

PSYCH 10 Final Project · Fall 2026

Week 2 · M1 · Finding a claim and working in a project folder

Full name _____

SUNet ID (the part before @stanford.edu) _____

Section (e.g., TUES 9:30) _____

Please write legibly so the teaching team can log your attendance: if we can't read your name, we can't give you credit.

Please hand this sheet in to your TA at the end of section. A copy is on the course website (psych10.github.io) for you to work from at home.

What you'll turn in

milestones/M1-claim-and-setup.Rmd and the PDF it knits to, from a project folder you made from the course template. Submit both files on Gradescope (the link is on Canvas).

Steps

1. **Make your project.** Download the course template (a .zip file) from the course website, psych10.github.io, unzip it into your psych10 folder (the one you made at the install clinic), and rename the folder `project-<your SUNet ID>`. Open it by double-clicking the .Rproj file. Opening it tells R where your project folder is, so a path like `data/raw/...` works the same on every laptop.
2. **Check that it knits.** Open `hello.Rmd` and press Knit and choose Knit to PDF (knitting turns an .Rmd file into a PDF). The file loads the course packages and makes one plot. If it doesn't knit, come to section or ask on Slack before doing anything else.
3. **Your claim.** Find the screenshot of the claim you picked in Friday's tutorial. If you'd rather work on a different one, find a new claim in the news that uses a number to say something about a group of people. In `M1-claim-and-setup.Rmd`, answer these in a line or two each:
 - What does the claim say, and what number does it give? Where did you find it (a link, and the date)?
 - What was counted or measured?
 - Who was included?
 - Why this one? One sentence on why you want to check it.
4. **Find a dataset you could check it against.** Look for data that measures something like what the claim counted, for a similar population. Any dataset that comes close will do for now; next week your group picks the one it will actually use. Good places to look: the organization behind the claim, government statistics sites (the Census Bureau, the CDC), Our World in Data, Pew Research Center, and R packages that come with data. An AI can suggest datasets, but it also invents ones that don't exist, so open every link yourself. Write down:
 - the dataset's name and a link;
 - who collected it;
 - one sentence on how you would use it to check the claim, and one thing it's missing.
5. **Load it, with the AI's help.** Download the file into `data/raw/` (or install the package), read it into R in your .Rmd, and print `dim()` and `str()` of it. The example below shows what this looks like. Ask an AI to help you with the code and with any error messages you get. Then write a few sentences in your own words: what one row of your dataset is (one person? one state in one year?), and what two of its columns measure, in what units. *If it won't load:* say what you tried and what went wrong, then load the example dataset below so the rest of the file still knits. You can still get full credit this way.
6. **Knit to PDF in a fresh session** (Session → Restart R, then Knit → Knit to PDF) and submit the .Rmd and the PDF.

An example

Say your claim was "Life expectancy around the world has gone up by about 20 years since the 1950s." The `gapminder` R package has life expectancy for 142 countries, every five years from 1952 to 2007, compiled by the Gapminder Foundation. You could check the claim against it, but notice two things: it stops in 2007, and it has one row per country per year, so an average across rows counts Iceland as much as India.

In the Console (the pane in RStudio where each line starts with >), once:

```
install.packages("gapminder")
```

In a chunk in your .Rmd:

```
library(gapminder)
dim(gapminder)
str(gapminder)
```

`dim()` says the data has 1704 rows and 6 columns. `str()` lists the columns: `country`, `continent`, `year`, `lifeExp`, `pop`, and `gdpPercap`. One row is one country in one year; `lifeExp` is life expectancy at birth, in years.

If your data is a file you downloaded instead, it goes in `data/raw/` and the chunk reads it from there with `here()`, which finds the file from the top of your project even though your `.Rmd` is in the `milestones/` folder:

```
library(here)
my_data <- read.csv(here("data", "raw", "my-file.csv"))
dim(my_data)
str(my_data)
```

Using AI on this milestone

Use the AI to find places to look for data, to install a package or read in a file, to explain the `str()` output, and to work through error messages. Do two things yourself: open the dataset's source to check that it's real, and write the answers about your claim and the sentences about the rows and columns in your own words. If the AI wrote a line you can't explain, ask it to explain the line, and log that.

How it's graded (1 point)

	Points
Project made from the template; <code>hello.Rmd</code> and the milestone knit	0.25
A claim about people that gives a number, with its source; what was counted and who was included are answered	0.25
A real dataset, linked, with who collected it and how you would use it to check the claim	0.25
The dataset loaded (or the attempt described and the example loaded); the sentences about the rows and columns are yours and correct	0.25

In section this week

Bring your laptop, your claim, and the dataset you found (or the ones you're choosing between). Your TA will help you get it loaded, and can help if your template won't open or knit. Next week's section starts by forming groups of three around related claims, so be ready to say what yours is about in a sentence.