Exercise: Practicing Data Wrangling with Heart Rate vs. Age Dataset

Objective:

Use R to explore, clean, and manipulate the dataset containing age and maximum heart rate (`hrmax`) values. Practice using key data wrangling functions such as `filter()`, `mutate()`, `arrange()`, `select()`, and `summarize()`.

Dataset Description:

The dataset has two columns:

- `age`: The age of the individual.
- `hrmax`: The maximum heart rate of the individual.

Dataset name: heartrate.csv

- This data is from an exercise study on maximum heart rates.
- Each row is a person.
- The two variables are the person's age in years and their maximum heart rate (hrmax) in beats per minute, as measured by a treadmill test.

Data is located on the class Github.

Tasks:

Task 1: Load and Preview the Data

1. Load the dataset using `read.csv()`.
2. View the first few rows using `head()`.

R Copy code

```r
# Load the dataset
heartrate <- read.csv("path/to/heartrate.csv")

# Preview the dataset
head(heartrate)
```

Task 2: Filter Rows

1. Filter the dataset to include only individuals aged **30 and above**.
2. Store the result in a new data frame `heartrate_30plus`.

```R
# Filter individuals aged 30 and above
heartrate_30plus <- heartrate |> filter(age >= 30)

# Preview the filtered dataset
head(heartrate_30plus)
```

Task 3: Create a New Variable

1. Add a new column, `target_hr`, which calculates the target heart rate as **85% of the maximum heart rate (`hrmax`)**.
2. Store this in a new data frame `heartrate_with_target`.

```R
# Add a column for target heart rate
heartrate_with_target <- heartrate |>
  mutate(target_hr = hrmax * 0.85)

# Preview the dataset with the new column
head(heartrate_with_target)
```

Task 4: Sort the Data

1. Arrange the dataset by `hrmax` in descending order.
2. View the top 5 rows.

```R
# Sort the data by hrmax in descending order
heartrate_sorted <- heartrate |> arrange(desc(hrmax))

# View the top 5 rows
head(heartrate_sorted, 5)
```

Task 5: Summarize the Data

1. Calculate the following summary statistics:
   - The average (`mean`) maximum heart rate (`hrmax`).
   - The minimum and maximum ages in the dataset.
2. Use `summarize()` for this task.

R Copy code

```r
# Summarize the data
heartrate_summary <- heartrate |>
  summarize(
    avg_hrmax = mean(hrmax),
    min_age = min(age),
    max_age = max(age)
  )

# View the summary
heartrate_summary
```

Task 6: Combine Multiple Wrangling Steps

1. For individuals aged 30 and above, calculate the average target heart rate (`target_hr`).
2. Use `filter()` , `mutate()` , and `summarize()` together in a pipeline.

R Copy code

```r
# Combine multiple steps
heartrate_summary_30plus <- heartrate |>
  filter(age >= 30) |>
  mutate(target_hr = hrmax * 0.85) |>
  summarize(avg_target_hr = mean(target_hr))

# View the result
heartrate_summary_30plus
```