2024 Spring Session 9: CCSQ D&A Roadmap and Vision

Date: Thursday, 5/16 4:00-5:00 pm ET
Speakers: Kenneth Howard & Alisha Hutson & Justyna Sardin

1. **Q- If you are removing SAS_VIYA_ANALYST role, will we need a HARP role other than Databricks Developer?**
   A- Currently, we only have the Databricks Developer role.

2. **Q- I was wondering about our single sign-on (SSO)/Security. I see that SECURITY_OFFICIAL role is not tied to SAS_VIYA in HARP. Should I assume that role will not be affected by the SAS Viya decommissioning?**
   A- The Security Official role should not be affected.

3. **Q- Does Centralized Data Repository (CDR) perform its own final action? It sounds like the direct connection to Integrated Data Repository (IDR) is not there yet.**
   A- Correct, the connection to IDR is not there yet. We mimic the IDR Final Action, just like most other downstream systems. The IDR implementation of the Final Action is deemed the agency standard.

4. **Q- What set of aliases are used in NCH data in CDR? standard alias? Also, is claim-segment still a thing in the CDR?**
   A- According to https://www.cms.gov/research-statistics-data-and-systems/files-for-order/identifiabledatafiles/downloads/sasidhha.pdf, **DB2 ALIAS: NCH_REC_VRSN_CD.** Claim segment information in CDR can be found on the Claims Data Confluence page.

5. **Q- Is there a way to get on pilot evaluation lists for new analytic features, as they come up?**
   A- We do not have a list. What we will do is ask for volunteers and interest of anyone on any specific project or pilot during the monthly Townhall.

6. **Q- Data processing and variable creation is a long-term, high-level of effort area needing resources. Are there any efforts by a team within CMS or external contractor to convert the raw data (Medicare Part A, B, C, D claims, Beneficiary Information in the Cloud (BIC), Master Data Management (MDM), Provider etc.) into a beneficiary-level analytic file (containing diagnosed conditions, demographics, Medicare or Medicaid coverage type, etc.) applicable for analytic end users so other teams can focus more energy on analysis (statistical, machine learning, etc.) rather than everyone within the CCSQ CDR consuming an intense amount of human and computing resources to process and join the raw data across many databases and tables in order to convert into a useful format? Due to limited resources during a project, many analyses can only create/include a handful of variables to be used in an analysis. Consequently, the results can be difficult to understand because useful**

variables are missing, i.e., missing variables may be impacting the ability to interpret analysis results.

Also, someone external to CMS has a limited technical understanding about which databases, tables, table variables, table variable values should be used reliably to create a given variable included in an analysis.
A- There has always been an ongoing effort on the part of our Data & Analytics team. Multiple examples of analytic files are available in CDR and beyond (i.e., in CMS at large). It should be noted that data analytic needs are so diverse that there is never a "one size fits all"; therefore, there will always be a need to do custom work. Please provide more information on what is meant by "useful variables missing" - all our datasets are documented in the Data Catalog posted on CDR Confluence Data & Analytics space.

The CDR has posted multiple training series on various analytic tools; we also have a Knowledge Base and Wiki's with valuable tips and tricks.

7. **Q- Is the Medicare Quality Assurance (MQA) file structure like claim-segment still a thing in CDR?**
A- Yes. We ingested the original NCH files and kept them updated.