Materials and Methods

### Materials and Methods

#### 3.1. Materials

Data for the seven metrics was accumulated from various open-source databases as listed in Table.1. (Appendix).

##### 3.1.1. Food Nutrient Adequacy

###### Non-Staple Food Energy

The data for this indicator was taken from [FAO’s food balance sheets](https://www.fao.org/faostat/en/" \l "data/FBS). It contained the per capita daily energy intake from 98 types of foods in 186 countries globally. Preliminary examination of the data showed a small to moderate amount of missing data for each item (0 - 20%, refer Fig. 1.). Due to the data being Missing Completely At Random (MCAR), a relatively small amount of missing values for most food items, and the fact that each country had at least one item with missing data, a replacement of 0 for each missing value was done. The resulting dataset was aggregated such that each of the 98 food-item’s daily calory intake was aggregated into one of the following major food types; *1.) Cereals, 2. Roots and Tubers, 3. Meat and Milk products and eggs, 4. Fruits and Plantains, 5.) Vegetables, 6.) Seeds and Nuts and Oil, 7.) Starches and Sugars, and 8.) Infant Foods* (Dimensions of Need - Staple Foods: What Do People Eat? n.d.) and (Gustafson et al., 2016). The resulting dataset had values from *46* high income countries, *45* upper middle income, *44* lower middle income, and *21* low income (Appendix, Table. 2.). Further exploration of the dataset revealed extreme values among different food groups in different income groups. To maintain a robust sample size, median was selected as the measure of central tendency to measure the contribution of each food group in daily kilo calory intake from each income group (also used in (Chaudhary et al., 2018)). The median values, after being processed into percentage values, were filtered to include food groups whose median daily calorie intake represented anything less than 15 percent of total median calorie intake from the particular income group. It was selected so, bearing in mind the insights derived from the dataset itself and official definitions of staple foods as defined by (Mäkelä & Rautavirta, 2018) and (FAO, n.d.). The results are presented in Table 3., where the daily median calorie intakes from each of the 8 major food groups are presented, together with the classification of whether the food group is a staple for that particular income group or not.

**Percent of Median Daily Kilo Calorie Intake by Income group: Staple Foods vs Non-Staple Foods**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Income group** | **Staple Foods** | **Percent** | **Non-Staple Foods** | **Percent** |
| Low Income | Cereals | 54.6% | Roots & Tubers | 13.2% |
| Seeds, Nuts, and Oils | 12.3% |
| Meat, Milk and products & Eggs | 6.8% |
| Vegetables | 5.3% |
| Starch & Sugars | 4.3% |
| Fruits & Plantains | 3.5% |
| Infant Food | 0% |
| Lower Middle Income | Cereals | 52.1% | Meat, Milk and products & Eggs | 12.9% |
| Seeds, Nuts, and Oils | 12.2% |
| Starches & Sugars | 8.6% |
| Vegetables | 5.7% |
| Roots & Tubers | 4.7% |
| Fruits & Plantains | 3.7% |
| Infant Food | 0% |
| Upper Middle Income | Cereals | 40.6% | Seeds, Nuts, and Oils | 13.9% |
| Starches & Sugars | 12.2% |
| Vegetables | 5% |
| Meat, Milk and products & Eggs | 20.7% | Fruits & Plantains | 4.3% |
| Roots & Tubers | 3.3% |
| Infant Food | 0.1% |
| High Income | Cereals | 31.9% | Starches & Sugars | 11.5% |
| Fruits & Plantains | 4.1% |
| Meat, Milk and products & Eggs | 30.3% | Vegetables | 3.9% |
| Seeds, Nuts, and Oils | 15.4% | Roots & Tubers | 2.8% |
| Infant Food | 0.1% |

###### Shannon Diversity

This indicator measures diversity of nutrients in a food system by measuring the quantity in weight of various food items available in that food system. Data for the annual per capita quantity supply of food items in countries globally, is readily available at FAO. Data for similar food items as the ones used in the first indicator was obtained from FAO’s food balance sheets. A preliminary examination of the data showed considerable missing values in some food items, and here, like in the previous indicator, the type and extent of missing data, together with previous knowledge on the domain of subject, determined the method of handling missing data. The type of missing data was both MNAR (Missing Not At Random) and MAR (Missing at Random), as some food items such as *Non-centrifugal sugars* had a high percentage of missingness in all four income groups, suggesting there being some reporting problems that caused the data for both 4 income groups to be missing (hence MNAR), and some, such as *millet and products*, have a high percentage of missingness in only three groups (High income, Low middle income and Upper middle income), suggesting that there is another factor influencing the missingness of this variable (hence MAR). It was decided to remove all food items with more than 40% of missing data in any of the 4 income groups as practical recommendation from (*What Should Be the Allowed Percentage of Missing Values?* 2015) and Hair, J. F. (2009) suggest. Imputation was done using the MICE and KNN multiple-imputation methods, and the results of the R statistic were compared to each other. The figure showing the percentage of missing values from each food group and the KNN, MICE R-statistic comparison can be found at the **supplimentary material file**, attached at the end of the document. Imputation by using MICE proved to be more robust (with the adjusted R-square of 0.89), and thus the dataset imputed by using it was the one used for the next procedures of data preparation. The data was then aggregated into seven food groups as the one used in the first indicator, and below is a table that contains a snapshot of that data.

**Preview: Annual Per Capita Supply of Food In-terms of Food Groups Globally (2019).**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Country** | **Income  group** | **Region** | **Cereals** | **Roots &  Tubers** | **Meat, Milk and  Products & Eggs** | **Vegetables** | **Fruits &  Plantains** | **Seeds, Nuts,  and Oils** | **Starch and  Sugars** | **Infant  Foods** |
| Spain | High  income | Europe &  Central Asia | 113.1 | 60.3 | 277.9 | 152.2 | 67.9 | 48.5 | 32.5 | 0.4 |
| Chile | High  income | Latin America  & Caribbean | 153.3 | 57.9 | 173.6 | 81 | 56.2 | 10.8 | 41.2 | 0 |
| Iceland | High  income | Europe  & Central Asia | 84.4 | 35.6 | 397.5 | 94.3 | 74.7 | 28 | 80.6 | 0.3 |
| Sweden | High  income | Europe  & Central Asia | 115 | 53.9 | 316.1 | 104.7 | 44.9 | 21.8 | 46.5 | 0.3 |
| Slovenia | High  income | Europe  & Central Asia | 140.1 | 38.1 | 255.2 | 109.3 | 96.6 | 16.6 | 44.7 | 0 |
| Norway | High  income | Europe  & Central Asia | 121.9 | 50.4 | 283.3 | 105.6 | 60.6 | 17.5 | 44.7 | 0.4 |
| Ireland | High  income | Europe  & Central Asia | 138.5 | 74.1 | 261 | 102.6 | 53.5 | 22.1 | 81.1 | 0.4 |
| Nauru | High  income | East Asia  & Pacific | 63.9 | 10.1 | 175 | 37.3 | 75.3 | 139.7 | 87 | 0.2 |
| Hungary | High  income | Europe  & Central Asia | 124 | 38.7 | 240.3 | 97.2 | 70.9 | 13.1 | 52.1 | 0.1 |
| Cyprus | High  income | Europe  & Central Asia | 139.3 | 16.1 | 222.5 | 97.2 | 54.8 | 25 | 56.4 | 1.2 |
| Germany | High  income | Europe  & Central Asia | 99.7 | 69 | 280.6 | 101.3 | 67.8 | 31.5 | 43.7 | 0 |
| New Zealand | High  income | East Asia  & Pacific | 114.3 | 47.8 | 256.3 | 134.4 | 56.7 | 20 | 61.1 | 0.3 |
| Croatia | High  income | Europe  & Central Asia | 113 | 39 | 236.5 | 305.3 | 55.6 | 19.1 | 86.1 | 0.2 |
| Denmark | High  income | Europe  & Central Asia | 111.3 | 62 | 368.2 | 108.9 | 50.5 | 16.4 | 54.5 | 0 |