MATH1324 Introduction to Statistics Assignment 2

Price Wars

Overview

In this assignment you will investigate which major supermarket, Coles or Woolworths, is cheaper. To complete this assignment, you will randomly sample matched products from each store and perform hypothesis testing to determine if there is a statistically significant difference between prices. **You are strongly encouraged to work in groups to ensure you collect adequate data.**

This assignment is worth 10% and must be uploaded to the **Assignment 2 Turnitin link by 22/09/2019**.

Data Collection

You will need to be mindful of a few issues. I'm sure you will discover more during the course of the investigation.

- What is an adequate sample size?
- How will you devise a way to randomly sample items?
- What price will you use?
- How will you ensure that products are matched between stores?

Groups

Students are permitted to work individually or in groups of up to 3 for Assignment 2. **Each group must fill out the following form before 15/09/2019 to register their group details.** Submit the details of your group here.

Group Registration Form

Submission Instructions

The report must be uploaded as a **PDF** with your code showing.

OR

The assignment 2 report must be completed using the R Markdown template provided here:

R Markdown Template - Assignment 2

This assignment is worth 10% and must be uploaded to the **Assignment 2 Turnitin link by 22/09/2019**. **One of the group members (not all) must submit a copy of the report with the details of full name and student number of group members in the above of the first page of assignment!** Group members that are not registered and do not submit a report will not be acknowledged. One group member's submission will be marked and given feedback. It will be the responsibility of the marked group member to share the group's feedback with the other group members. The other group members will receive a mark only.

Information for using the R Markdown package can be found here. The R Markdown template must be updated with the names and student numbers of your group. You must use the headings and chunks provided in the template. Failure to adhere to the template will result in a loss of marks.

The report must uploaded to Turnitin as a **PDF** with your code chunks showing. The easiest way to achieve this is to **Preview** your notebook → **Open in Browser** (Chrome) → Right click on the report in Chrome → Click **Print** and Select the **Destination** Option to **Save as PDF**.

Extensions will only be granted in accordance with the RMIT University Extension and Special Consideration Policy. No exceptions. Assignments submitted late will be penalised (see Course Information for further details).

Collaboration

You are permitted to discuss and collaborate on the assignment with your classmates and other groups. However, the write-up of the report must be an individual/group effort. Assignments will be submitted through Turnitin, so if you've copied from a fellow classmate/group, it will be detected. It is your responsibility to ensure you do not copy or do not allow another classmate/groups to copy your work. If plagiarism is detected, both the copier and the student/group copied from will be responsible. It is good practice to never share assignment files with other students/groups. You should ensure you understand your responsibilities by reading the RMIT University website on academic integrity. Ignorance is no excuse.

Report

The report will be in a reproducible R Markdown format with written sections, R code and output. The **report must be succinct** and will be composed of the following sections.

1. **Executive Summary [Plain text]**: In your own words, provide a brief summary of the investigation. Explain the aim of the investigation, the procedure, sample, variables, main findings and conclusions. Write this section last after you have performed hypothesis testing. (Word count Max: 300 words)

2. **Load Data and Packages [R Chunk]:** Report your R code used to load your data and the packages required to reproduce the report. (Not assessed)
3. **Summary Statistics [R Chunk]:** Use R to summarise the measurement variables from the investigation. Include an appropriate plot to help visualise the data. Describe the trend.
4. **Hypothesis Test [R Chunk]**: Use R to perform an appropriate hypothesis test to determine which supermarket is the cheapest. You need to explain your choice of hypothesis test, any assumptions and the significance level.
5. **Interpretation [Plain text]:** Interpret the results of the hypothesis test by stating its hypotheses, considering assumptions, interpreting the p-value and confidence intervals and commenting on the statistical significance of the findings.
6. **Discussion [Plain text]:** Discuss the findings of your investigation. What did you conclude? What were the strengths and limitations of your investigation? What improvements could be made or what would you do differently next time?

Assignment 2 Marking Rubric

Criteria	**Not acceptable (0)**	**Needs Improvement (5)**	**Excellent (10)**
Summary , Appearance of all plots you draw in the assignment, Method that you used for choosing the sample size (how you choose your sample size), method of collecting your data sets (25%)	No summary provided OR the summary failed to provide an accurate overview of the report. Plots have no good quality , no legend, no title, no figure number and etc.	A summary was provided, but there was room for improvement.	An accurate and complete summary of the report was provided.
Statistics (10%)	Statistical summaries presented were inappropriate or lacking.	Appropriate statistical summaries were presented, but there were some extra summaries needed.	Appropriate and insightful statistical summaries were presented.
Testing (30%)	The incorrect hypothesis test was used, or no hypothesis test was reported.	The hypothesis test used was correct, but was not correctly specified.	The hypothesis test used was correct and correctly specified.
Interpret (20%)	The interpretation of the hypothesis test was incorrect or misinterpreted.	The interpretation of the hypothesis test needed improvement.	The interpretation of the hypothesis test was accurate and well summarised.
Discuss (10%)	The discussion lacks insight into the investigation's findings. There is no critical discussion included.	Parts of the discussion were detailed, but improvement was needed relating the findings back	The discussion demonstrated a clear understanding of the results within the context of the

		to the problem and/or thinking critically about the investigation.	investigation and the limitations of the investigation performed.
Succinct (5%)	The report is too long and/or lacks clarity.	The report could be written more succinctly. There was unnecessary detail or discussion that distracted from the main findings.	The report is written succinctly and clearly.