



Strathmore
UNIVERSITY

STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES
AND ILAB AFRICA
MASTER OF SCIENCE IN DATA SCIENCE & ANALYTICS (MODULE II)
END OF SEMESTER EXAMINATION – EXAM and MARKING RUBRIC GUIDE
DSA 8404

DATE: DUE 20th December 2024

Project Description

Targeting customers is an important marketing decision for most companies. The aim of this project is to identify customer segments that a company can target in any marketing plans and the positioning of products. A company recently introduced a variety of discount offers to members. Customers need to spend a certain amount of money over a defined period of time. In return, the customer received a discount voucher. The company wants to know which members responded best to this marketing campaign by spending more money.

Data

The data comprises of three separate files described below.

portfolio.json contains information about the ten promotions.

- **reward** is the monetary value of the promotion.
- **channels** is the ways in which the promotion was advertised.
- **difficulty** is the amount the customer needs to spend in order to receive the reward.
- **duration** is the total number of days that the promotion was available.
- **offer_type** is the type of promotion. This is either a money off offer (discount), buy one get one free (BOGO) or a newsletter (informational).
- **id** is the promotion identification.

transactions.json contains information specific transaction across the promotion period.

- **person** is the customer membership identification.
- **event is the event** - either offer received, offer viewed, offer complete or transaction. •
value.offer.id is the promotion identification.
- **value.amount** is the amount spent in GBP for a given transaction.
- **value.offer_id** is the promotion identification.
- **value.reward** is the monetary value of the promotion.
- **time** time from the beginning of the promotion period.

profile.json contains information about customers.

- **gender** is the identified gender of the customer.
- **age** is the age of the customer at the time of the promotion period.
- **id** is the customer membership identification.
- **became_member** is the date when the customer became a member.
- **income** is the self-reported income of the customer at the time of the promotion period.

Learning objectives

Question of interest

1. Identify clusters of customers that completed more transactions and/or spent more money over the promotional period.
2. Provide an interpretation of clusters with visual representations of each cluster.
3. Provide a clear description of customer characteristics that the company should focus their attention to and explain why.

How to start

In order to complete this project, you must derive a set of informative variables that you can use to cluster customers into groups. Variables of interest may include, but certainly not limited to, the total average spend per customer, the total number of completed offers, or the number of transactions over the set promotion period.

Submission Requirements

You are asked to produce an executive report based on the analysis. The report should be no longer than 15 pages excluding references. The report should include your analysis and recommendations stemming from the analysis to a given audience e.g., marketing department or customer relations. The main objective of this assessment is to present a clear introduction to your project and present results and conclusions to a non-technical audience. You should assume the audience does not have access to your data and has little experience in the technical details. They should be able to understand the aims of the project and interpret the results and take recommendation actions using insights derived from the analysis.

The suggested layout for the report is as follows, though it is not compulsory to follow this structure but should be used as guidance.

- 1. Introduction:** The introduction should clearly explain the project background and questions of interest.
- 2. Ethical considerations:** Please mention any ethical considerations with respect to using these data and describe any concerns of data privacy.
- 3. Exploratory analysis:** You should include suitable and effective plots and summaries of your data that directly address the questions of interest.
- 4. Methods:** You should provide an appropriate description of the methods used in the context of your specific problem.
- 5. Results:** You should provide a comprehensive interpretation of your modelling results. Recall that you should target a non-technical audience. Your results need to include effective plots that illustrate your results.
- 6. Conclusion:** You should provide a summary of your results using non-technical language. Your audience should be able to understand your findings in relation to your questions of interest. It should include recommendations for actions to take.

| Criteria | Excellent (5) | Good (4) | Satisfactory (3) | Needs Improvement (2) | Inadequate (1) |
|------------------------------------|--|--|--|---|---|
| Introduction (5 points) | - Clearly defined objectives and scope of the project. Engaging and concise introduction. Presents a clear problem statement and context. | - Objectives and scope are defined but may lack some clarity. The introduction is mostly engaging and concise. Problem statement and context are presented but could be improved. | - Objectives and scope are somewhat defined but lack clarity. Introduction is somewhat engaging and concise. The problem statement and context are presented but need improvement. | - Objectives and scope are unclear or missing key elements. Introduction lacks engagement and conciseness. Problem statement and context are unclear or insufficient. | - Objectives and scope are missing. Introduction is absent or completely lacks engagement. Problem statement and context are absent or unclear. |
| Analysis (20 points) | - Comprehensive analysis of customer segments based on spending behavior. Effective use of data visualization techniques. In-depth insights into customer response to the marketing campaign. Clear identification of key customer segments. Robust statistical analysis, if applicable. | - Solid analysis of customer segments with some areas for improvement. Effective use of data visualization techniques, but some visuals could be enhanced. Valuable insights into customer response to the marketing campaign. Identification of key customer segments is clear. Adequate statistical analysis, if applicable. | - Analysis of customer segments is present but lacks depth. Basic use of data visualization techniques. Provides some insights into customer response to the marketing campaign. Identification of key customer segments is somewhat clear. Limited statistical analysis, if applicable. | - Limited analysis of customer segments with significant gaps. Limited use of data visualization techniques or ineffective visuals. Limited insights into customer response to the marketing campaign. Identification of key customer segments is unclear. Inadequate or missing statistical analysis, if applicable. | - No analysis of customer segments. No use of data visualization techniques or inappropriate visuals. No insights into customer response to the marketing campaign. No identification of key customer segments. No statistical analysis, if applicable. |
| Recommendations (15 points) | - Well-founded recommendations based on the analysis. Clear and actionable suggestions for marketing or customer relations. Recommendations directly address the company's objectives. Prioritization of recommendations if applicable. | - Sound recommendations with room for improvement. Suggests actionable steps for marketing or customer relations. Recommendations somewhat address the company's objectives. Some consideration of prioritization if applicable. | - Recommendations are present but lack depth or clarity. Offers vague or less actionable suggestions. Recommendations somewhat address the company's objectives. Limited consideration of prioritization if applicable. | - Recommendations are vague or insufficient. Offers no actionable suggestions. Recommendations do not clearly address the company's objectives. No consideration of prioritization if applicable. | - No recommendations provided. |

| | | | | | |
|---|--|---|--|---|--|
| Clarity and Organization (10 points) | - Clear and well-organized structure throughout the report. Logical flow of information. Effective use of headings, subheadings, and transitions. Concise and well-written content. | - Mostly clear and organized structure, with some minor issues. Generally logical flow of information. Adequate use of headings, subheadings, and transitions. Content is clear and well-written with few issues. | - Structure is somewhat clear and organized but needs improvement. Flow of information can be confusing at times. Limited use of headings, subheadings, and transitions. Content is somewhat clear but may be verbose or contain minor errors. | - Structure is unclear and disorganized. Flow of information is confusing. Lack of headings, subheadings, and transitions. Content is unclear and poorly written with significant issues. | - No discernible structure or organization. Content is entirely unclear and poorly written. |
| Overall Presentation (10 points) | - Professional formatting and layout. Appropriate use of visuals and tables. Error-free writing and language usage. Effective use of references, if applicable. | - Good formatting and layout. Appropriate visuals and tables used, but some room for improvement. Minor language or writing errors present. References are used effectively, if applicable. | - Adequate formatting and layout. Use of visuals and tables could be improved. Some language or writing errors present. References are present but not effectively used, if applicable. | - Formatting and layout need improvement. Visuals and tables are not effectively used. Multiple language or writing errors present. References, if applicable, are missing or not properly cited. | - Poor formatting and layout. No effective use of visuals or tables. Numerous language or writing errors present. No references provided, if applicable. |
| Total (60 points) | - Exceptional report that exceeds expectations. Demonstrates a deep understanding of the project's objectives and analysis. Provides valuable insights and actionable recommendations. | - A strong report that effectively meets the project's objectives. Demonstrates a good understanding of the analysis. Provides valuable insights and actionable recommendations. | - A satisfactory report that meets the basic project requirements. Demonstrates a basic understanding of the analysis. Provides some valuable insights and recommendations. | - A report with significant room for improvement. Demonstrates a limited understanding of the analysis. Provides few valuable insights and recommendations. | - An inadequate report that does not meet project requirements. Fails to demonstrate an understanding of the analysis. Provides no valuable insights or recommendations. |

Additional guidelines on data analytics report to a non-technical audience.

You can find some helpful resources for your project directed towards a non-technical audience. It's important that you are able to present a data analyse appropriate for the target audience, whether that be industry, academic, professional, general public etc.

- [Tips for writing assessments](#). Please read through this general guidance for writing a report specific to an academic assessment (much like this project).
- [Tips for writing a report to a specific audience ranging from an academic, technical audience to a journalistic general audience.](#)
- [What is a statistical story \(with examples\)](#).
- [How to write a statistical story](#).

Below are some examples of written reports that provide good examples of writing a report to a general, non-technical audience.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., ... & Cutler, D. (2016). The association between income and life expectancy in the United States, 2001-2014. *Jama*, 315(16), 1750-1766.

- This article begins by stating its objective and importance.
- The methods section covers data ascertainment.
- The data analysis section is short. They have not done anything majorly complex with respect to analyses, but it's clearly stated.
- The results are separated into clearly labelled small subsections which show their importance.

Makary, Martin A., and Michael Daniel. "Medical error—the third leading cause of death in the US." *Bmj* 353 (2016).

- This report has a very capturing title which draws an audience.
- The report does not follow a typical report structure that we might expect. Instead, there are a few short subsections with clear titles

Novembre, J., Johnson, T., Bryc, K., Kutalik, Z., Boyko, A. R., Auton, A., ... & Bustamante, C. D. (2008). Genes mirror geography within Europe. *Nature*, 456(7218), 98-101.

- This article has a bit more technical detail, more so that what you may aim for in this assessment.
- The reason I included this one is because of the visualisation of the results. The entire story of this work can be summarised in figure 1a. We just need to visually match the colours!

At the end they have a 'key messages' section which summarises their findings

Some additional videos with tips on presenting data

- [How to avoid death by PowerPoint](#)
- [Data Analysis Presentations](#)
- [Making data mean more through story telling](#)
- [Story telling with data](#)

More Resources

<https://www.spektrix.com/en-us/blog/how-to-segment-customers-in-the-arts-beginners-guide>

<https://cloud.google.com/blog/products/data-analytics/creating-actionable-customer-segmentation-models>

<https://seifip.medium.com/starbucks-offers-advanced-customer-segmentation-with-python-737f22e245a4>

<https://medium.com/@kristenkehrer/how-to-use-customer-segmentation-to-learn-f49e82b9a959>

<https://towardsdatascience.com/a-gentle-introduction-to-customer-segmentation-375fb4346a33>