MIS771 Descriptive Analytics and Visualisation

DEPARTMENT OF INFORMATION SYSTEMS AND BUSINESS ANALYTICS
DEAKIN BUSINESS SCHOOL
FACULTY OF BUSINESS AND LAW, DEAKIN UNIVERSITY

Assignment One

Tasks

Assignment one is an individual assignment with three tasks. The first task is to report on how to plan and deliver the assessment task on time. The second task is to analyse the given dataset and draw conclusions from the analysis. Finally, the third task is to convey the findings and conclusions in a written report to a person with very little or no knowledge of Business Analytics.

The assignment uses the dataset file *T22022A1.xlsx* and expects the use of techniques studied in Module-1.

Percentage of the final grade	30%
The Due Date and Time	**8.00 pm Thursday 18th August 2022**

Submission instructions

The assignment must be submitted by the due date electronically in CloudDeakin. Please note that we will NOT accept any paper, email copies, or part of the assignment submitted after the deadline.

Extensions

No extensions will be granted unless there are exceptional and most unusual circumstances outside the student's control.

A student who requires a time extension must use the Extension Request tool in CloudDeakin's Assessment menu. In addition, please provide relevant supporting documents, the plan for assignment one and a draft of the analysis.

Extension requests will not be considered after 5 pm on 18th August 2022.

Late submission

The following marking penalties will apply for late submissions without an approved extension:

- 5% will be deducted from available marks for each day, or part thereof, up to five days.
- Work submitted more than five days after the due date will not be marked and will receive 0% for the task.

Note: 'Day' means calendar day.

The Unit Chair may refuse to accept a late submission where it is unreasonable or impracticable to assess the task after the due date.

Academic misconduct

For information about academic misconduct, special consideration, extensions, and assessment feedback, please refer to the document *Your rights and responsibilities as a student in this Unit* in the first folder next to the Unit Guide in the Resources area of the CloudDeakin unit site.

Assurance of Learning

This assignment assesses the following Graduate Learning Outcomes and related Unit Learning Outcomes:

Graduate Learning Outcome (GLO)	Unit Learning Outcome (ULO)
GLO1: **Discipline-specific knowledge and capabilities** - appropriate to the level of study related to a discipline or profession. **GLO2**: **Communication -** using oral, written and interpersonal communication to inform, motivate and effect change **GLO5: Problem Solving** - creating solutions to authentic (real world and ill-defined) problems. **GLO6**: **Self-Management** - working and learning independently, and taking responsibility for personal actions	**ULO 1**: Apply quantitative reasoning skills to solve complex problems. **ULO 2**: Plan, monitor, and evaluate own learning as a data analyst. **ULO 3**: Deduce clear and unambiguous solutions in a form that they useful for decision making and research purposes and for communication to the wider public.

Feedback before submission

Students can seek assistance from the teaching staff to ascertain whether the assignment conforms to submission guidelines.

Feedback after submission

An overall mark, together with feedback, will be released via CloudDeakin, usually within 15 working days. Students should refer and compare their answers to the feedback to understand any areas of improvement.

Support

The Division of Student Life (see link below) provides all students with editing assistance. Students who wish to take advantage of this service must organise, plan and contact the Division of Student Life to schedule a booking well in advance of the due date of this assignment.
http://www.deakin.edu.au/about-deakin/administrative-divisions/student-life

Referencing
Any material used in this assignment that is not the student's original work must be acknowledged as such and appropriately referenced. Students can find information about plagiarism and other study support resources at the following website: http://www.deakin.edu.au/students/study-support

United Health Group Case Study

'Medical malpractice insurance is a type of professional liability insurance which protects the physicians and other licensed health care professionals (e.g. dentists, nurses) from liability associated with wrongful practice resulting in a bodily injury, medical expenses and property damage, as well as the cost of defending lawsuits related to such claims' (National Association of Insurance Commissioners[*].)

According to a study published by the Harvard School of Public Health in 2010, the cost of medical malpractice in the United States is $55.6 billion a year, which is 2.4 per cent of annual healthcare spending[†]. Another study published in the New England Journal of Medicine revealed that from 1991 to 2005, 7.4 per cent of all physicians licensed in the US had a malpractice claim against them annually.

Uncapped economic and noneconomic damages awarded by courts have increased premiums for medical malpractice insurance. Worst of all is the trend of specialists moving into defensive medicine. That is to reduce the number of high-risk patients they accept. See LeverageRX's 2019 Medical Malpractice Payout report for the current state of medical malpractice insurance at https://www.leveragerx.com/malpractice-insurance/2019-medical-malpractice-report.

To fix this broken medical system, some state and federal governments introduced new legislation – Medical Liability Reform (MLR). These reforms included a cap on noneconomic damages and specifying expert witness qualifications. See the American Medical Association report on MLR at https://www.ama-assn.org/system/files/mlr-now.pdf

The United Health Group (UHG): America's most prominent health insurance provider has collated a range of data and wants a better understanding of payouts for medical malpractice lawsuits. Its records show payout amounts and information about the presiding physician and the claimant for many mediated or settled lawsuits last year.

You are a Data Analyst working for United Health Group. Your Manager – Edmond Kendrick, has asked you to analyse the collected data. In particular, you are expected to perform descriptive and inferential analyses and produce a report based on your findings.

The data set contains numerous variables and details about the claims. The eight variables in the data table are described below:

Claimant ID	*Unique ID of the claimant*
Payout Amount	*Payout amount of the claim in dollars*
Severity	*The severity rating of damage to the patient (MILD, MEDIUM, SEVERE)*
Age	*Age of the claimant in years*
Private Attorney	*Whether a private attorney representing the claimant*
Marital Status	*Marital status of the claimant*
Speciality	*The specialty of the physician involved in the lawsuit*
Insurance	*Type of medical insurance carried by the patient*
Gender	*Patient Gender*

Edmond's email to you is reproduced on the next page.

---

[*] https://content.naic.org/cipr_topics/topic_medical_malpractice_insurance.htm
[†] https://www.hsph.harvard.edu/news/press-releases/medical-liability-costs-us/

Email from Edmond Kendrick

To: <<Your Name>>
From: Edmond Kendrick
Subject: Analysis of Claims (updated 03/08/2023)

Hi <Your Name>,

I have the following questions/issues relating to the claims dataset.

Theme -1: Analysing the impact of private attorneys.

1.1   Do claimants with private attorneys receive higher payouts than those without private attorneys?

1.2   Is private attorney representation less common in malpractice claims against dermatologists than orthopaedic surgeons?

1.3   Is there a difference in private attorney representation between ~~surgeon~~ *physician* specialities?

1.4   Design an experiment to see the effect of private attorney representation and insurance type on the amount claimed?

Theme -2: Analysing medical malpractice claims and medical liability reform.

2.1   Does the average payout significantly differ across the ~~surgeon~~ *physician* specialities?

2.2   Is the average payout for "SEVERE" claims higher for Orthopaedic surgeons compared to all others?

2.3   Obstetricians/Gynaecologists have been reducing the number of high-risk obstetric patients they accept due to fear of liability claims. Design an experiment to test the effectiveness of Medical Liability Reform (MLR) legislation on the acceptance rate of high-risk obstetric patients?

I look forward to your response on or before 18th August 2022.

Regards

Eddie

SUBMISSIONS

The assignment consists of **three** parts: **Assignment Planning and Execution Tables, Analysis and Report.** You must submit all three (your plan, data analysis and written report).

1. Guidelines for Assignment Planning and Execution Tables

This practical task aims to help you keep track of your progress with the assignment and submit it on time. To report how you plan your project and turn the plan into action, you must complete the tables provided in dot points as clearly as possible. Remember, effective planning, execution, and completing given tasks on time are essential self-management skills.

   Note: Dot point writing requires you to use 'point form', not complete sentences.

The assignment planning and execution details should be submitted in the appropriate tables provided. The tables should be in dot points. However, before filling in the tables, you are strongly encouraged to watch the pre-recorded workshop 'How to plan an assignment and turn the plan into action?'

   Note: Give the assignment planning and execution file the following name
   A1_Planning_YourStudentID.docx

2. Guidelines for _Data Analysis_

Read the case study and questions asked by Edmond carefully. Then spend some time reviewing the data to get a sense of the context. The analysis required for this assignment involves material covered in **Module 1,** with the corresponding tutorials being a helpful guide. Start the analysis by translating the worded business problems into testable propositions.

You can complete all data analysis using the Excel templates provided in practicals. In choosing the technique to apply for a given question, keep the following in mind:

- Are we dealing with a numerical variable or a categorical variable?
- Are we dealing with one population, two populations or more than two populations?
- Are we dealing with an independent population or a dependent population?
- Each question must be answered using the most appropriate technique(s) and justify your decision where applicable.
- Please formulate the hypotheses, and state them clearly in both notation and words in the Excel file.
- Even though a question(s) lead you to inferential techniques, consider conducting a descriptive analysis of the data first.

For all questions in the email, assume that:

- **95 % confidence level** is appropriate for confidence intervals and;
- **5.0 % level of significance** (i.e. α = 0.05) is appropriate for any hypothesis tests.

   ATTENTION!

- We expect you to estimate and report the difference if you have established a difference between two means or proportions.
- Suppose you have established a difference between two or more means or proportions. In that case, we expect you to follow up with an appropriate multiple comparison procedure.

You may need to make assumptions about the dataset to answer some questions. There will be technical/statistical assumptions that you need to make, for example, whether to use an equal or an unequal variance test. _You need to consider and incorporate any violations of assumptions such as unequal sample sizes as limitations of your analysis in both Excel and the report._

Each question from the email should be analysed in a separate worksheet (e.g. Q1.1, Q2.1 …). Therefore, you need to add these extra tabs.

Before submitting your analysis, ensure it is logically organised and removed any incorrect or unnecessary output. Marks will be penalised for poor presentation, disorganised/incorrect results, or unnecessary output.

Note: Name your Excel file in the following format **A1_*YourStudentID*.xlsx** (use a short file name while analysing).

3. Guidelines for the Report

Your report consists of three sections: **Introduction**, **Main Body**, and **Conclusion.** The report should be around 1,200 words.

The introduction begins by highlighting the primary purpose(s) of the analysis and concludes by explaining the report's structure (i.e., subsequent sections).

The report's main body summarises the key findings for each question, specifically addressing Edmond's questions. Use simple English to convey your conclusion and avoid statistical jargon. Use proper headings and do not include any Excel output in the report.

The conclusion should highlight the key findings of the analyses and **explain the main limitations** (if any).

When you have completed the report, it is a useful exercise to leave it for a day, return to it and then re-read. Does it flow easily? Does it make sense? Often, on re-reading, you become aware that you have made some points clumsily, and you find that you can re-phrase them much more clearly.

Note: Give the report the following name **A1_*YourStudentID*.docx** or **pdf.**

Rubric

PERFORMANCE	YET TO ACHIEVE MINIMUM STANDARD		MEETS STANDARD		EXCEEDS STANDARD	
Criteria	Not Attempted	Needs Improvement	Satisfactory	Good	Very Good	Exemplary
Analysis (Marks: 35%) *GLO1 GLO5*	Explores topic at a surface level and rarely uses appropriate data analysis tools. 0 – 10.4	Explores topic at a limited level and often uses irrelevant or inappropriate techniques to analyse data and/or there are many errors in the analysis. 10.5 – 17.4	Independently pursues substantial knowledge, explores topics in some depth and uses appropriate descriptive analysis and visualisation tools to analyse the data but there are some errors in the analysis. 17.5 – 20.9	Independently pursues substantial, additional knowledge, and explores a topic in depth, yielding mostly rich analysis using appropriate techniques but there are minor errors in the analysis. Some assumptions/limitations are considered. 21 – 24.4	Learning interests exist outside classroom requirements. Explores a topic in-depth and produces a very comprehensive analysis of the data using most appropriate techniques. Consistent use of correct Inferential analysis methods to answer questions. All assumptions/limitations are considered and communicated. 24.5-27.9	Learning interests exist and flourish outside classroom requirements. Explores a topic in depth and produces a skilful and comprehensive analysis of the data using many different techniques. All assumptions/limitations are addressed. 28-35
Interpretation (Marks: 45%) *GLO1 GLO2*	Does not communicate any of the main findings of the analysis in an accurate or useful way. Provide no insight and/or information beyond basic facts. 0 – 10.4	Interpretation and communication of findings is at a basic level or do not adequately explain the main findings of the analysis. Provide little insight and/or information beyond basic facts. 10.5 – 17.4	Explains most of the main findings of the analysis accurately. Shows evidence of independently applying learning to demonstrate comprehension in familiar and novel situations. 17.5 – 20.9	Provides a reasonable and accurate description of the most important features of the analysis along with appropriately qualified conclusions. Shows evidence of independently and creatively applying learning to demonstrate comprehension in familiar and novel situations. 21 – 24.4	Provides a very detailed and accurate description of the most important features of the analysis in appropriate language. Report clearly addresses main research questions asked in the assignment. Shows evidence of independently and creatively applying learning to demonstrate high level of comprehension in familiar and novel situations. 24.5-27.9	Provides an outstanding descriptions and inference of results. Conclusions are explained in clear language and insightful. The entire report is aimed to address main research questions. Shows evidence of independently and expertly applying learning to demonstrate outstanding level of comprehension in familiar and novel situations. 28-35
	The report is poorly structured and/or few sections missing with a poor use of language. 0 – 2.9	The report is poorly structured. Only few analysis insights are considered. Language is difficult to follow with many grammatical errors noted. 3 – 4.9	The report is clear and easy to follow. The 3 main elements of a well- structured statement (i.e., claim + evidence + conclusion) is considered across the entire report. 5 – 5.9	The report is well-structured with all sections included. All relevant insights are in the report. Communication is clear with NO grammatical errors noted. 6– 6.9	The report is on par with a professional report. All relevant analysis insights are included in the report Written communication is clear, easy to follow and has a structure. 7 – 7.9	The report is masterfully structured. All relevant analysis insights included. Report follows a consistent formatting style. Language is truly professional and easy to follow. 8 – 10
Assignment Planning and Execution (Marks: 20%) *GLO 6*	Takes no responsibility for maintaining accurate evidence of learning achievements from within formal course experiences. 0 – 5.9	Takes little responsibility for maintaining accurate evidence of learning achievements from within formal course experiences. 6 – 9.9	Takes responsibility for seeking improved learning and maintaining evidence of learning achievements from within formal course experiences, although there is some inconsistency in application. 10 – 11.9	Consistently takes responsibility for seeking improved learning and maintaining evidence of learning achievements from within formal course experiences. 12 – 13.9	Consistently takes responsibility for maintaining accurate and detailed evidence of learning achievements from within and beyond formal course experiences. 14 – 15.9	Consistently takes responsibility for maintaining accurate and compelling evidence of learning achievements from within and beyond formal course experiences. 16 – 20
Overall Description	Or Equivalent Fail (N) 0 – 49%		Or Equivalent Pass (P) 50% – 59%	Or Equivalent Credit (C) 60% – 69%	Or Equivalent Distinction (D) 70% –79%	Or Equivalent High Distinction (HD) 80% –100%