

Where Are My Agents? Automating Cohort Building with a Dynamic Cohort Builder

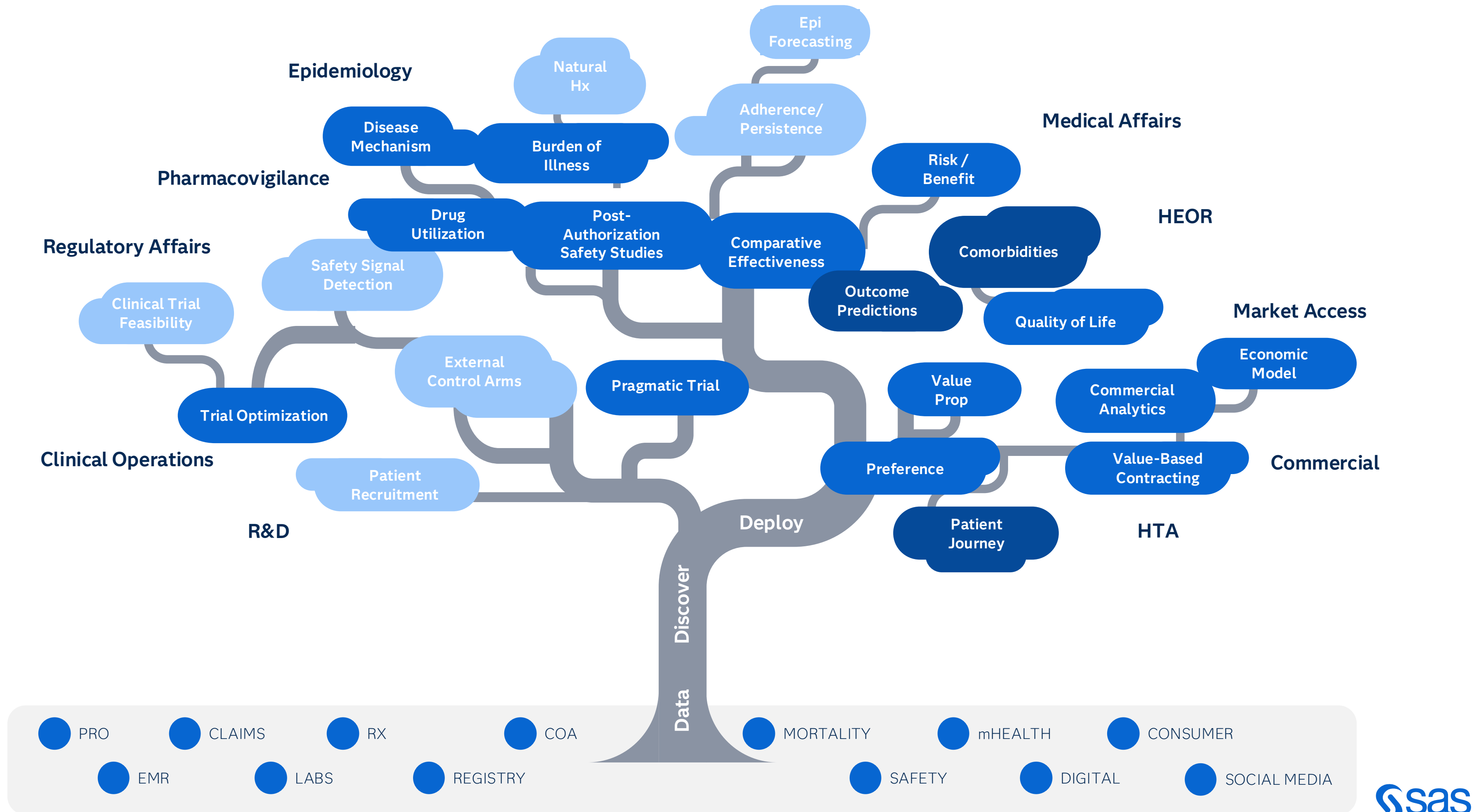
Sherrine Eid, Mary Dolegowski, Sundaresk Sankaran

PhUSE US Connect 2026



The AIM

Properly and accurately defining patient cohorts to affect validity and reliability



US FDA Regulation and Guidance on RWE

CDER Center for
Real-World Evidence
Innovation (CCRI)



In December 2024, CDER established
a new Center for Real-World Evidence Innovation

Date Published	Title
August 2017	<u>Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices</u>
July 2018	<u>Use of Electronic Health Records in Clinical Investigations</u>
December 2018	<u>Framework for FDA’s Real-World Evidence Program</u>
September 2021	<u>Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products</u>
October 2021	<u>Data Standards for Drug and Biological Product Submissions Containing Real-World Data</u>
November 2021	<u>Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products</u>
September 2022	<u>Submitting Documents Utilizing Real-World Data and Real-World Evidence to FDA for Drugs and Biologics</u>
February 2023	<u>Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products</u>
August 2023	<u>Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products</u>

Mastering the RWE Workflow



Data Management

- Data Ingestion
- Data Quality
- Data Prep
- Dynamic



Cohort Building

- Inclusion
- Exclusion
- Continuous
- Temporal



Analyses

- Causality
- Comparative Effectiveness
- Predictive
- Advanced
- Etc.

Generating Cohorts



Code

SAS, SQL, database, etc.

Custom Steps, Flows & Agents

SAS® Health Copilot
Gen AI

**Expert
Consulting**



The screenshot displays the SAS Health Copilot interface. On the left, a code editor shows SQL code for generating cohorts. On the right, a flow diagram illustrates the process from input to output.

```
proc fedsql sessref = mysession;

create table _or_filtered_Members as
select CONSISTENT_MEMBER_ID
from &rx_data.
where CODE_VALUE in &RX_filter.
UNION
select CONSISTENT_MEMBER_ID
from &code_data.
where CODE_VALUE in &code_filter.
;

create table _and_Filtered_Member_IDs as
select DISTINCT _RX_Members.CONSTISTENT_MEMBER_ID
from
  (select CONSISTENT_MEMBER_ID from &rx_data. where
   as _RX_Members
INNER JOIN
  (select CONSISTENT_MEMBER_ID from &code_data. where
   as _Code_Members
ON _RX_Members.CONSTISTENT_MEMBER_ID = _Code_Members.
;

create table Filtered_RX as
select RX.*
from &rx_data as RX
INNER JOIN _and_Filtered_Member_IDs as F
ON RX.CONSTISTENT_MEMBER_ID = F.CONSTISTENT_MEMBER_ID
;

create table Filtered_Service_Code as
select code.*
from &code_data as code
INNER JOIN _and_Filtered_Member_IDs as F
ON code.CONSTISTENT_MEMBER_ID = F.CONSTISTENT_MEMBER_ID
;

run;
```

The flow diagram on the right shows the following steps:

- RWE_SERVICE_CODES** (Input)
- Period and Frequency...** (Intermediate Step)
- Summary** (Output)
- Output** (Final Result)

Below the flow diagram, a list of generated cohort names is visible:

- Female Patients between 18 to 45 ye...
- Male Patients with Creatinine value g...
- YESTERDAY**
- List of Patients with Chronic Kidney D...
- Male Patients with Creatinine value g...
- Female Patients between 18 to 45 ye...
- How to trick an LLM
- Lymphoma Treatment Plan
- Female Patients between 18 to 45 ye...

SAS[®] Viya

The SAS[®] Platform for Life Sciences.



Automate

The mundane

Extend

Access to data &
analytics

Manage

Analytic assets

Eliminate

Redundancy

Custom Steps

Building Custom Steps

GUI Designer backed by SAS Code

Step 4 and 5.step x +

Designer Prompt UI Program Preview

Control Library

+ - :

Page 1

Filter controls

COMMON

- Check Box
- Color Picker
- Date and T...
- Drop-dow...
- File or Fol...
- Link
- List
- Numeric St...
- Option Table
- Radio Butt...
- Section
- Text

Input*

Enrollment Table*

Index Date*

Jun 1, 2008

Coverage Months*

6

Step ID/Name*

Step4

Output Table*

Properties: Input Table

ID:*

input_data

Label:

Input

☒ Required

☐ Make control read-only

☐ Hide control at runtime

Default library:

Default table:

Placeholder text:

Styles

Indent level:*

0

> Dependencies

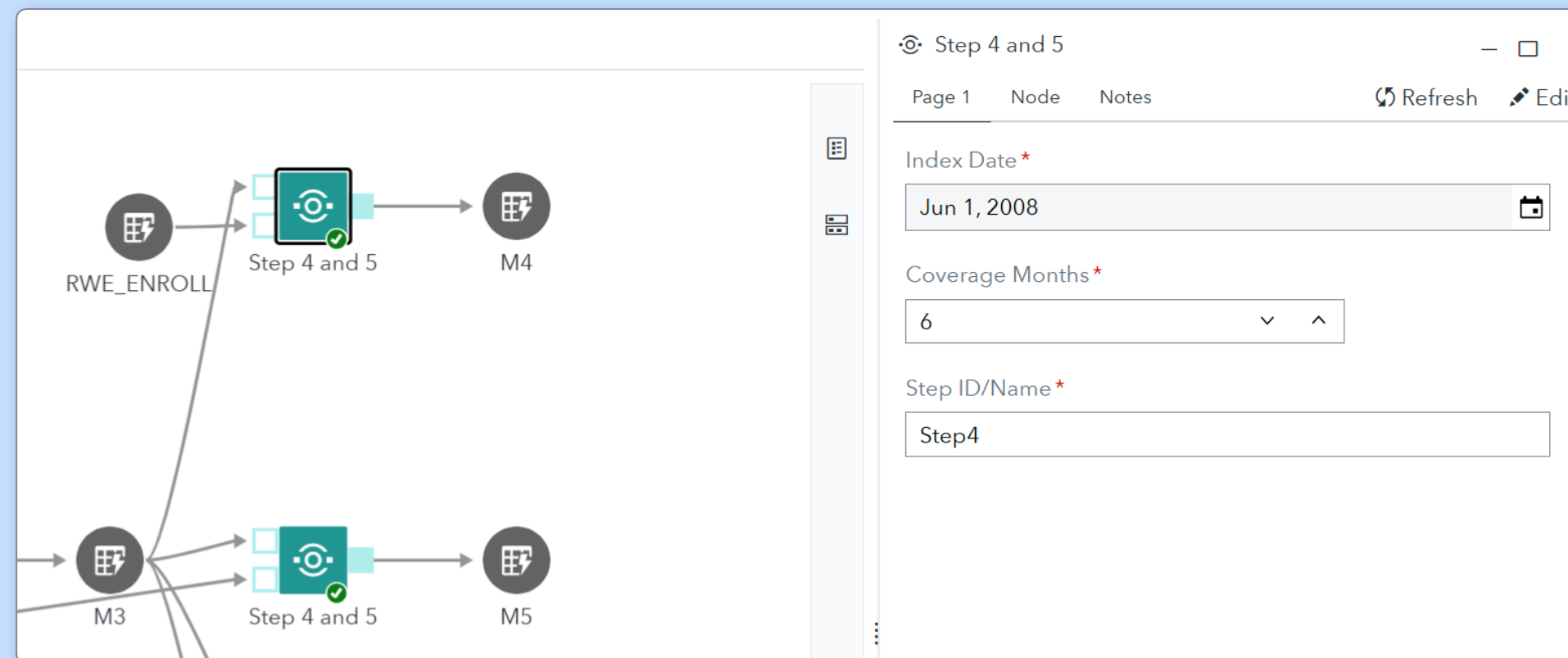
Step 4 and 5.step x +

Designer Prompt UI Program Preview

```
1  /* SAS templated code goes here */
2
3  proc fedsql sessref = mysession;
4
5      create table _&step_name. as
6      select distinct CONSISTENT_MEMBER_ID, COVERAGE_TYPE_CD
7      from (
8          select e.*
9          from
10             (select * from &enrollment_data.) as e
11             INNER JOIN
12             (select * from &input_data.) as m
13             ON e.CONSISTENT_MEMBER_ID = m.CONSISTENT_MEMBER_ID
14         ) as _step4_base
15         where (enrollment_end_dt - %sysfunc(INPUTN(&index_date., yymmdd10.)) >= %eval(&coverage_mc
16         ;
17
18     run;
19     quit;
20
21  data &step_output. (replace=yes partition=(CONSISTENT_MEMBER_ID)) / sessref = mysession;
22      set casuser._&step_name.;
23      by CONSISTENT_MEMBER_ID;
24      attrib test_MED test_DP length=$ 1;
25      retain test_MED test_DP;
26      if first.CONSISTENT_MEMBER_ID then call missing(test_MED, test_DP);
27      select(COVERAGE_TYPE_CD);
28          when('MED') test_MED = 'Y';
29          when('DP') test_DP = 'Y';
30          otherwise;
31      end;
32      if missing(test_MED) or missing(test_DP) then delete;
33      keep CONSISTENT_MEMBER_ID;
34  run;
```

Interacting with Custom Steps

In a flow or in a tab

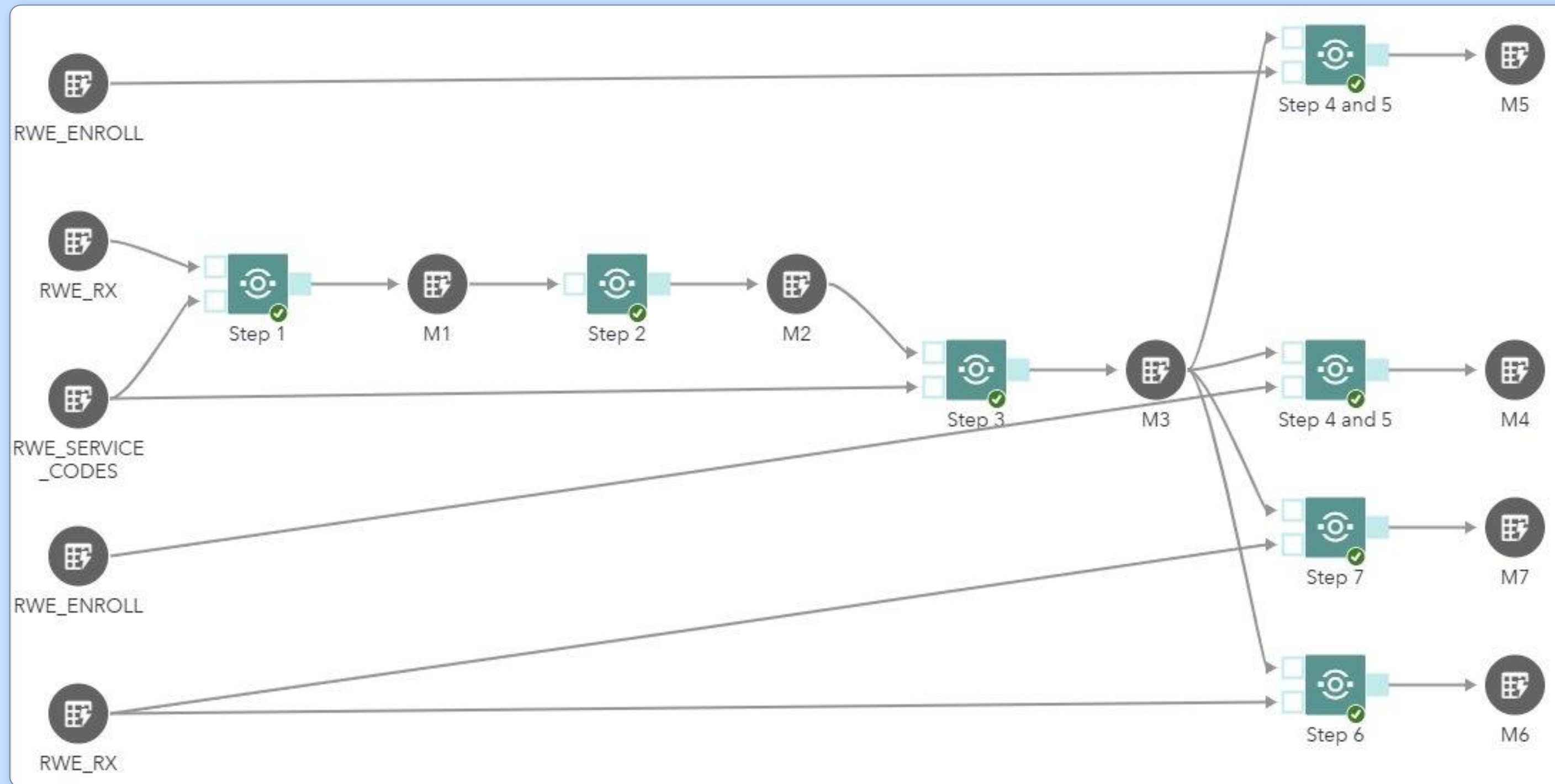


The screenshot shows the configuration panel for 'Step 4 and 5' in a tabbed view. The panel includes a 'Run' button, a 'Cancel' button, and two circular arrows for refreshing or undoing. The configuration fields are as follows:

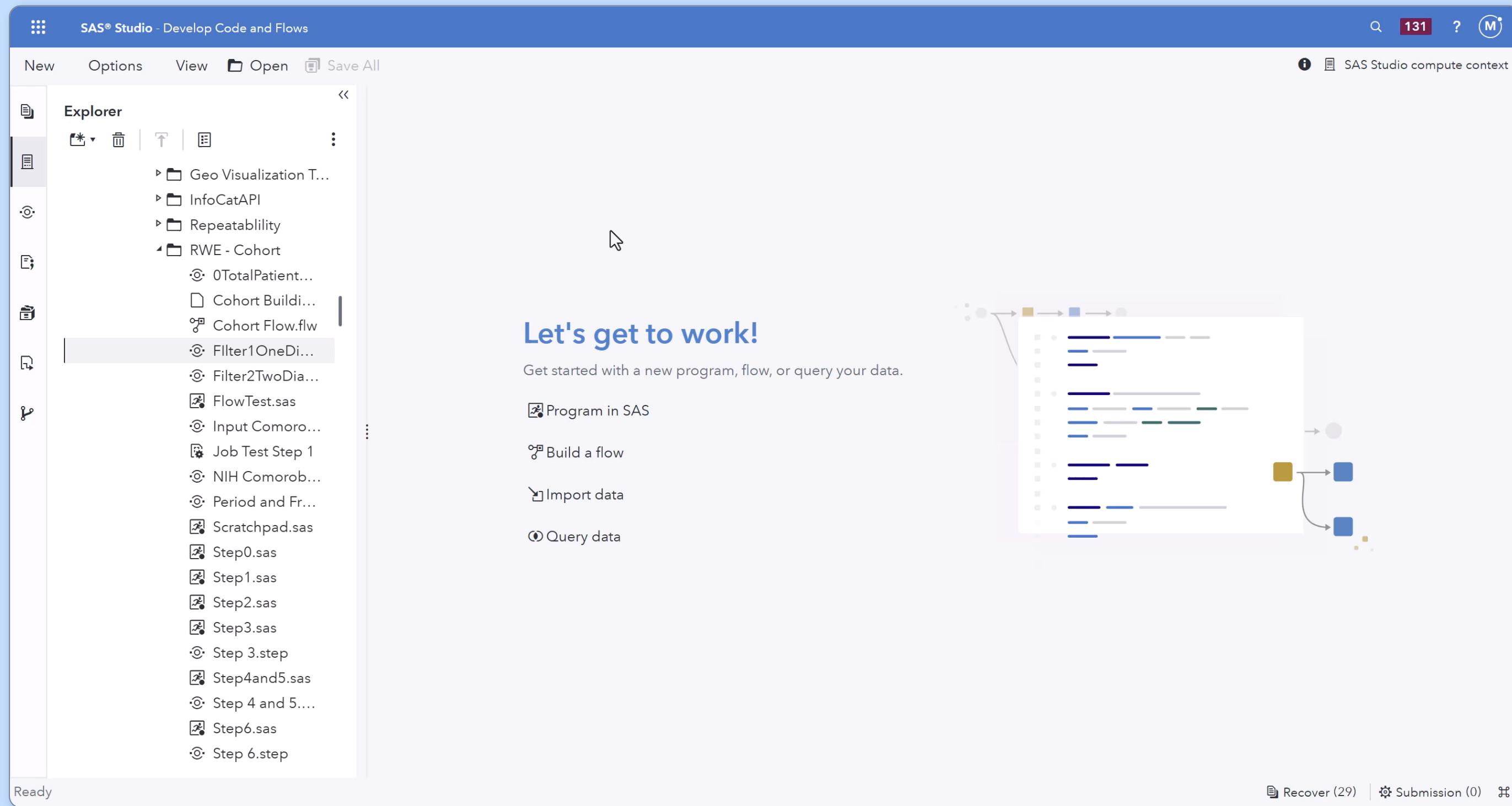
- Input ***: CASUSER.M3 (with a dropdown arrow and a help icon)
- Enrollment Table ***: RWE800K.RWE_ENROLL (with a dropdown arrow and a help icon)
- Index Date ***: Jun 1, 2008 (with a calendar icon)
- Coverage Months ***: 6 (with up/down arrows)
- Step ID/Name ***: Step4
- Output Table ***: CASUSER.M4 (with a dropdown arrow and a help icon)

Flows

Chain Repeatable Processes



Custom Steps



Agentic Building

MCP Tool Server

SAS® Retrieval Agent Manager - Generative AI for Enterprise Knowledge Base
Advanced view: ● ○ ☰ ? S

- Dashboard
- Sources
- Collections
- User Eval Tests
- Retrieval Settings
- Agents
- Automation
- Chat
- MCP Tools
- Code Templates

MCP Tool Servers ✚ ✎ ⌂ 🔍 🗑️

Name	Template	Server Status	Ping Status
Charts local	Charts acuvity (v2)	○ Starting	⚡ Disconnected
Financial RAG tools	RAG tools (v2)	✓ Running	✓ Connected
PKR_MCP_Server	PKR_Code_Template (v22)	■ Stopped	⚡ Disconnected
RWE_MCP	RWE_MCP_CodeBase (v5)	✓ Running	✓ Connected
Test_Tool_Server	Test_Code_Template (v58)	■ Stopped	⚡ Disconnected
Financial DB - ram-beta	GenAI Toolbox Prebuild (v2)	✓ Running	✓ Connected
Finance tools	Finance tools (v2)	✓ Running	✓ Connected

Name	Description
> step0	Given a member table and a desired summary table, output likely SAS code that calculates the total number of distinct pati...

Given a member table and a desired summary table, output likely SAS code that calculates the total number of distinct patients based on CONSISTENT_MEMBER_ID

Accessing the RWE Agent

The screenshot displays the SAS Retrieval Agent Manager interface, titled "SAS® Retrieval Agent Manager - Generative AI for Enterprise Knowledge Base". The interface is divided into a left sidebar and a main chat area.

Left Sidebar:

- Dashboard
- Sources
- Collections
- User Eval Tests
- Retrieval Settings
- Agents
- Automation
- Chat** (selected)
- MCP Tools
- Code Templates

Chat Area:

The chat area is titled "Chat" and shows a list of "Previous Chats" on the left, each starting with "For my RWE_2025 member ...". The main chat window is titled "RWE_SAS_Code_Agent".

User Input:

For my RWE_2025 member table, I want to find out the number of patients.

Agent Response:

Your code is **basically correct** for counting distinct patients from a **CAS table** using FedSQL 🙌. Below are a few **minor clarifications and best-practice tweaks**, plus an optional alternative depending on what you want.

✓ Your current code (validated)

```
/* Step 0 - count total patients in database.
   RWE data are stored as CAS tables, so FedSQL is required. */

cas mysession;
caslib _all_ assign;

proc fedsql sessref=mysession;
  create table casuser.summary as
  select
    count(distinct CONSISTENT_MEMBER_ID) as Total_Members
  from RWE_2025;
quit;

cas mysession terminate;
```

✓ This will:

- Start a CAS session
- Query the CAS table RWE_2025
- Count **distinct** CONSISTENT_MEMBER_ID
- Save the result to casuser.summary
- Properly terminate the CAS session

✓ Things to double-check

1. CASLIB of RWE_2025

Enter a query...

AI-generated queries and resulting data sets may be incorrect. Confirm the query captures your desired conditions.

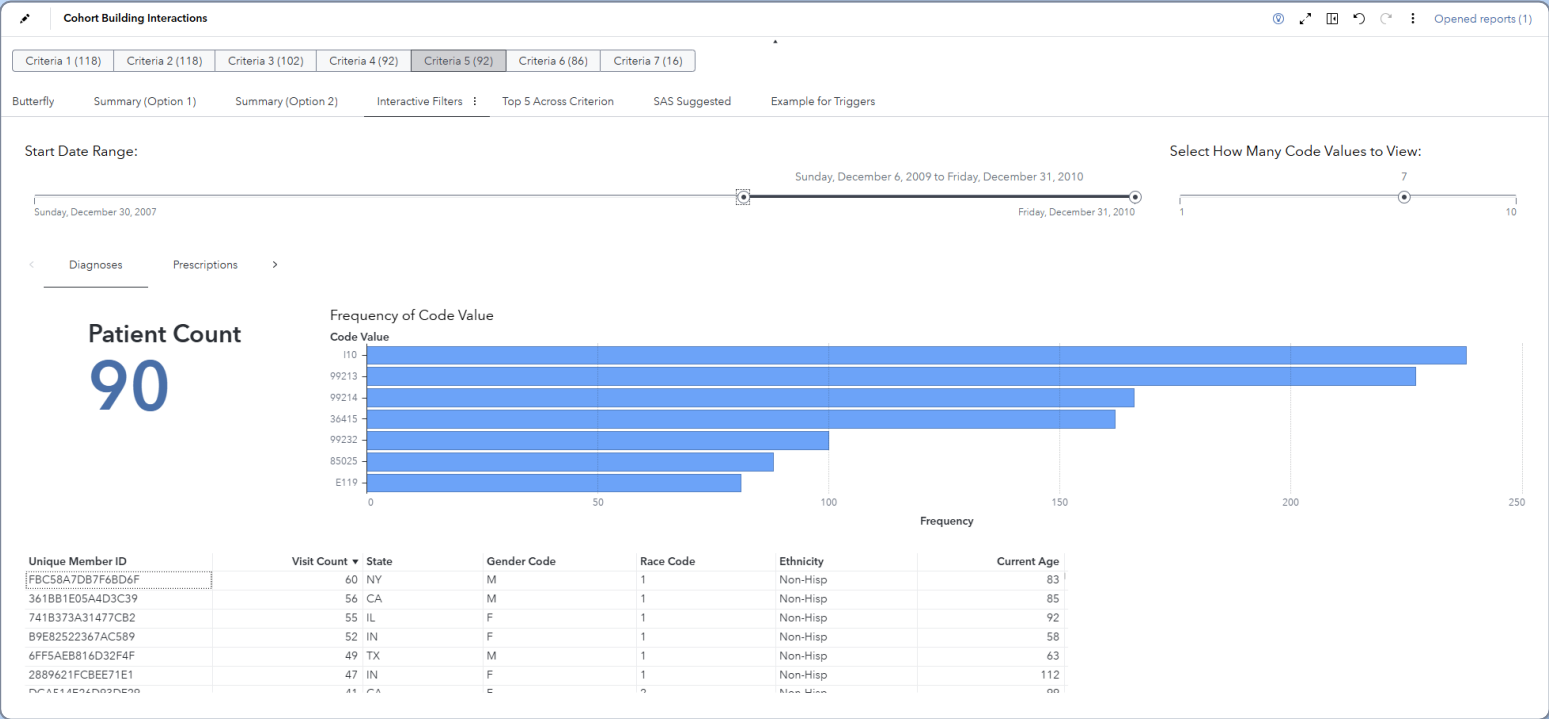
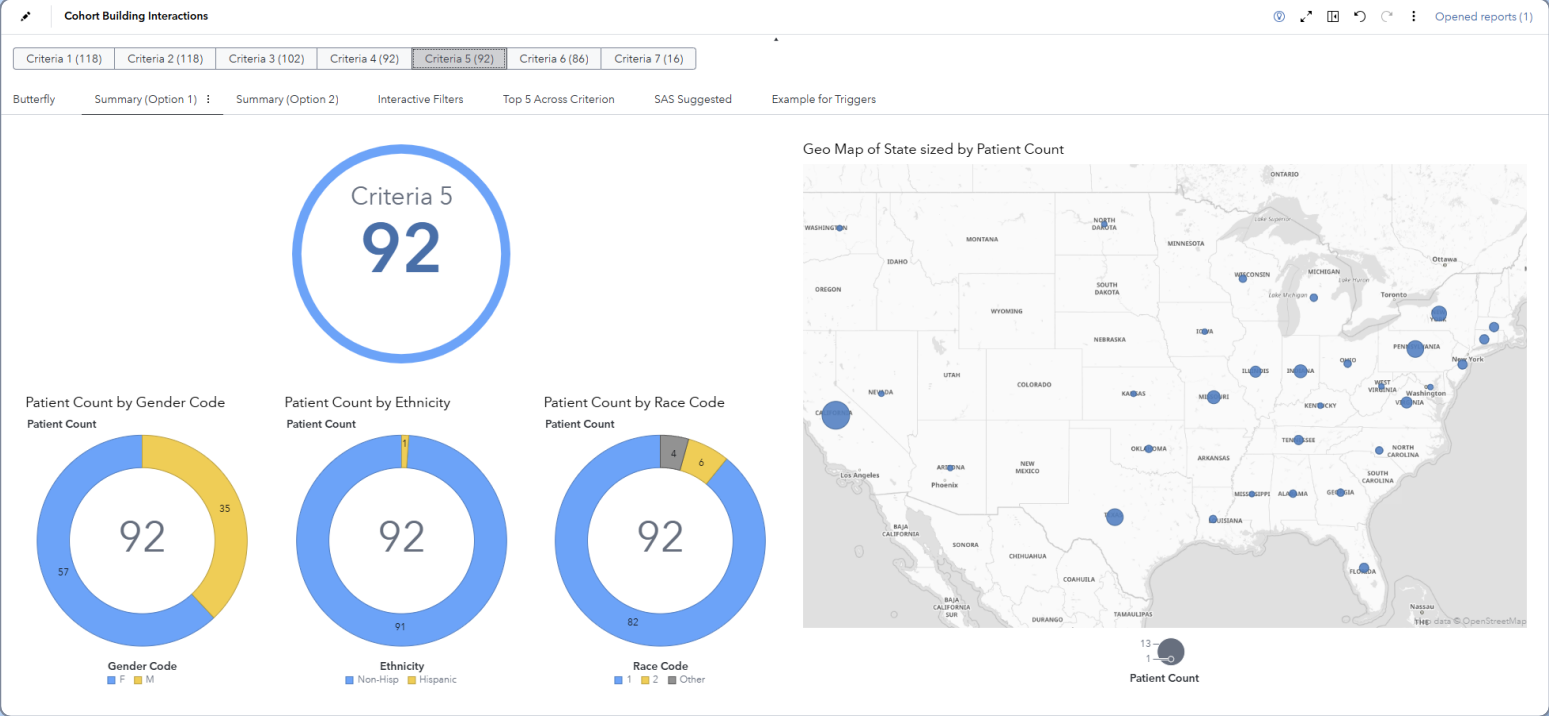
SAS Studio Flow vs. Conversational Agent + RAM

SAS Studio Flow Agent	
PROS	CONS
Deterministic execution paths	Rigid execution paths
Programs do not deviate from planned flow	Changes need to be hard-coded
Lineage and governance due to status as a SAS content “object”	SAS / SQL based paradigms for filter – requires conformation to syntax

Conversational Agent + RAM	
PROS	CONS
Flexible execution path	Hallucination risks
Agent decides flow on case-by-case basis	Need for stronger, explicit guardrails
Conversational interface accommodating wide range of filters	Need to track each execution path separately

Visual Analytics

Visual Analytics Dashboard



Visual Analytics



Stay in touch



Sherrine Eid, MPH
SAS Institute, Inc



Sherrine.Eid@sas.com



www.sas.com
www.linkedin.com/in/sherrineeid



Sundaresh Sankaran
SAS Institute, Inc



Sundaresh.Sankaran@sas.com



www.sas.com
www.linkedin.com/in/sundareshsankaran



Mary Dolegowski
SAS Institute, Inc



Mary.Dolegowski@sas.com



www.sas.com
www.linkedin.com/in/maryliang

