R for Research Capstone Project

EU StudyAssist

Project Overview

Introduction

Students will apply the skills learned throughout the bootcamp to analyze a dataset from any of the two projects provided below. They will use R to understand the data of the project they have choosen, perform exploratory data analysis (EDA), inferential statistics, and data visualization.

Project Format and Submission

The final output will include a final report written using Quarto and rendered to pdf and html. A sample report is available here. The assignment is to be submitted on your Github account in a new repository called **r-for-research-capstone-project**. Add xrander as project collaborator to ensure you get the best guide from instructors.

The Data

Project I: CO_2 Emission from Cereal Food Production

Figure 1: Rice farmers- source: linkedin.com

This data is an extract from Ritchie (2021). It presents the emission from converting forest land to rice, wheat, and other cereal products for different countries. The data for this project is available here here.

This data contains the following columns:

Variable	**Meaning**
entity	Country name
code	The country code
products	The cereal food
emission	Total CO_2 emission in tonnes
per_capita_emission	Deforestation emission per person

Objectives:

- Import your data
- convert emission from tonnes to kilotonne.
- Conduct exploratory data analysis (EDA).
- Visualize data using appropriate plots.

- Perform inferential statistical analysis to see if there's a difference in emission between the products.
- Communicate findings in a report.

Project II: Cuckoo Eggs Dimension

Figure 2: Cuckoo Egg - source:theconversation.com

The data for this project is from the DAAG package by Maindonald, Braun, and Braun (2015). It presents measurement on 120 eggs lain in the nest of 6 different species of cuckoos. The data for this project is available here.

This data contains the following columns:

Variable	Meaning
length	The length of eggs in millimeters
breadth	The breadth of eggs in millimeters
species	Six species of birth species

Objectives

- Import your data
- Conduct exploratory data analysis (EDA).
- Visualize data using appropriate plots.
- Perform inferential statistical analysis to see if there's a difference in breadth of eggs laid by the bird species.
- Perform inferential statistical analysis to see if there's a difference in length of eggs laid by the bird species.
- Communicate findings in a report.

References

Maindonald, John H, W John Braun, and Maintainer W John Braun. 2015. "Package 'DAAG'." *Data Analysis and Graphics Data and Functions*.

Ritchie, Hannah. 2021. "Carbon Emissions from Deforestation: Are They Driven by Domestic Demand or International Trade?" *Our World in Data*.