PEPFAR
U.S. President's Emergency Plan for AIDS Relief

DATIM Data Genie Application

Operating Unit by Implementing Mechanism (OU by IM), Priority Sub-National Unit by IM (PSNU by IM), Site by IM

User's Guide and Data Dictionary

Revised October 8th, 2020

PEPFAR
U.S. President's Emergency Plan for AIDS Relief

Table of Contents

Introduction

The DATIM Data Genie Application (Genie) contains three reports with in-process, structured MER data:

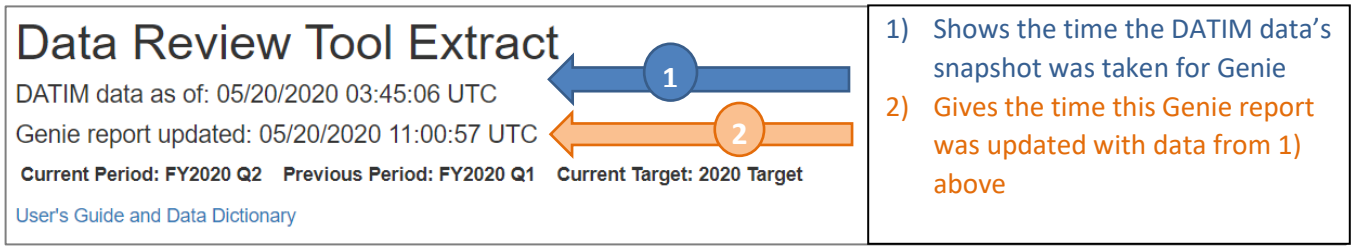
- Operating Unit by Implementing Mechanism (OU by IM),
- Priority Sub-National Unit by IM (PSNU by IM), and
- Site by IM.

These reports include the same data elements as the MER Structured Datasets (MSDs). The difference is that the MSD only contains approved data while the above-mentioned reports contain in-process and approved data. These three reports were created to reduce data access burdens, provide datasets in a standardized format, and facilitate analysis of in-process data. Users have the option to limit the number of rows returned by applying different filter options, and the extracted data (in tab-delimited format) can be imported into Excel or any statistical software package for analysis.

Data Availability

For "daily" reports, the Genie application synchronizes daily to reflect data entered in the DATIM system (www.datim.org). For "frozen" reports, the Genie application is updated after a DATIM data entry closure and is available on the date of the PEPFAR Panorama release. On the Genie landing page, there are 2 messages to take note of:

1. The "DATIM data as of" message indicates the last time DATIM data was pulled for the selected Genie report.
2. The "Genie report updated" message indicates the time the Genie application was updated to use the DATIM data that was taken from the time stated in the "DATIM data as of" message for the selected Genie report.



The following table shows when the synchronization process from DATIM to the Data Genie begins and ends in several time zones:

Sync from DATIM to Genie				
	Time Zone	DATIM backup	Update Start	Update End
Daily	UTC	3:45AM	8:00AM	12:00PM
	US Eastern Time	10:45PM EST; 11:45PM EDT	3:00AM EST; 4:00AM EDT	7:00AM EST; 8:00AM EDT
	Bangkok, Thailand	10:45AM	3:00PM	7:00PM
	Johannesburg, South Africa	5:45AM	10:00AM	2:00PM

**Eastern Standard Time (EST) is 5 hours behind UTC. Between the second Sunday in March and the first Sunday in November, the US observes Daylight Saving Time (EDT), which is 4 hours behind UTC.*

Data Availability Example

If a user enters data into DATIM before 0300 UTC, it should be included in Genie by 1200 UTC on the same day. This assumes there are no issues with the process that pulls data from DATIM into Genie.

However, if users enter data after 0300 UTC, they will need to wait until 1200 UTC on the following day to see their data in Genie. This is because the process for pulling DATIM data into Genie runs once per day.

The process to take data from DATIM to Genie occasionally has issues requiring a restart. This may cause a discrepancy between the data in Genie and the dates listed for each Genie report. If you think this is the case, please let us know by filing a technical support ("helpdesk") ticket here: https://datim.zendesk.com/hc/en-us.

Analytic Dataset Features

The OU by IM, PSNU by IM, and Site by IM reports include the following data features:

- A PEPFAR organizational hierarchy comprising levels relevant to program priorities: Region, Operating Unit, Country, Prioritization-level SNU (PSNU; as applicable), Community (as applicable), and Facility (as applicable)
- A column listing the specific indicator name (e.g., HTS_TST), regardless of disaggregate, to allow for easy display and filtering of data in analysis
- A column listing the PSNU Prioritization, which allows for easy filtering or grouping by prioritization level
- Category Option Combo Names which list all the disaggregation categories from the DATIM data entry screen. These are broken out into separate columns for age, sex, HIV status, TB status, CX status, modality, and other disaggregates
- Four age category columns for easy analysis: Age as Entered, Trends Fine, Trends Semi-Fine, and Trends Coarse age
- A filtering column to flag the Military sub national unit (i.e., _MIL SNU) for easy identification, removal, or analysis
- Target, Result, and Cumulative values for current Fiscal Years including the in-process (unapproved) reporting period
- Multi-OU filter selection capability in the PSNU by IM and OU by IM reports
- Frozen reports at the OU, PSNU, and Site by IM levels
- The Genie application includes columns to distinguish the data's approval level and approval descriptions

Role-Based Access to Data (Security Trimming)

When a user runs a report in Genie, the data included in the Genie report may be limited based on two factors described below. Limiting report data in this way is known as security trimming.

1) The permissions associated with an individual DATIM account, which are also called user roles. The different user roles are: Global, Inter-Agency, Agency, and Partner. If you are not sure of the user role associated with your DATIM account, please submit a DATIM Support Ticket.

2) The approval level of the data. MER data is entered by Implementing Partners and moves through an approvals process described in the table below. As of October 1, 2018, data with Approval Levels 1, 2 or 3 is considered "Approved," while data with Approval Levels 4, 5 and 99 is considered "Unapproved".

The table below shows which step each numerical approval level maps to in the data approval process and indicates which data each user role/type of account is able to view based on the approval levels. "Yes" means the data is viewable to users with that user role, while "No" means the data is not viewable to users with that user role.

Approval Level	Approval Level Description	Approved?	User Type			
			Global User	Inter-Agency User	Agency User	Partner User
1	Submitted by Global	Approved	Yes	Yes	Yes	Yes*
2	Submitted by HQ Agency	Approved (now used for Expenditure Reporting (ER))	Yes	Yes	Yes	Yes*
3	Submitted by Inter-Agency	Approved	Yes	Yes	Yes	Yes*
4	Submitted by Funding Agency	Unapproved	No	Yes	Yes	Yes*
5	Submitted by Implementing Partner	Unapproved	No	No	Yes	Yes*
99	Not submitted to workflow	Unapproved	No	No	Yes	Yes*

** Partner level users can only view their own data. Regardless of approval level, users with partner level accounts can only view their own data.*

The table below shows how security trimming is applied while considering both user roles and approval data of the data. While Global Users can see only approved data for all OUs, Inter-Agency users can see all approved and some unapproved data for their OU, Agency users can see unapproved data for their Agency as well as approved data for their OU, and Partners users can see all data for their IP, whether approved or unapproved.

User Role	Approved Data	Unapproved Data
Global	All OUs	None
Inter-Agency	User OU	Only Level 4 unapproved data for user OU
Agency	User OU	All unapproved for user agency
Partner	User Partner	All unapproved for user partner

Partners have access to the following Genie reports: The Data Review Tool (DRT), OU by IM, PSNU by IM, Site by IM, the SIMS Country x IM and SIMS 4 Site x IM.

Role-Based Access to Data (Security Trimming) Implications and Examples

Security trimming has the following implications:

1) Two users with different account types could run the same report in Genie at the same time but see different results. For example, a Partner user running a Site by IM report would only be able to see data for their Implementing Partner, but an Agency user running the same report would be able to see data for all IPs under that Agency.

2) Two users with different account types could run the same report in Genie at the same time but different results, even when looking at data for the same partner to which they both have access. For example, if an Implementing Partner has submitted data for the Funding Agency's approval (labeled with approval level 5) during a given period, then Inter-Agency and Global users would not be able to view or download that data for that period, but Agency and Partner users would.

However, once data is submitted by Inter-Agency (approval level 3), the data is considered approved for Global, Inter-Agency, Agency, and Partner users. As a result, these users will be able to view or download that data for a given period. Please note that, although Partner users can only see the data they have submitted, Agency users can see approved data for all agencies as well as unapproved data for their agency.

3) If any data is left unapproved (at levels 4, 5, or 99) following data entry close, an Inter-Agency, Agency, or Partner user would see different data in DATIM.org or a live Genie report compared to Panorama or a frozen Genie report. This is because Panorama and frozen Genie reports only reflect approved data. If you think you may have unapproved data from a prior reporting period, please submit a DATIM Support Ticket and copy your PEPFAR Program Manager. Further guidance on the data approval process can be found here on the DATIM support page.

Derived Data in the PEPFAR Data Analytics Platform - iHUB

Calculated Indicators

The PEPFAR Data Analytics Platform – Integration Hub (iHUB) creates calculated indicators based on data from DATIM. Calculated indicators are created for select indicators in the iHUB for Panorama, Genie, and other downstream systems. The calculations automatically create (a) data row(s) based on a specific disaggregate (e.g., positive or negative or for a specific set of disaggregates). These calculated indicators include:

- AGYW_PREV (Numerator)
- CXCA_SCRN (denominators)
- CXCA_SCRN_POS
- CXCA_TX (denominators)
- GEND_GBV_PhysicalEmotionalViolence
- GEND_GBV_SexualViolence
- HRH_CURR (DSD+TA, ClinicalCadre)
- HRH_CURR (DSD+TA, ClinicalSupportCadre)
- HRH_CURR (DSD+TA, LayCadre)
- HRH_CURR (DSD+TA, ManagementCadre)
- HRH_CURR (DSD+TA, OtherCadre)
- HRH_CURR (DSD+TA, SocialServiceCadre)
- HTS_INDEX_KNOWNPOS
- HTS_INDEX_NEG
- HTS_INDEX_POS
- HTS_TST_NEG
- HTS_TST_POS
- LAB (various)
- OVC_HIVSTAT_NEG
- OVC_HIVSTAT_POS
- OVC_SERV_ACTIVE
- OVC_SERV_GRADUATED
- OVC_SERV_OVER_18
- OVC_SERV_UNDER_18
- PMTCT_ART (denominators)
- PMTCT_EID_LESS_EQUAL_TWO_MONTHS
- PMTCT_EID_TWO_TWELVE_MONTHS
- PMTCT_EID_POS_2MO
- PMTCT_HEI_POS_2MO
- PMTCT_HEI_POS_ART
- PMTCT_STAT_POS
- TB_ART (denominators)
- TB_STAT_NEG_NEWLYIDENTIFIED_NEGATIVE
- TB_STAT_POS
- TB_STAT_POS_KNOWNATENTRY_POSITIVE
- TB_STAT_POS_NEWLYIDENTIFIED_POSITIVE
- TX_TB_D_NEG
- TX_TB_D_POS
- VMMC_CIRC_FollowUp
- MostCompleteAgeDisagg (MCAD; listed as a disaggregate, not an indicator)

Users who are interested in analyzing data using calculated indicators listed above can find them in the "Indicator" column of the MSD or Genie report.

Calculated Total Numerators and Total Denominators

Starting with FY2019 results data, all Total Numerators and Denominators are auto-calculated from data entered into specified disaggregates (*refer to the* "*MER Indicator Reference Guide (Version 2.3 FY19)*") rather than entered through the DATIM data entry screens.

Prior to FY2019, some Total Numerators and Denominators were entered into DATIM and some were provided through calculation. Unless the MER Indicator Reference Guide (such as "*MER Indicator Reference Guide (Version 2.3 FY19)*") specifies that an indicator should not have a total (e.g. SC_STOCK), a Total Numerator is provided for each Indicator by the PEPFAR Data Hub (PDH) for Genie, Panorama, and other systems downstream from DATIM. Some indicators also have denominator data and each of these will have a Total Denominator provided in PDH.

Additional Information for Analyses

Guidance for Data Extracts that contain Age Band Columns

Age Bands have become finer over time. In other words, the program's age bands have evolved from their baseline state in the years 2015 and 2016 to capture data across more age ranges. As an example, the age band 25-49 existed between 2015 and 2017, but finer age bands were created in 2018, so users can now access data within the 25-29, 30-34, 35-39, and 40-49 age bands.

To account for these changes year by year, Genie reports and MSDs contain the columns "TrendsSemiFine," "TrendsFine," "AgeAsEntered," and "TrendsCoarse." These columns may include age bands added during a given year and may include age bands from a previous year.

To analyze age-disaggregated results data for FY2017, FY2018, FY2019+:

- **FY 2017 – FY 2018Q1**: use age bands in "TrendsSemiFine"
- **FY 2018Q2 – FY 2018Q4**: use age bands in "TrendsFine"
- **FY 2019-current**: use age bands in "AgeAsEntered"
- **TrendsCoarse** can be used for all coarse analysis for <15, 15+ and Unknown ages.

Please refer to the table below to see which age bands were collected in each year for most indicators[1].

[1] Indicators such as OVC_SERV and VMMC_CIRC and a few others may have slightly altered age bands used for data entry in DATIM. More information can be found on DATIM support

Table 1: Evolution of PEPFAR Finer Age Bands for Results Reporting

Evolution of PEPFAR Finer Age Bands for Results Reporting							
FY 2015 – FY 2016		FY 2017		FY 2018		FY 2019 – FY 2021	
Age Band	*Sex*	*Age Band*	*Sex*	*Age Band*	*Sex*	*Age Band*	*Sex*
<1	M / F	<1	None	<1	None	<1	M / F
1-4	M / F	1-9	None	1-9	None	1-4	M / F
5-9	M / F					5-9	M / F
10-14	M / F	10-14	M / F	10-14	M / F	10-14	M / F
15-19	M / F	15-19	M / F	15-19	M / F	15-19	M / F
20-24	M / F	20-24	M / F	20-24	M / F	20-24	M / F
25-49	M / F	25-49	M / F	25-29	M / F	25-29	M / F
				30-34	M / F	30-34	M / F
				35-39	M / F	35-39	M / F
				40-49	M / F	40-44	M / F
						45-49	M / F
50+	M / F	50+	M / F	50+	M / F	50+	M / F

Most Complete Age-Sex Disaggregation for FY15-FY17

The Most Complete Age-Sex Disaggregate (MCAD) is a calculated disaggregate used for HTS_TST, HTS_TST_NEG, HTS_TST_POS, TX_NEW, and TX_CURR. ***While using the MCAD is the preferred method of analyzing data for these indicators when FY 2015-2017 data is included in an analysis, MCAD is not necessary when analyzing data that do not include any of those fiscal years***.

Most Complete Age-Sex Disaggregation: Origins

PEPFAR developed the MCAD algorithm because it was possible for an implementing mechanism (IM) to enter both Semi-Fine and Coarse data at a single site before FY18. It was unclear whether such data could be appropriately summed together, or whether the Semi-Fine and Coarse data were duplicative (e.g., if an 18-year-old beneficiary was accidently recorded both in the 15-19 and the 15+ age bands).

The MCAD algorithm was applied to the data in order to minimize the risk of over-counting age-disaggregated values.[2] The algorithm takes data from across the different age disaggregates (Semi-Fine or Coarse) and then stores an additional calculated disaggregate value, which is displayed as <15/15+ (Male, Female, or Unknown).

For FY2017 Analyses: When analyzing age/sex disaggregated data in FY17 (or earlier), the Most Complete Age Disaggregate (MCAD) is the preferred method for analyzing <15 or 15+ data. To use MCAD to analyze data in a pivot table, users should conduct analyses by:

1. filtering the Disaggregate column to contain: "MostCompleteAgeDisagg"

[2] Prior to FY2018, the MCAD selected the most complete age-sex data, selecting between Semi-fine and Coarse disaggregate choices for each site by IM for DSD or TA. The value is then stored as an additional calculated disaggregate value which is displayed as <15/15+, Male/Female/Unknown. This is based on a 'Proximal-Maximal' algorithm that chooses only one disaggregate, either Semi-Fine or Coarse entered at the Site-IM-DSD/TA level (and modality in case of HTS_TST), depending on which one is closer to the total numerator (proximal), or the in the absence of total numerator, the maximal value. In the case of FY17 Targets for HTS_TST and HTS_TST_POS, disaggregates are chosen for the MCAD according to MER Guidance about which targets should have been entered for DREAMS vs non-DREAMS districts.

2. using the indicator column to filter for the testing or treatment indicator of interest, and
3. using the "TrendsCoarse" column as a row, column, or filter in an Excel pivot table to display age bands.

Note: when using this method, you should not include the "standardizedDisaggregate" column in your pivot table.

Standardized Disaggregation

The Data Genie's reports and the MSDs contain a "standardizedDisaggregate" column, the value of which is based on the "disaggregate" column. The "standardizedDisaggregate" can:

1. facilitate easier analysis of the MER 2.0 HTS_TST indicator and,
2. indicate which disaggregates are comparable over different reporting periods, which is not necessarily intuitive, since some comparable disaggregates have different names.

Essentially, the "standardizedDisaggregate" column creates its values using the values in the "disaggregate" column, and then uses a common disaggregate name so data collected under the same disaggregate type may be meaningfully compared. As a result, users can analyze data using a single, standardized disaggregate rather than multiple disaggregates. Please see the following section for examples.

Standardized Disaggregate Examples

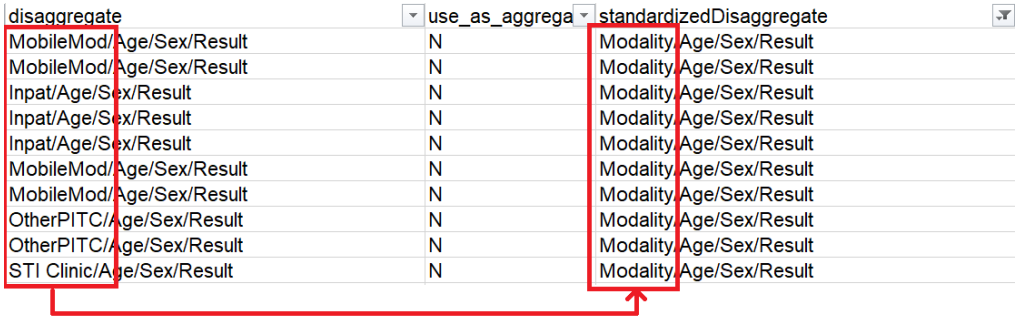
Some of the standardized disaggregates are described below. Please see Table 1 in Appendix A for additional details.

1. **What Periods:** FY2017
   Standardized Disaggregates: "Age" replaces "AgeLessThanTen" or "AgeAboveTen"
   Why: Facilitates analyses since FY2017 results and FY2018 targets are the only years with disaggregates that contained "AgeLessThanTen" or "AgeAboveTen" (screenshot below).

disaggregate	use_as_aggrega	standardizedDisaggregate
OtherPITC/AgeLessThanTen/Result	N	Modality/Age/Sex/Result
VCTMod/AgeLessThanTen/Result	N	Modality/Age/Sex/Result
VCTMod/AgeLessThanTen/Result	N	Modality/Age/Sex/Result
OtherPITC/AgeLessThanTen/Result	N	Modality/Age/Sex/Result
OtherPITC/AgeLessThanTen/Result	N	Modality/Age/Sex/Result
OtherPITC/AgeLessThanTen/Result	N	Modality/Age/Sex/Result

2. **What Periods:** All periods
   Standardized Disaggregates: The word "Modality" replaces each specific modality of HTS_TST, HTS_TST_POS, HTS_TST_NEG. For example, "Index/Age/Sex/Result" becomes "Modality/Age/Sex/Result".
   Why: Gives the ability to filter for all data that has a modality. Please note that results reported under the Malnutrition and Pediatric service delivery modalities are included under the Modality/Age

Aggregated/Sex/Results standardized disaggregate and the Modality/MostCompleteAgeDisagg standardized disaggregates.



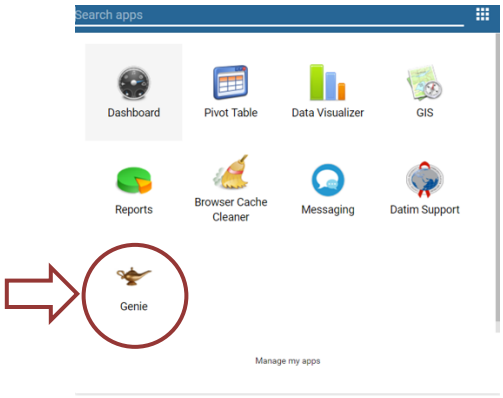
disaggregate	use_as_aggrega	standardizedDisaggregate
MobileMod/Age/Sex/Result	N	Modality/Age/Sex/Result
MobileMod/Age/Sex/Result	N	Modality/Age/Sex/Result
Inpat/Age/Sex/Result	N	Modality/Age/Sex/Result
Inpat/Age/Sex/Result	N	Modality/Age/Sex/Result
Inpat/Age/Sex/Result	N	Modality/Age/Sex/Result
MobileMod/Age/Sex/Result	N	Modality/Age/Sex/Result
MobileMod/Age/Sex/Result	N	Modality/Age/Sex/Result
OtherPITC/Age/Sex/Result	N	Modality/Age/Sex/Result
OtherPITC/Age/Sex/Result	N	Modality/Age/Sex/Result
STI Clinic/Age/Sex/Result	N	Modality/Age/Sex/Result

3. **What Periods:** Starting in FY2019 Q1 and forward
Standardized Disaggregates: Adds numerals to HTS_INDEX disaggregates
    1:Age/Sex/IndexCasesOffered
    2:Age/Sex/IndexCasesAccepted
    3:Age Aggregated/Sex/Contacts
    4:Age/Sex/Result
Why: Allows Panorama as well as analysts to sort the HTS_INDEX data into the order needed to show the cascade from offering index testing services to the results of the person identified via index testing.

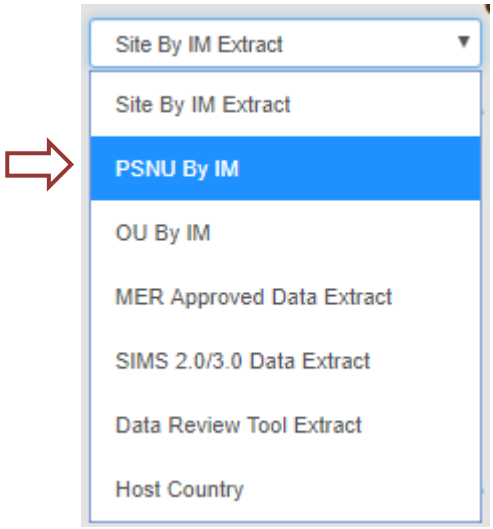
How to Extract for OU by IM, PSNU by IM and Site by IM

(Only one example is shown, but the same method applies to all three data extracts)

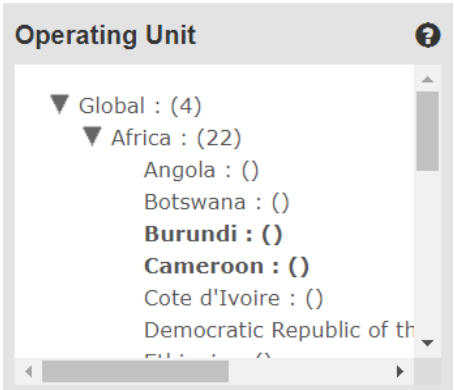
1. Login to https://datim.org
2. Select '**Genie'** app from the list of available apps.



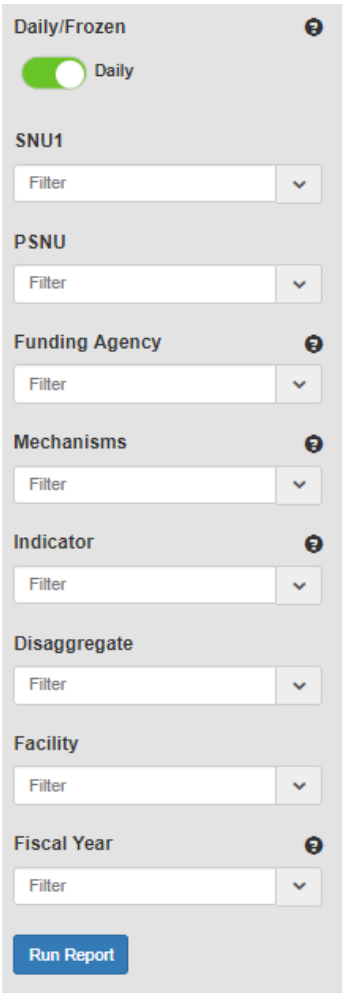
3. Select the desired report in the drop down. Visual shows the '**PSNU by IM Extract**' option.



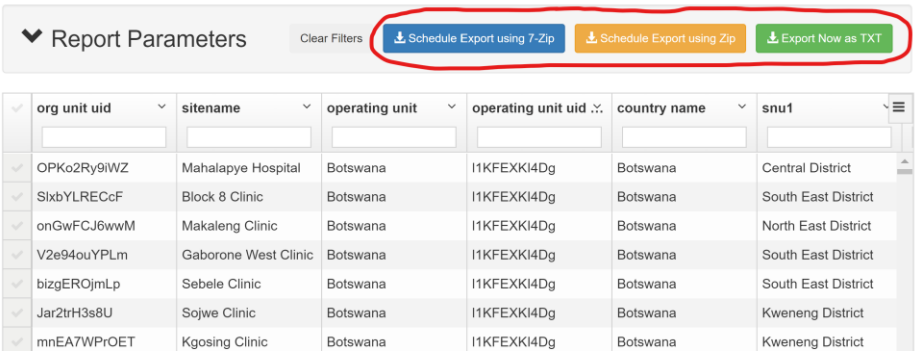
4. Select the Operating Unit(s) of choice from the organizational hierarchy tree. (For the 'OU x IM' and ' PSNU x IM' extracts, you may select multiple OUs.) Click any OU to select it, and click again to de-select it. The selection on the right would result in an export with data from both Burundi and Cameroon. Additionally, multiple OU's can be processed by selecting one or more of the higher level categories – Global, Africa, etc. Please exercise caution in selecting multiple OUs, as this may result in large datasets that take additional time to run.



5. Click on the '**Run Report**' option after refining your search results by the following (if applicable): SNU1, PSNU, Funding Agency, Mechanisms, Indicator, and Disaggregate filters. The 'Site x IM' extract will have an additional filter for Facility.



6. Users have the option to extract the data returned by selecting one of three export options (see below). This will create a tab-delimited output of all the rows returned, regardless of what is displayed on the Genie's main table in the browser window. In other words, even if you can only see one page's worth of observations in the browser window, either one of the export options will allow you to download a full dataset with the filters you've applied.



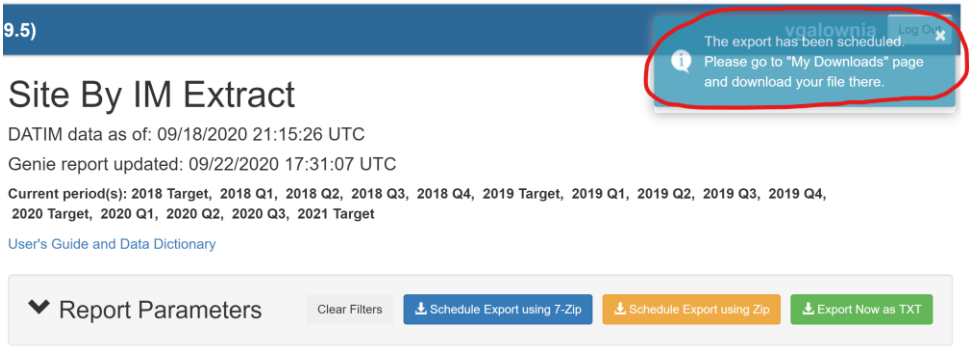
org unit uid	sitename	operating unit	operating unit uid	country name	snu1
OPKo2Ry9iWZ	Mahalapye Hospital	Botswana	I1KFEXKI4Dg	Botswana	Central District
SlxbYLRECcF	Block 8 Clinic	Botswana	I1KFEXKI4Dg	Botswana	South East District
onGwFCJ6wwM	Makaleng Clinic	Botswana	I1KFEXKI4Dg	Botswana	North East District
V2e94ouYPLm	Gaborone West Clinic	Botswana	I1KFEXKI4Dg	Botswana	South East District
bizgEROjmLp	Sebele Clinic	Botswana	I1KFEXKI4Dg	Botswana	South East District
Jar2trH3s8U	Sojwe Clinic	Botswana	I1KFEXKI4Dg	Botswana	Kweneng District
mnEA7WPrOET	Kgosing Clinic	Botswana	I1KFEXKI4Dg	Botswana	Kweneng District

Schedule Export using 7-Zip	Uses the 7-Zip format. This method is usually fastest, due to a smaller download size. For very large downloads, this is the best option to use if your computer has the software installed to support it. The software needed to support this file type can be found at www.7-zip.org.
Schedule Export using Zip	Uses the standard Zip format.
Export Now as TXT	Creates a standard, uncompressed text file.

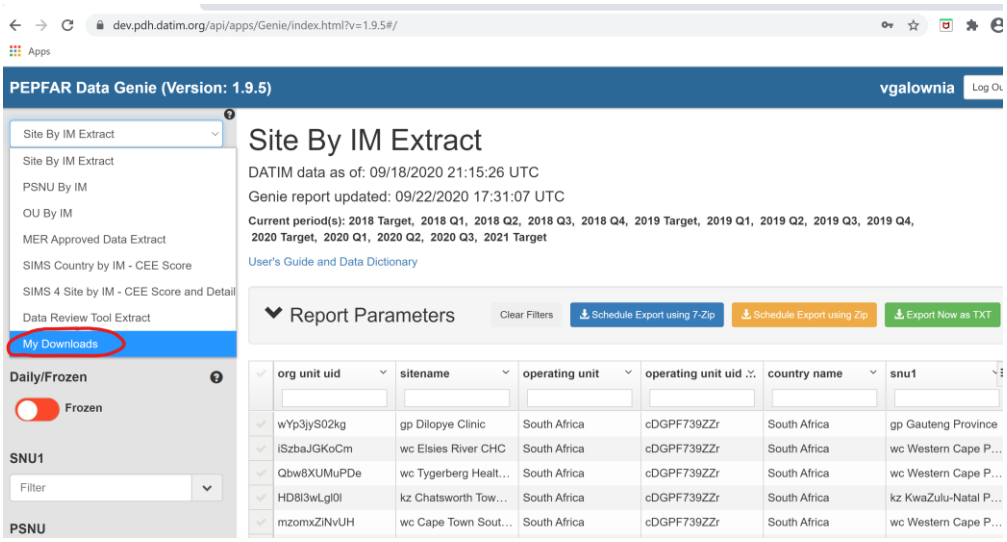
Note: If the download contains more than 9.5 million rows, the *Export Now as Text* option (green button) will be disabled.

Upon selecting either of the *Schedule Export* options, a message will display directing you to a new *My Downloads* page.

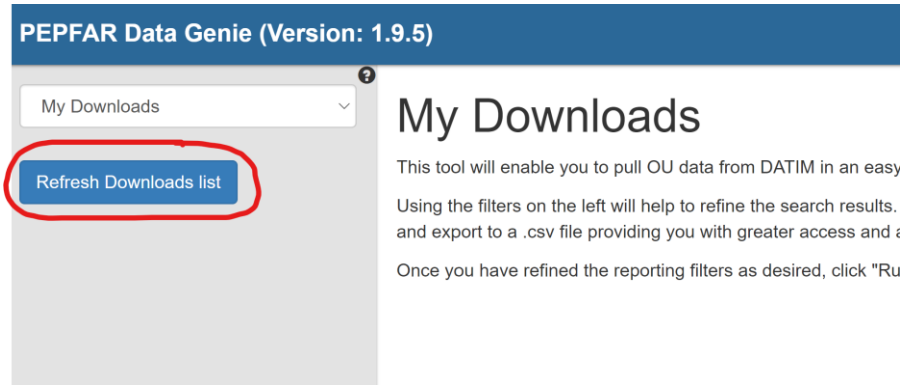
At this point, you can logoff Genie, if desired. The requested download will automatically be available once complete. It will remain available for the next 48 hours.



In the Report dropdown selector, select *My Downloads*. This will be the last option in the selection window.

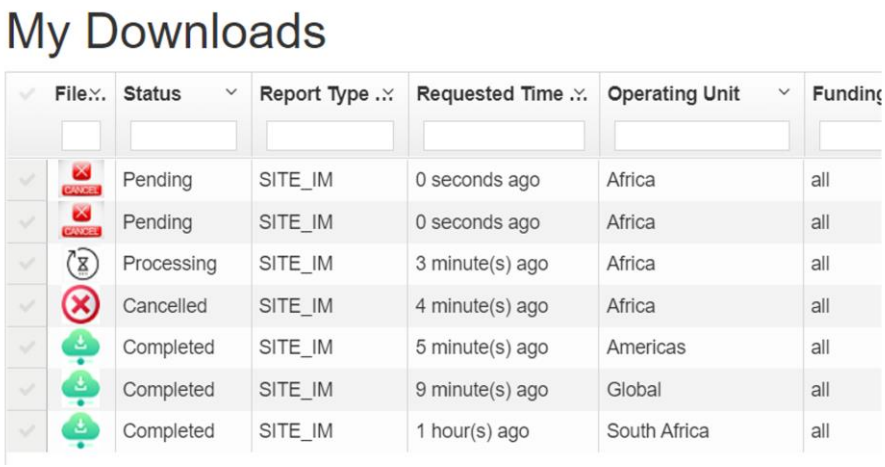


Upon selecting *My Downloads*, the My Downloads page will display. Click the *Refresh Downloads* list button to display your downloads.

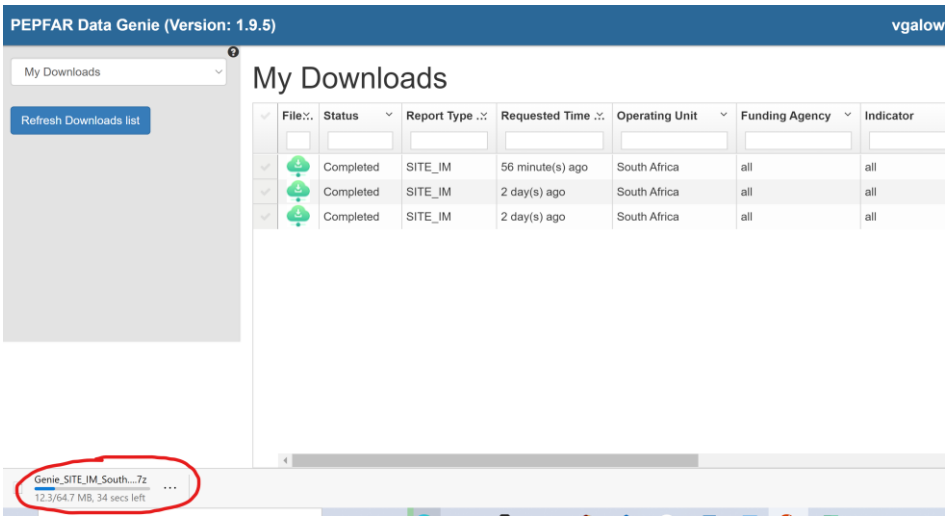


The Downloads page will display all downloads requested within the last 48 hours.

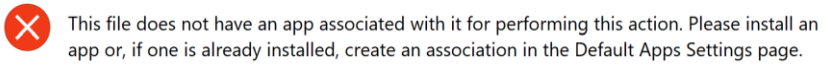
- Download status values will be Pending, Processing, Completed or Canceled.
- A download can be canceled while in the Pending state only. To do so, click the red cancel icon at the far left.
- Once a download is completed, it can be retrieved by clicking the green download icon.



Upon clicking the download icon, the export status will display in the lower left corner of the screen. Once complete, click *Open File*.



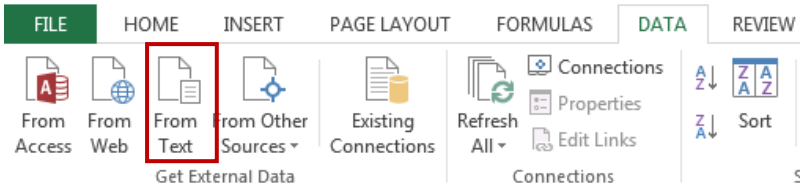
Note: If the 7-Zip option was used, and an error message that is similar to that shown below displays, the needed software for the 7-Zip format is likely not installed on your computer. The software needed to support this file type can be found at www.7-zip.org.



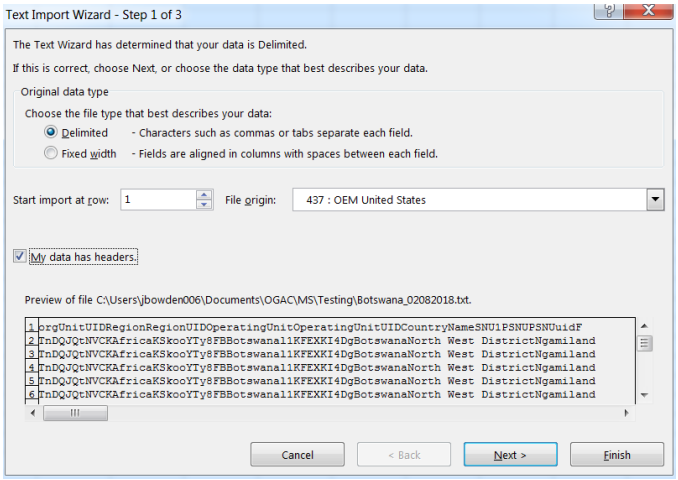
How to Use Excel to Analyze Data from the OU by IM, PSNU by IM or Site by IM Datasets for Excel 2016 and Earlier

When using Excel, users must import the data using Excel's built-in text import functionality to prevent data from importing incorrectly. This can be done by following these steps:

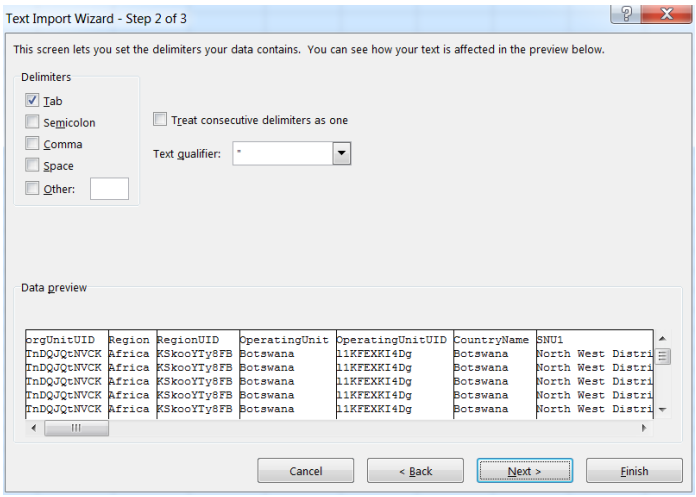
1. Open a blank Excel workbook
2. Select Data in the top ribbon
3. Select "From Text" in the Get External Data section of the "Data" tab



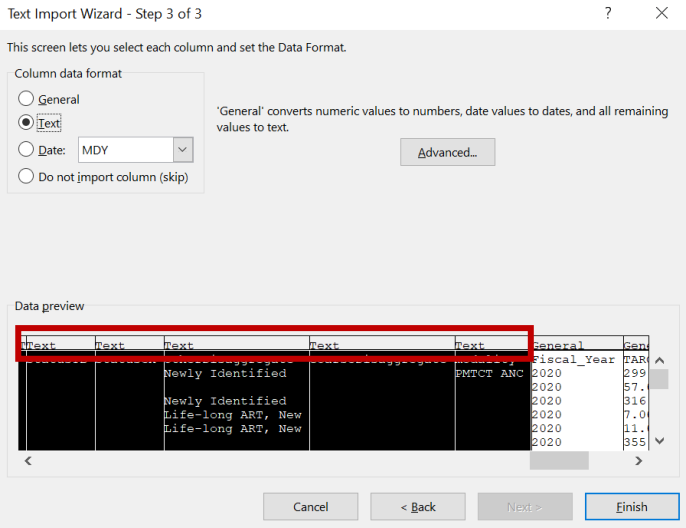
4. In the window that opens, select the file downloaded from Genie that you want to use.
5. In the Text Import Wizard window, select "Delimited" and "My data has headers". Then click "Next".



6. In the next window (Step 2 of 3) select "Tab" in the list of available delimiter options and click "Next".



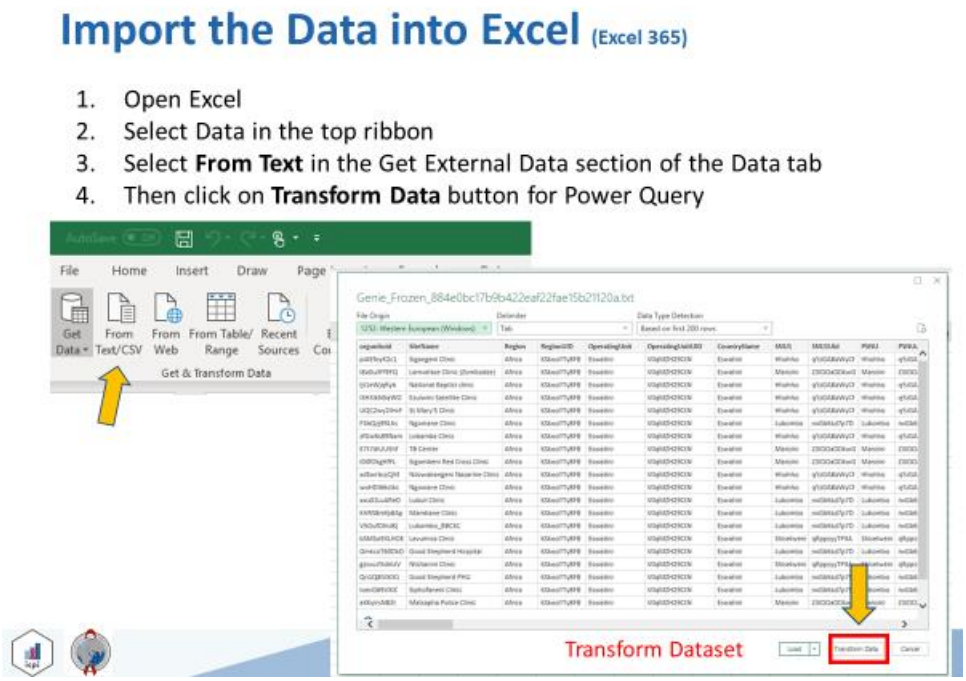
7. In the final window of the Text Import Wizard (Step 3 of 3), select the first column through the column titled "modality" (do not select any of the columns with result or target values) and click Text in the list of Column data format options. ***This will prevent age ranges from being read as dates***.



8. Click "Finish" and your data will import.

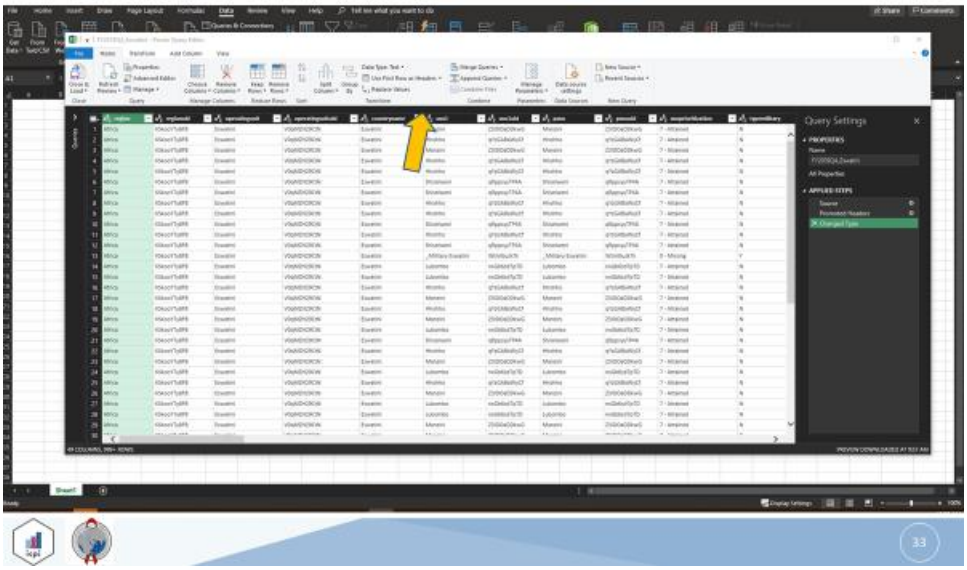
How to Use Excel to Analyze Data from the OU by IM, PSNU by IM or Site by IM Datasets for Later Versions of Excel

ICPI, the Interagency Collaborative for Program Improvement, has provided the slides below for users importing data with later versions of Excel.



Import the Data into Excel (Excel 365)

5. Make sure ***FiscalYear*** is text and the ***Qtr*** columns are numeric



Import the Data into Excel (Excel 365)

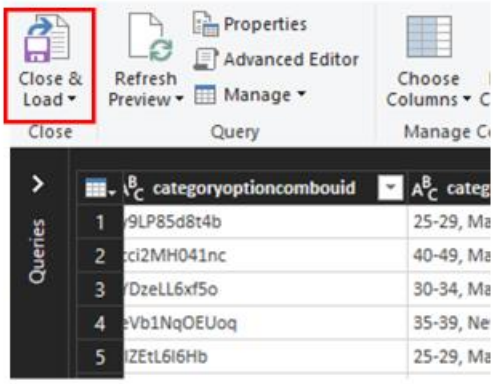
5. Make sure ***FiscalYear*** is text and the ***Qtr*** columns are numeric

ageasentered	trendsfine	trendssemifine	trendscoarse
25-29	25-29	25-49	15+
40-49	40-49	25-49	15+
30-34	30-34	25-49	15+
35-39	35-39	25-49	15+
25-29	25-29	25-49	15+
20-24	20-24	20-24	15+
15-19	15-19	15-19	15+
05-09	01-09	01-09	<15
30-34	30-34	25-49	15+
25-29	25-29	25-49	15+
10-14	10-14	10-14	<15
30-34	30-34	25-49	15+
25-29	25-29	25-49	15+
Unknown Age	Unknown Age	Unknown Age	Unknown Age
<01	<01	<01	<15

34

Import the Data into Excel (Excel 365)

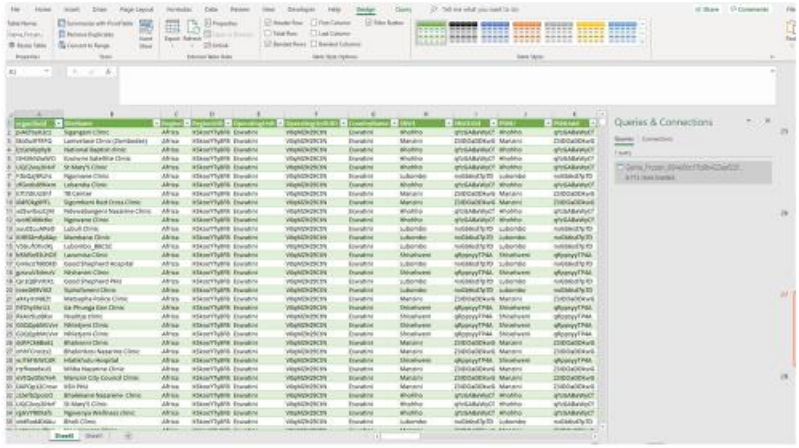
6. When columns are the correct type, click on **close & load** button at top left



35

Import the Data into Excel (Excel 365)

7. Data will load as a table and query. Import is done.

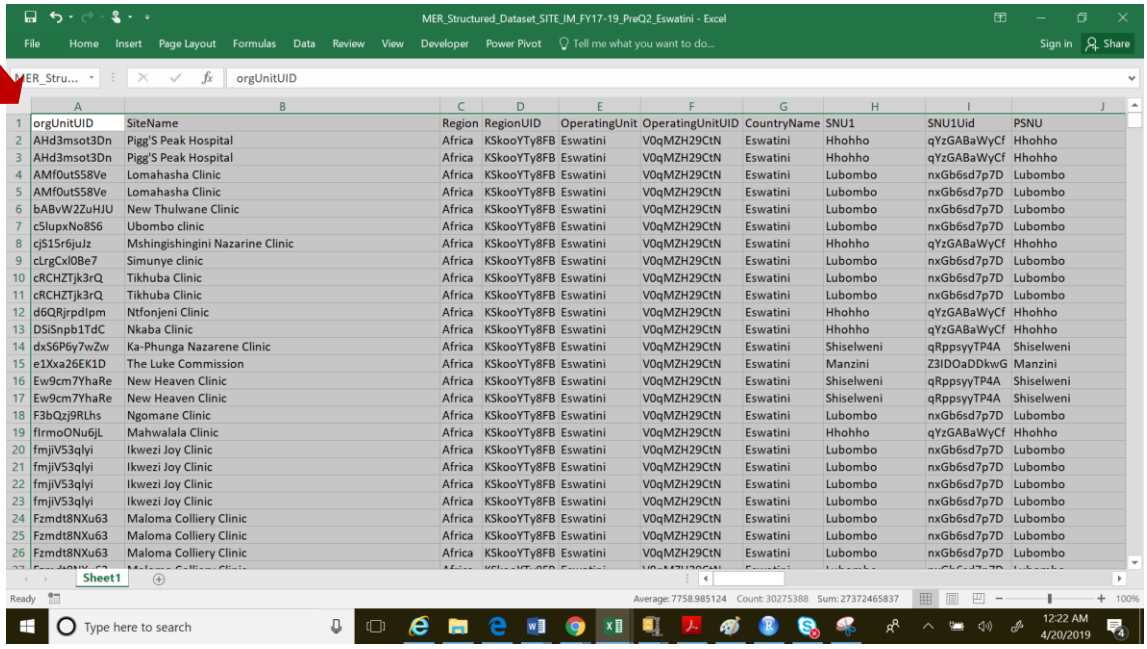


36

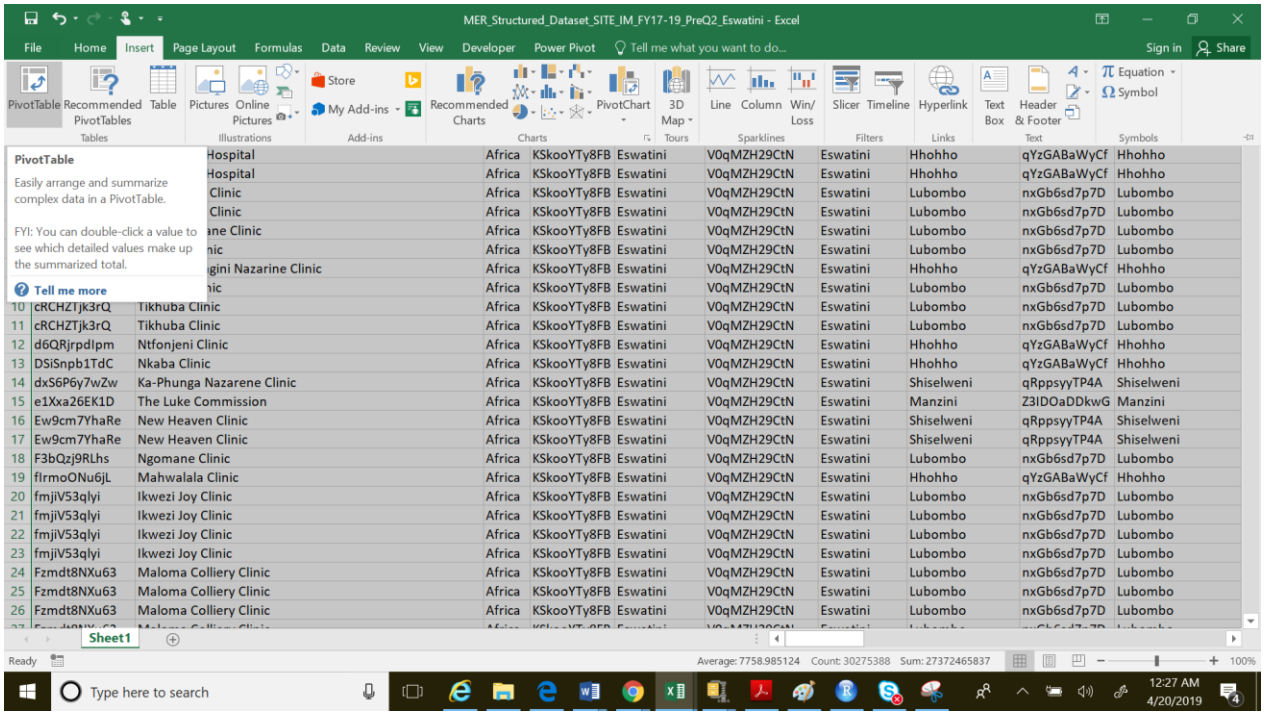
How to Create Pivot Tables from Genie Extracts

Pivot Tables are useful for summing or counting data across categories. There are many ways to build pivot tables so that they will show certain data. The process below outlines one possible approach.

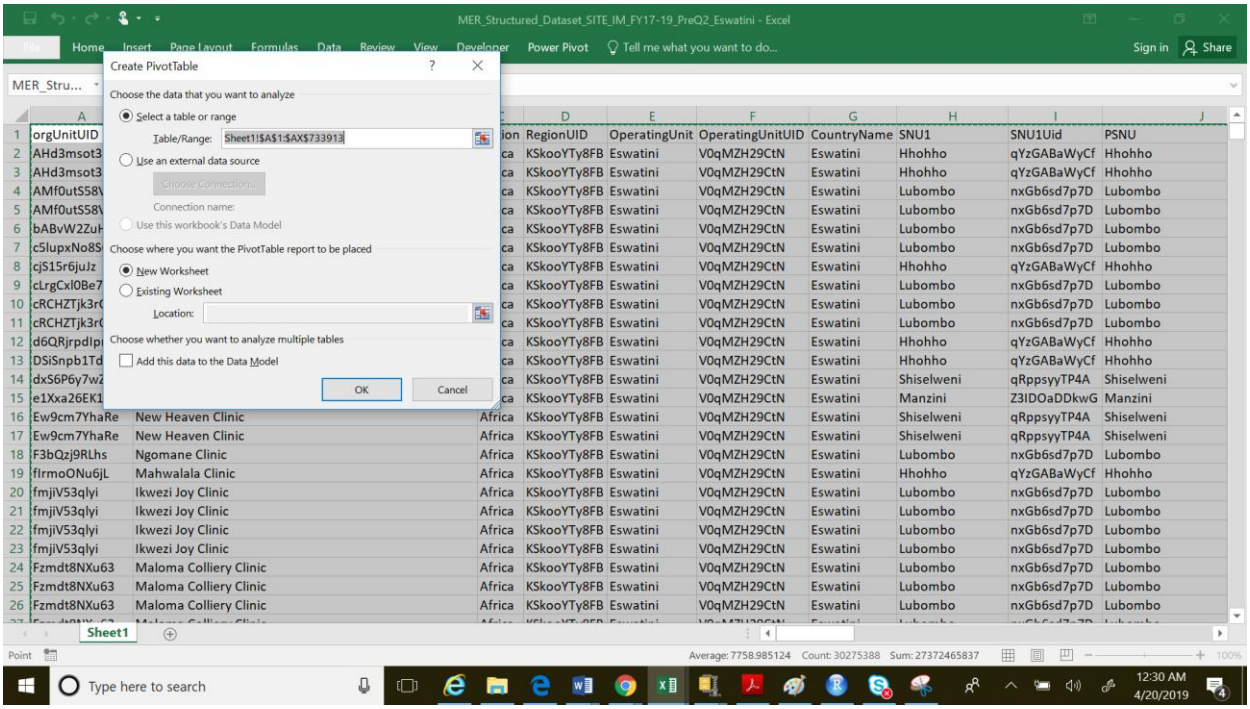
1. Select all of your imported Genie extract data by selecting the first cell with data in it. Then holding the **Shift** key, hit the **Right Arrow** key and then the **Down Arrow** key. Alternatively, you can click on the square in the upper-left corner of the spreadsheet, to the left of column "A", and above row "1."



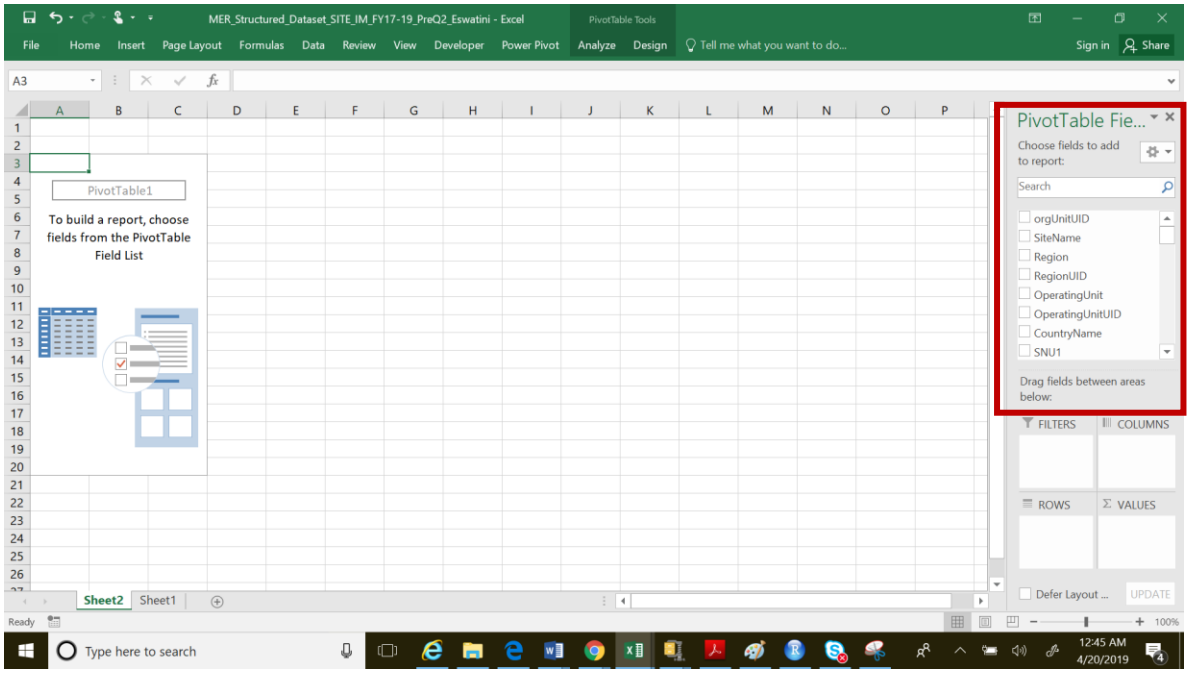
2. In the "Insert" tab, select "Pivot Table."



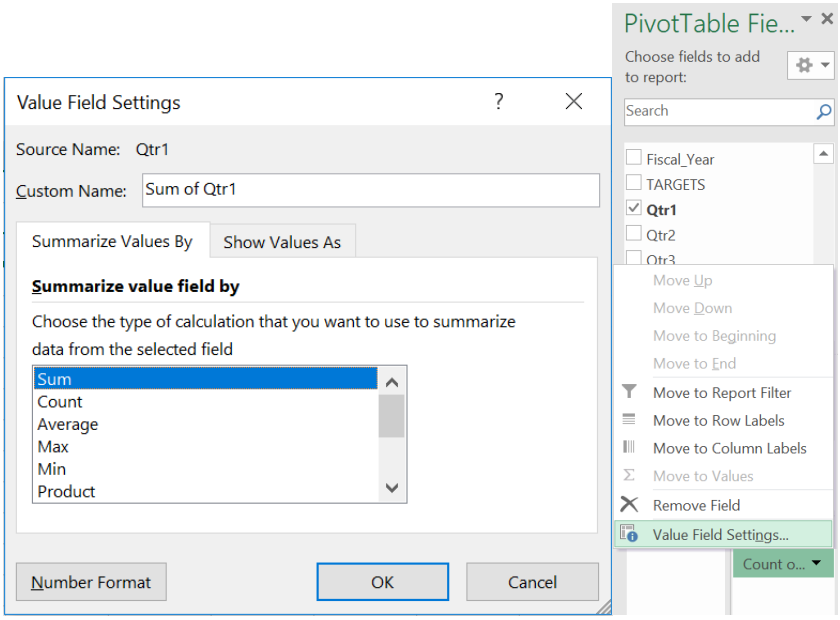
3. Select "Ok" from the window that appears.



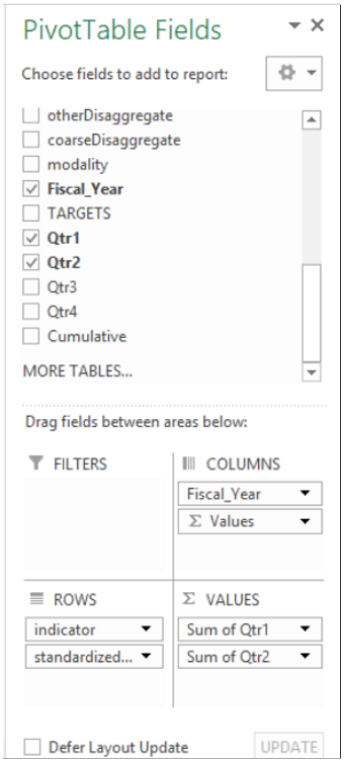
4. You will need to select which PivotTable Fields you are interested in from the list on the right side of the screen. (Note that the PivotTable Fields are the same as the column headers of the Genie report.)



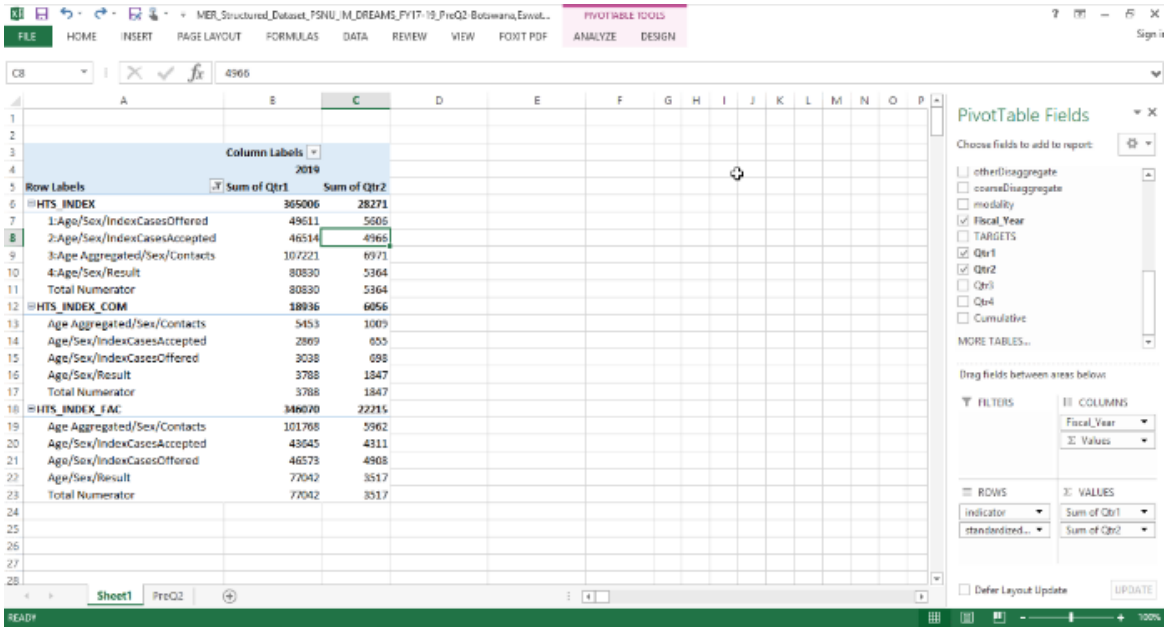
5. One potential pivot table "recipe" is the following:
    a. Drag the "Fiscal_Year" field from the list to the box labeled "Columns"
    b. Drag the "Sum of Qtr1" and "Sum of Qtr2" columns from the list to box labeled "Values"
    c. Click on "Sum of Qtr1", select "Value Field Settings,"
    d. Change the value field setting from "Count" (usually the default) to "Sum," then repeat steps "c" and "d" for "Sum of Qtr2"



    e. Drag "indicator" and "standardizedDisaggregate" from the list to the box labeled "Rows"



f. Your final pivot table will include the summed values of quarters 1 and 2 for whichever fiscal years and indicators you choose.



Reducing the File Size to Improve Performance in Excel

Excel may run very slowly on some computers when using large files. If this occurs, it is recommended to reduce the size of your dataset by refining your filter criteria in Genie (e.g. selecting only the specific SNU1, PSNU, indicators, disaggregates, and/or mechanism that are needed for your analyses).

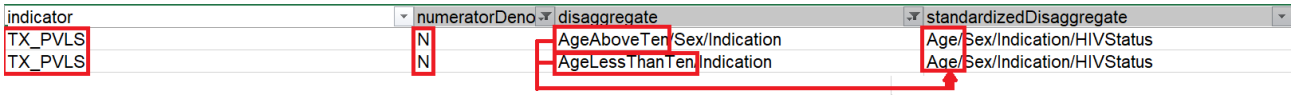
Filtering can also be done within Excel or an external statistical package (e.g., R, SAS, STATA, SPSS) after a report has been exported from Genie. For those interested in using statistical packages to filter or manipulate large datasets, ICPI has provided sample code which can be found in the ICPI Data Store on pepfar.net (link).

For additional guidance on how to correctly filter data within the analytical documents, refer to the Fact View webinars/PowerPoint slides posted to the training folder on pepfar.net. Please contact DATIM Support for any questions or inquiries.

Appendix A: Standardized Disaggregate Groupings

The table below shows the "recipes" for certain standardized disaggregates. Specifically, it shows which combinations of (1) indicators, (2) numerator or denominator data, and (3) original MER disaggregate comprise each of the standardized disaggregates. Please note that these "recipes" are not necessarily the only criteria for categorization under a given standardized disaggregate. Please also note that the table below shows only standardized disaggregates which change some part of the disaggregate (see example below); otherwise, the standardized disaggregate will be the same as the disaggregate, which can be found in the data element name.

For example, the standardized disaggregate "Age/Sex/Indication/HIVStatus" is applicable for the indicator TX_PVLS if it is: (1) numerator data and (2) either the AgeAboveTen/Sex/Indication or AgeLessThanTen/Indication disaggregate data. Please see the image below for an example.



indicator	numeratorDeno	disaggregate	standardizedDisaggregate
TX_PVLS	N	AgeAboveTen/Sex/Indication	Age/Sex/Indication/HIVStatus
TX_PVLS	N	AgeLessThanTen/Indication	Age/Sex/Indication/HIVStatus

Table 1: Standardized Disaggregate Groupings

	WHEN Conditions are met			
The Standardized Disaggregate is	**Indicator is one of**	**Numerator Denominator**	**Disaggregate contains**	**Disaggregate cannot contain**
1:Age/Sex/IndexCasesOffered	HTS_INDEX	N	Age/Sex/IndexCasesOffered	
2:Age/Sex/IndexCasesAccepted	HTS_INDEX	N	Age/Sex/IndexCasesAccepted	
3:Age Aggregated/Sex/Contacts	HTS_INDEX	N	Age Aggregated/Sex/Contacts	
4:Age/Sex/Result	HTS_INDEX	N	Age/Sex/Result	
Modality/MostCompleteAgeDisagg	HTS_TST, HTS_TST_POS, HTS_TST_NEG	N	Result/<5, MostCompleteAgeDisagg	
Modality/Age/Sex/Result	HTS_TST, HTS_TST_POS, HTS_TST_NEG	N	Result/<5	MostCompleteAgeDisagg

Modality/MostCompleteAgeDisagg	HTS_TST, HTS_TST_POS, HTS_TST_NEG	N	MostCompleteAgeDisagg	
Modality/Age/Sex/Result	HTS_TST, HTS_TST_POS, HTS_TST_NEG	N	AgeAboveTen/Sex/Result AgeLessThanTen/Result	
Modality/Age Aggregated/Sex/Result	HTS_TST, HTS_TST_POS, HTS_TST_NEG	N	Age Aggregated/Sex/Result Age Aggregated/Result	
Modality/Age/Sex/Result	HTS_TST, HTS_TST_POS, HTS_TST_NEG	N	/Age/Sex/Result /Age/Result	
Modality/Age/Sex/RTRI/HIVStatus	HTS_RECENT	N	/Age/Sex/RTRI/HIVStatus	
Total Numerator	TX_CURR, TX_NEW, TB_ART,TB_PREV,TX_TB	N	HIVStatus	
Total Denominator	TX_CURR, TX_NEW, TB_ART,TB_PREV,TX_TB	D	HIVStatus	
Total Numerator	PMTCT_STAT,VMMC_CIRC	N	Sex	
Age/Sex	OVC_SERV,TB_ART, TX_RET	N	AgeAboveTen/Sex AgeLessThanTen	
NewExistingArt/Sex/HIVStatus	PMTCT_ART	N	NewExistingArt/HIVStatus	
Age/Sex	PMTCT_STAT,VMMC_CIRC	N	Age	
Age/Sex/KnownNewResult	PMTCT_STAT, PMTCT_STAT_POS	N	Age/KnownNewResult	
Age/Sex/HIVStatus	TX_CURR, TX_NEW	N	Age/Sex,AgeAboveTen/Sex AgeLessThanTen	
Age Aggregated/Sex/HIVStatus	TX_CURR, TX_NEW	N	Age Aggregated/Sex	
KeyPop/HIVStatus	TX_NEW	N	KeyPop	
Age/Sex/Indication/HIVStatus	TX_PVLS	N	AgeAboveTen/Sex/Indication AgeLessThanTen/Indication	

Appendix B: Site by IM Data Dictionary[3]

Column Name	Column Description	Data Type (Characters)	Comments
orgunituid	11 digit alphanumeric Organization Unit (mixed case string uniquely identifying each Organization Unit) pertaining to the location for which data is reported.	String (11)	
sitename	Name of the organizational unit pertaining to the location for which data is reported.	String (200)	• The SiteName column uses the following logic: If SiteType is a facility then Facility Name If SiteType is a community then Community Name If SiteType is military PSNU then PSNU Name • For above-site indicators (e.g. HRH_CURR), this field is blank
operatingunit	Name of the Operating Unit pertaining to the location for which data is reported. Also known as *orgLevel3Name.*	String (32)	• For some indicators (e.g. LAB_PT, LAB_ACC), the Operating Unit will be the lowest level at which data is reported.
operatingunituid	11-digit, alphanumeric, mixed case string uniquely identifying each Operating Unit.	String (11)	

[3] The Site by IM dataset is inclusive of each field in this table. The OU by IM and PSNU by IM datasets do not include organizational hierarchy levels below the OU or PSNU levels respectively.

countryname	Name of the country pertaining to the location for which data is reported. For Regional Programs, this will differ from Operating Unit (e.g., for Asia Regional Program, this column would include Thailand, Laos, China, and _Military Asia Regional Program). In all other cases, this column will list the OU.	String (50)	• Helpful in combining data with historic MER/NGI results obtained from FACTS Info. • Caution is advised when using this column in isolation from Operating Unit, especially for _Military data. Because _Military data in Regional Programs is collected at a subnational level, selecting any Regional Program country using this column (e.g. China, Barbados, Bahamas) will inherently exclude all _Military data pertaining to that country and yield an underestimate of actual results or targets for that country.
snu1	The name of the sub-national unit immediately below the OU pertaining to the location for which data is reported. For _Military data, the SNU1 will be listed as _Military [*Operating Unit Name*]. Also known as *orgLevel4Name.*	String (50)	• For Operating Units that prioritize at the SNU1 level, and for all _Military locations, SNU1 will be identical to PSNU.
snu1uid	11-digit, alphanumeric, mixed case string uniquely identifying each SNU1	String (11)	•
psnu	Name of the Prioritization-level SNU pertaining to the location for which data is reported. The level at which prioritization occurs differs from OU to OU, varying from SNU1 to SNU3. This column correctly lists the name of the Prioritization-level SNU regardless of which level prioritization occurred at.	String (200)	• If data is reported at the _Military level, the PSNU will be listed as _Military [*Operating Unit Name*]. • If data is reported at the OU level, the PSNU will be left blank.
psnuuid	11-digit, alphanumeric, mixed case string uniquely identifying each PSNU.	String (11)	• Important for use in mapping outside of DATIM.

snuprioritization	The prioritization assigned to the Prioritization-level SNU (PSNU). Options include: – 1 - Scale-Up: Saturation – 2 - Scale-Up: Aggressive – 3 - NA – 4 - Sustained – 5 - Centrally Supported – 6 - Sustained: Commodities – 7 - Attained – 8 - Not PEPFAR Supported	String (50)	• Prioritizations may change each fiscal year depending on OU and PEPFAR priorities. This column reflects only the prioritization level from the fiscal year of the current reporting period (e.g. during the FY18 Q4 reporting period, this will show FY18 prioritization levels. During the FY19 Q1 reporting period, this will begin to show FY19 prioritization levels). • Many Countries have not reported prioritizations for all SNUs at the level at which prioritization occurs. These have been left blank in this dataset.
typemilitary	A yes or no field to indicate if the Site is military related	String (1)	Valid values are Y and N
dreams	A yes or no field to indicator if the organization is tagged as DREAMS	String (1)	Valid values are Y and N
primepartner	Name of the organization that is the prime partner for a given mechanism.	String(200)	• Caution is advised to use Implementing Mechanism Name instead of Prime Partner as the two are easily confused. A Prime Partner can be responsible for multiple implementing mechanisms across multiple organization units.
fundingagency	Name of the PEPFAR funding agency for the mechanism.	String(50)	• Dedup • DOD • HHS/CDC • HHS/HRSA • HHS/NIH • PC • State/AF • State/PRM • USAID

mech_code	Four or five digit, numeric value uniquely identifying each Mechanism. These are aligned across both FACTS Info and DATIM – as DATIM receives these directly from FACTS Info.	String (50)	• The two de-duplication mechanisms will be denoted as either '00000' (cross-mechanism de-duplication) or '00001' (DSD-TA de-duplication).
mech_name	Name of the Implementing Mechanism which collected the data.	String (200)	• While the mechanism code often does not change for a single mechanism across its lifetime, the name of the mechanism may change over time.
pre_rgnlztn_hq_mech_code	With COP 2019 mechanisms that are in regions will now be specific to country level, some of the old mechanism codes will change as they are moved/split-- this is the only trace back to the old code.	String (11)	
prime_partner_duns	Each partner is uniquely identified by the Data Universal Numbering System DUNS) starting with FY2020 data and soon the DUNS will be present for other prior years. Assists in analyzing partner performance.	String (20)	Can reference sam.gov for more information
award_number	Contract award numbers are given to Partners but may span across multiple mechanisms and OUs. Contract awards numbers will change over time as contracts end and new contracts begin.	String (250)	
communityuid	11-digit, alphanumeric, mixed case string uniquely identifying each Community.	String (11)	• Important for use in mapping outside of DATIM.
community	Name of the Community site.	String (200)	
communityprioritization	The prioritization assigned to the Community level. Options include: – 3 - Scale-Up – 4 - Sustained – 5 - Centrally Supported – 8 - Not PEPFAR Supported	String (50)	• Prioritizations may change each fiscal year depending on Country and PEPFAR priorities. This column provides the prioritization value for the current fiscal year only.
facilityuid	11-digit, alphanumeric, mixed case string uniquely identifying each Facility.	String (11)	• Important for use in mapping outside of DATIM.

facility	Name of the Facility site.	String (200)	
facilityprioritization	The prioritization assigned to the Facility level. Options include: – 3 - Scale-Up – 4 - Centrally Supported – 5 - Sustained – 8 - Not PEPFAR Supported	String (50)	• Prioritizations may change each fiscal year depending on Country and PEPFAR priorities. This column provides the prioritization value for the current fiscal year only.
sitetype	Indicates whether the organizational unit at which the data was reported is a facility, community, or _Military PSNU.	String(9)	• For above-site indicators (like HRH_CURR), this field is blank
dataelementuid	11-digit, alphanumeric, mixed case string uniquely identifying each DATIM data element name.	String (11)	This field is only displayed in Genie and will not be present in a MER Structured Dataset download from Panorama
indicator	The most commonly used method of referring to the MER indicators as outlined in PEPFAR MER guidance (e.g. HTC_TST, TX_CURR, PMTCT_STAT)	String (50)	• In the case of calculated indicators where a data element is structured as a disaggregate of a disaggregate (e.g., for KP_PREV, MSM/TG who are sex workers, which is a subset of MSM/TG) this column will list the full disaggregate (KP_PREV_MSMTGSW) as opposed to the actual indicator (KP_PREV). Other examples of this are: PMTCT_EID_POS_12MO, PMTCT_EID_POS_2MO, GEND_GBV_PEP, and all disaggregates for LAB_PT. These can easily be identified anywhere there is more than 1 underscore ('_') in the indicator name.
numeratordenom	Lists whether a data point pertains to the numerator ('N') or denominator ('D') of a MER indicator.	String (4)	
indicatortype	Lists whether a data point pertains to Direct Service Delivery ('DSD') or Technical Assistance ('TA').	String (4)	• Due to the unique nature of some indicators this column may also list "Not Applicable" in addition to "DSD" or "TA"

disaggregate	Lists the type of disaggregate the data point relates to, such as 'Age/Sex/Result', 'Known/New', 'ServiceDeliveryPoint', etc.	String (30)	• This is largely similar to the column by the same name available via DATIM Genie MER Approved Data Extract export, with the exception that when a data point pertains to a total numerator or total denominator value, this column has been enhanced to list 'Total Numerator' or 'Total Denominator', respectively. This was previously available only by using numeratorDenom where categoryOptionComboName is 'default'.
standardizeddisaggregate	Standardizes disaggregate names – particularly across testing modalities (for HTS_TST) and across time (when implied disaggregations were applied to some indicators in FY18Q1 within DATIM)	String (30)	• Standardizes the names of comparable disaggregate across time. This column is particularly useful for indicators for which implied disaggregations were added to the DATIM disaggregate name in FY18Q1. • See Standardized Disaggregation and Appendix A: Standardized Disaggregate Groupings for additional details.
categoryoptioncomobuid	11-digit, alphanumeric, mixed case string uniquely identifying each DATIM categoryOptionComboName.	String(11)	This field is only displayed in Genie and will not be present in a MER Structured Dataset download from Panorama
categoryoptioncombonam e	Lists the specific disaggregation the data point relates to, such as 'Male, <1, Positive', 'Known at Entry Positive', etc.	String (500)	

ageasentered	Lists the age range of service recipients described in the associated data point (where this is provided). This corresponds with the age group from the data entry screen. It is created as a byproduct of the 'categoryOptionComboName' column.	String (25)	• In previous versions of the Site x IM reports, this column was called "Age"; it was renamed in FY18 Q2 • To allow for easier ordering in visualizations and tables, this column structures ages as two digits, including a leading zero for all ages less than 10. • For age ranges in terms of months rather than years, ages are preceded by '[months]' to distinguish them from year ranges, as well as to allow most ordering algorithms to list these before any year ranges. • VMMC_CIRC also includes an age disaggregation called "02 months - 09 years" • This column (or the TrendsFine, TrendsSemiFine, or TrendsCoarse columns) should usually be used in combination with the 'disaggregate' or 'standardizedDisaggregate' columns, to prevent inadvertent duplication of data across disaggregate groups.
trendsfine	Displays age ranges that correspond to the Fine age bands that were introduced in FY18Q2.	String (25)	• See the Additional Fine, Semi-Fine, and Coarse Age Band Columns for additional information. • This column (or the AgeAsEntered, TrendsFine, TrendsSemiFine, or TrendsCoarse columns) should usually be used in combination with the 'disaggregate' or 'standardizedDisaggregate' columns, to prevent inadvertent duplication of data across disaggregate groups.
trendssemifine	Contains age ranges from both the Semi-Fine age bands and the Fine age bands (which are aggregated up to the corresponding Semi-Fine band).	String (25)	• See the Additional Fine, Semi-Fine, and Coarse Age Band Columns for additional information. • This column (or the AgeAsEntered, AgeFine, or AgeCoarse columns) should usually be used in combination with the 'disaggregate' or 'standardizedDisaggregate' columns, to prevent inadvertent duplication of data across disaggregate groups.

trendscoarse	Contains age ranges from the Coarse age bands as well as from the Semi-Fine and Fine age bands (which are aggregated up to the corresponding Coarse band).	String (25)	• See the Additional Fine, Semi-Fine, and Coarse Age Band Columns for additional information. • This column (or the AgeAsEntered, AgeFine, or AgeSemiFine, columns) should usually be used in combination with the 'disaggregate' or 'standardizedDisaggregate' columns, to prevent inadvertent duplication of data across disaggregate groups.
sex	Lists the sex of service recipients described by the associated data point (where this is provided). This is created as a byproduct of the 'categoryOptionComboName' column. Options include 'Female', 'Male', 'Unknown Sex' and blanks.	String (6)	• This column should usually be used in combination with the 'disaggregate' or 'standardizedDisaggregate' columns to prevent inadvertent duplication of data across disaggregate groups where sex may be used in both (e.g., Age/Sex Aggregated Result and Age/Sex/Result).
statushiv	Lists the HIV result status (positive or negative) of an indicator where provided or can be inferred – as is the case for indicators such as TX_NEW or TX_CURR. This is created as a byproduct of the 'categoryOptionCombo Name' column. Options include 'Positive', 'Negative', 'Unknown', or blanks.	String (8)	• This column should usually be used in combination with the 'disaggregate' or 'standardizedDisaggregate' columns to prevent inadvertent duplication of data across disaggregate groups. • There are two other columns that contain the more detailed phrases such as "Known Positive," "Newly Identified," etc. Those two other columns are Category Option Combo Name and Other Disaggregate.
statustb	Lists the TB result status for Tuberculosis Indicators	String(25)	
statuscx	Lists the Cervical Cancer status for Cervical Cancer (CXCA) Indicators	String(25)	
hiv_treatment_status	Lists status for HIV antiretroviral therapy (ART)	String (25)	• Values will include New or Already
otherdisaggregate	Any category of disaggregation not already captured in the 'Age', 'Sex', 'StatusHIV', or 'modality' columns is listed here.	String (500)	• This column should usually be used in combination with the 'disaggregate' or 'standardizedDisaggregate' columns to prevent inadvertent duplication of data across disaggregate groups.

otherdisaggregate_sub	This field has a limited use compared to the above field and captures specific options for AGYW_PREV, HRH_CURR, LAB_PTCQI and KeyPop disaggregates.	String (500)	• For AGYW_PREV data, this field will show the time in the DREAMS program • For HRH_CURR, this field will show the type of financial support • For LAB_PTCQI, this field will show the lab testing type • For data with a Key Population (KeyPop) disaggregate, this field will only show the KeyPop option. This field takes over from the prior "population" field.
modality	Lists the HTS modalities by the associated data point (where this is provided). This is created as a byproduct of the 'disaggregate' column. Options include: Index, IndexMod, OtherMod, MobileMod, HomeMod, Inpat, TBClinic, Pediatric, Malnutrition, PMTCT ANC, VCT, VCTMod, OtherPITC, VMMC, Emergency Ward, and STI Clinic. KeyPop is also included in this column even though it is not a service delivery point/modality and should not be summed along with other modalities. (Doing so will lead to duplicative results).	String (60)	• Displays the modality for HTS indicators. Use this column to filter data by modality (Service Delivery Points). • KeyPop is not included in the calculation of HTS_TST_POS or HTS_TST_NEG
fiscal_year	Gives the four-digit Fiscal Year for the data.	String(4)	
targets	Values for MER indicators' Targets.	Decimal (36,2)	Use this with the 'Fiscal_Year' column to see Targets for each specific year.
qtr1	Entered DATIM results for October – December periods of the fiscal year of interest. Only quarterly indicators will have values represented in this column.	Decimal (36,2)	

qtr2	Entered DATIM results for January – March periods of the fiscal year of interest. Only quarterly and semi-annual indicators will have values represented in this column.	Decimal (36,2)	
qtr3	Entered DATIM results for April – June periods of the fiscal year of interest.. Only quarterly indicators will have values represented in this column.	Decimal (36,2)	
qtr4	Entered DATIM results for July – September periods of the fiscal year of interest. Quarterly, semi-annual, and annual indicators will have values represented in this column.	Decimal (36,2)	
cumulative	The Annual Program Review (APR) value according to the MER Indicator guidance for the annual total.	Decimal (36,2)	
source_name	Captures if the data was entered/collected in PEPFAR systems or if the data was derived	String(15)	Possible values in official data: • **DATIM** - data entered/collected in the PEPFAR DATIM system • **Derived** - data created based on business rules per the MER Indicator Guide
approvallevel	Numeric value representing the approval level.	Decimal (36,2)	• 1, 2 and 3 are "Approved" and 4, 5 and 99 are in process and "Unapproved." See Security Trimming. This field is only displayed in Genie and will not be present in a MER Structured Dataset download from Panorama
approvalleveldescription	Description of the numeric approval level values.	String (45)	This field is only displayed in Genie and will not be present in a MER Structured Dataset download from Panorama