

# TSA - Final Project Instructions

Luana Lima

Spring 2022

## CREATE A REPOSITORY IN YOUR GITHUB ACCOUNT (optional)

1. Go to your user account on GitHub and navigate to the repositories tab.
2. In the upper right corner, click the green “New” button.
3. Name your repository with recommended naming conventions (suggestion: *LastName1LastName2LastName3\_ENV790\_TSA*). Write a short description of the purpose of the repository. Check the box to initialize the repository with a README. Add a .gitignore for R and add a GNU General Public License v3.0.
4. Invite other group members as collaborators to the repository.

## LINK YOUR REPO TO YOUR LOCAL DRIVE WITH RSTUDIO (optional)

1. Click the “Clone or download” button for your repository and then the “copy” icon. Make sure the box header lists “Clone with HTTPS” rather than “Clone with SSH.” If not, click the “Use HTTPS” button and then copy the link.
2. Launch RStudio and select “New Project” from the File menu. Choose “Version Control” and “Git.”
3. Paste the repository URL and give your repository a name and a file path.

## CHOOSE A DATASET AND A RESEARCH QUESTION

1. Choose a dataset of interest.
2. Describe what you want to do with the dataset and why.

## COMPLETE YOUR PROJECT REPORT

### General Guidelines

1. Use Rmd for your final report
2. Write in scientific style, not narrative style
3. Global options for R chunks should be set so that only relevant output is displayed. Turn on/off messages and warnings when applicable to avoid unnecessary outputs on the pdf.
4. Make sure your final knitted PDF looks professional. Format tables, size figures, chapters, etc.
5. Make sure the PDF file has the file name “LastName\_ENV790\_Project.pdf” and submit it to the dropbox in Sakai. You will only submit your PDF file.

### Contents of the Report

**Introduction, Motivation, Relevance, Objectives** Write a few paragraphs detailing the rationale for your study. This should include both the context of the topic as well as a rationale for your choice of dataset

(reason for location, variables, etc.). You may choose to include citations if you like or any other reference you may have used during the project (optional).

**Dataset information** Provide information on how the dataset for this analysis were collected (source), the data contained in the dataset (format). Describe how you wrangled/processed your dataset to get the time series object.

Add a table that summarizes your data structure (variables, units, ranges and/or central tendencies, data source if multiple are used, etc.). This table should inserted as a `kable` function in an R chunk. Just show the firts 10 rows of your data. Do not include the code used to generate your table.

**Analysis (Methods and Models)** Describe the analysis and tests that were performed. Described the components of the time series you identified. List any packages and functions used. Include visualizations of your dataset (i.e. time series plot, ACF, PACF, etc).

Format your R chunks so that graphs are displayed but code is not displayed. Accompany these graphs with text sections that describe the visualizations and provide context for further analyses.

Each figure should be accompanied by a caption, and referenced within the text if applicable.

**Summary and Conclusions** Summarize your major findings from your analyses in a few paragraphs and plots. What conclusions do you draw from your findings? Any insights on how to improve the model?