FOOD AVAILABILITY AND HEALTH ANALYSIS

Prepared by
Tyrell Cooper, Yongchan Lee, Jessica Akonor

DATA-201

SPRING 2025

Food and Agriculture Organization of the United Nations

THE WORLD BANK

Table of Contents

List of Visualizations

1. INTRODUCTION

Food availability, or food consumption, has a significant impact on human health[1]. Consuming healthy food has been shown to positively affect a person's overall well-being, while unhealthy eating habits often lead to negative health outcomes[2]

In our project, Food Availability and Health Analysis, we aim to answer the central question: How has food consumption affected general human health indicators across the globe? Rather than relying solely on written analysis, we use Tableau, a data visualization tool, to explore and present insights visually. This allows us to more clearly uncover patterns and relationships between food consumption and health on a global scale

---

[1] Kearney, John. "Food consumption trends and drivers." Philosophical transactions of the royal society B: biological sciences 365.1554 (2010): 2793-2807

[2] Holder, Mark D. "The contribution of food consumption to well-being." Annals of Nutrition and Metabolism 74.Suppl. 2 (2019): 44-52.

2. DATA SOURCES

As mentioned on Introduction, we want to extract insights from food consumption and human health. For extracting insights, datasets are required, and we will be using two datasets for achieving the goal: Food Availability and Health Data.

Food Availability dataset contains average daily per capita nutrient availability by country and food group. The dataset is sourced from Food and Agriculture Organization of the United Nations (FAO). FAO is one of the agency under the United Nation. FAO is trying to solve world's hunger problem and make sure that people are living in the food secured world. There are seven variables: Area code, Area, Food Group, Indicator, Year, Value, and Unit .

Area is a categorical variable that represents a country names. Similary, Area code indicates standatd country codes (ISO3). There are 190 different countries. It seems like 190 countries are all the countries in the world, but there are exception which are not recorded: West Sahara, Eritrea, Equatorial Guinea, French Guiana, Greenland, and Kosovo.

Food Group is a categorical variable that classifies all types of food people consume into distinct groups. There are 22 unique food groups represented in this variable, including examples such as: FG1: Cereals and their products, FG4: Milk and milk products, FG6: Fish, shellfish and their products, FG7: Meat and meat products, FG9: Vegetables and their products, and FG12: Sweets and sugars.

Indicator is a categorical variable that shows the type of nutrients from the food groups. There are 25 different nutrient types represented in this variable, including examples such as: Energy supply, Protein supply, Energy supply, Calcium supply, Vitamin C supply, Magnesium supply, Docosahexaenoic acid (DHA) supply, and Total saturated fatty acids supply.

Year is a discrete numerical variable the literally represent year. We will be using data from 2018 to 2022, so there are 5 different years: 2018, 2019, 2020, 2021, and 2022. *Value* is a continuous numerical variable that indicates measurement of the indicators. Each value is expressed in one of 4 Units of measurement: Grams per capita per day (g/cap/d), Kilocalories per capita per day (kcal/cap/d), Micrograms per capita per day (µg/cap/d), and Milligrams per capita per day (mg/cap/d). The table below is a structure of Food Availability dataset.

Area Code (ISO3)	Area	Food Group	Indicator	Year	Unit	Value
AFG	Afghanistan	Cereals and their products	Protein supply	2018	g/cap/d	44.8
AFG	Afghanistan	Cereals and their products	Protein supply	2019	g/cap/d	44
AFG	Afghanistan	Cereals and their products	Protein supply	2020	g/cap/d	44.5
AFG	Afghanistan	Cereals and their products	Protein supply	2021	g/cap/d	44.3
AFG	Afghanistan	Cereals and their products	Protein supply	2022	g/cap/d	45
AFG	Afghanistan	Cereals and their products	Fat supply	2018	g/cap/d	6.6
AFG	Afghanistan	Cereals and their products	Fat supply	2019	g/cap/d	6.6

Table 2-1. Structure of Food Availability Dataset

The Health Data is a dataset containing country-level health management indicators. The data is sourced from the World Bank. The World Bank is an international development organization that aims to reduce poverty by lending money to the government. Similar to the previously mentioned Food Availability dataset, it covers the years 2018 to 2022. This dataset includes six variables: Country, Country Code, Indicator, Year, and Value.

Country and *Country Code* are categorical variables trepresentssent the countries included in the dataset. The difference between the two is that Country shows the full name of the country, while Country Code displays its abbreviated form. Both variables contain 266 unique categories.

Indicator is a categorical variable that represents various health management statistics. It originally includes over 250 categories. However, after filtering out non-health related categories, 55 unique categories have left, such as Life Expectancy, Mortality Rate, Number of

Deaths, Prevalence of undernourishment (% of population), and Prevalence of Severe Food Insecurity, all of which are health-related metrics.

Year is a discrete numerical variable that includes years from 2018 to 2022, and *Value* is a numerical variable that shows the actual measurement corresponding to each indicator. The table below is a structure of the Health data.

Country	Country Code	Indicator	Year	Value
Aruba	ABW	Adolescent fertility rate (births per 1,000 women ages 15-19)	2018	24.049
Aruba	ABW	Age dependency ratio (% of working-age population)	2018	47.11667049
Aruba	ABW	Age dependency ratio, old (% of working-age population)	2018	19.58286617
Aruba	ABW	Age dependency ratio, young (% of working-age population)	2018	27.53380432
Aruba	ABW	Birth rate, crude (per 1,000 people)	2018	9.881
Aruba	ABW	Death rate, crude (per 1,000 people)	2018	8.597
Aruba	ABW	Fertility rate, total (births per woman)	2018	1.587
Aruba	ABW	Incidence of tuberculosis (per 100,000 people)	2018	5.4
Aruba	ABW	Life expectancy at birth, female (years)	2018	78.854
Aruba	ABW	Life expectancy at birth, male (years)	2018	73.096
Aruba	ABW	Life expectancy at birth, total (years)	2018	76.072
Aruba	ABW	Mortality rate, adult, female (per 1,000 female adults)	2018	59.643

Table 2-2. Structure of Health Dataset

As we can see from both datasets, there are overlapping variables: Country Code and Year. Based on these two key variables, we joined the two datasets. Using the merged dataset, we will visualize and analyze the relationship between food consumption and health

3. METHODOLOGY (using Tableau)

Tableau is a visual analytics platform that transforms the way we use data to solve problems—empowering individuals and organizations to make the most of their data[3]. Tableau not only provides users with the ability to interpret and edit data, but also offers tools to create a wide range of visualizations. In addition, its dashboard functionality allows users to build interactive visualizations, making it easier to explore and communicate insights.

In this project, we used Tableau to answer our research question. Specifically, we joined the Food Availability dataset and the Health dataset within Tableau to create a single, integrated dataset. Using this combined dataset, we extracted insights and built visualizations to support each team member's individual analysis. The detailed analyses presented by each team member are further explained in *4. ANALYSIS*.

[3] **Tableau.** *What Is Tableau?* Salesforce, https://www.tableau.com/why-tableau/what-is-tableau. Accessed 28 Apr. 2025.

4. ANALYSIS

Using the two datasets described earlier, we aim to answer our research question through detailed analysis. Rather than providing a single answer, each team member focuses on a different aspect of the analysis.

Tyrell conducts a Regional Analysis to investigate varying country rates of undernourishment and communicate how the consumption of energy and protein contributes to decreased rates of undernourishment. Yongchan performs a Life Expectancy Analysis to examine how life expectancy relates to food consumption and the nutrients provided by different foods, and to describe the food consumption patterns observed in countries with higher life expectancy. Jessica carries out a Trend Analysis to highlight sugar consumption levels and diabetes prevalence globally in the year 2021.

4.1. Regional Analysis (by Tyrell Cooper)

The central goal of this analysis was to identify which countries have the highest undernourishment (UN) percentages. My main variables were Health Value Indicator: *prevalence of undernourishment (% of population),* Food Value (composite), Food Value Indicator: *energy supply,* Country, Continents. The four graphs created and shown below showcase the relationship these variables in four different ways to boost user interpretabibilty of the data.

To start my exploration into the topic, here is a dual box and bar chart showing the differences between the continent's UN and their approximate composite consumption of food. This chart was created to show that there are noticeable differences between the ranges seen among the countries, specifically that some regions are marginally higher than others. All countries across Europe have UN percentages higher than half of all of African countries. The lowest ranked European country is Hungary with 11,047 (at the lower whisker). Half of all recorded African countries are below 9,918 with Namibia making it just above the median 9,948.

The bar charts in the background indicate the regions total food consumed

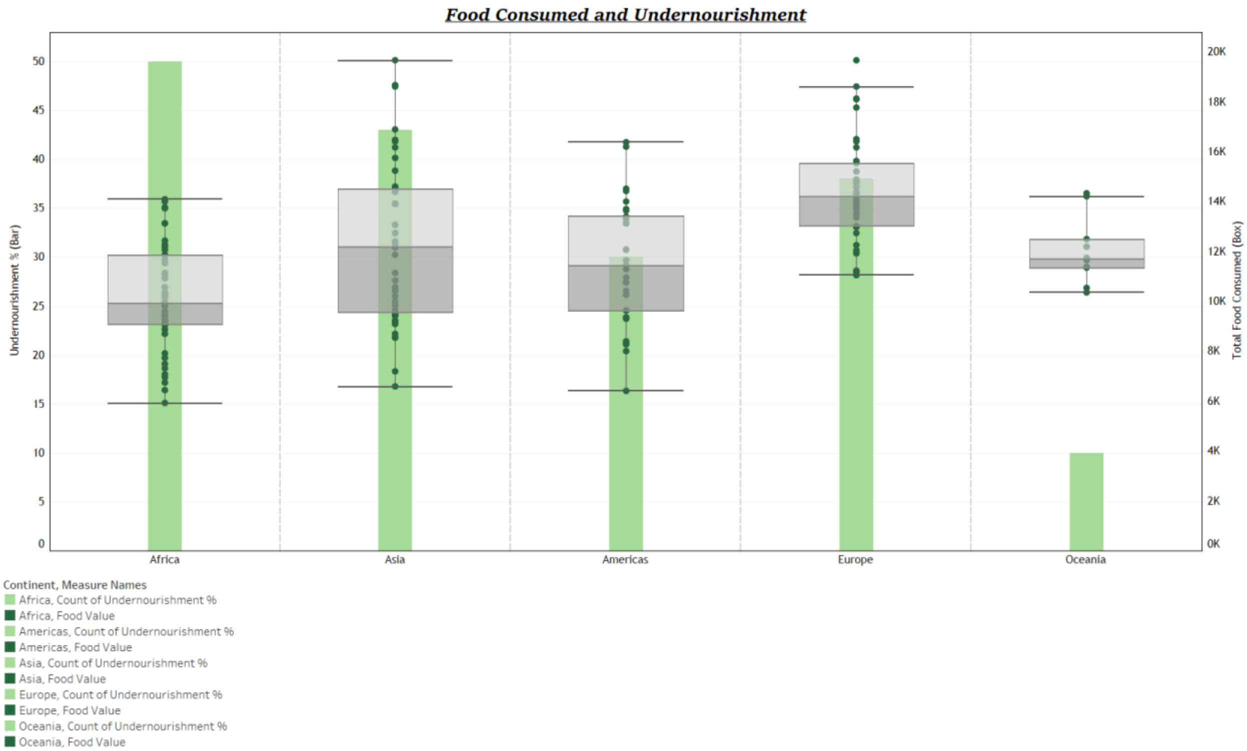


Figure 4-1. R1: Dual Bar/Box Plot showing the regional ranges of undernourishment and food consumption.

Next, I created a geographical map to showcase the physical scale of the country in relation to the country's undernourishment percentage. Here, I've focused on the Asian region. The specific range is between Isreal with lowest ~2% and Yemen with the highest ~40%.

Undernourishment Across the Globe

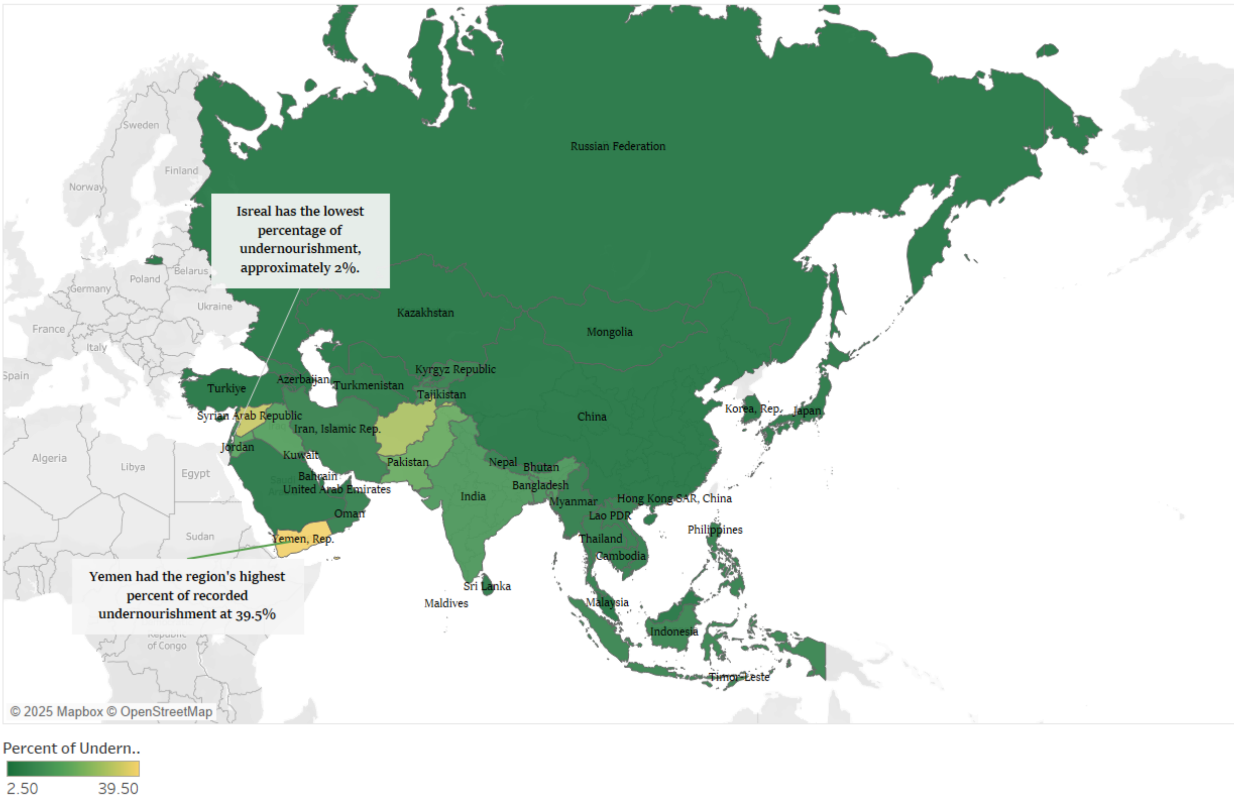


Figure 4-2. R2: Geographical map showing the regional specific range for percent of undernourishment.

Then, I wanted to investigate deeper into the relationship with the main contributing indicator of undernourishment (UN). dietary energy supply (DES). This indicator was chosen to strengthen the investigation with UN. With this plot, we can claim that there is a strong negatively linear correlation between DES and UN. The r-squared for the line of best fit is 0.57 with a p-value of <0.00001. This strength of correlation shouldn't be too surprising, but we can see clear groupings among the countries when a size scale and color scale are affixed to UN and

DES, respectively. Yemen, Syrian Arab Republic, and Afghanistan have relatively low DES with high UN. The opposite end consists of two countries, Isreal and Türkiye

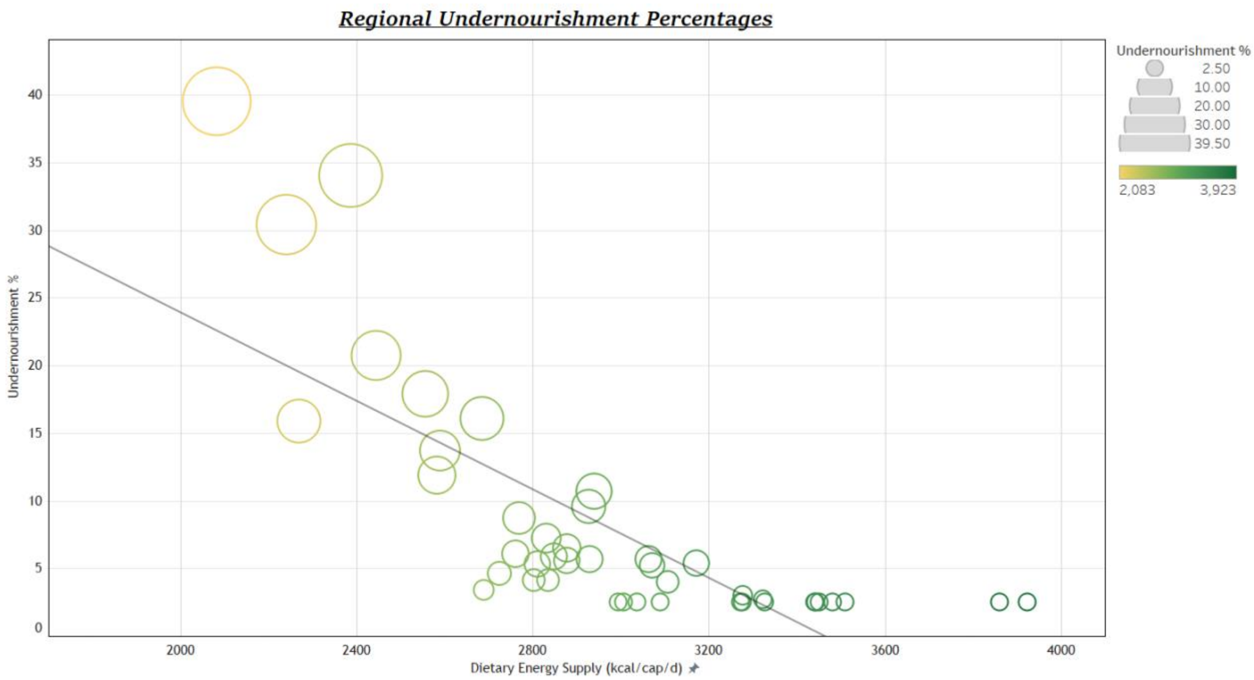


Figure 4-3. R3: Scatter plot showing the relationship (strong negative linear trend) between regional percent of undernourishment and energy supply.

And finally with a heat map, the composite total food consumed is shown for each country in the selected continent. The size and color represent the approximate total food value. I would have liked to add population to the graph to indicate the consumption rate.

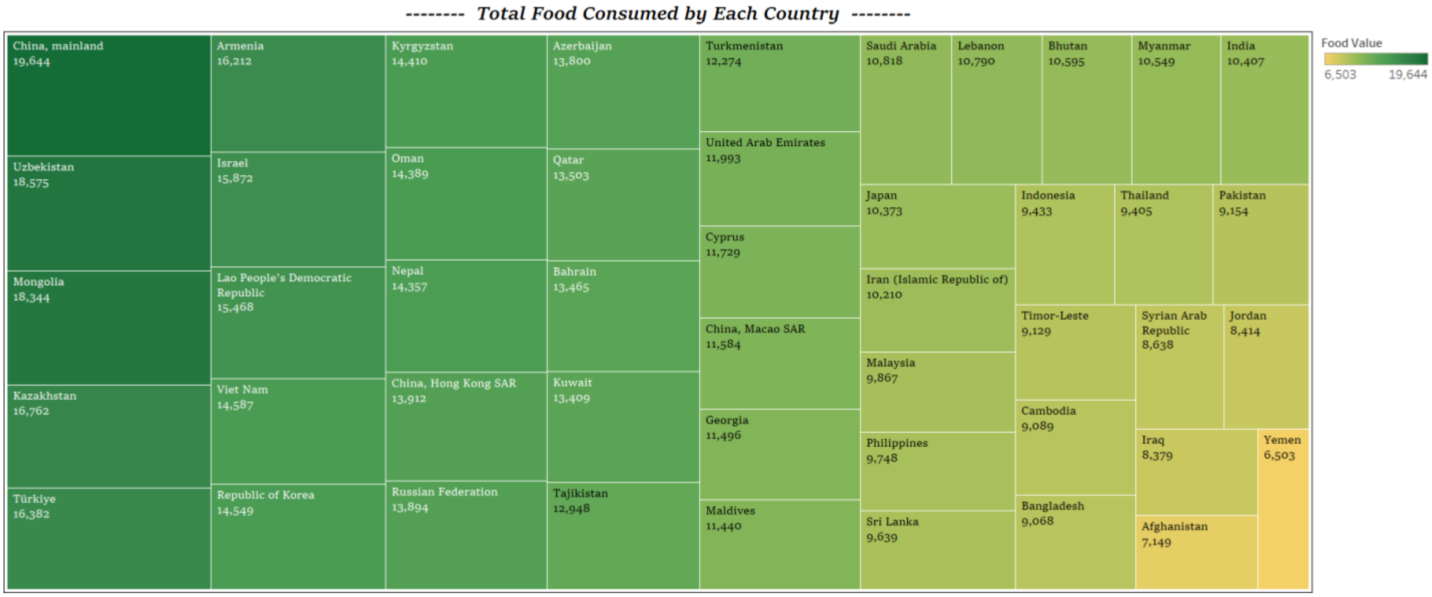


Figure 4-4. R4: Heat map showing the country breakdown of total food consumed (composite of all indicator values).

Key Takeaways from the Regional Analysis

- Dietary energy supply is highly correlated to undernourishment
- Clear groups among countries can be based on undernourishment percentages and energy supply
- Europe and Oceania have relatively lower undernourishment percentages, most countries below 25%
- Haiti and Somalia have the highest undernourishment percentages, reaching up to 50%

4.2. Life Expectancy Analysis (by Yongchan Lee)

Central question of this analysis is: "*Which types of food or nutrients are most closely associated with longer life expectancy?"* To manage the large number of countries in the dataset, I use new variables named *Continents* for the purpose of showing high-level (big region) patterns. Average life expectancy by continent will be shown through a **1) boxplotand**, and food consumption patterns across continents will be shown using a **2) heatmap**.

Beyond that, my main analysis focus moves to individual countries. I will use **3) scatterplots** to examine the relationship between specific nutrient supplies and life expectancy at the country level. Additionally, using a **4) bar chart** and **5) radar chart**, I will be comparing the top and bottom countries by life expectancy to explore how their dietary patterns differ. Even there is a variable *year*, I did not use it as a time series variable because life expectancy remained relatively stable from 2018 to 2022. So, I used the average values to focus more on cross-country differences and their relationship with food and nutrient supply.

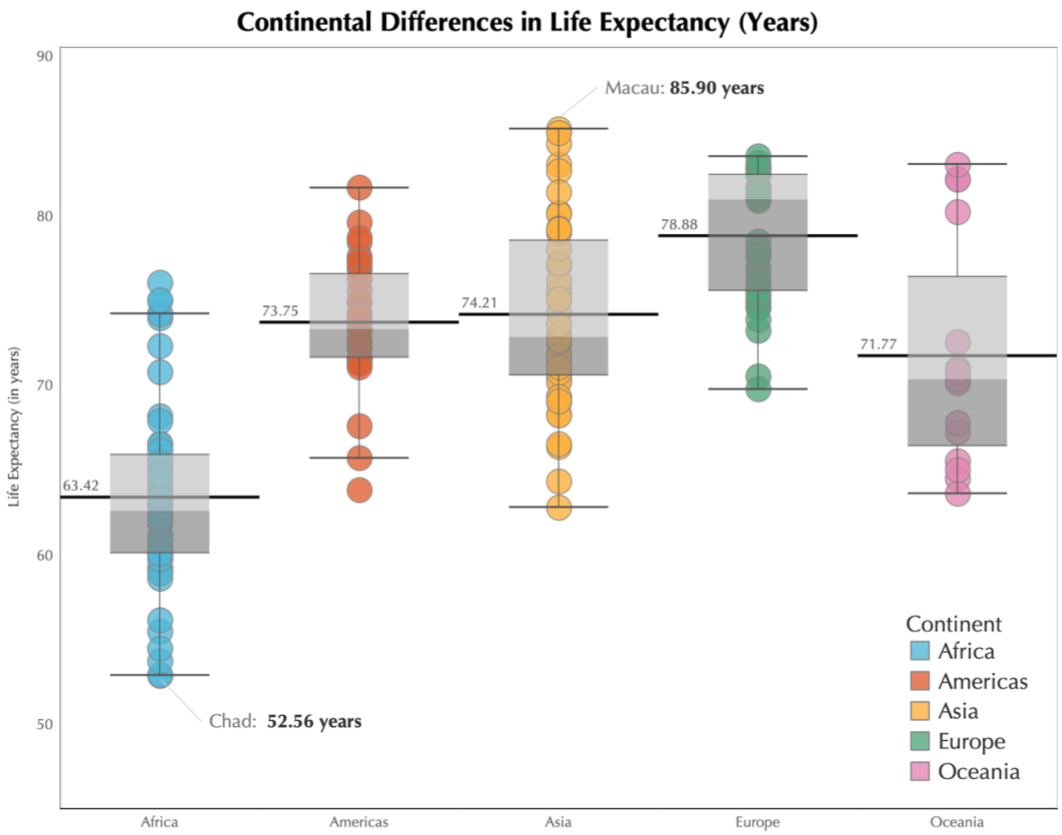


Figure 4-55. LE1: Continental Differences in Life Expectancy in Boxplot

The boxplot above shows average life expectancy by continent. The x-axis represents each continent, and the y-axis shows life expectancy in years. Each data point corresponds to the average life expectancy of a country from 2018 to 2022 within its respective continent. Additionally, the horizontal line within each box represents the mean life expectancy for that continent.

The continent with the lowest average life expectancy is Africa, at 63.42 years, followed by Oceania at 71.77 years, the Americas at 73.75 years, and Asia at 74.21 years. The highest average was recorded in Europe, with 78.88 years. However, one important insight is that the country with the highest life expectancy overall is not in Europe, but in Asia—Macau, with an average of 85.90 years. On the other hand, the lowest life expectancy was observed in Chad (Africa) at 52.56 years.

Another noteworthy insight is the wide gap in life expectancy among countries in Oceania, which contrasts with the more consistent life expectancy ranges seen in Africa, the Americas, Europe, and Asia. In the next figure, I will explore what food consumption by food groups patterns look like across these continents that might explain the differences in life expectancy.

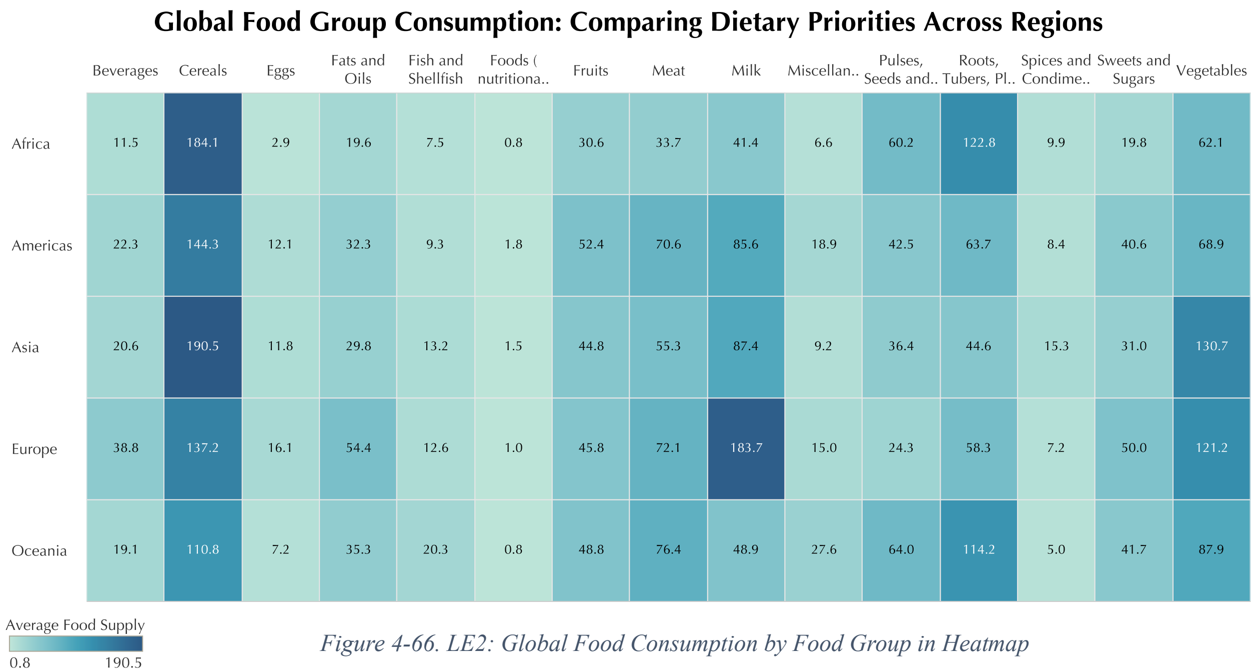


Figure 4-66. LE2: Global Food Consumption by Food Group in Heatmap

The heatmap above shows average food consumption by food group across different continents. One important point to clarify is the unit of measurement. As shown in *Table 2-1* in *2. DATA SOURCES*, the supply values represent nutrient supply within each food group, and each nutrient has its own unit of measurement (e.g., kcal, g, µg, etc.).

Because of this, the average values shown in the heatmap are based on a mix of units across nutrients, and therefore, cannot be represented using a single unit. Instead of labeling the values with specific units, we refer to them more generally as “consumption” values. For example, if a cell shows a value of 10, I interpret that as a relative consumption level of 10, without tying it to a specific unit.

First, I want to examine which food groups show relatively high consumption within each continent. Here, the comparison is not made between continents, but rather within each continent—comparing food groups against one another in the same region.

Starting with Africa, I observe higher levels of food consumption in the cereals, roots, nuts, and vegetables groups. In the Americas, consumption is relatively high in cereals, meat, and milk groups than other groups. For Asia, the main food groups with higher consumption include cereals, milk, and vegetables. Europe shows a similar pattern to Asia, with higher consumption in cereals, milk, meat, and vegetables groups. In Oceania, higher consumption is observed in cereals, meat, roots, and vegetables.

Across all continents, one common trend is the high consumption of cereals. In every region, the cereals group was either the highest or the second highest in terms of overall consumption. Another key finding is that Europe's milk consumption stands out significantly - at 183.7. This value is noticeably higher than that of any other continent.

Lastly, it is worth noting that the two continents with the highest average life expectancy, Europe and Asia, commonly share high consumption levels of both vegetables and milk. This suggests a potential link between these food groups and longer life expectancy.

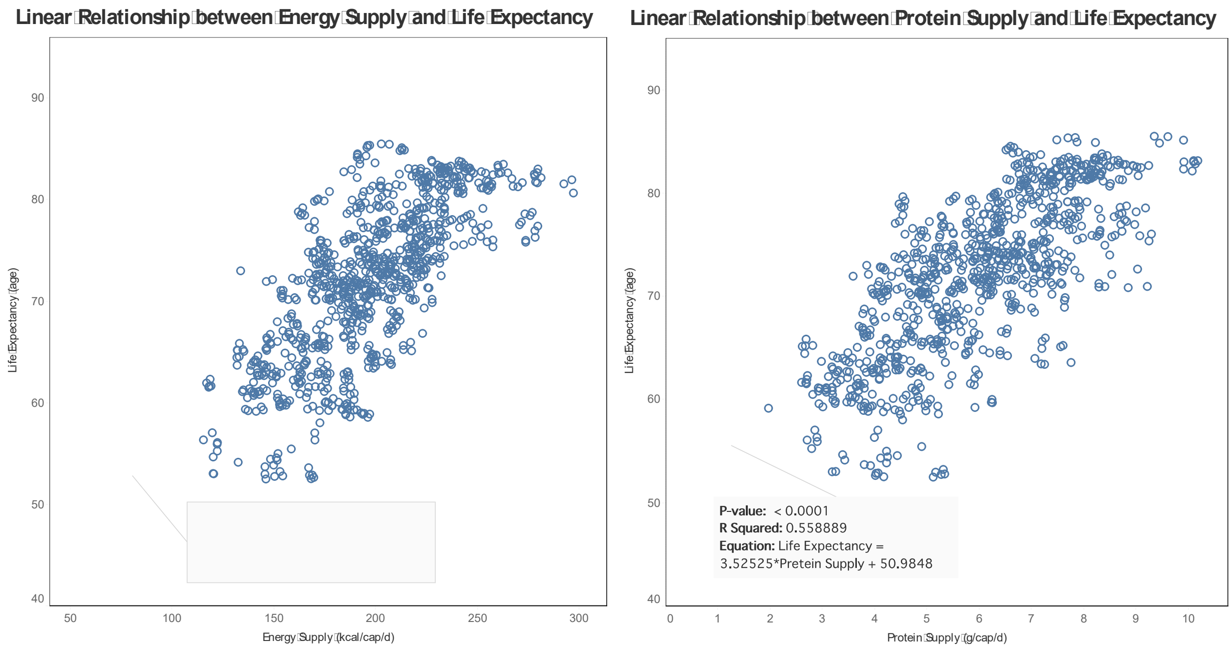


Figure 4-77. LE3: Linear Relationship between Nutrients and Life Expectancy in Scatterplot

So far, I have explored the relationship between food groups and life expectancy. Now, I want to shift the focus and ask: What kind of relationship exists between nutrient intake and life expectancy? To explore this, I analyzed two nutrients: Energy supply (measured in kilocalories per day) and Protein supply (measured in grams per day).

The two scatterplots visualize the relationship between each nutrient and life expectancy. Both plots reveal a positive correlation - in other words, countries that consume more energy and protein tend to have higher life expectancy.

According to the trend lines, a consistent increase of 10 kcal per day in energy supply is associated with a 1.6-year increase in life expectancy. Similarly, an additional 1 gram of protein per day correlates with a 3.5-year increase in life expectancy.

Both scatterplots have an R^2 value of approximately 0.5, which suggests that energy and protein intake together explain around 50% of the variance in life expectancy across countries.

Of course, we cannot conclude that there is a causal relationship, but this analysis does indicate a moderate correlation between these nutrients and life expectancy.

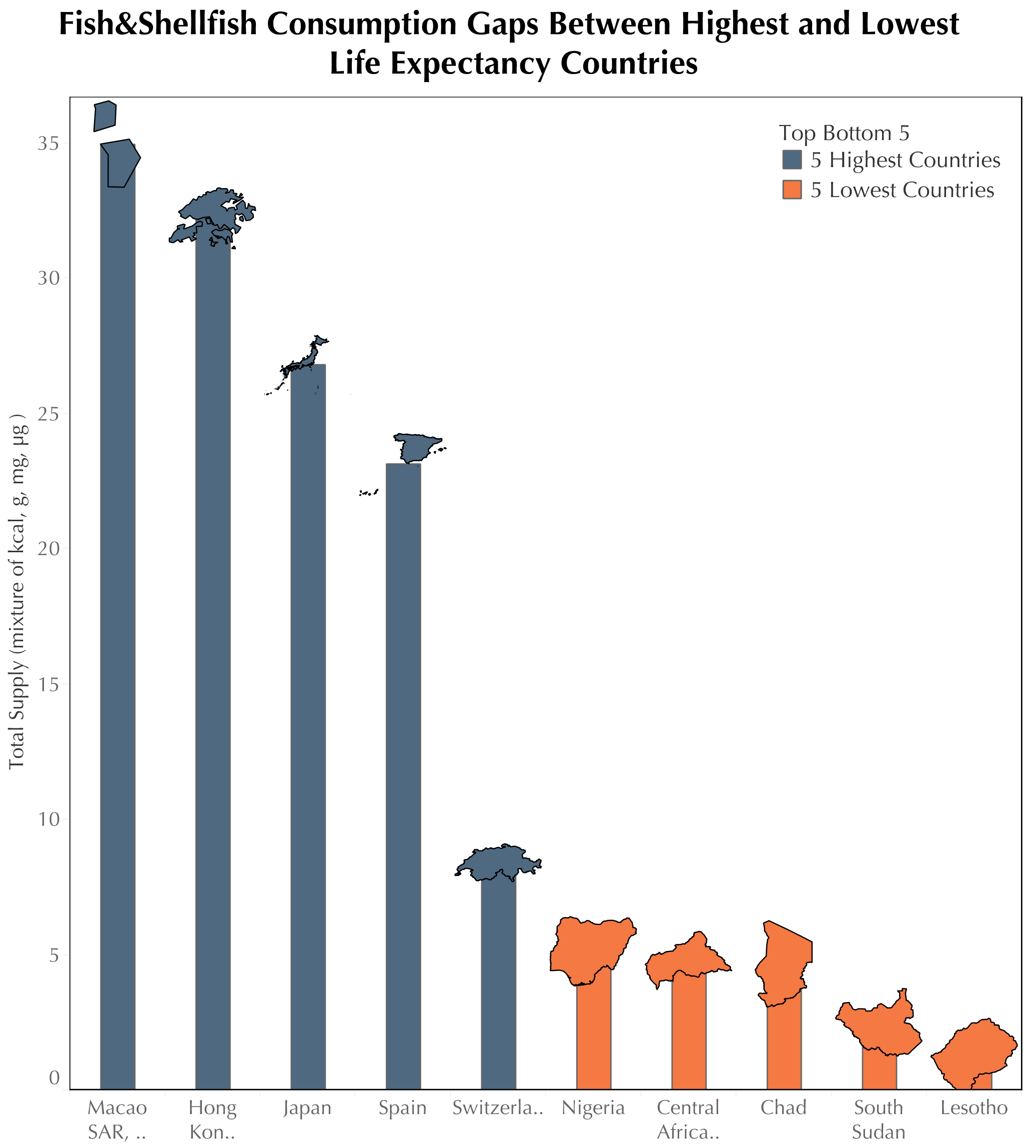


Figure 4-88. LE4: Fish and Shellfish Consumption Gaps between Highest and Lowest Life Expectancy Countries in Barchart

Until now, I have compared food consumption across continents and countries worldwide. In this section, I want to focus on analyzing the dietary patterns of the five countries with the highest life expectancy and the five with the lowest. After thoroughly examining food consumption by food group and by nutrient in these two groups, one important insight emerged:

The top 5 countries have significantly higher consumption of fish and shellfish compared to the bottom 5 countries.

The bar chart above clearly visualizes this finding. Countries with the highest life expectancy (blue bars) - Macau, Hong Kong, Japan, and Switzerland - show much higher total fish & shellfish supply than countries with the lowest life expectancy (orange bars) - Nigeria, Central African Republic, Chad, South Sudan, and Lesotho.

It is important to note that this difference cannot simply be attributed to whether a country is landlocked or coastal. For example, Nigeria has access to the sea, but its fish & shellfish consumption is lower than Switzerland, which is landlocked. This suggests that geographic location alone does not fully explain the difference in fish consumption patterns between these groups.

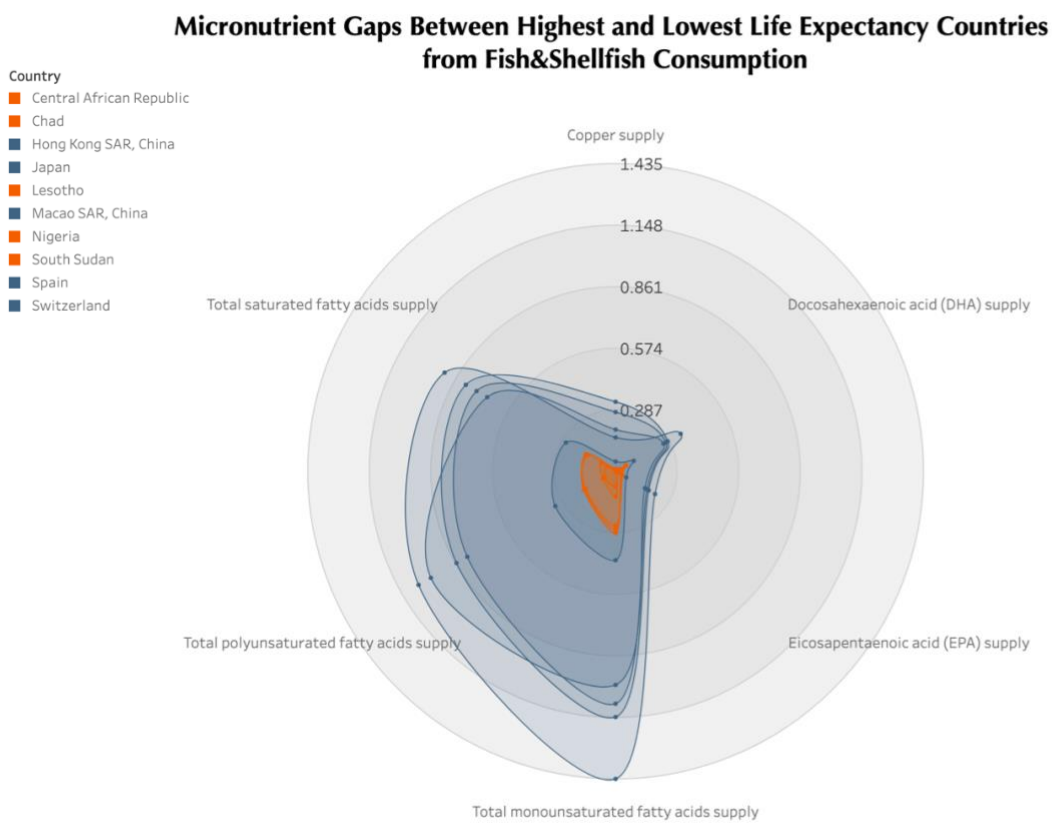


Figure 4-99. LE5: Micronutrients Gap between Highest and Lowest Life Expectancy Countries

Building on the previous analysis of food groups, we further examined the nutrient consumption patterns of the top 5 and bottom 5 countries in terms of life expectancy. One notable insight is that, beyond fish & shellfish, the top 5 countries outweighed the bottom 5 countries in their intake of six specific nutrients: Copper, Docosahexaenoic acid (DHA), Eicosapentaenoic acid (EPA), Total Monounsaturated Fatty Acids (MUFA), Total Polyunsaturated Fatty Acids (PUFA), and Total Saturated Fatty Acids (SFA).

This leads to a natural question: Why does the radar chart (Figure 4-5) focus on these six nutrients specifically within the fish & shellfish group? The answer is that, upon further inspection, I found that the majority of the supply of these nutrients comes from the fish & shellfish group in the top 5 countries. So, instead of looking at total nutrient intake from all food groups, the radar chart isolates nutrient supply specifically from fish & shellfish to reveal the key source of these nutritional advantages.

Importantly, this insight is supported by previous research. One study[4] examining global dietary trends and healthy aging found that higher intake of PUFA, particularly the omega-3 fatty acids EPA and DHA, was significantly associated with reduced functional impairments and chronic disease burden, and these positive features are leading to a healthier aging trajectory. These nutrients - primarily derived from fish - were found to be protective against unhealthy aging and positively correlated with higher healthy life expectancy.

When we combine this insight with Figure 4-4, one conclusion becomes clear: The top 5 countries not only consume more fish & shellfish than the bottom 5 but also receive significantly higher amounts of these six nutrients from that group.

---

[4] García-Esquinas, Esther, et al. "Dietary n-3 polyunsaturated fatty acids, fish intake and healthy ageing." International Journal of Epidemiology 48.6 (2019): 1914-1924

Looking at the radar chart, we see that Copper, DHA, and EPA appear lower than the other nutrients. However, this is expected because different nutrients naturally occur in different amounts. It does not mean they are unimportant, but it means fish & shellfish do not supply each nutrient in equal quantities.

What cannot be ignored is the clear visual trend: In all six nutrients, the blue line representing the top 5 countries is consistently and significantly larger than the orange line representing the bottom 5.

4.3. Trend Analysis (by Jessica Akonor)

This analysis is concerned with the relationship between the global consumption of sugar and the impact of sugar consumption on health conditions. Specifically, I aim to demonstrate the negative impacts of excessive sugar consumption on physical health. Diabetes is generally considered as the result of excessive sugar consumption. Thus, excessive sugar consumption could lead to a higher chance of getting the disease. The research question for this analysis is, “What are the effects of sugar consumption on health conditions?”

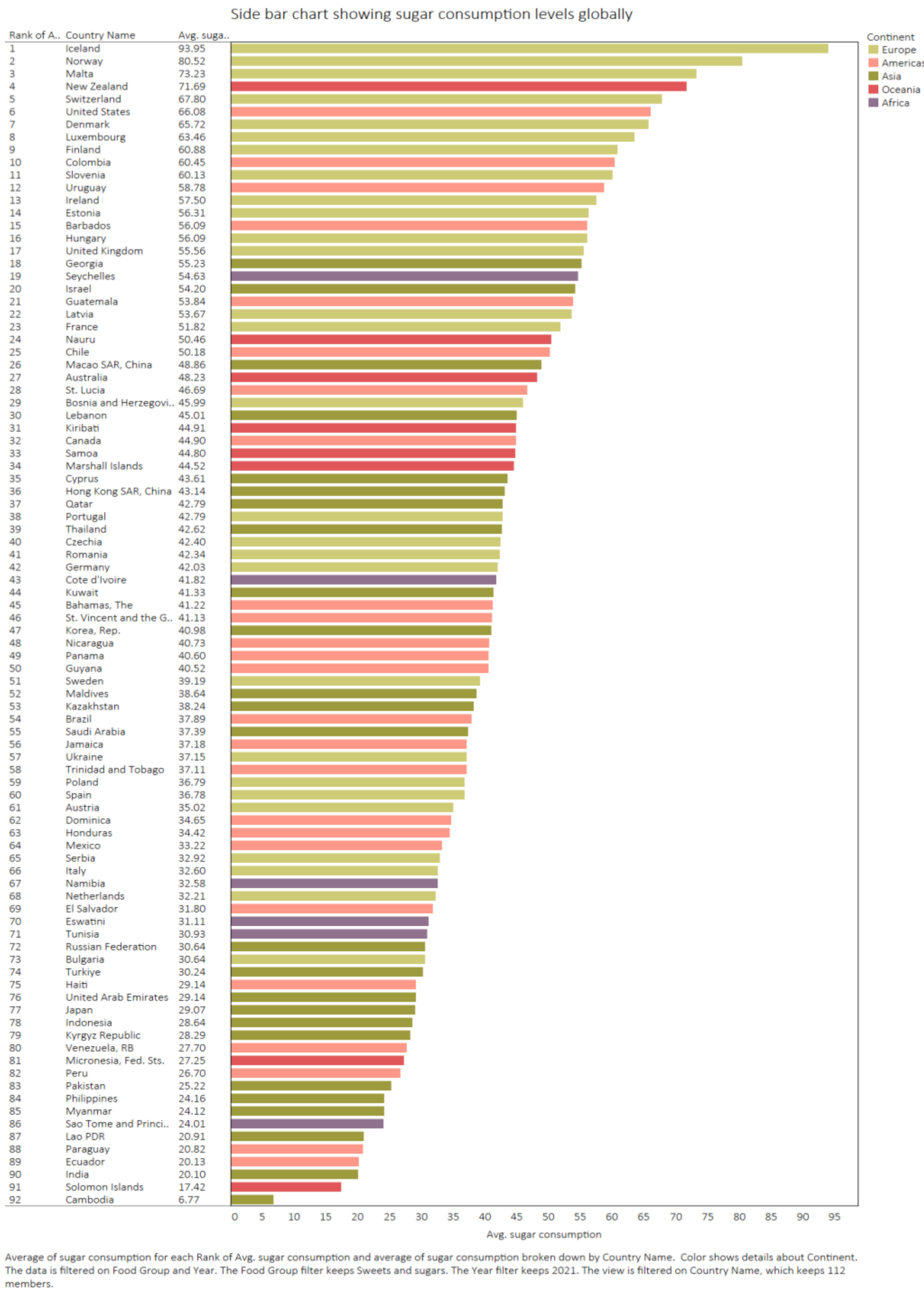


Figure 4-10.10 T1: Side Bar Chart of sugar consumption globally in the year 2021

As illustrated above in figure 4-6, we can observe that Iceland, Norway, Malta, New Zealand and Switzerland are the leading nations with respect to sugar consumption. From this chart, we can imply that the highest levels of sugar consumption can be found in countries or nations located in Europe. Thus, Europeans consume the most sugar in comparison to North American countries like the United States. Sugar consumption in Iceland is recorded at 93.95, Norway is 80.52 and Malta is recorded at 73.23. Average sugar consumption is measured in mg/cap/d. The country with the least average sugar consumption is Cambodia located in the continent of Asia recorded at 6.77.

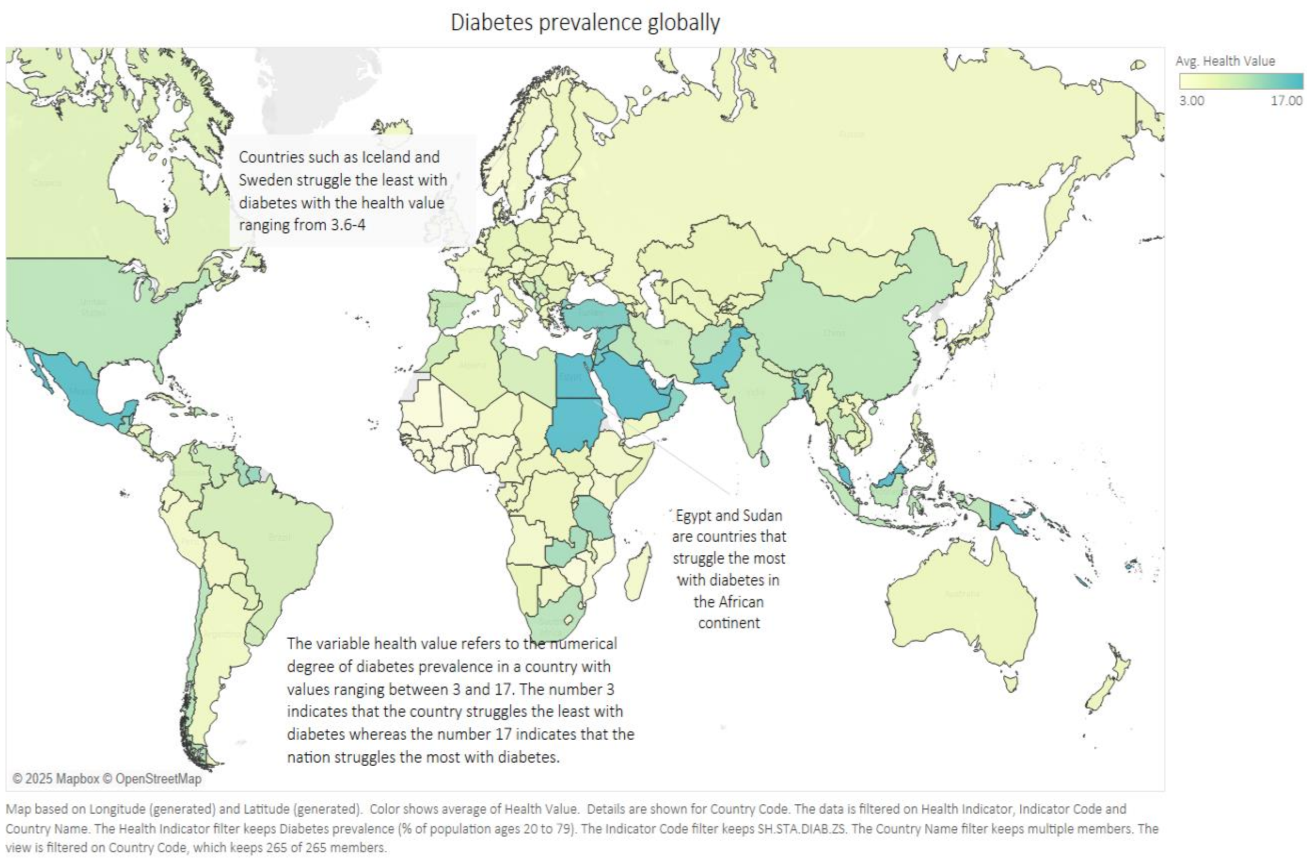


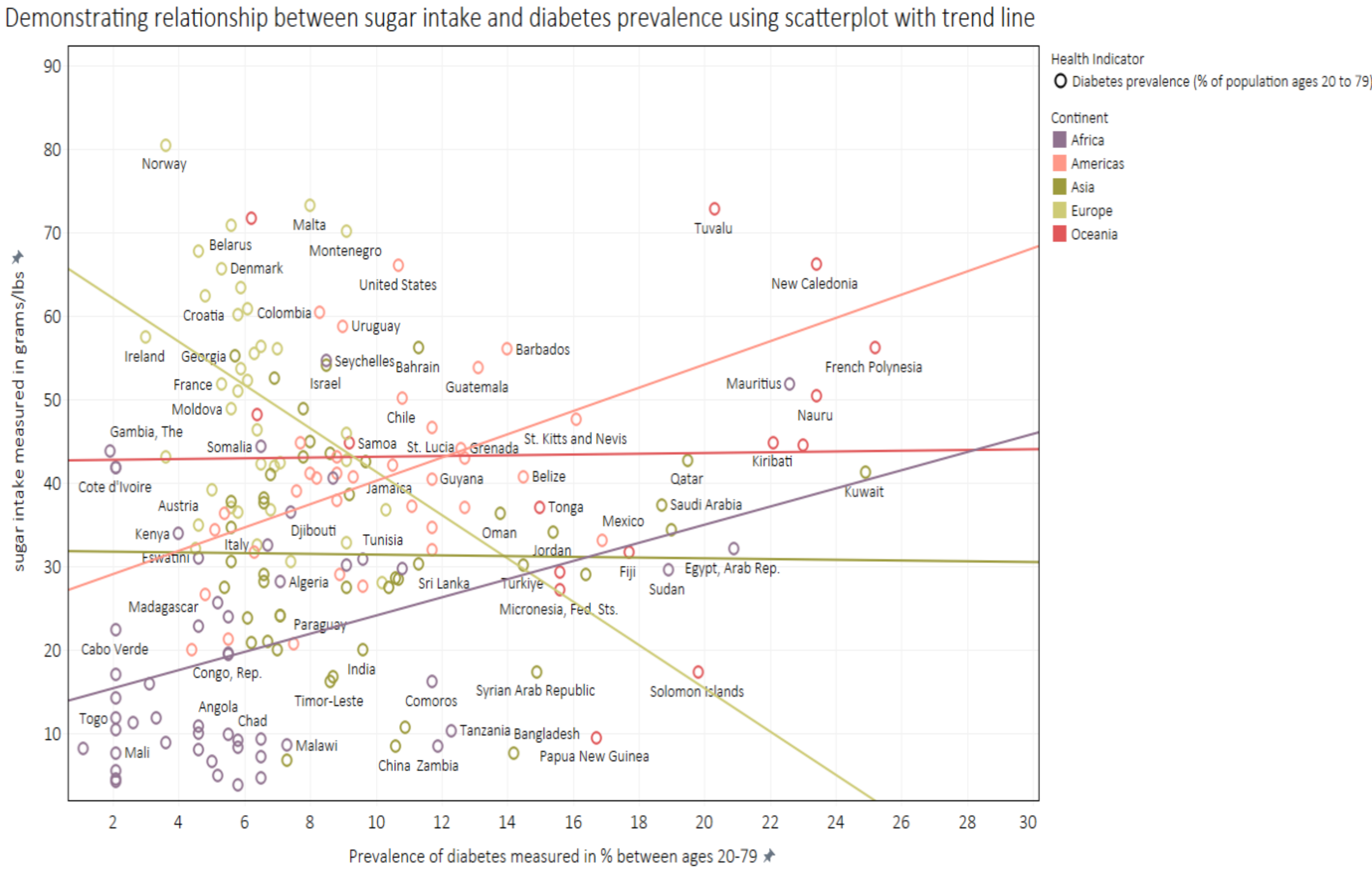
Figure 4-11.11 T2: Diabetes prevalence globally in the year 2021

From figure 4-7 above, we can observe the prevalence of diabetes on a global scale. This chart uses a sequential color scheme to symbolize the degree of diabetes prevalence in each country. I chose to use a geographic map to make it easier for observers to internalize and visualize the data given. Countries shaded in a light yellow or yellow color shows countries that are less likely to get diabetes with the least. Blue indicates that a country struggles with diabetes the most.

From the data given, countries such as Egypt, Sudan and Mexico struggle the most with diabetes even though they consume the least amounts of sugary goods globally. In Egypt, the health value is 20.9, Sudan is 18 and Mexico is 16.9. Even though countries such as Iceland and Norway consume goods with a large amount of sugar, they struggle the least with diabetes. In Norway, the health value is recorded at 3.6.

The scatterplot below in figure 4-8 demonstrates the relationship between sugar consumption and diabetes prevalence globally. This scatterplot is a visual representation of the way sugar consumption can result in diabetes globally. African countries such as Togo and Sierra Leone are shown in this diagram with purple rings. These purple rings are in the lower left section of the graph with sugar consumption ranging between 4.5 mg/cap/d to approximately 41.8 m/cap/d. In this scatterplot, Guinea Bissau is the African country with least sugar consumption rates with a smaller percentage of people who have diabetes. The health value indicator for diabetes prevalence is 2.1 and the sugar consumption is 4.23 mg/cap/d. The purple trend line indicates that the average sugar consumption for African countries in the continent is approximately 1.807 with the average health value being 13.263, the lowest health value indicator globally. Thus, African countries are less likely to consume sugary goods and get diabetes. We can also conclude that the diet culture in African countries incorporates more protein (in the form of fish and meat), vitamins (in the form of fruits and vegetables) and carbohydrates (in the form of rice and cassava-based dishes)

For European countries such as Iceland, Norway, Sweden and Switzerland, the trend line indicates that the average sugar consumption for European countries is -2.5 mg/cap/d with the average diabetes prevalence recorded at approximately 67.33. Figure 4-6 illustrated that the top consumers of sugary goods in Europe were Iceland and Norway yet on a global scale, their average consumption of sweets and sugar goods is recorded at -2.5 mg/cap/d although they are the most likely to fall sick with diabetes. These numbers reflect the fact that excessive sugar consumption can lead to diabetes, but it is only one factor out of many. Other factors that can be considered are exercise regimens and climate or weather conditions. European countries that receive the most rainfall compared to sunlight do not receive the necessary vitamins to strengthen their immune system. Hence, European countries are more prone to getting diabetes through other health factors outside of sugar consumption.



Average of Health Value vs. average of sugar consumption. Color shows details about Continent. Shape shows details about Health Indicator. The marks are labeled by Country Name. Details are shown for Country Name. The data is filtered on Indicator Code, Food Group and Year. The Indicator Code filter keeps SH.STA.DIAB.ZS. The Food Group filter keeps Sweets and sugars. The Year filter keeps 2021. The view is filtered on Health Indicator and Country Name. The Health Indicator filter keeps Diabetes prevalence (% of population ages 20 to 79). The Country Name filter keeps multiple members.

Figure 4-12.12 T3: Scatterplot illustrating the relationship between sugar consumption and the prevalence of diabetes globally in 2021

5. RECOMMENDATION

From Life Expectancy Analysis

- Increase consumption of fish and shellfish,
- Include more milk and dairy products
- Ensure sufficient energy and protein intake daily

From the Trend Analysis

- Countries in the global south (specifically African countries) should develop advanced infrastructure and health facilities to decrease diabetes prevalence in their countries.
- African countries like Guinea Bissau, Sierra Leone and Togo have the lowest rates of diabetes prevalence globally in addition to being least sugar consumers. This is because countries in Africa incorporate more carbohydrate and protein-based foods into their diets rather than sugar. African countries should continue eating balanced meals that incorporate vitamins from fruits, carbohydrates and protein from rice and meat.
- Even though Iceland and Norway are the largest consumers of sugar, they are the least likely to have diabetes. Global North countries such as Norway have established health networks and developed infrastructures to take care of diseases such as diabetes.
- European countries should consider incorporating more fruits, vegetables, protein and fibers to decrease the level of sugar consumption on a regional level and a global level.
- The trend lines in the scatterplot indicate that excessive sugar consumption can lead to diabetes but that is not the only reason. Regardless, established health structures and balanced diets can reduce diabetes and other sugar related diseases globally.