# Guided Project From Coursera: Sales and Management Analysis of Brew Heaven

## Prompt

You're a data analyst brought in by Brew Heaven, a bustling local coffee shop chain known for its cozy atmosphere and fantastic brews. Despite their charm, Brew Heaven faces a big challenge: they manage their sales, inventory, and customer feedback in separate Excel spreadsheets, which makes it hard to see the full picture and leads to missed opportunities. The owner, Emma, is frustrated with this disjointed process and wants a better way to understand what's selling well, what's running low, and how customers feel about their coffee. Your mission is to transform this situation by creating a sleek Power BI dashboard. This new tool will pull together all their data into one place, offering Emma clear insights into daily sales trends, how quickly items are flying off the shelves, and customer satisfaction. With this dashboard, Brew Heaven will be able to make smarter decisions about inventory, highlight their most popular brews, and ultimately provide an even better experience for their loyal customers.

Tasks of Prompt for creating Power BI Dashboard:

1. Import the data "final", transform it and add necessary relationships
2. On the first page, the "Sales" page, insert the following:
    a. Customers Table
    b. Evolution of sales over time
    c. Distribution of sales across products and categories
    d. Overall Sales Revenue
3. On the next page, the "Management" page, insert the following:
    a. Overall Rating
    b. Feedback table
    c. Distribution of stock level by category
    d. Distribution of reorder level by supplier
4. Add a Title, a Theme, and Publish

## Skills Presented

Power BI

- Power Query (for data cleaning, data transformation, and building relationships across different tables)
- ETL Process (Extracting data, Transforming it, and Loading it into Power BI for analysis)

- Dynamic Visualizations (Creating visualizations with tooltips and with the ability to change the information presented when clicking certain aspects of the dashboard for interactivity with Stakeholders)
- Accessible Visualizations (Choosing themes and color schematics that is accessible for people with certain visual impairments such as color-blindness)

## Initial Info on Data

Data Originated from Coursera Guided Project instructor: Hussein ElGhoul, a senior data engineering consultant with over 5 years of experience in data engineering and analytics.

There are 5 Tables in the "final" Excel file attached:

1. Orders:
    a. Rows: 50 Observations + 1 Header
    b. Columns: 6
        i. Order ID (categorical, integer, serves as row identifier)
        ii. Customer ID (categorical, integer)
        iii. Product ID (categorical, integer)
        iv. Quantity (quantitative, integer)
        v. Total Price (quantitative, number)
        vi. Order Date (categorical, date)
2. Feedback:
    a. Rows: 5 Observations + 1 Header
    b. Columns: 5
        i. Feedback ID (categorical, integer, serves as row identifier)
        ii. Customer ID (categorical, integer)
        iii. Rating (quantitative, integer)
        iv. Feedback (categorical, text)
        v. Feedback Date (categorical, date)
3. Inventory:
    a. Rows: 5 observations + 1 Header
    b. Columns: 4
        i. Product ID (categorical, integer, serves as row identifier)
        ii. Stock Level (quantitative, integer)
        iii. Reorder Level (quantitative, integer)
        iv. Supplier (categorical, text)
4. Customers:
    a. Rows: 5 observations + 1 Header
    b. Columns: 3
        i. Customer ID (categorical, integer, serves as row identifier)
        ii. Customer Name (categorical, text)
        iii. Email (categorical, text)

5. Products:
    a. Rows: 5 observations + 1 Header
    b. Columns: 4
        i. Product ID (categorical, integer, serves as row identifier)
        ii. Product Name (categorical, text)
        iii. Category (categorical, text)
        iv. Price (quantitative, number)

## Procedure of Analysis:

1. ETL Process
    a. Extracting "final" data from into Power BI
    b. Transforming data to account for inconsistencies
    c. Loading transformed data into Power BI for analysis
2. Creation of Dashboard for Project
    a. Creating the Sales page
    b. Creating the Management page
3. Formatting of Dashboard for Project
    a. Sizing
    b. Color schemes
    c. Labeling
4. Insights of Dashboard from Project
    a. Insights from Sales Page
    b. Insights from Management Page
    c. Additional Notes
5. Dashboard Revision
    a. Reasoning for Revision
    b. Proceeding with Revision
6. Insights from Dashboard Revision

- Reflections on the Project (not necessarily part of the analysis)

## ETL Process:

After downloading the data for analysis titled "Final" and scrolling through it, we understand the information that was elaborated in the Initial Info on Data section. From there, we needed to clean the data, and there were two processes in which we could do this. Either we could clean the data from the source, which is the excel file, or we could transform the data into Power BI before loading it in for analysis. Ideally speaking, cleaning the data from the source is the best practice, and would often be recommended considering that this allows anyone who wants to analyze the source data to access it in an already cleaned format and takes away future repetitive cleaning work. For the purposes of this project, we want to showcase our ability to use ETL processes from Power BI, and as such, we transform the data here.

Looking out for inconsistencies, one thing we could immediately spot is the misspelling of Coffee as “Cofe” in the Products tab. We use the replace values function to locate all instances of “Cofe” and replace it with “Coffee” so it matches all the other values in the same column that have “Coffee.” Additionally, we see that “Bean World” and “BeanWorld” are considered two different values when they should be the same. We can change one for the other, and I ultimately went with normalizing the values as “BeanWorld” because another value called “PastryPro” exists without a space. Ideally speaking, I would know whether Bean World should be listed with or without a space and work from there, but given what I know, I believe I made the correct choice and would double check with peers in the company Brew Heaven to confirm as such.

The last step before loading the data is creating the necessary relationships between tabs. Power BI was intelligent enough to already determine how each tab was related to each other and developed the relationships as follows:

- Inventory is one-to-one with Products, where Inventory’s Product ID aligns itself with Products’ Product ID.
- Products is one-to-many with Orders, where Products’ Product ID aligns itself with Orders’ Product ID.
- Orders is many-to-one with Customers, where Orders’ Customer ID aligns itself with Customers’ Customer ID.
- Customers is one-to-one with Feedback, where Customers’ Customer ID aligns itself with Feedback’s Customer ID.

Some things to note is that “Inventory” and “Products” both use “Product ID” as their primary keys while “Orders” uses “Order ID” as the primary key and “Product ID” as the foreign key. Additionally, “Customers” uses “Customer ID” as their primary key while “Orders” and “Feedback” both use it as foreign keys while “Feedback” uses “Feedback ID” as their primary key.

Once all the data has been appropriately transformed, we load this version of data to Power BI and work with this data and the established relationships between datasets to create the dashboards required for our task at hand.

## Creation of Dashboard for Project

First, we begin with the Sales page. According to the task, we needed to produce the Customers table, show the evolution of sales over time, show the distribution of sales across products and categories, and showcase the overall sales revenue. To produce the Customers table, there is a table visual that we can use, and from there, we simply link all the columns in the Customers dataset to the new visual table in the dashboard. Second, to show the evolution of sales over time, it goes without saying that most time-based data works best when using line charts, and as such, we create a line chart with the x-axis marking the dates of sales and the y-axis marking the total sales of each day.

Anything related to distribution can be represented by either a histogram when marking the counts of quantitative values or a bar/column chart when marking the counts of categorical data. Since we want the distribution of products and categories, we are using a bar chart. A column chart would have been fine, but a computer is naturally wider horizontally rather than vertically, and we will not have that many visualizations in the Sales page that will necessitate a small vertical column chart. From here, we have a choice between a stacked bar chart, a cluster bar chart, and a 100% stacked bar chart. The 100% bar chart is good when we want to focus only on proportions and percentages but doesn't help with giving context for absolute values. A cluster bar chart can help make side-by-side comparisons within distinct groups, but what we are trying to do is represent the data in a way where two groups are represented, but the groups showcase their own subgroups, since the Categories are made up by the individual Products, so the stacked bar chart best represents this. Lastly, for the overall sales revenue, a card that shows this information in one number works best to represent the information.

Next, the Management page. We need to produce an overall rating, the feedback table, the distribution of stock level by category and the distribution of restock level by supplier. Just like the overall sales revenue was best represented by a card, the overall rating is best represented by a card as well. One thing that had to be considered when creating this, however, is that when it comes to quantitative data being aggregated and summarized, the default method is to find the sum of all the quantitative values that apply to each group. This is something that will not work to properly represent ratings, as they are a system that exists between the values 1 and 5. As such, it is important to adjust the summarization of the values to refer to averages rather than the sum. The Feedback table is also created in a similar way to the Customers table in the Sales page, just with the information from the Feedback dataset.

As for the two distributions of stock level and restock level, these are distributions of categorical data, so we would use bar/column charts to represent the distributions. While I was inclined to use a bar chart due to the more horizontal nature of a computer's display screen, I understood that I would be working with multiple charts, so there were two options to consider if I used bar charts. The first option is that I could show the bar charts side by side to each other, but then that would take away the opportunity to better compare the stock and restock levels on a proportional basis. The next option would be to have the bar charts with one on top of the other so that it is easier to compare values, but that would result in the width of both bar charts being unnecessarily narrow, which not only looks awkward and disengaging, but also may make it harder to read for some. Both options did not seem good enough, so I considered column charts and ultimately decided that two column charts positioned side by side were the best fit for our task. Not only does it address the visual issue of the width of the columns, but it also provides an opportunity to observe side-by-side the stock and restock levels from a proportional perspective for each and intuitively notice possible correlations.

## Formatting of the Dashboard for Project

When you first create a visual on Power BI and leave it alone as is, there is a standard size to the visual that does not cover all of the page, a standard color/theme that won't automatically accommodate for any visual impairments, a title created based solely on the columns you chose to represent through your visual, and no other subtitles or axis titles for additional context. This is where personal formatting comes into place to make sure that the overall page itself is the most presentable and accessible.

When deciding to create the dashboard, one of the best practices is to have the visualizations themselves in the spotlight, and anything heavily text based such as tables and cards should be big enough to show what needs to be shown, but not so big that it overtakes possible space that could be used on the visualizations themselves. For the sales page, there are five parts: The Text Box containing the Title of the Page, the Card, the Table, the Line chart, and the Bar Chart. I decided that the three text-based items would take up about a quarter of the space to the left, while the two visualizations would take up the rest of the space. Within the left quarter of the page, there was enough space to properly show the Title box's and the Card's contents, but not enough space to show ALL of the contents of the Table. The contents of the Email column, the last column on the right, doesn't fully show the emails, but it does show the first part of the email before the "@" symbol, and it is understood that the next part after is "example.com". There is also a feature that allows you to scroll through the table horizontally to see this for yourself, and since this is not always information that a shareholder would need compared to the information presented in the other items, it felt appropriate enough to leave it as is rather than make it large enough to show all of the information and take up more space from any of the other two visuals on the right. From here, we just need to adjust the sizes accordingly to make sure all of the space is properly used by each item throughout the page.

The color/theme of a visualization should take the following into consideration: the possibility of color-blindness affecting the way a person is able to perceive the visualization itself, the connotation behind color and shading that can be used to further understand the information communicated by the visualizations, and the branding of the stakeholders along with their natural color scheme to allow the data to further connect with the audience. As such, the standard color scheme does not work for us because the shading of the visualization is way too similar to the point where it will not show itself well enough for those who are color-blind. Fortunately, there are other themes that take this into consideration, listed as the "accessible" versions of classic Power BI themes. Taking the stakeholders, Brew Heaven, into consideration, one thing that was brought up was that this is a local coffee shop. I reside in the urban area of Los Angeles, and as such, any themes that have an association with urban life would reasonably pair well with the sensibilities of the local coffee shop on my end, (unless the coffee shop explicitly communicates that there is a specific angle in their branding that they are attempting to capitalize on that skews away from the general vibe of the area it is located in, which it doesn't appear to do). As such, the accessible version of the theme "City Park" will work for taking the branding of

the coffee shop into consideration. Lastly, any connotation within differing colors does not 100% apply here for the following reason: There is no obvious inherent hierarchy between the individual products portrayed in the bar chart that should be signified by their colors, and all the colors used in the theme don't have inherent connotations when paired with the products. This is an aspect that you could get away with without taking into careful consideration if you need to create visualizations as quickly as possible. There are ways to more deeply consider this aspect of formatting the color scheme that we will get into later.

The titles were changed to better communicate what each item represented in more concise manners. There were instances where I felt there was some redundancy, such as the card having the title and the text underneath the value that essentially shared the same information. I could have left it without the title, and in my personal opinion, it would have looked better that way for the card itself, but if all the other items had titles of a specific format while the card did not, it would have stuck out in a way that was rather distracting, regardless of whether or not the individual presentation of the card was more aesthetically pleasing. As for the axis titles, while it did seem rather unnecessary on an intuitive level, it does help with clarifying the context of the values within the axes. For instance, the y-axis of the line chart normally shows only numbers, which we understand intuitively as the revenue figures. It doesn't illustrate which currency these values have, which is important to understand contextually. Another approach would have been to have this stated in the title itself and remove the axis title completely, but abusing this ability would lead to titles that may be more wordy than necessary just to give the proper context. As such, it is a good practice to add axis titles and make sure that it gives some context to your visualization without redundancies appearing. In addition to adding the axis titles with the specification of units representing US Dollars, I also made sure to format any other sales data to a currency format to make sure that the representation of sales data was consistent and well-understood.

For the management page, there were also five items: the title, the table, the card, and the two column charts. I felt that, unlike the Customers chart from the previous page, it was important to show more of the information of the Feedback chart to the Stakeholders, so I sized the Feedback chart to show most of the information, and doing so resulted in a long table that spanned most of the length of the page itself. As such, instead of taking the same approach where the text was all listed on the left and the visuals were on the right, we had all the text items show itself in the top quarter of the page above with the two column charts showing itself in the remainder of the space below. The theme selected applies to all of the pages in the dashboard, and as such, the color scheme transfers to the column charts appropriately. The key difference being that the colors of each column remain the same since the data being represented is the categories and the suppliers in separate charts. There aren't any subgroups recognized within these groups, such as the products in each category or supplier, similar to the bar chart on the Sales page.

## Insights of Dashboard from Project

The Sales Page:

- The business has five customers with emails on record.
- The total revenue thus far is $302.50
- The trend of sales over time fluctuates wildly but does not show any sign of overall growth or decay.
- While in terms of categories, Coffee provides more revenue than Pastries, in terms of products, the only Pastry item, the Croissant, brings more revenue than all other Coffee items.
- Out of all the Coffee items, the Latte brings in the most revenue while the Americano brings in the least revenue, with the Cappuccino and Espresso bringing in relatively the same amount of revenue in the middle.

The Management Page:

- Feedback shows mostly 4 and 5 ratings with one 3 rating.
    - The 4 and 5 ratings discuss coffee and the overall experience, with the 5s discussing excellence in the coffee and service and the 4s discuss how the space was a bit crowded or that the coffee was somewhat cold.
    - The only 3 rating discusses dissatisfaction with the pastry not being as fresh as expected.
    - The average overall rating is 4.2
- The Stock level and Reorder level are SUBSTANTIALLY higher for coffee than they are for pastries.
- The stock and reorder levels from an absolute perspective appear to have a ratio of about 2:1. This is something not exactly found by looking at the comparisons side by side visually, but rather by observing the tooltips associated with each column and comparing from there.
- Looking at the two charts visually, you can see that the ratios between Coffee and Pastries remain relatively the same within Categories and Suppliers.

Additional Notes:

There are some questions on the data that, in a real-world setting, I would ask before beginning with the data analysis that would further it down the line. First, there are only 5 customers we have on record, and they appear to have emails linked to their information. Are these the only five customers we have had ever, or was there a loyalty program and this information comes from this? Second, the data accounts for orders made, but it does not properly account for the lack of orders throughout the day. If there was a day when no products were sold and the store was open, this would be information that we would need to consider for analysis, and yet, it is not here. Do we have that information available? It

would be easy to make the adjustments necessary to show how daily performance went about because we only have a small dataset to work with, but in a real-world setting where there is a very large dataset, this would be the responsibility of a data engineering and management team to handle. Lastly, do we have information on the costs of the goods we are selling? This would help to understand things not just in terms of revenue but also in terms of profit, which is what I believe a stakeholder would probably care more about.

That being said, if all we cared about was revenue, there is one insight I gained not through the individual pages but by looking at both pages together. The pastries account for about a tenth of the stock and restock levels and yet bring in a little under a third of the overall revenue. If we wish to increase the revenue, here may be something to consider: The pastries aren't the focus, evident with the relatively low stock level, low restock level, the negative feedback of the pastries, and a rating of 3 out of 5 associated with this commentary, and yet, it is able to bring in almost a third of the revenue. If we gave more focus on the pastries by increasing the stock and restock levels by adding more pastries to the menu, made more of an effort to provide fresh pastries, and focused on marketing and selling the fresh pastries, we can expect the increase of revenue to be very substantial.

## Dashboard Revisions

With this, the tasks of the project itself have been completed. The task also called for the addition of a title, and I wasn't sure if they meant adding a title for the Sales and Management pages or adding a title page, so I created a title page just in case. A title page that simply gave the title of the Dashboard itself felt vapid, so I decided that I would add a stock image of coffee as a background, which would help connect the dashboard to the stakeholders of Brew Heaven. In doing so, the text box containing the written title would be in front of the picture, and instead of having a white or other colored text box slapped within the picture, it would be more aesthetically pleasing to have the background of the text box hidden so that all that appears is text in front of the picture itself. After that, I needed to select a color for the text to both match the brown background of the selected stock image and set the tone for the colors to anticipate for the rest of the dashboard. White was not going to work, so I went with the most common color in the first go around, mustard yellow.

From there, I decided to rework the dashboard for two distinct reasons. First, I decided to rework the color scheme of the dashboard as if the branding colors of Brew Haven were brown and mustard. Having the first color scheme works for all intents and purposes, as mentioned before, but I wanted to give more effort into considering the presentation and branding, so this gives me the opportunity to do just that in a way that maintains good visualization practices. Second, I decided that the information presented in the two column charts of the management page could be reworked into one clustered column chart with each group representing a different product matched with their respective supplier, and each column in each group would represent the stock level of the product, the quantity sold of the product, and the restock level of the product. I wanted to find more context with the restock level by not just comparing how much revenue the sold products

bring but also how many individual items were sold, and whether the restock level works well with resupplying our stock without much excess.

As such, I first approached the Sales page and individually changed the columns color by individual series in the products. The color selection recommendations are provided according to the series selected, and it showcased a series of five shadings for brown/mustard, so I was able to use them to represent one product each. The shading selected corresponds to the natural shading of the product you would normally see. Espressos tend to be very dark and very strong in flavor profile, so I selected the darkest shade to represent Espressos. Cappuccinos are often topped with a generous amount of whipped cream, and as such, the lightest shade represents Cappuccinos. The Croissant, not only typically being a yellow color, seemed to feel as if it should be representative of the whole pastry category, so giving it the most neutral shade felt the most appropriate. Lattes are often creamy in color, so the second lightest shade represents Lattes, and that leaves the Americano with the second darkest shade.

After making the adjustments for the Sales Page, I made the adjustments for the management page next. The two column charts were replaced by the clustered column chart with five groups of three columns. The colors were attributed not to the intersection of the product and the quantitative variable measured (i.e. the stock levels, restock levels, and quantity sold), but rather by the quantitative variables themselves. All the stock levels were of the same color across the clustered columns, and at the present skill level and understanding, I was not going to be able to change that within a reasonable amount of time, so I simply decided to just make the color alterations by variables. Having an understanding that in this business, first, you have an item to sell (stock level), then the item is sold for profit (quantity sold), and then the stock is replenished for more sales (restock level), I ordered the columns to represent this order of operations, and provided it with the appropriate shades from darkest to lightest. I could have selected a series of the three closest shades for the sake of simplicity and aesthetics; but considering that a higher contrast between colors would help in the accessibility for those with color-blindness, I elected to use the darkest, lightest, and most neutral shade to represent the three columns in each group.

## Insights from Dashboard Revision

After gaining more of an understanding of how many items were sold and comparing the quantity sold to the restock levels, I see that our restock levels for coffee products are twice as much as the quantity sold, and our restock levels for pastries are lower than the quantity sold. This tells us two things. First, if we assume that we are satisfied with the current stock levels of our coffee products, which we should consider that we have our quantity sold at about a fourth of our current stock level, there is an imbalance in terms of how much we spend on replenishing our coffee stock. Unless we anticipate a huge boom in coffee sales soon, we should not be getting more coffee products than we sell and take up more inventory without the proper turnover. As such, we should lessen how much we

spend restocking on Coffee products, especially Espresso and Americano, which shows that the restock level is almost double the quantity sold. Additionally, if we are to use the recommendations I made in the insights from the original project and give more focus towards pastries, our stock level is barely above the quantity sold, and as such, we need to increase our restocking of pastries, whether that is solely the croissant or adding additional products to our menu.

## Reflections on the Project

With this, the essential part of our project has been completed. As stated earlier, the purpose of this project is to demonstrate Power BI skills, ETL processes, data visualization practices aiding in stakeholder experience, and creating insights based on visualizations to aid in the executive decision-making of stakeholders. I feel that this project does exactly that, so I would consider this to be a success.

There is always room for improvement, however. One thing that should be addressed is the missing data. I am fairly aware of the fact that the Orders table only takes into consideration the orders that were made for specific products, but that unaltered information by itself does not provide us with the daily performance of products. As such, I do feel that the line chart presented in the dashboard is inaccurate, and what I would do to improve this is create a dataset that examines the performance of each product in each day given the information from the orders table and any tables related to that. From there, I would base my visualizations on this new dataset that focuses on presenting daily product sales and even create a clustered line chart where the trends of product sales over time are presented all at once to visualize and compare when certain items did well/poorly. Perhaps I could go about creating the dataset through certain Excel or SQL functions that aggregate the performance of each day from the Orders dataset. This may be something to consider as a means of expanding upon this project, but for now, I wished to demonstrate my ability to use Power BI for insights, follow through a guided project, and even find ways to improve it from a pure visualization standpoint.

Why did I choose to complete this guided project? Coursera presented this guided project as a project that you can complete if you were creating a project based on Power BI for the first time. While I was able to complete a course on Power BI that gave me the essential knowledge not just on how to use the tool but also how to consider ETL processes and stakeholder experience, there was not a full-on project that would help to demonstrate these acquired abilities from the class alone. The class is part of a group of classes that showcase a specialization on Power BI, and the last class is a project that considers all that I have learned from taking all the other classes in the specialization. While I am looking forward to completing those classes in the future and furthering my understanding of Power BI in Analytics, I wanted to take up a simple project that would be able to demonstrate that I am capable of using Power BI for analysis, even with beginner-level understanding, and this seemed like the best course of action to showcase this. It was fun not only completing the task at hand but also putting myself mentally in the position of a

data analyst providing insights for a local café. I look forward to learning more about Power BI, learning more skills along the way, taking up and even making up more projects that show more of my capabilities, and even helping companies in the future by giving insights to aid in executive decision making.

Last updated: 07/24/2025