BioSense Platform User Manual for
Data-Quality-on-Demand Programs

March 2020

Center for Surveillance, Epidemiology, and Laboratory Services
Division of Health Informatics and Surveillance

Technical Assistance: support.syndromicsurveillance.org

The National Syndromic Surveillance Program (NSSP) promotes and advances development of the cloud-based BioSense Platform, a secure integrated electronic health information system that hosts standardized analytic tools and facilitates collaborative processes. The BioSense Platform is a product of the Centers for Disease Control and Prevention (CDC).

CONTENTS

1 Overview and Prerequisites, 1

2 How to Access DQOD Programs, 2

3 User Input and Customization, 3

Required User Input, 3
How to Submit Programs in the Background, 5

4 Reports: Output and Interpretation, 6

Overview of Reports, Output, and Log Files, 6
How to Download Reports, 8
Data Completeness Reports, 8
Data Timeliness Reports, 15
Data Validity Reports, 16

5 Datasets, 18

Completeness and Timeliness, 18
Validity, 18

6 Logs, 19

Appendix A: Setting Up SAS Studio the First Time, 20

How to Copy Personal Info Snippets to Files (Home) Folder, 20
How to Customize Info Snippets and Run Utility Snippets, 20

Appendix B: Datasets and Variables, 23

Datasets, 23
Validity Processing Variables, 24

Appendix C: Site IDs and Short Names, 26

Appendix D: Priority 1 and 2 Data Elements, 27

Appendix E: Tips for Improving Run Efficiency, 29

User Manual for DQOD Programs

1. Overview and Prerequisites

The Data-Quality-on-Demand (DQOD) programs described in this manual offer real-time access to data quality measures and flexible ways to run reports. BioSense Platform users with access to SAS Studio can run data quality checks on their site data as needed, on their schedule.

The DQOD programs enable users to assess the following characteristics of their data:

- **Completeness**—Provides information on how frequently data elements from messages associated with a unique visit are populated;
- **Timeliness**—Measures the delay from the date and time the visit occurred to when the message arrived on the BioSense Platform; and
- **Validity**—Checks incoming data to make sure current PHIN VADS[1] standards are being met. Compliance is calculated for records and visits, with the visit-level data mirroring what is used in ESSENCE.

This guide assumes the user is comfortable using SAS Studio within the BioSense Platform. For more information, see the *BioSense Platform Quick Start Guide to Using SAS Studio*.

Prerequisites

- Users must have a BioSense Platform account that has been granted access to SAS Studio and the DataMart for their site. Users needing access should contact their site administrator.
- New users must have accessed SAS Studio and set up the User_Info.sas and Site_Info.sas programs with their encrypted password and other information. See Appendix A or the *BioSense Platform Quick Start Guide to Using SAS Studio* for setup instructions.

[1] PHIN VADS stands for Public Health Information Network Vocabulary Access and Distribution System (https://phinvads.cdc.gov/vads/SearchVocab.action). PHIN VADS provides standard vocabularies to CDC and its public health partners in one place to support the exchange of consistent information. PHIN VADS is a Web-based enterprise vocabulary system for accessing, searching, and distributing vocabularies used in public health and clinical care.

2. How to Access DQOD Programs

Data-Quality-on-Demand (DQOD) programs are available through SAS Studio. The main DQOD program is named ***DataQuality_OnDemand.sas*** and is stored in the shared repository folder inside the programs folder (**opt/sas/shared/repository/programs**). You can navigate to this area through "Server Files and Folders" and drill down through sas-cmp1, Folder Shortcuts, SAS, programs. Refer to Figure 1.

The DQOD main program will call SAS programs stored in a subfolder named DQOD_secondary_programs. You do not need to directly access the SAS programs inside the DQOD_secondary_programs folder.

Save a copy of the DataQuality_OnDemand.sas program in your Files (Home) folder. This will update the program and save your customizations described in the next section.

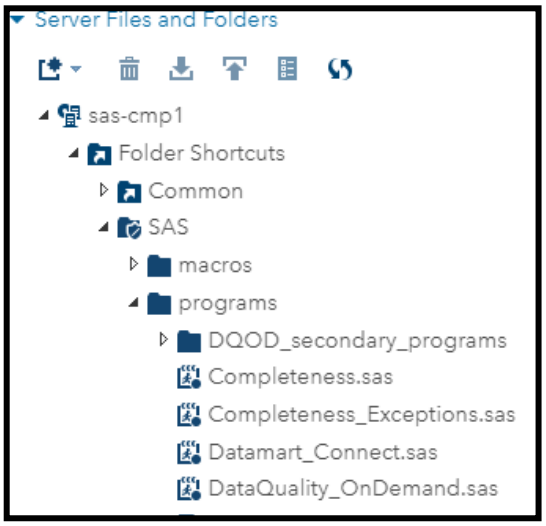


Figure 1. Location of DQ-on-Demand (DQOD) Programs

3. User Input and Customization

Required User Input

Go to the copy of DataQuality_OnDemand.sas saved in your Files (Home) folder. Follow steps 1 through 5 as seen in Figure 2. Once you are done, save and run the program, which can take a few minutes if you have selected a short interval such as a week. Longer intervals could take an hour or more.

Remember, you can launch the program in the background to improve efficiency. This makes the best use of available resources and allows you to continue other work in SAS Studio.

```
18 1.  Save DataQuality_OnDemand Program in your home directory [/opt/sas/shared/homes/%scan(&sysuserid.,1,@)]
19 2.  On LINES 36 and 37, Set site_id and site_short_name to that which is in your site_profile
20 3.  On LINES  39 and 40, Set the Start and End Date variables to the date range over which you would like to see
21         DQ Reports (consider setting this to no more than a few days the first time you run this program)
22 4.  OPTIONAL: On LINE  change dolog=N to dolog=Y if you want detailed logs saved
23 5.  Save and Run the program.
24
25 */
26
27 /* Directory Where DQOD Programs are stored and create subfolders to store output */
28 /*%let    sdir=/opt/sas/shared/homes/%scan(&sysuserid.,1,@)/public/NSSP/pgms/sas/dqod;*/
29 %let    sdir=/opt/sas/shared/repository/programs/DQOD_secondary_programs;
30 %put &sdir;
31 %include "&sdir/dqod_first.sas";    /* Creates DQOD and LOG, REPORT, and DATA subfolders in the users HOME directory */
32
33 /******************************
34     USERS INPUT BELOW
35 ******************************/
36 /* Site Information */
37 %let site_id= ;                                  /* EX: %let site_id=960;           See your site_profile
38 %let site_short_name= ;                          /* EX: %let site_short_name=AB;    See your site profile
39 /*Date Range over which data will be pulled to Create the DQOD reports */
40 %let start_date=20190101;                        /* EX:  20180901  Selected Data Range Start Date in YYYYMMDD format
41 %let end_date=20190107;                          /* EX:  20181116  Selected Data Range End Date   in YYYYMMDD format
42 %let dolog=N;                                    /* Detailed logs: Y=Write to Log, N=Do Not Write To Log
43 /* Choose whether the program should output Completeness, Timeliness, Validity, or ALL  */
44 %let report_type=C;                 /* EX: %let report_type=C Completeness = C, Timeliness = T, Validity = V, All=A */
45
```

Run Steps 1-5

Figure 2. Required User Input to DataQuality_OnDemand.sas

Site ID

On line 37 in the DataQuality_OnDemand program, enter the site ID listed in your site profile. For reference, a list of site IDs is available in Appendix C.

Site Short Name

On line 38 in the DataQuality_OnDemand program, enter the site short name listed in your site profile. This can be found in your User_Info.sas or Site_Info.sas files in your Files (Home) folder.

Start Date and End Date

On lines 40 and 41, enter start and end dates for the period you want this report to run. The dates must be in YYYYMMDD format (e.g., 20200202 for February 2, 2020).

Dolog

On line 42, you may choose to output detailed logs to a separate log folder. To do this, enter dolog=Y. Otherwise, enter dolog=N to have the logs appear in the main log window for SAS Studio.

Report_Type

On line 44, you can choose to run one report (either for completeness, timeliness, or validity) or run all three reports.

Reports Generated by Data-Quality-on-Demand Programs	
Type of Report	**Variable**
Completeness	report_type=C
Timeliness	report_type=T
Validity	report_type=V
All	report_type=A

Warning: Running all reports (report_type=A) **can take a great deal of time if you select a long interval such as a quarter or longer. Please consider only running the report you need.**

When running the Validity Report, you may further customize the output by setting a range of values for Validity processing using VLMIN and VLMAX variables.

VLMIN and VLMAX

The customization options, VLMIN and VLMAX, are **used in Validity processing only** and are located on lines 51 and 52 (Figure 3). By default, the validity report checks 61 different variables. However, you can change the values of VLMIN or VLMAX, thereby limiting the range of variables.

```
50 /* Validity Min and Max Columns for Processing and Reporting                    */
51 %let vlmin=1;            /* First Column Used in Validity processing            */
52 %let vlmax=61;           /* Last Column Used In Validity Processing             */
```

Figure 3. Validity Processing Customization Variables

You cannot pick and choose variables from the list. Instead, you must pick a range of values. The first 5 variables in the list of 61 are shown here. The full list is available in Appendix B.

"Short List" of Variables for Validity Processing			
vlno	**dqvarno**	**Name**	**name32**
1	110	C_Unique_Patient_ID_Source	C_Unique_Patient_ID_Source
2	270	C_Facility_ID_Source	C_Facility_ID_Source
3	300	Facility_Type_Code	Facility_Type_Code
4	340	Sending_Facility_ID_Source	Sending_Facility_ID_Source
5	460	C_Visit_Date_Source	C_Visit_Date_Source

Once you have updated your settings, save the program and click the **Running Man** icon to execute it.

How to Submit Programs in the Background

Depending on the time interval you choose and the number of reports you are processing, the job can run for many hours, sometimes even overnight. You may run the program in the background so that you can continue with other SAS Studio work. Refer to the *BioSense Platform Quick Start Guide to Using SAS Studio* for more information on how to submit background jobs and monitor their progress. Refer to Figure 4 for an example of the "Background Submit" option that pops up when you right click your program.

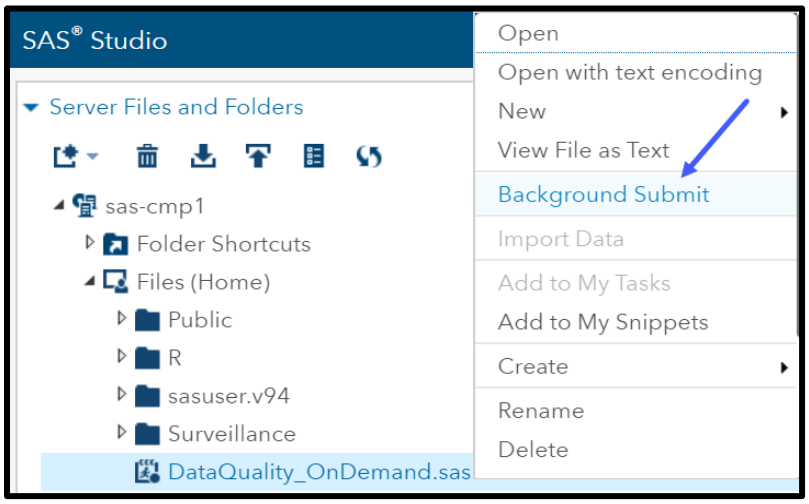


Figure 4. SAS Studio Background Submit of Programs

4. Reports: Output and Interpretation

Overview of Reports, Output, and Log Files

The DQOD programs will create a folder inside Files (Home) folder called **DQOD**. Inside this folder (Figure 5), three subfolders will be created: **data, log,** and **report.**

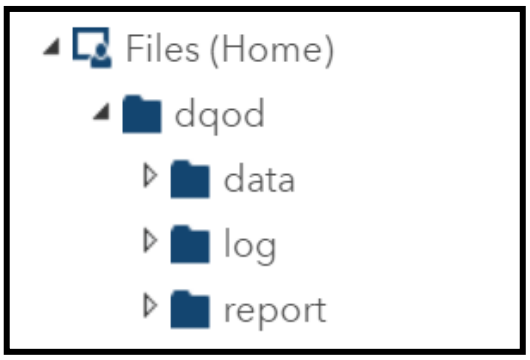


Figure 5. Folders Created by DQOD Programs

- The **data folder** will contain datasets based on the type of report that has been run for completeness, timeliness, or validity. You have the option to aggregate these datasets to a different period than the reports in the report folder.
- The **log folder** will contain detailed output logs.
- The **report folder** will contain one Excel-formatted (.xlsx) file for each report run for completeness, timeliness, and validity. These reports will be aggregated to a single, user-selected period.

Three separate reports can be created: Completeness, Timeliness, or Validity. If you choose to run all reports, all three will be displayed. If you choose to run one type of report, the other two will *not* be created. Reports are shown in the results section of your SAS Studio session and are also saved as .xlsx files in the report folder (Figure 6).

Figure 6. Reports Created by Using the "report_type=A" Option

The naming convention for reports is as follows:

dqod_[site_ID]_[dqv][Report Type]_[start_date][end_date]

Example: dqod_XX_0Completeness_190101190107.xls

Report-Naming Convention			
Naming Position	**Variable Name**	**Example**	**Notes**
1	DQOD	DQOD	Default "DQOD"
2	Site ID	XX	As entered in DataQuality_OnDemand program
3	DQV	0	Default 0 unless changed in DQODmain program
4	Report Type	Completeness	Completeness, Timeliness, or Validity
5	Start_Date	190101	As entered in DataQuality_OnDemand program (year only shows two digits in report file name)
6	End_Date	190107	As entered in DataQuality_OnDemand program (year only shows two digits in report file name)

SAS Studio will display an error message if the results exceed 3MB. To display the results, simply click **Display Anyway**, and the results will load (Figure 7).

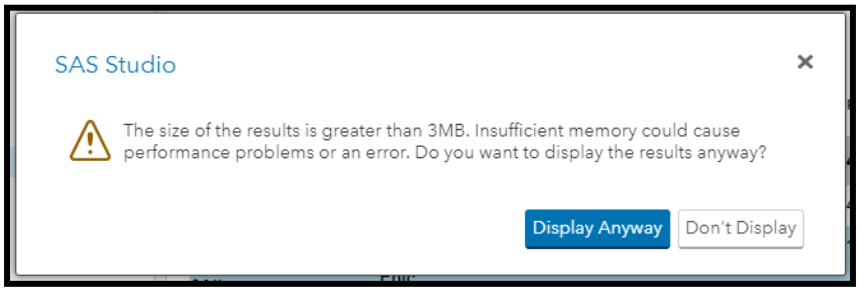
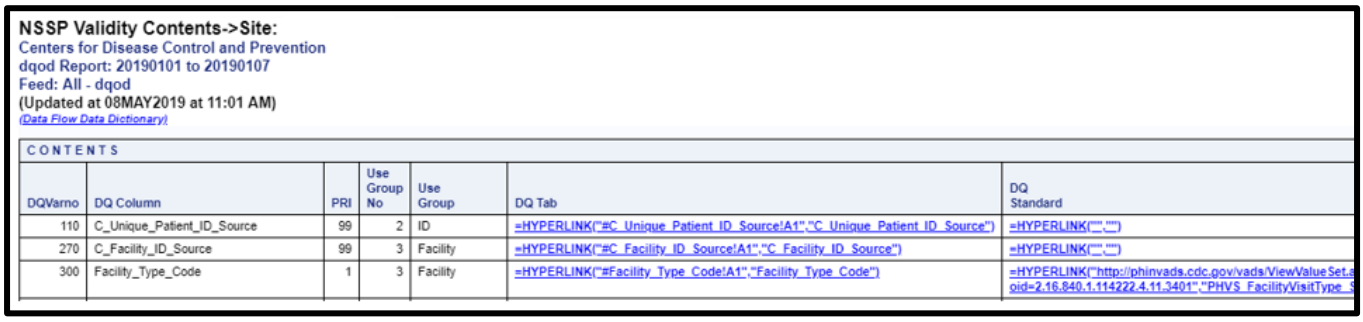


Figure 7. DQOD "Report-too-Large" Error Message

When you view the validity report in the SAS Studio results window, the hyperlinks will *not* display accurately. Once you **enable editing in Excel**, the hyperlinks will display accurately when you download the report (Figure 8).



NSSP Validity Contents->Site:
Centers for Disease Control and Prevention
dqod Report: 20190101 to 20190107
Feed: All - dqod
(Updated at 08MAY2019 at 11:01 AM)
(Data Flow Data Dictionary)

CONTENTS

DQVarno	DQ Column	PRI	Use Group No	Use Group	DQ Tab	DQ Standard
110	C_Unique_Patient_ID_Source	99	2	ID	=HYPERLINK("#C_Unique_Patient_ID_Source!A1","C_Unique_Patient_ID_Source")	=HYPERLINK("","")
270	C_Facility_ID_Source	99	3	Facility	=HYPERLINK("#C_Facility_ID_Source!A1","C_Facility_ID_Source")	=HYPERLINK("","")
300	Facility_Type_Code	1	3	Facility	=HYPERLINK("#Facility_Type_Code!A1","Facility_Type_Code")	=HYPERLINK("http://phinvads.cdc.gov/vads/ViewValueSet.a oid=2.16.840.1.114222.4.11.3401","PHVS_FacilityVisitType_S

Figure 8. Validity Report Hyperlink Display in SAS Studio Browser

How to Download Reports

Download the .xlsx version of the reports by right clicking on the report and selecting **Download File** (Figure 9).

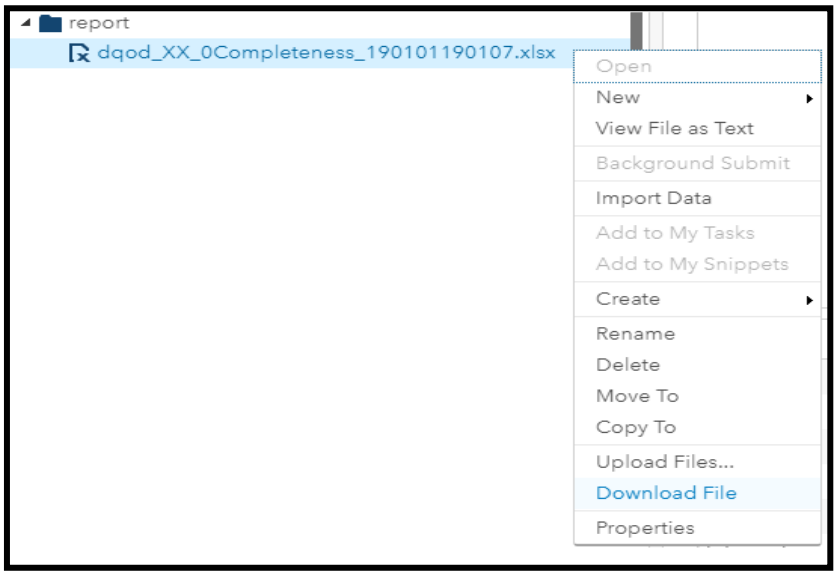


Figure 9. DQOD .xlsx Report Download

Data Completeness Reports

The DQ Completeness report provides information about the facility's ability to populate data elements from various messages or records being sent for a unique visit. The percent of completeness is obtained at the visit level, where individual messages are collapsed into a unique visit. A *unique visit* is defined as a unique C_BioSense_ID, where C_BioSense_ID is a concatenation of C_VisitDate, C_Biosense_facility_ID, and C_Unique_Patient_ID.

The DQ Completeness report contains standard tabs in addition to tabs for each feed. The standard tabs are XX_Summary, XX_Exceptions, XX_Cells, and XX_Cells_Red.

XX_SUMMARY Tab

The XX_Summary tab (Figure 10) shows total visits, patients, records, and minimum (min) and maximum (max) message arrival dates by feed and facility.

NSSP Completeness Summary Report->Site: X
Centers for Disease Control and Prevention
Completeness Report Month: 201711
Based On Processed Data Selected By C_Visit_Date from 2017-11-01 to 2017-11-30
Run On 22DEC2017 at 11:05 PM
(Data Flow Data Dictionary)

Feed Name	Facility Type	C Biosense Facility ID	Facility Name	Facility Count	Visits	Patients	Records	Min Arrived Date Time
$ALL Feeds	**$ALL Types**	**ALL**	**Facilities**	**25**	**261,102**	**197,407**	**1,679,955**	**01Nov17:0**
$ALL Feeds	Emergency Care	ALL	Facilities	13	158,763	119,407	1,357,721	01Nov17:00:0
$ALL Feeds	Medical Specialty	ALL	Facilities	1	5,124	3,259	10,848	01Nov17:07:0
$ALL Feeds	Primary Care	ALL	Facilities	11	97,215	74,741	311,386	01Nov17:07:0
Feed A	**$ALL Types**	**ALL**	**Facilities**	**1**	**12,467**	**10,669**	**356,388**	**01Nov17:0**
Feed A	**Emergency Care**	**ALL**	**Facilities**	**1**	**12,467**	**10,669**	**356,388**	**01Nov17:0**
Feed A	Emergency Care	1500	Facility 1	1	12,467	10,669	356,388	01Nov17:00:0

Figure 10. Completeness Report XX_Summary Tab

DATA QUALITY USE TIP: SUMMARY TAB
Use this tab to review the (1) last time of data processing for a specific feed or facility and (2) total records processed within the reported period.

XX_EXCEPTIONS Tab

The XX_Exceptions tab (Figure 11) provides the total records successfully processed and records sent to exceptions by feed and facility. The processed and exception numbers are separated into two sections that provide total records, percent processed/exceptions, min/max arrived dates, and min/max C_visit_date.

NSSP Exceptions Report->Site: X
Centers for Disease Control and Prevention
Exceptions Report Month: 201711
Based On Processed Data Selected By C_Visit_Date from 2017-11-01 to 2017-11-30
Run On 22DEC2017 at 11:05 PM
(Data Flow Data Dictionary)

Identification				Total Records Found In Processed, Exceptions and Filtered				Total Processed					Total Exceptions		
Feed Name	Facility Type	C Biosense Facility ID	Facility Name	Total Records	Filtered Records	% Total Records Filtered %	Total Read	Records	% Records of Total Read	Visits	Patients	Facilities	Record	% Records of Total Read	Visit
$ALL Feeds	**$ALL Types**	**ALL**	**Facilities**	**1,700,214**	**0**	**0.0%**	**1,700,214**	**1,681,503**	**98.9%**	**261,102**	**197,407**	**25**	**18,711**	**1.1%**	**4,698**
$ALL Feeds	Emergency Care	ALL	Facilities	1,359,269	0	0.0%	1,359,269	1,359,269	100.0%	158,763	119,407	13	0	0.0%	0
$ALL Feeds	Medical Specialty	ALL	Facilities	10,848	0	0.0%	10,848	10,848	100.0%	5,124	3,259	1	0	0.0%	0
$ALL Feeds	Primary Care	ALL	Facilities	311,386	0	0.0%	311,386	311,386	100.0%	97,215	74,741	11	0	0.0%	0
$ALL Feeds	Unknown	ALL	Facilities	18,711	0	0.0%	18,711	0	0.0%	0	0	0	18,711	100.0%	4,698
Feed A	**$ALL Types**	**ALL**	**Facilities**	**357,185**	**0**	**0.0%**	**357,185**	**357,185**	**100.0%**	**12,467**	**10,669**	**1**	**0**	**0.0%**	**0**
Feed A	**Emergency Care**	**ALL**	**Facilities**	**357,185**	**0**	**0.0%**	**357,185**	**357,185**	**100.0%**	**12,467**	**10,669**	**1**	**0**	**0.0%**	**0**

Figure 11. Completeness Report XX_Exceptions Tab

DATA QUALITY USE TIP: EXCEPTIONS TAB
Use this tab to assess how many records were sent to exceptions for a feed/facility during the reported period.

X_ByFacility Tab

Click the XX_ByFacility tab (Figure 12) to see an overall percent of completeness at the record and visit level for all variables ($ALL) or an individual view by feed and site. Variables are grouped by Use Group types (e.g., ID, Visit Info, Demographics).

NSSP Completeness Report->Site: X
Centers for Disease Control and Prevention
Completeness Report Month: 201711 ->No Alerting For Required Fields With < 90% Completeness
Based On Processed Data Selected By C_Visit_Date from 2017-11-01 to 2017-11-30
Run On 22DEC2017 at 11:05 PM
(Data Flow Data Dictionary)

Report Type:=Site_X_ByFacility

				Usegrpno, Usegrp, Dqvarno, Name, Required, HL7									
Feed_Name, Facility_Type, C_Biosense_Facility_ID, Facility_Name				00 Stats 0050 Visits/Recs X NA		00 Stats 0060 Facility Count X NA		02 ID 0080 C BioSense ID CR NA		02 ID 0090 C Processed BioSense ID CR NA		02 ID 0100 C Unique Patient ID CR PID.3.1	
Feed Name	Facility Type	C Biosense Facility ID	Facility Name	% Visits	% Recs	% Visit	% Rec	% Visit	% Recs	% Visit	% Recs	% Visit	% Rec
$ALL Feeds	**$ALL Types**	**ALL**	**Facilities**	**261,102**	**1,679,955**	**25**	**25**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100**
$ALL Feeds	Emergency Care	ALL	Facilities	**158,763**	**1,357,721**	**13**	**13**	100.0%	100.0%	100.0%	100.0%	100.0%	10
	Medical Specialty	ALL	Facilities	**5,124**	**10,848**	**1**	**1**	100.0%	100.0%	100.0%	100.0%	100.0%	10
	Primary Care	ALL	Facilities	**97,215**	**311,386**	**11**	**11**	100.0%	100.0%	100.0%	100.0%	100.0%	10
Feed A	**$ALL Types**	**ALL**	**Facilities**	**12,467**	**356,388**	**1**	**1**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100**
Feed B	$ALL Types	ALL	Facilities	**196,933**	**810,850**	**14**	**14**	100.0%	100.0%	100.0%	100.0%	100.0%	10
Feed A	**Emergency**	**ALL**	**Facilities**	**12,467**	**356,388**	**1**	**1**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100**

Figure 12. Completeness Report XX_ByFacility Tab

DATA QUALITY USE TIP: ByFacility TAB
This tab shows a horizontal view of percent completeness for data elements by feed, facility, and facility type.

All_Feeds Tab

Click the **All_Feeds tab** (Figure 13) to see a vertical view of percent completeness for data elements by feed, facility, and facility type. Here are some possibilities:

- Filter by priority, use group, required, or HL7 fields.
- Filter by color "red" for %Visits or %Recs or by Priority or Required fields.
- Identify fields with less than 80% completeness (highlighted in red).

Refer to the *NSSP Data Dictionary*, ESSENCE tab, to see more information on the variables used for the calculated field.

NOTE: Not all fields highlighted in red indicate values that are missing. Sometimes red denotes values received but invalid.

NSSP Completeness Report->Site: X
Centers for Disease Control and Prevention
Completeness Report Month: 201711
Based On Processed Data Selected By C_Visit_Date from 2017-11-01 to 2017-11-30
Run On 22DEC2017 at 11:05 PM
(Data Flow Data Dictionary)

Feed:=All_Feeds

						Column Order No, Feed_Name, Facility_Type, C_Biosense_Facility Facility_Name				
Dqvarno, Name, Pri, Usegrp, Required, HL7_Segment						1 $ALL Feeds $ALL Types ALL Facilities		2 $ALL Feeds Emergency Care ALL Facilities		2 $ALL Feeds Medical Specialty ALL Facilities
DQ Var No	Processed Column	PRI	Use Group	Re-quir-ed	HL7	% Visits	% Recs	% Visits	% Recs	% Visits
0050	Visits/Recs	1	Stats	X	NA	261,102	1,679,955	158,763	1,357,721	5,124
0060	Facility Count	1	Stats	X	NA	25	25	13	13	1
0080	C_BioSense_ID	99	ID	CR	NA	100.0%	100.0%	100.0%	100.0%	100.0%
0090	C_Processed_BioSense_ID	99	ID	CR	NA	100.0%	100.0%	100.0%	100.0%	100.0%
0100	C_Unique_Patient_ID	1	ID	CR	PID.3.1	100.0%	100.0%	100.0%	100.0%	100.0%
0110	C_Unique_Patient_ID_Source	99	ID	C	PID.3.1	100.0%	100.0%	100.0%	100.0%	100.0%
0120	First_Patient_ID	2	ID	R	PID-3.1	100.0%	100.0%	100.0%	100.0%	100.0%
0130	First_Patient_ID_Type_Code	3	ID	O	PID-3.5	100.0%	100.0%	100.0%	100.0%	100.0%

Figure 13. Completeness Report All_Feeds Tab

DATA QUALITY USE TIPS: $ALL TAB FILTERS

PRIORITY LEVEL
▪ Filter by "PRI" column to identify Priority 1 and 2 data elements.
  ○ P1: Priority 1 fields are the minimum required data elements for NSSP onboarding*
  ○ P2: Priority 2 fields are additionally required data elements (see *PHIN Messaging Guide for Syndromic Surveillance*)
▪ Look for fields with overall completeness less than (<) 90% that are highlighted in red. These fields may require further investigation.
 	
Investigation tips: Priority 1 data elements are essential for processing and analyzing data. "Minimum" refers only to data elements that let you assess the quality of foundational data. By filtering on Priority 1, you can quickly determine if incoming data have sufficient information to be processed successfully.

CLINICAL DATA (i.e., chief complaint, diagnosis)
▪ Filter on "Use Group" value of CC_Diagnostic. The fields listing include Chief Complaint and Diagnosis among other "diagnostic" type fields.
▪ View the percentages for Admit_Reason and Chief Complaint because both are used to populate the C_Chief_Complaint field.
  ○ For an example, an Admit_Reason of 40% completeness is not alarming if you have a Chief_Complaint completeness of 60%. Therefore, the CALCULATED chief complaint should have 100% completeness.
  ○ Use the Feed tab to identify specific facilities with a low percentage for Admit Reason.
▪ If Chief_Complaint_Code is low, it is okay if your feeds send TEXT to ensure Chief_Complaint_Text is more complete. As you review the diagnosis, keep in mind the timeframe in which reports are generated because the diagnosis might not have been received at time of processing.

DEMOGRAPHICS ▪ Filter on "Use Group" value of Demographics. ▪ Note the critical field Patient_Zip, which influences the "Regions" used in ESSENCE. Regions are similar to counties. ESSENCE infers the regions based on the patient ZIP code. ▪ Assess the percent completeness for calculated patient age (C_Patient_Age). Your data may have birthdate and Reported age, and both fields are used to generate the NSSP calculated age. ▪ Note that the HL7 message did not have Age_Calculated, which is usually part of the message. This is acceptable and common.
PATIENT CLASS ▪ Filter on "Use Group" value of Visitinfo. ▪ Verify Patient Class is complete.

XX_Cells Tab

The <Site>_Cells tab (Figure 14) includes everything in one report, including Overall ($All Feeds), specific feed, and specific facility-level data completeness.

- Use the Feed_Name column to filter on $All Feeds or on a specific feed of interest. Sites may create feed-specific reports for vendors simply by filtering and copying the data into a separate sheet.
- Use the Facility_Name to filter on a facility of interest.

NSSP Completeness Report->Site: XXX Site_X
Centers for Disease Control and Prevention
Completeness Report All Cells Month: 201802
Based On Processed Data Selected By C_Visit_Date from 2018-02-01 to 2018-02-28
Run On 03APR2018 at 10:38 AM
(Data Flow Data Dictionary)

Feed:=Site_X_Cells

Feed_Name	Facility_Type	C Biosense Facility ID	Facility_Name	DQ Var No	Processed Column	PRI	Use Group	Re-quir-ed	HL7	% Visit	Visits	% Recs	Recs
$ALL Feeds	**Emergency Care**	**All**	**Facilities**	**0050**	**Visits/Recs**	**1**	**Stats**	**X**	**NA**	**100.0%**	**157,753**	**100.0%**	**1,331,681**
$ALL Feeds	**Emergency Care**	**All**	**Facilities**	**0060**	**Facility Count**	**1**	**Stats**	**X**	**NA**	**100.0%**	**13**	**100.0%**	**13**
$ALL Feeds	Emergency Care	All	Facilities	0080	C_BioSense_ID	99	ID	CR	NA	100.0%	157,753	100.0%	1,331,681
$ALL Feeds	Emergency Care	All	Facilities	0090	C_Processed_BioSense_ID	99	ID	CR	NA	100.0%	157,753	100.0%	1,331,681
$ALL Feeds	Emergency Care	All	Facilities	0100	C_Unique_Patient_ID	1	ID	CR	PID.3.1	100.0%	157,753	100.0%	1,331,681
$ALL Feeds	Emergency Care	All	Facilities	0110	C_Unique_Patient_ID_Source	99	ID	C	PID.3.1	100.0%	157,753	100.0%	1,331,681

Figure 14. Completeness Report XX_Cells Tab

DATA QUALITY USE TIP: XX_Cells TAB
Create a Report for Specific Vendors/Feeds or Facilities Filter the Feed_Name column by a particular Feed or Facility_Type: ▪ When Facility_Type contains “ALLTypes,” the “ALLTypes” selection includes overall completeness for that feed. ▪ To view Emergency Facilities, filter by Facility_Type = “Emergency Care.”
Vendor Performance with Priority 1 or Priority 2 Data Elements ▪ To filter by Priority 1 or 2 data elements, select PRI = 1 or PRI=2. ▪ To view Emergency Facilities *only*, filter by Facility_Type = “Emergency Care.”

XX_Cells_Red Tab

The <Site>_Cells_Red tab (Figure 15) includes all red highlighted cells found throughout the report.

- Use this one-stop shop to find completeness issues at a glance.
- Check worksheet for overall feeds ($ALL Feeds), specific feed, and specific facility-level information for columns with <90% completeness. Use the new worksheet to find potential problematic cells at a glance.

NSSP Completeness Report->Site:XXX - Site_X

Centers for Disease Control and Prevention

Completeness Report Red Cells Only Month: 201802

Based On Processed Data Selected By C_Visit_Date from 2018-02-01 to 2018-02-28

Run On 03APR2018 at 10:38 AM

(Data Flow Data Dictionary)

Feed:=Site_X_Cells_Red

Feed_Name	Facility_Type	C Biosense Facility ID	Facility_Name	DQ Var No	Processed Column	PRI	Use Group	Re-quir-ed	HL7	% Visits	Visits	% Recs	Recs
Feed B	Primary Care	XXXX	Facility 5	1160	Chief_Complaint_Text	1	CC_Diagnostic	RE	OBX-2, OBX-3, OBX-5	81.1%	2,392	53.4%	5,941
Feed B	Primary Care	XXXX	Facility 5	1170	Chief_Complaint_Type	99	CC_Diagnostic	RE	NA	81.1%	2,392	53.4%	5,941
Feed B	Primary Care	XXXX	Facility 5	1180	C_Chief_Complaint	1	CC_Diagnostic	CRE	NA	81.1%	2,392	53.4%	5,941
Feed B	Primary Care	XXXX	Facility 5	1220	Diagnosis_Combo	99	CC_Diagnostic	RE	DG1.3.2DG1.3.3	71.2%	2,101	39.6%	4,401
Feed B	Primary Care	XXXX	Facility 5	1230	Diagnosis_Code	1	CC_Diagnostic	RE	DG1.3.1	71.2%	2,101	39.6%	4,401
Feed B	Primary Care	XXXX	Facility 5	1240	Diagnosis_Description	1	CC_Diagnostic	RE	DG1.3.2DG1.3.3	71.2%	2,101	39.6%	4,401
Feed B	Primary Care	XXXX	Facility 5	1250	Diagnosis_Segment	99	CC_Diagnostic	RE	DG1.3.2DG1.3.3	71.2%	2,101	39.6%	4,401
Feed B	Primary Care	XXXX	Facility 5	1260	Diagnosis_Type	2	CC_Diagnostic	RE	DG1.6.1	71.2%	2,101	39.6%	4,401
Feed B	Primary Care	XXXX	Facility 5	1780	Message_Profile_ID	99	Operations	R	MSH.21.1	0.0%	0	0.0%	0

Figure 15. Completeness Report XX_Cells_Red Tab

Data Timeliness Reports

The Data Quality Timeliness Report (Figure 16) measures the lag between the visit date/time and message arrival time on the BioSense Platform. The Data Timeliness reports and graphs note timeliness of visit-level data for 24- or 48-hour periods. To calculate lag time and thereby establish timeliness, SAS Studio uses the Arrived_Date_Time of the first message (for a visit) versus the date/time, C_Visit_Date_Time, of the patient visit. Subsequent messages for the same visit are not included in the calculation.

Site_XX_Summary Tab

The red, yellow, and green rows indicate the percent of visits received within 24 or 48 hours of the visit and the number of facilities associated with those visits.

NSSP Timeliness Summary Report->Site X
Centers for Disease Control and Prevention
Timeliness Report Month: 201711
Based On Processed Data Selected By C_Visit_Date from 2017-11-01 to 2017-11-30
Run On 24JAN2018 at 5:32 AM

	Within 1 day				Within 2 days			
Timeliness Performance Categorization (%)	Number of Facilities (N)	Cumulative Number of Facilities (N)	Percentage of Facilities (%)	Cumulative Percentage of Facilities (%)	Number of Facilities (N)	Cumulative Number of Facilities (N)	Percentage of Facilities (%)	Cumulative Percentage of Facilities (%)
00-30	1	1	4.0%	4.0%	1	1	4.0%	4.0%
30-80	6	7	24.0%	28.0%	0	1	0.0%	4.0%
>80	18	25	72.0%	100.0%	24	25	96.0%	100.0%

Basis by which the total percentage of visits for a facility was grouped among three ranges (0%-30%, 30%-80%, >=80%)
A. Red->Low Performance: Percentage of visits received is between 0% and 30%
B. Yellow->Intermediate Performance: Percentage of visits received is between 30% and 80%
C. Green ->High Performance: Percentage of visits received is above 80%

Figure 16. Timeliness Report Site_XX_Summary Tab

Site_XX_Detail Tab

The Detail tab (Figure 17) provides timeliness by facility_type, feed, and facilities for percent of visits received within the 24- and 48-hour periods. Within this tab, the mean and median lag days for each feed, facility type, and facility are also provided.

NSSP Timeliness Report->Site X

Centers for Disease Control and Prevention
Timeliness Report Month: 201711
Based On Processed Data Selected By C_Visit_Date from 2017-11-01 to 2017-11-30
Run On 24JAN2018 at 5:32 AM

Feed Name	Facility_Type	C Biosense Facility ID	Facility Name	Facility Count	Visits	% of visits received within 24 hours (1 day)	% of visits received within 48 hours (2 days)	Mean number of lag days	Median number of lag days	80% of visits were received within X days
$ALL Feeds	**$ALL Types**	**ALL**	**Facilities**	**25**	**261,150**	**91.548%**	**97.531%**	**0.71**	**0.32**	**0.44**
$ALL Feeds	Emergency Care	ALL	Facilities	13	158,811	86.109%	95.946%	1.00	0.34	0.73
$ALL Feeds	Medical Specialty	ALL	Facilities	1	5,124	99.922%	99.941%	0.27	0.26	0.41
$ALL Feeds	Primary Care	ALL	Facilities	11	97,215	99.991%	99.994%	0.28	0.30	0.40
Feed A	**$ALL Types**	**ALL**	**Facilities**	**0**	**12,467**	**100.0%**	**100.0%**	**0.01**	**0.01**	**0.02**
Feed A	**Emergency Care**	**ALL**	**Facilities**	**0**	**12,467**	**100.0%**	**100.0%**	**0.01**	**0.01**	**0.02**
Feed A	Emergency Care	1500	Facility 1	1	12,467	100.0%	100.0%	0.01	0.01	0.02
Feed B	**$ALL Types**	**ALL**	**Facilities**	**0**	**196,933**	**99.978%**	**99.993%**	**0.28**	**0.29**	**0.40**
Feed B	**Emergency Care**	**ALL**	**Facilities**	**0**	**94,594**	**99.967%**	**99.995%**	**0.27**	**0.29**	**0.41**
Feed B	Emergency Care	1510	Facility 1	1	5,017	100.0%	100.0%	0.25	0.26	0.38

Figure 17. Timeliness Report Site_XX_Detail Tab

Data Validity Reports

The DQ Validity Report identifies if incoming data conform with PHINVAD standards. Conformance is calculated at record and visit levels, with the visit level mirroring what is used in ESSENCE.

- Data either do or do not conform, and supporting reports specify the *value of conformance*. Missing data elements are categorized as *nonconforming*.
- DQ Validity reports provide record-level information. The reports include visit-level information by collapsing data across records and by using the same logic applied to downstream ESSENCE processing.
- NSSP encourages sites and facilities to achieve 100% compliance with data completeness and validation for all Priority 1 and Priority 2 data elements.
- Per onboarding guidance, all Priority 2 data elements are required to be compliant in validity and completeness within 12 months after a promotion to Production.
- DQOD users should focus on Priority 1 and 2 data elements to identify data validity issues.
- Users may want to download and save DQOD validity reports for comparisons of data validity levels over time or run the DQOD validity program for different time periods to identify changes in data validity levels. (See Figures 18 and 19.)
- The NSSP Validity Summary report provides links to the PHINVAD standards.
- Appendix D provides a table of the Priority 1 and 2 data elements for reference.

NSSP Validity Summary Report->Site X
Centers for Disease Control and Prevention
Validity Report Month: 201711
Based On Processed Data Selected By C_Visit_Date from 2017-11-01 to 2017-11-30
Run On 30JAN2018 at 1:05 AM
(Data Flow Data Dictionary)

DQ Column	DQ Standard
Facility_Type_Code	PHVS_FacilityVisitType_SyndromicSurveillance_V2
Discharge_Disposition	PHVS_DischargeDisposition_HL7_2x_V1
Patient_Class_Code	PHVS_PatientClass_SyndromicSurveillance_V1
C_Factype_Patient_Class	
C_MFT_Patient_Class	PHVS_PatientClass_SyndromicSurveillance_V1
C_Patient_Class	PHVS_PatientClass_SyndromicSurveillance_V1
C_Death	
Admission_Type	PHVS_AdmissionType_HL7_2x
Admit_Source	PHVS_AdmitSource_HL7_2x

Figure 18. Validity Summary Report Overview

NSSP Validity Summary Report->Site X
Centers for Disease Control and Prevention
Validity Detail Report Month: 201711
Based On Processed Data Selected By C_Visit_Date from 2017-11-01 to 2017-11-30
Run On 30JAN2018 at 1:05 AM
(Data Flow Data Dictionary)

Link To PHVS_Gender_SyndromicSurveillance_V1
(Click to return to the Summary Tab)

DQ Column:=Administrative_Sex

Feed Name	Facility Type	C Biosense Facility ID	Facility Name	Class	Value	Facility Cnt	Records	% Records	Visits	% Visits
$ALL Feeds	**$ALL Types**	**ALL**	**Facilities**	**0**	**$ALL_Values**	**25**	**1,747,094**	**100.0%**	**261,221**	**100.0%**
$ALL Feeds	**$ALL Types**	**ALL**	**Facilities**	**1**	**$ALL_Conforming**	**25**	**1,747,093**	**100.0%**	**261,220**	**100.0%**
$ALL Feeds	$ALL Types	ALL	Facilities	1	F	25	1,001,926	57.348%	154,140	59.008%
$ALL Feeds	$ALL Types	ALL	Facilities	1	M	25	744,932	42.638%	107,058	40.984%
$ALL Feeds	$ALL Types	ALL	Facilities	1	O	25	11	0.001%	3	0.001%
$ALL Feeds	$ALL Types	ALL	Facilities	1	U	25	224	0.013%	19	0.007%
$ALL Feeds	**$ALL Types**	**ALL**	**Facilities**	**2**	**$ALL_NonConforming**	**1**	**1**	**0.0%**	**1**	**0.0%**
$ALL Feeds	$ALL Types	ALL	Facilities	2	$Missing	1	1	0.0%	1	0.0%

Figure 19. Validity Detail Report

5. Datasets

Completeness and Timeliness

A sample of datasets created by the DQOD programs is listed below. A full list of datasets can be found in Appendix B. In the DQOD-file naming scheme for files 1 through 24, File_Type is the identifier string.

Sample Datasets Created by Data-Quality-on-Demand Programs		
File Number	**File Type**	**Description**
1	CCR	Chief Complaint Rankings
2	CLFT	Completeness details—Transposed
3	CLSUM	Completeness Summation of the Completeness Details—Transposed by Overall, Vendor, Parent_Organization, Feed_Name, and Facility levels for the time aggregations selected
4	CLVR	Completeness details by all columns selected
5	DCR	Diagnosis Code Rankings

Each dataset has the following naming convention:

DQ&site_id. _&dqv.[File_Type]._&fbe Ex: DQ860_0CCR_181116190215

Where:

- **site_id:** Site ID entered in the DataQuality_OnDemand program
- **dqv:** Data selection version number (default is 0)
- **fbe**: Last 6 digits of the start_date and end_date concatenated together. For example, if start_date is 20181116 and end_date is 20190215, then FBE will be 181116190215.

Validity

Validity summation datasets follow a different naming convention. Validity datasets *always* contain VS for the file type, for example:

DQ&site_id._&dqv.VS&dqvarno. _&fbe. Ex: DQ860_0VS110_181116190215

Where:

- **dqvarno:** Numeric identifier for one of the 61 columns analyzed for validity
- **site_id**: Site ID entered in the DataQuality_OnDemand program
- **fbe:** Last 6 digits of the start_date and end_date concatenated together. For example, if start_date is 20181116 and end_date is 20190215, then **FBE** will be 181116190215.

Here is another example: 110 is the numeric identifier for C_Unique_Patient_ID_Source. The validity summation file for the column is: DQ&site_id._&dqv.VS110_&fbe. Ex: DQ860_0VS110_181116190215

Short List of Variables for Use in Validity Processing			
vlno	**dqvarno**	**name**	**name32**
1	110	C_Unique_Patient_ID_Source	C_Unique_Patient_ID_Source
2	270	C_Facility_ID_Source	C_Facility_ID_Source
3	300	Facility_Type_Code	Facility_Type_Code
4	340	Sending_Facility_ID_Source	Sending_Facility_ID_Source
5	460	C_Visit_Date_Source	C_Visit_Date_Source

6. Logs

If the Dolog=Y option is chosen on line 41 of the DQOD program, detailed logs will be recorded in the log folder. The logs files in the table below will be created for each type of report.

Completeness	**Timeliness**	**Validity**
DQODmain	DQODmain	DQODmain
DQODdtinit	DQODdtinit	DQODdtinit
DQODvp	DQODvp	DQODvp
DQODclrpt	DQODmrpt	DQODvldt
DQODclsum	DQODmsum	DQODvlrpt

Appendix A: Setting Up SAS Studio the First Time

First-time SAS Studio users must follow a few steps to use DQOD programs. Here is a high-level summary of the steps to take:

1. In the Folder Shortcuts, expand the **SAS** folder, then the snippets folder. Copy **Site_Info.sas** and **User_Info.sas** snippets to your Files (Home) folder.
2. Modify and execute **Password_Encode.sas** to encrypt your password, then set up the **User_Info.sas** and **Site_Info.sas** snippets. Except for your encrypted password, your site administrator can provide the information needed for these files.

The next two steps are not required for DQOD but are included for completeness:

3. Set up and connect to your site-specific folder.
4. Connect to the DataMart library. (This is done dynamically during the DQOD run.)

How to Copy Personal Info Snippets to Files (Home) Folder

In the Navigator pane on the left, expand the **sas-cmp1 > Folder Shortcuts > SAS > snippets** folder and right-click on each of the following snippets. In the sub-menu popup, click on **Copy To** and select your **Files (Home)** folder to copy these:

- **User_Info.sas**
- **Site_Info.sas**

How to Customize Info Snippets and Run Utility Snippets

These snippets are required to use DQOD programs and many other provided programs in your SAS environment.

Password_Encode.sas
Passwords stored on the BioSense Platform are *not* allowed to be saved in clear text. All passwords saved on the BioSense Platform must be encrypted or entered every time you need to make a data connection. Here's how it's done:

Go to **Folder Shortcuts > SAS > snippets** and open the **Password_Encode.sas** utility snippet. (You open a snippet by double-clicking it or right-clicking and selecting **Open**.)

Look down several lines and you will see instructions. Follow these instructions closely.

In this line:

```
PROC PWENCODE IN='MyUnencryptedPW1' METHOD=SAS005;
```

replace **MyUnencryptedPW1** with your current AMC unencrypted password, then click the **Running Man** icon to execute this code.

When you run this, it will automatically switch to the **LOG** tab. Once there, scroll down until you see a line beginning {SAS005} followed by a long string of hexadecimals.

For example: {SAS005}A61FC901B27CCD53A87111AB1AD8A56E262525E8D2B64A6F

Paste this string to your saved copy of **User_Info.sas.** (Instructions follow below.)

NOTE: When you close the **Password_Encode.sas** file, you will be asked *Do you want to save Password_Encode.sas?* We recommend that you DO NOT save this file because it still contains your unencrypted password.

User_Info.sas

The User_Info snippet contains your username, encrypted password, and your site_short_name (e.g., AK, AL, AZ). This snippet will be integrated into various SAS macros and programs. The long string created when you run **Password_Encode.sas** beginning with {SAS005} is your encrypted/encoded password. Copy this string to the clipboard, and then open your copy of **User_Info.sas**—the one you saved to your **Files (Home)** folder—and locate the line that reads:

```
%LET PW = %STR({SAS005}XXXXXXXXXXXXX);
or your previously encrypted password
```

Replace the string within the parentheses with the string you copied to your clipboard. **Be sure to include the {SAS005} prefix.**

Before you save **User_Info.sas**, also follow the instructions to enter your SAS UserID (same as your AMC User ID) and site_short_name. (Remember to remove any < > brackets.) When all three lines have been updated, save this file to your **Files (Home)** folder. **NOTE:** You do not need to run this snippet.

For example:

```
%LET UserID=Your_AMC_UserID;
%LET PW =
%STR({SAS005}A61FC901B27CCD53A87111AB1AD8A56E262525E8D2B64A6F);
%LET site_short_name=XX;
```

NOTE: When your AMC password expires or is changed, you must re-run the **Password_Encode.sas** to encrypt your new password and save it in the **User_Info.sas** snippet in your Files (Home) folder. Then run the **DataMart_Connect.sas** program again to connect to the DataMart.

Site_Info.sas

This file contains your site-related information that other SAS programs need for processing data. Open your copy of the **Site_Info.sas** snippet you saved in your Files (Home) folder. This code snippet contains five variables or parameters that you must assign values to. If you are unsure of any of these, please contact your site administrator for the correct values. Remove the < > signs and names within them, enter the correct values for your site, and then save this file in your Files (Home) folder.

You must have **Site_Info.sas** saved in your Files (Home) folder and make sure its contents are up to date. Once saved, the location for the **Site_Info.sas** snippet should appear with the **User_Info.sas** snippet (Figure 20).

NOTE: Other SAS programs use variables stored in **Site_Info.sas** (e.g., programs distributed by the NSSP Analytic Data Management team). Blank fields can trigger SAS errors.

Figure 20. Save snippets in Files (Home) folder.

Create_Site_Folder_Shortcut.sas

This snippet will connect you to your site's shared folder. After initial access and execution, the shared folder will always be available when you log in to SAS Studio. (This program should only be run one time.)

NOTE: You must run the **Create_Site_Folder_Shortcut.sas** snippet *before* your site folder will appear.

Setting Up the Shared Folder—Follow these steps after you update and save **User_Info.sas** and **Site_Info.sas** in your Files (Home) folder:

1. Open the **Create_Site_Folder_Shortcut.sas** snippet in the SAS snippets folder, and then run it. (Click the **Running Man** icon.)
2. **Refresh the folder view**. You may need to click the refresh icon in the Navigator pane if the newly created folder shortcut does not appear (Figure 21).

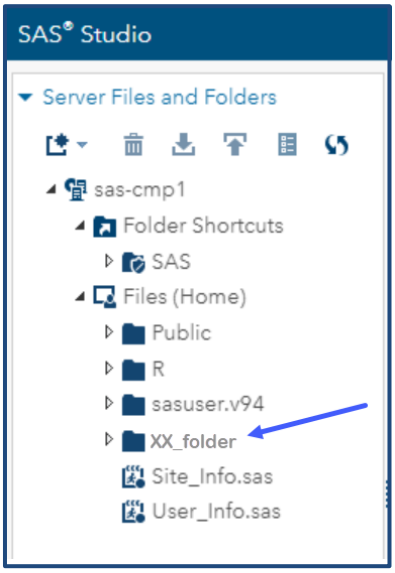


*Figure 21. Upon refresh, site folder should appear under **Files (Home)** folder.*

REMINDER: The program and data files saved in your site's shared folder may be run by all users on your site but not by users from other sites.

When run, **Create_Site_Folder_Shortcut.sas** creates a link to your site-shared folder on the file server. This program should only be run one time. When this code completes you will see the new folder under Folder Shortcuts named "XX_folder" where XX stands for your site's short name (e.g., AL, NY or FL).

On the Fileserver, its directory path is "/opt/sas/shared/sites/XX."

DataMart_Connect.sas

You will find the **DataMart_Connect.sas** program in the **sas-cmp1 > Folder Shortcuts > SAS > programs** folder. It contains the default SAS LIBNAME statement used to connect your session to the BioSense_Platform DataMart (Figure 22). You may run this program from the global SAS programs folder shortcut.

Notice that this program requires your **User_Info.sas** snippet to access your SAS UserID and encrypted password. Be sure you have saved **User_Info.sas** to your Files (Home) folder and the information in that snippet is current before running this program.

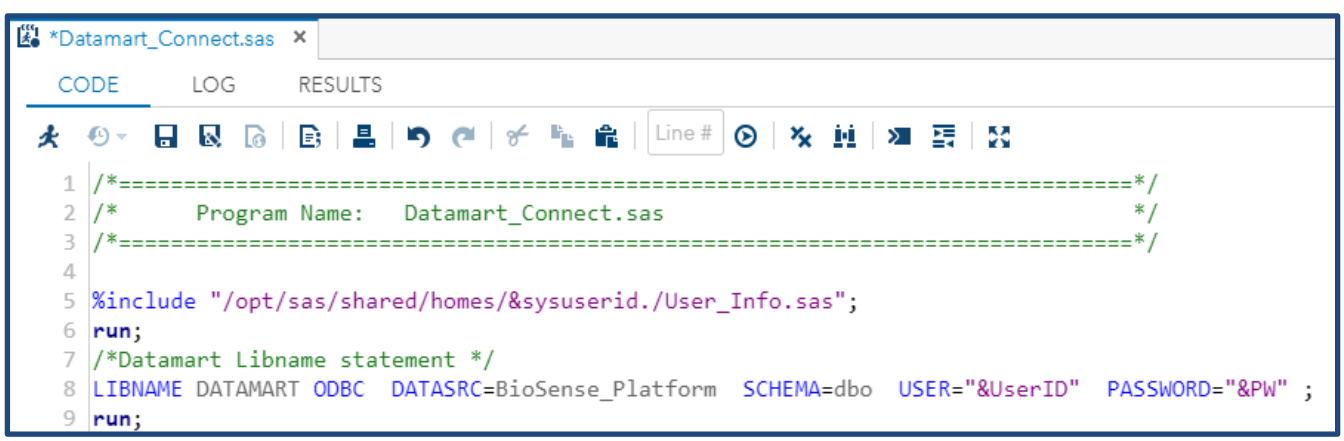


*Figure 22. **DataMart_Connect.sas** program includes saved values for username and encrypted password from **User_Info.sas**.*

Appendix B: Datasets and Variables

Datasets

Full List of Output Datasets		
File Number	**File Type**	**Description**
1	CCR	Chief Complaint Rankings
2	CLFT	Completeness details—Transposed
3	CLSUM	Completeness Summation of Completeness Details—Transposed at Overall, Vendor, Parent_Organization, Feed_Name, and Facility levels at the time aggregations were selected
4	CLVR	Completeness details by all columns selected for completeness
5	DCR	Diagnosis Code Rankings
6	EXCBEARVD	Exceptions beginning and ending Arrived_Date_Time associated with the Exceptions Processed table for C_Visit_Date range selected
7	EXCCNT	Exception Code and Exceptions Count for the C_Visit_Date range selected
8	IDSUM	Processed Table Summary Date associated with the Summary Tab of the Completeness Report
9	PRDBEARVD	Processed Table beginning and ending Arrived_Date_Time associated with the Processed Table C_Visit_Date range selected
10	RAWBEARVD	Raw Table beginning and ending Arrived_Date_Time associated with the Raw Table for C_Visit_Date range selected
11	RAWMS	Message_Status Code and Record Count for the Arrived_Date_Time associated with the C_Visit_Date range selected
12	RAWSUM	Summation of Processed, Exceptions, and Raw Visits and Records that populate the Exceptions Tab of the completeness report
13	TMSTATS	Timeliness Statistics
14	VLDT	Validity Detail—61 columns selected for Validity processing for the C_Visit_Date range selected
15	VPERR	Exceptions selected from the Exceptions Table for the C_Visit_Date range selected
16	VPERRC	Exceptions selected from the Exceptions Table for the C_Visit_Date range selected—Transposed
17	VPEXC	Visit Profile Exceptions
18	VPPRD	Visit Profile Production
19	VPRAW	Visit Profile Raw—This is a view
20	VS	Validity Summation—One file for each of the 61 columns for which validity is processed. See Standard Validity Columns for details.
21	st1plkup	St1_Processed_Lkup table column listing
22	timeagg	Year, Month, Week, and Day consolidated aggregation dates
23	tzstdst	Time Zone and Daylight-Saving Time Adjustment Table
24	st1_processed_lkup	St1_Processed_Lkup Table detail listing
25	Standard_Validity_Columns	St1_Processed_Lkup Table detail listing for the 61 columns used for Validity

Validity Processing Variables

To use SAS macros that loop through the list of variables and perform the calculations in the DQOD programming, each variable is assigned a number:

- **Vlno**: Numeric identifier for variables used in the validity portion of the DQOD program. The **vlno** is used in the **vlmin** and **vlmax** option customization described in Section 3: VLMIN and VLMAX.
- **Dqvarno:** Numeric identifier for one of the 61 columns analyzed for validity. The user will see these numbers embedded as part of the output datasets: DQ&site_id._&dqv.VS&**dqvarno**. _&fbe

Full List of Variables Available for Use in Validity Processing			
vlno	**dqvarno**	**name**	**short_name**
1	110	C_Unique_Patient_ID_Source	C_Unique_Patient_ID_Source
2	270	C_Facility_ID_Source	C_Facility_ID_Source
3	300	Facility_Type_Code	Facility_Type_Code
4	340	Sending_Facility_ID_Source	Sending_Facility_ID_Source
5	460	C_Visit_Date_Source	C_Visit_Date_Source
6	550	Discharge_Disposition	Discharge_Disposition
7	580	Patient_Class_Code	Patient_Class_Code
8	590	C_FacType_Patient_Class	C_FacType_Patient_Class
9	600	C_MFT_Patient_Class	C_MFT_Patient_Class
10	610	C_Patient_Class	C_Patient_Class
11	620	C_Patient_Class_Source	C_Patient_Class_Source
12	640	C_Death	C_Death
13	650	C_Death_Source	C_Death_Source
14	680	Admission_Type	Admission_Type
15	690	Admit_Source	Admit_Source
16	710	Hospital_Service	Hospital_Service
17	840	Administrative_Sex	Administrative_Sex
18	850	Age_Calculated	Age_Calculated
19	860	Age_Reported	Age_Reported
20	870	Age_Units_Calculated	Age_Units_Calculated
21	880	Age_Units_Reported	Age_Units_Reported
22	890	C_Patient_Age	C_Patient_Age
23	900	C_Patient_Age_Source	C_Patient_Age_Source
24	910	C_Patient_Age_Units	C_Patient_Age_Units
25	920	C_Patient_Age_Years	C_Patient_Age_Years
26	970	Race_Code	Race_Code
27	980	Race_Description	Race_Description
28	1000	Ethnicity_Code	Ethnicity_Code
29	1010	Ethnicity_Description	Ethnicity_Description
30	1040	Patient_State	Patient_State
31	1050	Patient_Zip	Patient_Zip

Full List of Variables Available for Use in Validity Processing			
vlno	**dqvarno**	**name**	**short_name**
32	1060	C_Patient_County	C_Patient_County
33	1070	C_Patient_County_Source	C_Patient_County_Source
34	1080	Patient_Country	Patient_Country
35	1170	Chief_Complaint_Type	Chief_Complaint_Type
36	1180	C_Chief_Complaint	C_Chief_Complaint
37	1190	C_Chief_Complaint_Source	C_Chief_Complaint_Source
38	1230	Diagnosis_Code	Diagnosis_Code
39	1260	Diagnosis_Type	Diagnosis_Type
40	1270	Diagnosis_Priority	Diagnosis_Priority
41	1350	Initial_Acuity_Code	Initial_Acuity_Code
42	1490	Initial_Temp	Initial_Temp
43	1500	Initial_Temp_Units	Initial_Temp_Units
44	1520	Initial_Pulse_Oximetry	Initial_Pulse_Oximetry
45	1530	Initial_Pulse_Oximetry_Units	Initial_Pulse_Oximetry_Units
46	1550	Diastolic_Blood_Pressure	Diastolic_Blood_Pressure
47	1560	Diastolic_Blood_Pressure_Units	Diastolic_BPU
48	1580	Systolic_Diastolic_Blood_Pressure	Systolic_Diastolic_BP
49	1590	Systolic_Diastolic_Blood_Pressure_Units	Systolic_Diastolic_BPU
50	1610	Systolic_Blood_Pressure	Systolic_Blood_Pressure
51	1620	Systolic_Blood_Pressure_Units	Systolic_BPU
52	1640	Pregnancy_Status_Code	Pregnancy_Status_Code
53	1670	Smoking_Status_Code	Smoking_Status_Code
54	1700	Body_Mass_Index	Body_Mass_Index
55	1720	Height	Height
56	1730	Height_Units	Height_Units
57	1750	Weight	Weight
58	1760	Weight_Units	Weight_Units
59	1790	Message_Structure	Message_Structure
60	1800	Message_Type	Message_Type
61	1830	Trigger_Event	Trigger_Event

Appendix C: Site IDs and Short Names

Site_Short_Name	Site_ID
AK	857
AL	858
AR	859
AZ	860
CA	861
CA_El_Dorado	862
CA_Fresno	949
CA_Long_Beach	950
CA_Monterey	863
CA_Napa	864
CA_Nevada	865
CA_Riverside	866
CA_Sacramento	867
CA_San_Diego	868
CA_San_Mateo	869
CA_Santa_Clara	870
CA_Santa_Cruz	871
CA_Solano	872
CA_Stanislaus	873
CA_Sutter	874
CA_Tulare	875
CA_Yolo	955
CA_Yuba	876
CO	877
CO_NCR	879
CT	880
DC	881
DE	882
DOD	883
FL	884
GA	885
HI	886
IA	887
IA_Linn_Co	888
ID	889
IL	890
IN	892
IN_Marion	893
KS	894

Site_Short_Name	Site_ID
KY	895
LA	896
MA	899
MD	901
ME	902
MI	903
MN	904
MO	906
MS	905
MT	907
NC	908
ND	909
NE	910
NH	911
NJ	912
NM	913
NV	914
NY_City	915
NY_State	916
OH	917
OK	918
OR	919
PA	920
PR	939
RI	922
SC	923
SD	924
TN	925
TX	926
TX_Region23	929
TX_Region65	928
UT	930
VA	931
VI	940
VT	933
WA	934
WI	936
WV	937
WY	938

Appendix D: Priority 1 and 2 Data Elements

PRIORITY 1 DATA ELEMENTS				
Processed Column	**Priority**	**Use Group**	**Requirement**	**HL7**
Admit_Reason_Code	**1**	CC_Diagnostic	RE	PV2-3.1, PV2-3.
Admit_Reason_Description	**1**	CC_Diagnostic	RE	PV2-3.2, PV2-3.
C_Chief_Complaint	**1**	CC_Diagnostic	CRE	NA
Chief_Complaint_Text	**1**	CC_Diagnostic	RE	OBX-2, OBX-3, OBX-5
Diagnosis_Code	**1**	CC_Diagnostic	RE	DG1.3.1
Diagnosis_Description	**1**	CC_Diagnostic	RE	DG1.3.2 DG1.3.3
C_Patient_Age	**1**	Demographics	CRE	OBX.5.1
C_Patient_Age_Years	**1**	Demographics	CRE	OBX.5.1
Patient_Zip	**1**	Demographics	RE	PID.11.5
C_Biosense_Facility_ID	**1**	Facility	CR	MSH-4.1
C_Facility_ID	**1**	Facility	CR	MSH-4.1
Facility_Type_Code	**1**	Facility	R	OBX-2, OBX-3, OBX-5
Sending_Facility_ID	**1**	Facility	R	MSH-4.1
Treating_Facility_ID	**1**	Facility	R	EVN.7.2
C_Unique_Patient_ID	**1**	ID	CR	PID.3.1
Visit_ID	**1**	ID	R	PV1.19.1
Arrived_Date_Time	**1**	Operations	CR	NA
Admit_Date_Time	**1**	Visitdate	R	PV1.44.1
C_Visit_Date	**1**	Visitdate	CR	NA
C_Visit_Date_Time	**1**	Visitdate	CR	NA
C_Death	**1**	VisitInfo	CRE	PID.30.1
C_FacType_Patient_Class	**1**	VisitInfo	CR	PV1.2.1
C_Patient_Class	**1**	VisitInfo	CR	PV1.2.1
Patient_Class_Code	**1**	VisitInfo	R	PV1.2.1

PRIORITY 2 DATA ELEMENTS				
Processed Column	**Priority**	**Use Group**	**Requirement**	**HL7**
Chief_Complaint_Code	**2**	CC_Diagnostic	RE	OBX-2, OBX-3, OBX-5
Diagnosis_Type	**2**	CC_Diagnostic	RE	DG1.6.1
Administrative_Sex	**2**	Demographics	RE	PID.8.1
Age_Reported	**2**	Demographics	RE	OBX.5.1
Age_Units_Reported	**2**	Demographics	RE	OBX.6.2
C_Patient_County	**2**	Demographics	CRE	PID-11.9
Ethnicity_Code	**2**	Demographics	RE	PID-22.1
Ethnicity_Description	**2**	Demographics	RE	PID-22.2
Patient_City	**2**	Demographics	RE	PID.11.3
Patient_Country	**2**	Demographics	RE	PID.11.6
Patient_State	**2**	Demographics	RE	PID.11.4
Race_Code	**2**	Demographics	RE	PID-10.1
Sending_Facility_ID_Source	**2**	Facility	R	MSH-4.1
First_Patient_ID	**2**	ID	R	PID-3.1
Medical_Record_Number	**2**	ID	R	PID.3.1
Message_Profile_ID	**2**	Operations	R	MSH.21.1
Message_Structure	**2**	Operations	R	MSH.9.3
Message_Type	**2**	Operations	R	MSH.9.1
Processing_ID	**2**	Operations	R	MSH.11.1
Trigger_Event	**2**	Operations	R	MSH.9.2
Version_ID	**2**	Operations	R	MSH.12.1
Message_Date_Time	**2**	Visitdate	R	MSH.7.1
Recorded_Date_Time	**2**	Visitdate	R	EVN.2.1
Discharge_Date_Time	**2**	VisitInfo	RE	PV1.45.1
Discharge_Disposition	**2**	VisitInfo	RE	PV1.36.1

Appendix E: Tips for Improving Run Efficiency

Whenever you run the DQOD programs—

- Consider using standard timeframes, such as 1 week, 1 month, or 1 quarter. This will simplify comparisons over time.
- Limit your runs to only the report needed. If you just need the Timeliness Report, do not select to run all reports. Because of the resources needed to run all reports, the general rule is "Do not generate what you do not need."
- Consider using the **Background Submit** button (see How to Submit Programs in the Background for detailed instructions). This uses available resources more efficiently and will allow you to work on other tasks in SAS Studio while it runs.
- Consider running the DQOD programs later in the day, taking into account the time difference between East Coast and West Coast.
- Keep in mind that longer timeframes require significant processing resources. You can break a long timeframe into multiple runs of shorter duration, then combine the results. Although some manual effort is required, this method is more efficient because of the additional time required to extract data over long time periods.

Warning: If you attempt to match numbers from aggregated short timeframe results to the results from one longer run, they may not match. This is due to message arrived dates that are used when collapsing messages into patient visits records.