THE UNIVERSITY OF SYDNEY

School of Economics

ASSIGNMENT

Semester 1 - 2023

ECMT1010 Introduction to Economic Statistics

Due: 11.59PM Friday 19 May 2023

Instructions

i. Review the entries for 'Late submission', 'Special consideration', and 'Academic integrity' in the ECMT1010 unit of study outline.

ii. Enter your answers using the Word template available under the ECMT1010 Canvas module 'Assignment'.

iii. Save your answers as a .PDF file named 123456789.pdf where 123456789 is your 9-digit University of Sydney SID. *Do not put your name on your answers. Do not include a cover sheet.*

iv. Submit the .PDF of your answers under the ECMT1010 Canvas module 'Assignment'. *Work submitted after 11.59PM Friday 19 May 2023 is subject to a penalty of 5% per calendar day late. Work submitted after 11.59PM Monday 29 May 2023 will receive a mark of 0.*

v. Use your assigned data set (see below) is available under the ECMT1010 Canvas module 'Assignment'. Enter your data set number (#) using the box provided in the Word template. *Use of the wrong data set is automatically reported to the Educational Integrity Coordinator, Faculty of Arts and Social Sciences.*

vi. There are 10 questions worth 2 marks each for a maximum of 20 marks. Answer all questions. The assignment is anonymously graded (provided you don't put your name on it).

vii. Show numerical answers to 3 decimal places. Carry out all tests at a 5% level of significance.

viii. When communicating statistical results, it is important to be accurate and concise. Keep your comments, conclusions, comparisons, etc., to one or two sentences. *Excessively long responses indicate a lack of understanding and will be penalised accordingly.*

Aim: The assignment illustrates the use of various statistical methods and software (e.g., Excel, StatKey) to analyze economic data.

Data description: Your assigned data set is extracted from the Hollywood Movies dataset in the 3rd edition of the Lock et al. [...] ts of the movie name, Rotten Tomatoes score, budget, and world gross (the [...] s of US dollars) from a randomly-selected sample of 101 movies released [...] 18.

- Your assigned [...] nder the Canvas module 'Assignment' in the Excel file `Movies#.csv` [...] **igit of your SID**).

- The first row contains the variable names; the remaining rows contain the movie data. The `Movie` column contains the name of the film, `RottenTomatoes` is the percentage of critical reviews that are more positive than negative on the review aggregation website of the same name, `Budget` is the production cost of the movie (in millions of US dollars), and `WorldGross` is the cinema ticket revenue generated by the movie (in millions of US dollars).

QUESTIONS

1. Find out and report the number of outliers the Budget variable has, briefly showing your steps. [2 marks]

2. Using appropriate software, produce a scatterplot of Budget against RottenTomatoes for your sample. Compute the sample correlation. Comment on the association between the two variables. [2 marks]

3. You want to use your data to investigate whether movie critics are biased in favour of big-budget movies. While you realize you cannot establish causality in this setting, you nonetheless decide to investigate the possibility that larger movie budgets are associated with more favourable critical reviews. Write down the null and alternative hypotheses accordingly, taking care to define your notation clearly. [2 marks]

4. Using a randomization distribution, carry out the hypothesis test from the previous step, reporting your test statistic and p-value. Do you find any evidence to support the possibility that critics are biased towards big-budget movies? [2 marks]

5. Another interesting question is whether people are more likely to see movies with favourable critical reviews. While (again) you cannot establish causality, you decide to investigate the possibility that more favourable critical reviews are related to higher ticket revenue. Set up the null and alternative hypotheses accordingly, taking care to define your notation clearly. [2 marks]

6. Using the appropriate approximation distribution, carry out the hypothesis test, showing your test statistic, and p-value. Do you fund any evidence to support the possibility that larger audiences attend movies with favourable reviews? [2 marks]

7. You are also interested in whether a film's budget can be used to predict its ticket revenue. Write down the population regression model that would allow you to make such a prediction. Take care to define your notation clearly. [2 marks]

8. Using appropriate [...] the regression model and report your results. Interpret the regression slop[...] [...]

9. Test whether the i[...] in your model is an effective predictor in the estimated regression model. [...] ur null and alternative hypotheses, the test statistic, decision rule, and [...]. [2 marks]

10. Use the model to predict the ticket revenue for a blockbuster film with a budget of $500 million. How confident are you in this prediction? Briefly explain your reasoning. [2 marks]

Due: 11.59PM Friday 19 May 2023