

Data Immersion

Achievement 6

Levels and trends in child malnutrition

Advanced Analytics &
Dashboard Design

Objective

The JME includes estimates of prevalence and numbers affected for stunting, overweight, wasting and severe wasting among children under five years of age at country, regional and global levels. Global, regional and country annual trends from 2000-2022 are available for stunting and overweight. The objective is to build an interactive dashboard that will visually showcase well-curated results of an advanced exploratory analysis conducted in Python to measure nutritional imbalance such as undernutrition (assessed from stunting, wasting and underweight) or overweight.

Why the analysis

Every child has the right to good nutrition. Well-nourished children grow and develop to their full potential. They are better equipped to lead healthy lives, to be free from poverty, to learn and participate, and to continue thriving across the life course, with benefits that continue over generations. The past decade has seen important gains in improving maternal and child nutrition, including a one-third decline in the proportion of children suffering from stunting. Yet the triple burden of malnutrition – stunting, wasting and overweight – continues to jeopardize children's ability to survive and thrive.

All forms of malnutrition are preventable. To stop malnutrition before it starts, children and their families need access to nutritious diets, essential services and positive practices to set them on the path to survival and thrive. But today, these vital pathways to good nutrition are under growing threat as many countries plunge deep into a global food and nutrition crisis fuelled by poverty, conflict, climate change and the enduring secondary effects of the COVID-19 pandemic. As the world responds to the crisis, urgent action is critical to protect maternal and child nutrition – especially in the most affected regions – and secure a future where the right to nutrition is a reality for every child.

Key Questions and Objectives

- Identify the number of deaths and affected due to malnutrition.
- Which countries are affected most by child nutrition?
- Does progress to reduce stunting has equal across regions and sub-regions?.
- Find the malnutrition and country's population proportion.

- SDG target by 2030, end all forms of malnutrition, including achieving, by 2025, the internationally agreed targets on stunting and wasting in children under 5 years of age. Will this target be achievable by the trends in dataset?

Scope

SDG target by 2030, end all forms of malnutrition, including achieving, by 2025, the internationally agreed targets on stunting and wasting in children under five years of age.

Data source

- The UNICEF, WHO and the World Bank inter-agency team update the Joint Child Malnutrition Estimates (JME) every other year.
- The JME includes estimates of prevalence and numbers affected for stunting, overweight, wasting and severe wasting among children under 5 years of age at country, regional and global level. Global, regional and country annual trends.
- Data Source: United Nations Children's Fund (UNICEF), World Health Organization (WHO), World Bank (WB)
- Data Collection: Data is collected various national organisation's health survey departments. The list of organisations are defined here.

Source	Short source for major survey types.
DHS	Demographic and Health Survey (from the official Demographic and Health Survey Programme available at https://dhsprogram.com/)
DHS-style or MICS-style survey	Survey with MICS or DHS in the titled but which is not and official MICS or DHS/not available at the above websites
LSMS	The Living Standards Measurement Study (from official World Bank Group supported surveys available at: http://iresearch.worldbank.org/lsmssurveyfinder.htm)
MICS	Multiple Indicator Cluster Survey (from official UNICEF-supported surveys available at http://mics.unicef.org/)
NNS	National Nutrition Survey
SMART	Standardized Monitoring and Assessment of Relief and Transitions

Nutrition Surveillance	Estimates originating from nutrition monitoring and surveillance systems which are nationally representative
ONS	Other National Survey
PAPFAM	Pan Arab Project for Family Health survey

Data Contents: The JME includes estimates of prevalence and numbers affected for stunting, overweight, wasting and severe wasting among children under 5 years of age at country, regional and global level. Global, regional and country annual trends from 2000-2022 are available for stunting and overweight. Refer Data Profiling (section 1. About the data and data type) for more information.

- **Data Relevance:** The data is part UN's Sustainable Development Goals (SDG) and accurate to their standards. Data can be used for study, research and analysis. Data contains relevant columns for the analysis and meets project objectives.
 - **Objective 1:** Identify the number of deaths and affected due to malnutrition(stunting, overweight, wasting and severe wasting).
 - **Hypothesis:** Malnutrition and mortality among children's counts reducing each year.
 - **Datasource**
 - UNICEF_Global_Databases_Mortality dataset contains mortality rate in each geographic regions with year.
 - Limitation: Data isn't recorded every year.
 - UNICEF_Global_Joint-Malnutrition-Estimates contains global stunting, wasting and overweight information.
 - **Objective 2:** Which countries are affected most by child nutrition(stunting, overweight, wasting and severe wasting)?
 - **Hypothesis:** Countries with low income have the highest child malnutrition count.
 - **Datasource:**
 - UNICEF_Global_Joint-Malnutrition-Estimates

- UNICEF_Global_Databases_Overweight
- UNICEF_Global_Databases_Stunting
- UNICEF_Global_Databases_Databases_Wasting

All the above table have enough columns to find the trends in child nutrition growth in each country with period over the years.

- **Objective 3:** Does progress to reduce stunting has equal across regions and sub-regions?
- **Hypothesis:** Progress to reduce stunting has been equal across regions and sub-regions.
- **Datasource:**
 - UNICEF_Global_Joint-Malnutrition-Estimates
 - UNICEF_Global_Databases_Overweight
 - UNICEF_Global_Databases_Stunting
 - UNICEF_Global_Databases_Databases_Wasting

All the above table have enough columns to find the trends in child nutrition growth in each country with period over the years.

- **Objective 4:** Find the trend in countries income and malnutrition growth.
- **Hypothesis:** countries income and malnutrition growth worldwide affected with countries income.
- **Datasource:**

- **Objective 5:** SDG target by 2030, end all forms of malnutrition, including achieving, by 2025, the internationally agreed targets on stunting and wasting in children under 5 years of age. Will this target achievable by the trends in dataset?
- **Hypothesis:** More intensive efforts are required if the world is to achieve the global target of reducing the number of children with stunting to 89 million by 2030.
- **Datasource:**
 - UNICEF_Global_Joint-Malnutrition-Estimates
 - UNICEF_Global_Databases_Overweight
 - UNICEF_Global_Databases_Stunting
 - UNICEF_Global_Databases_Databases_Wasting

All the above table have enough columns to find the trends in child nutrition growth in each country with period over the years.

- **Data Bias and Gaps:**

{Need more analysis}

- **Completeness of Data:**

{Need more analysis}

- **Limitations and Ethics**

- Data is collected by UN from various countries reports.
- The data could be possibly politically biased to provide countries reputation.
- Collected data have clear transparency from which organisation is collected, in which year and survey timing. And the data can be accessed by any one globally.
- The data doesn't contain any personal information or to target specific community.
- All the data is collected from the UNICEF WHO database and the data can be used to for research, study or analysis purpose.

About the data and Data Types

Wasting refers to a child who is too thin for his or her height. Wasting results from recent rapid weight loss or the failure to gain weight. A child who is moderately or severely wasted has an increased risk of death, but treatment is possible.

Table:

- UNICEF_Databases_Severe_Wasting
- UNICEF_Global_Databases_Databases_Wasting

Column Name	Description	Type
ISO	Three-digit alphabetical codes International Standard ISO 3166-1 assigned by the International Organization for Standardization (ISO). http://www.iso.org/iso/home/standards/country_codes	Characters(3), Qualitative- Nominal
Countries and areas	The UNICEF Global databases contain a set of 202 countries as reported on through the State of the	Characters(20), Qualitative- Nominal

	World's Children Statistical Annex 2023	
UNICEF Regions	EAP: East Asia and the Pacific EECA: Eastern Europe and Central Asia ESA: East and Southern Africa LAC: Latin America and the Caribbean MENA: Middle East and North Africa NA : North America SA : South Asia WCA: West and Central Africa WE: Western Europe	Characters(4), Qualitative- Nominal
UN Regions		Characters(4), Qualitative- Nominal
WHO Regionns		Characters(4), Qualitative- Nominal
World Bank Income Groups		Characters(4), Qualitative- Nominal
Survey Years	Represents the year(s) in which the data collection (e.g. survey interviews) took place	Character Qualitative - Ordinal
Year	Year assigned to each survey	Number Quantitative-Continuous Time variant

Short Source	Short source for major survey types. Demographic and Health Survey (from the official Demographic and Health Survey Programme available at https://dhsprogram.com/)	Characters
Full Source Title		Characters
Latest Estimate		Characters
Estimate Type	The source of the estimate, whether as reported, reanalyzed, etc. Reanalyzed : Dataset has been obtained, standardized and reanalyzed to conform to the definition. Lower and Upper Confidence limits are based on a 95% confidence level External Reanalysis : Dataset reanalyzed by external partners Reported: Point Estimate, and when available, other information including sample size, Lower and Upper Confidence limits, are based on results from report.	Characters Qualitative- Nominal
<p>National, Sex, Residence , Age Group , Wealth Quintile Wealth Quintile Grouping, Mother's Education, Area / Wealth Quintile, Area / Wealth Quintile Grouping Sex / Age Group .Subnational Region:</p> <p>The above column contains all the estimate values for the above categories, For exaple National contains records Point Esimate, Lower Limit, Upper Limit, Sample Size and foot note.</p> <p>Value:</p> <ul style="list-style-type: none"> Percentage of children under 5 years of age falling below -3 standard deviations severe wasting from the median weight-for-height of the reference population Percentage of children under 5 years of age falling below -2 standard deviations wasting from the median weight-for-height of the reference population 		
Point Estimate	The prevalence of the wasting (severe) in the population	Continuous-Discrete Time variant

Lower Limit	95% lower confidence interval of the point estimate	Continuous-Discrete Time variant
Upper Limit	95% upper confidence interval of the point estimate	Continuous-Discrete Time variant
Sample Size	The weighted sample size which the point estimate is based on	Continuous-Discrete Time variant
Footnote	<p>A variety of notes that relate to the point estimates as follows:</p> <p>a: Alternate definition (Numbered footnote contains details)</p> <p>b: Age-adjusted estimate</p> <p>c: Converted estimate</p> <p>d: Adjusted National-Rural to National</p> <p>p: Age calculated using century month codes</p> <p>q: Oedema data were collected in the survey</p> <p>r: Height/length modality (standing/lying) was not collected</p> <p>s : Oedema data was collected in the survey and not considered in the analysis</p> <p>y : Unweighted sample size is less than 25, point estimate, lower limit and upper limit have been suppressed</p> <p>z : Unweighted sample size is between 25 and 49</p>	Qualitative- Nominal

Stunting r refers to a child who is too short for his or her age. Children affected by stunting can suffer irreversible physical and cognitive damage accompanying stunted growth. The devastating consequences of stunting can last a lifetime and even affect the next generation.

Table:

- UNICEF_Global_Databases_Stunting

Column Name	Description	Type
ISO	Three-digit alphabetical codes International Standard ISO 3166-1 assigned by the International Organization for Standardization (ISO). http://www.iso.org/iso/home/standards/country_codes	Characters(3),Qualitative-Nominal
Countries and areas	The UNICEF Global databases contain a set of 202 countries as reported on through the State of the World's Children Statistical Annex 2023	Characters(20), Qualitative- Nominal
UNICEF Regions	EAP: East Asia and the Pacific EECA: Eastern Europe and Central Asia ESA: East and Southern Africa LAC: Latin America and the Caribbean MENA: Middle East and North Africa NA : North America SA : South Asia WCA: West and Central Africa WE: Western Europe	Characters(4), Qualitative- Nominal
UN Regions		Characters(4), Qualitative- Nominal
WHO Regionns		Characters(4),

		Qualitative- Nominal
World Bank Income Groups		Characters(4), Qualitative- Nominal
Survey Years	Represents the year(s) in which the data collection (e.g. survey interviews) took place	Character Qualitative - Ordinal
Year	Year assigned to each survey	Number Quantitative-Continuous Time variant
Short Source	Short source for major survey types. Demographic and Health Survey (from the official Demographic and Health Survey Programme available at https://dhsprogram.com/)	Characters
Full Source Title		Characters
Latest Estimate		Characters
Estimate Type	The source of the estimate, whether as reported, reanalyzed, etc. Reanalyzed : Dataset has been obtained, standardized and reanalyzed to conform to the definition. Lower and Upper Confidence limits are based on a 95% confidence level External Reanalysis : Dataset reanalyzed by external partners Reported: Point Estimate, and when available, other information including sample size, Lower and Upper Confidence limits, are	Characters Qualitative- Nominal

	based on results from report.	
<p>National, Sex, Residence , Age Group , Wealth Quintile Wealth Quintile Grouping, Mother's Education, Area / Wealth Quintile, Area / Wealth Quintile Grouping Sex / Age Group .Subnational Region:</p> <p>The above column contains all the estimate values for the above categories, For exaple National contains records Point Esimate, Lower Limit, Upper Limit, Sample Size and foot note</p>		
Point Estimate	The prevalence of the stunting (moderate and severe) in the population	Continuous-Discrete
Lower Limit	95% lower confidence interval of the point estimate	Continuous-Discrete
Upper Limit	95% upper confidence interval of the point estimate	Continuous-Discrete
Sample Size	The weighted sample size which the point estimate is based on	Continuous-Discrete
Footnote	<p>A variety of notes that relate to the point estimates as follows:</p> <p>a: Alternate definition (Numbered footnote contains details)</p> <p>b: Age-adjusted estimate</p> <p>c: Converted estimate</p> <p>d: Adjusted National-Rural to National</p> <p>p: Age calculated using century month codes</p> <p>q: Oedema data were collected in the survey</p> <p>r: Height/length modality (standing/lying) was not collected</p> <p>s : Oedema data was collected in the survey and not considered in the analysis</p> <p>y : Unweighted sample size is less than 25, point estimate, lower limit and upper limit have been suppressed</p> <p>z : Unweighted sample size is between 25 and 49</p>	Qualitative- Nominal

--	--	--

Overweight: refers to a child who is too short for his or her age. Children affected by stunting can suffer irreversible physical and cognitive damage accompanying stunted growth. The devastating consequences of stunting can last a lifetime and even affect the next generation.

Table:

- UNICEF_Global_Databases_Overweight

Column Name	Description	Type
ISO	Three-digit alphabetical codes International Standard ISO 3166-1 assigned by the International Organization for Standardization (ISO). http://www.iso.org/iso/home/standards/country_codes	Characters(3),Qualitative-Nominal
Countries and areas	The UNICEF Global databases contain a set of 202 countries as reported on through the State of the World's Children Statistical Annex 2023	Characters(20), Qualitative- Nominal
UNICEF Regions	EAP: East Asia and the Pacific EECA: Eastern Europe and Central Asia ESA: East and Southern Africa LAC: Latin America and the Caribbean MENA: Middle East and North Africa NA : North America	Characters(4), Qualitative- Nominal

	SA : South Asia WCA: West and Central Africa WE: Western Europe	
UN Regions		Characters(4), Qualitative- Nominal
WHO Regionns		Characters(4), Qualitative- Nominal
World Bank Income Groups		Characters(4), Qualitative- Nominal
Survey Years	Represents the year(s) in which the data collection (e.g. survey interviews) took place	Character Qualitative - Ordinal
Year	Year assigned to each survey	Number Quantitative-Continuous Time variant
Short Source	Short source for major survey types. Demographic and Health Survey (from the official Demographic and Health Survey Programme available at https://dhsprogram.com/)	Characters
Full Source Title		Characters
Latest Estimate		Characters
Estimate Type	The source of the estimate, whether as reported, reanalyzed, etc.	Characters Qualitative- Nominal

	<p>Reanalyzed : Dataset has been obtained, standardized and reanalyzed to conform to the definition. Lower and Upper Confidence limits are based on a 95% confidence level</p> <p>External Reanalysis : Dataset reanalyzed by external partners</p> <p>Reported: Point Estimate, and when available, other information including sample size, Lower and Upper Confidence limits, are based on results from report.</p>	
--	--	--

National, Sex, Residence , Age Group , Wealth Quintile Wealth Quintile Grouping, Mother's Education, Area / Wealth Quintile, Area / Wealth Quintile Grouping Sex / Age Group
 .Subnational Region:

The above column contains all the estimate values for the above categories, For exaple National contains records Point Esimate, Lower Limit, Upper Limit, Sample Size and foot note

Point Estimate	The prevalence of the wasting (severe) in the population	Continuous-Discrete
Lower Limit	95% lower confidence interval of the point estimate	Continuous-Discrete
Upper Limit	95% upper confidence interval of the point estimate	Continuous-Discrete
Sample Size	The weighted sample size which the point estimate is based on	Continuous-Discrete
Footnote	<p>A variety of notes that relate to the point estimates as follows:</p> <p>a: Alternate definition (Numbered footnote contains details)</p> <p>b: Age-adjusted estimate</p> <p>c: Converted estimate</p> <p>d: Adjusted National-Rural to National</p> <p>p: Age calculated using century month codes</p>	Qualitative- Nominal

	<p>q: Oedema data were collected in the survey</p> <p>r: Height/length modality (standing/lying) was not collected</p> <p>s : Oedema data was collected in the survey and not considered in the analysis</p> <p>y : Unweighted sample size is less than 25, point estimate, lower limit and upper limit have been suppressed</p> <p>z : Unweighted sample size is between 25 and 49</p>	
--	---	--

Data Profiling

- **Profile and clean a data set ready for analysis**
- **Improve the integrity of data in each data set**

1. Convert Excel data into CSV format and extract relevant columns that are required for analysis.
2. Data are integrated into different .csv format under 02 data/Prepared data/01 CSV/

Data Analysis Criteria for this Project:

Be open source. - source. - Include non-anonymized column names.	Data Source https://data.unicef.org/resources/dataset/malnutrition-data/
Come from an authentic/authoritative	The UNICEF, WHO and the World Bank inter-agency team update the Joint Child Malnutrition Estimates (JME) every other

	year. And the data can be used for analysis, study or the research purpose.
Include non-anonymized column names.	Databases containing national country values from survey and administrative sources, disaggregated by various background characteristics such as place of residence, geographic location, age, and sex, when available.
Be recent (ideally, no more than 3 years old. However, this factor is not essential - if you've found a perfect data set for your purposes, it could be older too, but not more than 10 years old).	The UNICEF, WHO and the World Bank inter-agency team update the Joint Child Malnutrition Estimates (JME) every other year. The data collected on May 2023 from WHO database.
Be recent (ideally, no more than 3 years old. However, this factor is not essential - if you've found a perfect data set for your purposes, it could be older too, but not more than 10 years old).	The data contains all the essentials requirement for this course UNICEF, WHO and the World Bank inter-agency team update the Joint Child Malnutrition Estimates (JME) every other year
Contain at least 2-3 continuous variables (apart from index variables, ID variables, dates, years, etc).	Yes, the data set contains continuous variable
Contain at least 2-3 categorical variables (apart from index variables, ID variables, dates, years etc).	The table contains year and geographical information.
Contain at least 1,500 rows.	Yes
Include a geographical object of some kind: for instance, a column relating to a country, continent, or something similar. If the information from the data set refers to the US, for example, there should be a column containing the names/abbreviations of the states. Note: there should be at least a couple of different values in this column. This is important for the geospatial analysis you'll be conducting. Of course, you can also use data sets with latitude and longitude	The data is collected from worldwide and contains all regional geographical information like country and continental information.

Analysis Criteria

Data Analyst will be conducting the following analyses in this Achievement (note that not all results need to be included in the final dashboard):

- Exploratory analysis through visualizations (scatterplots, correlation heatmaps, pair plots and categorical plots)
- Geospatial analysis using a shapefile
- Regression analysis

- Cluster analysis
- Time-series analysis
- Analysis narrative and
- final results (presented in a dashboard)

Dashboard Requirements

As explained above, the dashboard needs to tell the story of your analytical journey. it should contain a curation of the key results that discovered as you worked through the Achievement.

Final dashboard must: -

- Be designed with a use-case in mind (answering key guidance questions).
- Be created in Tableau Public. - Be interactive.
- Adhere to visual design best practices.
- Include an introduction page that describes the project (data and purpose).
- Include relevant result(s) of initial visual exploratory analysis.
- Include an explanation for how the results of the exploratory analysis resulted in defining research questions and/or hypotheses.
- Contain a geospatial component.
- Address the defined questions/hypotheses using advanced analytical techniques.
- For example: - Regression analysis - Cluster analysis - Time-series analysis -
- Include a results summary page explaining how the results do or don't address the initial research questions/hypotheses.
- It should include details on the limitations of the project.
- Include a proposal of next steps for further analysis.