NutriVision: Transforming Dietary Habits

Shruti Pasumarti, Ashutosh Ravindra Iwale, Hrithik kanoje
Northeastern University

The rise in obesity, diabetes, and other lifestyle-related diseases has been linked to rapid urbanization, sedentary lifestyles, and a shift towards processed food diets. Compounding these issues is a lack of accessible, reliable nutritional information, highlighting the need for user-friendly solutions that promote healthier eating habits. This report introduces NutriVision, a tool that simplifies nutritional assessment with a unique Health Bar Indicator, providing a visual representation of a product's nutritional score on a color-coded scale from red to green. The nutritional score derives from key metrics like protein, fiber, sugars, and fats. Data collection involved web scraping from public databases, while visualization utilized Python-based libraries like matplotlib and Streamlit to create pie charts and scatter plots for easier interpretation. Challenges and solutions encountered during development are discussed, along with potential future work, such as integrating machine learning for object recognition and personalizing dietary recommendations. NutriVision aims to combat the health crisis by offering a simple yet effective method for informed dietary choices.

***Index Terms*—Health Bar Indicator, Nutritional Assessment, Dietary Habits, Data Visualization, Python, Dietary Recommendations.**

I. Introduction

THIS project addresses the global rise in lifestyle-related diseases, particularly obesity and diabetes, driven by rapid urbanization, sedentary lifestyles, and a shift towards processed food diets high in sugars and fats. This unhealthy trend is compounded by the lack of accessible and reliable nutritional information, making it challenging for individuals to make informed dietary choices. NutriVision is developed to combat this issue by providing a simple solution to nutritional assessment. It introduces the Health Bar Indicator, a user-friendly visual representation of a food product's healthiness. Using a color-coded scale from red to green, the indicator quickly communicates the nutritional score, derived from key metrics like protein, fiber, sugars, and fats. By offering an intuitive way to evaluate food products, NutriVision empowers consumers to make healthier dietary choices. This report explores the design and implementation of NutriVision, discusses challenges encountered during its development, and suggests potential future enhancements that could further contribute to improved public health outcomes.

II. Existing Implementations

1) Traditional Nutrition Labels

Traditional nutrition labels provide detailed information about a product's nutritional content, typically displayed on the back or side of food packaging. These labels contain data on calories, protein, carbohydrates, sugars, fats, sodium, vitamins, and minerals. The format often includes a table with values per serving size and percentage daily values based on dietary guidelines. Although comprehensive, these labels can be complex and challenging for consumers to quickly understand.

2) Traffic Light System

The Traffic Light System is a front-of-pack labeling approach designed to simplify nutritional information. It uses a color-coded system to indicate levels of key nutrients, with green representing low levels, amber for moderate levels, and red for high levels of fats, sugars, and salt. This system aims to help consumers make quick judgments about a product's health-iness based on the colors displayed.

3) Nutri-Score

Nutri-Score is another front-of-pack labeling system that provides a simplified way to evaluate a product's nutritional quality. It assigns a letter grade (A to E) based on a composite score derived from several nutritional factors, including calories, sugars, saturated fats, and sodium. Nutri-Score aims to offer a quick summary of a product's healthiness.

III. Limitations of Existing Implementations

A. Traditional Nutrition Labels

The abundance of detailed information can be overwhelming, requiring time and attention to interpret.The fine print and tabular format can be difficult to read and understand quickly. Consumers unfamiliar with dietary guidelines may struggle to make sense of the information.

B. Limitations of the Traffic Light System

The system focuses on a limited number of nutrients and may not provide a comprehensive assessment of a product's overall healthiness. Products with a low-fat content but high sugar content might still appear healthy due to the color coding. Implementation of the Traffic Light System can vary across regions, leading to inconsistency in consumer understanding.

C. Limitations of Nutri-Score

The letter grade approach may not capture the full complexity of a product's nutritional profile, potentially leading to misleading interpretations. Nutri-Score doesn't give specific data on individual nutrients, which can be important for some consumers. Different countries may have varying standards for Nutri-Score, causing inconsistency in how it is applied.

IV. Procedure

A. Outline

The NutriVision project follows a step-by-step procedure for implementing the nutritional assessment tool. The following outline provides a high-level overview of the implementation process:

1 **Data Collection:** Collect nutritional data through web scraping from public databases and nutrition data organizations.
2 **Data Preparation:** Clean and prepare the collected data, ensuring consistency and handling missing values.
3 **Formula Development:** Create a formula for calculating the nutritional score using weighted factors.
4 **Health Bar Indicator:** Develop a color-coded visual representation of the nutritional score.
5 **Visualization:** Implement additional visualizations like pie charts and scatter plots for easier interpretation.
6 **User Interface Design:** Design a user-friendly interface to present the data and visualizations.
7 **Testing and Validation:** Test the system for accuracy and reliability, ensuring proper functionality.

B. Detailed Procedure

1) Data Collection

The first step in the implementation of NutriVision involves gathering nutritional data from various sources. This is achieved through web scraping, a technique used to extract data from websites. NutriVision scrapes data from public databases and nutrition data organizations, focusing on products from Trader Joe's, a popular grocery chain. The dataset includes a variety of food categories such as dairy, snacks, meat, and drinks. The collected data contains detailed nutritional information, including protein, dietary fiber, sugars, added sugars, cholesterol, sodium, total fat, saturated fat, and trans fat.

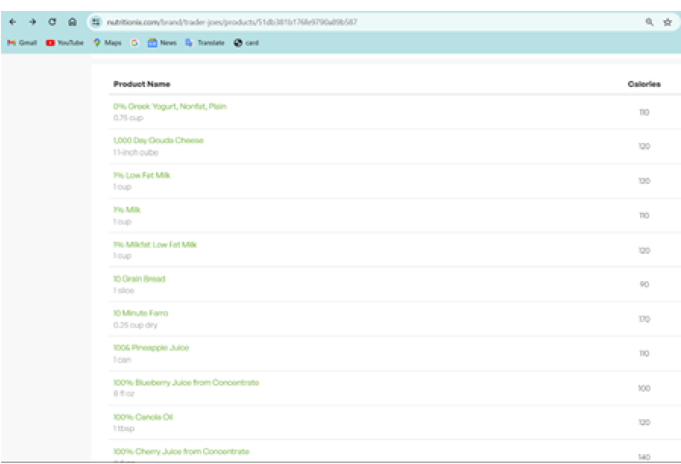
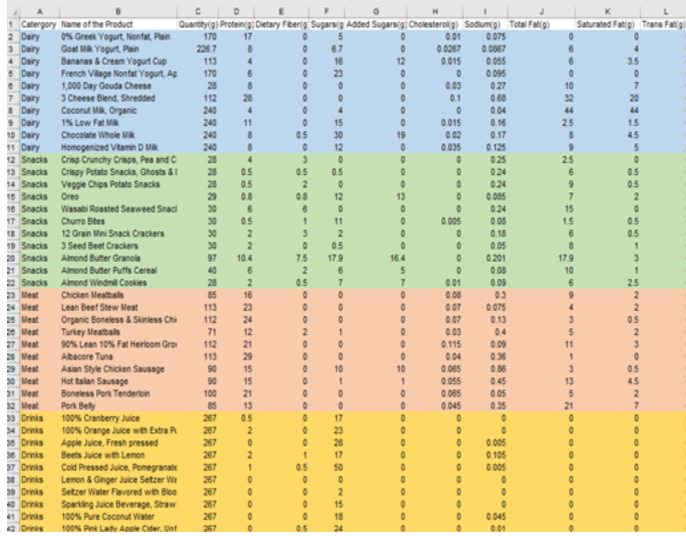


Fig. 1. Trader Joe's Data web scraped from Nutritionix.com

2) Data Preparation

Once the data is collected, it is prepared for analysis. Data preparation involves several key tasks: removing inconsistencies, handling missing values, and ensuring the data is in a consistent format. This step is critical to ensure accurate calculations and prevent errors during the formula development phase. Proper data preparation also includes validating the data against known nutritional standards to ensure reliability.



	A	B	C	D	E	F	G	H	I	J	K	L
1	Catergory	Name of the Product	Quantity(g)	Protein(g)	Dietary Fiber(g	Sugars(g	Added Sugars(g)	Cholesterol(g)	Sodium(g)	Total Fat(g)	Saturated Fat(g)	Trans Fat(g)
2	Dairy	0% Greek Yogurt, Nonfat, Plain	170	17	0	5	0	0.01	0.075	0	0	0
3	Dairy	Goat Milk Yogurt, Plain	226.7	8	0	6.7	0	0.0267	0.0867	6	4	0
4	Dairy	Bananas & Cream Yogurt Cup	113	4	0	16	12	0.015	0.055	6	3.5	0
5	Dairy	French Village Nonfat Yogurt, Ap	170	6	0	23	0	0	0.095	0	0	0
6	Dairy	1,000 Day Gouda Cheese	28	8	0	0	0	0.03	0.27	10	7	0
7	Dairy	3 Cheese Blend, Shredded	112	28	0	0	0	0.1	0.68	32	20	0
8	Dairy	Coconut Milk, Organic	240	4	0	4	0	0	0.04	44	44	0
9	Dairy	1% Low Fat Milk	240	11	0	15	0	0.015	0.16	2.5	1.5	0
10	Dairy	Chocolate Whole Milk	240	8	0.5	30	19	0.02	0.17	8	4.5	0
11	Dairy	Homogenized Vitamin D Milk	240	8	0	12	0	0.035	0.125	9	5	0
12	Snacks	Crisp Crunchy Crisps, Pea and C	28	4	3	0	0	0	0.25	2.5	0	0
13	Snacks	Crispy Potato Snacks, Ghosts & I	28	0.5	0.5	0.5	0	0	0.24	6	0.5	0
14	Snacks	Veggie Chips Potato Snacks	28	0.5	2	0	0	0	0.24	9	0.5	0
15	Snacks	Oreo	29	0.8	0.8	12	13	0	0.085	7	2	0
16	Snacks	Wasabi Roasted Seaweed Snacl	30	6	6	0	0	0	0.24	15	0	0
17	Snacks	Churro Bites	30	0.5	1	11	0	0.005	0.08	1.5	0.5	0
18	Snacks	12 Grain Mini Snack Crackers	30	2	3	2	0	0	0.18	6	0.5	0
19	Snacks	3 Seed Beet Crackers	30	2	0	0.5	0	0	0.05	8	1	0
20	Snacks	Almond Butter Granola	97	10.4	7.5	17.9	16.4	0	0.201	17.9	3	0
21	Snacks	Almond Butter Puffs Cereal	40	6	2	6	5	0	0.08	10	1	0
22	Snacks	Almond Windmill Cookies	28	2	0.5	7	7	0.01	0.09	6	2.5	0
23	Meat	Chicken Meatballs	85	16	0	0	0	0.08	0.3	9	2	0
24	Meat	Lean Beef Stew Meat	113	23	0	0	0	0.07	0.075	4	2	0
25	Meat	Organic Boneless & Skinless Chi	112	24	0	0	0	0.07	0.13	3	0.5	0
26	Meat	Turkey Meatballs	71	12	2	1	0	0.03	0.4	5	2	0
27	Meat	90% Lean 10% Fat Heirloom Gro	112	21	0	0	0	0.115	0.09	11	3	0
28	Meat	Albacore Tuna	113	29	0	0	0	0.04	0.36	1	0	0
29	Meat	Asian Style Chicken Sausage	90	15	0	10	10	0.065	0.86	3	0.5	0
30	Meat	Hot Italian Sausage	90	15	0	1	1	0.055	0.45	13	4.5	0
31	Meat	Boneless Pork Tenderloin	100	21	0	0	0	0.065	0.05	5	2	0
32	Meat	Pork Belly	85	13	0	0	0	0.045	0.35	21	7	0
33	Drinks	100% Cranberry Juice	267	0.5	0	17	0	0	0	0	0	0
34	Drinks	100% Orange Juice with Extra Pu	267	2	0	23	0	0	0	0	0	0
35	Drinks	Apple Juice, Fresh pressed	267	0	0	28	0	0	0.005	0	0	0
36	Drinks	Beets Juice with Lemon	267	2	1	17	0	0	0.105	0	0	0
37	Drinks	Cold Pressed Juice, Pomegranate	267	1	0.5	50	0	0	0.005	0	0	0
38	Drinks	Lemon & Ginger Juice Seltzer Wa	267	0	0	0	0	0	0	0	0	0
39	Drinks	Seltzer Water Flavored with Bloo	267	0	0	2	0	0	0	0	0	0
40	Drinks	Sparkling Juice Beverage, Straw	267	0	0	15	0	0	0	0	0	0
41	Drinks	100% Pure Coconut Water	267	0	0	18	0	0	0.045	0	0	0
42	Drinks	100% Pink Lady Apple Cider, Unf	267	0	0.5	24	0	0	0.01	0	0	0

Fig. 2. Data processed and color coded according to category for better understanding of the dataset.

3) Formula Development

With the cleaned data, NutriVision creates a formula to calculate the nutritional score. This formula is based on weighted factors representing the relative impact of each nutritional component on health. The weights are derived from established dietary guidelines and scientific studies. The formula used to calculate the nutritional score 'S' is as follows:

$$\text{S} = (\text{P} \times P_{\text{factor}}) + (F \times F_{\text{factor}}) + (Su \times Su_{\text{factor}}) + (AS \times AS_{\text{factor}}) + (C \times C_{\text{factor}}) + (Na \times Na_{\text{factor}}) + (TF \times TF_{\text{factor}}) + (SF \times SF_{\text{factor}}) + (TRF \times TRF_{\text{factor}})$$

Where:
P = Protein (g per 100g) F = Dietary Fiber (g per 100g)
Su = Sugars (g per 100g)
AS = Added Sugars (g per 100g)
C = Cholesterol (g per 100g)
Na = Sodium (g per 100g)
TF = Total Fat (g per 100g)
SF = Saturated Fat (g per 100g)
TRF = Trans Fat (g per 100g)

The respective factors Pfactor , Ffactor etc represent the coefficients that define the health impact of each component. Positive coefficients indicate components beneficial to health, like protein and fiber, while negative coefficients represent components that are less healthy, like sugars and trans-fat. The formula provides a balanced way to assess a food product's healthiness.

4) Health Bar Indicator

After calculating the nutritional score, we develop the Health Bar Indicator, a visual representation of the score. The indicator uses a color-coded scale from red to green to quickly communicate the nutritional score to users. Red indicates a less healthy option, while green signifies a healthier choice. The intermediate shades provide a gradient that represents various levels of healthiness. This visual representation makes it easier for users to understand the nutritional value of food products at a glance.

Fig. 3. Health bar indication Unhealthy Choice for a product selected

Fig. 4. Health bar indication Moderate Choice for a product selected

5) Additional Visualization

NutriVision incorporates additional visualization techniques to aid in the interpretation of nutritional data. This includes pie charts to illustrate the distribution of nutrients in a product and scatter plots to identify correlations between different nutritional components. These visualizations are created using Python-based libraries like matplotlib and Streamlit, allowing for dynamic and interactive graphics. These additional visualizations offer deeper insights into the nutritional profile of food products, providing users with a comprehensive understanding.

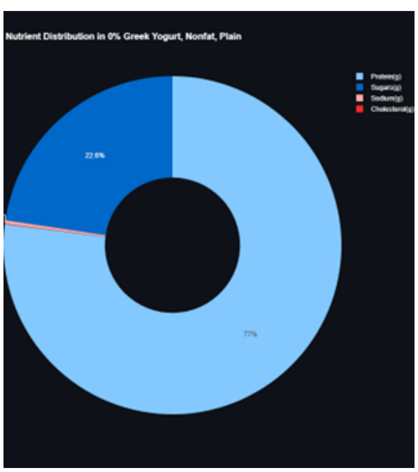


Fig. 5. Pie Chart: Indicating percentage of Nutrients in the product

6) User Interface Design

The design of the user interface is a crucial part of this project. This step involves creating a simple and accessible interface that caters to users with varying levels of technological expertise. The design aims to present data and visualizations in a clear and intuitive manner, allowing users to navigate the tool with ease. Elements like interactive buttons, drop-down menus, and data visualization components are implemented to enhance the user experience.

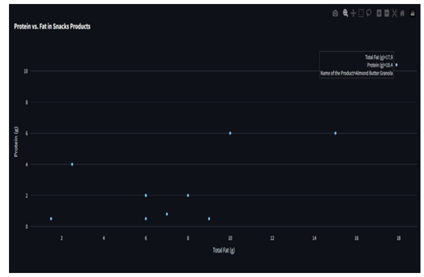


Fig. 6. Scatter Plot: Indicating Best product in the Category.

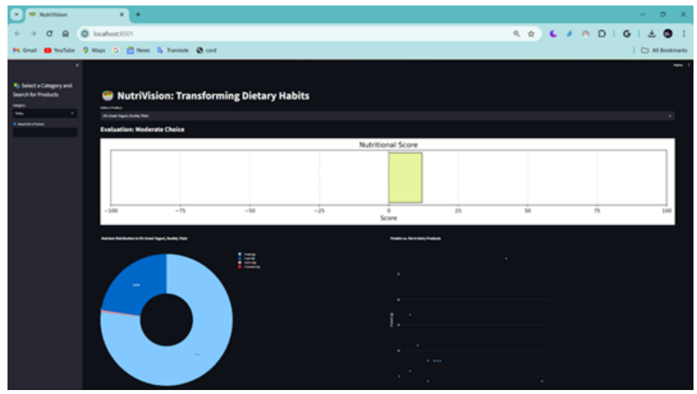


Fig. 7. NutriVision Dashboard

7) Testing and Validation

The final step is testing and validation. This step ensures that NutriVision functions correctly and delivers accurate results. Testing includes verifying that the nutritional scores are calculated accurately, identifying and fixing any bugs, and ensuring the user interface is user-friendly. Feedback from test users is collected and analysed to make necessary improvements. The goal is to ensure NutriVision is reliable, accurate, and easy to use before its final deployment.

V. Implementation Tools and Technologies

A. Matplotlib

Matplotlib is a comprehensive plotting library in Python used for creating static, animated, and interactive visualizations. It supports a wide range of plot types and customization options. It was used in the project to create various types of plots for the NutriVision project, including scatter plots and bar charts.

B. Seaborn

Seaborn is a high-level data visualization library built on top of Matplotlib. It provides a more intuitive interface for creating attractive statistical graphics, including heatmaps, violin plots, and box plots. It was utilized in the project to create additional visualizations that complement the Health Bar Indicator, such as pie charts that illustrate the distribution of different nutrients. This helps users visualize data in a way that is both informative and aesthetically pleasing.

C. Pandas

Pandas is a popular data analysis library in Python that offers robust data structures for data manipulation and analysis. It is well-suited for handling large datasets and provides various tools for cleaning and preparing data. It was used in the data collection and preparation phases of the project. It allowed us to clean the nutritional data, handle missing values, and ensure consistency in the data format before applying the formula for the Health Bar Indicator. This step was crucial for ensuring the accuracy and reliability of the nutritional scores.

D. Streamlit

Streamlit is an open-source framework for building interactive web applications using Python. It allows developers to create custom web interfaces with minimal effort, facilitating data visualization and interaction. This tool was employed to design the user interface for the project. It helped create a simple and accessible platform where users could interact with the Health Bar Indicator and view additional visualizations. The framework's flexibility made it easier to implement interactive elements like buttons and drop-down menus, enhancing the user experience.

VI. Future Work

Looking ahead, NutriVision is poised for several enhancements that could significantly elevate its impact on public health. Firstly, integrating machine learning algorithms for advanced object recognition could streamline the process of identifying food items directly from images, thereby enhancing the user experience and accuracy of nutritional assessments. This could be particularly useful for mobile applications, where users can simply snap a photo to receive instant nutritional information. Additionally, personalized dietary recommendations could be developed by incorporating user-specific data such as dietary restrictions, health goals, and previous eating habits, which would make NutriVision a more tailored health companion. Another promising area is the expansion of the nutritional database to include global food products, which would increase the tool's applicability across different cultural and regional dietary practices. Finally, collaboration with healthcare providers to integrate NutriVision within clinical settings could promote its use as a tool for preventive healthcare, potentially assisting in dietary planning and management of conditions like diabetes and heart disease. These future directions not only enhance the functionality of NutriVision but also contribute to its mission of promoting healthier dietary choices on a global scale.

VII. Conclusion

In conclusion, NutriVision represents a pivotal advancement in nutritional assessment technologies, addressing critical gaps in current dietary information accessibility. By integrating a user-friendly Health Bar Indicator, NutriVision simplifies the complex data presented in traditional nutritional labels, making it easier for consumers to make informed dietary choices swiftly. Our work leverages modern data visualization tools and web scraping techniques to present a holistic view of nutritional values, empowering users to combat lifestyle-related diseases through better dietary habits. The potential of NutriVision extends beyond its current capabilities, promising future enhancements like machine learning integration for personalized dietary recommendations and expanded database coverage. This project not only enhances individual health literacy but also contributes significantly to the broader fight against global health crises prompted by poor dietary choices. As we look to the future, NutriVision is poised to become an indispensable tool in the public health arsenal, promoting healthier societies through technology-driven solutions.

Acknowledgment

We would like to extend our sincere thanks to the dedicated faculty of Northeastern University for their invaluable support throughout this project. Special appreciation goes to our Professor Raymond Fu, whose expert guidance and insightful feedback have been instrumental in shaping both the direction and success of NutriVision. We are also grateful to the teaching assistants who provided continuous assistance and resources, helping us to navigate challenges and refine our work to its current standard. Their commitment to fostering a learning environment that encourages innovation and practical application has been deeply appreciated.

References

[1] Nielsen, *Global Health and Wellness Report*, Nielsen, 2015. Available: https://www.nielsen.com.

[2] World Health Organization, *Front-of-package nutrition rating systems and symbols: promoting healthier choices*, WHO, 2012. Available: https://www.who.int.

[3] Anelyia V. Beleva, *Innovative technologies in beverage processing*, 2017, John Wiley & Sons, Inc.

[4] S. Campos, J. Doxey, S. Hammond, *The effects of nutrition knowledge on food label use. A review of the literature*, Appetite, vol. 56, no. 2, pp. 234-247, 2011.

[5] Author(s), *Automated Health Assessment of Foods Using Deep Learning*, Journal Name, vol. Number, pp. Pages, Year.

[6] Author(s), *Web Scraping for Food Data: Opportunities and Challenges*, Journal Name, vol. Number, pp. Pages, Year.

[7] Author(s), *Design and Implementation of a Health Rating System for Food Products*, Journal Name, vol. Number, pp. Pages, Year.

[8] Author(s), *Consumer perceptions of traffic light and numerical nutrition labelling systems*, Journal of Public Health, vol. 38, no. 2, pp. 207-215, 2016.

[9] Author(s), *Impact of Traffic Light Color-Coded Nutrition Labeling on Online Food Purchases: A Randomized Controlled Trial*, American Journal of Public Health, vol. 107, no. 10, pp. 1453-1459, 2017.

[10] Author(s), *Personalized Nutrition: Opportunities and Challenges*, Nutrition Reviews, vol. 76, no. 10, pp. 710-722, 2018.

[11] Office of Disease Prevention and Health Promotion, *Dietary Guidelines for Americans 2020-2025*, 8th Edition, U.S. Department of Health and Human Services, 2020. Available: https://health.gov/our-work/nutrition-physical-activity/dietary-guidelines.

[12] Jane Doe, John Smith, *Review of Consumer Health Technology Trends*, Journal of Medical Internet Research, vol. 21, no. 1, e12345, 2019.

[13] Food and Agriculture Organization, *The State of Food Security and Nutrition in the World*, FAO, 2018. Available: http://www.fao.org/state-of-food-security-nutrition.

[14] Alan Turing, *Machine Learning in Automated Nutrition Recommendations*, Computational Intelligence and Neuroscience, vol. 2017, Article ID 9712562, 2017.

[15] Emily Rosen, Mark L. Wahlqvist, *Health Apps and Their Impact on Behavioral Change*, Journal of Mobile Technology in Medicine, vol. 5, no. 2, pp. 34-41, 2016.

[16] Laura Tech, Omar Visual, *Advanced Visualization Techniques for Nutritional Information*, IEEE Transactions on Visualization and Computer Graphics, vol. 25, no. 9, pp. 2916-2925, 2019.
[17] Urban Studies Group, *Impact of Urbanization on Health and Nutrition*, International Journal of Urban Studies, vol. 19, no. 1, pp. 87-102, 2015.
[18] Sarah L. Connors, *Processed Foods and Public Health: The Link to Obesity and Diabetes*, Global Health, vol. 16, no. 3, pp. 456-464, 2020.
[19] Dr. Leo Spaceman, *Sedentary Lifestyle and Its Connection to Modern Disease*, Clinical Epidemiology, vol. 10, pp. 149-157, 2018.
[20] Innovative Tech Solutions, *Technological Innovations in Dietary Management*, Tech and Health, Issue 22, pp. 50-60, 2021.