MIS771 Descriptive Analytics and Visualisation

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

Assignment Two

Background

This is an **individual** assignment. You need to analyse a given dataset, and then interpret and draw conclusions from your analysis. You then need to convey your findings in a written report to an expert in Business Analytics.

Percentage of the final grade	30%
The Due Date and Time	11.59 pm **Sunday 17th May** 2020

Submission instructions

The assignment must be submitted by the due date, electronically in CloudDeakin. When submitting electronically, you must check that you have submitted the work correctly by following the instructions provided in CloudDeakin. Please note that we will NOT accept any paper or email copies, or part of the assignment submitted after the due date.

Information for students seeking an extension BEFORE the due date

If you wish to seek an extension for this assignment before the due date, you need to apply directly to the Unit Chair by completing the Assignment and Online Test Extension Application Form before Friday 5 pm 15th May 2020. Please make sure you attach all supporting documentation and a draft of your assignment. The request for extension needs to occur as soon as you become aware that you will have difficulty in meeting the due date.

Please note: Unit Chairs can only grant extensions up to **two weeks** beyond the original due date. If you require more than two weeks, or have already been provided with an extension by the Unit Chair and require additional time, you must apply for Special Consideration via StudentConnect within 3 business days of the due date.

Conditions under which an extension will usually be considered include:

- **Medical** – to cover medical conditions of a severe nature, e.g. hospitalisation, serious injury or chronic illness.

  Note: temporary minor ailments such as headaches, colds and minor gastric upsets are not serious medical conditions and are unlikely to be accepted. However, serious cases of these may be considered.

- **Compassionate** – e.g. death of a close family member, significant family and relationship problems.

- **Hardship/Trauma** – e.g. sudden loss or gain of employment, severe disruption to domestic arrangements, a victim of crime.

Note: misreading the due date, assignment anxiety, or multiple assignments will not be accepted as grounds for consideration.

Information for students seeking an extension AFTER the due date

If the due date has passed; you require more than two weeks extension, or you have already been provided with an extension and require additional time, you must apply for Special Consideration via StudentConnect. Please be aware that applications are governed by University procedures and must be submitted within three business days of the due date or extension due date.

Please be aware that in most instances the maximum amount of time that can be granted for an assignment extension is three weeks after the due date, as Unit Chairs are required to have all assignment submitted before results/feedback can be released back to students.

Penalties for late submission

The following marking penalties will apply if you submit an assessment task after the due date without an approved extension:

- 5% will be deducted from available marks for each day, or part thereof, up to five days.
- Work that is submitted more than five days after the due date will not be marked; you 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.

Additional information: For advice regarding academic misconduct, special consideration, extensions, and assessment feedback, please refer to the document "Rights and responsibilities as a student" in the "Unit Guide and Information" folder under the "Resources" section in the MIS771 CloudDeakin site.

The assignment uses the dataset file ***A2.xlsx,*** which can be downloaded from CloudDeakin. Analysis of the data requires the use of techniques studied in Module-2.

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

You 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**. You are expected to refer and compare your answers to the feedback to understand any areas of improvement.

The Case Study

TassPaperMill (TMP), a subsidiary of Pinnon Paper Industries (PPI), is an Australian company with a long history of manufacturing paper products. In 2019, TPM produced 6,000 tonnes of products and sold more than 5,800 tonnes of products to local and overseas markets.

TPM sells paper products to two market segments: (1) Newspaper and (2) Magazine, through either directly to customers or various intermediaries.

Despite solid financial performance over the last two decade, TPM is forecasting a significant shift in business climate within the next five years. TPM attributes the downturn to change in end-consumers preferences (i.e. online and social media). Now more than ever, TPM management feels the need to build a strong customer base. Besides, they would like to put in place a formal procedure to forecast turnovers using historical data.

Consequently, TPM commissioned ANALYTICS7 (a Market Research Company you work for) to conduct a large-scale survey of TPM customers to understand their characteristics, perceptions, and intent better.

Data Collection Process (Conducted by ANALYTICS7)

ANALYTICS7 invited the purchasing managers of TPM customers to participate in an online survey. TPM also provided data held in their data warehouse and the decision support system to ANALYTICS7 for the study.

Dataset (accessible via A2.xlsx file)

The dataset consists of 200 customer records. There are three different groups of information in this dataset. The first group of information comes from the TPM data warehouse and includes information about TPM customers such as brand loyalty in years; type; region, and distribution channel. The second group of information relates to customers' perceptions of TPM on various factors (TPM customers were asked to rate TPM on these attributes using a 1- 10 Scale.) The third group of information relates to the quantity purchased by the customers and whether they have signed a contract nominating TPM as their preferred supplier.

A complete listing of variables is provided in ***A2.xlsx*** file.

Your Role in ANALYTICS7

You are a modeller at ANALYTICS7. The team leader (Hugo Barra, with a PhD in Data Science and a Master Degree in Digital Marketing) has asked you to lead the modelling component for the TPM project and report the findings to him. The minutes of your team meeting are below.

Your task is to review and complete the modelling activities as per the document.

Form 210-3

ANALYTICS7 Team Meeting

	ANALYTICS7 727 Collins St, Docklands VIC 3008 Phone: (+61 3 212 66 000) infor@analytics7.com.au	**Reference**	AP-210 TPM Project
		Revised	24th April 2020
		Level	Expert Analysis

Meeting Chair	Hugo Barra				
Date	24 April 2020	**Time**	11:00 AM	**Location**	ANALYTICS7 L4.340
Topic	TPM Research Project – Analytics Details				

Meeting Purpose:	**Specifying and Allocating Data Analytics Tasks**	
Discussion items:	• Variable(s) description. • Modelling quantity ordered from TPM. • Modelling the likelihood of a customer signing a contract with TPM. • Forecasting turnover in the upcoming four quarters. • Producing a technical report.	
Detailed Action Items	**Who:** Modeller	**What:** 1. Providing an overall summary of the following two variables: 1.1. Order_Qty. 1.2. Contract. 2. Identify potential factors that may influence Order_Qty: 2.1. Identify a list of possible factors that influence Order_Qty. 2.2. Build a model to estimate Order_Qty. 3. Hugo has created a separate regression analysis and found that the perception of product quality is a significant predictor of the quantity ordered. Prior research shows that the strength of the relationship between quality and quantity may vary according to brand image. That is, customers often associate the brand image with quality products. Therefore, Hugo believes that the relationship between product quality and quantity ordered should be stronger for those who have a favourable perception of TPM brand. Model the interaction between the variables to test Hugo's assumption. Comment whether there is sufficient evidence conclude that the interaction term is statistically significant in the model.

		4. A model to predict the likelihood of customers signing a contract with TPM. 4.1. Hugo has completed the initial analysis for this task. He has narrowed down the key predictors to *Product Quality*, *Product Line*, *Brand Image*, *Price Flexibility*, and *Competitive Pricing.* Continue and refine his work and develop a model to ascertain the likelihood of customers signing a contract with TPM. 4.2. Hugo is specifically interested in understanding the probability of customers who meet the following criteria to sign a contract with TPM. Those who: a) Feel neutral (a score of 5 on the relevant scales) towards TPM's brand image and its product line. b) Varying levels of perception towards product quality (scores from 1 to 10) and price flexibility scores of 5 and 10. Hugo believes the perceptions around the ability to negotiate the price and the quality of the product would define TPM's success in getting customers to sign a contract with TPM as their preferred supplier. Accordingly, your job is to visualise the predicted likelihood of customers signing contracts with TPM with the attributes described above. 5. A time-series model to forecast TPM's turnover for the next four quarters. 6. A written report detailing ALL aspects of your analysis. The report should be as detailed as possible and should describe ALL key outputs of the analysis. The results of the analysis should drive the recommendations to TPM management.
Next meeting	Monday 18th May	

Appendix: Explanatory Notes

To accomplish allocated tasks, you need to examine and analyse the dataset (**A2.xlsx**) thoroughly. Below are some guidelines to follow:

Task 1 – Summarising dependent variables

The purpose of this task is to analyse and explore the key features of these variables individually. At the very least, you should thoroughly investigate relevant summary measures/charts and graphs of these variables. Proper visualisations should be used to illustrate key features.

Your **technical report** should describe ALL critical aspects of each variable.

Task 2. – Model building (Order_Qty)

You should follow an appropriate model building process. All steps (**including pre and post model diagnostics**) of the model building process should be included in your analysis. You can have as many Excel worksheets (tabs) as you require to demonstrate different iterations of your regression model (i.e., 2.2.a., 2.2.b., 2.2.c. etc.). You must make, and document, reasonable/realistic/practical assumptions about the parameters you are working within Task 2.

Your **technical report** should clearly explain why the model might have undergone several iterations. Also, you must provide a detailed interpretation of ALL elements of the **final** model/regression output.

Task 3. – Interaction effect

To accomplish this task, you need to develop a new regression model using **ONLY** the factors discussed in the team meeting (Item 3). In other words, this section of the analysis is separate from the regression model constructed in Task 2. You must make, and document, reasonable/realistic/practical assumptions about the parameters you are working within Task 3.

Your **technical report** should clearly explain the role of each variable included in the model. A suitable visualisation technique should be provided. Make sure you interpret all relevant outputs in detail and provide managerial recommendations based on the results of your analysis.

Task 4.1 – Model building (likelihood of signing a contract)

You should follow an appropriate model building process. All steps (**including pre and post model diagnostics**) of the model building process should be included in your analysis. You can have as many Excel worksheets (tabs) as you require to demonstrate different iterations of your regression model. You must make, and document, reasonable/realistic/practical assumptions about the parameters you are working within Task 4.

You are required to discuss all details of your predictive model/logistics regression output.

Task 4.2. – Visualising and interpreting predicted probabilities

Your **technical report** must include the predicted probability visualisation and be supplemented by practical recommendations. These recommendations should answer the following question:

"How a change in Product Quality and Price Flexibility may affect the likelihood of signing a contract with TPM for those customers who have neutral feelings towards Brand Image and Product Line?"

Task 5 – Forecasting Turnover

Past quarterly turnovers are given in the Excel file. Your job is to develop a suitable forecasting model to predict turnover for **the next four quarters**.

In your **technical report**, you must explain the reason for selecting the forecasting method to predict future turnover. The report also must include a detailed interpretation of the **final** model (e.g. a practical interpretation of the time-series model…etc.)

Task 6. – Technical report

Your **technical report** must be as comprehensive as possible. ALL aspects of your analysis and final outputs must be described/interpreted in detail. Remember, your audience are experts in analytics and expect **a very high standard of work** from your report. High standards means **quality content** (demonstrated attention to details) as well as an **aesthetically appealing report**.

Note: The use of technical terms is acceptable in this assignment.

Your report should include an **introduction** as well as a **conclusion**. The introduction begins by highlighting the main purpose(s) of analysis and concludes by explaining the structure of the report (i.e., subsequent sections). The conclusion should highlight the key findings and explain the main limitations. There is no requirement for a table of content or an executive summary.

Submission Guide

The assignment consists of **two** documents: 1) *Analysis* and 2) *Technical Report*.

1) Analysis

The analysis should be submitted in the appropriate worksheets in the Excel file. Each step in the model buildings should be included in a separate tab (e.g. 2.2.a., 2.2.b., …; and 3.2.a. 3.2.b., …). Add more worksheets if necessary.

Before submitting your analysis, make sure it is logically organised, and any incorrect or unnecessary output has been removed. Marks will be deducted for poor presentation or disorganised/incorrect results. Your worksheets should follow the order in which tasks are allocated in the minutes of the team meeting document.

Note: Give the Excel file the following name **A2_YourStudentID.xlsx** (use a short file name while you are doing the analysis).

2) Technical Report

Your technical report consists of four sections: **Introduction**, **Main Body**, **Conclusion**, and **Appendices**. The report should be approximately 2,500 words.

Use proper headings (i.e., 1., 2.1., 2.2., …) and titles in the main body of the report. Use sub-headings where necessary.

Visualisations / statistical output allowed in the report are:

1. Interaction effect plots
2. Predicted probability plots.

All other visualisations should ideally be in the ***Appendices*** (appendices are not included in the word count).

Make sure these outputs are **visually appealing**; have **consistent formatting style,** and **proper titles** (title, axes titles etc.); and are **numbered correctly**. Where necessary, refer to these outputs in the main body of the report.

Note: Give the report the following name **A2_YourStudentID.docx**.