

This is an **individual** graded event. While receiving help is permitted with proper documentation, over-reliance on the assistance from others may result in grade deduction.

Answer the questions in the spaces provided. All code and figures will be submitted on a separate document using the “Compile Report Document” button found at the top of your script in RStudio, ensure you select “MS Word”. Ensure each section of code has a committed label with its corresponding problem, example “#Problem1”. If you run out of room for an answer, you may attach additional sheets of paper.

Introduction:

In this lab we will discover how statistical reasoning could be used to settle a legal dispute between the City of New York and a contractor. Statistics and probability have a long history of application to legal issues. In fact, some of the earliest applications of probability were to resolve legal issues.

In March of 1978, Brink’s Inc. was awarded the contract to collect coins from approximately 70,000 parking meters in the City of New York. In response to an anonymous tip that not all money collected by Brink’s was being delivered to the city’s finance depository, the city began an investigation of parking meter collections. Through the use of “salted” coins (coins treated with a fluorescent substance and inserted into specific meters) and surveillance cameras, five Brink’s employees were arrested and subsequently convicted of grand larceny and criminal possession of stolen property. When they were arrested, they had in their possession \$4,500 in coins stolen that day from parking meter collections. (For a complete background and references see De Groot et al. (1986), *Statistics and the Law*, John Wiley & Sons, New York).

As a result, a civil suit was filed by the City of New York alleging that Brink’s had failed to honor its contract and acted negligently. The city was seeking monetary compensation from Brink’s for losses incurred by the criminal activities of its employees. Brink’s was subsequently found guilty of negligence and breach of contract. The question remained as to what the actual dollar amount of the damages were. Although the exact calculation of monetary loss is impossible to determine, the law allows for the introduction of testimony regarding the estimation of such loss as a “...matter of just and reasonable inference.”

Before you proceed, read “Statistical Analysis of Brink’s Data” by Bentow and Afshartous.

1. **Step 1. Ask a research question.** What is one research question being investigated in this study?

Data Analysis Lab 5

Name: _____

2. Step 2. Design a study and collect data.

Inside of the file “metertheft.csv” are the following variables.

Variable	Description
month	Numerical value assigned to the month of collection
priv	Amount collected by a private company (in dollars)
govt	Amount collected by the government (in dollars)
brinks	Indicates if Brink’s was the company that collected the money during a given month (1 = yes, 0 = no)

- (a) What are the observational units in the data that you believe are related to the court case?

- (b) Describe an appropriate population for this case.

3. Step 3. Explore the data.

- (a) Create a scatter plot that compares the amount of money the government collected (x-axis) and the amount the private company collected (y-axis). Additionally, color the points based on whether or not Brink’s collected the money. Ensure you properly label the plot.
 - i. Describe the relationship between the variables as revealed in the scatter plot(*Hint*: Remember to comment on the direction, strength, and form of association as well as unusual observations.)

 - ii. Would you say that a straight line could summarize the relationship between the amount of money collected by the government and the private company?

Data Analysis Lab 5

Name: _____

- iii. Calculate and interpret the correlation coefficient for the relationship between the amount of money the government and private company collected.

- iv. When you look at the data do you believe that Brink's was collecting enough money? Why or why not?

4. Step 4. Draw inferences beyond the data.

- (a) One approach to answering the original question about how much money the Brink's employees stole is to reason like this: we can use the honest contractors and compare them to the government to see what amount an honest contractor should collect whenever the government collects a certain amount.
 - i. Create a data frame of only months that Brink's did not collect money:

**NonBrinks = MeterData% > %
filter(brinks==0)**

- ii. Using this data frame, create a least squares regression line using the amount the government collected to estimate the amount an honest contractor should collect. Write the equation of your line. $[\hat{y} = \hat{\beta}_0 + \hat{\beta}_1(x_1)]$

- iii. Based on your model, conduct a hypothesis test to determine if there is a relationship between your explanatory and response variable. Ensure you state your null and alternative hypothesis, p-value, and your conclusion.

Data Analysis Lab 5

Name: _____

- iv. How much variation in your response variable is explained by the linear relationship with the explanatory variable? What is this statistic called?

5. Step 5. Formulate Conclusions.

- (a) Next, we will use our model to determine how much money the private companies were expected to collect based on the amount the City collected. The following code will use your model to calculate the expected amount:

First, we must create a data frame of the months Brink's collected money:

```
Brinks = MeterData% > %  
  filter(brinks==1)
```

Next, we will calculate the expected values. You could do this by hand or let R do the heavy lifting for you:

```
expected.amount = predict(model,Brinks)
```

Note: In this example, the model is named "model." This may change based on your model name.

Next, we will find the total amount Brink's should owe back to the city. To do this we will subtract what they should have collected from what they actually collected:

```
sum(expected.amount - Brinks$priv)
```

- (b) Based on your model, how much should Brink's pay the city?

Name:

(a) Suppose you have been hired as a consultant by the City of New York to determine the amount of money that Brink's should pay. Write one paragraph to the judge explaining how you reached your decision. Assume the judge has had a single introductory statistics course. Keep in mind the employees were found guilty. It is your job to provide a statistically valid estimate of how much was stolen. Be sure to incorporate what William Fairley thought were threats to any causal interpretations concerning the differences in revenue collected over the two periods.

- Page 5