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

Agnes Kiragga

Report generated on: January 16, 2026

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

Overview

Identification

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

Version

VERSION DESCRIPTION

PRODUCTION DATE
2025-07-14

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.

A further objective was to examine changes in food purchasing and nutritional composition before, during and after the COVID-19 pandemic restrictions.

UNITS OF ANALYSIS
Supermarket transaction records.

Scope

NOTES
- Transaction Level Data: Food Item Details (Specific products purchased), Quantity, Price, Date of Purchase, Location of Purchase, Payment Method(Cash, credit card, digital payment, etc.), Basket Composition

KEYWORDS
Supermarkets, Ultra-Processed Foods, Processed Foods, Non-communicable Chronic Diseases, COVID-19, SARS-CoV-2, Micronutrient

Coverage

GEOGRAPHIC COVERAGE
County coverage: Nairobi

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, analysis and documentation of study
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 Centre (APHRC)	Data Governance Officer
Mambe Shem	African Population and Health Research Centre (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
2025-07-14

DDI DOCUMENT VERSION
Version 1.0 (July 2025)

DDI DOCUMENT ID
DDI-KEN-APHRC-SUPERMARKET-CD-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 2 supermarket chains each with 1 branch.

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.

The standardized form is provided as external resources data.
V1-V27 the questions are found in the “Study abstraction tool”
V28-V30 are generated food classifications (user developed) and are not in any resource
V31 the questions are found in the “NOVA-Classification-Reference-Sheet”
V32-V59 the questions are found in the “Kenya Food Composition Tables 2018”

Data Collection

Data Collection Dates

Start	End	Cycle
2018-01-12	2023-12-31	Supermarket C
2018-11-03	2023-12-31	Supermarket D

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.

The standardized form is provided as external resources data.
V1-V27 the questions are found in the “Study abstraction tool”
V28-V30 are generated food classifications (user developed) and are not in any resource
V31 the questions are found in the “NOVA-Classification-Reference-Sheet”
V32-V59 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. Additionally, a COVID period variable was created and coded "Pre-covid", "Covid" and "Post-covid" based on Kenya government official dates for start and end of pandemic curfew on 27th March 2020 and 20th October 2021 respectively.

Data Appraisal

Estimates of Sampling Error

Not Applicable

File Description

Variable List

clean_analysis_supermarket_c_d_sample

Content	
Cases	444238
Variable(s)	59
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 description
V5	price	Unit price of product	contin	numeric	Unit price of product
V6	quantity	Quantity of product purchased	contin	numeric	Quantity of product purchased
V7	total	Sales invoice total price	contin	numeric	Sales invoice total price
V8	trnref	Transaction id	contin	numeric	Transaction id
V9	sdatetime	Transaction date	discrete	character	Transaction date
V10	paymentmode	Payment mode	discrete	character	Payment mode
V11	branch	Branch id of supermarket	discrete	character	Branch id of supermarket
V12	transaction_id	Supermarket branch transaction id	discrete	character	Supermarket branch transaction id
V13	dob_new	Date of birth of shopper	discrete	character	Date of birth of shopper
V14	supermarket_name	Supermarket name	discrete	character	Supermarket name
V15	branch_name	Branch name of supermarket	discrete	character	Branch name of supermarket
V16	county_name	Location	discrete	character	Location
V17	sub_county_name	Sub-county	discrete	character	Sub-county
V18	month_date	Month date	discrete	character	Month date
V19	week_date	Week date	discrete	character	Week date
V20	week_number	Week number from 1st january	contin	numeric	Week number from 1st january
V21	day_year	Day of the year	contin	numeric	Day of the year
V22	total_new	Total product price	contin	numeric	Total product price
V23	customer_type	Shopper type recorded as loyalty and non-loyalty	discrete	character	Shopper type recorded as loyalty and non-loyalty
V24	year	Year	discrete	numeric	Year
V25	quarter_date	Quarter of year	discrete	numeric	Quarter of year
V26	covid_period	Covid period groups	discrete	numeric	Covid period groups

V27	item_type	Product type recorded as food item and non-food item	discrete	character	Product type recorded as food item and non-food item
V28	class_name	Food category	discrete	character	Food category
V29	subclass_name	Food category sub-groups	discrete	character	Food category sub-groups
V30	food_group	Developed food groups classification	discrete	character	Developed food groups classification
V31	nova	Nova food classification	discrete	character	Nova food classification
V32	energy_k_j	Energy (kJ)	contin	numeric	Energy (kJ)
V33	energy_kcal	Energy (kcal)	contin	numeric	Energy (kcal)
V34	water_g	Water (g)	contin	numeric	Water (g)
V35	protein_g	Protein (g)	contin	numeric	Protein (g)
V36	fat_g	Fat (g)	contin	numeric	Fat (g)
V37	carbohydrate_available_g	Carbohydrate available (g)	contin	numeric	Carbohydrate available (g)
V38	fibre_g	Fibre (g)	contin	numeric	Fibre (g)
V39	ash_g	Ash (g)	contin	numeric	Ash (g)
V40	calcium_ca_mg	Calcium (mg)	contin	numeric	Calcium (mg)
V41	iron_fe_mg	Iron (mg)	contin	numeric	Iron (mg)
V42	magnesium_mg_mg	Magnesium (mg)	contin	numeric	Magnesium (mg)
V43	phosphorus_p_mg	Phosphorus (mg)	contin	numeric	Phosphorus (mg)
V44	potassium_k_mg	Potassium (mg)	contin	numeric	Potassium (mg)
V45	sodium_na_mg	Sodium (mg)	contin	numeric	Sodium (mg)
V46	zinc_zn_mg	Zinc (mg)	contin	numeric	Zinc (mg)
V47	selenium_se_mcg	Selenium (mcg)	contin	numeric	Selenium (mcg)
V48	vit_a_rae_mcg	Vitamin A-RAE (mcg)	contin	numeric	Vitamin A-RAE (mcg)
V49	vit_a_re_mcg	Vitamin A-RE (mcg)	contin	numeric	Vitamin A-RE (mcg)
V50	retinol_mcg	Retinol (mcg)	contin	numeric	Retinol (mcg)
V51	b_carotene_equivalent_mcg	b-carotene equivalent (mcg)	contin	numeric	b-carotene equivalent (mcg)
V52	thiamin_mg	Thiamin (mg)	contin	numeric	Thiamin (mg)
V53	riboflavin_mg	Riboflavin (mg)	contin	numeric	Riboflavin (mg)
V54	niacin_mg	Niacin (mg)	contin	numeric	Niacin (mg)
V55	dietary_folate_eq_mcg	Dietary Folate Equivalent (mcg)	contin	numeric	Dietary Folate Equivalent (mcg)
V56	food_folate_mcg	Food folate (mcg)	contin	numeric	Food folate (mcg)
V57	vit_b12_mcg	Vitamin B12 (mcg)	contin	numeric	Vitamin B12 (mcg)
V58	vit_c_mg	Vitamin C (mg)	contin	numeric	Vitamin C (mg)
V59	cholesterol_chole_mg	Cholesterol (mg)	contin	numeric	Cholesterol (mg)

Shopper id (id)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 1

Valid cases: 0
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_c_d_sample

Overview

Type: Discrete
Format: character
Width: 7

Valid cases: 444238
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_c_d_sample

Overview

Type: Discrete
Format: character
Width: 1

Valid cases: 0
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_c_d_sample

Overview

Type: Discrete
Format: character
Width: 50

Valid cases: 444238
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 description

Post question

N/A

Interviewer instructions

N/A

Unit price of product (price)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 4-7000

Valid cases: 444238
Invalid: 0
Minimum: 4
Maximum: 7000
Mean: 129.1
Standard deviation: 173.7

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

Unit price of product

Post question

N/A

Interviewer instructions

N/A

Quantity of product purchased (quantity)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.06-1600

Valid cases: 444238
Invalid: 0
Minimum: 0.1
Maximum: 1600
Mean: 1.8
Standard deviation: 3.9

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

Quantity of product purchased

Post question

N/A

Interviewer instructions

N/A

Sales invoice total price (total)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 4-169655

Valid cases: 444238
Invalid: 0
Minimum: 4
Maximum: 169655
Mean: 898.6
Standard deviation: 1637.8

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

Sales invoice total price

Post question

N/A

Interviewer instructions

N/A

Transaction id (trnref)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 12
Decimals: 0
Range: 100085-6360092

Valid cases: 444238
Invalid: 0
Minimum: 100085
Maximum: 6360092
Mean: 3121227.7
Standard deviation: 1751452.5

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_c_d_sample

Overview

Type: Discrete
Format: character
Width: 11

Valid cases: 444238
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

Transaction date

Post question

N/A

Interviewer instructions

N/A

Payment mode (paymentmode)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 1

Valid cases: 0
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 mode

Post question

N/A

Interviewer instructions

N/A

Branch id of supermarket (branch)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 7

Valid cases: 444238
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

Branch id of supermarket

Post question

N/A

Interviewer instructions

N/A

Supermarket branch transaction id (transaction_id)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 15

Valid cases: 444238
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

Supermarket branch transaction id

Post question

N/A

Interviewer instructions

N/A

Date of birth of shopper (dob_new)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 1

Valid cases: 0
Invalid: 0

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

Date of birth of shopper

Post question

N/A

Interviewer instructions

N/A

Supermarket name (supermarket_name)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 1

Valid cases: 444238
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 name

Post question

N/A

Interviewer instructions

N/A

Branch name of supermarket (branch_name)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 2

Valid cases: 444238
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

Branch name of supermarket

Post question

N/A

Interviewer instructions

N/A

Location (county_name)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 7

Valid cases: 444238
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

Location

Post question

N/A

Interviewer instructions

N/A

Sub-county (sub_county_name)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 9
Valid cases: 444238
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_c_d_sample

Overview

Type: Discrete
Format: character
Width: 11
Valid cases: 444238
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

Month date

Post question

N/A

Interviewer instructions

N/A

Week date (week_date)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 11
Valid cases: 444238
Minimum: NaN
Maximum: NaN

Description

This question seeks to get the week 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

Week date

Post question

N/A

Interviewer instructions

N/A

Week number from 1st january (week_number)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-53

Valid cases: 444238
Invalid: 0
Minimum: 1
Maximum: 53
Mean: 27.2
Standard deviation: 15.1

Description

This question seeks to get the week number in the calender year 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

Week number from 1st january

Post question

N/A

Interviewer instructions

N/A

Day of the year (day_year)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-366

Valid cases: 444238
Invalid: 0
Minimum: 1
Maximum: 366
Mean: 187.4
Standard deviation: 105.4

Description

This question seeks to get the day in the calender year 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

Day of the year

Post question

N/A

Interviewer instructions

N/A

Total product price (total_new)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 4-152000

Valid cases: 444238
Invalid: 0
Minimum: 4
Maximum: 152000
Mean: 177.1
Standard deviation: 352.2

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

Total product price

Post question

N/A

Interviewer instructions

N/A

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

Overview

Type: Discrete
Format: character
Width: 15

Valid cases: 444238
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

Shopper type recorded as loyalty and non-loyalty

Post question

N/A

Interviewer instructions

N/A

Year (year)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: numeric
Width: 10
Decimals: 0
Range: 2018-2023

Valid cases: 444238
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 of year (quarter_date)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 1-4

Valid cases: 444238
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 of year

Post question

N/A

Interviewer instructions

N/A

Covid period groups (covid_period)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: numeric
Width: 12
Decimals: 0
Range: 1-3

Valid cases: 444238
Invalid: 0

Description

This question seeks to get the Covid period group

Universe

Individuals and supermarket transaction records.

Source of information

Individuals and supermarket transaction records.

Pre question

N/A

Literal question

Covid period groups

Post question

N/A

Interviewer instructions

N/A

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

Overview

Type: Discrete
Format: character
Width: 9

Valid cases: 444238
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

Product type recorded as food item and non-food item

Post question

N/A

Interviewer instructions

N/A

Food category (class_name)
File: clean_analysis_supermarket_c_d_sample

Overview

Type: Discrete
Format: character
Width: 44

Valid cases: 444238
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_c_d_sample

Overview

Type: Discrete
Format: character
Width: 48

Valid cases: 444238
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_c_d_sample

Overview

Type: Discrete
Format: character
Width: 41

Valid cases: 444238
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_c_d_sample

Overview

Type: Discrete
Format: character
Width: 37

Valid cases: 444238
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_c_d_sample

Overview

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

Valid cases: 415194
Invalid: 29044
Minimum: 1
Maximum: 3700
Mean: 904.3
Standard deviation: 624.4

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 2-3692

Valid cases: 415194
Invalid: 29044
Minimum: 2
Maximum: 3692
Mean: 624.3
Standard deviation: 662.6

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.1-99.95

Valid cases: 427049
Invalid: 17189
Minimum: 0.1
Maximum: 100
Mean: 37.6
Standard deviation: 31.9

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_c_d_sample

Overview

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

Valid cases: 401763
Invalid: 42475
Minimum: 0.1
Maximum: 84.4
Mean: 6.1
Standard deviation: 4.3

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_c_d_sample

Overview

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

Valid cases: 365836
Invalid: 78402
Minimum: 0.1
Maximum: 100
Mean: 12.1
Standard deviation: 13.1

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_c_d_sample

Overview

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

Valid cases: 408623
Invalid: 35615
Minimum: 0.1
Maximum: 101.3
Mean: 44.5
Standard deviation: 24.8

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_c_d_sample

Overview

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

Valid cases: 232996
Invalid: 211242
Minimum: 0.1
Maximum: 52.3
Mean: 4
Standard deviation: 5.2

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_c_d_sample

Overview

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

Valid cases: 279849
Invalid: 164389
Minimum: 0.1
Maximum: 99.8
Mean: 1.9
Standard deviation: 4.8

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_c_d_sample

Overview

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

Valid cases: 424282
Invalid: 19956
Minimum: 1
Maximum: 4280
Mean: 96.2
Standard deviation: 186.4

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_c_d_sample

Overview

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

Valid cases: 382000
Invalid: 62238
Minimum: 0
Maximum: 123.6
Mean: 2.1
Standard deviation: 4.4

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-430

Valid cases: 419981
Invalid: 24257
Minimum: 1
Maximum: 430
Mean: 34.6
Standard deviation: 52

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_c_d_sample

Overview

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

Valid cases: 395315
Invalid: 48923
Minimum: 1
Maximum: 8410
Mean: 140.3
Standard deviation: 326

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_c_d_sample

Overview

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

Valid cases: 423833
Invalid: 20405
Minimum: 1
Maximum: 10200
Mean: 299.8
Standard deviation: 620.7

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_c_d_sample

Overview

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

Valid cases: 413752
Invalid: 30486
Minimum: 1
Maximum: 38500
Mean: 285.6
Standard deviation: 1687.1

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.01-8

Valid cases: 359474
Invalid: 84764
Minimum: 0
Maximum: 8
Mean: 0.8
Standard deviation: 0.7

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_c_d_sample

Overview

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

Valid cases: 357164
Invalid: 87074
Minimum: 0.2
Maximum: 254
Mean: 6.5
Standard deviation: 7.9

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-5490

Valid cases: 134303
Invalid: 309935
Minimum: 1
Maximum: 5490
Mean: 95.7
Standard deviation: 364.8

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-11000

Valid cases: 204302
Invalid: 239936
Minimum: 1
Maximum: 11000
Mean: 125.1
Standard deviation: 657.2

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 1-5930

Valid cases: 170443
Invalid: 273795
Minimum: 1
Maximum: 5930
Mean: 72.8
Standard deviation: 108.1

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_c_d_sample

Overview

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

Valid cases: 266730
Invalid: 177508
Minimum: 1
Maximum: 65800
Mean: 293.5
Standard deviation: 3423.7

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_c_d_sample

Overview

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

Valid cases: 306182
Invalid: 138056
Minimum: 0
Maximum: 3.8
Mean: 0.2
Standard deviation: 0.2

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_c_d_sample

Overview

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

Valid cases: 352770
Invalid: 91468
Minimum: 0
Maximum: 272
Mean: 0.3
Standard deviation: 4.8

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.05-42.9

Valid cases: 332230
Invalid: 112008
Minimum: 0.1
Maximum: 42.9
Mean: 2.3
Standard deviation: 4.7

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.4-4000

Valid cases: 336652
Invalid: 107586
Minimum: 0.4
Maximum: 4000
Mean: 23.7
Standard deviation: 99

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.4-4000

Valid cases: 335450
Invalid: 108788
Minimum: 0.4
Maximum: 4000
Mean: 19.7
Standard deviation: 89.5

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.05-12

Valid cases: 157101
Invalid: 287137
Minimum: 0.1
Maximum: 12
Mean: 0.5
Standard deviation: 0.7

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.1-295

Valid cases: 80732
Invalid: 363506
Minimum: 0.1
Maximum: 295
Mean: 7
Standard deviation: 12.6

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_c_d_sample

Overview

Type: Continuous
Format: numeric
Width: 10
Decimals: 0
Range: 0.1-418

Valid cases: 247734
Invalid: 196504
Minimum: 0.1
Maximum: 418
Mean: 22.7
Standard deviation: 50.6

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	16/01/2026
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	Study abstraction tool.pdf

Other materials

Metadata of supermarkets.pdf

Title	Metadata of supermarkets.pdf
Author(s)	African Population and Health Research Center
Date	16/01/2026
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	Metadata of supermarkets.pdf

Kenya Food Compostion Tables 2018.pdf

Title	Kenya Food Compostion Tables 2018.pdf
Author(s)	African Population and Health Research Center
Date	16/01/2026
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	Kenya Food Compostion Tables 2018.pdf

NOVA-Classification-Reference-Sheet.pdf

Title	NOVA-Classification-Reference-Sheet.pdf
Author(s)	African Population and Health Research Center
Date	16/01/2026
Country	Kenya, Ethiopia, Nigeria, Mozambique
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	16/01/2026
Country	Kenya
Language	ENGLISH
Contributor(s)	Dr. Agnes Kiragga
Publisher(s)	African Population and Health Research Center (APHRC)
Filename	FAO Nova.pdf