Prepare for the Analysis

Primary Research Question

How has the men's shotput world record changed over time? What about the women's world record?

Breakdown Your Analysis

Let's break this analysis into its required steps:

1. Create a subset of the dataset that contains only the World Record cases for men's shotput.
2. Create a subset that contains only the World Record cases for women's shotput.
3. Create a scatterplot of year and record shotput distance: one for men and one for women.
4. Confirm from these plots that a linear model is appropriate.
5. Run a linear model for each event and then interpret the results.

Here is the code you will use:

```
#Subset the data
menshot <- WR[WR$Event=='Mens Shotput',]
womenshot <- WR[WR$Event=='Womens Shotput',]

#Create scatterplots
plot(menshot$Year,menshot$Record,main='Mens Shotput World Records',xlab='Year',ylab='World Record Distance (m)',pch=16)
plot(womenshot$Year,womenshot$Record,main='Womens Shotput World Records',xlab='Year',ylab='World Record Distance (m)',pch=16)

#Run linear models
linFit(menshot$Year, menshot$Record)
linFit(womenshot$Year,womenshot$Record)
```

problem

1/1 point (graded)

1) What is the best description of what will be included in the new dataframe "**menshot**"?

- Only those columns in WR that include data from the men's shotput.
- All rows and columns from WR.
- Only those rows in WR that include the event Mens Shotput. ✔

Submit You have used 1 of 1 attempt

problem

1/1 point (graded)

2) Which variable will be on the x-axis of each scatterplot?

- Sex (male or female)
- Year ✔
- Distance

Submit You have used 1 of 1 attempt

problem

1/1 point (graded)

3) Which function will we use to fit a linear model to the world record data?

- linFit ✔
- menshot
- plot

Submit You have used 1 of 1 attempt

problem

1/1 point (graded)

4) What is the dependent variable in our linear models?

- Sex (male or female)
- Shotput distance ✔
- Year

Submit You have used 1 of 1 attempt

problem

1/1 point (graded)

5) Suppose we wanted to subset our dataset for only world records that were from 1990 and later. What caused the error below? (You may want to examine the dataset in R for help.)

```
WR<-WorldRecords
recent<-WR[Year>=1990,]

Error in [.data.frame (WR, Year >= 1990, ) : object
'Year' not found
```

- We forgot to tell R to look in the "recent" dataset for "Year" (ie. recent$Year).
- We forgot to tell R to look in the "WR" dataset for "Year" (ie. WR$Year).> ✔
- We need quotation marks around 1990 even though it is numerical.
- The "Year" variable is spelled differently in our dataset.

Submit You have used 1 of 1 attempt

© All Rights Reserved