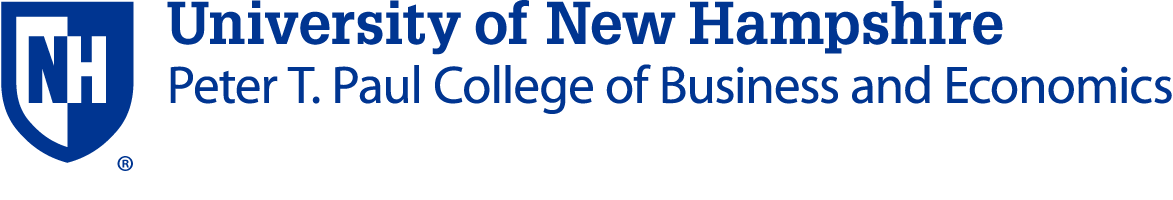
****

**LAB II - CAUSAL FORESTS REVISITED**



In November, 2019, after a Halloween night shooting at a “party house” in Orinda, California, Airbnb pushed a number of safety measures and rules.[[1]](#footnote-0) In this lab, we will quantify the causal impact of Airbnb’s safety measures on the booking rates. You will develop causal forest models to estimate the causal impact. Like you did in Lab I, you need to code your treatment: This time, you need a different binary variable: *Code Airbnbs in and after Nov. 2019 as 1; and otherwise 0.*

You can access the in-class lab on Posit Cloud. You can also find the Airbnb dataset inside the data folder. Do not blindly copy and paste the code. Check out [the documentation](https://grf-labs.github.io/grf/reference/causal_forest.html) and extensive [tutorials](https://grf-labs.github.io/grf/articles/grf_guide.html) to understand what you are doing. Add a visualization of your causal forest if you like. Be curious and explore. This is a *semi-*open ended assignment and you are the learner. You are also the solver of a problem and answerer of a question nobody on earth knows the answer to.

### On popular request, some instructions:

* The panel data here is a combination of time (month) and Airbnb ID
* If you will use a subset of your data, set the seed to 3.14159 and sample by group
  + You can’t keep some months but not the other months for the same Airbnb
    - *An Airbnb goes into your subsample or it doesn’t go into your subsample*
* In any causal forest you will build, set clusters = id in causal\_forest() function
* Do **not** split your data until you are asked to (So, no training vs. test sets to start with)
* Set the seed to 3.14159 in causal\_forest() and grow 5000 trees for your models

### Questions with even more instructions:

1. In a table, list the variables you decided to include in your models and explain why.
2. Estimate the conditional average causal effect on the booking rate (and standard error)
   * Use average\_treatment\_effect() to calculate the treatment effect
     + target.sample would depend on the effect you will be seeking
       - *You’ll have to set this to “overlap” if you receive a warning or NaN*
     + Set method argument to AIPW (Augmented Inverse-Propensity Weighting)
     + Use subset argument to answer questions on treatment heterogeneity
3. The causal effect you found above is a doubly robust estimate of the average treatment effect. But you did not split the data and set aside a completely isolated test set to start with. To confirm your results are robust, split your data and use predict() to calculate individual effects in the test set, and take an average of the individual effects to find the average treatment effect. Is the new effect consistent with the effect you found above?

***Hint:*** *Robustness checks almost never produce exactly the same results. Your hope here is to find results that are in the same direction and close in magnitude to your effect size.*

1. See the hints in Q2 and calculate the heterogeneity in the average treatment effect for:
   * Airbnbs with few or many bedrooms (Use bedrooms)
   * Airbnbs that accommodate more or less people (Use accommodates)
   * Low or high minimum nights requirements (Use minimum\_nights)
   * Higher or lower rated Airbnbs (Use review\_scores\_rating)
   * Whether Airbnbs are the entire house or not (Use room\_type)
   * For three states and cities in different parts of the country (Use state and city)
   * One source of heterogeneity you think is interesting to check (Dealer’s choice!)
2. In a single paragraph, interpret, explain, and discuss your findings. What do you think is your most interesting finding and why? If it is counterintuitive, how do you explain it?

*In addition, use bullet points at the end of your write-up to list your key takeaways and lessons learned from the analysis. Feel free to connect the dots with the previous lab.*

### Deliverables:

You will submit **two separate PDF documents on Canvas**:

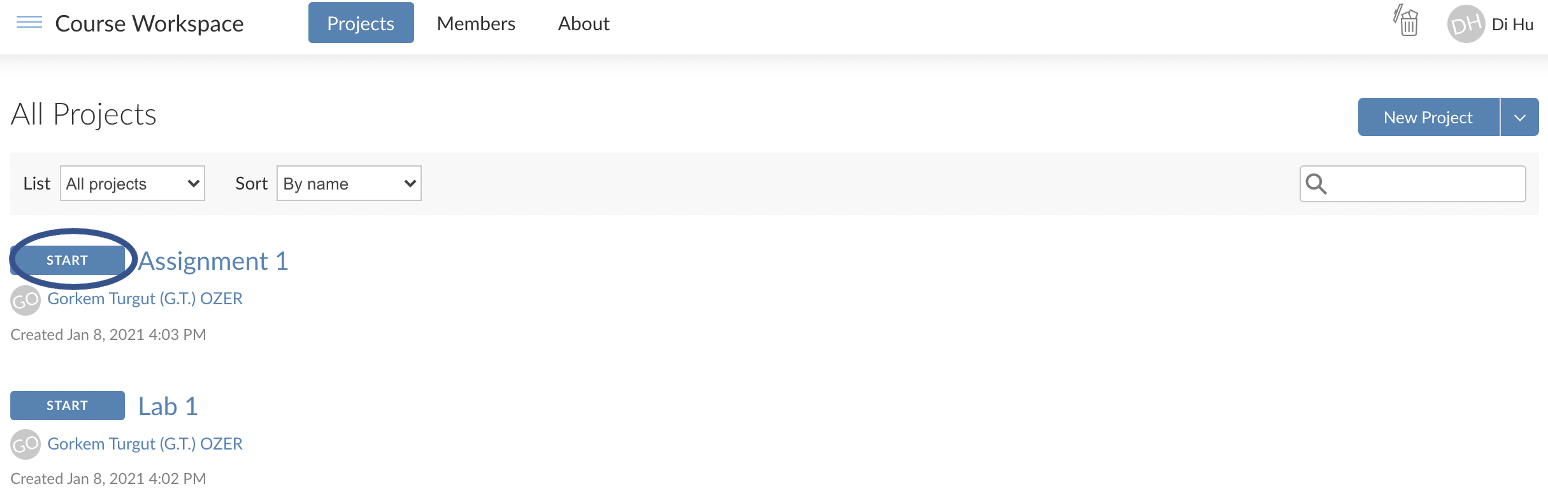
1. The first PDF will include your answers to the lab questions. You may want to create your own copy of the Google Doc, or download it as a MS Word document. Do whichever you like. When you are done answering the questions, please save it using the following naming convention: **LastName-FirstName-Answers.PDF** *-This is the main submission file*
2. The second PDF will be the code file you will generate in Posit Cloud as a PDF file when you are done working on the code. Please see the guidelines starting from the next page. Name it as as **LastName-FirstName-Code.PDF** *-This is the supplemental file*

**Step 0:**

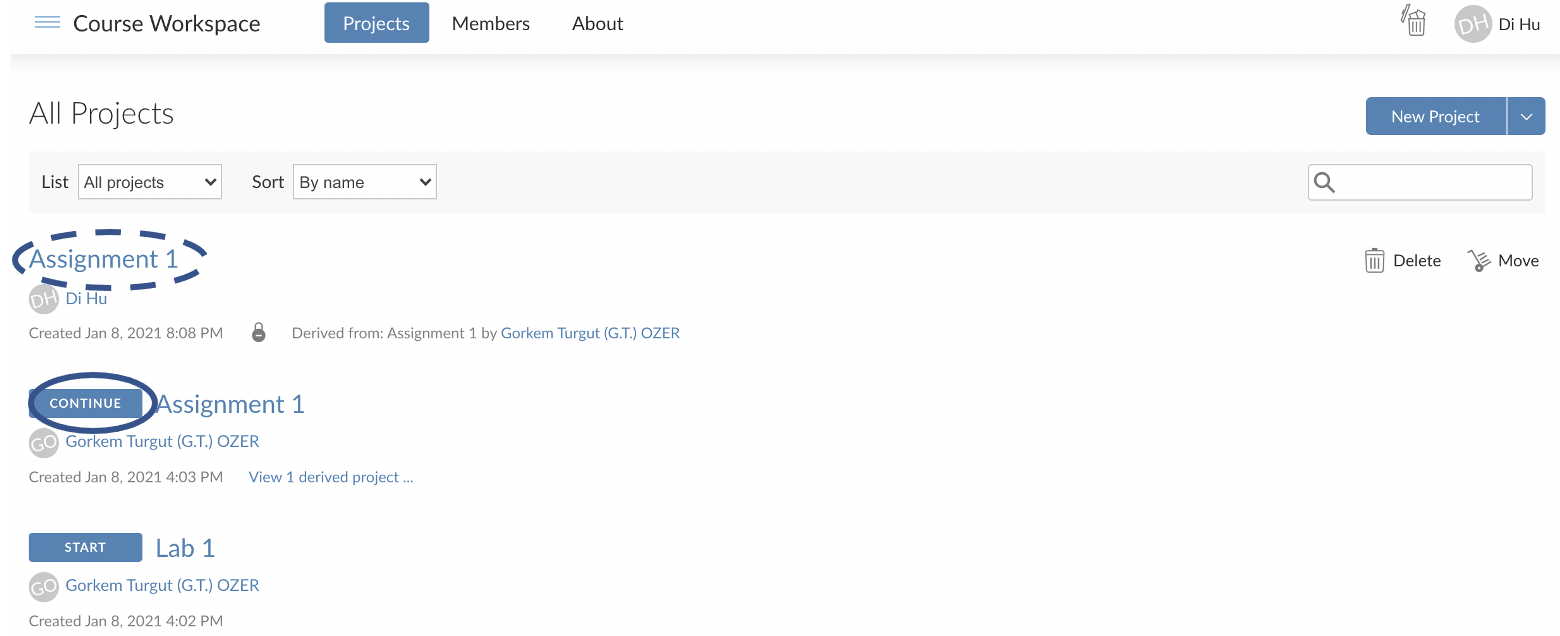
First, read [**How to use Posit Cloud**](https://docs.google.com/document/d/1CcAbFVQYwOJb5AiYusQ68_BX1WIPnz6-7N1_dWYs44M/edit?usp=sharing) in its entirety.

**Step 1:** In Posit Cloud, either start with the lab (where 10k Airbnb data is uploaded) or use your project. Start with clicking the ‘Start’ button to make your own derived copy (or to continue your work on your copy by clicking the ‘Continue’ Button). The assignment will open in an IDE.

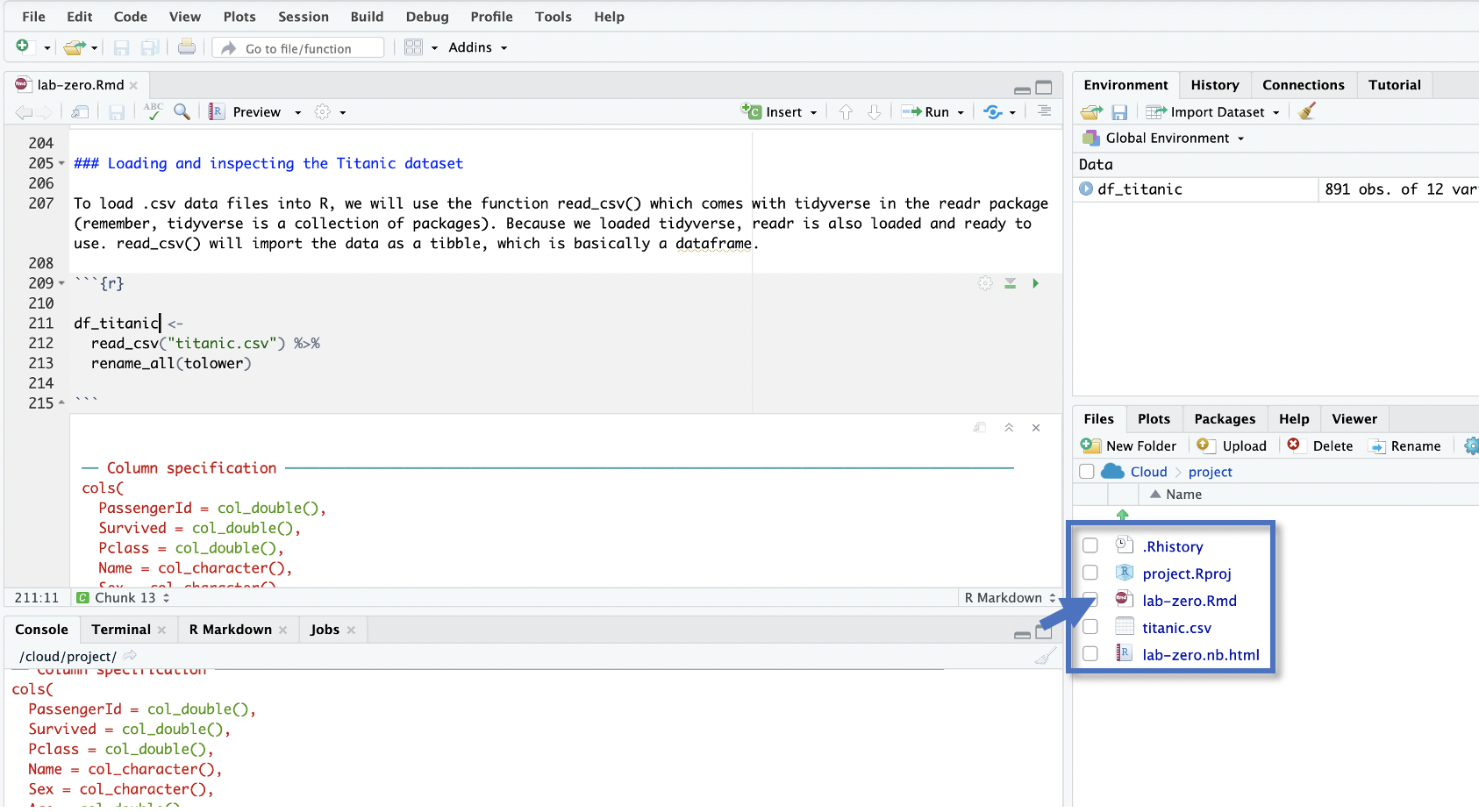
Start working on a lab:



If you already started once, continue working on a lab:



**Step 2:** When the lab opens, look to the right bottom of the window, and you will see several files and folders there. The name of the file will correspond to the name of the lab.



**Step 3:** As you know, you’ll enter your R commands between special markers called “[chunks](https://rmarkdown.rstudio.com/lesson-3.html).” Labs may come with some comments and an empty chunk to start with. A good practice is to use a different chunk for each question (or subquestion in a long question).

```{r}

*Your R commands here*

```

**Step 4:** When you are done answering the questions, style your code using **StyleR**. Install styler and style your code file following the example [here](https://www.tidyverse.org/blog/2017/12/styler-1.0.0/).

**Step 5:** After you fix any styling issues you may have, follow the instructions on the next page to create a **PDF** of your R Notebook with the naming convention **LastName-FirstName-Code.PDF**

**Step X:** Submit the two PDFs you now have through Canvas under the respective assignment:

1. LastName\_FirstName-**Answers**.PDF (will include your answers to the questions)
2. LastName\_FirstName-**Code**.PDF (will include your report including your code)

**How to create a PDF of your code in an R Notebook**

~~0. First, run the following two lines of code in your R Console to install the necessary components:~~

~~tinytex::install\_tinytex() => This should not be needed. Try it if you have trouble.~~

1. To create a PDF (or Word) version of your work from an R Notebook, you need to have the following lines of code at the top of your R notebook. You can copy and paste from below and change the title. When you are using a provided lab/assignment template, these lines should already be in the file.

---

title: "Your Title"

output:

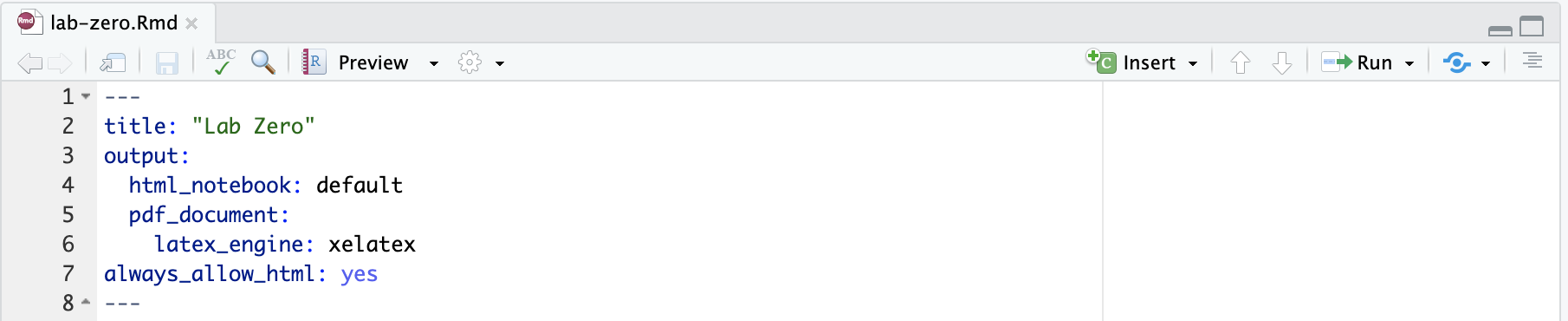
html\_notebook: default

pdf\_document:

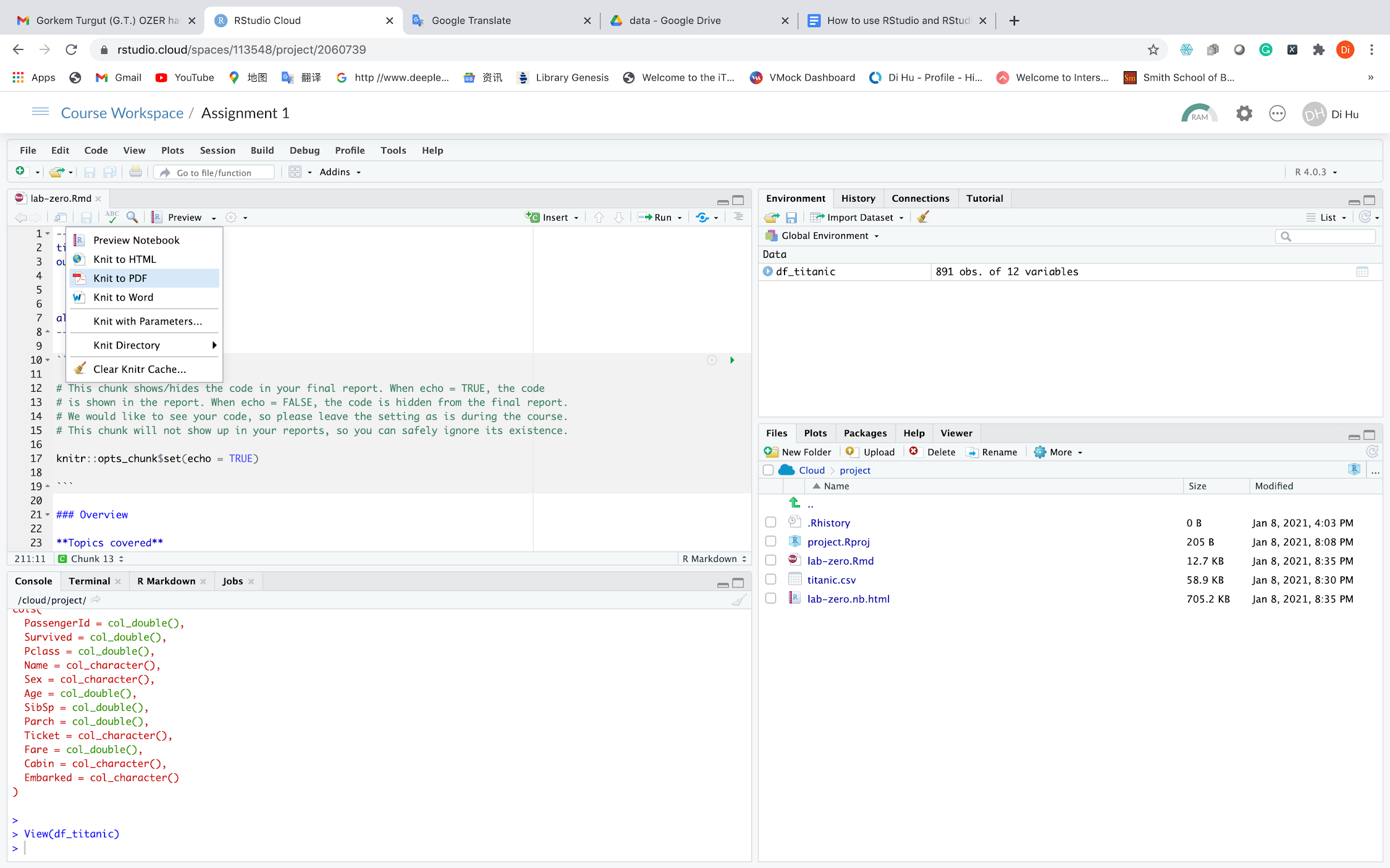
latex\_engine: xelatex

always\_allow\_html: yes

---



2. Save your file after adding these lines. Then, click the down arrow next to the “Preview.” This will show the options as shown below. Choose the PDF (or PDF) option to create a report of your work in seconds:

****

3. Here you go:

- If you choose the PDF option, your PDF report will open in a new window automatically for you to save.

- If you choose the Word option (not needed for the course), you will see a prompt to download the file.

1. <https://www.cnbc.com/2019/11/06/airbnb-is-pushing-major-safety-changes.html> [↑](#footnote-ref-0)