Kenya - Analysis of Supermarket Grocery Data for Prediction of Nutritional and Health Outcomes at the Population Level - Supermarket A

Agnes Kiragga

Report generated on: January 14, 2026

Visit our data catalog at: https://microdataportal.aphrc.org/index.php

Overview

Identification

ID NUMBER
DDI-KEN-APHRC-SUPERMARKET-A-2023-V1.0

Version

VERSION DESCRIPTION

PRODUCTION DATE
2025-06-23

NOTES
Not Applicable

Overview

ABSTRACT
Rates of overweight, obesity, and chronic diseases such as cardiovascular diseases, hypertension, type 2 diabetes and certain cancers (bowel, lung, prostate and uterine) are on the rise in most sub-saharan Africa (SSA) countries like kenya. These increases can be largely attributed to the shift toward unhealthy diet patterns and increased access to processed foods that are high in fat, sugar, and sodium. The influx of supermarkets in east africa and the replacement of traditional foods for processed foods places this region in a vulnerable position for greater increases in chronic disease rates. Consumer purchasing history from supermarkets can provide valuable insight to food intake over time and the present and future effects on chronic diseases. Purchasing data from supermarkets is available yet underutilized in SSA.

The study aimed to harmonize and increase accessibility to grocery data, use statistical methods to explore purcharing patterns and predict the effects of nutrition on chronic diseases, and inform policy on the various influences on consumer purchases.

UNITS OF ANALYSIS
Individuals and supermarket transaction records.

Scope

NOTES
- Transaction Level Data: Food Item Details (Specific products purchased), Quantity, Price, Date of Purchase, Location of Purchase, Customer Demographics (age, gender when collected via loyalty programs), Payment Method(Cash, credit card, digital payment, etc.), Basket Composition

The standardized form is provided as external resources data.

V1-V24 the questions are found in the "Study abstraction tool"

V25-V27 are generated classifications (user developed) and are not in any resource

V28 the questions are found in the "NOVA-Classification-Reference-Sheet"

V29-V56 the questions are found in the "Kenya Food Composition Tables 2018"

KEYWORDS

Supermarkets, Ultra-Processed Foods, Processed Foods, Non-communicable Chronic Diseases

Coverage

GEOGRAPHIC COVERAGE

National coverage: Nairobi, Nakuru, Kajiado, Machakos and Kirinyaga.

UNIVERSE

The survey covers transaction records of individuals who made purchases in supermarkets.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
Agnes Kiragga	African Population and Health Research Centre (APHRC)

OTHER PRODUCER(S)

Name	Affiliation	Role
Steve Cygu	African Population and Health Research Center (APHRC)	Co-Investigator - Study coordination and Co-lead
Maureen Ng'etich	African Population and Health Research Center (APHRC)	Co-Investigator - Study coordination and Co-lead
Lindsey English	African Population and Health Research Center (APHRC)	Co-Investigator - Supporting methods of data mapping, analysis, and nutrition policy
Reinpeter Momanyi	African Population and Health Research Center (APHRC)	Co-Investigator - Supporting methods of data mapping and analysis
Elizabeth Kimani	African Population and Health Research Center (APHRC)	Co-Investigator - Support methods of data analysis and co-lead in policy analysis
Gershim Asiki	African Population and Health Research Center (APHRC)	Co-Investigator - Support methods of data analysis and co-lead in policy analysis

FUNDING

Name	Abbreviation	Role
African Populattion and Health Research Center	APHRC	Funder (Big Idea)

OTHER ACKNOWLEDGEMENTS

Name	Affiliation	Role
Bonface Ingumba	African Population and Health Research Center (APHRC)	Data Governance Officer
Shem Mambe	African Population and Health Research Center (APHRC)	Data Documentation Officer

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
African Population and Health Research Center	APHRC		Documentation of the DDI

DATE OF METADATA PRODUCTION

2025-06-23

DDI DOCUMENT VERSION
Version 1.1 (June 2025)

DDI DOCUMENT ID
DDI-KEN-APHRC-SUPERMARKET-A-2023-V1.0

Sampling

Sampling Procedure

The study is a cross-sectional exploratory study with a phased approach employing quantitative secondary data collection from a third-party information management solution provider. The third party provider employs an open integrated point of sale and store information retail system that connects retail touch points and sales channels in several counties in Kenya.

Sampling was conducted after a census of all supermarkets subscribed to the third party system was done. Only those counties with supermarkets subscribed to the platform were sampled. A sample of large, medium sized and small supermarkets were selected to participate in the study. The supermarket sizes were determined as follows; large supermarkets (supermarkets with a cumulative total of more than 8 branch networks). Medium size supermarkets will be those with 3-8 branch networks in the counties and smaller supermarkets are those with 1-2 branch networks.

Grocery data was received from 10 supermarket chains.

Deviations from Sample Design

Not Applicable

Response Rate

Not Applicable

Weighting

Not Applicable

Questionnaires

Overview

A standardized form was developed to guide in extration of information from 3rd party information provider for supermarket purchase data. Variables of interest includes supermarket name, supermarket branch, location of supermarket, invoice id, customer id, customer demographics (gender, age), date and time of purchase, product name purchased, unit price per item, number of items purchased, payment method used by customer for purchase etc.

Secondary data collected will not be identifiable as it will be anonymized at the supermarket and client level.

The standardized form is provided as external resources data.
V1-V24 the questions are found in the “Study abstraction tool”
V25-V27 are generated classifications (user developed) and are not in any resource
V28 the questions are found in the “NOVA-Classification-Reference-Sheet”
V29-V56 the questions are found in the “Kenya Food Composition Tables 2018”

Data Collection

Data Collection Dates

Start	End	Cycle
2020-11-07	2023-12-31	Supermarket A

Data Collection Mode

Other [oth]

Questionnaires

A standardized form was developed to guide in extration of information from 3rd party information provider for supermarket purchase data. Variables of interest includes supermarket name, supermarket branch, location of supermarket, invoice id, customer id, customer demographics (gender, age), date and time of purchase, product name purchased, unit price per item, number of items purchased, payment method used by customer for purchase etc.

Secondary data collected will not be identifiable as it will be anonymized at the supermarket and client level.

The standardized form is provided as external resources data.
V1-V24 the questions are found in the “Study abstraction tool”
V25-V27 are generated classifications (user developed) and are not in any resource
V28 the questions are found in the “NOVA-Classification-Reference-Sheet”
V29-V56 the questions are found in the “Kenya Food Composition Tables 2018”

Supervision

Not Applicable

Data Processing

Data Editing

Not Applicable

Other Processing

The extracted grocery data was in the form of csv files and was saved into a local database using postgresql version 15.2 and imported into r version 4.3.3 for cleaning and pre-processing.

Data pre-processing techniques applied included: transactions and demographics alignment, dealing with missing values, checking for data consistency, quality assurance checks and filtering non-food items.

After data pre-processing, we applied the NOVA food classification and combined the purchase data with Kenya Food Composition Tables (KFCT). We further developed a classification of nineteen food groups from the food purchases.

Data Appraisal

Estimates of Sampling Error

Not Applicable

File Description

Variable List

clean_analysis_supermarket_a_sample

Content	
Cases	473890
Variable(s)	56
Structure	Type: Keys: ()
Version	
Producer	
Missing Data	

Variables

ID	NAME	LABEL	TYPE	FORMAT	QUESTION
V1	id	Shopper id	discrete	character	Customer ID
V2	county	County acronym	discrete	character	County
V3	gender	Gender	discrete	character	Gender
V4	description	Product name description	discrete	character	Product name
V5	price	Unit price of product	contin	numeric	The unit price per item
V6	quantity	Quantity of product purchased	contin	numeric	Number of items purchased by customer
V7	total	Sales invoice total price	contin	numeric	Total price including tax
V8	trnref	Transaction id	contin	numeric	Transaction id
V9	sdatetime	Transaction date	discrete	character	Date and Time of purchase
V10	paymentmode	Payment mode	discrete	character	Payment method used by customer for purchase
V11	branch	Branch id of supermarket	discrete	character	Supermarket branch ID
V12	transaction_id	Supermarket branch transaction id	discrete	character	Invoice ID
V13	dob_new	Date of birth of shopper	discrete	character	Age
V14	supermarket_name	Supermarket name	discrete	character	Supermarket ID
V15	branch_name	Branch name of supermarket	discrete	character	Supermarket branch ID
V16	county_name	Location	discrete	character	County
V17	sub_county_name	Sub-county	discrete	character	Sub-county
V18	month_date	Month date	discrete	character	Date and Time of purchase
V19	total_new	Total product price	contin	numeric	The unit price per item
V20	customer_type	Shopper type recorded as loyalty and non-loyalty	discrete	character	Type of customer
V21	year	Year	discrete	numeric	Year
V22	quarter_date	Quarter date	discrete	numeric	Quarter date
V23	age	Age (years)	contin	numeric	Age (years)
V24	item_type	Product type recorded as food item and non-food item	discrete	character	Category
V25	class_name	Food category	discrete	character	Food category

V26	subclass_name	Food category sub-groups	discrete	character	Food category sub-groups
V27	food_group	Developed food groups classification	discrete	character	Developed food groups classification
V28	nova	Nova food classification	discrete	character	Nova food classification
V29	energy_k_j	Energy (kj)	contin	numeric	Energy (kj)
V30	energy_kcal	Energy (kcal)	contin	numeric	Energy (kcal)
V31	water_g	Water (g)	contin	numeric	Water (g)
V32	protein_g	Protein (g)	contin	numeric	Protein (g)
V33	fat_g	Fat (g)	contin	numeric	Fat (g)
V34	carbohydrate_available_g	Carbohydrate available (g)	contin	numeric	Carbohydrate available (g)
V35	fibre_g	Fibre (g)	contin	numeric	Fibre (g)
V36	ash_g	Ash (g)	contin	numeric	Ash (g)
V37	calcium_ca_mg	Calcium (mg)	contin	numeric	Calcium (mg)
V38	iron_fe_mg	Iron (mg)	contin	numeric	Iron (mg)
V39	magnesium_mg_mg	Magnesium (mg)	contin	numeric	Magnesium (mg)
V40	phosphorus_p_mg	Phosphorus (mg)	contin	numeric	Phosphorus (mg)
V41	potassium_k_mg	Potassium (mg)	contin	numeric	Potassium (mg)
V42	sodium_na_mg	Sodium (mg)	contin	numeric	Sodium (mg)
V43	zinc_zn_mg	Zinc (mg)	contin	numeric	Zinc (mg)
V44	selenium_se_mcg	Selenium (mcg)	contin	numeric	Selenium (mcg)
V45	vit_a_rae_mcg	Vitamin a-rae (mcg)	contin	numeric	Vitamin a-rae (mcg)
V46	vit_a_re_mcg	Vitamin a-re (mcg)	contin	numeric	Vitamin a-re (mcg)
V47	retinol_mcg	Retinol (mcg)	contin	numeric	Retinol (mcg)
V48	b_carotene_equivalent_mcg	B-carotene equivalent (mcg)	contin	numeric	B-carotene equivalent (mcg)
V49	thiamin_mg	Thiamin (mg)	contin	numeric	Thiamin (mg)
V50	riboflavin_mg	Riboflavin (mg)	contin	numeric	Riboflavin (mg)
V51	niacin_mg	Niacin (mg)	contin	numeric	Niacin (mg)
V52	dietary_folate_eq_mcg	Dietary folate equivalent (mcg)	contin	numeric	Dietary folate equivalent (mcg)
V53	food_folate_mcg	Food folate (mcg)	contin	numeric	Food folate (mcg)
V54	vit_b12_mcg	Vitamin b12 (mcg)	contin	numeric	Vitamin b12 (mcg)
V55	vit_c_mg	Vitamin c (mg)	contin	numeric	Vitamin c (mg)
V56	cholesterol_chole_mg	Cholesterol (mg)	contin	numeric	Cholesterol (mg)

Shopper id (id)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 5

Valid cases: 163317
Invalid: 0

Description

This question seeks to get the anonymized individual customer identification number created with unique identifiers

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Customer ID

Post question

N/A

Interviewer instructions

N/A

County acronym (county)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 3

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the country name where the shopper is in

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

County

Post question

N/A

Interviewer instructions

N/A

Gender (gender)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 6

Valid cases: 163029
Invalid: 0

Description

This question seeks to get the gender of the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Gender

Post question

N/A

Interviewer instructions

N/A

Product name description (description)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 53

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the name of the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Product name

Post question

N/A

Interviewer instructions

N/A

Unit price of product (price)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 4-6875

Valid cases: 473890
Invalid: 0
Minimum: 4
Maximum: 6875
Mean: 120.2
Standard deviation: 148.8

Description

This question seeks to get the unit price of the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

The unit price per item

Post question

N/A

Interviewer instructions

N/A

Quantity of product purchased (quantity)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.003704-480

Valid cases: 473890
Invalid: 0
Minimum: 0
Maximum: 480
Mean: 1.5
Standard deviation: 2.1

Description

This question seeks to get the quantity of the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Number of items purchased by customer

Post question

N/A

Interviewer instructions

N/A

Sales invoice total price (total)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-238547.8

Valid cases: 473890
Invalid: 0
Minimum: 1
Maximum: 238547.8
Mean: 1187.6
Standard deviation: 2067.3

Description

This question seeks to get the total price of the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Total price including tax

Post question

N/A

Interviewer instructions

N/A

Transaction id (trnref)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 42-19606573

Valid cases: 473890
Invalid: 0
Minimum: 42
Maximum: 19606573
Mean: 9492413.9
Standard deviation: 5607792.1

Description

This question seeks to get the transaction id of the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Transaction id

Post question

N/A

Interviewer instructions

N/A

Transaction date (sdatetime)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 473890
Minimum: NaN
Maximum: NaN

Description

This question seeks to get the date and time of purchase for the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Date and Time of purchase

Post question

N/A

Interviewer instructions

N/A

Payment mode (paymentmode)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 18

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the mode of payment for the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Payment method used by customer for purchase

Post question

N/A

Interviewer instructions

N/A

Branch id of supermarket (branch)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 7

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the supermarket branch ID for the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Supermarket branch ID

Post question

N/A

Interviewer instructions

N/A

Supermarket branch transaction id (transaction_id)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 16

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the invoice ID for the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Invoice ID

Post question

N/A

Interviewer instructions

N/A

Date of birth of shopper (dob_new)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 162419
Minimum: NaN
Maximum: NaN

Description

This question seeks to get the age of the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Age

Post question

N/A

Interviewer instructions

N/A

Supermarket name (supermarket_name)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 1

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the supermarket name for the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Supermarket ID

Post question

N/A

Interviewer instructions

N/A

Branch name of supermarket (branch_name)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 3

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the supermarket branch name for the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Supermarket branch ID

Post question

N/A

Interviewer instructions

N/A

Location (county_name)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 9

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the county name where the supermarket is located

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

County

Post question

N/A

Interviewer instructions

N/A

Sub-county (sub_county_name)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 17

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the sub-county name where the supermarket is located

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Sub-county

Post question

N/A

Interviewer instructions

N/A

Month date (month_date)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 473890
Minimum: NaN
Maximum: NaN

Description

This question seeks to get the month and date for the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Date and Time of purchase

Post question

N/A

Interviewer instructions

N/A

Total product price (total_new)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.999945-35760

Valid cases: 473890
Invalid: 0
Minimum: 1
Maximum: 35760
Mean: 152.3
Standard deviation: 223.2

Description

This question seeks to get the product price for the product purchased by the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

The unit price per item

Post question

N/A

Interviewer instructions

N/A

Shopper type recorded as loyalty and non-loyalty (customer_type)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 15

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the type of customer the shopper was

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Type of customer

Post question

N/A

Interviewer instructions

N/A

Year (year)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: numeric
Width: 10
Decimals: 0
Range: 2020-2023

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the year the purchase was done

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Year

Post question

N/A

Interviewer instructions

N/A

Quarter date (quarter_date)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 1-13

Valid cases: 473890
Invalid: 0

Description

This question seeks to get the quarter the purchase was done

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Quarter date

Post question

N/A

Interviewer instructions

N/A

Age (years) (age)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 18-121

Valid cases: 162419
Invalid: 311471
Minimum: 18
Maximum: 121
Mean: 41.8
Standard deviation: 10.9

Description

This question seeks to get the computed age of the shopper

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Age (years)

Post question

N/A

Interviewer instructions

N/A

Product type recorded as food item and non-food item (item_type)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 9

Valid cases: 473890
Invalid: 0

Description

This question seeks to know if the item purchased was a food or non food item

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Category

Post question

N/A

Interviewer instructions

N/A

Food category (class_name)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 44

Valid cases: 473890
Invalid: 0

Description

This question seeks to know the Food category of the item purchased

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Food category

Post question

N/A

Interviewer instructions

N/A

Food category sub-groups (subclass_name)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 48

Valid cases: 473890
Invalid: 0

Description

This question seeks to know the Food sub-groups of the item purchased

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Food category sub-groups

Post question

N/A

Interviewer instructions

N/A

Developed food groups classification (food_group)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 41

Valid cases: 473890
Invalid: 0

Description

This question seeks to know the Developed food groups classification of the item purchased

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Developed food groups classification

Post question

N/A

Interviewer instructions

N/A

Nova food classification (nova)
File: clean_analysis_supermarket_a_sample

Overview

Type: Discrete
Format: character
Width: 37

Valid cases: 473890
Invalid: 0

Description

This question seeks to know the Nova food classification of the item purchased

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Nova food classification

Post question

N/A

Interviewer instructions

N/A

Energy (kj) (energy_k_j)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-3700

Valid cases: 448634
Invalid: 25256
Minimum: 1
Maximum: 3700
Mean: 963.6
Standard deviation: 794.7

Description

This question seeks to know the amount of kilojoule contained in the purchased product per 100grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Energy (kj)

Post question

N/A

Interviewer instructions

N/A

Energy (kcal) (energy_kcal)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 2-3082

Valid cases: 448634
Invalid: 25256
Minimum: 2
Maximum: 3082
Mean: 492.1
Standard deviation: 540.3

Description

This question seeks to know the amount of calories contained in the purchased product per 100grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Energy (kcal)

Post question

N/A

Interviewer instructions

N/A

Water (g) (water_g)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.2-99.95

Valid cases: 428188
Invalid: 45702
Minimum: 0.2
Maximum: 100
Mean: 43.1
Standard deviation: 33.8

Description

This question seeks to know the amount of water contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Water (g)

Post question

N/A

Interviewer instructions

N/A

Protein (g) (protein_g)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.1-84.4

Valid cases: 421029
Invalid: 52861
Minimum: 0.1
Maximum: 84.4
Mean: 6.4
Standard deviation: 5.5

Description

This question seeks to know the amount of protains contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Protein (g)

Post question

N/A

Interviewer instructions

N/A

Fat (g) (fat_g)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.1-100

Valid cases: 364876
Invalid: 109014
Minimum: 0.1
Maximum: 100
Mean: 12.7
Standard deviation: 20.9

Description

This question seeks to know the amount of fats contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Fat (g)

Post question

N/A

Interviewer instructions

N/A

Carbohydrate available (g) (carbohydrate_available_g)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.1-101.3

Valid cases: 426403
Invalid: 47487
Minimum: 0.1
Maximum: 101.3
Mean: 42.5
Standard deviation: 28.6

Description

This question seeks to know the amount of Carbohydrates contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Carbohydrate available (g)

Post question

N/A

Interviewer instructions

N/A

Fibre (g) (fibre_g)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.1-52.3

Valid cases: 236004
Invalid: 237886
Minimum: 0.1
Maximum: 52.3
Mean: 4.8
Standard deviation: 4.7

Description

This question seeks to know the amount of fibre contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Fibre (g)

Post question

N/A

Interviewer instructions

N/A

Ash (g) (ash_g)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.05-99.8

Valid cases: 314217
Invalid: 159673
Minimum: 0.1
Maximum: 99.8
Mean: 4.2
Standard deviation: 14.4

Description

This question seeks to know the amount of ash contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Ash (g)

Post question

N/A

Interviewer instructions

N/A

Calcium (mg) (calcium_ca_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-4280

Valid cases: 452724
Invalid: 21166
Minimum: 1
Maximum: 4280
Mean: 75.1
Standard deviation: 180.6

Description

This question seeks to know the amount of calcium contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Calcium (mg)

Post question

N/A

Interviewer instructions

N/A

Iron (mg) (iron_fe_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.01-123.6

Valid cases: 387751
Invalid: 86139
Minimum: 0
Maximum: 123.6
Mean: 2
Standard deviation: 3.9

Description

This question seeks to know the amount of iron contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Iron (mg)

Post question

N/A

Interviewer instructions

N/A

Magnesium (mg) (magnesium_mg_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-420

Valid cases: 441432
Invalid: 32458
Minimum: 1
Maximum: 420
Mean: 39.1
Standard deviation: 60.4

Description

This question seeks to know the amount of magnesium contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Magnesium (mg)

Post question

N/A

Interviewer instructions

N/A

Phosphorus (mg) (phosphorus_p_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-8410

Valid cases: 427822
Invalid: 46068
Minimum: 1
Maximum: 8410
Mean: 152.5
Standard deviation: 330.6

Description

This question seeks to know the amount of Phosphorus contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Phosphorus (mg)

Post question

N/A

Interviewer instructions

N/A

Potassium (mg) (potassium_k_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-10200

Valid cases: 453645
Invalid: 20245
Minimum: 1
Maximum: 10200
Mean: 368.5
Standard deviation: 995.8

Description

This question seeks to know the amount of Potassium contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Potassium (mg)

Post question

N/A

Interviewer instructions

N/A

Sodium (mg) (sodium_na_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-38500

Valid cases: 454910
Invalid: 18980
Minimum: 1
Maximum: 38500
Mean: 817.6
Standard deviation: 4673.9

Description

This question seeks to know the amount of Sodium contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Sodium (mg)

Post question

N/A

Interviewer instructions

N/A

Zinc (mg) (zinc_zn_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.01-16

Valid cases: 388377
Invalid: 85513
Minimum: 0
Maximum: 16
Mean: 0.9
Standard deviation: 0.8

Description

This question seeks to know the amount of Zinc contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Zinc (mg)

Post question

N/A

Interviewer instructions

N/A

Selenium (mcg) (selenium_se_mcg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.2-254

Valid cases: 340319
Invalid: 133571
Minimum: 0.2
Maximum: 254
Mean: 7.8
Standard deviation: 11.9

Description

This question seeks to know the amount of Selenium contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Selenium (mcg)

Post question

N/A

Interviewer instructions

N/A

Vitamin a-rae (mcg) (vit_a_rae_mcg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-19534

Valid cases: 172369
Invalid: 301521
Minimum: 1
Maximum: 19534
Mean: 109.2
Standard deviation: 533.8

Description

This question seeks to know the amount of Vitamin a-rae contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Vitamin a-rae (mcg)

Post question

N/A

Interviewer instructions

N/A

Vitamin a-re (mcg) (vit_a_re_mcg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-19601

Valid cases: 221860
Invalid: 252030
Minimum: 1
Maximum: 19601
Mean: 106
Standard deviation: 511

Description

This question seeks to know the amount of Vitamin a-re contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Vitamin a-re (mcg)

Post question

N/A

Interviewer instructions

N/A

Retinol (mcg) (retinol_mcg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-19466

Valid cases: 191149
Invalid: 282741
Minimum: 1
Maximum: 19466
Mean: 97.7
Standard deviation: 498.7

Description

This question seeks to know the amount of Retinol contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Retinol (mcg)

Post question

N/A

Interviewer instructions

N/A

B-carotene equivalent (mcg) (b_carotene_equivalent_mcg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-65800

Valid cases: 162317
Invalid: 311573
Minimum: 1
Maximum: 65800
Mean: 175.6
Standard deviation: 1410.2

Description

This question seeks to know the amount of B-carotene equivalent contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

B-carotene equivalent (mcg)

Post question

N/A

Interviewer instructions

N/A

Thiamin (mg) (thiamin_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.01-3.78

Valid cases: 270612
Invalid: 203278
Minimum: 0
Maximum: 3.8
Mean: 0.2
Standard deviation: 0.3

Description

This question seeks to know the amount of Thiamin contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Thiamin (mg)

Post question

N/A

Interviewer instructions

N/A

Riboflavin (mg) (riboflavin_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.01-272

Valid cases: 348010
Invalid: 125880
Minimum: 0
Maximum: 272
Mean: 0.5
Standard deviation: 8.5

Description

This question seeks to know the amount of Riboflavin contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Riboflavin (mg)

Post question

N/A

Interviewer instructions

N/A

Niacin (mg) (niacin_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.1-42.9

Valid cases: 317485
Invalid: 156405
Minimum: 0.1
Maximum: 42.9
Mean: 3
Standard deviation: 5.3

Description

This question seeks to know the amount of Niacin contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Niacin (mg)

Post question

N/A

Interviewer instructions

N/A

Dietary folate equivalent (mcg) (dietary_folate_eq_mcg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-4000

Valid cases: 349921
Invalid: 123969
Minimum: 1
Maximum: 4000
Mean: 55.8
Standard deviation: 149.7

Description

This question seeks to know the amount of Dietary folate equivalent contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Dietary folate equivalent (mcg)

Post question

N/A

Interviewer instructions

N/A

Food folate (mcg) (food_folate_mcg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-4000

Valid cases: 348455
Invalid: 125435
Minimum: 1
Maximum: 4000
Mean: 28.5
Standard deviation: 99.2

Description

This question seeks to know the amount of Food folate contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Food folate (mcg)

Post question

N/A

Interviewer instructions

N/A

Vitamin b12 (mcg) (vit_b12_mcg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.05-112.84

Valid cases: 240494
Invalid: 233396
Minimum: 0.1
Maximum: 112.8
Mean: 0.6
Standard deviation: 2.9

Description

This question seeks to know the amount of Vitamin b12 contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Vitamin b12 (mcg)

Post question

N/A

Interviewer instructions

N/A

Vitamin c (mg) (vit_c_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.15-156

Valid cases: 89383
Invalid: 384507
Minimum: 0.2
Maximum: 156
Mean: 6.3
Standard deviation: 13.4

Description

This question seeks to know the amount of Vitamin c contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Vitamin c (mg)

Post question

N/A

Interviewer instructions

N/A

Cholesterol (mg) (cholesterol_chole_mg)
File: clean_analysis_supermarket_a_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.2-977

Valid cases: 121296
Invalid: 352594
Minimum: 0.2
Maximum: 977
Mean: 32.6
Standard deviation: 63.7

Description

This question seeks to know the amount of Cholesterol contained in the purchased product per 100 grams/ml

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Cholesterol (mg)

Post question

N/A

Interviewer instructions

N/A

Documentation

Questionnaires

Study abstraction tool.pdf

Title	Study abstraction tool.pdf
Author(s)	African Population and Health Research Center
Date	14/01/2025
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	Study abstraction tool.pdf

Technical documents

momanyi-et-al-2025-analyzing-demographic-grocery-purchase-patterns.pdf

Title	momanyi-et-al-2025-analyzing-demographic-grocery-purchase-patterns.pdf
Author(s)	African Population and Health Research Center
Date	14/01/2025
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	momanyi-et-al-2025-analyzing-demographic-grocery-purchase-patterns.pdf

Other materials

Metadata of supermarkets.pdf

Title	Metadata of supermarkets.pdf
Author(s)	African Population and Health Research Center
Date	14/01/2015
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	Metadata of supermarkets.pdf

NOVA-Classification-Reference-Sheet.pdf

Title	NOVA-Classification-Reference-Sheet.pdf
Author(s)	African Population and Health Research Center

Date	14/01/2025
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	NOVA-Classification-Reference-Sheet.pdf

FAO Nova.pdf

Title	FAO Nova.pdf
Author(s)	African Population and Health Research Center
Date	14/01/2025
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	FAO Nova.pdf