R_Hospital_Quality_Assignment3

Overview

This project was completed as part of **Programming Assignment 3** from the **R Programming course** by Johns Hopkins University on Coursera.
It uses data from the **Hospital Compare** website to analyze **30-day mortality rates** for three conditions:
- Heart attack
- Heart failure
- Pneumonia

The goal is to write R functions to find and rank hospitals by outcome and create a histogram of heart attack mortality rates.

Dataset

All data files are located in the `data/` folder:
- `outcome-of-care-measures.csv`: 30-day mortality and readmission rates.
- `hospital-data.csv`: Hospital information.
- `Hospital_Revised_Flatfiles.pdf`: Codebook with variable descriptions.

> ⚠️Note: The original assignment zip file from Coursera was unzipped into the `data/` folder.

Setup

1. Clone the repository.
2. Set your R working directory to the `data/` folder where the CSV files are located:

```
# Example path on local machine
setwd("E:/coursera_r/assignment3/rprog_data_ProgAssignment3-data")
```

3. Ensure all CSV files and the PDF are present in `data/`.

Scripts

1. `plot_heart_attack_mortality.R`

- Reads outcome data and plots a histogram of 30-day death rates for heart attack.
- Run with:

```
source("plot_heart_attack_mortality.R")
```

2. `best.R`

- Function: `best(state, outcome)`
- Returns the hospital with the **lowest 30-day mortality rate** in a state for a specific outcome.
- Usage:

```
source("best.R")
best("TX", "heart attack")
```

3. `rankhospital.R`

- Function: `rankhospital(state, outcome, num)`
- Returns the hospital with the specified rank in a state for a given outcome.
- `num` can be `"best"`, `"worst"`, or an integer.
- Usage:

```
source("rankhospital.R")
rankhospital("TX", "heart failure", 4)
```

4. `rankall.R`

- Function: `rankall(outcome, num)`
- Returns a data frame of hospitals with the given rank in **every state**.
- `num` can be `"best"`, `"worst"`, or an integer.
- Usage:

```
source("rankall.R")
rankall("heart attack", "best")
```

Notes

- Warnings about NAs when converting columns to numeric are expected and normal.
- Ties in mortality rates are broken alphabetically by hospital name.
- Example outputs are included in the scripts.

Repository Structure

```
R_Hospital_Quality_Assignment3/
│
├─ README.md
├─ best.R
├─ rankhospital.R
├─ rankall.R
├─ plot_heart_attack_mortality.R
├─ data/
│   ├─ outcome-of-care-measures.csv
│   ├─ hospital-data.csv
```

```
│   └─ Hospital_Revised_Flatfiles.pdf
└─ .gitignore
```

This single README.md explains the whole project and how to run all scripts.