

Courseware Course Info Discussion Syllabus Download R and RStudio R Tutorials Readings Contact Us
Progress Office Hours Community

Reflect on the Question

Analyze the Data

Draw Conclusions

Primary Research Question

Which variable has the strongest linear relationship with Earnings: Ride Percentage or Cup Points?

Analysis

Let's break this analysis into the different steps that you will need to take to construct a complete answer. Be sure to:

1. Make a histogram to visualize the distribution of Earnings.
2. Generate the appropriate descriptive statistics for this distribution.
3. Make a correlation matrix for Earnings, RidePer and Cup Points.
4. Plot a scatterplot for Earnings with each variable of interest. **Put Earnings on the y-axis.** Check for outliers.
5. Determine which variable has the strongest linear relationship with *Earnings*.


(2/3 points)

Earnings Distribution

What is the **shape** of the Earnings distribution?

Help

☐ positively skewed 

☒ negatively skewed 

What was the **average** amount earned by a bull rider? (Choose the appropriate measure of center; report without a \$ sign and round to the nearest whole number.)

Answer: 111148

What was the **highest** amount earned by a bullrider? (Report without a \$ sign and round to the nearest whole number.)

Answer: 1464476

Help

Hide Answer

You have used 1 of 1 submissions

(2/2 points)

Make a Scatterplot of Earnings and Ride Percentage

Does the scatterplot show a **linear** relationship?

Yes

What is the **correlation** of Earnings with Ride Percentage? (report to three decimal places)

[Show Answer](#)*You have used 1 of 1 submissions*

(1/2 points)

Create a Scatterplot of Earnings and Cup Points

Help

Does the scatterplot show a **linear** relationship?

☒ No ☐ Yes

What is the **correlation** of Earnings with Cup Points? (report to three decimal places)

Answer: 0.674[Hide Answer](#)*You have used 1 of 1 submissions*

(5/5 points)


Outliers and Influential Points

An outlier can have a significant impact on the correlation coefficient. Sometimes it is important to remove these points to examine the size of this impact. Run this code to **identify** the extreme data value in Earnings:

```
# identify specific case  
which(bull$Earnings == max(bull$Earnings))
```

The extreme earnings data point belonged to the rider that came in _____ Place. (Please spell your answer; do not use numerals.)

Where does this data point fall in the scatterplot? (**Make sure that Earnings is on the y-axis**)

- ☒ Above the line 
- ☐ Below the line
- ☐ On the line

Let's **remove** this data point from the dataset to assess what kind of impact, if any, it had on our correlation analysis. Run this code:

```
#Subset the data  
nooutlier <- bull[-1,]
```

Then **rerun** the correlation matrix and the scatterplots to see the difference. Make sure to use the new dataframe (nooutlier)

that you just created.

After removing the outlier, what was the **new correlation** of Earnings and Ride Percentage? (Round to three decimals)

0.814

0.814

After removing the outlier, what was the **new correlation** of Earnings and Cup Points? (Round to three decimals)

0.904

0.904

We would say that this data point was an **influential point** because it

- ☐ caused the underlying relationship to be non-linear.
- ☐ inflated the relationship between Earnings and the other variables.
- ☐ made the earnings of the other bull riders look less impressive than they really were.
- ☒ masked the strength of the relationships between Earnings and the other variables



[Show Answer](#)*You have used 1 of 1 submissions*[Help](#)

EdX offers interactive online classes and MOOCs from the world's best universities. Online courses from MITx, HarvardX, BerkeleyX, UTx and many other universities. Topics include biology, business, chemistry, computer science, economics, finance, electronics, engineering, food and nutrition, history, humanities, law, literature, math, medicine, music, philosophy, physics, science, statistics and more. EdX is a non-profit online initiative created by founding partners Harvard and MIT.

© 2015 edX Inc.

EdX, Open edX, and the edX and Open edX logos are registered trademarks or trademarks of edX Inc.

[Terms of Service and Honor Code](#)

[Privacy Policy \(Revised 10/22/2014\)](#)

7 of 7

About edX

[About](#)

[News](#)

[Contact](#)

[FAQ](#)

[edX Blog](#)


[Donate to edX](#)

[Jobs at edX](#)


Follow Us

 [Twitter](#)

 [Facebook](#)

 [Meetup](#)

 [LinkedIn](#)

 [Google+](#)