**Data Visualisation Assignment**

Assignment Brief

**Introduction**

This is an individual assignment. Please read this assignment brief carefully in order to do the task. You need to read a case study ‘Happy Cow Ice Cream: Data-Driven Sales Forecasting’ and analyse and visualise two datasets for your report. The assignment is to be carried individually and carries **30%** of the unit marks.

Task

Sarah has obtained two datasets from the Happy Cow store located at the University of Hong Kong from the IT support team.

Clean up the dataset and format it into the time series structure. Using Tableau (preferred) – explore and visualise the dataset by answering the following questions:

1. Read the accompanying case study and explore the provided datasets.
2. Explore and visualise the sales performance of the three consumer groups (students, staff and tourists).
3. Both Mary and Prem, the senior sales assistant, believe that different groups of flavours sell better at different times of the year. Do the data back this up? Please propose your groupings and visualise them to generate insights into the ice cream sales. Regarding flavour groups, does grouping give a better level of analysis than individual flavours?
4. What outliers can be identified from the daily sales of Happy Cow? Please define the outliers and explain how to address them.
5. Given the characteristics of the Happy Cow dataset, what are the purposes of the time series analysis (ie predictive versus descriptive)?
6. Submit a written report (no more than 1000 words) providing visuals and accompanying explanation in support of each question.
7. Case Study provided of ‘Happy Cow Ice Cream: Data-Driven Sales Forecasting’.
8. Dataset 1 (DailySales.xls): Daily sales of several ice cream flavours over five months across three types of consumers (students, staff and others / tourists).

* Dataset 2 (HourlySales.xls): hourly sales over seven months (April to October).

This assignment tests whether you have achieved the learning outcomes of the module:

|  |  |
| --- | --- |
| **PC 5** | Apply systems theory principles in order to differentiate between data and information for decision making and operationalise data / information distinctions in sample case studies. |
| **PC 6** | Acquire, analyse and visualise data in order to reflect on its decision making potential, the information generated, its relevance and validity, and its synthetic potential in new situations. |
| **PC 7** | Critically evaluate different information visualisation approaches and apply them using software (eg Tableau or other) by using sample datasets provided, in order to interpret business decision potential. Learn how to communicate with information visualisation experts and design pathways to information for decision making. |

Grading Criteria

The grading criteria will be applied to your work. The table below is to help you understand the grading criteria and provide you with a frame of reference for your effort.

| **Categories** | **Criteria** | **Marks** |
| --- | --- | --- |
| **The business question:** questions were answered | * Answers to the questions: all business questions were adequately answered with compelling and interesting visuals and explanations. | Excellent (90-100 %) Very good (70-90%) Good (50-70 %) Limited (30-50%) Poor (10-30 %) |
| **Effective communication:** ​clear presentation of results through visualisation | * Results: intended message is communicated clearly to the intended audiences. * Independent message: visualisation tells a story with limited (or no) support from other textual explanations. * Fair representation: data is accurately represented without distortion and message is consistent. | Excellent (90-100 %) Very good (70-90%) Good (50-70 %) Limited (30-50%) Poor (10-30 %) |
| **Creativity and innovation:** new direction in field to approach to visualising the data | * Answers to questions: Clearly conveyed by the visualisation. * Innovation: visualisation itself is innovative and creative. * Inspiration: provocative, compelling and memorable content, message or design. * Unique approach: representation of data is bold and original. | Excellent (90-100 %) Very good (70-90%) Good (50-70 %) Limited (30-50%) Poor (10-30 %) |
| **Design and aesthetics:** appropriate use of colour and design | * Title, headings, labels: appropriate size, location, spelling and content. * Choice of visual: appropriate for the audience and the message being conveyed. * Design: aesthetically pleasing, limited clutter, good use of colour contrast. * Clarity: appropriate balance of function and design. | Excellent (90-100 %) Very good (70-90%) Good (50-70 %) Limited (30-50%) Poor (10-30 %) |
| **Quality of report:**report is well structured and of quality | * Quality: structured report including introduction, body and conclusion, details of processes, methods, working materials, assumption and discussion of results. | Excellent (90-100 %) Very good (70-90%) Good (50-70 %) Limited (30-50%) Poor (10-30 %) |

Submission

For this coursework, you need to submit the following:

* A written report providing answers for each question with accompanying visuals (\*.pdf or \*.word formats).
* Your report must not exceed 1000 words (plus or minus 10%).

Please submit this assignment electronically via Canvas in Word doc or PDF format.

**Assignment 1 Questions and Answers**

1. **What do we have to do for assignment 1?**

The first assignment is to analyse data related to the Happy Cow store located at the University of Hong Kong.

More specifically, the two datasets that have been uploaded to Canvas are:

* Dataset 1 (DailySales.xls): Daily sales of several ice cream flavours over six months (April-September) across three types of consumers (students, staff and others / tourists).
* Dataset 2 (HourlySales.xls): hourly sales over seven months (April to October).

There is a case study report that goes into more detail that can be accessed via Canvas, but in short, the key insights required by analysing the data are to:

* Explore and visualise the sales performance of the three consumer groups (students, staff and tourists).
* Mary and Prem, the senior sales assistant, believe that different groups of flavours sell better at different times of the year. The task here is to show whether the data back this up. Here you would propose your groupings and visualise them to generate insights into the ice cream sales. This will help answer the question of whether grouping provides a better level of analysis than individual flavours.
* What outliers (unusual outcomes compared to others) can be identified from the daily sales of Happy Cow? The goal here is to define the outliers and explain how to address them.
* Answer the query on whether the purposes of the time series analysis (i.e., predictive versus descriptive)?

The above should be provided in a written report (no more than 1000 words) with visuals and explanations.  

1. **Why is there already some analysis and charts in the data file?**

The first question (comparing the three groups) has been completed in Excel but this is just an example, and you should visualise it again using Tableau and also provide the interpretation within your report.  

1. **Do we need to use academic sources in our reports?**

There is no need to use academic and/or any other reference source within the report.

1. **For assignment 1, is a Dashboard needed?**

Dashboard is not required. The only thing required is to answer the questions by any means (including the dashboard sheet needed for that question).

The emphasis of the assignment is on the visuals and the explanations accompanying them as a means of providing insight to the requested questions.

1. **What should we do with the negative values, are they naturally occurring or does some data cleaning need to be performed?**

The negative values and other issues with the data are there purposely. Students need to clean the data.