
About Term Project

Prepare a report with RMarkdown in RStudio by selecting a data set (any data set).

- Term project will be the analysis of a data set you choose.
- The data set can be an existing data or you can collect data by conducting a survey. If you choose to conduct your own survey please make sure that you ask the appropriate questions to collect different types of data.
- You can choose the data according to your interests or you can use the data you used in other courses.
- The purpose of the term project is to apply the techniques we learned in the class.
- The aim of the project is not to perform a detailed analysis. Therefore, you do not need to calculate every statistics or plot every possible graph for each variable.
- The aim of the project should be to ask meaningful questions about the data set you have and to provide meaningful answers to these questions with data analysis.
- You do not need to apply all the statistical methods we learned in the class. If you want, you can also use the methods that we did not cover.
- It is also useful to discuss the shortcomings of the methods you use and how one can improve the analysis.
- You should discuss the reliability and validity issues related to your data and the appropriateness of the statistical analysis to this data.
- Do not visualize all of your data with a single graph. You need to create various graphs with your data.
- Do not forget that a single well-prepared graph with attention to detail is better than a lot of poorly prepared graphs.
- It doesn't matter which additional packages you use, but you must use the **tidyverse** packages.
- Pay attention to spelling rules as much as possible. The Elements of Style is a great book and I highly recommend you to obtain a copy of this book. This will be a great reference guide in your professional life.
- All analysis must be done in RStudio.

Data

- To be successful in the project, it is important that the data set is relatively small but the data is easily accessible and large enough to explore multiple relationships.
- **It is recommended that you have at least 50 observations and 10 to 20 variables in your data set.**
- **It is recommended that the data set includes categorical variables, discrete numeric variables and continuous numeric variables.**
- Do not reuse the data sets used in the examples or assignments in the class.

- If you are having problems, seek help before it is too late.
- Below you can find links to some sites that may be of interest to you:

AEA Resources	İBB Açık Veri Portalı	TOBB Sanayi Veritabanı
Awesome Public Datasets	-	TUIK
CRAN Official Statistics Task View	OECD	US Census
Eurostat	OpenIntro datasets	USDA ARMS
FAO	Penn World Tables	USDA NASS
FRED	PRISM Data Archive Project	World Bank
Harvard Dataverse	SBB	WTO
ILO	TCMB EVDS	
IMF	thelittledataset.com	
IPUMS	TidyTuesday	

- I will not accept any data you find on Kaggle datasets

Important Dates

Project Proposal and Final Project submission will be done by uploading a zip file to the ekampus system along with the Github repo link. If you do not upload a zip file to the system and do not provide a Github repo link, you will be deemed not to have entered the midterm and final exams.

Your Make-Up Exam will be in the form of both a Project Proposal (Project) submission and an additional Exam.

1. **Proposal Due:** TBA
2. **Project Due:** TBA

Project Proposal

- You must upload your project folder (YourStudentID.zip file) to *ekampus.ankara.edu.tr* until TBA.
- You are expected to add the data set you plan to use into the `/data/` folder.
- You are expected to add the bibliographic information of the literature you plan to use into the `biblio.bib` file.
- In your project proposal, you are expected to explain your data analysis plan and data set.
- Your project proposal should be the introduction section of your project. You can continue to make changes to the introduction section until the final submission of the project, ie the end of the term.
- **Section 1 - Introduction:** In this part, you are expected to explain the following points.
 - Your research question and where & how you found the data set,
 - Observations, variables, etc. included in the data.

- Literature relevant to your research question. Briefly discuss the articles (the word article refers to **academic** articles, not blog posts, not reddit posts, not newspaper articles but **academic** articles that are published in academic journals) you read by giving reference. **Do not summarize each article individually under a separate title.** It is expected that you cite at least six articles in the literature review.
- Your project proposal **should not be more than 2 pages (excluding references).**

Grading

Data	30 points
Project Proposal	40 points
Resources, organization, code quality, spelling	30 points
Total	100 points

Project

- **You must upload the final version of your project (YourStudentID.zip file) to *ekampus.ankara.edu.tr* until TBA.**
- In the final form of your project, you are expected to finish the analysis that you promised in your proposal.
 - **Section 2 - Data:** Summarize the content of your data with a summary statistics table and provide information about the variables included in the data.
 - **Section 3 - Methods and Data Analysis:**
 - * Provide information about the outcome (dependent, Y) variable and predictive (explanatory, X) variables that you will use to answer your research question.
 - * Provide information about the comparison groups you will use, if any.
 - * Perform a preliminary data analysis, including some summary statistics and visualizations (graphs), as well as some explanations to help you learn more about your data.
 - * Provide information about the method(s) that you believe will be useful in answering your questions.
 - * Provide information about the results of the methods you used (eg regression results if you have used regression analysis) and add the necessary tables.
 - **Section 4 - Conclusion:** Summarize the results of your analysis. Discuss to what extent your results responded to the research question you identified at the beginning and how this work could be improved in the future.
- You are expected to write a one-page summary of your final project. This summary should provide information about the data set you used, your research questions, methodology, and findings.
- The summary is worth 10 points, it will be evaluated according to whether it follows the guidance given above and whether it is concise but detailed.
- **Each section should not be more than 2 pages (excluding references, tables and figures).**
- **All results presented in the project must be derived with the R code.** The answers/results given without the R code will be considered incomplete. If you perform calculations manually instead of using R and then report results from those calculations, you will not be able to receive points for them. All code reported in your final project document should work properly. Please do not include any code that causes error messages or is not related to analysis.

Grading

Project	50 points
Summary	10 points
Reproducibility	20 points
Resources, organization, code quality, spelling	20 points
Total	100 points

Grading Criteria Your project will be evaluated according to the following criteria:

- *Content* - What is the quality of the research and/or policy question and how do the data relate to those questions?
- *Accuracy* - Are statistical procedures correctly applied and explained?
- *Writing and Presentation* - What is the quality of statistical presentation, writing and explanations?
- *Creativity and Critical Thinking* - Has the project been carefully thought out? Are the limitations carefully considered? Does it appear that time and effort went into the planning and implementation of the project?

Folder Organization

Your project folder should have the following folders and files:

- `/bibliography/`: Bibliography Folder
- `/bibliography/biblio.bib`: Bibliography File
- `/csl/`: Bibliography Style Folder
- `/data/`: The Folder Containing the Data Set (CSV, RDS, etc.)
- `/proposal/`: Project Proposal Folder
- `/proposal/proposal.Rmd` + `/proposal/proposal.pdf`: Project Proposal Text
- `README.Rmd` + `README.md`: README Files
- `proposal.Rproj`: RStudio Proposal File

Since style and form are important in this project, please make sure everything looks good and that your data and code are correctly formatted.