

Problem Set 1

POS 5737

Due **September 08, 2025** before class

You have to submit this problem set electronically. See final problem set submission section below about what you have to submit via an email.

Task 1 (2 point)

In this task you need to install LaTeX on your local machine. LaTeX is a software system for typesetting documents (for example this problem set has been created in LaTeX).

1. For Windows OS, you can use MikTeX installer - more information [here](#)
2. For Mac OS, you can use MacTeX installer - more information [here](#)
3. After you successfully installed the LaTeX core, you need to install an editor to actually type LaTeX files. For LaTeX/TeX editor, you can use TexShop or TexStudio (I am sure there are more options on the market now).

I use a more a general editor (in a sense that it can be used to write any types of text and code, not just .tex files) - **Sublime** (more information [here](#)). If you do choose to work with Sublime, you have to install a separate pdf previewer.

- For Mac OS the most common is **Skim**. The full process of integrating Sublime/Skim with LaTeX is nicely outlined [here](#) (the article is a bit old, but it looks like the steps outlined in it are still valid).
- For Windows users, I found that you can integrate Sublime with Sumatra. See the details [here](#). Maybe you will be able to find even better option.

You need LaTeX to be installed on your local machine so that you were able to compile R Markdown codes. I would know that you successfully installed LaTeX if you would be able to compile R Markdown document in R in the next steps of the problem set.

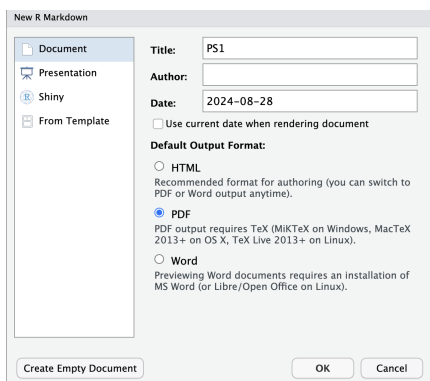
Extra Task: Create an account on Overleaf ([here](#)), which allows you to use LaTeX virtually, without installing it to your local machine. You will need to have access to the internet to work with Overleaf. Overleaf also allows you to collaborate with multiple people while working on the same document (similar to what you experience with the Google Doc). You can get an extended version of overleaf subscription through FSU. Read more information [here](#).

Optional Task: Explore how to create L^AT_EX documents in this [short guide](#)

Task 2 (2 point)

In this task you need to create a new R Markdown document. To do so you need to start R Studio and proceed with the following steps:

1. File → New File → R Markdown
2. You will see the following pop up window:



3. Choose a title of your choice; choose PDF as the Output format; click OK
4. You have an R Markdown document created now
5. Save .Rmd document with a proper name in your working folder

Now you created an R Markdown document, it is time to start working in it. R Markdown allows to generate text/pdf documents with incorporated chunks of your R code and the results of running that R code.

6. Explore how to R Markdown works [here](#)
7. You can notice that R Markdown resembles a regular R code but is structured around **chunks** that are marked in the following way:

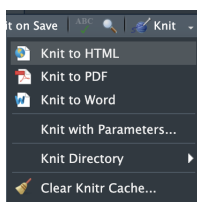


You can add more chunks using Code → Insert Chunk

You always write your R code within the chunks and any text that you want to include in your document outside of the chunks.

To troubleshoot each individual R code chunks you can push a little green button **Run** for that specific chunk. You can also run individual lines of code within chunks, just like you would do it in your regular R script.

8. To compile the final document you have to use **Knit** button at the top menu of the R Markdown script. (Remember at the beginning we requested to knit the final document to PDF and for that precisely you need to have all the LaTeX tools installed on your local machine. However, Knit button has a little arrow below that opens options (see the illustration) and there you can choose to Knit your document to HTML or MS Word as well.)



Task 3 (5 point)

Now I ask you to work with some data wrangling and conduct some data analysis, but you will have to write your code and all the comments and interpretations to the code within the R Markdown document. One of the main tasks is to be able to successfully compile a final PDF(!) document that will include your R codes and all the necessary comments and descriptions.

1. Download dataset *kdrama.csv* from the email that I sent you. This is a dataset with top-ranked Korean Dramas as per the MyDramaList website downloaded from [Kaggle](#). How many Korean Dramas are included in this list? Write an R code to respond to this question.
2. Use an R code to return a list of variables that are included in this dataset.
3. Use an R code to identify what the mean value of total number of episodes for all the kdramas in the dataset is.
4. Plot a histogram of the shows rating.
5. How many shows have a rating higher than 9 points? Show how you calculated that in R.
6. Rename variable **Year.of.release** to simply **Year** without creating a new variable.
7. How many shows in this dataset were released in 2020-2022? Demonstrate with an R code.
8. What type does the variable **Duration** have? Show with an R code.
9. Recode variable **Duration** so that it would be a numerical variable measuring duration in minutes. Plot the histogram of the recoded variable.
10. Create a new dataset that will include shows with **Original.Network** being Netflix.
11. What is the average rating score for the shows that have Netflix as an Original Network.

Task 4 (1 point)

Finally, let's learn how to work with Github. Create an account at <https://github.com>. After you successfully created an account, follow the tab **Repositories** and create a new repository. You can name it as you wish. Please keep this repository public (you can later delete it). In the repository description specify that this is a practice repository for Methods II class. Upload your knitted PDF document with all the calculations and descriptions for Task 3 to the that newly created repository.

Note: You can read more about how to work with github [here](#).

Final Problem Set Submission

As part of this problem set submissions, you have to send me **3 elements** via an email:

- .Rmd document with your code and solutions for Task 3
- PDF document that results from a knitted R Markdown document
- a link to your public Github repository where you uploaded your knitted PDF document.