

Module 2



USING SECONDARY DATA SOURCES

Module 2 Overview



In this module you will learn how to:

- Identify and access the available data sets;
- Determine the data sets that are likely to be most useful in your project;
- Evaluate key strengths and drawbacks of each data set;
- Decide when collecting primary data is necessary
- Case study, links and appendices available as a resource for assessment teams

Value of Secondary Data



- Before undergoing primary data collection, important to make use of the wealth of information available from local, state, Federal, and private sources.
- This “secondary” data often provides essential insights rather rapidly, and in a more standardized format such that information can be compared across regions in the country.
- We recommend exploring available secondary data sources before determining whether or not you will need to collect primary data.

Case Study: Bureau of Labor Statistics

(<http://www.bls.gov/cex/>)



- BLS interviews about 120,000 households each year
 - Track what they spend and list all of their consumer purchases
 - Information about their household spending (primary data), only shared and made public as aggregated information (secondary data) without individual identifiers.
- BLS Consumer Expenditure Survey
 - Categorizes the survey results by income level, region, race, ethnicity and other relevant attributes.
 - Look up how much money was spent buying food each year by an average household in one particular region of the country.
 - Allow you to calculate a reasonable approximation of the amount residents of your community spend each year

Questions to Guide Data Collection



- Which data will tell a story that moves your food systems work forward the most effectively?
- How precise do the data need to be to serve your purpose?
- How recent do the data need to be to be useful and persuasive?
- If the data are a few years old, what may have changed in your community since the data were compiled?
- How close a fit are the readily available data to the questions you are trying to answer?

Quality and Appropriateness of Data



- Who collected these data, and for what purpose?
 - Does the source introduce any bias into the data set?
- Is this data set appropriate?
 - When you show your findings to local stakeholders, do the data reflect their experiences?
- Does the data align with the how you plan to categorize and report the information in your study?
- Can your team identify patterns in time-series data that would illuminate new trends?
 - How often is data compiled? Was it recently reported?
- Are you able to map data you retrieve from website?

More Considerations for Data



- Is the data set you are looking at a compilation of raw data from specific respondents?
 - Has it been processed through a mathematical model to represent averages or aggregate numbers?
- How large is the original sample in the data?
 - Does the source explicitly list the error ranges, or offer other explanations that allow you to interpret reliability?
 - Does the data's accuracy diminish when you study smaller geographic areas, like a neighborhood or a city?

Production Data



- Website provides matrix of sources, characteristics and limitations (ADD link when available)
- Most data is from government sources, some annual, but many sources are updated every five years
- Most data is available at state and county level, but currently, little standardized data is available at the operation or parcel level
- New innovations in GIS mapping, big data and precision agriculture may lead to more interesting data availability in the near future

Food Handling, Processing, Marketing, and Distribution Data



- Website provides matrix of sources, characteristics and limitations (ADD link when available)
 - Prices, volumes, operation characteristics are available but spread across a broader range of sources
- These data are also commonly available from government sources, USDA and beyond
 - This topic encompasses the broader economy so the Business Census provides data “beyond the farmgate”
- Local directories, private data aggregators (industry trade reports) and state agencies are key as well
 - Some secondary data will not be free or publicly available

Food Consumption



- Website provides matrix of sources, characteristics and limitations (ADD link when available)
- Bureau of Labor Statistics is a key source as discussed in previous slide
- USDA Economic Research Service (ERS) reports
 - Explores changes in behavior
- USDA ERS also compiles and combines some data sources to estimate other topics of interest
 - Food Waste
 - Cost of eating different dietary choices

Waste Recycling



- Website provides matrix of sources, characteristics and limitations (ADD link when available)
- Municipalities often keep data on compostable wastes that are collected by local waste hauling services
 - Of growing interest in the food security discussion
 - Local government personnel may measure how much of this waste matter is converted into compost, but not common
- Some schools or other institutional food services also track the waste they generate

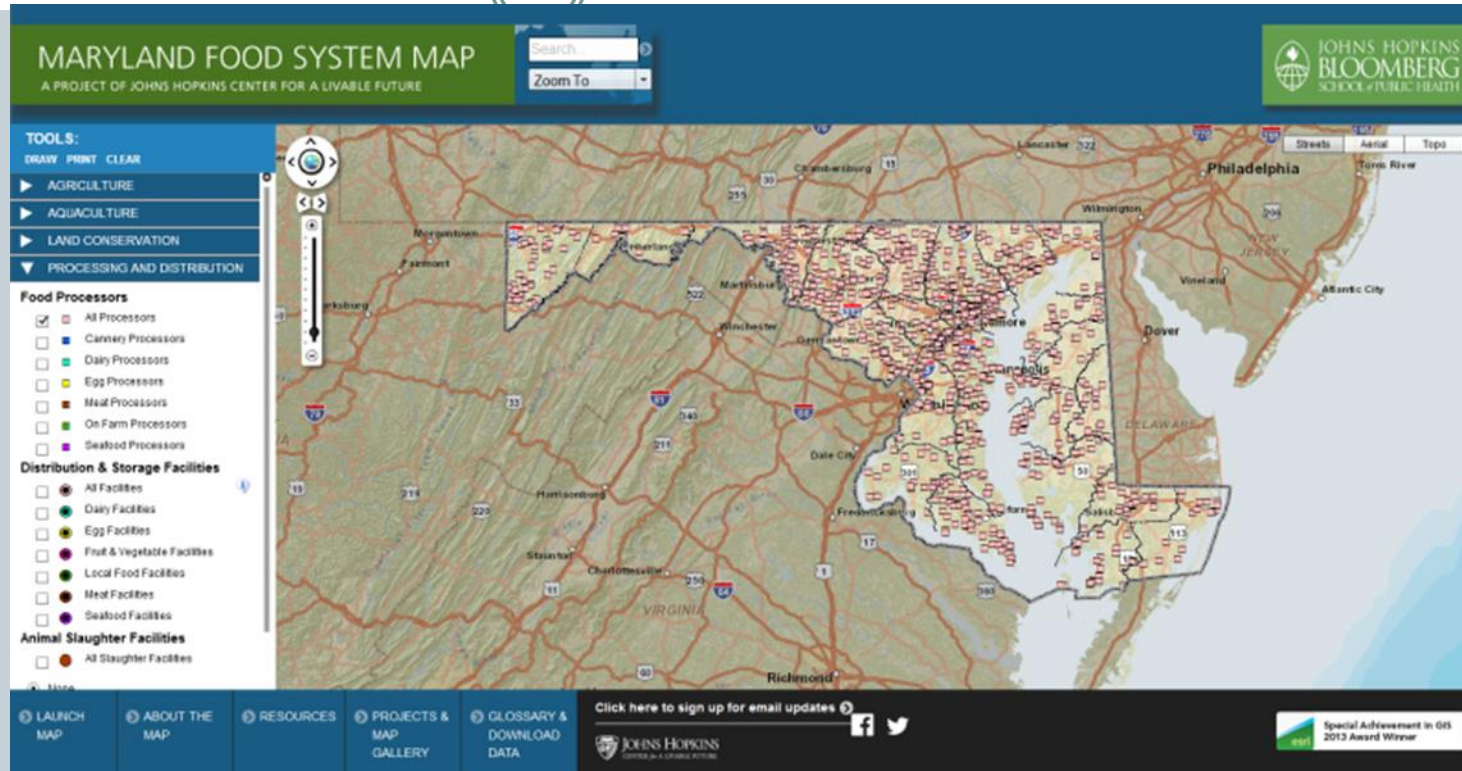
Demographic and Economic Contexts



- A variety of household, ag and business censuses are of value to these discussions
- Are the demographics of the farming population in the region similar to the entire state? The U.S.?
 - Do most farmers have off-farm jobs?
 - What is the average age of farmers in the region?
- What is the average household income level in your study area? How does this level compare to the entire state? The U.S.?
 - Is there any difference in average household income levels between urban and rural areas?

Maryland Food System Map

The Maryland tool designed to assist local food leaders and educators to visualize the current landscape of Maryland's food system from farm-to-plate.



This interactive, GIS-based mapping tool and database allows users to overlay layers of data on a map of Maryland's food system including how food is grown, processed, distributed, sold, and consumed.

Benefits of Secondary Data



- Widely available
- Quick to access
- Relatively inexpensive to compile
- Developed according to professional, standardized protocols
- Often provide time-series data useful in identifying patterns and emerging trends as well as comparisons across different areas (cities, counties, states)
- Comparable to other studies as the data would be drawn from similar sources

Cautions about Secondary Data



- Data compiled at a national scale may not suit local conditions
 - National data may not be robust when paired down to the local level
 - Findings should be ground-truthed with local stakeholders
- Data may not address the questions you wish to answer
- Data may seem more precise than they actually are, and interpretation must be done carefully
- Data from one source may not be totally compatible with data from another source

Other Issues to Consider



- Comparisons across time may not be valid
 - Data protocols may have changed, so read the fine print
- Comparisons across geography may not be valid
 - Climatic, political and policy differences are important to consider
- In rapidly changing situations, conditions may have changed since data were compiled
 - Commodity price fluctuations, weather events and macroeconomic conditions sometimes change rapidly
 - Truth numbers a bit using the most current data to check validity and maintain credibility of the project

Module 2 Takeaways



- The U.S. collects and makes available a huge amount of data, from the local to the national level
 - Secondary data sets are often the first source you will consult to get a basic sense of local conditions.
- Data sets are highly useful but learn about their strengths and limitations, and interpret with care
 - Different data sources may measure the same quantities differently, and data collection protocol may change over time
- Local food initiatives have used secondary sets in a diverse variety of ways to provide context for food system or economic impact assessments