Peer-graded Assignment: Data Analysis Project

ⓘ It looks like this is your first peer-graded assignment. Learn more

You passed!
Congratulations. You earned 50 / 50 points. Review the feedback below and continue the course when you are ready. You can also help more peers by reviewing their submissions.

Review assignments

Instructions

My submission

Shift in American views on expenditures on dealing with drug addiction: Evidence from the GSS

Discussions

Submitted on December 23, 2020

Shareable Link

PROMPT

Upload the HTML file of your project.

Shift in American views on expenditures on dealing with drug addiction: Evidence from the GSS
Shift in American views on expenditures on dealing with drug addiction: Evidence from the GSS
Our research checks if the American average opinion on the level of expenditures on dealing with drug addiction has changed between 1994 and 2012, where in 1996 some major reforms in the marijuana use reforms have taken place.

RUBRIC

Did the author use the R Markdown to complete their project and upload the resulting HTML file?

- **1 point**
  Yes
- 0 points
  No

AS

M(

DATA - Generalizability: *Describe how the observations in the sample are collected, and the implications of this data collection method on generalizability.*

Did the author describe the sampling method, mentioning that a random sample was used?

1 point
Yes

AS

M(

0 points
No

DATA - Generalizability: *Describe how the observations in the sample are collected, and the implications of this data collection method on generalizability.*

Did the author decide that the results can be generalized to the US population?

1 point
Yes

AS

M(

0 points
No

DATA - Generalizability: *Describe how the observations in the sample are collected, and the implications of this data collection method on generalizability.*

If the author discussed potential sources of bias, comment on this discussion.

DIEGO ALEJANDRO SANCHEZ RODRIGUEZ

I think that everything is fine.

Anubhav Sharma

...

MACATOL, YANNIS BELLE (LAURENTE)

The author mentioned a source of bias, particularly the different years the question is asked.

DATA - Causality: *Describe how the observations in the sample are collected, and the implications of this data collection method on causality.*

Did the author decide that no random assignment was used, and hence causality cannot be inferred?

1 point
Yes

AS

M(

0 points
No

RESEARCH QUESTION: Is the research question phrased in a non-causal way?

1 point
Yes

AS

M(

0 points
No

RESEARCH QUESTION: Is the research question well defined / not vague?

Note: "Well defined" means it is obvious from the research questions which variables will be involved in the analysis.

1 point
Yes

AS

M(

0 points
No

RESEARCH QUESTION: Is it clear why this research question is of interest to the author and/or the reader?

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

PLOTS: Do the plots address the research questions?

Note: There is no requirement on minimum number of plots to be provided. A single plot can be sufficient, as long as it addresses the research question, or multiple plots may be needed.

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

PLOTS: Are the plots constructed correctly?

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

PLOTS: Are the plots formatted well? (Size not too large, not too small, etc.)

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

SUMMARY STATISTICS: Do the summary statistics address the research question?

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

SUMMARY STATISTICS: Are the summary statistics calculated correctly?

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

SUMMARY STATISTICS: Are the summary statistics formatted well? (Not taking up pages and pages, etc.)

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

NARRATIVE: Is each plot and R output accompanied by a narrative?

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

NARRATIVE: Does the narrative interpret the visuals and summary statistics correctly?

1 point
Yes

AS

M(

0 points
No

EDA (Exploratory Data Analysis)

NARRATIVE: Does the narrative address the research question?

2 points
Yes

AS

M(

0 points
No

INFERENCE: *Statistical inference via hypothesis testing and/or confidence interval.*

Are the hypothesis stated clearly and matches the research question?

4 points
Yes hypotheses are stated clearly, and they match the research question

AS

M(

2 points
Yes hypotheses are stated clearly, but they do not match research question

0 points
Not present or not clear

INFERENCE: *Statistical inference via hypothesis testing and/or confidence interval.*

Are the conditions checked in context of the data (not just a generic bullet point list of the conditions, but reasoning through them for the given dataset)?

4 points
All conditions checked correctly and in context of the data

AS

M(

2 points
Correct conditions are listed as a generic list, but not checked in context of the data

0 points
Not present or incorrect

INFERENCE: *Statistical inference via hypothesis testing and/or confidence interval.*

Are the appropriate method(s) the writer will be using stated? Did the author provide a discussion of why they chose these methods, and described how they work? Note that in this part the author should display a thorough and conceptual understanding of how the methodology works, however the write-up does not need to be as detailed as if they were teaching the method to someone with no background in statistics.

4 points
Present and correct

AS

M(

0 points
Not present or incorrect

INFERENCE: *Statistical inference via hypothesis testing and/or confidence interval.*

Was the correct code used and output provided for all required techniques?

See below for which situation requires which technique:

One numerical and one categorical variable (with only 2 levels): hypothesis test + confidence interval

- parameter of interest = difference between two means (theoretical or simulation)
- parameter of interest = difference between two medians (simulation only)

One numerical and one categorical variable (with more than 2 levels): hypothesis test only, compare means across several groups, no defined parameter of interest, ANOVA and pairwise tests (theoretical only)

Two categorical variables (each with only 2 levels): hypothesis test + confidence interval

- parameter of interest = difference between two proportions (theoretical if success-failure condition met, simulation if not)

Two categorical variables (either one or both with more than 2 levels): hypothesis test only, compare proportions across several groups, no defined parameter of interest, Chi-square test only (theoretical if expected sample size condition met, simulation if not)

6 points
All required code and output present and correct

AS

M(

3 points
Only some of the required techniques are presented, but others are missing

0 points
Not present or incorrect

INFERENCE: *Statistical inference via hypothesis testing and/or confidence interval.*

Are correct interpretations and conclusions for all output provided? Note that this portion should be evaluated based on criteria stated earlier about which technique is required when). This includes some or all of

conclusions of hypothesis tests,

interpretations of p-values as conditional probabilities, and

interpretations of confidence intervals

depending on the methods used. All interpretations must be in context of the data and the research question.

- **7 points**
  All conclusions/interpretations present, correct, and in context of the data and the research question

AS

M(

- 4 points
  All conclusions/interpretations present and correct, but not put in context of the data and the research question
- 2 points
  Conclusions/interpretations are provided and correct for only some of the required techniques
- 0 points
  Not present or incorrect

INFERENCE: *Statistical inference via hypothesis testing and/or confidence interval.*

Is whether or not results from hypothesis test and confidence interval agree stated? Or, if doing ANOVA or chi-square testing, did the author state that no other methods were applicable and hence there's nothing to compare?

- **3 points**
  Present and correct

AS

M(

- 0 points
  Not present or incorrect

OVERALL: *Organization and readability.*

The document follows the organization of parts outlined in the template.

- **1 point**
  Yes

AS

M(

- 0 points
  No

OVERALL: *Organization and readability.*

The narrative uses correct grammar and clearly and succinctly addresses the research question.

2 points
Yes

AS

M(

0 points
No

OVERALL: *Organization and readability.*

The code is clear, readable, well organized, and uses syntax and packages taught in the course.

2 points
Yes

AS

M(

OVERALL ASSIGNMENT RUBRIC

Please make any general constructive comments on this project that the student would find helpful.

0 points
No

DIEGO ALEJANDRO SANCHEZ RODRIGUEZ

Great job

Anubhav Sharma

...

MACATOL, YANNIS BELLE (LAURENTE)

The project is already good although the author could improve in conciseness.

Is there anything in the author's writing that is distracting in their writing style? Please provide comments that might be helpful for the author in the subsequent phase of the project.

DIEGO ALEJANDRO SANCHEZ RODRIGUEZ

Everything is fine

Anubhav Sharma

...

MACATOL, YANNIS BELLE (LAURENTE)

None in particular, although it would be better if the points presented would be in shorter paragraphs or sentences.

Based on having evaluated this author's project, what have you learned that might help you with your own project?

DIEGO ALEJANDRO SANCHEZ RODRIGUEZ

I have learned a new form of presenting an RMarckdown file. Thank you.

Anubhav Sharma

An awesome project for everyone like me who is new to R and statistics. This is a great reference to us.

MACATOL, YANNIS BELLE (LAURENTE)

I learned that I should be meticulous when doing data analysis on my own projects.

Edit submission