

---

# MCBS Advanced Tutorial on Using Community and Facility Data

# Version Control Log

<b>Date</b>	<b>Version</b>	<b>Revisions</b>
1/6/20	1.0	Initial version released based on data year 2017.
12/18/20	2.0	Updated select content and updated based on data year 2018.
12/21/21	3.0	Updated select content and updated based on data year 2019.
7/15/22	4.0	Updated to incorporate corrections to data year 2019 analytic examples.

# Section 1: Introduction





## Learning Objectives

- After completing this Medicare Current Beneficiary Survey (MCBS) advanced tutorial, you will be able to answer the following questions:
  - What are the similarities and differences in sample design, questionnaire instruments, data collection, and data processing methods for the Community and Facility questionnaire components?
  - Where can you find data for the Community component and the Facility component in the MCBS Limited Data Set (LDS) data files?
  - Which LDS variables can you use to identify a beneficiary's residence status?
  - What are the analytic guidelines for combining data collected by the Community and Facility components?



# Advanced Tutorial Outline

- Section 1: Introduction
- Section 2: Overview of Community and Facility Data in the MCBS
- Section 3: Analytic Guidelines for Combining Community and Facility Data
- Section 4: Analytic Examples for Cross-sectional Analysis Using Community and Facility Data
- Section 5: Analytic Examples for Longitudinal Analysis Using Community and Facility Data
- Appendix: SAS<sup>®</sup> Code for Analytic Examples

# Introduction to the MCBS

- The Medicare Current Beneficiary Survey (MCBS) is a continuous, multi-purpose longitudinal survey of a nationally representative sample of the Medicare population.
- Interviews are traditionally conducted in-person using computer-assisted personal interviewing (CAPI). However, due to the COVID-19 pandemic, data collection switched to phone-only interviews in March 2020 and throughout 2021.
- The MCBS consists of a sample of beneficiaries aged 65 and over and beneficiaries aged 64 and below with certain disabling conditions, residing in the United States.
- The MCBS is sponsored by the Office of Enterprise Data and Analytics (OEDA) of the Centers for Medicare & Medicaid Services (CMS) and is conducted through a contract with NORC at the University of Chicago (NORC).
- The MCBS is designed to aid CMS in administering, monitoring, and evaluating the Medicare program. A leading source of information on Medicare and its impact on beneficiaries, the MCBS provides important information on beneficiaries that is not otherwise collected through operational or administrative data from the Medicare program and plays an essential role in monitoring and evaluating beneficiary health status and health care policy.



## Inclusion of Facility Beneficiaries in the MCBS

- To obtain an accurate representation of all Medicare beneficiaries, the MCBS sample includes all beneficiaries regardless of residence status. The MCBS follows beneficiaries into and out of long-term care facilities to maintain a comprehensive profile of their health care utilization and expenditures.
- About 5-8% of the sample are beneficiaries who live in a long-term care facility (henceforth referred to as Facility) or alternate between living in the community and living in a facility.
- Unlike beneficiaries living in the community, beneficiaries living in facilities or their proxies do not complete an interview. Instead, an interviewer conducts the interview with appropriate facility staff and abstracts some information from medical records to reduce burden on facility staff.
  - Note: Due to disclosure concerns, data on beneficiaries living in facilities are not included in some MCBS products such as the Public Use Files.

# MCBS Definition of Facility

- A facility interview is conducted when the beneficiary lives in a long-term care or other residential facility that meets the following MCBS definition of a facility:
  - Is a place or unit of a larger place with three or more beds, and either:
    - Is certified by Medicare as a Skilled Nursing Facility (SNF);
    - **or** is certified by Medicaid as a Nursing Facility or an Intermediate Care Facility for the Mentally Challenged;
    - **or** is licensed as a Personal Care Home, Board and Care Home, Assisted Living Facility, Domiciliary Care Home or Rest Home by a state or local government agency; **or** provides 24 hours a day, 7 days a week supervision by a person willing and able to provide personal care;
    - **or** provides personal care services to residents (personal care may include assistance with eating, dressing, preparing meals, etc.).
- For the purposes of this tutorial, data collection for beneficiaries living in facilities is referred to as the “Facility component” and data collection for beneficiaries living in the community is referred to as the “Community component.” Further, data collected from the Facility component are referred to as “Facility data” and data collected from the Community component are referred to as “Community data.”

# Introduction to the Advanced Tutorial on Using Community and Facility Data

- This tutorial is intended to provide an overview of the differences between MCBS data collected from beneficiaries living in the community and those living in facilities and analytic guidance on when and how these data should and should not be combined.
- Analytic decisions about whether to include all beneficiaries regardless of residence status, or those living only in the community or only in facilities, are driven by both the research question and data limitations. However, since the underlying data collection approach differs based on the beneficiary's residence status, caution must be observed when combining data across these populations to address questions requiring analysis of all Medicare beneficiaries.
- In addition to reviewing the [\*MCBS New User's Tutorial\*](#), this advanced tutorial builds on information provided in the MCBS [\*Data User's Guides\*](#), [\*Methodology Reports\*](#), [\*Questionnaire User Documentation\*](#), and other documentation, which provide documentation of the differences between Community and Facility data.

# MCBS Documentation and Resources

- CMS provides a wide array of MCBS documentation that is publically available on the CMS MCBS website. This documentation contains more in-depth descriptions of the topics covered in this tutorial.
  - [Tutorials, including the \*New User Tutorial\* and topical Advanced Tutorials](#)
    - It is recommended that new MCBS data users refer to the *New User Tutorial* before consulting this or other Advanced Tutorials.
  - Additional data documentation including [Data User's Guides, Methodology Reports, and codebooks](#) for the Limited Data Set (LDS) files.
  - Annual [Questionnaires and Questionnaire User Documentation](#).
  - Annual [Chartbooks](#) and data tables.
  - Annual [Bibliographies](#), which include annotations starting in 2020.
  - Annual [Early Looks](#) and [topical infographics](#).

## CMS Website:

<https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/index>

### Medicare Current Beneficiary Survey (MCBS)

[Questionnaires](#)

[Data Documentation and Codebooks](#)

[Data Tables](#)

[Bibliography](#)

[Data Briefs and Tutorials](#)

## Section 2: Overview of Community and Facility Data in the MCBS



# Key Differences in How Community and Facility Data are Collected

No differences in sampling methods. Sampling for the MCBS is person-based, not setting-based.			
<b>Sampling</b>  <b>Data Collection</b>	<b>Key Difference</b>	<b>Community Component</b>	<b>Facility Component</b>
	<b>Survey Respondent</b>	Administered to the beneficiary or a proxy.	Administered to facility staff. Interviewers abstract some information from medical records to reduce burden on facility staff.
	<b>Questionnaire Scope</b>	Core sections collect data on socio-demographics, health insurance, utilization, cost, experiences with care, and health status. Topical sections collect data on housing characteristics, health behaviors, and knowledge and decision-making about Medicare.	Core sections collect data on socio-demographics, health insurance, utilization, cost, and health status. Facility questionnaire does not include topical sections.  Questionnaire is designed to align with the information available to facility staff. For example, it collects fewer details about the beneficiary's income and assets and supplemental insurance compared to Community.
	<b>Questionnaire Reference Periods</b>	Because of the differences in questionnaire administration, the reference periods for similar items can differ between the Community and Facility components. For example, questions about flu shot and pneumonia shot have different reference periods for the Community and Facility components.	
	<b>Utilization Data Collection</b>	Collects event-level utilization data for all service types the beneficiary received.	Collects counts per month of utilization for services provided inside and outside the facility.
	<b>Costs Data Collection</b>	Collects event-level cost data for all services types, which is linked to the utilization events.	Collects per-diem costs associated with each facility stay; costs for services outside of the facility are not collected.

# Key Differences in How Community and Facility Data are Processed

Data Processing	Key Difference	Community Component	Facility Component
	<b>Administrative Sources</b>	Medicare Enrollment Database	Medicare Enrollment Database, Certification and Survey Provider Enhanced Reports (CASPER), and Minimum Data Set (MDS).
	<b>Claims Matched to Survey Data</b>	All Fee-for-Service (FFS) claims	Only Skilled Nursing Facility (SNF) and Inpatient FFS claims.
	<b>Claims Added to Survey Data</b>	Data from all types of FFS claims are used to fill in gaps in the survey-collected data.	
	<b>Imputation</b>	Imputation is conducted for items each respective questionnaire would have collected. For example Facility only collects total income so only total income is imputed. The more detailed income and assets variables collected for Community only are imputed for Community only. Historic Medicare Advantage (MA) Encounter data are used to estimate aggregate payment amounts for medical events not covered by Medicare (e.g., Dental) for beneficiaries with MA coverage living in the community.	

## Takeaway:

- The Community component collects costs and utilization for all events, whereas the Facility component collects utilization for events inside and outside the Facility but only collects costs for the Facility inpatient stay. Therefore, the lack of non-FFS claims results in larger gaps in the Facility data than in the Community data.

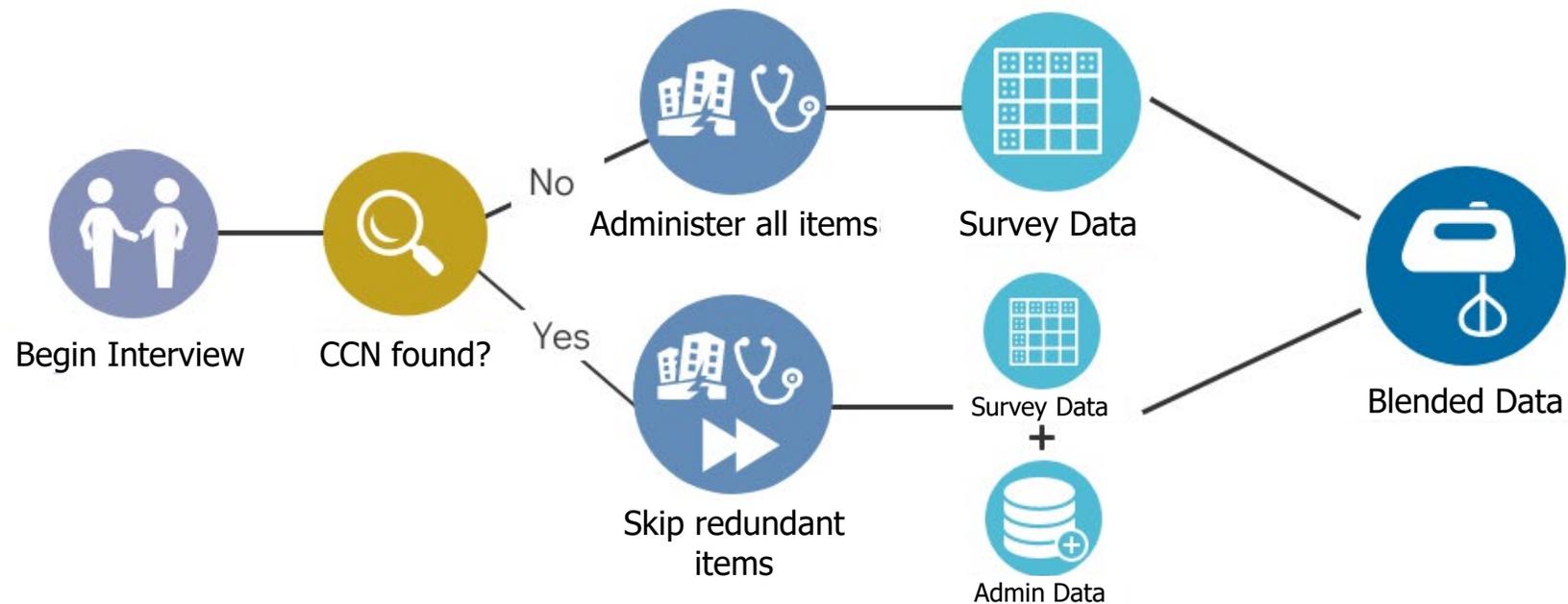


## Introduction to the Facility Redesign

- In Fall 2019 the MCBS Facility instrument was redesigned to skip items redundant with Certification and Survey Provider Enhanced Reports (CASPER) and Minimum Data Set (MDS) administrative data that are regularly reported to CMS, so as to shorten the Facility instrument for interviews conducted at Medicare- or Medicaid-certified facilities.
- The redesigned instrument provides a link to the administrative data used to replace the skipped questionnaire items by collecting the CMS Certification Number (CCN), a unique number that is assigned to all facilities certified by Medicare and/or Medicaid, via a questionnaire lookup tool.

# Overview of the Facility Redesign

- For facilities with a CCN, the Facility instrument skips more than 100 questionnaire variables that are redundant with CASPER and MDS in the Facility Questionnaire (FQ) and Health Status (HS) sections.





## Overview of the Facility Redesign Administrative Data Match

- For beneficiaries for whom the facility respondent reported a CCN, the survey-reported data are later merged with CASPER and MDS administrative data during data processing.
- CASPER administrative data on facility characteristics are linked directly to the FQ Survey Data directly on the CCN.
- MDS administrative data are first linked by CCN and BASEID, then to any MDS assessments that were reported in the survey.
- For detailed data matching protocols, reference the *2019 Methodology Report* Special Section: Facility Redesign.
- Notes in the codebooks will indicate when a variable has blended data.

## Impact of the Facility Redesign on the FACCHAR Segment

- The Facility Characteristics segment (FACCHAR), which is constructed using data from the Facility Questionnaire (FQ) section, is impacted by the Facility Redesign.
- FACCHAR provides information about survey-collected facility stays, and the administrative Provider of Service (POS) file, which provides facility characteristics pertaining to SNF stays.
- For beneficiaries for whom the facility respondent reported a CCN, 12 FQ variables used to create FACCHAR are skipped during data collection and are then populated with CASPER administrative data in data processing using the CCN.
  - Some of these variables from CASPER are not directly comparable to the survey-reported items.
  - For additional information on how CASPER data is used in the FACCHAR segment in these cases, reference the *Survey File Data User's Guide* data file notes.

## Impact of the Facility Redesign on the FACASMNT Segment

- The Facility Assessments segment (FACASMT), which is constructed using data from the Health Status (HS) section, is impacted by the Facility Redesign.
- FACASMT contains assessment information conducted while the beneficiary was living in a Medicare approved or non-Medicare approved facility.
- For beneficiaries for whom the facility respondent reported a CCN, more than half of the variables in FACASMNT are skipped during data collection and are then populated with MDS administrative data in data processing using the BASEID and CCN.
- For a list of variables on FACASMNT populated with MDS administrative data during data processing, reference the *Survey File Data User's Guide* data file notes.
- FACASMT rows populated with MDS data can be linked to corresponding rows in the LDS segment MDS3.

# Overview of MCBS Survey File LDS Segments by Data Source

- Because of the differences in data collection and data processing, it is important to know which LDS file segments contain data sourced from the Community component, Facility component, or administrative data. The table below contains this information for the 2019 data year. Please reference the *Survey File Data User's Guides* for other years and full definitions of the LDS segments.

● **Community Component**

■ **Facility Component**

◆ **Administrative Data**

## Survey File Segments

- |              |               |              |              |
|--------------|---------------|--------------|--------------|
| • ACCESSCR ● | • FACASMNT ■◆ | • INCASSET ● | • PNTACT ●   |
| • ACCSSMED ● | • FACCHAR ■◆  | • MAPLANQX ● | • PREVCARE ● |
| • ADMNUTLS ◆ | • FALLS ●     | • MCREPLNQ ● | • RESTMLN ●■ |
| • ASSIST ●   | • FOODINS ●   | • MDS3 ◆     | • RXMED ●    |
| • CHRNCND ●  | • GENHLTH ●   | • MENTHLTH ● | • RXPARTD ●  |
| • CHRNCDFL ◆ | • HISUMRY ●◆  | • MOBILITY ● | • SATWCARE ● |
| • CHRNPAIN ● | • HITLINE ●■◆ | • NAGIDIS ●  | • USCARE ●   |
| • DEMO ●■◆   | • HHCHAR ●    | • NICOALCO ● | • VISHEAR ●■ |
| • DIABETES ● | • INTERV ●■   | • OASIS ◆    |              |

# Overview of the Facility Redesign FACASMT and MDS3

- The FACASMNT segment contains blended administrative (MDS) and survey-reported data. The MDS3 segment contains administrative (MDS) data only.
- FACASMNT rows populated with MDS data can now be linked to corresponding rows in the MDS3 segment using the unique key BASEID, TRGT\_DT, and A2300.
  - Example SAS code for Merging FACASMNT and MDS3 can be found on Slide 32

Data Type	Facility Assessment (FACASMNT)	Minimum Data Set (MDS3)
<b>Source</b>	Blended administrative (MDS) and survey-reported (facility staff may pull information from electronic health records or systems to answer the survey questions)	Administrative (MDS)
<b>Population</b>	Represents all Facility residents, not just those in nursing homes	Represents all residents of nursing homes certified to participate in Medicare or Medicaid only
<b>Reference period</b>	Throughout the year	Could be multiple assessments during the year, time periods may differ based on what happened to each individual
<b>Unit of observation</b>	One per beneficiary	One per beneficiary per assessment

# Overview of MCBS Survey File LDS Segments by Universe

- As shown on slide 19, for some topics in the Survey File LDS (such as health and functional status) both Community and Facility data are available but included in separate LDS segments. The table below provides an overview by topic of the LDS segments that include data for each beneficiary universe.
- If the analysis includes all Medicare beneficiaries and the data are in separate Survey File LDS segments, data users will need to appropriately merge segments before analyzing the data. Question text, response categories, and reference period may differ between Community and Facility variables, requiring recoding in order to combine the variables for analysis.
- The table below contains this information for the 2019 data year. Please reference the *Survey File Data User's Guides* for other years and full definitions of the LDS segments.

Topic	Segments with Community Data	Segments with Facility Data	Segments with Data for All Beneficiaries
<b>Health Status</b>	GENHLTH; FALLS; CHRNCND; MENTHLTH; OASIS	FACASMNT; MDS3	
<b>Preventive Care</b>	PREVCARE	FACASMNT	
<b>Functional Status &amp; Assistance with Long-Term Care Needs</b>	ASSIST; NAGIDIS; OASIS; MOBILITY	FACASMNT; MDS3	
<b>Demographics and Socio-Economic Status</b>	INCASSET		DEMO
<b>Health Insurance Coverage</b>			HISUMRY; HITLINE; ADMNUTLS



## Overview of MCBS Cost Supplement File LDS Segments

- The Cost Supplement LDS segments contain both survey-reported data and administrative claims data on health service utilization, costs, and payers.
- While Cost Supplement segments contain data sourced from the Community component, Facility component, or administrative data, the data source varies by payer.
- Survey-reported data are matched to Medicare FFS administrative claims and Medicare Part D Prescription Drug Event data to correct for under-reporting of events and payments in the survey.
- Survey-reported data for services covered by other payers, such as MA, Medicaid, private payers, and other payer sources are not matched to administrative data.
- Remember that the MCBS definition of a Facility includes many types of long-term care facilities, some qualify to be paid for by Medicare and some do not.

# 2019 Cost Supplement File Segments by Data Source

Event-Level LDS Segment	Beneficiary Residence Status	Community			Facility			Included in LDS Summary Segments (SS & PS)
	Sources for Utilization and Cost Data	Survey Reported	Claims*		Survey Reported	Claims*		
	Beneficiary Health Insurance Status		FFS	Part D		FFS	Part D	
	Service Type							
PME	Prescribed medicine events	X		X			X	X
DUE	Dental utilization events	X	+			+		X
VUE	Vision utilization events	X	+			+		X
HUE	Hearing utilization events	X	+			+		X
OPE	Outpatient hospital events	X	X			X		X
IPE	Inpatient hospital events	X	X			X		X
IUE	Institutional events	X	X			X		X
MPE	Medical services, equipment, and supplies	X	X	X		X		X

Costs for services received **outside** of a Facility that are paid for by FFS or Part D are contained in the other, non FAE, event-level segments.

\*Claims payment information is only available for events paid for by Medicare FFS and Part D. Administrative data pertaining to events paid for by MA, Medicaid, private insurance, or out of pocket are not available to the MCBS. Historic Medicare Advantage Encounter Data are used during imputation for some event types. + FFS Medicare does not cover most procedures or supplies for dental, vision, and hearing segments. It only covers certain services that are received in a hospital. The segment includes claims data for the few services Medicare does cover, but the vast majority of events are Community survey-reported only.

For the FAE segment, home health, and hospice events, see the next slide. For other data years please see the *Cost Supplement Data User's Guide*.

# 2019 Cost Supplement File Segments by Data Source, Continued

Event-Level LDS Segment	Beneficiary Residence Status		Community			Facility		Included in LDS Summary Segments (SS & PS)	
	Sources for Utilization and Cost Data		Survey Reported	Claims*		Survey Reported	Claims*		
	Beneficiary Health Insurance Status			FFS	Part D		FFS		Part D
	Service Type								
FAE	Nursing home or other long-term care facility	• Cost of Facility <b>Stay</b>				X	X <sup>^</sup>		X
		• Services provided <b>Within</b> the Facility				X			
		• Services received <b>Outside</b> of the Facility				X			
Not available at event-level	Home health events		X	X			X		X
Not available at event-level	Hospice care events		X	X			X		X

Costs for services received **outside** of a Facility that are paid for by FFS or Part D are contained in the other, non FAE, event-level segments.

\*Claims are only available for events paid for by Medicare FFS and Part D. Administrative data pertaining to events paid for by MA, Medicaid, private insurance, or out of pocket are not available to the MCBS. Historic Medicare Advantage Encounter Data are used during imputation for some event types.

<sup>^</sup> Skilled Nursing Facility (SNF) and In-patient (IP) Medicare FFS claims are used during data processing of the FAE segment to impute missing Medicare payment amounts and to check the accuracy of the lengths of stays, when available. However claims-only records are not added to FAE whereas they are added for the other event-level segments. See Appendix C for more information on claims matching and imputation processes.

<sup>†</sup> Medicare FFS pays for hospice care for all beneficiaries, including those covered by MA.

# Identifying Beneficiaries by Residence Status in the Survey File LDS

- There are two ways to identify a beneficiary's residence status:

	<b>Interview Type</b>	<b>Beneficiary Classification Type</b>
<b>Survey File LDS Segment</b>	DEMO	RESTMLN
<b>Variable Name</b>	INT_TYPE	D_TYPE
<b>Code Frame</b>	C = Completed only Community interviews in the data year F = Completed only Facility interviews in the data year B = Completed both Community and Facility interviews in the data year	C = Lived in the Community the full year F = Lived in a Facility the full year B = Spent at least one day in both settings
<b>Definition</b>	Interview type as defined by the questionnaire completed for the beneficiary.	Beneficiary classification as defined by the number of days the beneficiary spent in each setting.

## Differences between INT\_TYPE and D\_TYPE

- For most beneficiaries, INT\_TYPE and D\_TYPE are consistent. However, because INT\_TYPE reflects the type of **interviews** conducted in a year and D\_TYPE reflects the **days** spent in the two settings, it is possible for the INT\_TYPE and D\_TYPE to be different if a beneficiary was administered only one type of interview but spent days in both settings.
  - A beneficiary could have spent a week in a facility but was only administered Community interviews, resulting in INT\_TYPE=C and D\_TYPE=B.

## Which Definition of Community and Facility should I use?

- Whether you use INT\_TYPE or D\_TYPE in your analysis depends on your research question.
  - If the research question requires distinguishing between the type of interviews administered, INT\_TYPE provides this information.
  - If the research question requires knowing the number of days spent in a setting, D\_TYPE provides this information.
- Because the Community and Facility components have differences in type and extent of data collected, researchers need to account for these differences and identify the best universe for analysis. The way to identify these universes is by using INT\_TYPE. Therefore, the rest of the tutorial is going to focus on using **INT\_TYPE** to define the universe.
  - For more information on using INT\_TYPE, see Section 10, “Data File Notes” in the *MCBS Survey File Data User’s Guide*.

# MCBS COVID-19 Community Supplements

- Two COVID-19 Community Supplements were fielded in 2020. An additional COVID-19 Community Supplement was fielded in Winter 2021. In Summer 2021, the COVID-19 Supplement content was incorporated into the main MCBS for subsequent interviews as appropriate. The data from the Community supplements is available as Public Use Files (PUFs) and as part of the LDS.
- A COVID-19 Facility Supplement was also administered to facility staff on behalf of beneficiaries living in a facility in Fall 2020 and Winter 2021. Due to disclosure concerns, those data will not be released as PUFs but will be available in the 2020 and 2021 LDS files, respectively.
- Additional information can be found in the *MCBS Advanced Tutorial on the COVID-19 Supplement Data*.

COVID-19 Supplements	Date of Survey Administration	Planned LDS File Release	Release Date of LDS File	Planned Public Use File Release	Release Date of Public Use File
<b>COVID-19 Summer 2020 Community Supplement</b>	June- July 2020	2019 Survey File COVIDSUM Segment	Summer 2021	COVID-19 Summer 2020 PUF	Released 10/16/20
<b>COVID-19 Fall 2020 Community Supplement</b>	October- November 2020	2019 Survey File COVIDFAL Segment	Summer 2021	COVID-19 Fall 2020 PUF	Released 1/15/21
<b>COVID-19 Winter 2021 Community Supplement</b>	March- April 2021	2020 Survey File COVIDWIN Segment	Summer 2022	COVID-19 Winter 2021 PUF	Released 7/27/21

## Section 3: Analytic Guidelines for Combining Community and Facility Data





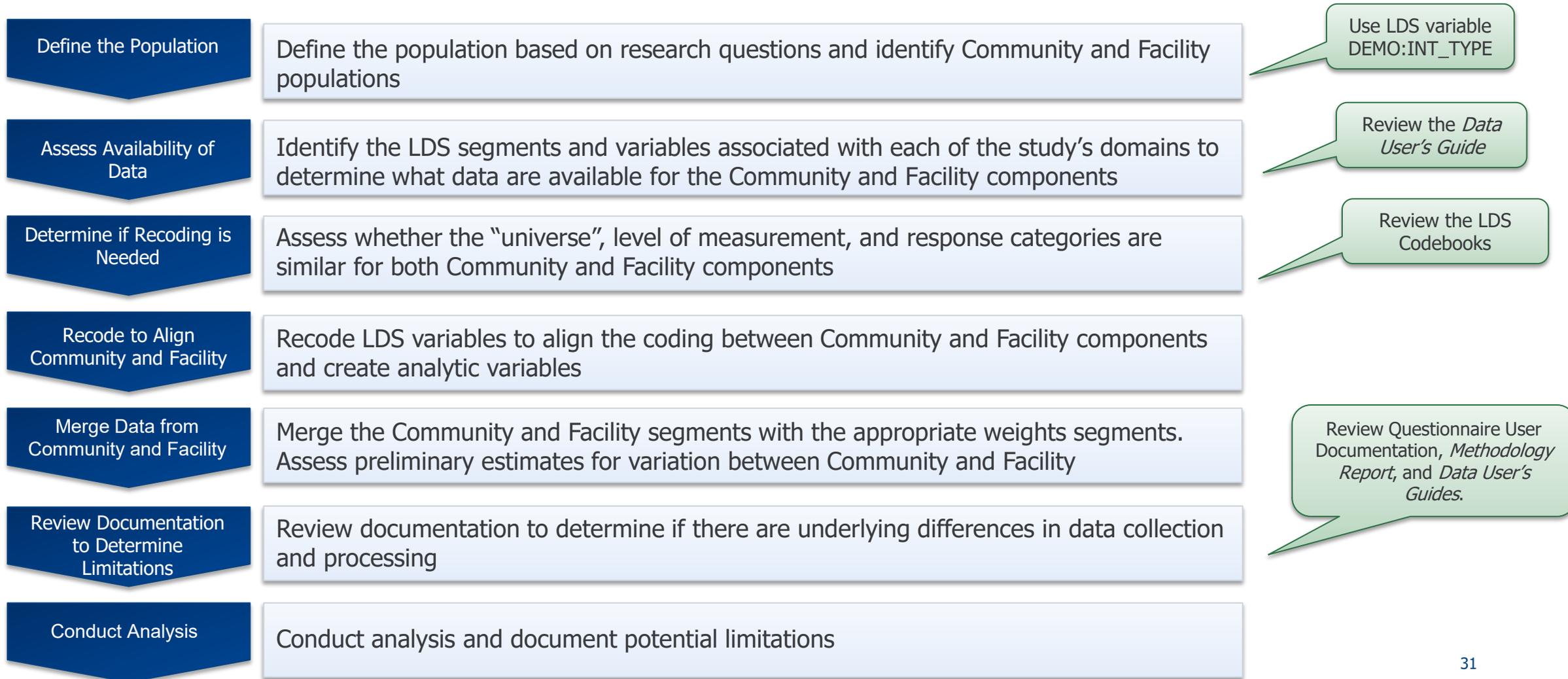
# Analytic Scenarios for Combining Data for the Community and Facility Populations

Data users may encounter the following analytic scenarios when conducting analysis using MCBS data:

1. Data are available from the Community or Facility components but not both
2. Data are available from both Community and Facility components and there are no differences in the LDS variable coding, data collection, and processing
3. Data are available from both the Community and Facility components but the LDS variable constructs and/or coding are different
4. Data are available from both the Community and Facility components and the LDS variable constructs and coding are similar, but data collection and/or processing methods are different

The steps on the next slide provide guidance on the recommended step-by-step procedures for combining data from Community and Facility components.

# Recommended Steps for Combining Community and Facility Data



## Example: SAS Code for Merging FACASMNT and MDS3

- As introduced on slide 20, the FACASMNT and MDS3 segments can now be merged, allowing data users to leverage the full MDS assessment information provided on MDS3.
- SAS code for creating a combined dataset of MDS and FACASMNT:

```
data facasmnt_mds3;  
  merge surveyyy.FACASMNT (keep=_ALL_)  
        surveyyy.MDS3 (keep=_ALL_);  
  by BASEID TRGT_DT A2300;  
run;
```

# Section 4: Analytic Examples for Cross-sectional Analysis Using Community and Facility Data



# Combining Data from the Community and Facility Components to Conduct Analysis

The analytic examples\* in the slides to follow show how the step-by-step guidance introduced earlier can be applied to the following analytic scenarios:

- Data are available for the Community or Facility component but not for both (**Example 1**)
- Data are available for both Community and Facility components and there are no differences in the LDS variable coding, data collection, and processing (**Example 2**)
- Data are available for both Community and Facility components but LDS variable constructs and coding are different (**Example 3**)
- Data are available for both Community and Facility components, and the LDS variable constructs and coding are similar, but data collection and processing methods are different (**Examples 4**)

\*SAS code for all examples can be found in the Appendix at the end of this tutorial.

# Example 1: Data are available for the Community or Facility component but not for both

**Research Question:** What is the level of satisfaction with the amount paid for prescription drugs for Medicare beneficiaries in calendar year 2019?

Define the Population

The research question is applicable to all beneficiaries. The target population includes beneficiaries enrolled in Medicare Part D who are also continuously enrolled in 2019 and still alive, and entitled in Summer 2020.

Assess Availability of Data

'Satisfaction with amount paid for RX drugs' (**MCAMTPAY**) – the outcome variable - is in the **RXMED** segment, which is a topical segment collected for Community beneficiaries only.

Found in the  
*Survey File Data  
User's Guide*

# Example 1: Data are available for the Community or Facility component but not for both

Review documentation to apply proper weights

Applying topical weights:

- Special non-response adjustment weights are included in the **RXMED** segment to account for survey non-response from the Fall to Summer data collection period. This includes the weight **RXSCWT** and replicate weights **RXSC1-RXSC100**
- When merging with appropriate weights, all observations in the weights segment file should be preserved

Found in the *Survey File Data User's Guide*

Limit the analysis based on availability of data

**RXMED** is a topical segment administered to Community in the Summer round. There is no satisfaction measure for Facility because the measure asks for beneficiary's opinions, which are not accessible to the facility staff. The *Data User's Guide* explains that there could be a small number of **INT\_TYPE=F** cases in RXMED because of when Summer topical sections are conducted versus released for analysis.

Found in the *Survey File Data User's Guide*

As a result, conduct analysis with data where **INT\_TYPE=C** and document that the study is limited to beneficiaries with data for the Community component who had Part D insurance coverage during 2019.

# Example 1: Data are available for the Community or Facility component but not for both

Conduct Analysis and Document Limitations

**Level of Satisfaction with the Amount Paid for Prescription Drugs Among Beneficiaries with Community Component Data and Enrolled in Medicare Part D (2019)**

MCAMTPAY	Description	Unweighted N	Weighted N (Standard Error)	Percent - % (Standard Error)
1	Very Satisfied	2,346	13,650,341 (315,731)	33.0 (0.7)
2	Satisfied	3,500	20,672,735 (344,495)	50.0 (0.7)
3	Dissatisfied	557	3,525,301 (164,888)	8.5 (0.4)
4	Very dissatisfied	172	1,219,221 (113,749)	2.9 (0.3)
5	No experience	366	2,264,448 (151,796)	5.5 (0.4)

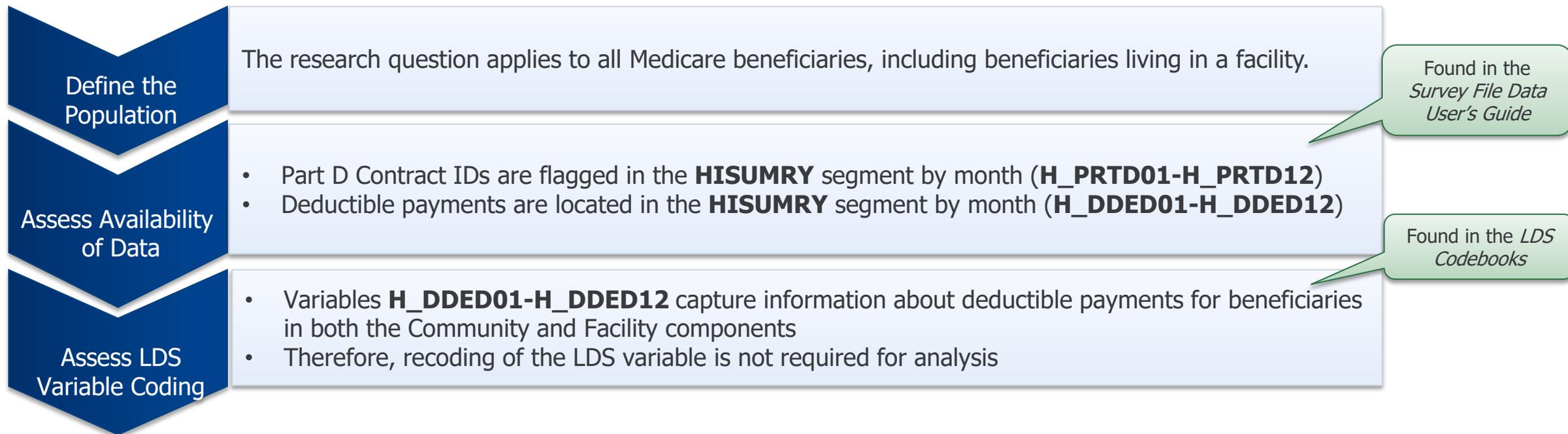
SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, Survey File, 2019.

**Analytic Note:** Beneficiaries with Facility component data were excluded. Data collection occurred during the Summer of 2020. Only beneficiaries with Community component data who were enrolled in Medicare Part D in the past year (H\_PARTD) were included.

Please refer to Appendix, Example 1 for detailed guidance on constructing the analytic file and the SAS code for producing estimates.

## Example 2: Data are available for both Community and Facility components and there are no differences in the LDS variable coding, data collection, and processing

**Research Question:** What is the average Part D plan deductible cap for Medicare beneficiaries enrolled in Part D during 2019?



## Example 2: Data are available for both Community and Facility components and there are no differences in the LDS variable coding, data collection, and processing

Review Documentation to Assess Underlying Differences in Data Collection and Processing

Information on Part D enrollment and plan deductible payment caps is sourced from administrative data, which covers all beneficiaries. Therefore, there are no differences in data collection and processing between the Community and Facility components.

Found in the *Survey File Data User's Guide*

Review documentation to apply proper weights

Applying full sample weights:

- Full sample weights for the continuously enrolled population are included in the **CENWGTS** segment. This includes the weight **CEYRSWGT** and replicate weights **CEYRS001-CEYRS100**
- When merging with appropriate weights, all observations in the weights segment file should be preserved

Found in the *Survey File Data User's Guide*

## Example 2: Data are available for both Community and Facility components and there are no differences in the LDS variable coding, data collection, and processing

Conduct Analysis  
and Document  
Limitations

Average Annual Part D Plan Deductible Cap for Medicare Beneficiaries Enrolled in Medicare Part D (2019)

Unweighted N	Mean Deductible - \$ (Standard Error)
10,168	205.0 (4.1)

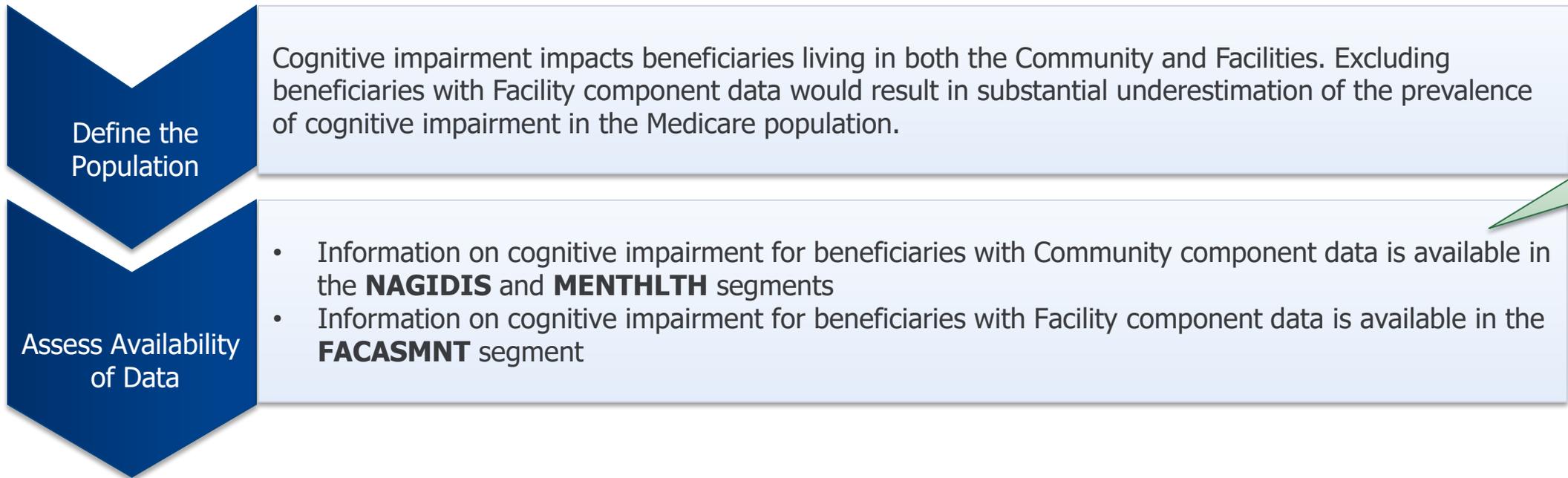
SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, Survey File, 2019.

**Analytic Note:** Information on Medicare Part D enrollment and deductible payments is sourced from administrative data.

Please refer to [Appendix, Example 2](#), found at the end of this tutorial, for detailed guidance on constructing the analytic file and the SAS code for producing estimates

## Example 3: Data are available for both Community and Facility components; LDS variable constructs are different

**Research Question:** What is the prevalence of cognitive impairment among all beneficiaries aged 65 and over in 2019?



Found in the  
*Survey File Data  
User's Guide*

## Example 3: Data are available for both Community and Facility components; LDS variable constructs are different

Assess LDS  
Variable Coding

- The **NAGIDIS** and **MENTHLTH** segments contain questions related to the beneficiary's ability to concentrate and include data from the Patient Health Questionnaire (PHQ-9)
- The **FACASMNT** segment contains questions related to the beneficiary's memory, recall, decision-making, capacity to understand, and overall mental status as assessed through the Brief Interview for Mental Status (BIMS)
- In order to properly restrict and analyze the data, information on demographics and weights for all beneficiaries is needed:
  - The **DEMO** segment contains information on interview type (**INT\_TYPE**) and beneficiary age (**H\_AGE**)

Found in the *LDS Codebooks*

## Example 3: Data are available for both Community and Facility components; LDS variable constructs are different

Recode Community variables to create binary indicator of cognitive impairment

Cognitively impaired (=1) if:

- beneficiary has "difficulty concentrating/remembering/deciding"

OR

- has "trouble concentrating for more than half the days or nearly every day" as per the Patient Health Questionnaire (PHQ-9)

Recode Facility variables to create binary indicator of cognitive impairment

Cognitively impaired (=1) if:

- beneficiary has issue with memory, recall, or decision-making

OR

- has a diagnosis related to cognitive impairment

OR

- is moderately to severely cognitively impaired as per the BIMS (0 – 12)

## Example 3: Data are available for both Community and Facility components; LDS variable constructs are different

Review documentation to properly restrict, weight, and analyze

- When merging with appropriate weights, all observations in the weights segment file should be preserved.
- When generating estimates:
  - Based on the research question, restrict data to all beneficiaries aged 65 and over and use the full sample weights for beneficiaries ever enrolled in Medicare during the survey period.
  - Use appropriate survey commands to generate preliminary estimates stratified by residence status and compare the rates to external benchmarks.

Found in the  
*Survey File Data  
User's Guide*

## Example 3: Data are available for both Community and Facility components; LDS variable constructs are different

### Prevalence of Cognitive Impairment among Medicare Beneficiaries, Aged 65 and Over, by Residence Status (2019)

Conduct Analysis  
and Document  
Limitations

Residence Status	COGIMP	Unweighted N	Weighted N (Standard Error)	Percent - % (Standard Error)
Community	Yes	1,811	6,591,384 (202,749)	13.1(0.4)
	No	9,838	43,613,800 (224,564)	86.9(0.4)
Facility	Yes	635	984,566 (49,226)	69.2(2.0)
	No	257	437,628 (34,939)	30.8(2.0)
Total	Yes	2,499	7,715,919 (194,985)	14.9(0.4)
	No	10,121	44,115,923 (218,499)	85.1(0.4)

SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, Survey File, 2019.

NOTE: Estimates are not presented for the residence status of "Both", although they are included in the total.

**Limitations:** Data collection methods are substantially different for the Community and Facility components. Cognitive status is self-reported or administered to a proxy for beneficiaries with Community component data. Cognitive status is merged during data processing from the MDS, abstracted from medical records, or reported by facility staff for beneficiaries with Facility component data.

**Please refer to Appendix, Example 3, found at the end of this tutorial, for detailed guidance on constructing the analytic file and the SAS code for producing estimates**

## Example 4: Data are available for both Community and Facility components, and the LDS variable constructs and coding are similar, but data collection and processing methods are different

**Research Question:** What was the average annual spending per beneficiary on inpatient and outpatient care for all Medicare beneficiaries enrolled in the program during 2019?

Define the Population

The research question applies to all Medicare beneficiaries, including those living in Facilities.

Found in the *Survey File Data User's Guide*

Assess Availability of Data

- The LDS segment **HISUMRY** has information for all beneficiaries on whether they had some or no enrollment in a group health plan, such as Medicare Advantage, during the year (**H\_GHPSW**)
- The LDS segment **PS** contains information for all beneficiaries on the adjusted sum of inpatient (**PAMTIP**) and outpatient events (**PAMTOP**)

Found in the *Cost Supplement Data User's Guide*

## Example 4: Data are available for both Community and Facility components, and the LDS variable constructs and coding are similar, but data collection and processing methods are different

### Assess LDS Variable Coding

- The **HISUMRY** segment shows that the variable **H\_GHPSW** captures information on group health participation for both the Community and Facility components
- Therefore, recoding of the LDS variable is not required for analysis

Found in the *LDS Codebooks*

### Review Documentation to Assess Underlying Differences in Data Collection and Processing

The Facility component does not collect costs associated with events occurring outside of the facility. FFS claims are used to fill this gap, however since claims for other payers are not available, any costs and utilization associated with other payers for events occurring outside of the facility are mostly unavailable.

Preliminary assessment of **H\_GHPSW** stratified by type of interview (**INT\_TYPE**) reveals that the estimates for Facility may be systematically underestimating the costs and utilization of beneficiaries with Facility component data.

Found in the *Cost Supplement Data User's Guide*

## Example 4: Data are available for both Community and Facility components, and the LDS variable constructs and coding are similar, but data collection and processing methods are different

**Group Health Participation by Residence Status (2019)**

Residence Status	MA participation – Unweighted N	MA participation – Weighted N (Standard Error)	MA Participation – % (Standard Error)	FFS only – Unweighted N	FFS only – Weighted N (Standard Error)	FFS only – % (Standard Error)
Community	3,927	23,580,326 (2,979,134)	39.1 (5.0)	5,065	36,771,688 (3,089,009)	60.9 (5.0)
Facility	242	549,216 (340,839)	30.1 (16.2)	565	1,277,364 (535,125)	69.9 (16.2)

SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, Survey File and Cost Supplement File, 2019.

**Analytic Note:** Whenever the Data User’s Guide points to a substantial difference in data collection and processing methods between Community and Facility components, it is helpful to stratify output for your variable of interest by residence status in order to quantify these differences and guide further analysis.

## Example 4: Data are available for both Community and Facility components, and the LDS variable constructs and coding are similar, but data collection and processing methods are different

Conduct Analysis  
and Document  
Limitations

**Average Annual Spending per Beneficiary on Inpatient and Outpatient Care for All Medicare Beneficiaries Enrolled in the Program (2019)**

Residence Status	Medicare Advantage / Fee-For-Service	Adjusted Sum for Inpatient Events – Weighted Mean (\$)	Adjusted Sum for Outpatient Events – Weighted Mean (\$)
Community	MA	2,600	1,402
	FFS	3,181	2,785
Facility	MA	979	268
	FFS	6,683	3,688

SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, Survey File and Cost Supplement File, 2019.

**Limitations:** Data collection and processing methods are substantially different for the Community and Facility components. For Community, all survey data on MA and FFS events are included in the data and FFS events are matched with claims data. The Facility component does not collect costs for services outside of the facility, therefore only FFS events from the claims data are available for these beneficiaries. Therefore, the data underrepresents cost and utilization for beneficiaries living in Facilities.

**Please refer to Appendix, Example 4, found at the end of this tutorial, for detailed guidance on constructing the analytic file and the SAS code for producing estimates**

# Section 5: Analytic Examples for Longitudinal Analysis Using Community and Facility Data



# Combining Data from the Community and Facility Components to Conduct Longitudinal Analysis

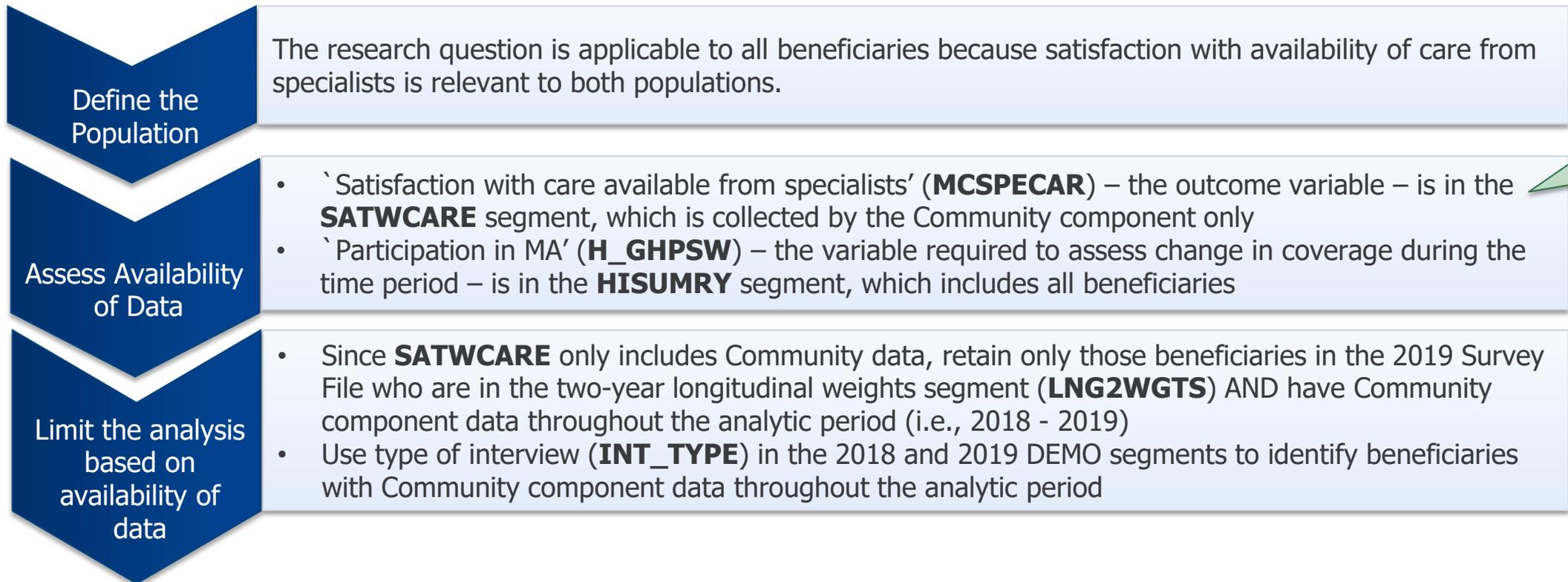
Data users may need to take additional steps when data from the Community and Facility components are combined to conduct longitudinal analyses. The analytic examples in the following slides illustrate some of the additional steps that data users will need to take to appropriately combine the data, conduct the analysis, and document potential limitations.

For beneficiaries who transition between living in the Community and living in a Facility during the analytic period, LDS variables related to the research question:

- may only be available for the Community or Facility component but not for both (**Example 5**); or
- may be available for both Community and Facility components, but data collection and/or processing methods are different (**Example 6**).

## Example 5: Data are available for the Community or Facility component but not for both

**Research Question:** Was there a change in satisfaction with availability of care from specialists for beneficiaries who switched from FFS to MA between 2018 and 2019 compared to those who were enrolled in MA during both years?



Found in the  
*Survey File  
Data User's  
Guides*

## Example 5: Data are available for the Community or Facility component but not for both

Conduct Analysis  
and Document  
Limitations

### Change in Satisfaction with Availability of Care from Specialists Among Beneficiaries Living in the Community – MA Beneficiaries who Switch from FFS

	Unweighted N	Weighted N	Satisfaction with Availability of Care from Specialists	
			Dissatisfied / Very Dissatisfied in 2018 Percent - % (Standard Error)	Dissatisfied / Very Dissatisfied in 2019 Percent - % (Standard Error)
<b>Beneficiaries enrolled in FFS in 2018 and MA in 2019</b>	294	2,175,954	1.9 (0.6)	2.5 (1.1)
<b>Beneficiaries enrolled in MA in 2018 and 2019</b>	3,241	18,594,977	6.1 (0.5)	5.5 (0.6)

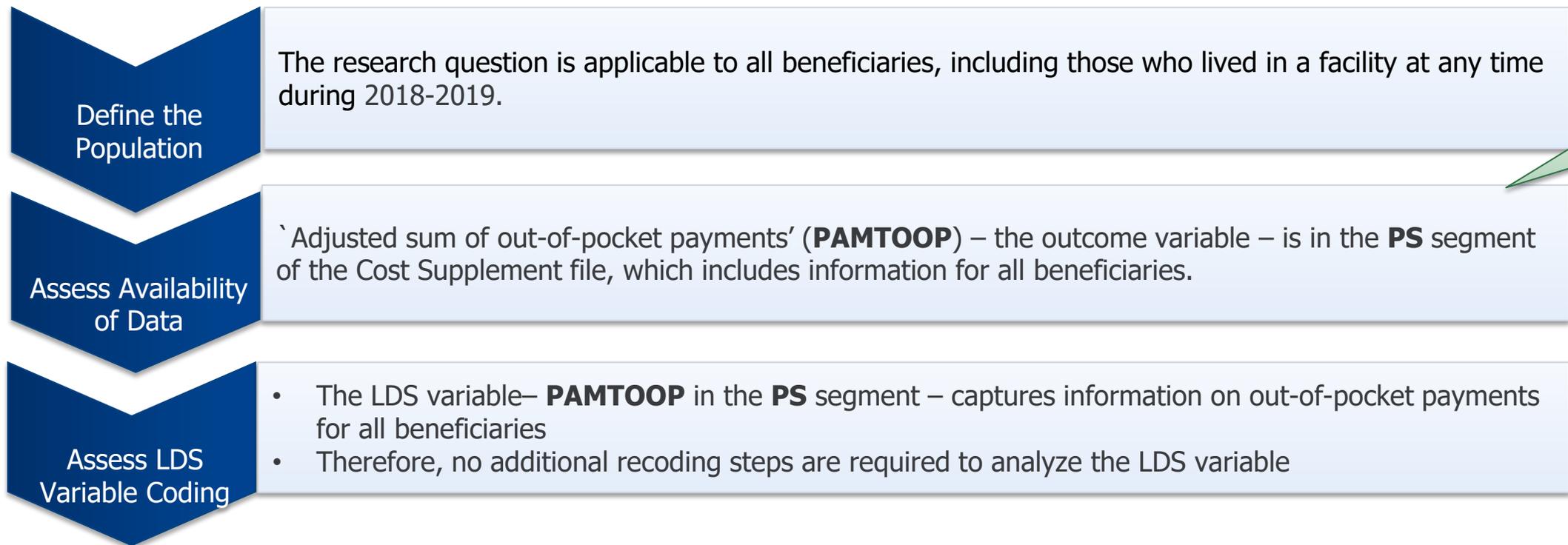
SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, Survey File, 2018 & 2019.

**Analytic Note:** Beneficiaries with Facility component data during 2018 - 2019 were excluded from the analysis.

Please refer to **Appendix Example 5**, found at the end of this tutorial, for detailed guidance on constructing the analytic file and the SAS code for producing estimates

## Example 6: Data are available for the Community and Facility component, but data collection and/or processing methods are different

**Research Question:** What is the median percent change in out-of-pocket payments for Medicare beneficiaries between 2018 and 2019?



## Example 6: Data are available for the Community and Facility component, but data collection and/or processing methods are different

Found in the  
*Survey File Data  
User's Guides*

Review  
Documentation to  
Assess Underlying  
Differences in Data  
Collection and  
Processing

The Community component collects out-of-pocket payments for all health care services rendered to the beneficiary and survey responses are reconciled with all FFS claims. The Facility component does not collect costs associated with events occurring outside of the facility so the only source of data on these events are FFS claims and imputed payers.

As expected, out-of-pocket payments are substantially higher for beneficiaries with Facility component data compared to those with Community component data. Therefore, stratifying the estimates based on whether beneficiaries remained in the same setting or had a change in their residence status is recommended.

### Median Out-of-Pocket Payment for Medicare Beneficiaries by Residence Status (2018-2019)

Residence Status	Median Out-of-Pocket Payment in 2018 - \$ (Standard Error)	Median Out-of-Pocket Payment in 2019 - \$ (Standard Error)
Community	1,008 (37.4)	997 (36.8)
Facility or Both	16,533 (2,092.4)	13,153 (927.3)

SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, Survey File and Cost Supplement File, 2018 & 2019.  
NOTE: Median income in 2018 and 2019 are shown in 2019 dollars.

## Example 6: Data are available for the Community and Facility component, but data collection and/or processing methods are different

Conduct Analysis and Document Limitations

**Median Percent Change in Out-of-Pocket Payments for Medicare Beneficiaries between 2018-2019**

Residence Status during 2018-2019	Weighted N (Unweighted N)	Median Out-of-Pocket Payment in 2018 - \$ (Standard Error)	Median Change in Out-of-Pocket Payment between 2018-2019 - \$ (Standard Error)	Median Percent Change in Out-of-Pocket Payment between 2018-2019 - % (Standard Error)
Remained in Community	54,774,804 (5,284)	996 (36.8)	47.6 (12.5)	9.1 (2.5)
Remained in Facility	1,550,155 (454)	16,422 (2,043.8)	-82.4 (215.9)	-1.2 (1.5)
Transitioned from Community to Facility	489,141 (58)	3,034 (794.1)	13,825 (5,049.0)	308.1 (244.1)

SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, Survey File and Cost Supplement File, 2018 & 2019.  
NOTE: Median out-of-pocket payment in 2018 is shown in 2019 dollars. Change in out-of-pocket payments and percent change in out-of-pocket payments are inflation-adjusted. Due to small sample sizes, beneficiaries who transitioned from the Facility to the Community component between 2018-2019 are excluded.

**Limitations:** Since the Facility component does not collect costs associated with events occurring outside of the facility and the only source of data for these events are FFS claims, out-of-pocket costs associated with such services are likely underestimated for beneficiaries in the Facility component data.

Please refer to **Appendix Example 6**, found at the end of this tutorial, for detailed guidance on constructing the analytic file and the SAS code for producing estimates

## Key Takeaways

- About **5-8% of the MCBS sample** are beneficiaries who live in a Facility or who alternate between living in the Community and living in a Facility
- Refer to the *Data User's Guides* to **ascertain whether differences in data collection and processing methods between** the Community and Facility components will influence analysis and interpretation
- Follow the **recommended steps for combining and weighting data** from the Community and Facility components
- Whenever there are substantial differences in data collection and processing methods between components, it is helpful to **stratify** output for your variable of interest **by interview type** in order to quantify these differences and guide further analysis
- Always **document** limitations!

# Thank you!

If you have any questions, please contact CMS at the following email address: [MCBS@cms.hhs.gov](mailto:MCBS@cms.hhs.gov).

The MCBS is authorized by section 1875 (42 USC 139511) of the Social Security Act and is conducted by NORC at the University of Chicago for the U.S. Department of Health and Human Services. The OMB Number for this survey is 0938-0568.

**Copyright Information:** This communication was produced, published, and disseminated at U.S. taxpayer expense. All material appearing in this report is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated.



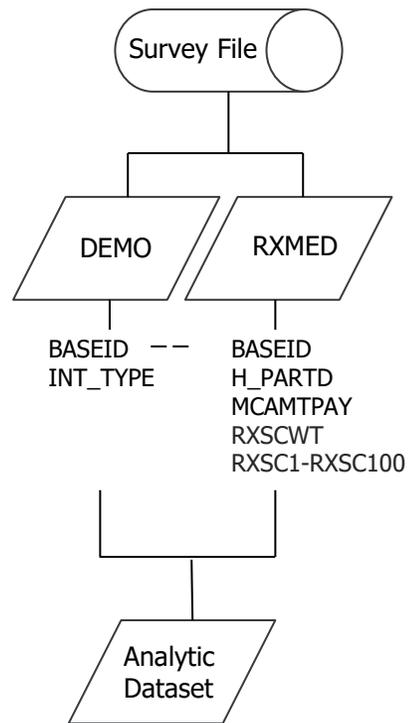
# Appendix: SAS® Code for Analytic Examples



# Example 1: Flow Chart for Merging Segments

**Data  
Segments**

**Variables &  
Topical Weights**



# Example 1: SAS code for Setup and Analysis

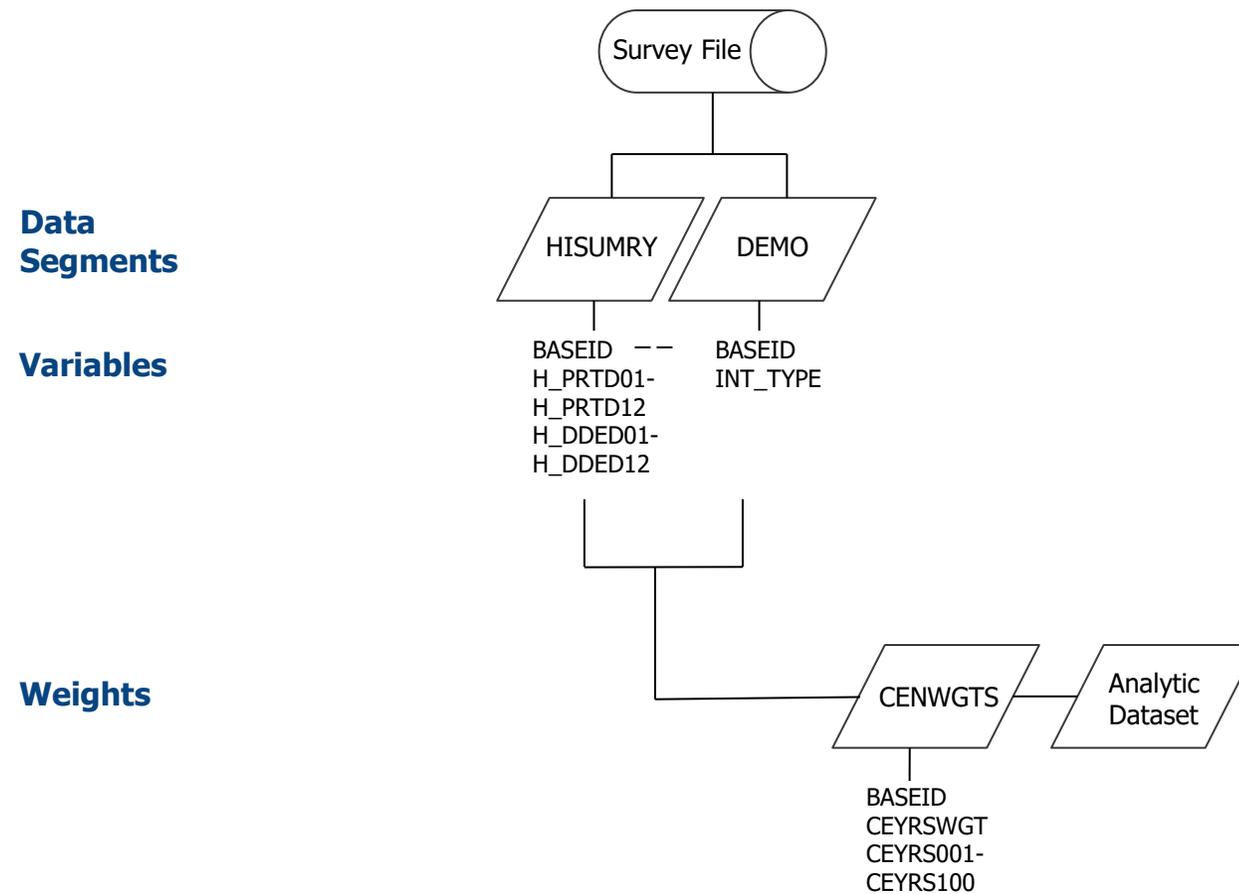
- Create dataset of beneficiaries with Community component data using variables related to satisfaction

```
data rx_satis;  
  merge surveyyy.DEMO (keep=BASEID INT_TYPE)  
        surveyyy.RXMED (keep=BASEID H_PARTD MCAMTPAY RXSCWT RXSC1-RXSC100);  
  by BASEID;  
run;
```

- Use topical weights and balanced repeated replication (BRR) method to produce estimates only for beneficiaries with Community component data in the weight file. Beneficiaries with Facility component data at any point during the survey period were excluded from the analysis.

```
proc surveyfreq data=rx_satis varmethod=brr (fay=.30);  
  tables MCAMTPAY;  
  weight RXSCWT;  
  repweights RXSC1-RXSC100;  
  where INT_TYPE="C" and RXSCWT ne . and H_PARTD = 1;  
run;
```

## Example 2: Flow Chart for Merging Segments



## Example 2: SAS code for Setup and Recode

- Create dataset of all beneficiaries using variables related to Part D deductible caps by month

```
data partd;  
  merge surveyyy.CENWGTS (in=a)  
        surveyyy.DEMO (keep=BASEID INT_TYPE)  
        surveyyy.HISUMRY (keep=BASEID H_PRTD01-H_PRTD12 H_DDED01-H_DDED12);  
  if a then output;  
  by BASEID;  
run;
```

- Create variable that averages monthly Part D deductible cap, based on MCBS Codebook

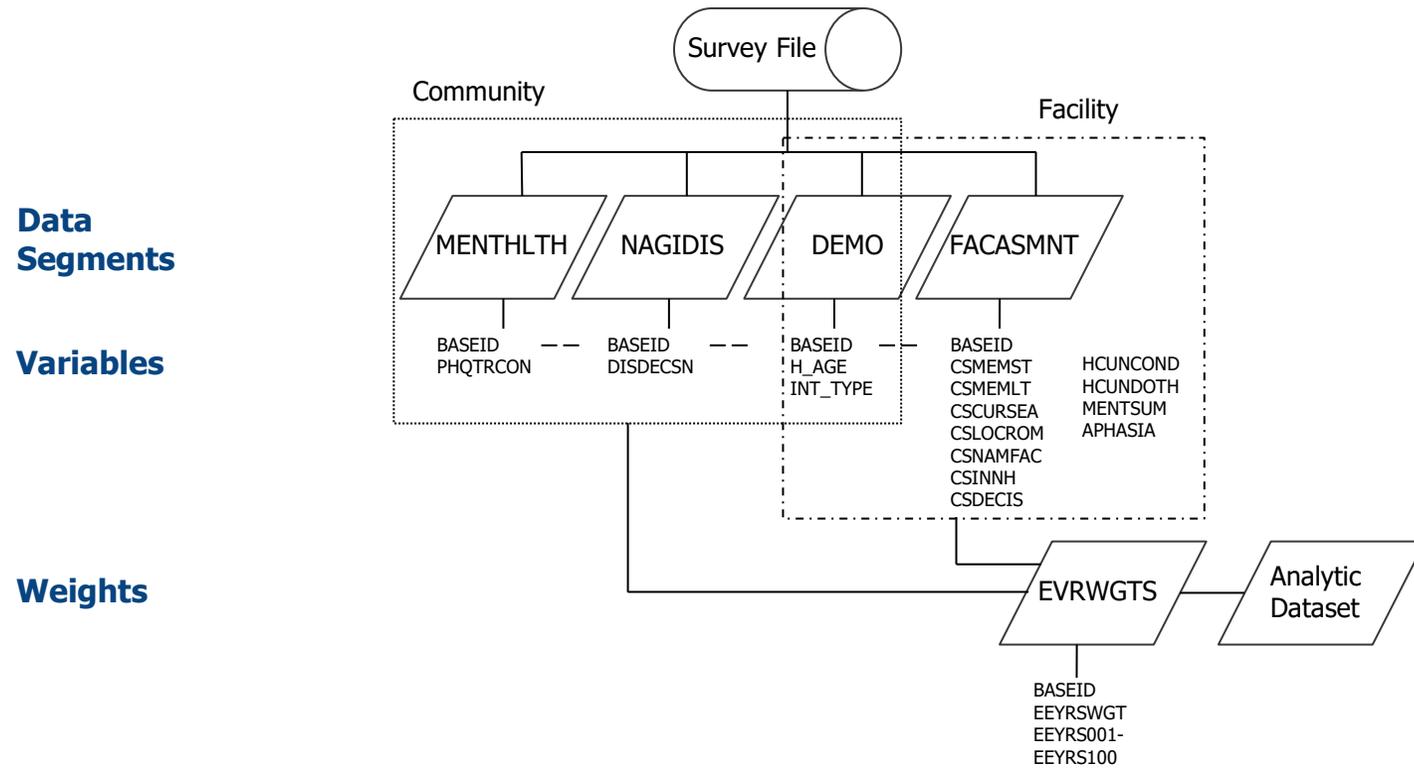
```
data partd_final;  
  set partd;  
  DEDUC=mean(of H_DDED01-H_DDED12);  
run;
```

## Example 2: SAS code for Analysis

- Use weights and BRR method to produce estimates of mean Part D deductible cap

```
proc surveymeans data=partd_final varmethod=brr (fay=.30) plots=none;  
  var DEDUC;  
  weight CEYRSWGT;  
  repweights CEYRS001-CEYRS100;  
run;
```

# Example 3: Flow Chart for Merging Segments



## Example 3: SAS code for Setup

- Create dataset of beneficiaries with Community component data using variables related to cognitive impairment

```
data comm;  
  merge surveyyy.DEMO (keep=BASEID H_AGE INT_TYPE in=a)  
        surveyyy.NAGIDIS (keep=BASEID DISDECSN in=b)  
        surveyyy.MENTHLTH (keep=BASEID PHQTRCON in=c);  
  if a and b or a and c then output;  
  by BASEID;  
run;
```

- Create dataset of beneficiaries with Facility component data using variables related to cognitive impairment

```
data fact;  
  merge surveyyy.DEMO (keep=BASEID H_AGE INT_TYPE in=a)  
        surveyyy.FACASMNT (keep=BASEID CSMEMST CSMEMLT CSCURSEA CSLOCROM CSNAMFAC CSINNH  
                           HCUNCOND HCUNDOTH CSDECIS MENTSUM APHASIA in=b);  
  if a and b then output;  
  by BASEID;  
run;
```

## Example 3: SAS code for Recode

- Recode Community variables as a binary indicator of cognitive impairment (Yes=1; No=0) based on responses to survey or PHQ-9

```
data comm_recode;
  set comm;
  if DISDECSN = 1 or PHQTRCON in (3:4) then COGIMP = 1;
  else if DISDECSN = 2 or PHQTRCON in (1:2) then COGIMP = 0;
run;
```

- Recode Facility variables as a binary indicator of cognitive impairment (Yes=1; No=0) based on condition, diagnosis, or BIMS scale

```
data fac1_recode;
  set fac1;
  COGIMP_ABILITY = 0;
  if CSMEMST = 1 then COGIMP_ABILITY + 1; if CSMEMLT = 1 then COGIMP_ABILITY + 1;
  if CSCURSEA = 0 then COGIMP_ABILITY + 1; if CSLOCROM = 0 then COGIMP_ABILITY + 1;
  if CSNAMFAC = 0 then COGIMP_ABILITY + 1; if CSINNH = 0 then COGIMP_ABILITY + 1;
  if HCUNCOND in (2,3) then COGIMP_ABILITY + 1; if HCUNDOTH in (2,3) then COGIMP_ABILITY + 1;
  if CSDECIS in (2,3) then COGIMP_ABILITY + 1;
  COGIMP_DX = 0;
  if APHASIA = 1 then COGIMP_DX + 1;
  if MENTSUM in (0:12) or COGIMP_ABILITY ge 1 or COGIMP_DX ge 1 then COGIMP = 1;
  else COGIMP = 0;
run;
```

## Example 3: SAS code for Merge and Analysis

- Merge Community and Facility data with weights file, preserving only BASEIDs that exist in the weights file

```
data ment_merged;  
  merge surveyyy.EVRWGTS (in=a)  
        work.comm_recode  
        work.facl_recode;  
  if a then output;  
  by BASEID;  
run;
```

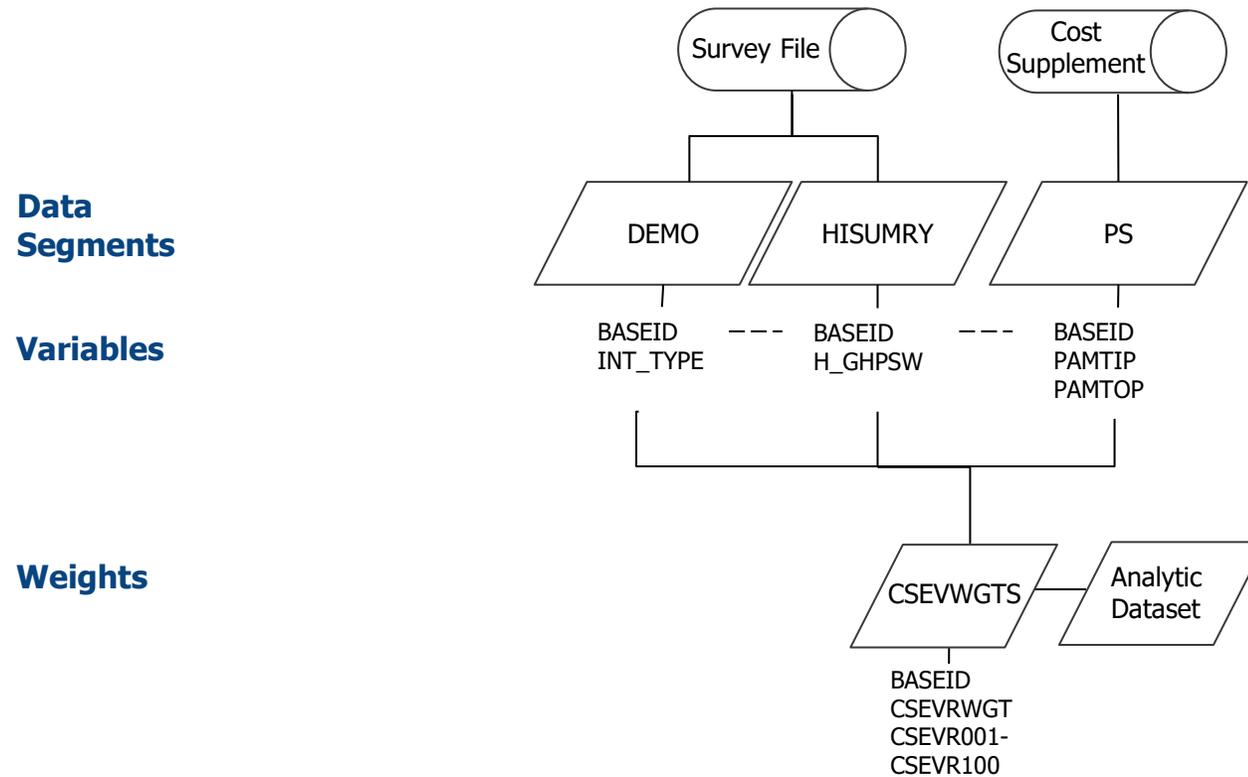
- Restrict file to beneficiaries aged 65 and over

```
data ment_final;  
  set ment_merged;  
  where H_AGE ge 65;  
run;
```

- Use weights and BRR method to produce COGIMP prevalence estimates stratified by INT\_TYPE

```
proc surveyfreq data=ment_final varmethod=brr (fay=.30);  
  tables INT_TYPE*COGIMP / row;  
  weight EEYRSWGT;  
  repweights EEYRS001-EEYRS100;  
run;
```

# Example 4: Flow Chart for Merging Segments



## Example 4: SAS code for Setup and Recode

- Create dataset for all beneficiaries using variables related to group health plan participation and inpatient and outpatient cost events

```
data cost_merged;
  merge surveyyy.DEMO (keep=BASEID INT_TYPE)
        surveyyy.HISUMRY (keep=BASEID H_GHPSW)
        costyy.PS (keep=BASEID PAMTIP PAMTOP);
  by BASEID;
run;
```

- Merge dataset with weights for Cost Supplement File

```
data cost_final;
  merge costyy.CSEVWGTS (in=a)
        work.cost_merged;
  if a then output;
  by BASEID;
run;
```

## Example 4: SAS code for Analysis

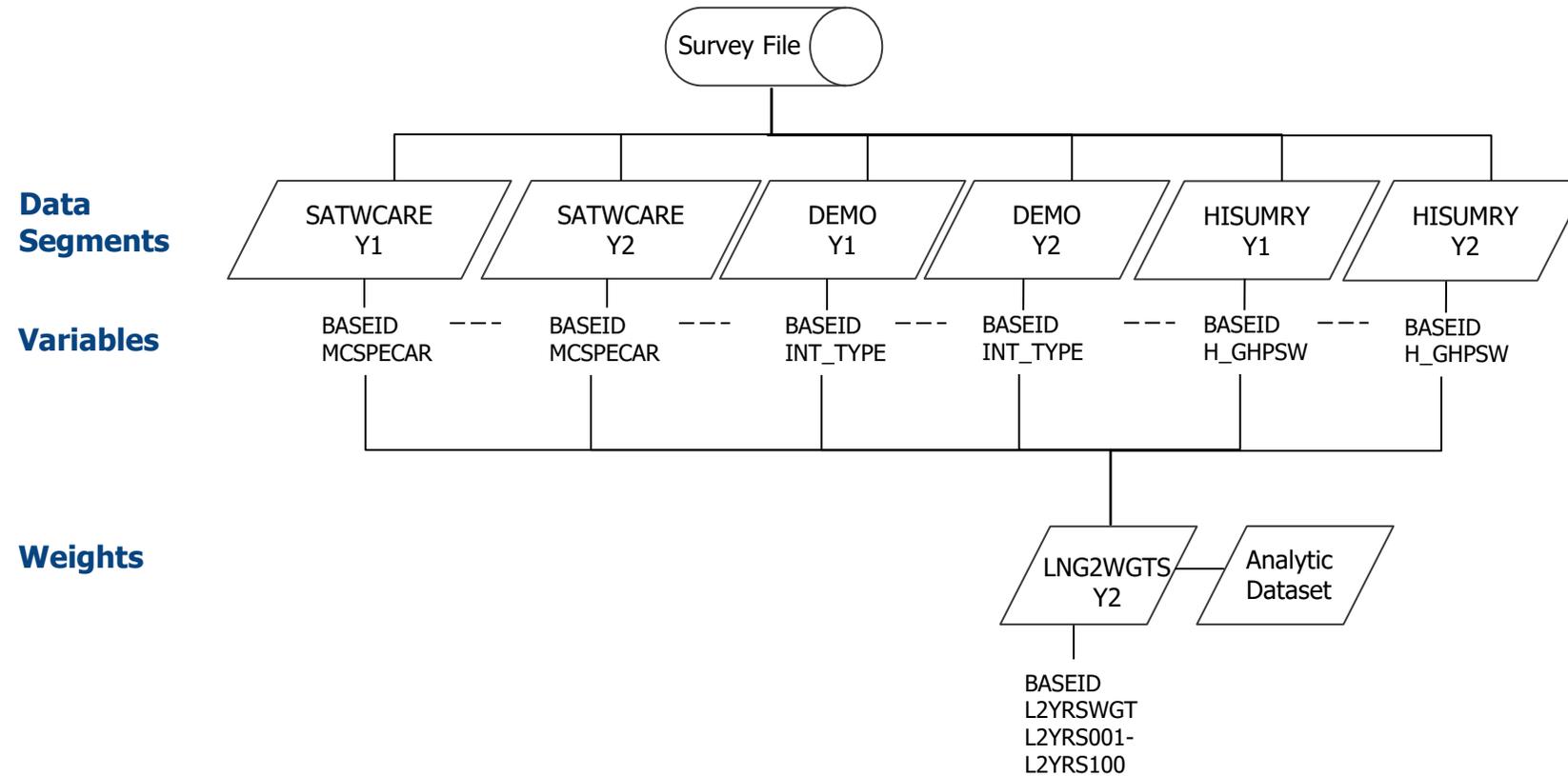
- Make a preliminary assessment of the H\_GHPSW variable stratified by interview type

```
proc surveyfreq data=cost_final;  
  tables INT_TYPE*H_GHPSW / row;  
  weight CSEVRWGT;  
  repweights CSEVR001-CSEVR100;  
run;
```

- Use weights and BRR method to produce estimates stratified by interview type and group health participation

```
proc sort data = cost_final;  
  by INT_TYPE H_GHPSW;  
run;  
  
proc surveymeans data=cost_final varmethod=brr (fay=.30) plots=none;  
  var PAMTIP PAMTOP;  
  weight CSEVRWGT;  
  repweights CSEVR001-CSEVR100;  
  by INT_TYPE H_GHPSW;  
run;
```

# Example 5: Flow Chart for Merging Segments



## Example 5: SAS code for Setup

Create longitudinal analytic dataset by merging the Survey File segments from two years (in this example, y1=2018 and y2=2019) to the two-year backward longitudinal weights segment from year 2. Subset the analytic file to beneficiaries belonging to the Community component in both years, enrolled in MA during year 2, and included in the two-year backward longitudinal weights file from year 2.

```
data lex5;
  merge surveyy2.Lng2wgts (in=a)
        surveyy2.SATWCARE (keep=BASEID MCSPECAR rename=(MCSPECAR=MCSPECARY2))
        surveyy1.SATWCARE (keep=BASEID MCSPECAR rename=(MCSPECAR=MCSPECARY1))
        surveyy2.DEMO (keep=BASEID INT_TYPE rename=(INT_TYPE=INT_TYPEY2))
        surveyy1.DEMO (keep=BASEID INT_TYPE rename=(INT_TYPE=INT_TYPEY1))
        surveyy2.HISUMRY (keep = BASEID H_GHPSW rename=(H_GHPSW=H_GHPSWY2))
        surveyy1.HISUMRY (keep = BASEID H_GHPSW rename=(H_GHPSW=H_GHPSWY1));
  by BASEID;
  if a and INT_TYPEY1 = "C" and INT_TYPEY2 = "C" and MCSPECARY1 in (1,2,3,4,5) and
     MCSPECARY2 in (1,2,3,4,5) and H_GHPSWY2 = 1
  then output;
run;
```

## Example 5: SAS code for Recode and Analysis

- Recode `Satisfaction with care available from specialists' (MCSPECAR) to create a dichotomous variable to identify beneficiaries who were dissatisfied or very dissatisfied with care available from specialists (Yes=1; No=0).

```
data lex5;
  set lex5;
  dsatspcy1 = (MCSPECARY1 in (3,4));
  dsatspcy2 = (MCSPECARY2 in (3,4));
run;
```

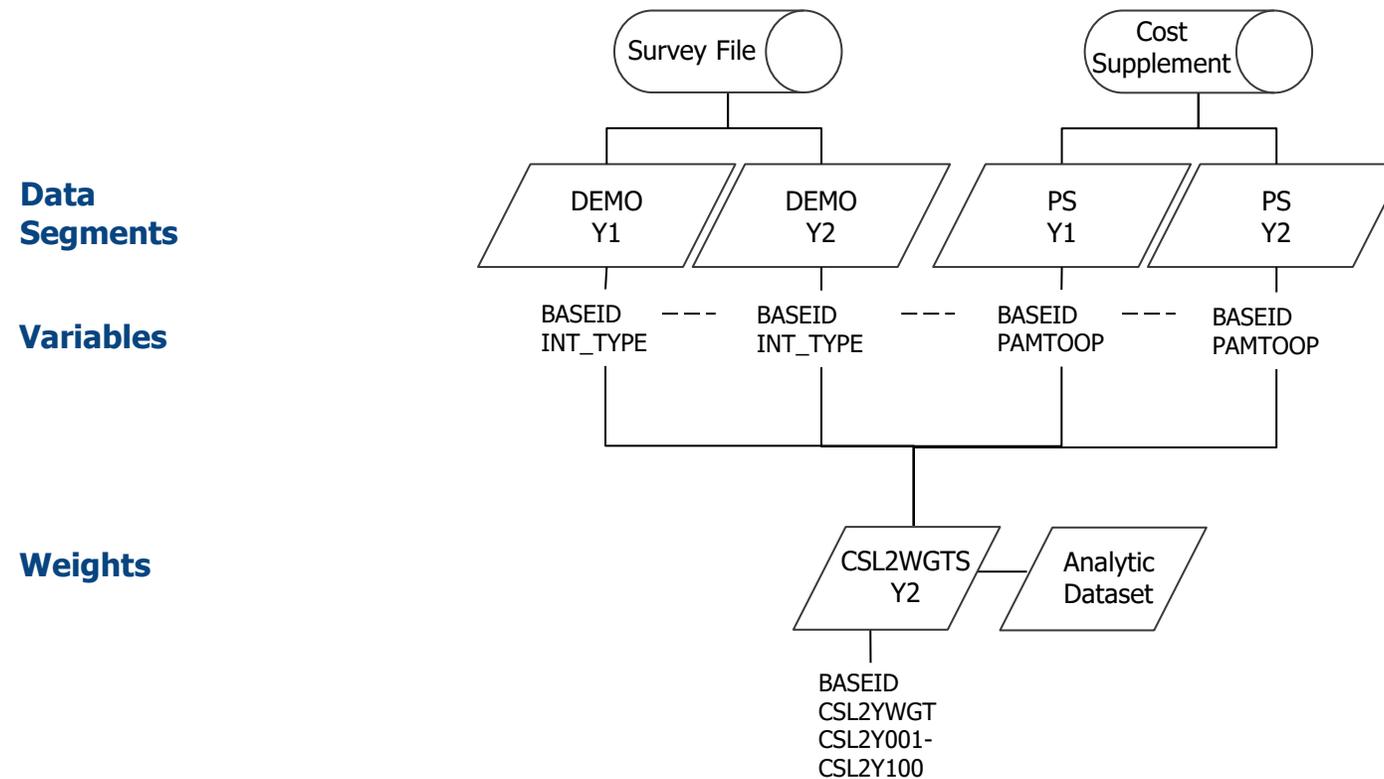
- Utilize two-year backward longitudinal weights and BRR variance estimation method to estimate the proportion of beneficiaries who are dissatisfied or very dissatisfied with care available from specialists during year 1 and year 2. Stratify the estimates by beneficiaries who transition into MA from FFS between year 1 and year 2 and beneficiaries who are enrolled in MA during year 1 and year 2 .

```
proc surveymeans data=lex5 varmethod=brr (fay=.30) plots=none;
  var dsatspcy1 dsatspcy2;
  domain H_GHPSWy1;
  weight L2YRSWGT;
  repweights L2YRS001-L2YRS100;
run;
```

- Output weighted and unweighted N values.

```
proc surveyfreq data=lex5 varmethod=brr (fay=.30);
  table H_GHPSWy1;
  weight L2YRSWGT;
  repweights L2YRS001-L2YRS100;
run;
```

# Example 6: Flow Chart for Merging Segments



## Example 6: SAS code for Setup and Preliminary Assessment

- Create longitudinal analytic dataset by merging the Survey File and Cost Supplement segments from two years (in this example, y1=2018 and y2=2019) to the two-year, cost-supplement, backward longitudinal weights file from year 2. Subset the analytic file to beneficiaries included in the year 2 two-year backward longitudinal weights file.

```
data lex6;
  merge csy2.csl2wgts (in=a)
        sfy2.DEMO (keep=BASEID INT_TYPE rename=(INT_TYPE=INT_TYPEY2) )
        sfy1.DEMO (keep=BASEID INT_TYPE rename=(INT_TYPE=INT_TYPEY1) )
        csy2.ps (keep=BASEID PAMTOOP rename=(PAMTOOP = OOPY2))
        csy1.ps (keep=BASEID PAMTOOP rename=(PAMTOOP = OOPY1));
  by BASEID;
  if a then output;
run;
```

- Calculate change in out-of-pocket payments and percent change in out-of-pocket payments between 2018-2019 for each beneficiary in the sample. Inflation-adjust 2018 dollars to 2019 dollars using a CPI conversion factor.

```
data lex6;
  set lex6;
  OOPY1a=OOPY1*(251.107/245.120);
  CHNG_OOP= OOPY2 - OOPY1a;
  if OOPY1a > 0 then PCHNG_OOP= (CHNG_OOP/OOPY1a) * 100;
  COMMY1 = ( INT_TYPEY1 in ("C"));
  COMMY2 = ( INT_TYPEY2 in ("C"));
run;
```

## Example 6: SAS code for Preliminary Assessment and Analysis

- Conduct preliminary assessment of the extent to which out-of-pocket payments differ across Community and Facility components

```
proc surveymeans data=lex6 varmethod=brr (fay=.30) plots=none median;  
  var OOPY1a OOPY2;  
  domain COMMY1 COMMY2;  
  weight CSL2YWGT;  
  repweights CSL2Y001-CSL2Y100;  
run;
```

- Utilize two-year backward longitudinal weights and BRR variance estimation method to estimate the median change in out-of-pocket payments for Medicare beneficiaries in the Community and Facility components during year 1 and year 2.

```
proc surveymeans data=lex6 varmethod=brr (fay=.30) plots=none mean median;  
  var OOPY1a CHNG_OOP PCHNG_OOP;  
  domain COMMY1 * COMMY2;  
  weight CSL2YWGT;  
  repweights CSL2Y001-CSL2Y100;  
run;
```

- Output weighted and unweighted N values.

```
proc surveyfreq data=lex6 varmethod=brr (fay=.30);  
  table COMMY1 * COMMY2;  
  weight CSL2YWGT;  
  repweights CSL2Y001-CSL2Y100;  
run;
```