**Global Sales Data Analytics**

|  |  |
| --- | --- |
| Data | 09-11-2022 |
| Team ID | PNT2022TMID45267 |
| Project Name | Global Sales Data Analytics |

**Understanding The Dataset**

6 STEPS TO ANALYZE A DATASET

1. Clean Up Your Data

[Data wrangling](https://online.hbs.edu/blog/post/data-wrangling)—also called data cleaning—is the process of uncovering and correcting, or eliminating inaccurate or repeat records from your dataset. During the data wrangling process, you’ll transform the raw data into a more useful format, preparing it for analysis.

It’s imperative to clean your data before beginning analysis. This is particularly important if you’ll be presenting your findings to business teams who may [use the data for decision-making purposes](https://online.hbs.edu/blog/post/data-driven-decision-making). Teams need to have confidence that they’re acting on a reliable source of information.

2. Identify the Right Questions

Once you’ve completed the cleaning process, you may have a lot of questions about your final dataset. There’s so much potential that can be uncovered through analysis.

Identify the most important questions you hope to answer through your analysis. These questions should be easily measurable and closely related to a specific business problem. If the request for analysis is coming from a business team, ask them to provide explicit details about what they’re hoping to learn, what they expect to learn, and how they’ll use the information. You can use their input to determine which questions take priority in your analysis.

3. Break Down the Data Into Segments

It’s often helpful to break down your dataset into smaller, defined groups. Segmenting your data will not only make your analysis more manageable, but also keep it on track.

For example, if you’re attempting to answer questions about a specific department’s performance, you’ll want to segment your data by department. From there, you’ll be able to glean insights about the group that you’re concerned with and identify any relationships that might exist between each group.

4. Visualize the Data

One of the most important parts of data analysis is [data visualization](https://online.hbs.edu/blog/post/data-visualization-techniques), which refers to the process of creating graphical representations of data. Visualizing the data will help you to easily identify any trends or patterns and obvious outliers.

By creating engaging visuals that represent the data, you’re also able to effectively communicate your findings to key stakeholders who can quickly draw conclusions from the visualizations.

There’s a variety of [data visualization tools](https://online.hbs.edu/blog/post/data-visualization-tools) you can use to automatically generate visual representations of a dataset, such as Microsoft Excel, Tableau, and Google Charts.

5. Use the Data to Answer Your Questions

After cleaning, organizing, transforming, and visualizing your data, revisit the questions you outlined at the beginning of the data analysis process. Interpret your results and determine whether the data helps you answer your original questions.

If the results are inconclusive, try revisiting a previous step in the analysis process. Maybe your dataset was too large and should have been segmented further, or perhaps there’s a different type of visualization better suited to your data.

6. Supplement with Qualitative Data

Finally, as you near the conclusion of your analysis, remember that this dataset is only one piece of the puzzle.

It’s critical to pair your quantitative findings with qualitative information, which you may capture using questionnaires, interviews, or testimonials. While the dataset has the ability to tell you what’s happening, qualitative information can often help you understand why it’s happening.