HUNGRY CITIES PARTNERSHIP

Discussion Paper No. 60

DECEMBER 2023

OPENING ACCESS TO URBAN FOOD SECURITY DATA IN AFRICA FROM THE HUNGRY CITIES PARTNERSHIP

GARETH HAYSOM,[1] BRUCE MCDOUGALL,[2] SAMUEL OWUOR,[3] NDEYAPO NICKANOR,[4] INES RAIMUNDO,[5] LAWRENCE KAZEMBE,[6] LYNN WOOLFREY[7] AND JONATHAN CRUSH[8]

SERIES EDITOR: JONATHAN CRUSH

[1] African Centre for Cities, University of Cape Town, South Africa, garethhaysom@gmail.com
[2]
[3] University of Nairobi, Kenya, samowuor@uonbi.ac.ke
[4] University of Namibia, Namibia, nnickanor@unam.edu
[5] Eduardo Mondlane University, Mozambique, inesmacamo@gmail.com
[6] University of Namibia, Namibia, lKazembe@unam.na
[7]
[8] Balsillie School of International Affairs, Canada; University of the Western Cape, South Africa, jcrush@balsillieschool.ca

Abstract

The Hungry Cities Partnership (HCP) initiated a 7-year programme of food system research in eight cities of the Global South in January 2015: Mexico City, Mexico; Kingston, Jamaica; Windhoek, Namibia; Cape Town, South Africa; Maputo, Mozambique; Nairobi, Kenya; Bangalore, India; and Nanjing, China. This paper describes the research process in the four African cities of Cape Town, Maputo, Nairobi, and Windhoek. The HCP data provide rich and context-specific information useful in avoiding generalizations in theory and inserting a Southern and African perspective into wider food studies and global policy debates. The paper describes and compares the survey methodology in the four cities and then, as a legacy for other researchers, the process of preparing the data sets for open access at an institutional repository, Datafirst.

Keywords

food studies, urban food security, food systems

Suggested Citation

Haysom, G., McDougall, B., Owuor, S., Nickanor, N., Raimundo, I., Kazembe, L., Woolfrey, L. and Crush, J. (2023). *Opening Access to Urban Food Security Data in Africa from the Hungry Cities Partnership*. HCP Discussion Paper No. 60, Waterloo and Cape Town.

This is the 60th discussion paper in a series published by the Hungry Cities Partnership (HCP), an international research project examining food security and inclusive growth in cities in the Global South. The multi-year collaborative project aims to understand how cities in the Global South will manage the food security challenges arising from rapid urbanization and the transformation of urban food systems. The Partnership is funded by the Social Sciences and Humanities Research Council of Canada (SSHRC).

Introduction

The United Nations Department of Economic and Social Affairs (UN-DESA) estimates that over the next 30 years, virtually all the anticipated three billion increase in the human population will occur in cities of the South, with the largest growth taking place in Africa. Southern cities will absorb 95% of urban growth in the next two decades and, by 2030, will be home to almost 4 billion people, or 80% of the world's urban population (UN-DESA, 2018). Acuto and Parnell (2016) have argued that understanding the implications of rapid urbanization is hampered by the general absence of current data on African urban areas. This is particularly evident when it comes to the urban food system and food security of new urbanites where there is a general paucity of food security-specific data on the urban scale (Crush, 2016).

Rural bias has meant that food security issues have generally been ignored in research on the effects of turbocharged urbanization in Southern and East Africa. In the 1980s, Lipton (1984) argued that donor and national funding decisions on African food systems and food security displayed considerable urban bias. In the last two decades, this tendency has been reversed and replaced by rural bias which now dominates all funding decisions (Crush and Riley, 2017). As a result, food insecurity and the (mal)functioning of the food system in these regions is poorly understood. The absence of food system data from urban areas has profound implications for development planning, particularly as "measurement drives diagnosis and response" (Haysom and Tawodzera, 2018).

Against this backdrop, the Hungry Cities Partnership (HCP) initiated a 7-year programme of food system research in eight cities of the Global South in January 2015 including. Mexico City, Mexico; Kingston, Jamaica; Windhoek, Namibia; Cape Town, South Africa; Maputo, Mozambique; Nairobi, Kenya; Bangalore, India; and Nanjing, China. The HCP uses food as a lens to better understand the issues of urban development, governance, informality, inclusive growth, and food security. The Partnership sought to engage these intersecting challenges through the following areas of inquiry:

- Rapid Urbanization, Food Insecurity, and Inclusive Growth
- Reshaping Informal Food Systems through Inclusive Growth
- Youth Entrepreneurship in the Informal Food Economy
- Competition and Inclusive Growth in the Urban Food Economy

This paper describes the research process in the four African cities of Cape Town, Maputo, Nairobi, and Windhoek. The focus on African cities is motivated by several considerations. First, at the global governance scale, the Sustainable Development Goals (SDGs) and the aspirations of the New Urban Agenda (NUA) have brought cities back into development discourse (NUA, 2016). The SDGs form part of a global reporting process, which generates data at the national level. Urban food security is seldom considered in national development agendas and resourcing and the focus is often shifted away from African cities. Second, urban food policies require data at the city scale, and useful data in most African cities are scarce. Third, in Southern and East Africa, there is a general trend for national governments to report on SDG 2 – Zero Hunger (the so-called Hunger Goal) using data from small, mostly rural samples. Relevant, scale-appropriate urban data provide a necessary foundation for urbanization of the SDGs and for NUA's national urban plans (Pieterse et al., 2018). The HCP data provide rich and context-specific information useful in avoiding generalizations in theory and inserting a Southern and African perspective into wider food studies and global policy debates.

Methodology

The HCP had the following research objectives in its mission to address the urban food security data challenges:

- Examine the levels and determinants of the four dimensions of food security (availability, access, utilization, and stability) in a range of large cities across the global South and the relationship between food insecurity, income generation, and unemployment;
- Generate a significant body of comparative, interdisciplinary knowledge on the organization, structure, and potential for inclusive growth in the informal food economy and implications for food security;
- Understand the implications of formal food sector competition for entrepreneurship, innovation, job creation, and inclusive growth in the urban informal food economy;
- Critically assess opportunities for women and youth to be incorporated into urban food systems as entrepreneurs and employees with decent jobs;
- Evaluate national, regional, and municipal policies that enable or constrain enterprise development, entrepreneurship, and innovation in the informal economy.

To meet these objectives, the HCP adopted a mixed-methods approach using quantitative surveys and in-depth qualitative interviews. In this paper, we focus on the former, which involved two city-scale surveys in each of the eight partner cities: (a) Household Food Security Survey; and (b) Informal Food Vendor Survey.

Survey Instruments

Household Food Security Survey

The survey was administered to more than 6,800 households between 2016 and 2018 in the four African cities, representing an estimated 10 million urban residents in Southern and East Africa. The Cape Town household survey was conducted by Citizen Surveys for the African Food Security Urban Network (AFSUN). The data was made available to HCP for analysis and report writing. The Household Food Security Survey instrument was a modified version of a questionnaire originally developed by AFSUN for a survey of 11 Southern African cities (Frayne et al., 2018) (see HCP Open Data Site). The survey used the household as the primary unit of investigation and sought to understand the extent of household food security across all dimensions including availability, accessibility, utilization, stability, and safety. Household level data was collected on demographics and economics, household food purchasing dynamics, and the frequencies of patronage of different food retail options, socio-political processes at the household scale, consumption patterns, who makes the food choices in the household and nutritional discounting taking place in households, and detail on the foods purchased and consumed. Furthermore, attitudinal questions on issues such as supermarket perceptions and attitudes to urban agriculture were included in the survey.

The household survey used a validated cross-cultural set of metrics to determine levels of food insecurity. This allows comparisons with similar surveys in other regions and between urban and rural food security in the same region. The linking of food security measures means that urban food security outcomes across income categories can be captured and comparisons made between higher, middle, and lower income areas in cities. These metrics included (a) the Household Food Insecurity Access Scale (HFIAS), (b) the Household Food Insecurity Access Prevalence Scale (HFIAP); (c) the Household Dietary Diversity Score (HDDS); and (d) the Months of Adequate Food Provisioning (MAHPF) indicator (Bilinsky and Swindale, 2007;

Coates et al., 2007; Swindale and Bilinsky, 2006). The utility of these and other tools for measuring urban food security has been extensively discussed (Ballard et al., 2013; Coates, 2013; Jones et al., 2013; Haysom and Tawodzera, 2018). To assess the relationship between poverty and food insecurity, the Lived Poverty Index (LPI) of the Afrobarometer was used (Mattes, 2008; Mattes et al., 2016).

Informal Food Vendor Survey

The HCP informal vendor survey instrument was developed collaboratively and supported and refined through strategic input from external experts. The questionnaire included 12 sections and a total of 685 questions to collect data on the demographic characteristics of food vendors; their entrepreneurial motivations; business financing; enterprise character, operations, challenges and strategies; interactions with policy actors; and aspirations.

The survey sought to strike a balance between the issues of inclusive growth and informal governance. At the same time, it attempted to mediate differing views on the informal sector as a whole, not just informal food vendors. For this reason, the questions were generally informed by the definitions provided by the 2013 Report of the International Conference of Labour Statistics (ILO, 2013). However, every effort was made to avoid falling into a specific informal sector 'camp', and as such, the questions were further informed by definitional categories based on the work of the organization Women in Informal Employment: Globalizing and Organizing (WIEGO) (Chen, 2012).

To be eligible for inclusion in the study, the enterprise had to be selling food and employing fewer than five people. Informal food vendors are not a homogenous group. There are a variety of vendor types including fruit vendors on the street, vendors at the market, and vendors operating from shipping containers in neighbourhoods (known as spazas or tuck-shops). This meant that care had to be taken to ensure that these different types were captured in the survey.

Sampling Procedure

Household Food Security Survey

Previous survey work by AFSUN had sampled poorer, low-income neighbourhoods in three of the cities (Battersby, 2011; Pendleton et al., 2012; Raimundo et al., 2014). To achieve a more general picture and generate insights into food security inequality, the HCP adopted a city-wide sampling strategy using a stratified, random probability sample of a pre-agreed number of households per city. The sampling frame was based on the most recent National Census Enumeration Area (EA) sample frame from the respective countries and cities, updated with the latest population estimates. Every effort was made to ensure that a sample was representative across each city. However, it is necessary to exercise caution when claiming that these surveys are representative. In some instances, the final selection of households was influenced by external factors such as the form and density of a neighbourhood, the access to adult respondents in the targeted households, and security concerns (both for enumerators and respondents).

During the data collection period, external factors influenced the timing of surveys and data collection. In South Africa, for example, there were two outbreaks of xenophobic violence. Given the size of the migrant population in Cape Town, this meant that surveys and schedules had to be adapted to avoid these periods. Maputo experienced severe climatic events that affected survey schedules. Nairobi, encountered spikes in civic unrest which further influenced survey scheduling.

The survey was administered to a senior adult member of the household, that is, someone who could speak for the household. A household was defined as including everyone who regularly "ate from the same pot". There were 215 primary questions in the household survey questionnaire. Most of the questions pertained to the household, but the household profile section included questions for each member of the household. This allowed analysis of household and city dynamics by gender and age.

Informal Food Vendor Survey

Every effort was made to generate data that reflected the entire informal food sector in each city. However, caution must be exercised when claiming representativeness. A major limitation with vendor sampling is the absence of informal vendor censuses from which to create a sampling frame. As a result, a slightly different sampling strategy was adopted in each city (see below).

For security reasons, the surveys were only administered between 8am and 5pm. Therefore, traders operating outside of those times were excluded. Additionally, only a small number of informal food vendors operate in permanent structures. Many operate precariously and are often subject to clearances and official actions (Balbuena and Skinner, 2020). They therefore tend to be mobile or move according to the dynamics associated with clearances, permit inspections, and other constraints. This has two implications for the representativeness of the data collected. First, the constant movement of traders impacts on return visits and replication. Second, traders experiencing precariousness avoid exposure and, as a result, are reluctant to participate in surveys, and this affects sampling strategies.

Household Sample

Cape Town: The sampling frame used was based on the Census 2011 Enumeration Area (EA) frame, updated with the latest population estimates. The EAs and households in Cape Town were stratified by low, middle, and high household income categories. In each of the randomly selected EAs, six households were systematically selected. To capture various longitudinal trends, poor households in three poor neighbourhoods where previous AFSUN research was conducted were oversampled (Khayelitsha, Philippi, and Dunoon) (Table 1). The survey generated a stratified random probability sample of 2,514 households.

Maputo: The sampling strategy involved a two-stage process, First, 19 wards (bairros) were randomly selected in the five mainland districts in the City of Maputo. The sample sizes for these 19 wards were approximated using proportionate allocation sampling. The households were then systematically selected according to proportionate allocation. The total sample size was 2,071 households in the 19 *bairros* (Table 2).

Nairobi: To generate as representative a city-wide sample as possible, the survey was conducted in randomly selected administrative sublocations spread across all administrative districts (or sub-counties) and divisions of Nairobi City County. The households were located in 23 administrative locations and sublocations, covering all the administrative divisions and districts of Nairobi City County. The number of sampled households was determined using a multistage proportional-to-population size (PPS) random sampling procedure. The number of households sampled in each chosen sublocation was proportional to the total number of households in that sublocation. Depending on the form and density of the sublocation, a random sampling procedure was used to select the sampled households in residential neighbourhoods (estates) in the sublocation. The Nairobi household survey covered a total of 1,434 households in 23 areas/wards (Table 3)

TABLE 1: Cape Town Household Sample

Suburb/Type of area	No. of sampled households	% of sample
Khayelitsha	390	15.5
Philippi	384	15.3
Dunoon	200	8.0
Other low-income areas	844	33.6
Middle-income areas	498	19.8
High-income areas	198	7.9
Total	2,514	100.0

TABLE 2: Maputo Household Sample

Ward/bairros	No. of sampled households	% of sample
PolanaCaniço B	205	9.9
Maxaquene B	186	9.0
Laulane	172	8.3
Hulene A	166	8.0
Magoanine A	164	7.9
Zimpeto	156	7.5
Mafalala	109	5.3
Xipamanine	108	5.2
Malhangalene B	95	4.6
Malanga	86	4.2
Aeroporto A	85	4.1
Chamanculo A	83	4.0
Costa do sol	79	3.8
Magoanine B	78	3.8
Inhagoia B	77	3.7
Alto mae A	70	3.4
Jardim	67	3.2
Central A	46	2.2
PolanaCimento B	39	1.9
Total	2,071	100.0

TABLE 3: Nairobi Household Sample

Location	No. of sampled households	% of sample
Kawangware	192	13.4
Umoja	160	11.2
Huruma	112	7.8
Embakasi	111	7.7
Riruta	94	6.6
Lindi	74	5.2
Hamza	65	4.5
Zimmerman	62	4.3
Pangani	58	4.0
Roysambu	55	3.8
Hazina	50	3.5
Highridge	50	3.5
South C	49	3.4
Komarock	46	3.2
Makongeni	43	3.0
ShauriMoyo	41	2.9
Uhuru	40	2.8
Ngara East	30	2.1
Kenyatta/Golf Course	27	1.9
Kileleshwa	24	1.7
Karen	21	1.5
Bondeni/Gorofani	17	1.2
Spring Valley	13	0.9
Total	1,434	100.0

Windhoek: Households were sampled in the 10 Windhoek constituencies using a two-stage sampling design. First, primary sampling units (PSUs) were randomly selected with probability proportional to size from a master frame developed and demarcated for the 2011 Population and Housing Census. Within the 10 constituencies, a total of 35 PSUs were selected covering the whole of Windhoek, and 25 households were systematically selected in each PSU. The sampled PSUs and households were located on maps, which were used to target 875 households (Table 4).

Informal Vendor Sample

Informal food vendors are not a homogenous group. There are a variety of vendor types, such as roadside fruit vendors, street-side restaurant vendors, market stall vendors, and vendors operating out of shipping containers in neighbourhoods. This meant that care needed to be taken to ensure that these different types were noted and then captured in the survey. The target sample size was 400 in Windhoek and 1,000 in the other three, much larger cities. In each case, the target was exceeded.

Cape Town: A strategy of 'maximum variation' sampling was used to ensure a mix of commercial, formal residential, informal residential, mixed formal and informal residential, and industrial retail sites. In these areas, the main street served as the primary site of research. Informal food vending businesses were randomly selected. The final sample size was 1,025 vendors.

Maputo: A strategy of maximum variation sampling was used to capture a mix of commercial, formal residential, informal residential, mixed residential, dumba-nengues (concentrations of informal traders), Class C and D markets, and industrial retail sites. In residential and industrial neighbourhoods, the main street served as the primary site of research. In the markets, the food selling areas were targeted. Maputo has four different types of marketplace. Class A and B markets are provided with infrastructure (including toilets and drainage), while Class C markets are not. Class D markets are more informal and are not recognized by city authorities. All vendors were randomly selected. If a surveyed market fell within an already selected neighbourhood, this neighbourhood was excluded and an alternative chosen. The final sample size was 1,022 vendors.

Nairobi: To achieve city-wide coverage, the survey was conducted in the eight administrative divisions of Nairobi. At least three residential neighbourhoods were randomly selected in each of the eight administrative divisions for the study. Subsequently, respondents were randomly selected based on the form and density of the location of business enterprises in the residential neighbourhood, the category of the food vendor, the types of food items sold by the vendor, and their willingness to participate in the survey. The final sample size was 1,265 vendors.

TABLE 4: Windhoek Household Sample

Constituencies	No. of sampled households	% of sample
Moses Garoeb	150	17.1
Windhoek East	150	17.1
Samora Machel	125	14.3
Tobias Hainyeko	125	14.3
Khomasdal	75	8.6
Windhoek West	75	8.6
Katutura East	50	5.7
John Pandeni	50	5.7
Katutura Central	50	5.7
Windhoek Rural	25	2.9
Total	875	100.0

TABLE 5: Number of Records by City and Survey*

	Cape Town	Maputo	Nairobi	Windhoek
Households	2,504	2,071	1,414	863
Members	80,131	9,922	4,936	3,922
Vendors	1,025	1,022	1,265	470
Employees	518	64	271	470

* The Cape Town survey did not generate details of household members. Therefore, this number is an estimate based on 2016 demographic data from the City of Cape Town at https://municipalities.co.za/demographic/6/city-of-cape-town-metropolitan-municipality

Windhoek: The smaller size of the city meant that it was possible to perform a census of all food vendors that captured the type of enterprise and the spatial location. A total of 2,421 informal food vendors were mapped. A sample of eight key food vending areas was drawn from across the city. Within these areas, informal food vending businesses were randomly selected through a random number generator on the handheld tablet used to administer the survey. The final sample size was 470 vendors.

Table 5 summarizes the number of usable records by city and survey. Any discrepancies between these figures and those in the sampling sections above are a result of data cleaning.

Technical Validation

The possibility of errors in the data was reduced by input validation procedures in the data collection phase. Important among these was that qualitative questions were asked using predefined lists of answers. Using such lists creates a neat and comparable categorization of the possible responses to a question across the four cities, while also reducing the possibility of typological errors or incorrect/incoherent answers. To avoid popular or significant answers being overlooked, respondents were also allowed to answer each question in free-text if their answer did not match well with any of the available options.

The validation of daily data collection was carried out. Survey supervisors assessed daily responses and observed incomplete or erroneous responses in interviews from specific households or enterprises. The enumerators were asked to correct the responses or discard those interviews if corrections were not possible. This was an infrequent occurrence and was noted. The survey supervisors also assessed the time and location information for each interview, validating both the duration of the interview and the variation in the recorded responses. Enumerators were advised that these tests would be carried out, and no instances of fraudulent entries were observed. Responses were also checked by the survey supervisor to make sure they were numeric and in range. For example, to ensure ages were positive numbers and not impossibly large.

The use of handheld tablets and the associated survey design with safety measures in the HCP surveys helped in technical validation and integrity confirmation. For all surveys carried out using tables, data on the location (place) of interviews, start and end time of interviews – speed of survey completion, and time at which surveys were carried out was recorded. These recorded data, embedded within the reports, enabled the statistical assessment of each survey to ensure and confirm its integrity. All surveys were found to be suitably robust.

Open Access Data

Location

The data and supporting documents from the eight surveys of the four African cities have been made publicly available under an agreement with the DataFirst Open Data Portal at the University of Cape Town and are presented in diagrammatic form in Figure 1. This section details how the Hungry Cities Partnership (HCP) data files and documents were prepared and made available on DataFirst's

open data portal. The files were prepared with a single overarching goal in mind: creating data that are as useful as possible to the end user. To that end, the DataFirst standard data preparation protocols were used, based on international best practices. Examples of standard practices were confirmed by consulting the data preparation guide of the well-known ICPSR research data repository at the University of Michigan (ICPSR, 2012).

The HCP survey data files are uploaded along with explanatory metadata and supporting documents: questionnaires, technical documents, and fieldwork reports. Data files for each city survey are available in .csv and Stata formats. Additionally, DataFirst provides lists of linked citations to research publications based on HCP data. DataFirst provides a descriptive guide to each data set on-line under various headings and subheadings. Appendix A shows an edited version of the Study Description for the Cape Town surveys household and vendor surveys as an example.

Data Preparation

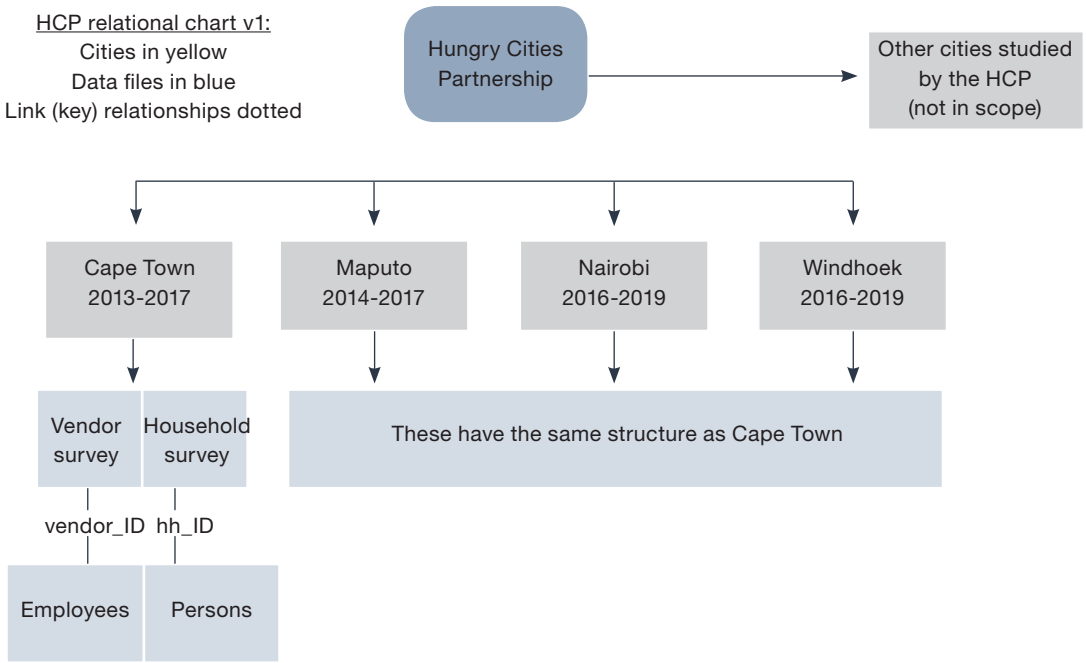
Several principles from DataFirst protocols were applied to the HCP data:

(a) *Conversion to a platform-agnostic format:* According to open data principles, DataFirst releases data in the.csv format, as csv is a 'platform-agnostic' format, meaning that it can be read on any platform (e.g Windows, Apple) without requiring the installation of additional software. DataFirst also releases data files in Stata format, because Stata is popular among researchers and is the preferred format for data cleaning. DataFirst also trains researchers in data analysis using Stata. Stata also allows import of data files in other formats. Raw data files were read from SPSS (.sav) and Excel (.xlsx), cleaned, and converted appropriately.

(b) *Hierarchical, succinct, and consistent variable naming:*

(i) Variable naming is somewhat of an art as there are competing desiderata. First, it is useful to keep

FIGURE 1: Data Files from the Hungry Cities Partnership African City Surveys



a variable name (varname) short as it serves as a short reference to the full question as stated on the questionnaire. Shorter names save time, and some data analysis software programmes truncate long variable names. For example, Stata by default limits varnames to a maximum of 32 characters. However, it is useful to include the question number in the varname, as well as a brief description of what was stored. The descriptive part is beneficial for two reasons: 1) the variable will show up in a keyword search and 2) it reminds the user what the variable stores. The variable names in the HCP data files were kept short by removing section letters from the questionnaire (which were unnecessary) and by abbreviating descriptive words.

(ii) Hierarchical naming systems help group similar questions and/or possible answers, improving usability. For the descriptive part of the varname, an adaption of the *prefix, root, suffix* system of the ICPSR (2012) was used. In this system, the question number is always the prefix. The *root-suffix* part is best explained by an example. A set of questions in the HCP all started with 'have you experienced the following problem:'. These questions all received the root word *problem* in the varname, directly after the question number. If three of these problem-related questions were about violence, a suffix 'violence' was added. Then the specific type of violence would be added as another suffix. The result is the following naming system: **qnumber_root_suffix1_suffix2**.

(iii) Variables were also named consistently across cities - i.e., if the same question was asked in questionnaires for two different cities, it got the same name, unless the question *number* differed, in which case this would be reflected in the qnumber. Fortunately, this scenario was rare, as the numbering was mostly consistent across cities. Items and ideas that appeared repeatedly in the data were also given consistent abbreviated names. A list of foods and their abbreviated names has been released with the data.

(iv) When a respondent answered 'other' to a question and had specified what they meant, an additional string variable would exist to store their answer. This variable was named x_oth, where the x is the name of the original nonstring variable.

(c) *Variable labels:* Variable labels contain full-text questions from the questionnaire, up to 80 characters, which is the maximum allowed by Stata. This is consistent with the ICPSR standards. If feasible, DataFirst recommends implementing shorter, self-explanatory variable labels, which include the question number for ease of reference. However, in the case of the HCP, which included more than 8,000 variables, re-labelling manually was impractical.

(d) *Missing values and dates:* It is best practice to avoid changing values within a data file as much as possible. This practice reduces the chance that the data scientist introduces mistakes or biases, and thus improves transparency. However, it is beneficial to standardize the appearance of missing values and nonresponses across variables and data files (ICPSR, 2012)34. In the HCP surveys, users could respond with 'do not know' or 'refused', and the enumerator could also write 'missing'. DataFirst harmonized these so that *refused*, *do not know*, and *missing* and are always stored numerically using -1, -2 and -3, respectively. These negative values are useful as an analyst can restrict their attention to positive values with simple code. In string variables, these were simply recorded as *refused*, *do not know,* and *missing*. The date variables were reformatted to be stored in machine-readable format.

(e) *Encoding variables and harmonizing lists:* DataFirst's approach to storage is to be as space and computationally efficient as possible without incurring data loss. To this end, they encode categorical data, as numeric storage is more efficient. Perhaps the largest part of the preparation work on the HCP was in creating the lists necessary to encode. A decision was necessary on whether the data for the variables for each city should have values that match the numbering on the questionnaire for *that city* or if it was better to harmonize the values between cities. For example, if the marital status question had a 'single' represented by 1 in the Cape Town questionnaire, but the Maputo questionnaire had this response coded as 2, harmonization would mean changing the code for this response to 1 in

Maputo. The ICPSR guide is ambiguous on what is best practice in this scenario, arguing both that it is best for values to match those in the questionnaires that they came from and that identical variables should have the same values and labels across files (ICPSR, 2012:22, 35).

A decision was made to harmonize across cities and files for various reasons. A user comparing a variable across two or more cities (which is the point of asking the same questions repeatedly) can do so without confusion. If the lists were not harmonized and instead the values for answers to specific questions jumped around between cities, analysts confused by or unaware of this could produce wrong results and attempts to merge data files across cities would create broken datasets. Harmonizing binary (yes/no) variables to 0 and 1 also improves usability, as analysts can use these as dummy variables. Lastly, the harmonization work done at DataFirst now saves future analysts having to repeat the process, eliminating huge amounts of work.

Cutting-edge tools were used for the encoding process in correspondence with Statalist. As a result of this correspondence, a new command was written that has been added to the Stata Baum repository (Klein, 2020). The command, called encodelabel, allows the user to import a list from one data file, find all answers that correspond to that list in a second data file, and encode them to match, while adding to the list wherever new values are found (in the second data file) that are not yet present in the list. All categorical variables were encoded this way, after being checked for typos to ensure that new list entries were not made unnecessarily.

(f) *Disclosure control:* In line with best practices, all GPS and location data were removed from the version of the data to be made public. The same is true for direct identifiers such as individual names and contact details. The free text responses were examined for potentially disclosive information (ICPSR, 2012).

(g) *Identification variables:* Non-disclosive identification variables (ID variables) were included so that users can match employees to vendors, and household members to households. These IDs were created either by using the existing random IDs made by the data collection team, or where those were absent, by using the index position of the household/vender in the data file.

(h) *Metadata and sampling description:* Once the data files were finalized, a metadata record could be created for each data set. DataFirst creates metadata that is complaint with Data Documentation Initiative (DDI) standards, the same standards used by the ICPSR and other well-established data repositories. Our metadata include information on scope, geographic coverage, sampling, data collection, and data processing (a summary of this document). Our metadata also highlights data quality concerns to help analysts reuse the data.

The technical validity of the HCP datasets relies on the following quality attributes: (1) the questionnaires have been tested previously and found to be reliable, (2) the surveys have reliable sampling properties, (3) the data have been thoroughly checked and cleaned and ready-to-use for analyses, (4) the data are accompanied by useful metadata; and (5) the data have broad geographical coverage of the cities concerned.

(i) *Code availability:* No computational code was developed to generate these data. The code used to prepare and clean the data is available from DataFirst on request from support@data1st.org.

Lessons Learned for Future Research Design

There are challenges in preparing research-ready data from a multi-country survey engaging context-specific variations. This section discusses ways in which future quantitative research projects can learn from the HCP experience, to improve the usability and comparability of the data files they produce, while also reducing the amount of work necessary in the data preparation phase. This is deemed essential given the data-related challenges discussed earlier, but also in the necessary ongoing

process of building data repositories focussing on urban data in Africa and the global South.

(a) *Question bank:* If different questionnaires are going to include similar questions (for example, across the cities of the HCP), it is beneficial to draw these from a common question bank. This ensures that the formatting and wording of questions are the same across multi-site projects. All members of the survey preparation team would have access to the question bank and would be able to see the process by which the questionnaires were developed.

(b) *Codebook - List of questions, items and item codes:* There are situations where it is necessary to 'tweak' a question across questionnaires, depending on the context. Often it is necessary to change a specific word or phrase in the question; for example, in the HCP the food items changed by city so that the foods being asked about were actually consumed in that city. In this situation, the recommendation is to keep a detailed record of the words substituted for the relevant questions in different survey sites.

In a second scenario, the text of the question is unchanged, but the possible answer categories may differ. In this situation, the recommendation is to maintain formal documentation of the values that were altered in each case, including the order and number of items on the list. Furthermore, if there are common items in the lists, comparability can be increased by arranging them so that the common items always appear in the same positions. For example, 'missing', 'do not know', and 'refused' can always be listed last.

We recommend that a comprehensive codebook is available that details the questions and codes for a given site, as well as across sites, as this has several advantages for data preparation: it helps one to understand what should be in the data files, it makes encoding simpler by providing master lists, and it makes it easier to track variables and items across data files. A codebook also simplifies the variable-naming process by providing a summary of the variables in the data.

(c) *Pre-defined naming system:* Variables in this project were named as described in the previous section using a semi-manual process, which is time-consuming and open to human error. This process required comparing variables and their answers across all files and cities and naming them accordingly, with matching variables receiving matching names and matching answers receiving the same code values. It is suggested that projects have a pre-defined prefix-suffix-root-type naming system that includes question numbers. Names of the variables, the variable labels, and the list of possible responses with values could be automatically created based on the codebook.

(d) *Policy on dealing with non-response:* The HCP benefitted from having a pre-defined set of non-response options, as mentioned above. How these are recorded and stored can also be planned. In numeric variables, it is recommended that values be out of range (ICPSR, 2012)34, ideally as negatives if the true (non-missing) values of the variable are non-negative. Values of 97, 98 and 99 which are sometimes used for missing values can be confusing if the variable could potentially take this value (for example, age). String variables should have consistent lettering for non-response options – "do not know" should not also appear as "don't know".

(e) *Policy on dealing with 'other' categories:* The same is true for how the answers of 'other' should be recorded. It is recommended that a policy be developed beforehand regarding:

1) whether there is a flag variable indicating that "other" was given or just the string variable

2) the maximum length of the possible answer

3) the wording of the text requesting the response, and

4) the naming of the additional "other" variable.

Providing the question number in the variable name for the 'other' responses would help to clearly link these responses to the question from which they arose.

(f) *ID management:* Where there is a parent-child relationship between data files, such as household to household member, ID variables are crucial so that individual household members can be matched to their households. ID numbers can be generated using a reliable system that is non-disclosive, with zero input from the fieldworker. Allowing field-workers to influence ID risks mistakes, such as duplicate IDs, which can undermine the matching system. The principle of creating an ID system beforehand is important because reverse-engineering an ID system and trying to create IDs after the fact can be difficult if not impossible. Fortunately, for the one HCP survey for which IDs were an issue, Cape Town, it was possible to create IDs using the index position in the data files, which was stored automatically by the data collection software.

(g) *Dates:* The most important date for the end user and for curation is the data reference date (the date of the interview). For this reason, it is advisable to include a 'date of interview' variable for when the data was collected. The format is less important as almost any format can be re-worked to be human-readable, but this is impossible if date data does not exist. The same is true of location variables, although this is less of an issue as it can more easily be inferred from context.

(h) *Project and document centralization:* It is recommended that data curation or management be considered during the life of the project, and if possible a central data management hub be created for project teams. This is increasingly important when the project involves different partner organizations and teams working somewhat autonomously. Online curation hubs can help connect team members and facilitate the sharing of key data management documents related to policies and standards on data handing and data naming. Ideally, the hub could be administered by a full-time data manager or curator who gives input throughout, including at the questionnaire design phase.

Usage Notes

The African city datasets from the HCP surveys are available as public use data from the DataFirst Open Data Portal. Data are shared as Stata files, but also in a platform-agnostic format (.csv) for access to the wider research and data science community. The harmonization process undertaken on the data allows these data to be compared across locations, which adds to their research value. The detailed provenance and usage metadata supports data reuse. The data have been anonymized to protect the personal information of respondents, and data users agree to adhere to privacy principles along with attribution requirements when they download the data.

Researchers must register once on DataFirst's data site to download data. Meta-data records for each data set provide usage information and tips on how to cite the data in research publications. Researchers should notify DataFirst of their research output based on the data so that their research can be highlighted on the data site in linked citations to the data. Data users with data queries can contact DataFirst's helpdesk at support@data1.st.org.

As of June 2023, the four city datasets had nearly 3,000 downloads. However, these downloads represent the selection of any information on the dataset landing page. Between May 2021 and June 2023 over 50 users downloaded the full datasets for specific cities in a manner that suggests active use. These users range from city departments to academic institutions.

City	Downloads
Cape Town	729
Maputo	761
Nairobi	887
Windhoek	414
Totals	2,791

Hungry Cities Outputs (To Date)

To illustrate the range of uses to which the Hungry Cities Partnership data has been put to date, this section lists all refereed publications on the four cities produced by the HCP team to mid-2023.

Charamba, V., Nickanor, N., & Kazembe, L. (2019). Validation of the HCP Survey Tool for Measuring Urban Food Insecurity: An Item Response Theory Approach. HCP Discussion Paper No. 37, Waterloo and Cape Town.

Crush, J., & McCordic, C. (2017). "The Hungry Cities Food Purchases Matrix: Household Food Sourcing and Food System Interaction" *Urban Forum* 28: 421-433.

Crush, J., Caesar, M., & Haysom, G. (2018). *The State of Household Food Security in Cape Town, South Africa*. Hungry Cities Report No. 12, Cape Town and Waterloo.

Crush, J., Kazembe, L., & Nickanor, N. (2023). "Opportunity and Survival in the Urban Informal Food Sector of Namibia" *Businesses* 3(1): 129-149.

Crush, J., Nickanor, N., & Kazembe, L. (2018). "Informal Food Deserts and Household Food Insecurity in Windhoek, Namibia" *Sustainability* 11(1): 37.

Haysom, G. (2020). Urban Food Governance Perspectives in Changing African and Southern Cities, HCP Discussion Paper No. 39, Waterloo and Cape Town.

Kazembe, L., Nickanor, N., & Crush, J. (2019). "Informalized Containment: Food Markets and the Governance of the Informal Food Sector in Windhoek, Namibia" *Environment & Urbanization* 31(2): 461-480.

Kazembe, L., Nickanor, N., & Crush, J. (2022). "Food Insecurity, Dietary Patterns, and Non-Communicable Diseases (NCDs) in Windhoek, Namibia" *Journal of Hunger & Environmental Nutrition* 17(3): 425-444.

McCordic, C.(2017). Household Food Security and Access to Medical Care in Maputo, Mozambique. HCP Discussion Paper No. 7, Waterloo and Cape Town.

McCordic, C., & Abrahamo, E. (2019). "Family Structure and Severe Food Insecurity in Maputo and Matola, Mozambique" *Sustainability* 11(1): 267.

McCordic, C., & Frayne, B. (2021). "The Network of Household Barriers to Achieving SDG 1, 2 and 3 in Maputo, Mozambique" *Journal of Hunger & Environmental Nutrition* 16(6): 739-750.

McCordic, C., & Raimundo, I. (2019). "The Inclusion of Young Entrepreneurs in the City of Maputo's Food Markets" *Urban Forum* 30: 461-475.

McCordic, C., Frayne, B, & Sgro, J. (2020). The Role of Infrastructure Access in Urban Household Vulnerability to Food Insecurity in Southern Cities. HCP Discussion Paper No. 42, Waterloo and Cape Town.

McCordic, C., Frayne, B., & Sunu, N. (2022). "Assessing Supermarket Patronage in Matola, Mozambique" *Urban Transformations* 4(1): 11.

McCordic, C., Riley, L., & Raimundo, I. (2021). "Household Food Security in Maputo: The Role of Gendered Access to Education and Employment" *Development Southern Africa* 38(5): 816-827.

McCordic, C., Frayne, B., Sunu, N., & Williamson, C. (2022). "The Household Food Security Implications of Disrupted Access to Basic Services in Five Cities in the Global South" *Land* 11(5): 654.

Nickanor, N., Crush, J., & Kazembe, L. (2021). "The Informal Food Sector and Cohabitation with Supermarkets in Windhoek, Namibia" *Urban Forum* 30: 425-442.

Nickanor, N., Tawodzera, G., & Kazembe, L. (2023). "The Threat of COVID-19 on Food

Security: A Modelling Perspective of Scenarios in the Informal Settlements in Windhoek" *Land* 12(3): 718 2023

Nickanor, N., Kazembe, L., Crush, J. & Wagner, J. (2018). *The Urban Food System of Windhoek, Namibia.* Hungry Cities Report No. 8, Cape Town and Waterloo.

Nickanor, N., Kazembe, L., Crush, J., & Wagner, J. (2021). "Revisiting the African Supermarket Revolution: The Case of Windhoek, Namibia" *Development Southern Africa* 38(2): 230-247.

Onyango, E., Crush, J., & Owuor, S. (2021). "Preparing for COVID-19: Household Food Insecurity and Vulnerability to Shocks in Nairobi, Kenya" *PLoS One* 16(11): e0259139.

Onyango, E., Crush, J., & Owuor, S. (2021). "Migration, Rural-Urban connectivity, and Food Remittances in Kenya" *Environments* 8(9): 92.

Onyango, E., Crush, J., & Owuor, S. (2023). "Food Insecurity and Dietary Deprivation: Migrant Households in Nairobi, Kenya" *Nutrients* 15(5): 1215.

Owuor, S. *The State of Household Food Security in Nairobi, Kenya.* Hungry Cities Report No. 11, Cape Town and Waterloo.

Owuor, S. (2019). "Urbanisation and Household Food Security in Nairobi, Kenya" In Nagao, M., Masinja, J., & Alhassam, A. (Eds), *Sustainable Development in Africa: Case Studies* (Denver: Spears Media Press), pp. 161-174.

Owuor, S. (2020). *Inclusive Growth and Informal Food Vending in Nairobi, Kenya*. Hungry Cities Report No. 21, Cape Town and Waterloo.

Raimundo, I., McCordic, C., & Chikanda, A. (2018). *The State of Household Food Security in Maputo, Mozambique.* Hungry Cities Report No. 10, Cape Town and Waterloo.

Raimundo, I., Wagner, J., Crush, J., Abrahamo, E. & McCordic, C. (2020). *Inclusive Growth and Informal Vending in Maputo's Food Markets*. Hungry Cities Report No. 18, Cape Town and Waterloo.

Riley, L.. & Caesar, M. (2018). "Urban Household Food Security in China and Mozambique: A Gender-Based Comparative Approach" *Development in Practice* 28(8): 1012-1021.

Tawodzera, G. (2019). "The Nature and Operations of Informal Food Vendors in Cape Town" *Urban Forum* 30(4), 443-459.

Tawodzera, G., & Crush, J. (2019). *Inclusive Growth and the Informal Food Sector in Cape Town*. South Africa Hungry Cities Report No. 16, Cape Town and Waterloo.

Wagner, Hinton, L., McCordic, C., Owuor, S., Capron, G., & Gonzalez, S. (2019). "Do Urban Food Deserts Exist in the Global South? An Analysis of Nairobi and Mexico City" *Sustainability* 11(7), 1963.

Young, G. (2020). Reconceptualizing Informal Economic Governance: Implications from Cape Town, South Africa. HCP Discussion Paper No. 38, Waterloo and Cape Town.

Acknowledgements

The research and publication of this report was funded by the International Development Research Centre (IDRC) Open Research Data Initiative and the IDRC and Social Sciences and Humanities Research Council (SSHRC) project under the International Partnerships for Sustainable Societies (IPaSS) Program. The Windhoek household food security survey was funded by the Open Society Foundation South Africa (OSF-SA). We also wish to acknowledge the work carried out by Dr Wade Pendleton and Dr Cameron McCordic in the initial development of the surveys and associated coding, loading onto and piloting of tablet use and then subsequent support to city-level research partners. The work of all city-level partners and enumerators who collected data and supported this process is also duly acknowledged.

References

1. Acuto, M. and Parnell, S. (2016). "Leave no city behind." *Science*. 352, 873 (2016).
2. Ballard, T. J., Kepple, A. W. and Cafiero, C. (2013). *The food insecurity experience scale: development of a global standard for monitoring hunger worldwide*. Technical Paper version 1.1. Food and Agriculture Organization.
3. Balbuena, P. and Skinner, C. (2019). "Where are the inclusive cities? Street vendors globally face increasing hostility". In *Street Vendors and Public Space: Essential insights on key trends and solutions*. WIEGO. https://www.wiego.org/sites/default/files/resources/file/Street%20Vendors%20and%20Public%20Space%20-%20An%20EBook.pdf
4. Battersby, J. (2011). *The State of Urban Food Insecurity in Cape Town*. AFSUN Urban Food Security Series No. 11, AFSUN, Cape Town and Kingston.
5. Bilinsky, P. and Swindale, A. (2007). *Months of adequate household food provisioning (MAHFP) for measurement of household food access: indicator guide.* Food and Nutritional Technical Assistance Project, Academy for Educational Development.
6. Chen, M. A. (2012). *The informal economy: Definitions, theories and policies*. WIEGO Working Paper, 1, 90141-4.
7. Coates, J. (2013). "Build it back better: Deconstructing food security for improved measurement and action." *Global Food Security*. 2, 188-194.
8. Coates, J., Swindale, A. and Bilinsky, P. (2007). *Household Food Insecurity Access Scale (HFIAS) for Measurement of Household Food Access: Indicator Guide (Version 3).* Washington, DC: FANTA.
9. Crush, J. (2016) Hungry Cities of the Global South. HCP Discussion Paper No. 1, Waterloo and Cape Town.
10. Crush, J. and Riley, L. (2017). Rural bias and urban food security. Urban food systems governance and poverty in African cities. Hungry Cities Discussion Paper No. 11, Waterloo and Cape Town.
11. Haysom, G. and Tawodzera, G.(2018). "Measurement drives diagnosis and response": Gaps in transferring food security assessment to the urban scale." *Food Policy*. 74, 117-125.
12. International Labour Organisation (ILO)(2013). *Statistics of work, employment and labour underutilization.* 19th International Conference of Labour Statisticians, Geneva. Report II. ILO, Department of Statistics.
13. Inter-university Consortium for Political and Social Research (ICPSR) (2012). *Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle* (5th ed.). Ann Arbor, MI. https://www.icpsr.umich.edu/files/deposit/dataprep.pdf (2012).\
14. Jones, A.D., Ngure, F.M., Pelto, G. and Young, S.L. "What are we assessing when we measure food security? A compendium and review of current metrics." *Advances in Nutrition.* 4, 481–505.
15. Klein, D. (2020). ENCODELABEL: Stata module to encode string variable into categorical variable. EconPapers. Available: https://econpapers.repec.org/software/bocbocode/s458862.htm.
16. Lipton, M. (1984). "Urban bias revisited." *The Journal of Development Studies*. 20, 139-166.
17. Mattes, R. (2008). "The Material and Political Basis of Lived Poverty in Africa: Insights from the Afrobarometer." In Møller, V., Huschka, D. & Michalos, A, *Barometers of Quality of Life Around the Globe: How Are We Doing?* 161-185).
18. Mattes, R., Dulani, B. and Gyimah-Boadi, E. (2016). *Africa's Growth Dividend? Lived Poverty Drops Across Much of the Continent.* Policy Paper No. 29. Afrobarometer.
19. NUA (New Urban Agenda) (2016).*United Nations Conference on Housing and Sustainable Urban Development, Habitat 111.* Quito, Ecuador.
20. Pendleton, W., Nickanor, N., and Pomuti, A. (2012). *The State of Urban Food Insecurity in Windhoek, Namibia*. AFSUN Urban Food Security Series No. 14, Cape Town and Kingston.
21. Pieterse, E., Parnell, S. and Haysom, G. (2018). "African Dreams: locating urban infrastructure in the 2030 sustainable developmental agenda". *Area development and policy*. 3,149-169 (2018).
22. Raimundo, I., Crush, J., and Pendleton, W. (2014). *The State of Urban Food Insecurity in Maputo, Mozambique*. AFSUN Urban Food Security Series No. 20, Cape Town and Kingston.
23. Swindale, A. and Bilinsky, P. (2006). *Household dietary diversity score (HDDS) for measurement of household food access: indicator guide.* Washington, DC: FANTA, Academy for Educational Development.
24. UN DESA (2018). *World Urbanization Prospects: The 2018 revision.* Report number ST/ESA/SER.A/420. United Nations Department of Economic and Social Affairs, New York.

APPENDIX A

HUNGRY CITIES PARTNERSHIP SURVEYS, 2013-2019: Cape Town

A.STUDY DESCRIPTION (edited)

1. Identification

SURVEY ID NUMBER

zaf-acc-hcp-cpt-2013-2017-v1

TITLE

Hungry Cities Partnership Survey 2013-2017, Cape Town

SUBTITLE

Cape Town

COUNTRY

Name: South Africa Country Code: zaf

STUDY TYPE

Household Survey [hh]

ABSTRACT

This study covers Cape Town, one of the four African cities surveyed between 2013 and 2019 by the African Centre for Cities. The African Centre for Cities is based at the University of Cape Town and is a partner of the Hungry Cities Partnership (HCP). The HCP studies include household data on food insecurity, household food purchasing dynamics, nutritional discounting taking place in households, foods consumed, and multidimensional measures of poverty. The household data is complimented with household member data and food retailer (vendor) data, including information on vendor employees. The Hungry Cities Partnership is an international network of cities and city-based partner organizations which focusses on the relationships between rapid urbanization, informality, inclusive growth, and urban food systems in the Global South.

KIND OF DATA

Sample survey data

UNIT OF ANALYSIS

Households and individuals

2. Version

VERSION DESCRIPTION

V1: Cleaned and anonymized for public use

VERSION DATE

2020-11-06

3. Scope

NOTES

The household surveys adopted the USAID-aligned Food and Nutrition Technical Assistance modules, detailed by the measures of the Household Food Insecurity Access Scale, Household Food Insecurity Access Prevalence Scale, Household Dietary Diversity Score and the Months of Adequate Food Provisioning. Surveys provide data on food insecurity, household food purchasing dynamics, nutritional discounting that occurs in households, and foods consumed. In addition to this, the surveys also provide insights into the levels of multidimensional poverty, through the use of the Lived Poverty Index. This combination highlights the connections between food insecurity and lived poverty.

4. Coverage

GEOGRAPHIC COVERAGE

The household sample is designed to be representative of the city of Cape Town.

GEOGRAPHIC UNIT

In the public release the lowest geographic level is the city (Cape Town). In the secure version of the data, the lowest geographic/administrative unit at which disaggregated data is available is the Enumeration Area. GPS data is also available in the secure version.

UNIVERSE

Households in Cape Town

5. Producers and Sponsors

PRIMARY INVESTIGATORS

The Primary Investigator for the household survey was the African Food Security Urban Network (AFSUN)

FUNDING AGENCY/SPONSOR

Social Sciences and Humanities Research Council (SSHRC) (household survey)

International Development Research Centre (IDRC) (vendor survey)

6.Sampling

SAMPLING PROCEDURE

Household sampling: the sample for the 2013 AFSUN Food Security Study was designed to be two-stage and stratified, using a random probability sample of 2,500 Cape Town households. The number areas

were taken from the Statistics SA master lists and used as the primary sampling unit. Households were the standard sampling unit. Stratification was done by income group of the household. Some areas were over-sampled to improve accuracy. In each of the drawn EAs, six households were systematically selected, with the exception of the EAs in DuNoon (where 10 households were systematically selected). Starting points were allocated to ensure coverage of the entire EA. The household was defined by everyone who regularly "ate from the same pot". Vendor sampling: The survey team documentation reads as follows: A strategy of maximum variation sampling was used to ensure a mix of commercial, formal residential, informal residential, mixed formal and informal residential and industrial retail sites. In these areas, the main street served as the primary site of research. Informal food vending businesses were randomly selected. In total, 1,018 food vendors were interviewed over a three-week period. For more on sampling, see the study documentation.

DEVIATIONS FROM THE SAMPLE DESIGN

In cases, xenophobic violence made vendor interviews dangerous in some areas.

WEIGHTING

Household: Design weights were calculated based on the survey design. These were adjusted post hoc to account for non-random patterns of nonresponse. The adjustment was done using 2015 mid-year estimates as the auxiliary data, and the CALMAR method. See the technical documents for more information. The vendor data is not weighted, although the collection team says it "tried to take a representative sample".

7.Data Collection

DATES OF DATA COLLECTION

Start	End	Cycle
2013	2017	Both surveys
2017	2017	Vendor survey
2013	2013	Household survey

DATA COLLECTION MODE

Face-to-face [f2f]

SUPERVISION

This project received ethics approval by the Ethics in Research Committee of the Faculty of Commerce at the University of Cape Town on the 13th of April 2015. Note: approval for the AFSUN household survey was also obtained from the Research Ethics Board of Wilfrid Laurier University, Canada.

DATA COLLECTION NOTES

The household data was collected in 2013, the vendor data in 2017.

DATA COLLECTORS

AFSUN/Citizen Surveys

8. Questionnaires

There are two questionnaires, a household questionnaire and a vendor questionnaire. The household questionnaire has a subsection for household members (persons), and the vendor questionnaire has a subsection for employees. The answers to these subsections are supplied in separate data files that can be matched to (merged with) the questionnaire as necessary. Vendor surveys were administered to the person directly responsible for the running of the business using handheld tablets. The household survey was administered to a senior adult member of the household, someone who could speak for the household. Note that for the household questionnaire, the Question 8 section changed slightly for Cape Town, in that the answers are not stored in a 'wide' format like the other cities. Rather, if a respondent provided more than one answer, additional variables were created. This is why the dataset has less variables and the question 8 section looks different. Only up to three locations were recorded in Section 8, even if the respondent mentioned more than three sources of food.

9. Data Processing

Datafiles were received by DataFirst in SPSS (.sav) and Excel (.xlsx) format. Variables had to be named, and variable labels were taken from question text. The variables were named according to the number of questions and the subject matter, in a hierarchical fashion. An effort was made to keep the number of questions consistent between cities where the same questions were asked for the 2013-2019 surveys. For the vendor data, Cape Town, Maputo, and Nairobi had almost identical questionnaires, and so the question numbers were naturally the same across these cities (harmonized). For the household data, Maputo, Nairobi, and Windhoek were similar and could be harmonized. This means that users could try stacking these datafiles. The Cape Town household questionnaire was different from the others and variable names would require adjustment to match the other cities. The missing values of 97, 98, and 99 were converted to -97, -98, and -99. There were some wrong question numbers in the vendor data questionnaires (typos) that were corrected.

10. Data Appraisal

It seems that there is slight mismatch between the Cape Town household questionnaire provided and the lists in the datafile, for an example see the question 15 income sources. In the Cape Town household data, data was not collected for the question 10.c and 10.d, about crops and time to travel to crops. In general, the lists change subtly between cities, for example the lists of foods in question 8 of the household data. As such the user should take caution when comparing across cities, and refer to the questionnaires. When the lists differed, list item letters (a-z) were left in the variable name as a second way for the user to check that the data match the questionnaire in the expected way. In Cape Town an answer to questions 15a and b "support from relatives" was captured although it does not reflect in the questionnaire.

11. Access Policy

CONTACTS

DataFirst Support, University of Cape Town. Email: support@data1st.org

ACCESS CONDITIONS

Public use files, available to all

CITATION REQUIREMENTS

Hungry Cities Partnership, African Centre for Cities, University of Cape Town. Hungry Cities Partnership Survey 2013-2017, Cape Town. [dataset]. Version 1. Cape Town: HCP [producer], 2020. Cape Town: DataFirst [distributor], 2020.DOI: https://doi.org/10.25828/9pn8-5b51

12. Metadata Production

PRODUCERS

DataFirst, University of Cape Town

B. DATA DESCRIPTION

1. Data Files

	Cases	Variables
cpt-2017-employee-v1	99	20
cpt-2017-vendor-v1	99	1932
cpt-2013-hh-v1	999	521

2. Downloads

Questionnaires

Cape Town HCP Food Retail Survey

Download [PDF, 18.11 MB]

AFSUN Food Security Household Survey

Download [PDF, 951.38 KB]

Technical Documents

Household survey on food security in the City of Cape Town

Download [PDF, 558.5 KB]

hcp-2013-2019-food-lists.xlsx

C. RELATED PUBLICATIONS

https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/844

Similar study descriptions are available for the Maputo, Nairobi and Windhoek are available at:

Maputo: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/842

Nairobi: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/842

Windhoek: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/844