

Team Name: The Back Right Corner

Introducing Naniar: How to put the “NA” values to use!

Member Duty:

Luke Syverud: Code in R, Clean Data, Presentation Creation

Elizabeth Ruiz: Research Package, Code in R, Presentation Creation

Maxwell Creager-Roberts: Presentation creation/Story-building

Kaia Lui: Research Package, Research valuable data sets, Presentation Creation

Isaac Kujak: Research Package, Research valuable data sets, Presentation Creation

Description of the Data Set

- We are going to use the [Income Predictor Dataset - US](#) to showcase how datasets that have important data, but contain some missing values can still be utilized to showcase trends without having to remove large sums of the dataset by entirely dropping NA values. This dataset contains 15 columns of data with over 30,000 rows of data. Thus, it is a large dataset to use, with some missing or meaningless values.

Goal of the Project

- The goal of this project is to introduce the package “naniar” to the class and show some of its functionality and implementation. The package allows us to plot “NA” values from datasets instead of throwing them out, while also giving us valuable information on where the missing data is located. We want to showcase that datasets that are important but have lots of missing values can still be utilized to show trends and still utilize the entirety of the dataset instead of removing large parts of the data. Also, if one variable has a few missing values, this package allows us to use that variable instead of removing it entirely by giving estimates to the “NA” values.