Data Nutrient Label

_____

About
This dataset is a set of death caused by blood pressure, high sodium, low whole grains, alcohol, low in fruits, unsafe water, secondhand smoking, low birth weight, child wasting, unsafe sex, low in nuts/seeds, pollution, low in vegetables, first hand smoking, high glucose, vitamin deficiency and iron deficiency. In relation to the inflation rates for 302 countries in the years 2007 - 2019.

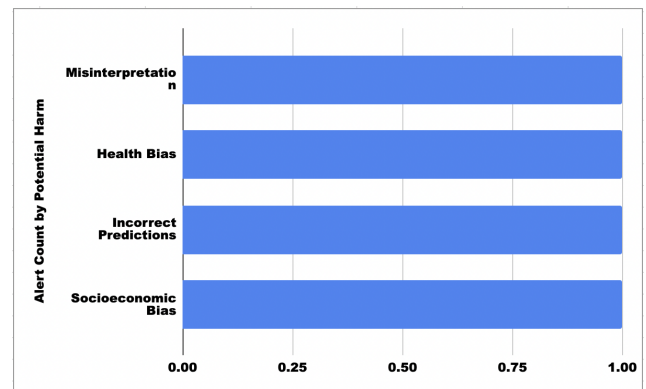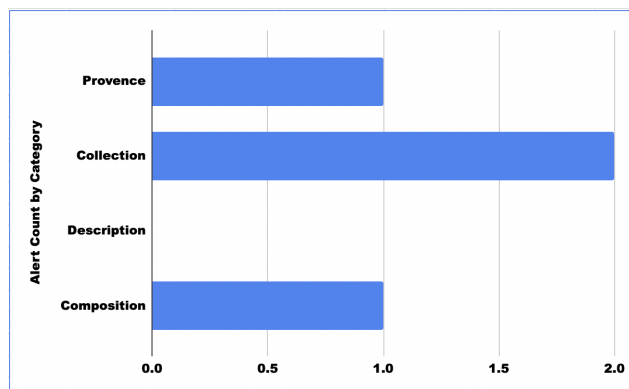**Data Creation Ranges:** January 2007 - December 2019
Created By: The World Bank & Our World in Data
Content: List of CSV Files
Sources: https://ourworldindata.org/obesity,
https://datacatalog.worldbank.org/search/dataset/0060165/Monthly-food-price-inflation-estimates-by-country

Alert Count	5
Provence	**1**
Misinterpretation	1
Collection	**2**
Health Bias	1
Incorrect Predictions	1
Description	**0**
Composition	**1**
Socioeconomic Bias	1

Use Cases:

1. What is the cost of healthy diets in each country?
2. What is the affordability of healthy diets in each country?
3. What is the population's physical access to sufficient quantities of locally available items for an active and healthy life?
4. What is the population's economic access to sufficient quantities of locally available items for an active and healthy life?
5. What are the key indicators on the cost and affordability of an energy-sufficient diet?
6. What are the key indicators on the cost and affordability of a nutrient-adequate diet?

Badges:

Dataset Information

Workflow

1. **TELL US ABOUT THIS DATASET.**
   This dataset is a set of food prices in relationship to mortality rates in each country.
2. **IS THERE AN INTENDED PURPOSE FOR THE DATASET? WHAT DOMAIN WAS IT DESIGNED FOR?**
   The purpose of this dataset is to assess the relationship between food prices/affordability in comparison to death rates in each country.
3. **WHAT IS THE LICENSE UNDER WHICH THE DATASET IS MADE AVAILABLE?**
   Public Domain
4. **ARE THERE ANY ADDITIONAL RESTRICTIONS UNDER WHICH THE DATASET IS MADE AVAILABLE? ARE THERE ANY LEGISLATION, CODES OF PRACTICE OR GUIDANCE FROM THE JURISDICTION OR DOMAIN IN WHICH THE DATA WAS COLLECTED ABOUT THE USE OF THIS DATA? ARE THERE TASKS FOR WHICH THE DATASET CANNOT BE USED?**

PLEASE PROVIDE A DESCRIPTION OR GUIDANCE.

No

5. **ARE THERE TASKS FOR WHICH THE DATASET WOULD BE CAUTIONED AGAINST BEING USED? IF SO, PLEASE PROVIDE A DESCRIPTION.**

No

6. **HAS THE DATASET BEEN USED FOR ANY TASKS ALREADY? IF SO, PLEASE PROVIDE A DESCRIPTION AND LINKS TO PAPERS OR SYSTEMS USING THE DATASET**

No

Composition

1. **DOES THE DATASET HAVE A METADATA REPOSITORY OR DATA DICTIONARY? IF YES, PLEASE PROVIDE THE LINK AND IF NOT, PLEASE EXPLAIN WHAT EACH FIELD MEANS.**

No data dictionary exists for this dataset.

2. **IN THE CASE OF EXPLICITLY RELATED INSTANCES, ARE THESE RELATIONSHIPS INDICATED? (E.G., RELATING AN INSTANCE ABOUT A USER TO INSTANCES ABOUT THEIR POSTS)**

No

3. **WHAT DATA DOES EACH INSTANCE CONSIST OF? FOR EXAMPLE, DOES IT CONSIST OF RAW DATA (E.G., UNPROCESSED TEXT OR IMAGES) OR FEATURES? IN EITHER CASE, PLEASE PROVIDE A DESCRIPTION.**

This contains raw data

4. **HOW MANY INSTANCES ARE THERE IN TOTAL (OF EACH TYPE, IF APPROPRIATE)?**

There are 35 different variables

5. **IS THERE A LABEL OR TARGET ASSOCIATED WITH EACH INSTANCE?**

No.

6. **IS ANY INFORMATION MISSING FROM INDIVIDUAL INSTANCES? IF SO, WHAT IS MISSING AND WHY?**

Some places have NA as data for that was probably not available during that time of year, years that are in the 1900's data is not as available

7. **IS YOUR DATASET A SAMPLE? IF SO, WHAT WAS THE SAMPLING STRATEGY USED (E.G. DETERMINISTIC, PROBABILISTIC WITH SPECIFIC SAMPLING PROBABILITIES), AND DOES IT ACCURATELY REPRESENTS THE INTENDED OUTPUT?**

No

8. **WAS ANY PREPROCESSING/CLEANING/LABELING OF THE DATA DONE? IF SO, PLEASE PROVIDE A DESCRIPTION.**

Yes we did have to parse out the date to take year in order to combine the two data sets

9. **IF THE DATA WAS PROCESSED, WAS THE RAW DATA SAVED IN ADDITION TO THE PREPROCESSED/CLEANED/LABELED DATA? IF YES, PLEASE PROVIDE A LINK TO THE RAW DATA.**
   N/A

10. **ARE THERE RECOMMENDED DATA SPLITS? IF SO, PLEASE PROVIDE A DESCRIPTION OF THESE SPLITS, EXPLAINING THE RATIONALE BEHIND THEM.**
    No

11. **ARE THERE ANY ERRORS, SOURCES OF NOISE, OR REDUNDANCIES IN THE DATASET? IF SO, PLEASE PROVIDE A DESCRIPTION**
    No.

12. **DOES THE DATASET INCLUDE INFORMATION FROM UPSTREAM SOURCES? IF SO, NAME THESE SOURCES, THEIR STABILITY, AND ANY KNOWN USAGE LIMITATIONS.**
    No

13. **IF THE DATASET INCLUDES INFORMATION FROM UPSTREAM SOURCES, IS THIS DATASET TECHNICALLY DEPENDENT UPON THOSE SOURCES?**
    N/A

14. **IS THERE CONFIDENTIAL DATA INCLUDED IN THIS DATASET?**
    No

15. **DOES THE DATASET IDENTIFY ANY SUBPOPULATIONS (E.G., BY AGE, GENDER)? IF YES, PLEASE INDICATE THEIR DISTRIBUTIONS WITHIN THE DATASET WHERE POSSIBLE / AVAILABLE.**
    No

Provenance

1. **WHO CREATED THE DATASET AND ON BEHALF OF WHICH ENTITY?**
   This data was created by The World Bank Group

2. **HOW IS / WAS THE COLLECTION AND MANAGEMENT OF THIS DATASET FUNDED?**
   Privately-funded

Collection

1. **OVER WHAT TIMEFRAME WAS THE DATA COLLECTED? IF THE DATA CONTINUES TO BE UPDATED, PLEASE INDICATE.**
   The dataset was collected over the time period of 2017 to 2021.

2. **WHAT MECHANISMS OR PROCEDURES WERE USED TO COLLECT THE DATA?**
   The database version Food Prices for Nutrition 1.0 houses indicators estimated in July 2022, when relevant affordability indicators were calculated based on the available Poverty and Inequality Platform (PIP) data from the World Bank expressed in 2011 purchasing power parity terms (PPP).

3. **WHAT IS THE RELATIONSHIP BETWEEN THE DATASET COLLECTOR AND OWNERS OR MANAGERS OF THE DATASET?**
   The original dataset is raw data collected by the World Bank.

4. **IF INDIVIDUALS' DATA IS INCLUDED IN THIS DATASET, DID THOSE INDIVIDUALS CONSENT TO THE COLLECTION AND USE OF THEIR DATA? IF SO, PLEASE DESCRIBE CONSENT PROCEDURE.**
   Not applicable, the data is public and available to use by all.

5. **IF INDIVIDUALS' DATA IS INCLUDED IN THIS DATASET, WAS THIS DATA ALTERED TO ENSURE HIGHER LEVELS OF PRIVACY? IF YES, PLEASE DESCRIBE ANY PRIVACY PROCEDURES FOLLOWED WITH REGARDS TO THIS DATASET (ANONYMIZATION EFFORTS, PRIVACY PROTOCOLS, SUPPRESSION TECHNIQUES, ETC).**
   No, the information has not been altered since there was no data specifically endangering anyone's identity or private information.

6. **HAS THE DATA BEEN REVIEWED FOR QUALITY?**
   Yes.

7. **WERE ANY ETHICAL REVIEW PROCESSES CONDUCTED (E.G., BY AN INSTITUTIONAL REVIEW BOARD)? IF SO, PLEASE PROVIDE A DESCRIPTION OF THESE REVIEW PROCESSES, INCLUDING THE OUTCOMES, AS WELL AS A LINK OR OTHER ACCESS POINT TO ANY SUPPORTING DOCUMENTATION.**
   No.

Management

1. **HOW IS THE DATA BEING MANAGED AT REST AND IN TRANSIT?**
   Internally

2. **HOW CAN THE DATA BE ACCESSED? IF THIS IS GOING TO CHANGE OVER TIME, PLEASE INDICATE THIS.**
   Data can be accessed through the website provided where the public can then download it.

3. **WILL THE DATASET BE DISTRIBUTED TO ANY OTHER INDIVIDUALS/THIRD PARTIES (E.G. COMPANIES, INSTITUTIONS, ORGANIZATIONS)? IF YES, WHEN WILL THE DATASET BE DISTRIBUTED AND HOW WILL THE DATASET BE DISTRIBUTED (E.G. GITHUB, TARBALL ON THE WEBSITE, ETC)?**
   No it will not.

4. **WILL THE DATASET BE UPDATED (E.G., TO CORRECT LABELING ERRORS, ADD NEW INSTANCES, DELETE INSTANCES)? IF SO, PLEASE DESCRIBE HOW OFTEN, BY WHOM, AND HOW UPDATES WILL BE COMMUNICATED TO USERS (E.G., MAILING LIST, GITHUB)?**
   Data is updated yearly with estimations based on previous years' data.

5. **IS THERE A MECHANISM THROUGH WHICH AN INDIVIDUAL CAN REQUEST THEIR PERSONAL INFORMATION BE REMOVED?**
   Not applicable.