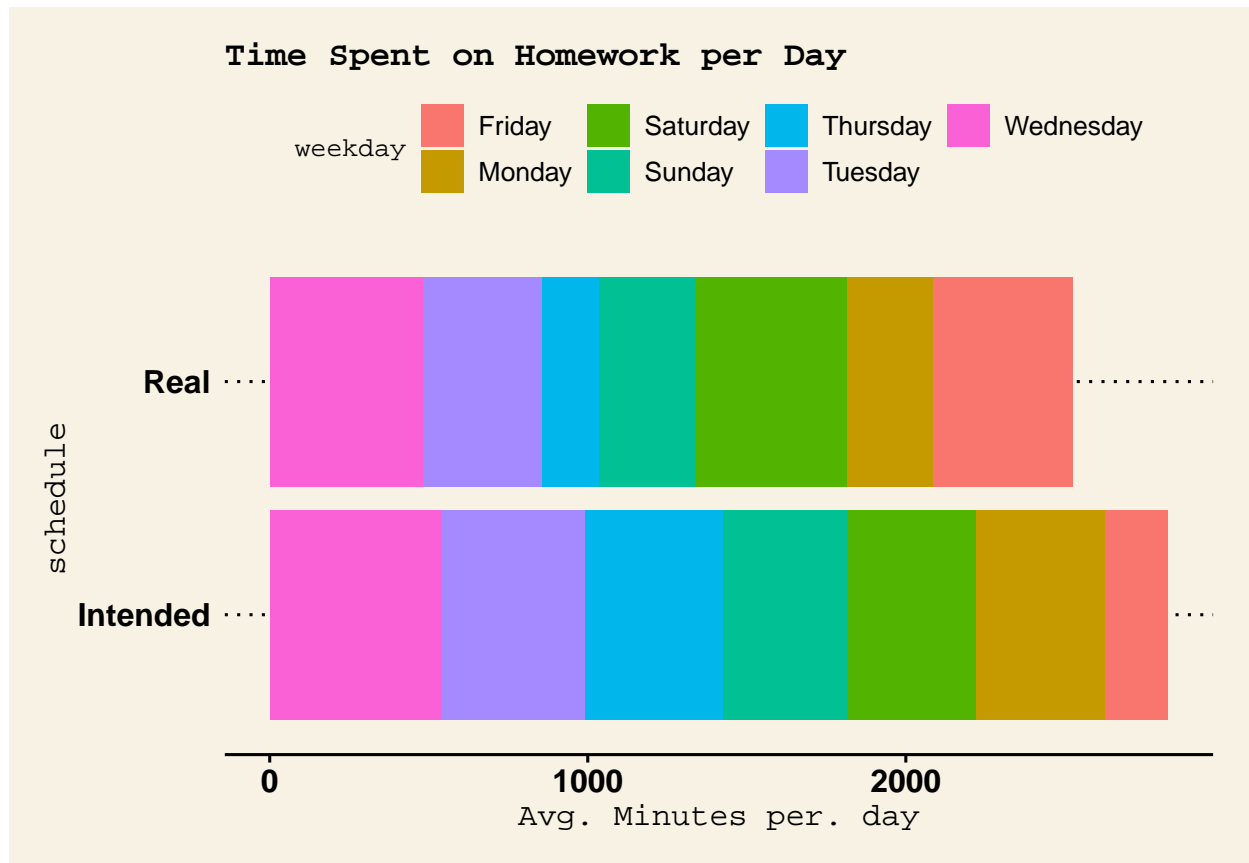


Table 1: Time Doing Homework by Schedule per. Day

Weekday	Real Avg Minutes/day	Intended Avg Minutes/day
Friday	440	200
Monday	270	405
Saturday	480	405
Sunday	300	390
Thursday	180	435
Tuesday	375	450
Wednesday	480	540

## Multi-tasking

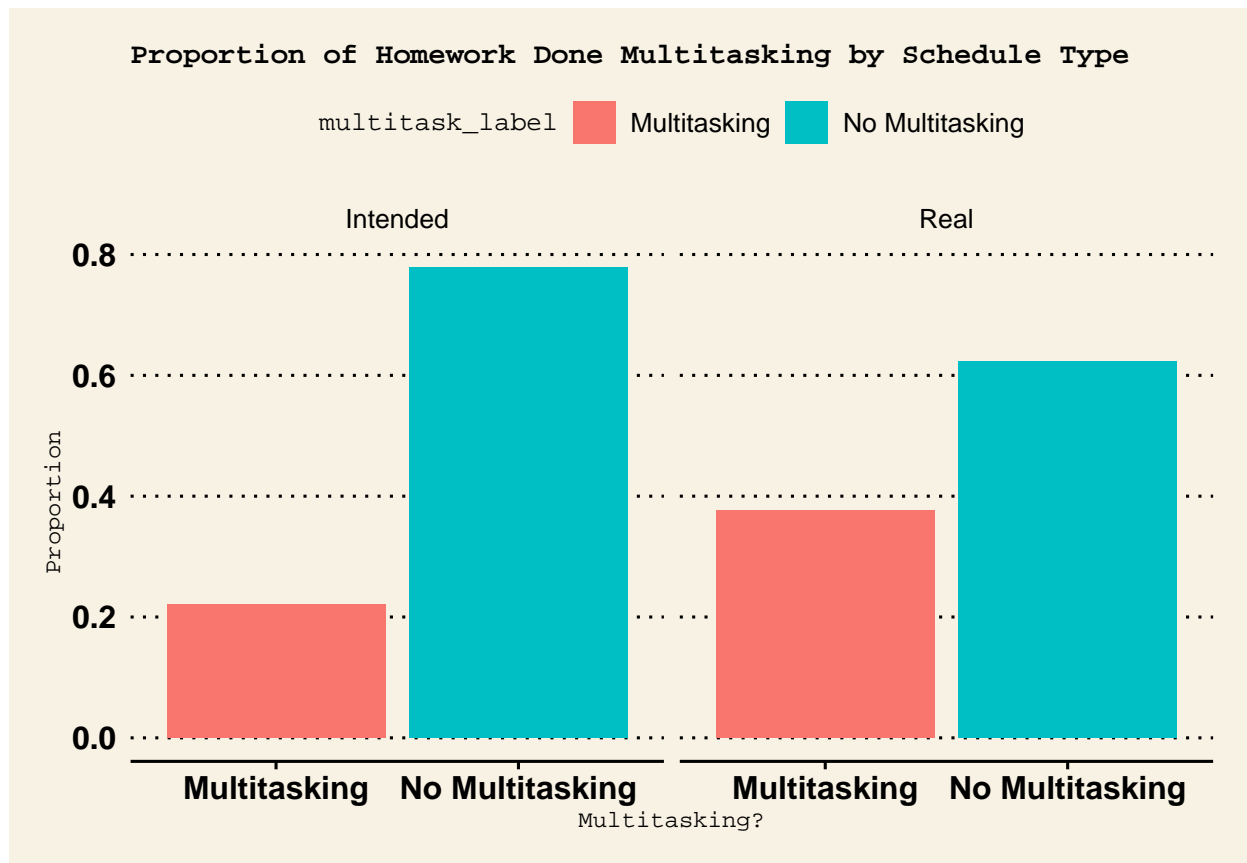


Table 2: Doing Homework with and w/o Multitasking

	Real Proportion	Intended Proportion
No Multitasking	0.6229508	0.7794872
Multitasking	0.3770492	0.2205128

## Hobbies

