Climate-Friendly Food Choices: Analyzing the UBC AMS and Food Services Menu for Sustainability

Prepared by:
Vivaan Wadhwa
Sustainable Food Systems Data and Reporting Analyst
SEEDS Sustainability Program

Prepared for:
UBC Food Services, AMS Food & Beverage, and AMS Sustainability

The University of British Columbia
January 3, 2025

Cover Photo: UBC Brand and Marketing
The Spruce / Nisanova Studio
UBC Brand and Marketing

Disclaimer: UBC SEEDS Sustainability Program provides students with the opportunity to share the findings of their studies, as well as their opinions, conclusions and recommendations with the UBC community. The reader should bear in mind that this is a student research project and is not an official document of UBC. Furthermore, readers should bear in mind that these reports may not reflect the current status of activities at UBC. We urge you to contact the research persons mentioned in a report or the SEEDS Sustainability Program representative about the current status of the subject matter of a report.

SEEDS Sustainability Program

Table of Contents

List of Figures

Introduction

The Climate Friendly Food Label Project is an initiative aimed at promoting sustainable food choices within the University of British Columbia (UBC) community. As part of the Climate Action Plan 2030, this project seeks to reduce greenhouse gas emissions associated with campus food services by providing transparency and actionable insights on the environmental impact of different food options. Through the development and implementation of climate-friendly food labels, the project empowers consumers to make informed, sustainable decisions.

During my co-op term as a Data Analyst for the UBC SEEDs Program, I contributed to key aspects of the Climate Friendly Food Label Project. My role encompassed refining baseline emission data, developing methodologies for calculating food-related emissions, and creating data-driven tools to enhance the project's outreach and effectiveness. Additionally, I supported the development of a Plant-Based Dashboard and a Climate Friendly Food Label Desktop Application to streamline the analysis and visualization of emissions data for UBC Food Services.

This report outlines the methodologies, processes, and tools developed during my co-op term. It covers revisions to baseline data, a review of label calculation methods, the architecture of automated processes, and the development of digital solutions for emissions tracking and classification. The insights and tools generated aim to facilitate the ongoing success of UBC’s Climate Action Plan 2030 and support the university’s broader sustainability goals.

Baseline Revision

The baseline calculation method has been rectified to address initial mistakes that resulted in approximate, rather than precise, values. Previously, the baseline was being calculated using the formula:

$$\frac{\sum_{Restaurant} TotalSales_{Restaurant} \cdot AverageEmission_{Restaurant}}{Total\ Sales}$$

Equation 1: Previous Calculation Equation

where the "restaurant" referred to the dining halls: Open Kitchen, Gather, and Feast. This approach was inaccurate and did not align with the original intent of the baseline.

The revised formula now calculates the baseline by considering individual dishes within the restaurants and their respective emissions:

$$\frac{\sum_{Restaurant} \sum_{Dish} TotalSales_{Dish} \cdot Emission_{Dish}}{Total\ Sales}$$

Equation 2: Original Calculation Equation: This was what was intended initially to be the baseline

This adjustment ensures the baseline is calculated with appropriate weighting for each dish's sales and emissions. Following the implementation of the revised method, the baseline has shifted by approximately 5%, with only a few items moving by one color tier and no items have shifted by two color tiers. While these changes are relatively minor, this correction is essential to ensure accuracy and to make sure that the baseline is repress entative of the original intent.

Label Calculations in Review

Dining Hall Calculations

The label was calculated for all three dining halls this year too. The statistics for this can be found below:

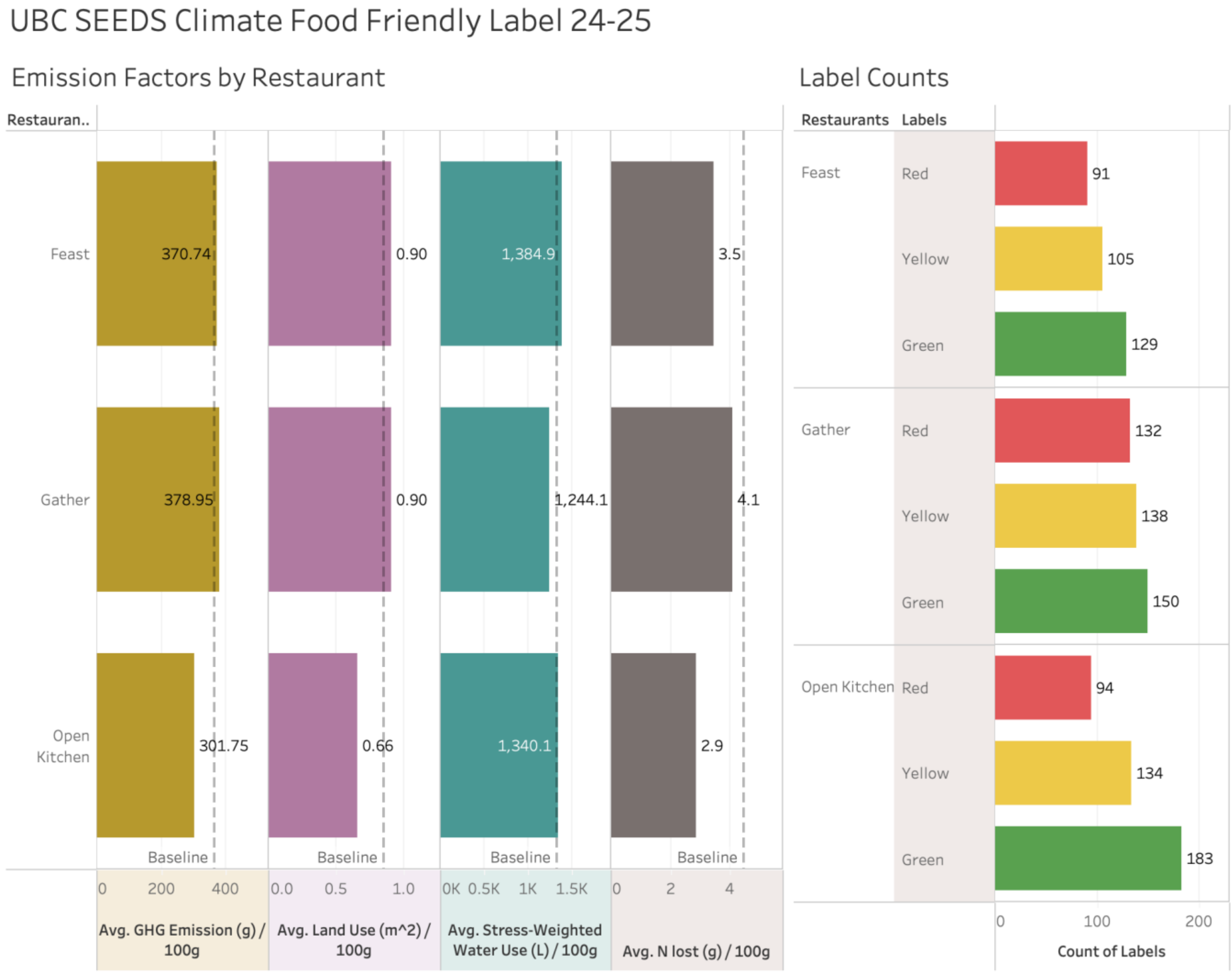


Figure 1: Dining Hall Calculation Summary: Showcases label counts and average emissions values for all 3 dining halls. The dotted line represents the baseline for each factor

All venues offer more green labelled items than yellow or red items.

- Open Kitchen performs better in terms of label distribution, having the most green labels (183) and the least red labels (94). This may indicate that Open Kitchen offers more environmentally friendly options compared to the others.
- Gather shows an even distribution across the three labels, suggesting a balance between sustainable and less sustainable options.

- Feast has the lowest number of green labels compared to the others but still maintains a good balance between Red, Yellow, and Green.

Overall, the data suggests that Open Kitchen is leading in providing more climate-friendly food options, whereas Feast has more room for improvement based on the distribution of sustainability labels.

AMS Restaurant Calculations

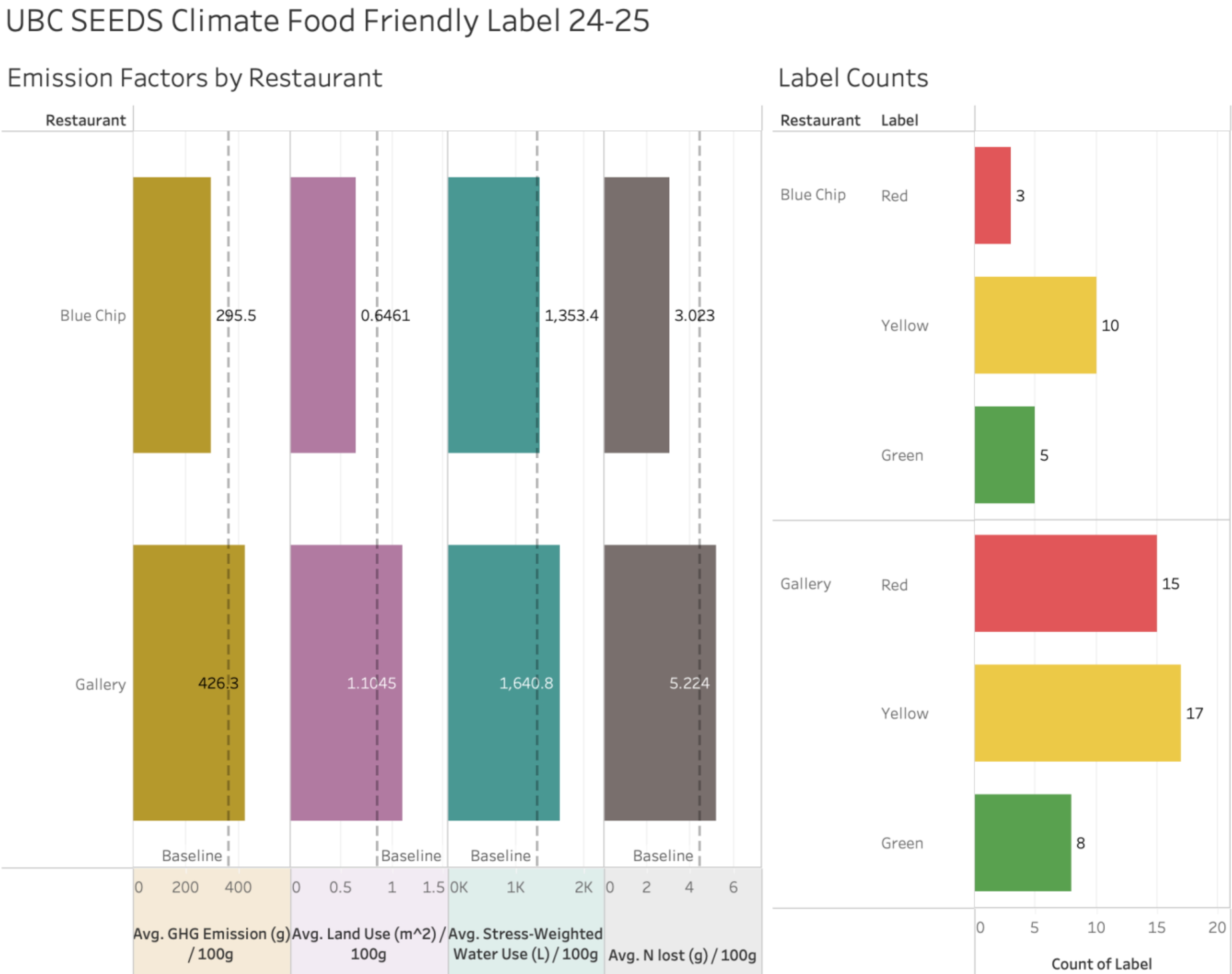


Figure 2: AMS Restaurants Calculation Summary: Showcases label counts and average emissions values for The Gallery Restaurant and Blue Chip Cafe. The dotted line represents the baseline for each factor

With a more limited menu compared to the dining halls, I conducted an analysis of the recipes for Blue Chip Cafe and the Gallery Restaurant in the AMS Nest.

Blue Chip Cafe

The analysis for Blue Chip Cafe was restricted to in-house items for which recipes were available. This excludes many menu items, including some of the cafe's most popular offerings, as they are not made in-house. As a result, the data provides an incomplete picture of the cafe's overall environmental impact, making it difficult to draw definitive conclusions about its menu.

Gallery Restaurant

For the Gallery Restaurant, the analysis revealed that the average emissions for its menu items are higher than the baseline, and the number of green-labeled items is lower than yellow and red items, with green labelled items representing only 20% of items offered. This suggests room for improvement in offering more environmentally sustainable options.

Other Outlets

Apart from the outlets mentioned above, the label was also calculated for the following: **Harvest Feastival, LFS Dinner, Perugia Café, Mercante Pizzeria** and **Sage Catering**

Future Considerations

Over the course of this term, some key concerns and potential improvements to the workflow have been identified. These require further research and consideration to optimize the accuracy and effectiveness of the label calculation process.

1. **Equal weights of environmental impacts**: Currently, we have given equal weights to all the environmental impacts (Green House Gas emissions, Nitrogen Lost, Stress-Weighted Water use and Land use). This approach, while straightforward, is still subjective and different weighting could be used to reflect different priorities from the CFFS-AT and UBC stakeholders.
2. **Revision of Item Categories**: The current categorization of items is overly broad, leading to inconsistencies and a lack of granularity in the labeling process. Several items do not fit neatly into a single category, which can reduce the accuracy and credibility of the labels. More granularity in high or medium intensity categories like tree nuts and seeds, fish (finfish), stimulants & spices misc., etc. could be seeked

AMS Process Pipeline

With the transition to Optimum Control v5 Premier, the extracted data for AMS has undergone significant changes, notably the absence of a unique identifier and the presence of unintelligible columns. This necessitated a comprehensive restructuring of the data ingestion pipeline to accommodate the new format.

The current workflow heavily relies on existing data to calculate emission data for various items and preparations. Therefore, establishing a unique identifier system for AMS items became essential, ensuring that newly assigned IDs do not overlap with existing data.

To address these challenges, I implemented a solution that began with cleaning and preprocessing the data. This involved reading CSV files and consolidating relevant information from various sources by dropping unnecessary columns and normalizing item descriptions for enhanced consistency.

To prevent ID overlaps with existing data, I developed a systematic approach for assigning unique identifiers based on normalized descriptions, employing both exact and fuzzy matching techniques. This process included creating a mapping of existing items to their normalized descriptions for accurate ID assignment.
For items that could not be matched, I introduced a random ID generation function to ensure uniqueness.

UBC Food Services Plant-Based Dashboard

To effectively communicate the environmental impact of food choices, I developed a website that allows users to explore and compare emissions data between animal-based recipes and their plant-based counterparts. You can find it here.

The website currently uses data from UBC's dining halls, which ensures consistency and relevance for the campus community. The comparisons feature pairs of recipes, such as Spicy Soy Salmon and Spicy Soy Tofu, with detailed insights into their greenhouse gas (GHG) emissions and other sustainability metrics.

On the homepage, users are presented with a clean, interactive bar chart highlighting a list of recipe comparisons. Each pair is displayed with their respective emission values

Figure 3: Landing Page for Plant-based Dashboard: The various bars display the comparison between a plant based and animal based dish. Clicking on a bar leads to a in-depth comparison page

Upon selecting a comparison, users are directed to a dedicated page providing an in-depth analysis of the two recipes. This page features several visualizations that break down and compare key metrics such as GHG emissions, nitrogen loss, land use, and water consumption. The visualizations are designed to be both visually appealing and easy to interpret, ensuring that users can quickly understand the data.

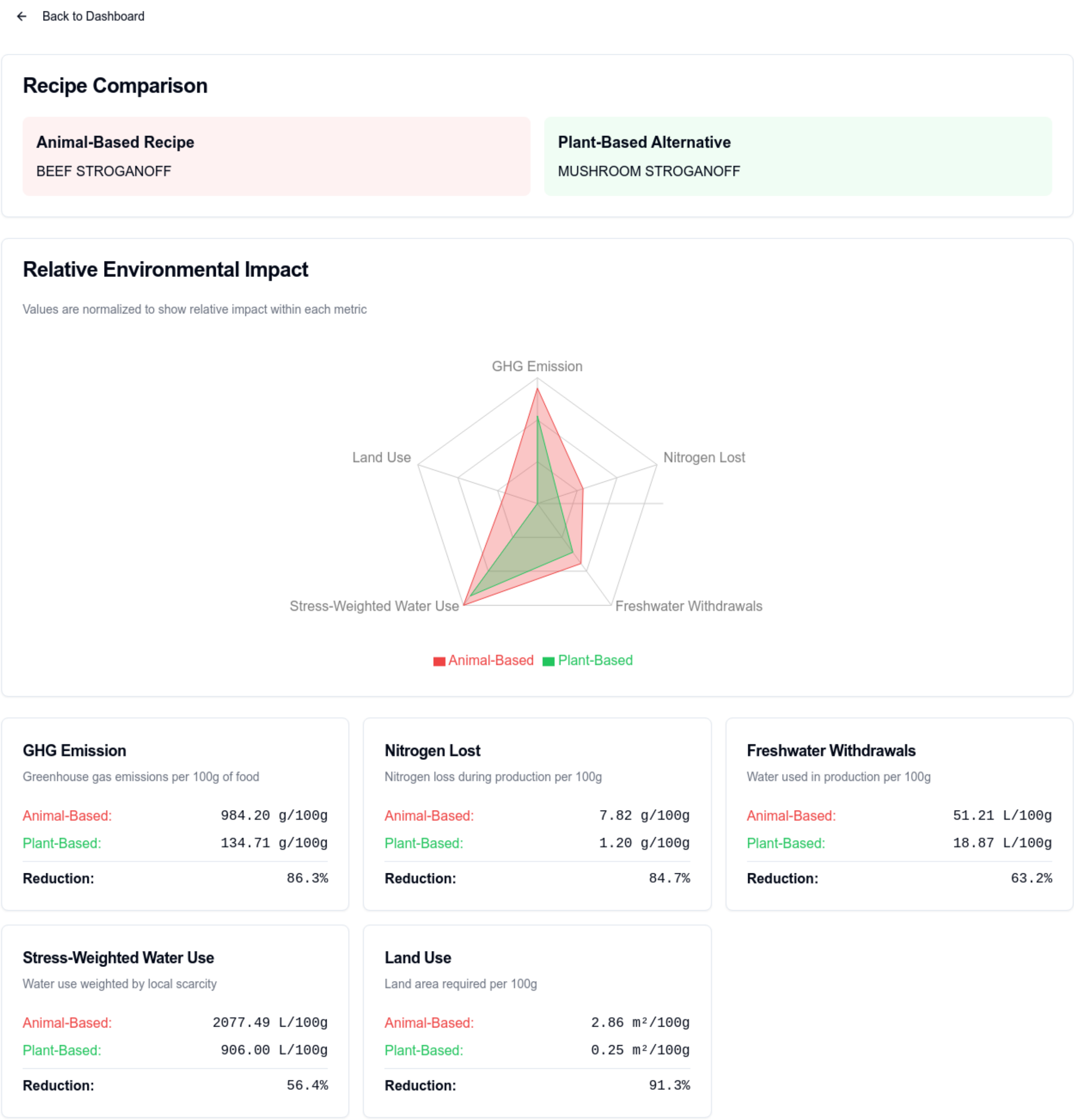


Figure 4: In-Depth Comparison Page: This page showcases the significant difference between the two dishes using a radar chart and values for each emission factor and the reduction percent

The goal of this initiative is to highlight the significant differences in environmental impact, empowering users to make informed and sustainable dietary decisions. While the website was initially designed with the goal of informing users, its focus should shift toward actively driving them to adopt more sustainable options.

Expansion

The dashboard developed to promote Plant-Based choices is an excellent way to educate the people in dining halls about more sustainable choices. It can be improved further by incorporating the following features

1. **Search Feature**: Introducing a search functionality will significantly improve user experience by allowing individuals to look up their favorite dishes and explore plant-based alternatives.
2. **Segregation by Restaurant**: To provide users with a more personalized and relevant experience, the dashboard can be divided based on the outlet names. Currently, this includes the three dining halls. Features of this enhancement include:
    a. **Restaurant-Specific Pages**: Users can navigate directly to a specific dining hall's menu and see the available dishes along with their plant-based alternatives.
    b. **Localized Recommendations**: The dashboard can display popular plant-based options unique to each dining hall, encouraging adoption through familiarity and availability.
    c. **Navigation Menu**: A clear and user-friendly navigation menu that lists dining hall names for easy access.

Climate Friendly Food Label Desktop Application

This term marked the initiation of a long-term project aimed at developing a desktop application to reduce the technical knowledge required now so more focus can be put towards the development and research behind the workflow. The application is designed to streamline the label generation process, reduce manual effort, and ensure consistency across different outlets.

Overview and Objectives

The primary goal of the application is to provide a user-friendly and efficient tool for label calculation. By building this application, individual outlets could eventually calculate labels for their menu items independently, reducing reliance on centralized manual workflows. This shift will also enable future individuals in this position to focus on refining the label methodology and improving its relevance and impact or automate some of the manual workflows.

This approach offers several advantages, including increased efficiency, scalability, and empowering outlets to take greater ownership of their sustainability efforts. By digitizing the process, it lays the groundwork for continuous improvement of the label system. The application will gather valuable data and user feedback, facilitating iterative enhancements and providing developers and analysts with insights to devise innovative solutions.

Drawbacks

While the development of an application marks an exciting step forward for the Climate Friendly Food Label Project, it also presents several challenges that must be considered:

1. **Development Timeline:** Creating a fully functional application is a significant undertaking, requiring extensive hours for development and rigorous testing. With only one developer handling this project alongside other responsibilities, the estimated timeline for completion ranges between 7 to 10 months.

2. **Maintenance:** One of the goals of this application is to empower future students to focus on research or enable outlets to compute the labels independently. However, as with any software, it is impossible to anticipate and resolve all potential issues during development. Errors may arise over time, which could prove challenging for users without technical expertise to address effectively.
3. **Technical Transition:** Over time, the underlying framework used to build the application will inevitably become outdated or deprecated. When this occurs, SEEDS will need to invest in hiring skilled students to migrate the application to updated frameworks or technologies. While this scenario is likely a decade away, it will bring its own set of complexities and resource requirements.

Current Progress

To date, the following components of the application have been designed and partially implemented:

Home Page

- A simple and intuitive landing page with two main buttons: Upload Data and Explore Outlets.

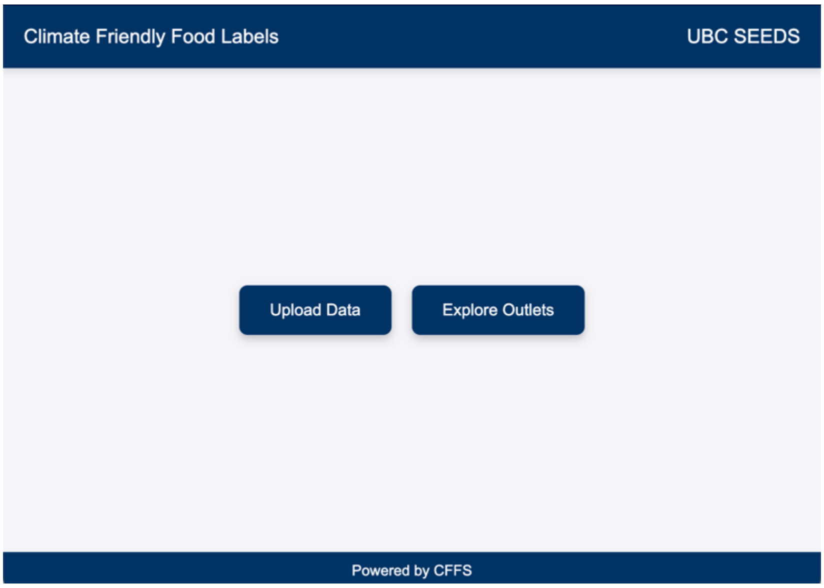
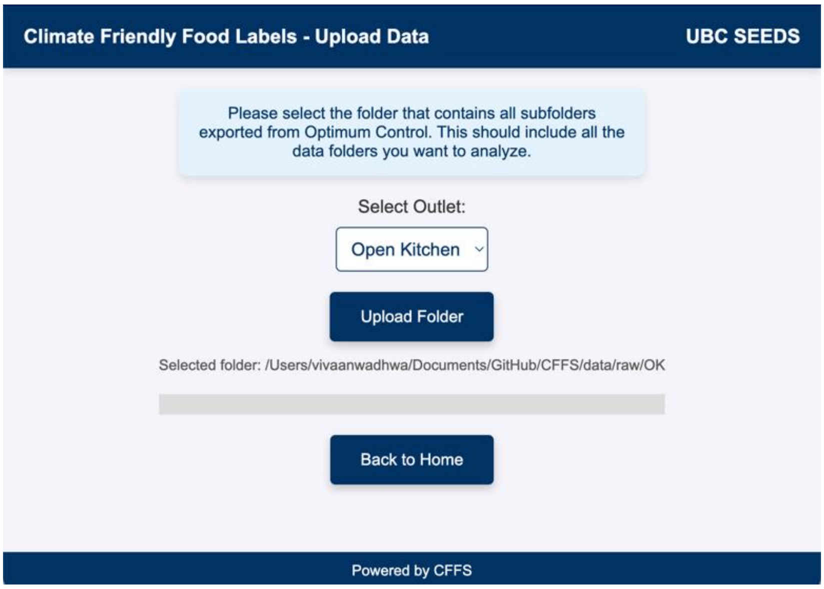


Figure 5: Home Page for the application

Upload page

- This page allows users to select a folder to upload data, ensuring seamless data ingestion into the system.
- Basic validations and file checks are included to ensure compatibility with the application’s requirements.

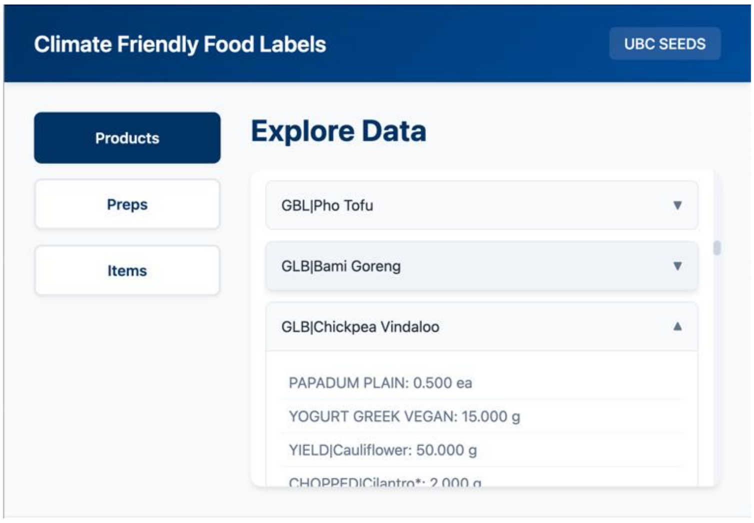
Figure 6:
Upload Page:
An outlet selecting dropdown and option to choose folder

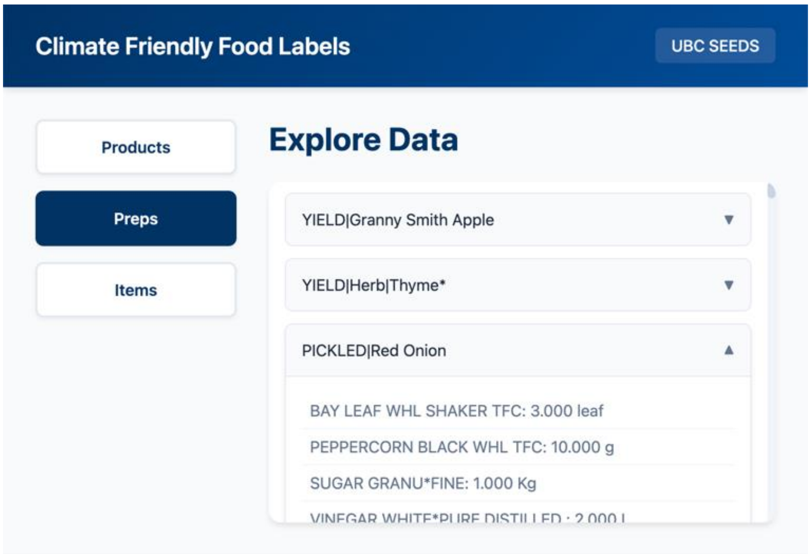


Explore Outlets Page

- Displays a list of all products in an outlet.
- Each product entry includes buttons that allow users to view details about individual preparations and items.

Figure 7 and 8:
Explore oulet Pages:
Lists all the various Products, Preps and Items. The recipe for the respective Products and Preps can be seen by expanding them





The foundational elements of the Climate Friendly Food Label desktop application have been successfully established, with detailed technical documentation outlining the app's architecture and functionality. The next phase of development will focus on implementing the workflow scripts and expanding the app's features. By prioritizing backend functionality before frontend design, the development process can proceed more efficiently.

Future Considerations and Next Steps

1. **Complete Scripts**: Developing the remaining scripts is critical to advancing the application. These include:
    a. Scripts for Label Calculations: Updating items and preparations with nonstandard weights and assigning items their emissions category IDs. These scripts will complete the core label calculation process.
    b. Additional Scripts: Other functionalities, such as viewing outlet labels and displaying visualizations, will require separate scripts to be developed.
2. **Testing**: Rigorous testing using pytest is essential for ensuring the reliability of the scripts and custom classes. This will identify potential bugs early and ensure the application performs as expected under various scenarios.
3. **Design Pages**: The next set of frontend pages needs to be designed and seamlessly integrated with their respective backend scripts. These include:

   a. Pages for updating items and preparations with nonstandard weights.
   b. Pages for assigning emissions category IDs to items.

   Mock-ups for these pages can be found in the repository as well.

4. **Maintenance Solutions**: Planning for the application's long-term usability is crucial. Two potential solutions for maintaining the app include:
   a. **Comprehensive Documentation**: Providing detailed guides and troubleshooting steps for common issues.
   b. **Ongoing Support**: Employing a part-time resource to address technical challenges and ensure the app remains functional.

Emission Category Classification

Context and rationale

Currently, we rely solely on manual assignment or generic Large Language Models (LLMs) such as ChatGPT without specific fine-tuning or domain optimization to assign the various items into emission categories. However, it can yield suboptimal results for the following reasons:

1. **Scalability**: Manual assignment is time-consuming and error-prone, especially for large datasets, making it impractical for consistent classification.
2. **Generalized AI Limitations**: Untrained LLMs may lack the domain-specific understanding required for accurate categorization and could struggle with subtle nuances such as distinguishing between overlapping terms like "White Wine Vinegar" and "Paste Chili Bean".
3. **Model Customization**: Traditional machine learning models can be tailored to this specific problem through preprocessing, hyperparameter tuning, and targeted techniques for class imbalance—offering better performance in structured datasets.

Current Results

The task of classifying items into emissions categories based on their descriptions and inventory groups is being addressed using three models: **Logistic Regression, Multilayer Perceptron (MLP) Classifier, and Multinomial Naive Bayes (NB)**. Current performance metrics for each model are as follows:

- **Multinomial Naive Bayes (NB): 72.1% accuracy**
- **Logistic Regression: 68.17% accuracy**
- **MLP Classifier: 67.04% accuracy**

Challenges

Despite these results, several factors are limiting the overall accuracy:

1. Class Imbalance

   a. Approximately 16% of the dataset consists of items in the "wheat/rye (bread, pasta)" category, creating a significant imbalance.
   b. Some categories are underrepresented, with only one or two entries, making it difficult for models to identify consistent patterns.
2. Inconsistent and Ambiguous Descriptions
   a. Items often have descriptions with overlapping identifiers. For instance, "White Wine Vinegar" could belong to both "wines" and "vinegars," depending on its usage.
   b. Proper preprocessing of descriptions is essential to reduce ambiguity and improve classification accuracy.

Next Steps

1. **Improved Preprocessing**
   a. **Goal**: Standardize and clean item descriptions to reduce ambiguity.
   b. **Approach**:
      i. Use natural language processing (NLP) techniques to extract context-sensitive keywords.
      ii. Incorporate domain knowledge to disambiguate terms like "vinegar" vs. "wine" based on usage patterns.
2. **Addressing Class Imbalance**
   a. **Goal**: Improve model performance on underrepresented classes.
   b. **Approach**:
      i. Apply resampling techniques like SMOTE or under sample dominant classes.
      ii. Adjust loss functions using class weights (e.g., weighted cross-entropy) to prioritize minority classes during training.
3. **Expand Dataset**

a. **Goal**: Enhance generalizability across categories.

b. **Approach**:

    i. Collect additional entries for underrepresented categories.

    ii. Leverage synthetic data generation techniques to simulate plausible new entries.

4. **Explore LLM Based Solution**

    a. LLMs like GPT offer a significant advantage by understanding contextual nuances in complex descriptions.

    b. **Proposed Application**: Fine-tune a GPT model on emissions categories, enabling it to accurately classify ambiguous terms. For instance, the model could differentiate between "White Wine Vinegar" as a "vinegar" when used in cooking versus a "wine" in fermentation processes.

    c. **Mitigating Class Imbalance**: LLMs can use contextual learning to handle minority classes better and even generate synthetic data to support traditional models.

Suggestions For Data Extraction Timeline

The timeline for data extraction and label calculation varies between UBC Food Services (UBCFS) and the AMS, based on data structure, preparation requirements, and the scale of operations. Below are the proposed timelines and considerations for efficient workflow:

1. **UBC Food Services (UBCFS)**
    a. **Outlets (excluding Dining Halls)**: For outlets outside the dining halls, it is ideal to receive the data at least **two weeks before any deadlines**. Since the data from these outlets is usually well-structured, the time is mainly allocated to performing comprehensive consistency checks to ensure accuracy before submission. Consistency check encompass going over label assignments and comparing that to expert expectation and double checking that counterintuitive results have indeed been assigned the right categories, that non-standard units have been converted appropriately, and that there are no data entry issues in the data supplied by the food business.
    b. **Dining Halls and Large-Scale Outlets**: For the three main dining halls or any outlet of similar scale, the data should be provided **at least one month in advance if not more**. The larger volume and complexity of data for these outlets necessitate additional time for thorough verification and calculations.
2. AMS
    a. **General Timeline**: Given the less structured nature of AMS data and the occasional issues arising from the current ID assignment process, the data should be made available at least three to four weeks before the deadline. This timeline accounts for additional preprocessing and error correction efforts.
3. **Challenges with Not-made--In-House Recipes**: Currently, there is no established method for calculating Climate Friendly Food Labels for items not made in-house, as these lack precise recipes or ingredient lists. Without accurate measurements and

ingredient breakdowns, calculating the accurate emissions data for such items is not feasible. The best option available right now is an approximation using online recipes for similar products, but this is potentially inaccurate. This limitation should be considered in planning timelines and communication with outlets.

Recommendations

1. **Collaborative Planning**: Schedule regular coordination meetings with outlet managers to ensure timely data submission.
2. **Future Improvements**: Explore ways to incorporate proxy data or estimations for not-made--in-house items to expand label calculation coverage.

Conclusion

The Climate Friendly Food Label Project represents a significant step toward achieving UBC's Climate Action Plan 2030 goals. By enhancing transparency around the environmental impact of food choices, this initiative empowers the UBC community to make more sustainable decisions. My contributions as a Sustainable Food Systems Data and Reporting Analyst have focused on improving data accuracy, refining emission calculation processes, and developing user-friendly tools for data visualization and analysis.

Acknowledgments

I would like to express my deepest gratitude to Ms. Laure Dupuy and Dr. Juan D. Martinez for their unwavering support and guidance throughout my time at UBC SEEDs. I am also deeply thankful to the CFFS Action Team and the various client outlet for their patience, collaboration, and understanding. This project has provided me with an incredible opportunity to contribute directly to UBC's sustainability initiatives and has been an enriching and rewarding experience for me.