Assignment 2
Report Manual

Before you start writing: copy the whole manual, paste it to your favorite text editor and then start checking the statistics.

Task Description

Copy the task description which you received with the assignment and paste it here:

Summary & Visualization

Describe the data sample which your group generated for the assignment. What does is consist of? What does all the variables mean? Present this in a way which is easy for a reader to understand below:

Visualize the data in an appropriate way. You are asked to present the histograms of the distributions of the variables `age`, `fleeing` and `signs of mental illness` included in your data. Write a summary of the conclusions based on the figures consisting of the different histograms. Do not forget to reference to the figures which you base your summary on. Do not forget to write a descriptive figure text about the figures you present (Click **here** for more information about figures). Present all this below:

Question 1

Assumptions

Present the relevant assumptions that have to be made in order to answer the question based on the data sample which your group generated. Consider the following:

- Distribution (Continuous/Discrete)

- Sample size

Present your assumptions below:

Method and Notation

Present your method of finding the average age of the data which you have generated. In the presentation you shall include:

- Notation of all variables with symbols

- Symbolic formula for the calculation of the mean

Continue by describing how you determine the lower bound for the mean age. Include the following:

- Notation of all variables with symbols

- Notation of the distribution of the investigated variable

- Symbolic formula for the calculation of the confidence interval

Present the methodology below:

Result and Discussion

Present the result of your calculations regarding the average age as well as the lower bound for the confidence interval. After presenting the result, give a short probabilistic interpretation of the confidence interval you have presented. What does it actually say in the context of the whole data set? Answer the questions presented in the assignment:

Question 2

Assumptions

Present the relevant assumptions that have to be made in order to answer the question based on the data sample which your group generated. Consider the following:

- Distribution (Continuous/Discrete)

- Sample size

Remember what distribution the variable follows and what that implies for the calculations. Present your assumptions below:

Method and Notation

Present your method of finding the relative proportion of the unarmed victims in your sample. Furthermore, construct the confidence interval. Include the following in your calculations:

- Notation of all variables with symbols

- Notation of the distribution of the investigated variable

- Symbolic formula for the calculation of the relative proportion, as well as the confidence interval

Present the methodology for finding the proportion and the confidence interval below:

Result and Discussion

Present the results of the proportion of unarmed victims together with the confidence interval. Shortly comment on the results and what they mean in the context of the data set:

Question 3

Assumptions

Present the relevant assumptions that have to be made in order to answer the question based on the data sample which your group generated. Consider the following:

- Distribution (Continuous/Discrete)

- Sample size

Remember what distribution the variable follows and what that implies for the calculations. Present your assumptions below:

Method and Notation

Define the variable of interest as the proportion of victims of shootings being mentally ill. Make a clear notation of what you call the variable of interest:

Present your null hypothesis together with the alternative hypothesis. Do this in clear notations so that a potential reader certainly understands it:

Present your method of finding the relative proportion of victims suffering from mental illness and once again be very clear with notations of what you have calculated:

Perform a relevant hypothesis test to verify the statement in the assignment. Clearly state how the hypothesis testing was made by notations as used before when defining the variable as well as null-hypothesis:

Result and Discussion

Clearly state the results of your hypothesis testing and use the notations you have defined:

Comment on the result and what it means in terms of the statement presented in the assignment. Is the statement correct? Comment below:

Question 4

Assumptions

Present the relevant assumptions that have to be made in order to answer the question based on the data sample which your group generated. Consider the following:

- Distribution (Continuous/Discrete)

- Sample size

Remember what distribution the variable follows and what that implies for the calculations. Present your assumptions below:

Method and Notation

Define the variable as the proportion of victims of shootings that were fleeing from the police. Make a clear notation of what you call the variable of interest:

Present your null hypothesis together with the alternative hypothesis. Do this in clear notations so that a potential reader certainly understands it:

Present your method of finding the relative proportion of victims fleeing from the police and once again be very clear with notations of what you have calculated:

Perform a relevant hypothesis test to be able to say whether or not fleeing from the police increases the chances of being shot. Clearly state how the hypothesis testing was made by notations as used before when defining the variable as well as null-hypothesis:

Result and Discussion

Clearly state the results of your hypothesis testing and use the notations you have defined:

Comment on the result and what it means in terms of the intuitive thought presented in the assignment. Was the thought correct? Comment below:

Question 5

Method and Notation

Use the full data set reminiscent of the population to calculate the following:

- The average age of the population

- The proportion of unarmed victims in the population

- The proportion of mentally ill victims within the population

- The proportion of victims who were fleeing within the population

All calculations should include clear notations of the variable of interest.

Result and Discussion

Present the results from the calculations with all the variables and their respective values:

Now use these true population values and compare these to the confidence intervals as well as conclusions in Question 1 to 4. Can you verify any of the conclusion you made based on your own data sample? If not, what do you think is the main reason for this? Write your answers and analysis below: