Data Frames in R: Analyzing the Monty Hall Problem

Preparation

The Monty Hall Problem is a famous probability puzzle that challenges our intuition about statistics and decision-making. Named after Monty Hall, the host of the American television game show "Let's Make a Deal," this problem presents a scenario where a contestant must choose between three doors. Behind one door is a prize, while the other two conceal goats. After the contestant makes an initial choice, the host, who knows what's behind the doors, opens another door to reveal a goat. The contestant is then given the option to stick with their original choice or switch to the remaining unopened door. The question is: which is the best strategy?

Before we dive into working with R, let's gain an illustrative understanding of the problem. Watch the video below from the 1994 German television show "Geh aufs Ganze," where a contestant faces this very dilemma.

- Video "Geh aufs Ganze" (1994)

To learn about data frames in R, watch the lecture on representing data from Harvard's CS50 Introduction to Programming with R, as well as the short video on data frames:

- Video CS50R - Representing Data
- Video CS50R - Short on Data Frames

Now you should be prepared to solve the following exercises using R and RStudio.

Exercise 1: Data as tables

Daat frames represent data in a tabular format, similar to a spreadsheet. Each row corresponds to an observation, while each column represents a variable. Data frames are a fundamental data structure in R, allowing you to store, manipulate, and analyze data efficiently. Let's revisit the Monty Hall Problem using data frames.

1. Load the historical data from the game show “Let’s Make a Deal” into a data frame in R.
2. …