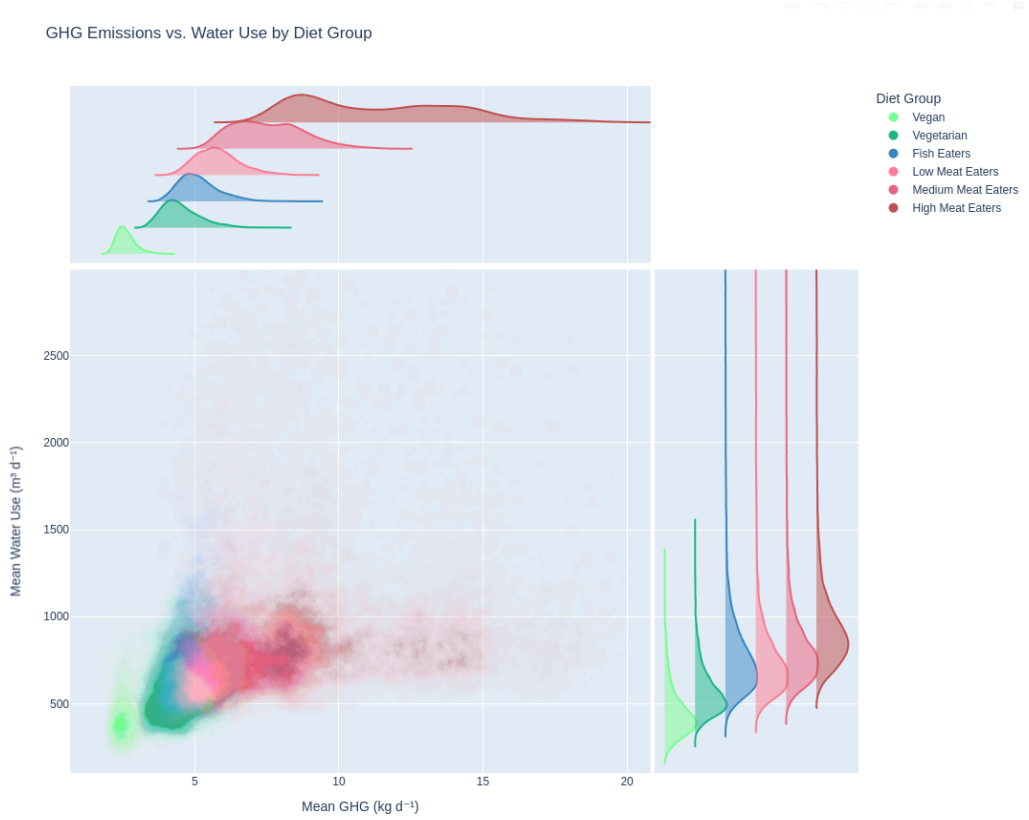
COMP4037 CW2: Visualisation of Environmental Impact of Various Diets

- **Image:**



- **Visual Design Type:** Bubble plot with marginal half-violins.
- **Name of Tool:** Python with Plotly, Pandas, and Jupyter.
- **Diet Groups**: This section shows all six diet groups. You can interactively disable and enable these by clicking on the appropriate legend item.
- **Variables**
    - **Mean greenhouse gas emission per day** (kg/day) Chosen because it is one of the main things the researchers said they were interested in measuring.
    - **Mean water use** (m^3/day) - Chosen because it was one of the least correlated environmental measures with the others, and it showed a slightly different distribution with fish-eaters compared to the other metrics.
    - **Size of the simulated population -** This helps keep things to scale by ensuring that small simulated groups don't disproportionally crowd out the other data.
    - **The dietary group -** Chosen because this is the main thing the study wants to compare the environmental impact of.
- **Visual Mapping:**

    - X-axis: Mean greenhouse gas emission per day (kg/day)
    - Y-axis: Mean water use (m^3/day)
    - Bubble area: Size of the simulated population
    - Color: The dietary group, using a custom colour map (generated through coolors.co), chosing colours that are commonly associated with the diet, while creating a smooth gradient, and ensuring sufficient contrast..
- **Unique Observation:** High meat-eaters (>100g/day) form two clusters: Group A (~13kg/day GHG) and Group B (~8.5kg/day). Group B's water use is slightly higher and resembles medium meat-eaters (50-99g/day). This is shown at 1:22 in the demo video.
- **Data Preparation:** To prepare the data, I simply joined it with a table I made for the different diet groups, doing a dataframe merge. I also shuffled the data points so it wouldn't be biased by, for example, showing specific points on top of others. The diet group is a table containing information about each diet:
    - `key` - the key for the diet group in the dataset.
    - `rank` - the order to have the groups in, chosen to correlate with the approximate overall environmental impact.
    - `color` - the hex code for the colour to draw the bubbles as.
    - `label` - the text to display on the legend of the chart.

- (**Optional**) **URL** to **screen-capture demo:**
  A screen capture demo is available on YouTube: https://youtu.be/WeSi4IhMcQE.
  The video demonstrates the chart's interactivity: viewers can click legend items to show or hide specific diet groups and use zoom and pan functions to explore the distribution in greater detail, reducing visual clutter from overlapping points.
- **(Optional) URL to source code:** https://github.com/cd256/comp4037-cw2
    - '`water_gpg_scatter.ipynb`' - the notebook for creating the chart above.
    - '`water_gpg_scatter.py`' - the Python version of that notebook created by `nbconvert`.
    - 'analysis.ipynb' contains code for creating various charts I used to explore the data.
    - 'inputs/' contains the input CSV files for my analysis, including the Scarborough et. al 2023 results. [1].
    - 'outputs/' contains intermediate outputs, and isn't used to generate the final chart.
    - '`fish.py`', '`fooddb.py`', '`survey.py`' and '`lca.py`' all contain preprocessing for the fish data, the FoodDB database, the resulting dataset found by Scarborough et al [1], and the LCA dataset, respectively.

Discussion

I began processing them to show an overall source, but I struggled to find high-quality diet survey information to merge with the current data. Although the plot is a scatter bubble plot I fould it contained richer and more detailed information on other plots I attempted to create (found in analysis.ipynb and analysis.py), including tree-maps, and violin plots.

References

[1] P. Scarborough *et al.*, ‘Vegans, vegetarians, fish-eaters and meat-eaters in the UK show discrepant environmental impacts’, *Nat. Food*, vol. 4, no. 7, pp. 565–574, Jul. 2023, doi: 10.1038/s43016-023-00795-w.