

Business Statistics Assessment IB94X0 2020-2021 #1

Tim Mullett

- Instructions
- The Scenario
- Question 1
- Question 2

options(width=100)

Instructions

Please complete the following questions. Your answer to each question should have two separate sections. In *Section 1* section, write out your answers using complete sentences. Include descriptive statistics in the text, or in tables or figures as appropriate. Tables and figures should be of publication quality (i.e., fully labelled, etc.). Integrate inferential statistics into your description of the results. Your answer should be short and concise. There should be no R code in Section 1.

Section 2 should include the complete R code that you used. Add comments to explain what the code does. The code should show all of the commands that you used, enough for me to replicate exactly what you did. Check that the code runs in one smooth go when you knit the R Markdown together in a fresh R session. You can include figures here that you used to explore the data that you don't wish to include in the first section. I will use the second section to help identify the sources of any mistakes. The first section should stand alone without the second section.

Use both null hypothesis significance testing and the estimation approach.

While there is a word limit of 3,500 words, your answers should be *much shorter than this*, perhaps 100–200 words per question. You get credit for a clear and concise report, and writing more words than necessary is not required.

Use the template “9999999.zip” which contains “9999999.Rmd” for your assessment. Replace everywhere 9999999 with your own student number. Knit the file to make the html file, which is what we will mark. Zip just the Rmd and html files by *right clicking on the folder to zip it*. Submit this zip file. Do not include data files or project files.

If you have questions, please be sure to post them to the module forum (see the “Forums” tab on the module page on my.wbs). If you are stuck coding, post a minimal working example (<https://gist.github.com/hadley/270442>).

The Scenario

You work for a large retailer that sells food and a variety of other products. The retailer has conducted a trial of a new store layout and signage design. It is hoped that this redesign will increase sales. The trial was implemented in roughly half of stores, which were selected at random. You have data on the sales in each store for the last full reporting period prior to the trial being implemented (“sales_1”), and for the first full reporting period after the change was made (“sales_2”). The retailer operates 3 different types of store, the

type for each store is indicated in “outlettype”. The variable “intrial” indicates whether the outlet was selected to be in the trial (TRUE) or not (FALSE), and the staff_turnover variable indicates the proportion of staff working at that outlet that left during the period the data covers.

Question 1

Did the new store layout and signage change average sales? Examine this in terms of the change in GPB and the change as a percentage. Which measure do you prefer, and why?

Question 2

Now look in more detail at the effect of the trial on sales. Was there a different effect in different outlet types, and does adding staff turnover as a predictor improve the model?
