DATA ANALYSIS DOCUMENTATION

The tools utilized in this case study include Python's Pandas library for data manipulation and exploratory data analysis, and Tableau for data visualization.

Fundamental Takeaways

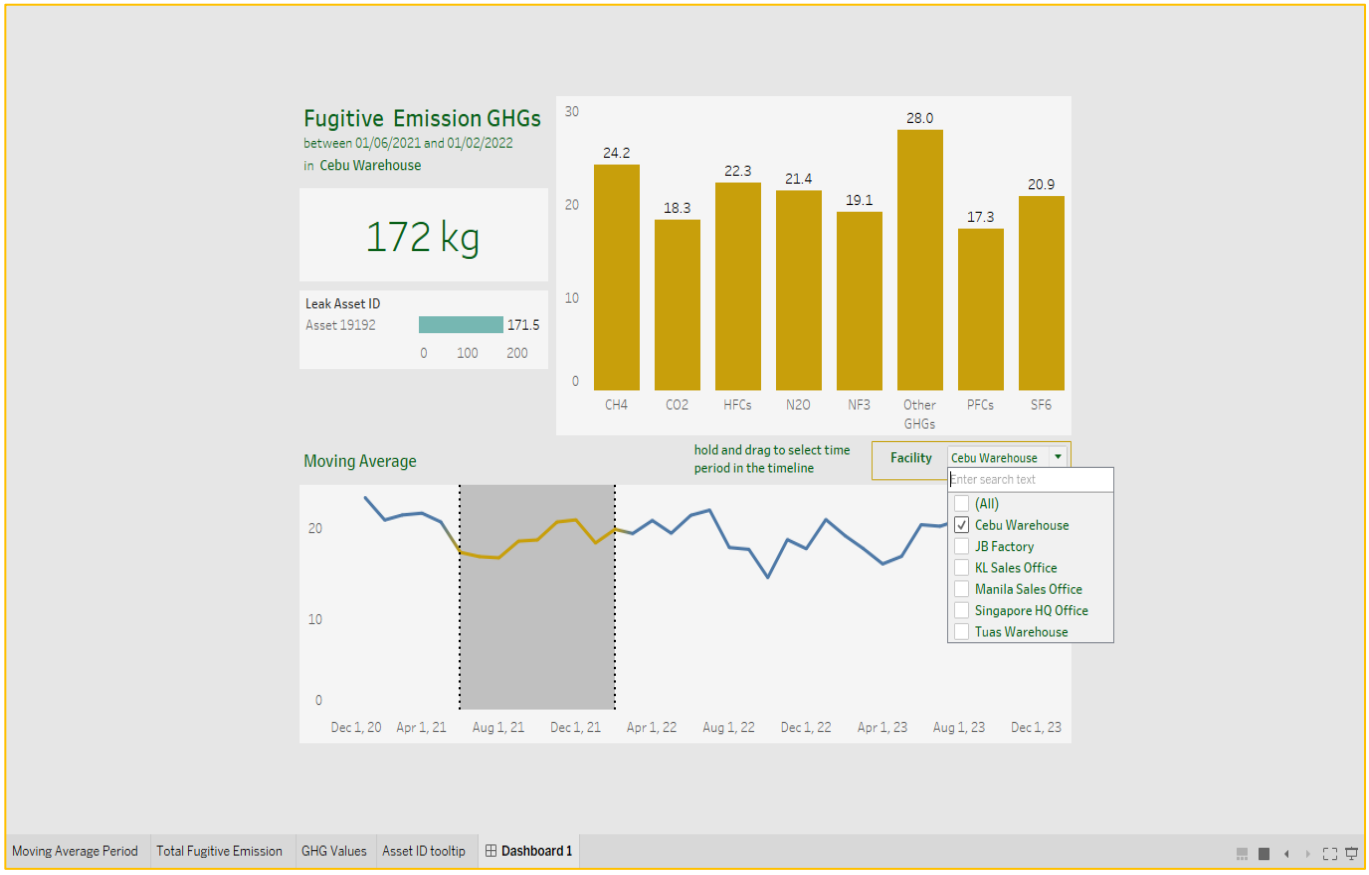
1. **Utilizing the Pandas Library:** During data manipulation, I opted to use Python's Pandas library due to its convenience in combining multiple Excel files efficiently.
2. **Dropping Irrelevant Columns:** I removed columns that were not necessary for the analysis, particularly those with monotonous values such as "Description," "Data Quality Type," "Emission Factor Library," "Is Market Based," "Industrial Process Type," "Leak Source," "Leak Description," and the various GHG unit columns.
3. **Consolidating Date Columns:** Since both the "Consumption Start" and "Consumption End" columns contained identical values, I retained only one, renaming it as "Consumption Period" for clarity.
4. **Excluding the Emission Rate Column:** The "Emission Rate" column was also dropped, as there was no distinct unit associated with it, and I discovered that the total GHG emissions did not align with the emission rate provided.
5. **Focusing on Quantitative Data:** I removed the GHG percentage columns, opting to retain only the actual GHG values to maintain a more quantitative approach to the analysis.
6. **Data Grouping in Tableau:** For visualization and dashboard creation, I used Tableau's pivot function to group all the GHG emission columns, allowing me to work with them simultaneously.
7. **Trend Analysis with Moving Average:** I developed a moving average visualization to uncover trends in GHG emissions over time. A set action was incorporated to dynamically select a specific time range, serving as the main control for the entire dashboard.
8. **Summation of GHG Emissions:** I created a visualization to display the total GHG emissions for the selected time period, providing a clear overview of emissions within the specified range.
9. **GHG Emissions Breakdown by Type:** A bar chart was created to display the values of each GHG, revealing which gas had the largest emission levels.
10. **HVAC Leak Asset Analysis:** Another bar chart was developed to display HVAC leak asset numbers along with their respective GHG emission values, identifying the assets responsible for the highest emissions.
11. **Facility-Based Filtering:** A filter for the facilities was implemented and applied to all visuals, enabling dynamic selection of individual facilities and uncovering insights specific to each one.

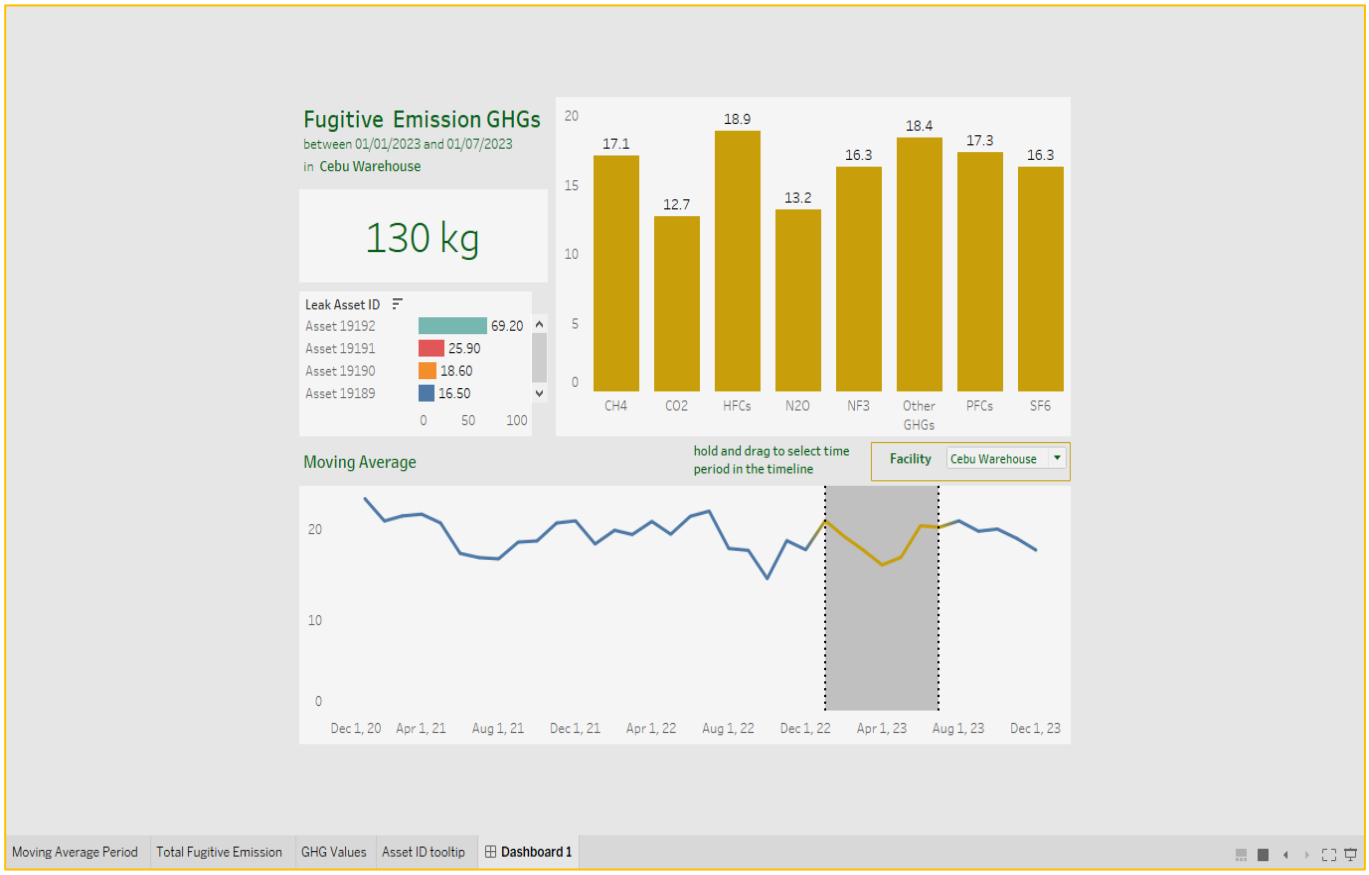
12. **Dashboard Insights:** The dashboard created from this analysis offers valuable insights that align with the objectives and goals set in the case study, providing a comprehensive view of GHG emissions and contributing factors.

(See dashboard control below)

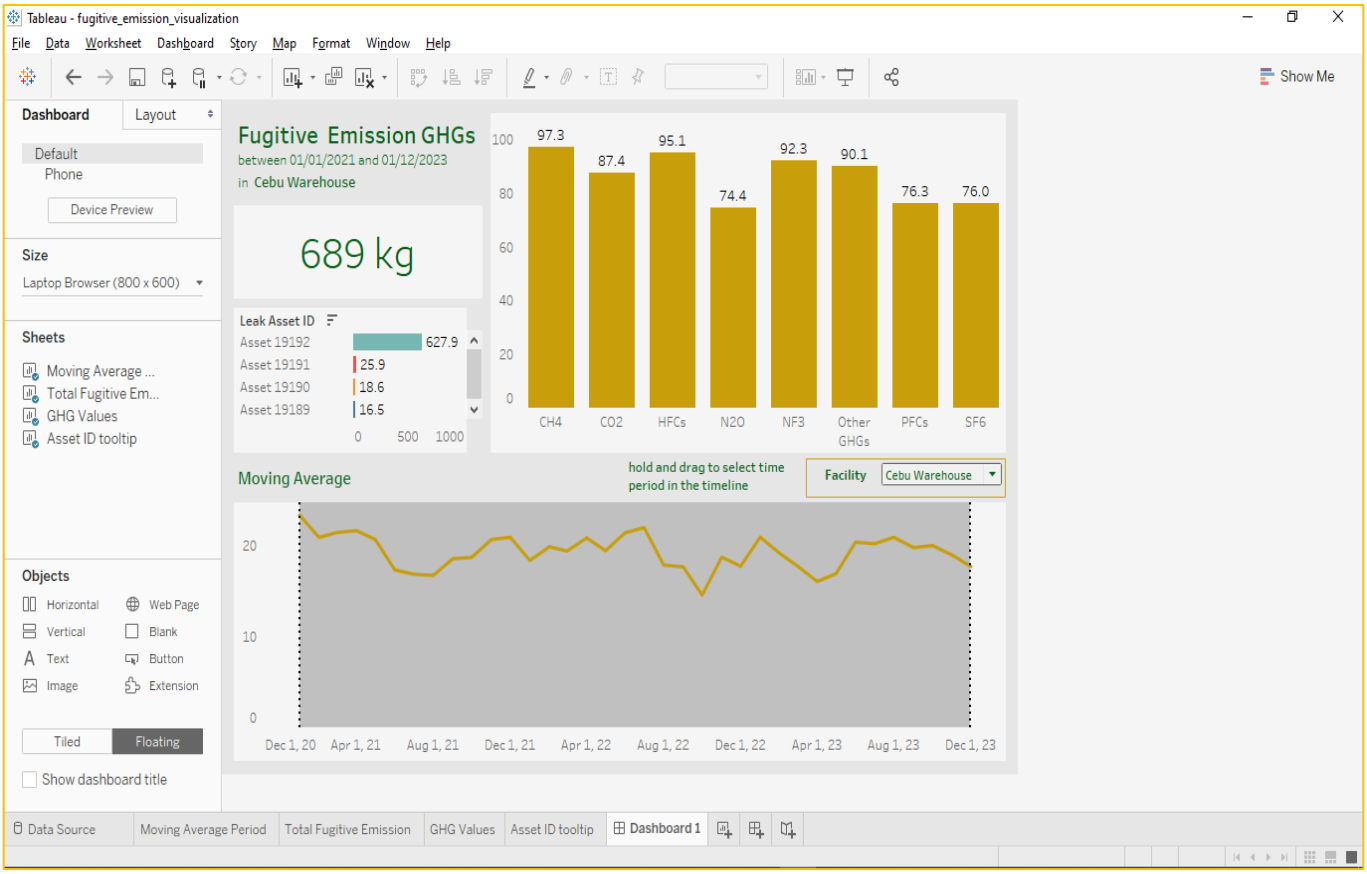
Dashboard Control

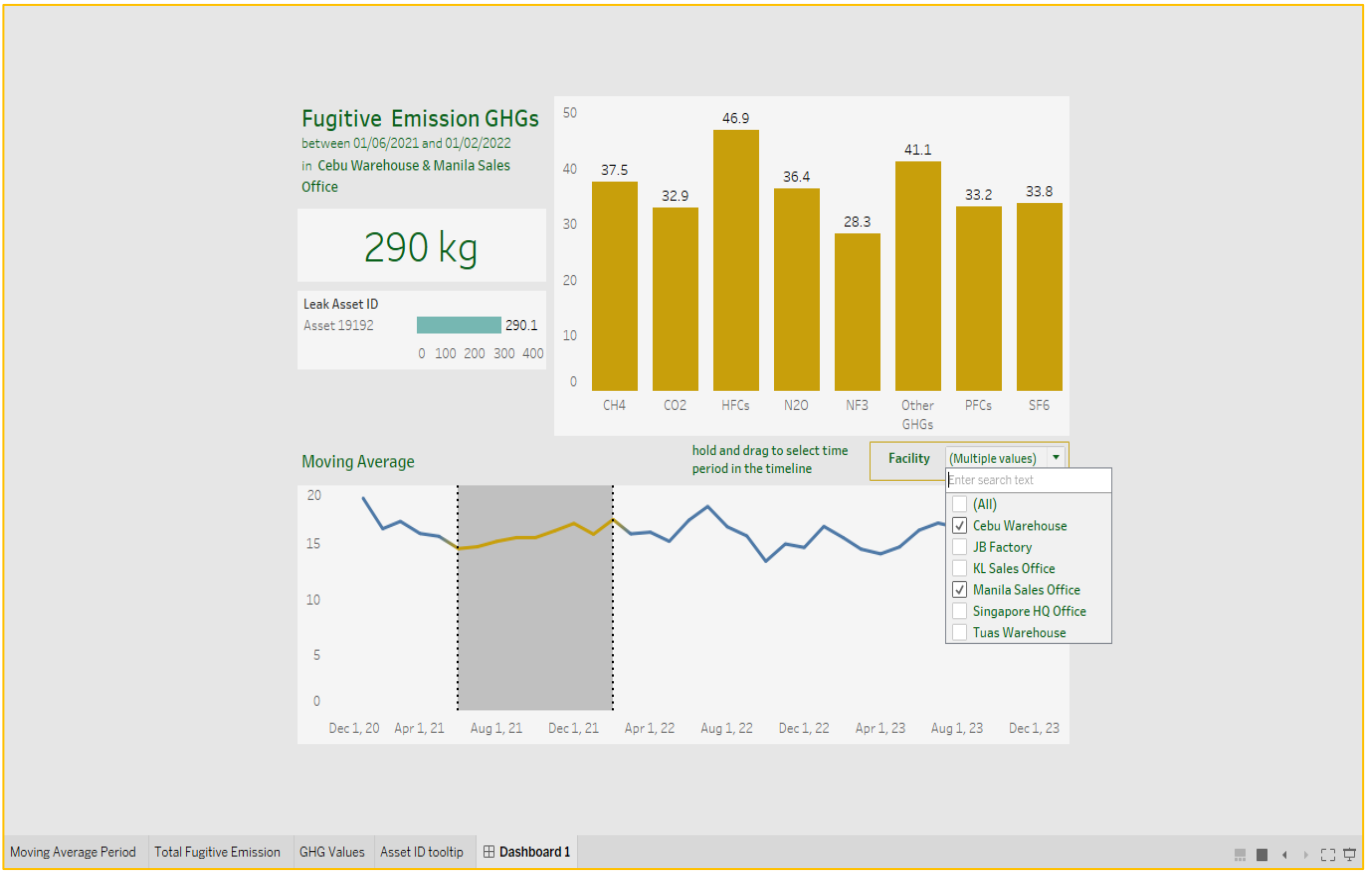
1. **Selecting a Time Period**: In the moving average section, click and drag to select a time range on the timeline. Adjust your selection to the desired time period for analysis.



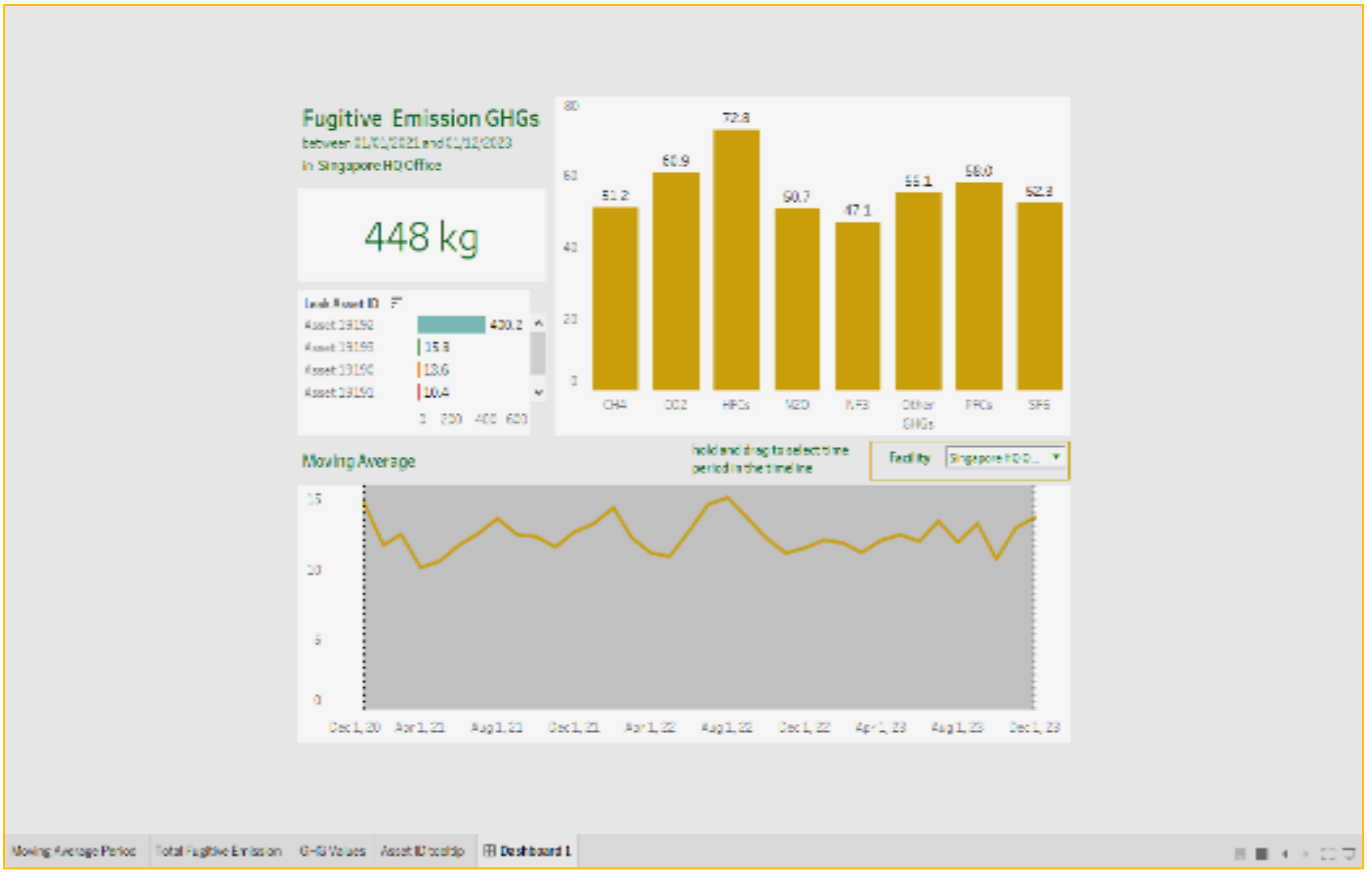


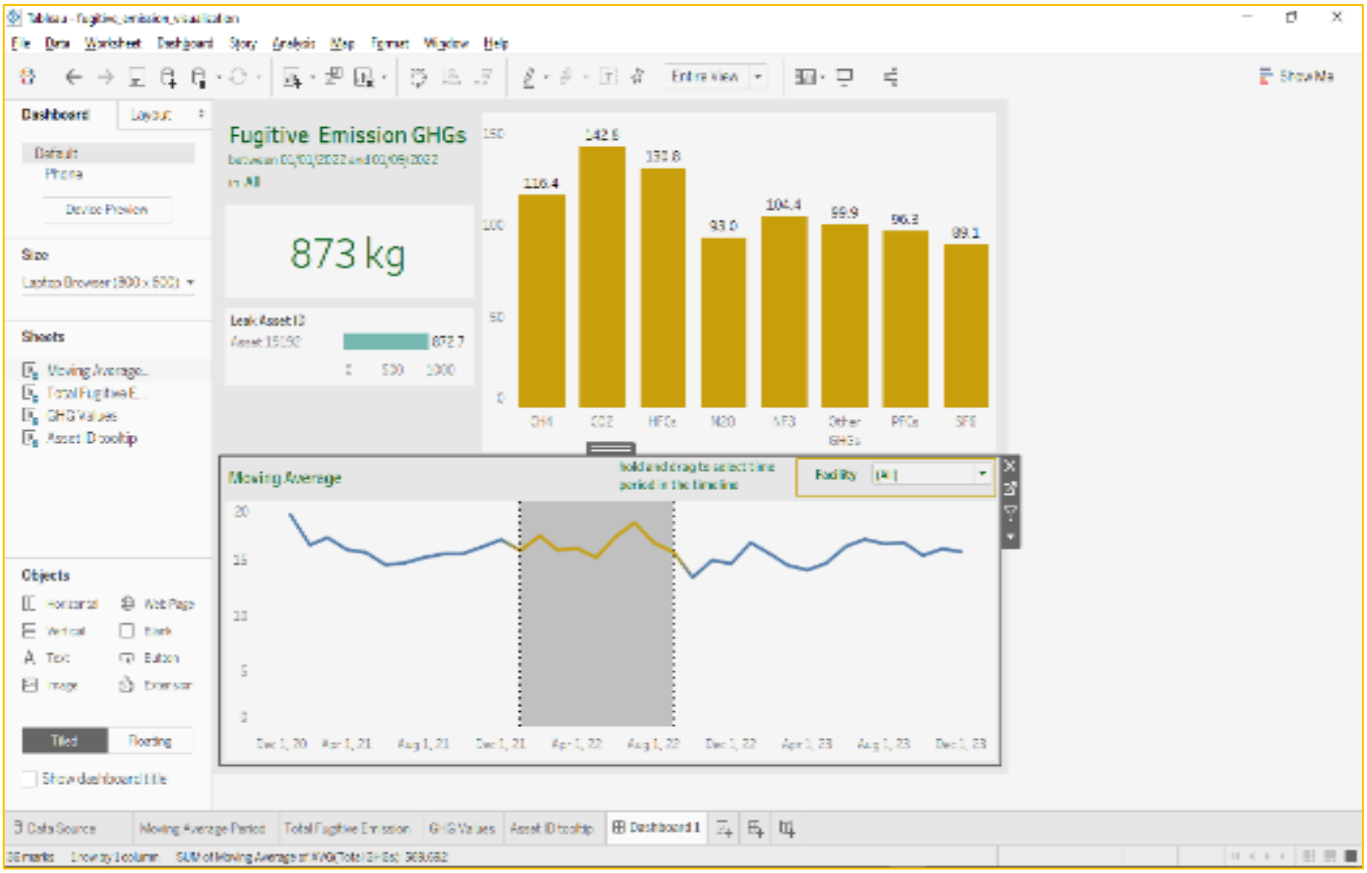
2. **Facility Filter:** Use the facility filter to select the specific facility you wish to analyze. You can also choose multiple facilities for a broader analysis.



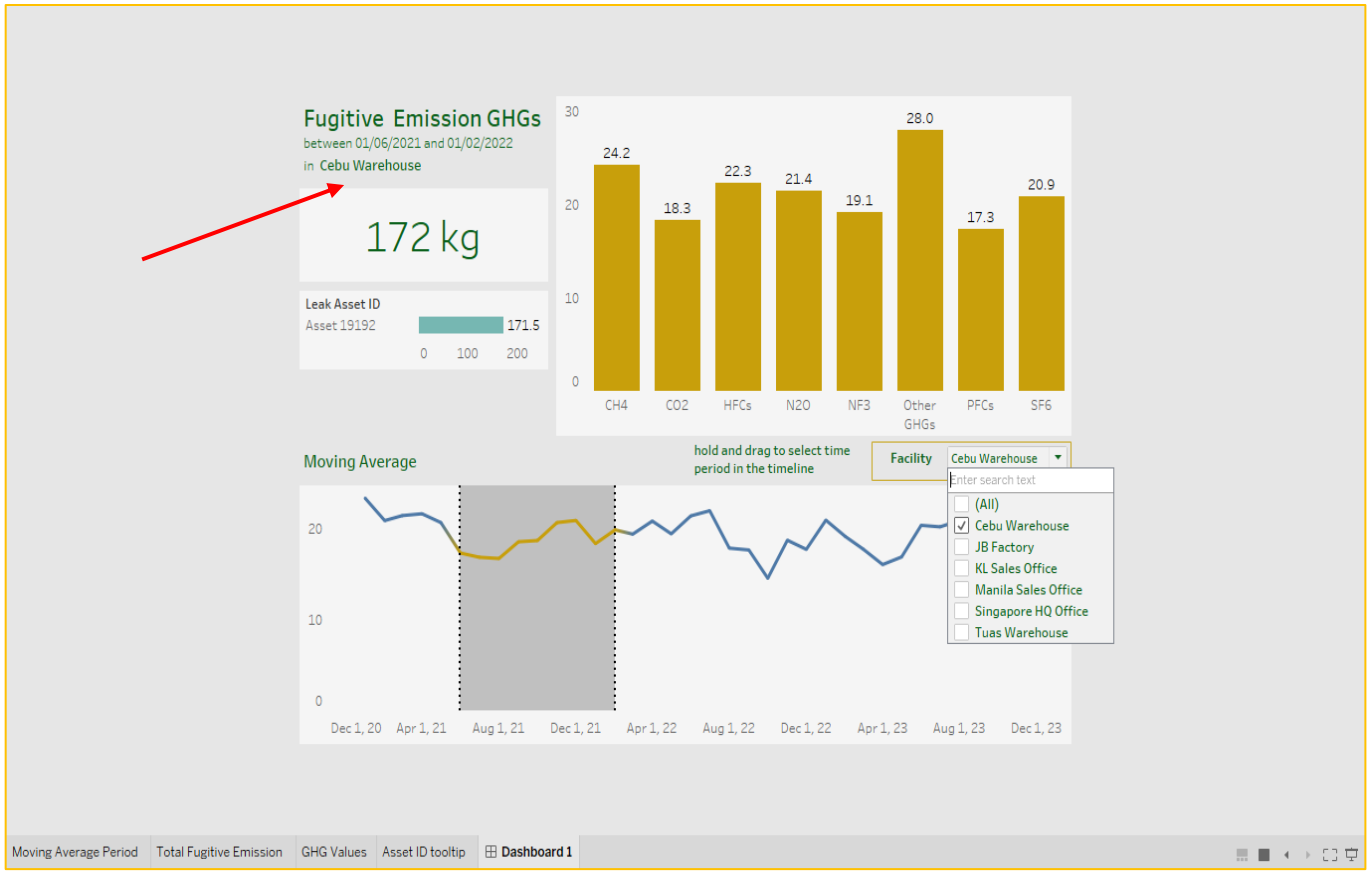


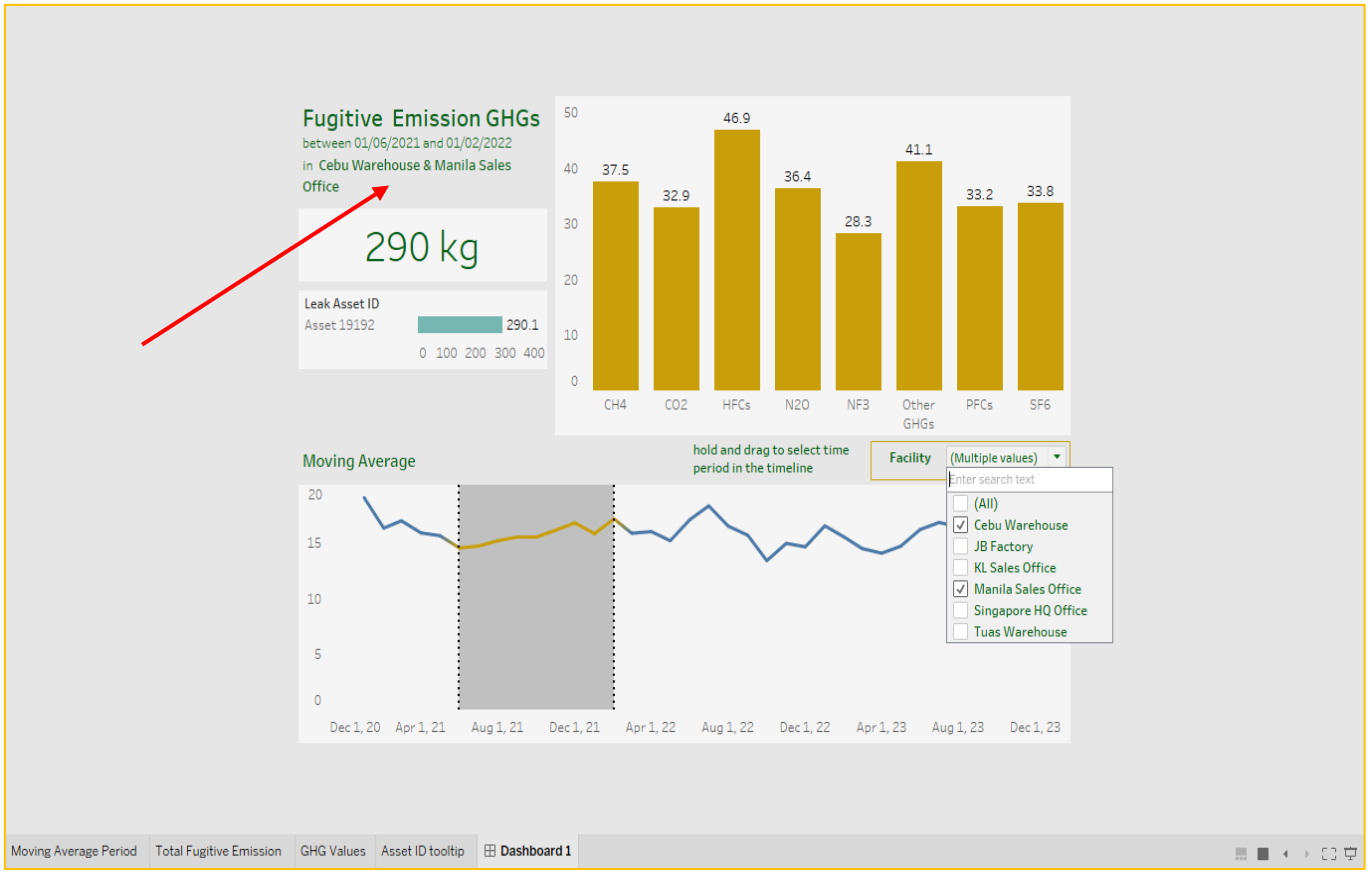
3. **Dynamic Dashboard Control**: Combine the moving average and facility filter controls to dynamically adjust the dashboard, enabling deeper insights and analysis.





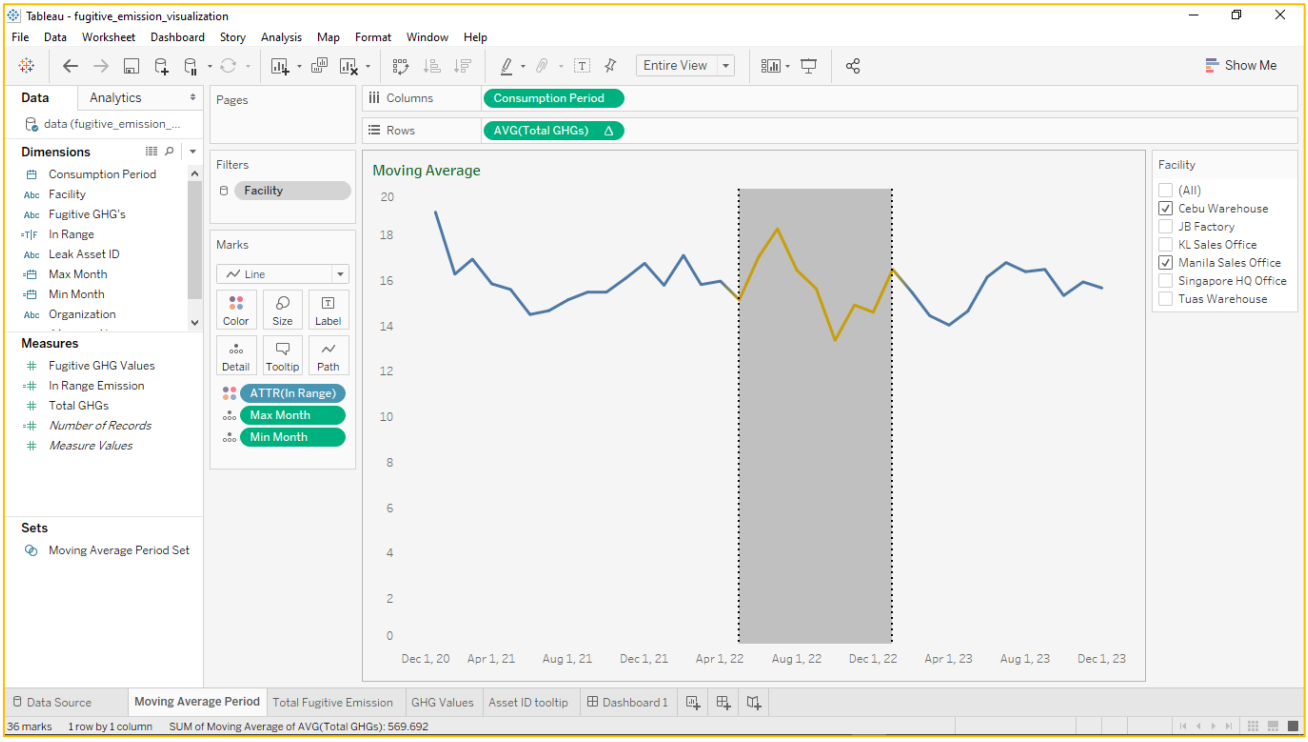
4. **Dynamic Text Display**: The text displays within the dashboard are configured to update automatically based on changes made to the selected controls.



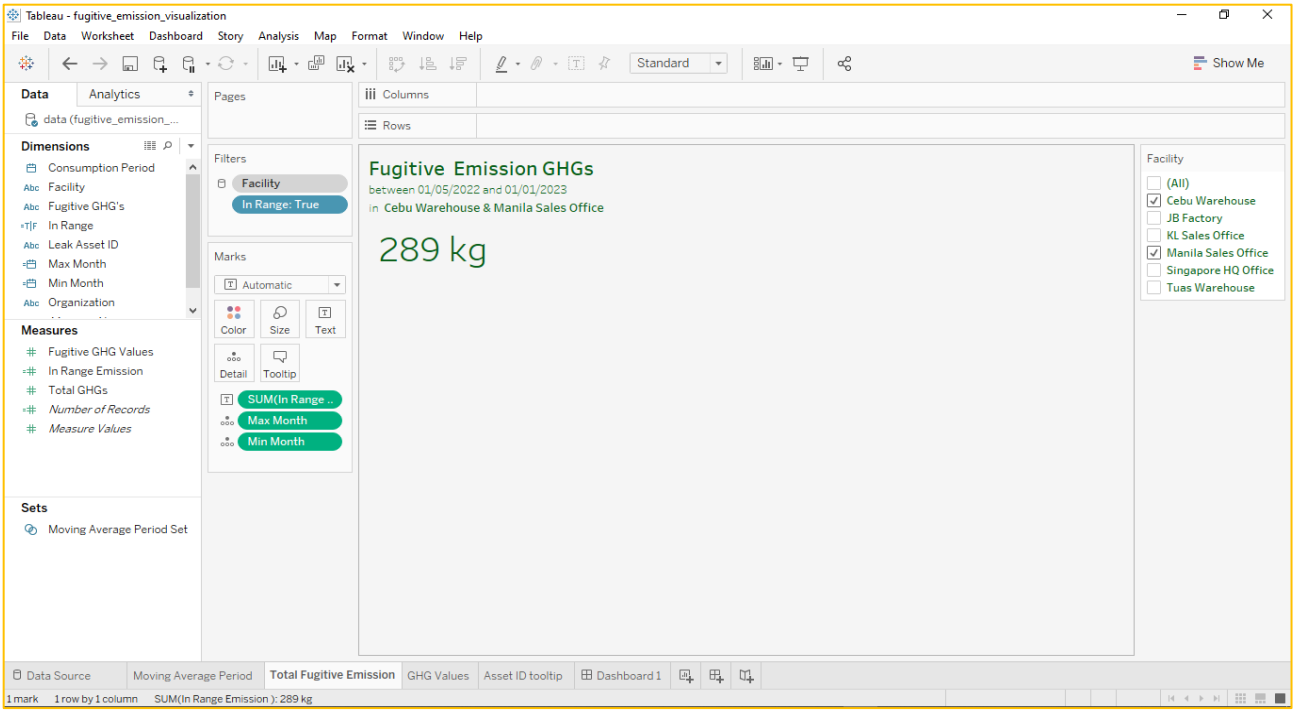


Worksheet Pictures: (TABLEAU)

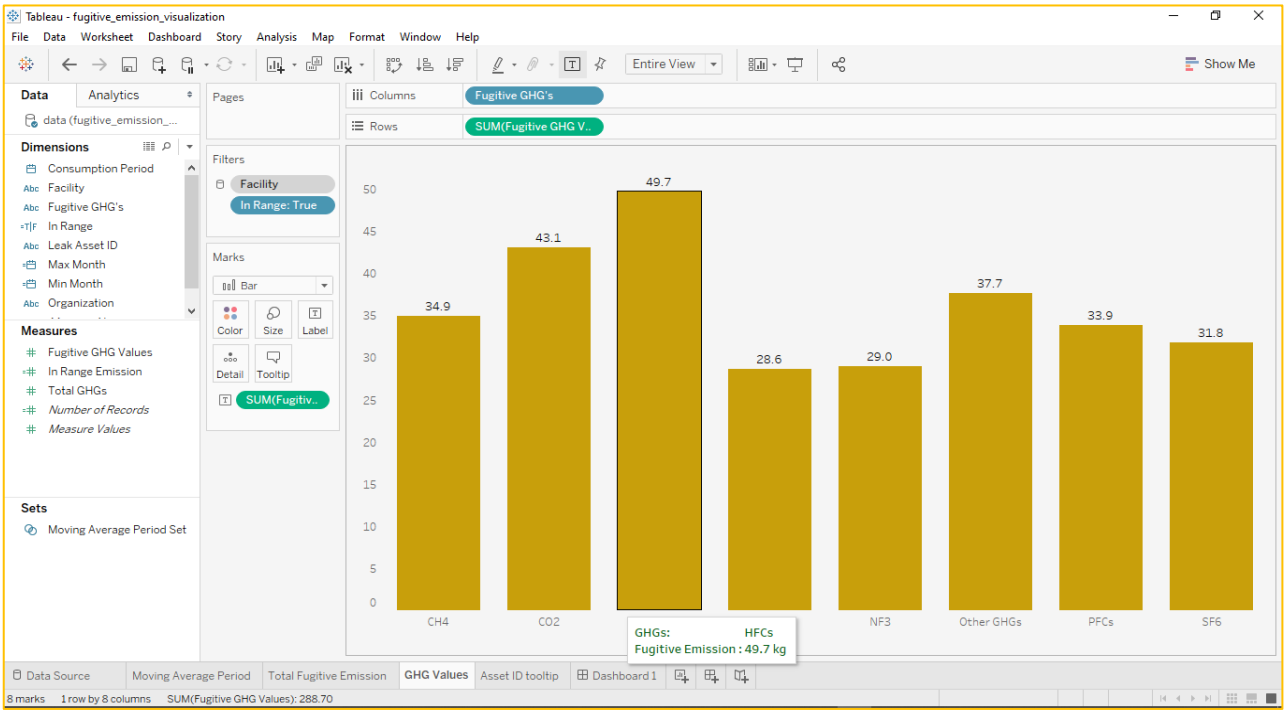
Moving Average Worksheet



Total Fugitive GHG Emission Worksheet



Fugitive GHG Emission Values Worksheet



HVAC Leak Asset Vaues Worksheet