**Handout 1 - Final Project Overview**

| **UK Data Analyst Program - Final Project** | |
| --- | --- |
| **Project Description** | Olist Store is the largest department store in Brazilian marketplaces. Olist connects small businesses from all over Brazil to channels without hassle and with a single contract. The Brazilian eCommerce public dataset of orders (from 2016 to 2018) made at Olist Store has been provided to your company for analysis.  Your manager is asking you to analyse the provided dataset using your data analytics skills, prepare a PowerBI dashboard with visualisations that answer key questions, as well as prepare and present a final report. |
| **Learning Objectives** | * Visualise a company’s customer demographics, sales trend, orders by categories, orders changes by year, etc. * Utilise created visualisations to provide recommendations for improvements * Create a final report. * Deliver a group presentation, clearly communicating analytical results. |
| **BSM** | * All |
| **Duration** | * Final Project: 12 hours working sessions + 15 minute presentation |
| **Team Structure** | * A group of 3 - 5 members |
| **Project Instructions and Deliverables** | **Project Instructions**  Your group is tasked with doing exploratory data analysis with the Olist Store dataset, to generate business insights and provide recommendations for improvement. You will need to use the results of your analysis to create a usable PowerBI dashboard with clear, concise, and informative data visualisations that drive recommended improvements for the client. You will present your results and recommendations in a final report (flexible format - can be a PowerPoint slide deck, a written document, etc).  The Olist Store dataset has information from 100,000 orders made at multiple marketplaces in Brazil. Its features allow you to view an order from multiple dimensions: from order status, price, payment, and freight performance to customer location, product attributes, and finally reviews written by customers. A geolocation dataset that relates Brazilian zip codes to latitude and longitude coordinates is also integrated into the dataset.  After a customer purchases a product from Olist Store, a seller gets notified to fulfil that order. Once the customer receives the product, or the estimated delivery date is due, the customer gets a satisfaction survey by email where he can give a note for the purchase experience and write down some comments.  Here are some ideas of the kinds of analyses your group might conduct for Olist Store:   1. **Feedback Sentiment Analysis.** Evaluate the polarity of the tweets as customer feedback positive, negative or neutral. 2. **Clustering.** Some customers did not write a review. But for those that did, why are they happy or mad? 3. **Sales Prediction:.**With purchase date information you will be able to predict future sales. 4. **Delivery Performance**. You will also be able to work through delivery performance and find ways to optimise delivery times.   **Required Deliverables**  When complete, you will need to submit your groups’ full project portfolio for a group technical score. This portfolio should include the following deliverables:   * A simple project proposal * Power BI Desktop source files (dataset and report)   + Data source file, if not in PowerBI Desktop source files * Final presentation (flexible format; can be PowerPoint slides, a written report document, etc.) * Comprehensive PowerBI Dashboard ready for presentation   The class will have a total of twelve hours for groups to complete project tasks and deliverables; this includes time to research the presentation. Each group will deliver a 15 minute presentation using their final report, followed by 5-10 minutes of Q&A. Time is allocated for final presentation prep and rehearsal before the actual presentation. |
| **References** | **Data Schema**  The Olist data is divided into multiple datasets for better understanding and organisation. Please refer to the following data schema when working with it:  Data Schema  Notice!   * An order might have multiple items. * Each item might be fulfilled by a distinct seller.   **Helpful external references:**   * [Sentiment Analysis in Power BI](https://www.youtube.com/watch?v=mhe9Vs3jQes) * [Text Mining and Sentiment Analysis: Power BI Visualisations](https://www.red-gate.com/simple-talk/sql/bi/text-mining-and-sentiment-analysis-power-bi-visualizations/) * [Forecasting in Power BI](https://www.youtube.com/watch?v=mMd2rKK7dWc) * [Power BI Clustering](https://www.youtube.com/watch?v=LsFDWS77ER4) |
| **Assessment** | * **Technical Assessment:**   + A technical score will be given to each group. This score will be determined by the instructor using the [Final Project Rubric](https://docs.google.com/spreadsheets/d/1lR6qF0---TlvCJQa3hmpmOKoz5DsXbwnsv0Z1CtfGpc/edit?usp=sharing). * **BSM Assessment:**   + After the project, participants will use the [BSM Evaluation Criteria](https://docs.google.com/document/d/1MzfcotTUaBx5uWeoGH-ip0e8IntbraV-MGS0P0LWS6Q/edit?usp=sharing) as a guiding tool to complete the [BSM Self-Assessment Form](https://docs.google.com/document/d/1qUM_ZgboPI9yeSpvC_6bnn_wxNj1r0XmSnXnBHvzsOk/edit?usp=sharing) and [BSM Peer-Assessment Form](https://docs.google.com/document/d/1ZppO_oGopKikFlJgChavh_W13U-kbrQnk6wa8qvHOMI/edit?usp=sharing). |