

Excel Report

Trimester 1, 2021

1 OVERVIEW OF THE ASSIGNMENT

This assignment will test your skill to present and summarise data as well as to make basic statistical inferences in a business context. You will use the results and any feedback given in the first assignment (Assessment 3, Excel Report) and produce a single report in a word document. You will need to construct interval estimates, perform suitable hypothesis tests and regression analysis, and make conclusion and suggestion for management action.

Your report should be written in a word document (or other word processing alternative) and should be submitted to Turnitin following the requirement explained below.

2 TASK DESCRIPTION

There are two datasets involved in this assignment: Dataset 1 and Dataset 2, which are the same datasets used in the first assignment (Assessment 3, Excel Report). Please refer to Assessment 3 Description for the details about these datasets. All data processing and calculation should be performed in Excel or Statkey (<http://www.lock5stat.com/StatKey>), hence you should not use a statistical table to find critical value or p-value. Specific instructions as to which computer tools should be used for each section will be given during tutorials.

Your tasks are to answer the following research questions given in Section 2 to Section 6 below using dataset 1 or dataset 2 as indicated in each section. To answer each question, you will need to first present the relevant numerical summary (summary statistics) and graphical display and perform suitable statistical analysis to make inferences and to provide a conclusion.

Your tasks are described below.

1. Section 1: Introduction

Provide a brief and clear introduction about the report (e.g. the objective(s) of the report, the datasets involved, etc.).

Find 1-3 articles (minimum one article, maximum three articles) which are **relevant** to any of the research questions given in Section 2 – Section 6 and write a proper literature review. Your literature review should include in-text citation and you will need to add a reference list at the end of your report.

2. Section 2: Do you believe that 99.8% of Google Play Apps (with more than 1 million installs) are free?

Using Dataset 1, provide the frequency and the proportion (either as a decimal or a percentage) for each category of the variable *Free*. You also need to provide a graphical display that easily shows the proportion of each category.

Then, construct a 95% confidence interval of the population proportion of free app.

Finally, write a comment about your findings and answer the question using the confidence interval

3. Section 3: Is the average rating of Google Play Apps (with more than 1 million installs) less than 4.1?

Using Dataset 1, describe the rating distribution of Google Play Store apps. You need to provide numerical summary (sample size, mean, standard deviation and median) as well as graphical display which shows the outliers, if any.

Then perform a suitable hypothesis test to answer the research question above at 5% level of significance.

Finally, write a comment about your findings and answer the question.

4. Section 4: Is there a difference in the rating of Google Play Store Apps of different categories?

Using Dataset 1, first filter the variable Category to include only *Entertainment*, *Shopping*, and *Tools*. Then provide the numerical summary of the variable *rating* grouped by the three different categories. You also need to provide graphical display which shows any outliers.

Then perform a suitable hypothesis test to answer the research question above at 5% level of significance.

Finally, write a comment about your findings and answer the question.

5. Section 5: Can we predict the price of an app based on its rating?

First, filter Dataset 1 to include only paid apps, then provide a graphical display to describe the relationship between the rating of the apps and their prices.

Next, perform a regression analysis and provide the regression output.

Finally, interpret the correlation coefficient, the coefficient of determination and the relevant p-values and use them to answer the research question above.

6. Section 6: Exploration of Two Categorical Variables

Using Dataset 2 that you collected in the previous assignment, describe the relationship between the two variables. You need to provide both numerical summary and a graphical display.

Then, perform a suitable hypothesis test to answer the research question that you propose in the previous assignment. Use a 5% significance level.

7. Section 7: Conclusion

Write a summary of all the findings in the previous sections and then write concluding statements that would benefit a stake holder (e.g. phone users, mobile apps developer, etc.) to take management action.

Finally, suggest further research by discussing an interesting topic or a research question that can be further explored related to the datasets and/or the findings.

3 SUBMISSION REQUIREMENT

Deadline to submit the report: end of Week 11, Sunday 23rd May 2021, 23:59 AEST (Sydney Time)

You need to submit a word document file (or a pdf) to Turnitin. Your document should show all computer outputs (numerical summary & graphs) and discussion. You do **not** need to submit the dataset. You do **not** need to submit the Excel file.

You should submit a correct file, in any case of submitting an incorrect file, resubmission *may be* approved for a valid reason, but this may attract mark deduction.

4 MARKING CRITERIA

Students are advised to read the marking rubric provided on Moodle, as well as detailed marking criteria based on this rubric.

5 DEDUCTION, LATE SUBMISSION AND EXTENSION

Late submission penalty: - 5% of the total available marks per calendar day unless an extension is approved. This means 0.75 marks (out of 15 marks) per day.

For extension application procedure, please refer to the FAQ on Moodle (<https://koi.edu.au/wp/faq-academic-area/>). The form to apply for an extension can also be found on Moodle (<http://koi.edu.au/wp/policies-forms-2/>). **Please do NOT email your lecturer or tutor to seek an extension, as they are not authorised to grant an extension.**

6 PLAGIARISM

Please read Section 3.4 Plagiarism and Referencing, from the Subject Outline. Below is *part* of the statement:

“Students plagiarising run the risk of severe penalties ranging from a reduction through to 0 marks for a first offence for a single assessment task, to exclusion from KOI in the most serious repeat cases. Exclusion has serious visa implications.”

“**Authorship** is also an issue under Plagiarism – KOI expects students to submit their own original work in **both assessment and exams**, or the original work of their group in the case of a group project. All students agree to a statement of authorship when submitting assessments online via Moodle, stating that the work submitted is their own original work.

The following are examples of academic misconduct and can attract severe penalties:

- **Handing in work created by someone else (without acknowledgement)**, whether copied from another student, written by someone else, or from any published or electronic source, is fraud, and falls under the general Plagiarism guidelines.
- **Students who willingly allow** another student to copy their work in any assessment may be considered to assisting in copying/cheating, and similar penalties may be applied. ”