

About best function and ranking

Sowjanya

01/03/2020

R Markdown

The data for this assignment come from the Hospital Compare web site (link run by the U.S. Department of Health and Human Services). The purpose of the web site is to provide data and information about the quality of care at over 4,000 Medicare-certified hospitals in the U.S. This dataset essentially covers all major U.S. hospitals. This dataset is used for a variety of purposes, including determining whether hospitals should be fined for not providing high quality care to patients (see [link] (<http://goo.gl/jAXFX>)) for some background on this particular topic.

Finding the best hospital in a state: Write a function called `best` that take two arguments: the 2-character abbreviated name of a state and an outcome name. The function reads the `outcome-of-care-measures.csv` file and returns a character vector with the name of the hospital that has the best (i.e. lowest) 30-day mortality for the specified outcome in that state. The hospital name is the name provided in the `Hospital.Name` variable. The outcomes can be one of “heart attack”, “heart failure”, or “pneumonia”. Hospitals that do not have data on a particular outcome should be excluded from the set of hospitals when deciding the rankings. Handling ties. If there is a tie for the best hospital for a given outcome, then the hospital names should be sorted in alphabetical order and the first hospital in that set should be chosen (i.e. if hospitals “b”, “c”, and “f” are tied for best, then hospital “b” should be returned).

Ranking hospitals by outcome in a state: Write a function called `rankhospital` that takes three arguments: the 2-character abbreviated name of a state (`state`), an outcome (`outcome`), and the ranking of a hospital in that state for that outcome (`num`). The function reads the `outcome-of-care-measures.csv` file and returns a character vector with the name of the hospital that has the ranking specified by the `num` argument. For example, the call `rankhospital(“MD”, “heart failure”, 5)` would return a character vector containing the name of the hospital with the 5th lowest 30-day death rate for heart failure. The `num` argument can take values “best”, “worst”, or an integer indicating the ranking (smaller numbers are better). If the number given by `num` is larger than the number of hospitals in that state, then the function should return `NA`. Hospitals that do not have data on a particular outcome should be excluded from the set of hospitals when deciding the rankings.

Ranking hospitals in All states The function called `rankall` that takes two arguments: an outcome name (`outcome`) and a hospital ranking (`num`). The function reads the `outcome-of-care-measures.csv` file and returns a 2-column data frame containing the hospital in each state that has the ranking specified in `num`. For example the function call `rankall(“heart attack”, “best”)` would return a data frame containing the names of the hospitals that are the best in their respective states for 30-day heart attack death rates. The function should return a value for every state (some may be `NA`). The first column in the data frame is named `hospital`, which contains the hospital name, and the second column is named `state`, which contains the 2-character abbreviation for the state name. Hospitals that do not have data on a particular outcome should be excluded from the set of hospitals when deciding the rankings. Handling ties. The `rankall` function should handle ties in the 30-day mortality rates in the same way that the `rankhospital` function handles ties.