STA101 - STATISTICS 1

## **Assessment:** *Assignment 1- Data Analytics Report 1*

|  |
| --- |
| Assessment @ a glance  |
| **Task**  | Use the Data Science Road Map, your Excel & Descriptive Statistics skills to analyse two data sets to solve a problem  |
| **Learning** **outcomes** | Present and describe information effectivelySuggest ways to improve decision making processes |
| **Value** | **15%**  |
| **Submission** | Learnline *Submit Here*Week 7 (see *Learnline* for details) |
| **Volume** **&** **Documentation** | Two documents are required:* An *Excel* workbook showing data analysis and a data dashboard
* A 2- 4-page report in *Word* summarising outcomes and decisions (see required template in task details)
 |
| **Group** or **Individual** | Individual |
| **Criteria** | See separate document |

## **Overview**

This assignment will enable you to practice your data analytics skills learned to date. You need to implement the Data Science Road Map and respond to the scenario on page 2.

### ***SCENARIO***

You are Aya Baxter and you are working as a data analyst in a tertiary institution. Your manager has been asked by a Senior staff member responsible for student engagement to analyse a data set of an ‘at risk’ first year university class (often with a high failure rate) to understand the demographics and the potential for all students to successfully complete the unit. The institution is looking to implement some student support strategies but first need to understand the student cohort in some detail.

To: Aya Baxter

From: Monica Dave

Subject: Student Engagement project

Dear Aya,

The Pro Vice Chancellor (PVC) Student Engagement is gathering data regarding some of the problematic subjects we teach in first year across the institution. She is particularly interested in a data set from a first year Statistics unit which has had some very high fail rates in the past.

Could you please review the data sets extracted from the student system and the Learning Management system and provide some advice including a brief report to help with planning intervention strategies to ensure that students can be successful.

You will need to include a range of descriptive statistics and data visualisations (in the form of a data dashboard) so that the PVC and colleagues will be able to quickly understand the situation when in their next meeting which is scheduled for Week 8 of this semester.

It would be great if you could get this work to me no later than Friday, Week7

I look forward to your response

Monica Dave

## **Preparation**

In order to prepare a reply to the email, you will need to examine and analyse the datasets thoroughly. There are two separate excel files that will need to be reviewed.

Remember the process to use and the importance of the first step: Identifying the problem. The data will be messy, you need to clean the data and decide what will be best for your analysis. After cleaning the data, you will need to do some explorations about what its telling you. Consider the variables against which data is being recorded and the various relationships between variables that may be of help. A consideration of the ethical aspects of this request should also be included.

Demonstrate your use of a pivot table to manipulate the data as part of your analysis. Use the pivot table to create **an Interactive data dashboard** that includes at **least 3 charts** from which decision making is possible.

## **Presentation**

The structure for the report is as follows:

1. Title page
2. Problem: outline the problem you are trying to solve based on the scenario outlined below and the data sets you receive.
3. Data Wrangling: outline the steps you took to clean the data initially and the processes you used to understand what was in the data set you received. Include screen captures etc to show the process used.
4. Features extracted: based on your problem, what features did you choose to work with and why?
5. Ethical considerations: use the factors from the Business Ethics Canvas to write about the ethical aspects you considered
6. Modelling & Analysis: what visualisation and statistical processes did you employ to find answers to the problem you identified. Why did you choose these?
7. Present Results: this should include a data dashboard created in EXCEL and contain both visual, numerical and textual data. Data in your report should refer to data in the spreadsheet
8. References (where appropriate)

## **Criteria**

1. Data Analysis Skills (Use of the Data Science Road Map process to structure the analysis; discuss the ethics😉
2. Demonstration of Statistical Thinking skills (ability to use the language, concepts and processes of statistics to address the problem presented)
3. Demonstration of EXCEL SKILLS (pivot tables, data dashboards, functions)
4. Communication skills – ability to write a coherent report
5. Innovation/Originality – for going above and beyond!