

**FINAL PROJECT: TABLEAU SUPERSTORE ANALYSIS**

In this final project, your company’s leadership has requested an analysis of revenue data, as sales continue to grow but profitability declines. The data set is in an Excel spreadsheet called “Sample – Superstore.xls,” which was included as part of the Tableau installation. This final project option focuses on your ability to perform and provide visual analysis instead of working with a large data set.

After considering the analytics workflow, review the questions in the final presentation section below and prepare results, findings, and recommendations for a stakeholder presentation.

This project will be a summation of the data and analytical skills you’ve learned throughout this course, with an emphasis on visual analysis and data storytelling for a non-analytical audience. It’s also intended to provide an opportunity to experience the “analytical flow” of data discovery as you visually interrogate the data using Tableau’s drag-and-drop VizQL capabilities. And while there is a list of requirements below, it’s often just as interesting to explore for unknown insights.

**DELIVERABLES AND TIMELINE**

**1. Final Presentation**

* **Due** by 3pm on Wednesday, December 19th.
* **Format**: PowerPoint or Google Slides or PDF
* **Description**:
  + Your team will create a presentation running 5 to 7 minutes that delivers the most important insights to key stakeholders within your company.
  + Review your goals, methods, and findings.
    - Include all relevant information and data.
    - Describe any cleaning methods used on your data.
  + List recommendations based on your data.
  + Present your findings to the class.
  + Allow an additional two minutes minimum for questions.

**REQUIREMENTS**

**1. Exploratory Analysis**

* Deal with missing values.
  + There are no missing values in the Tableau “Sample – Superstore.xls” Excel file. However, as an added bonus, you *can* delete a number of cities and states in the orders data set and then conduct a lookup or blend in another data set of zip code-based cities and states to demonstrate your data-cleaning skills. A set of current zip codes by city and state are available in Excel format [here](http://www.pier2pier.com/links/files/Countrystate/USA-Zip.xls).
  + Remember, you can replace missing data with new values (impute new values) — such as the mean or median of a column — or remove rows that are missing values.
  + Create a new column that calculates the cost of sales for each customer order.
  + Use Excel to either remove/impute missing values or create a JOIN or blend in Tableau.
  + Explain your reasoning for choosing one option over the other.
    - IMPORTANT! Keep a record of your choices, revisions, or modifications when cleaning or transforming data and briefly mention them in your presentation.
* Clean the data (*if you choose to alter the sample file*).
  + Look for any outliers or incorrectly recorded data.
  + Determine if it’s better to remove or include outliers in your analysis.
  + Correct any data-formatting issues. Do columns need to be joined or split?
  + Apply any other necessary normalization or cleaning techniques.
* Explore the data.
  + The goal is to understand the data well enough to begin to see distinct segments within the set.
* Create distinct customer/user segments.
  + Create 2–5 distinct data segments that partition the majority of your data. This range can vary, but it needs to be at least two and more than five becomes a bit difficult to track.
  + The data segments may also be a combination of several different, newly created variables.
* Summarize each data segment.
  + Just as people have distinct personalities, your data segments should also have distinct characteristics and attributes.
  + Use Excel’s statistical and aggregation functions, PivotTables, plots, graphs, and any other methods to summarize and describe your data segments.
* Organize your insights.
  + Organize your insights so that you can navigate them quickly.
  + You might want to put all of your newly created tables and graphs on a separate worksheet in your workbook.
  + Your tables and graphs should have the appropriate titles and labels. They don't need to be absolutely finalized, but they should make sense to someone who isn't familiar with your data. Make sure to include titles, axis labels, column names, and row names where appropriate.

**2. Final Presentation**

* Describe the data sets you worked with.
* Describe the presence of NULL values and how you handled them (*if applicable*).
* Describe your cleaning methods (*if applicable*).
* Create a set of Tableau individual worksheets, then build and conduct a visual analysis to answer the following **questions from your stakeholders**:
  + What products, by category or subcategory, are unprofitable?
  + Is this related to one or more product manufacturers?
  + In a map, show whether or not this is a local or regional issue.
  + Use a scatterplot with trend lines to explore the underlying reasons for why profits are lower for certain products and locations. Attempt to determine the root cause, whether it’s correlated to expedited shipping costs, specific customer segments, or other factors found in the data set.
  + Demonstrate how the profit trend lines change if outliers are addressed.
  + From the scatterplot, create a cross tab that lists unprofitable customer orders.
* Clearly state the problem statement or goal of your analysis:
  + Think about the original question you were trying to answer.
  + With that in mind, identify the insights most important to answering that question.
    - For example, if your original goal was to study your customers' order profitability, you would focus on the attributes that make up the revenues and costs associated with each order.
* Create a dashboard that pulls together three related individual worksheets (e.g., a scatterplot that isolates the primary profitable issue, a color-coded profitability map, and a cross tab of customer orders that illustrates the profitability issue).
* Use 4–5 Story Points in a presentation that summarizes the stakeholder questions as well as your resulting analysis and insights.
  + Bonus! Publish your Story Points view to Tableau Public.
* Finish by summarizing your top insights, recommended actions, or next-level analysis:
  + Answer the main stakeholder questions and then follow with any other insights.
  + Focus on the most important insights and their actionable implications.
  + People's time is valuable; make the most of it!

**GETTING STARTED**

* Include the appropriate charts, graphs, and aggregations to support your insights.
* Your presentation should demonstrate your proficiency with Tableau. Avoid giving a step-by-step recounting of your entire process but do include highlights from any individual worksheets you created that contributed to your analysis.
* **When in doubt, focus on sharing any “aha” moments.**
* Finally, make sure you're familiar with every aspect of your analysis, so you can be ready for questions. Don’t spend a lot of time talking about the intricacies of the data but be familiar enough to answer specific questions from the class.

**RESOURCES**

* Search for [Tableau Superstore demos on YouTube](https://goo.gl/NHeDlR) (start at 1:29 in this video).
* [Here is a shorter version with other visual analyses](https://goo.gl/Rre7ck).
* Set up a [free account](https://public.tableau.com/s/) on Tableau Public to publish Story Points.

**REQUIREMENTS & EVALUATION**

* Your presentation will be graded using the Requirements section as a rubric.
* The Excel workbook containing the master data and data overview will be evaluated using the Requirements section as a rubric.

**RUBRIC**

* For all requirements, project deliverables will be evaluated using a simple point scale.
* In addition to numeric feedback, instructors will provide comments on all required portions.

|  |  |
| --- | --- |
| **Score** | **Expectations** |
| **0** | *Incomplete.* |
| **1** | *Partial credit but does not meet expectations.* |
| **2** | *Meets expectations.* |
| **3** | *Surpasses expectations.* |

**Description:**

* A “1” means you have met some but not all of the project requirements.
* A “2” means you have completely satisfied all requirements.
* A “3” indicates performance above and beyond these requirements and will not apply to most items.