Generating, auditing and tweaking the Census Database and EOY Reports

Introduction

This document will cover how to create the annual census using data from EDW, Prime, and how to incorporate additional data to fix data entry errors in either of the first two systems, to meet the goal of 100% accurate reporting data.

The following queries are all in SVN in the following folder: ap\portal\trunk\db\build-2013-census.

Where a filename located in this folder is used, text will be **bolded.** 2013 will be used to represent the current year.

Using data already on the server

In 2012, the first two steps were getting the census.edwdump<year> and census.pa<year> tables built using data from external sources.

Now, we use the census.edwdump and census.pa tables already being populated, each paycalc, indicated by LoadDate.

To select a LoadDate to use, simply open **fill-census2013.sql** and change the line:

set @LoadDate = '2013-10-31';

to the LoadDate you wish to use. (tip: the load date is usually the last October load date)

Also, open **fill-census2013separated.sql** and set the following two dates to match the reporting period:

AND maxtime.separation_date >='2012-11-01'
AND maxtime.separation_date <='2013-10-31')

Also, open **fill-census2013new.sql** and set the following two dates to match the reporting period:

ods_new_position_view.new_position_date >= '2012-11-01' AND
ods_new_position_view.new_position_date <= '2013-10-31' AND

Build the Census dataset

When I refer to the census, I refer to the table of active EDW appointments, which includes protected class data, appointment pay data, and the employee's topadmin appointment, which we are using as the population for the annual end of year reports.

The census data is the basis for the Affirmative Action Plan, The AP FactSheet, New Hires Report, Separations Report, and The Roster.

To create this dataset, you only need to run **batch-rebuildcensus.sql**, after you've added/removed the configuration data outlined in the "Configuring the Census" section of this document.

Requirements: You need the census db and the prime_prod db in order to run these scripts.

Build on Vagrant:

C:\wamp\www\ap\roster-reports\trunk\db\build-2013-census>mysql --show-warnings -uroot -ppassword --host=192.168.50.5 < **batch-rebuildcensus.sql** >> import.log 2>&1

Build on Stage/Prod:

su2 apis

cd /data/admin

sh /data/admin/deploy_build_2013_census.sh

cd /data/admin/build-2013-census

mysql -u <netid>_root -p

tee import.log

source batch-rebuildcensus.sql;

Fill in census2013separated.Age_on_separation field

This requires birthdate so that we can calculate ther person's age on separation. Since we don't store birthdate (only age) we need to federate data in from EDW to do this.

1. open this access database, in svn: ap\portal\trunk\db\build-2013-census\age_on_separation.accdb
2. Run the query: EOY_age_on_separation_mktable (it's a make table because access will try to run this query again on export, and it takes a long time to run)
3. Click Yes three times to verify that you want to run the query and replace the data in the access table with new data.
4. Double-click the age_on_separation table.
5. Click on the External Data tab, then export to text file.
6. Put the text file in the directory where you checked out svn: ap\portal\trunk\db\build-2013-census by clicking browse, then click ok
7. Delimited, Next.
8. Select Tab, check Include Field Names on First Row and click Next
9. Click Finish, Close, then close Microsoft Access.
10. On your desktop, open a command line and change to your checkout directory.
11. Run: C:\wamp\www\ap\portal\trunk\db\build-2013-census>php -f tabdelim-to-sql-update.php > age_on_separation.sql
12. Check in the .sql file.
13. Redeploy and build the census.

Q: Why do we deploy and build the census, then do the access thing and redeploy it all again?

A: We need to know who has separated, and their separation date, to know what ages to calculate. Even if we incorporated all the separations logic into the access database, it still wouldn't take into account anyone that we manually enter into separations on the census.

Auditing the Census Database

We have the following audit queries defined. They are stored in svn: ap\portal\trunk\db\verify-2013-census

Null Gender (SexCode)

svn: ap\portal\trunk\db\verify-2013-census**gender-null-factsheet.sql**

This answers the question: Is anyone who is on the factsheet missing a gender?

How do we resolve issues? If someone shows up on this list, we assign them a binary gender. This is done by editing **fill-update-all-sexcode. sql**

New Hire Appointment Dates

svn: ap\portal\trunk\db\verify-2013-census**new-hire-appointment-dates.sql**

This answers the question: Do we have people whom we have reported in the previous year as new hires, in the current reporting year new hire population?

How do we resolve issues? If someone is showing up on this list and shouldn't be, they should be removed from census2013new using "Removing someone from New Hires" in this document.

In 2013, the only people who showed up were supposed to be there, as they changed categories from Instructional to Professorial during the reporting period, and this is considered a new hire in our End of Year reporting.

Prime Promotions Not In Census

svn: ap\portal\trunk\db\verify-2013-census**prime-promotions-not-in-census.sql**

This answers the question: Is there anyone who has been promoted in Prime, who are not included in the Promotions End of Year report?

Possible resolutions:

1) Have me manually add the Person to the Promotions Report
2) Do nothing, the person is on the Separations Report but not the Promotions Report

How do we resolve issues? People will show up on this audit if they have an award of tenure as a promotion action in Prime. Award of tenure is not reportable on the End of Year report, so ignore these (the Rank for both years will be the same).

If the census2013_jcc field is null, that means the person likely needs to be added to the census (use "Adding people to the Census" in this document)

You may also need to refer to "Forcing someone to be included in Promotions" in this document, or using the specialcases table to change the report, outlined in "Configuration Tables For Portal Reports."

This query contains the sql logic used in the front end to determine promotions. If function getNumberOfProfessorialInstructionalPromotions($year) or function getNumberOfClinicalAffiliatePromotions($year) in svn: ap\portal\trunk\www\application\portal\models\factsheet_model.php is updated, then this query may need to be changed to reflect that as well.

The most common problem here is that people were promoted during the year but were not active on the census date. This case is easy to spot because census2013_jcc will be null. Determine if these people should be reported, and add them to the census population using the instructions in this document.

You may also see that someone's jcc has changed in Prime (previous_jcc/promoted_jcc columns), but census2012_jcc and census2013_jcc fields don't match the promoted fields. Use the instructions in this document to change the persons jcc in the census population.

Prime Separations Not In Census

svn: ap\portal\trunk\db\verify-2013-census**prime-separations-not-in-census.sql**

This answers the question: Is there anyone who has a Separation personnel action in Prime, who is not listed in the separations report?

How do we resolve issues? If someone is listed in the audit and we want them to be included in the End of Year Separations report, first check to make sure that they are not in the census (census2013_jcc will be null). Refer to "Removing people from the Census" in this document; if they have a Separation personnel action in Prime that should fix it, otherwise refer to "Forcing someone to be included in Separations."

To be included on the census separations report, faculty have to pass the diff test and have a separation personnel action:

The diff test is that they were in the Factsheet population last year AND either:

1. Active, but were not in the factsheet population this year,
2. Active, but changed tracks
3. Not-active (no entry in the census table)

So, for example, if they are separated in Prime but are showing up as Active in the census, then follow the instructions on this page to delete them from the census, and rebuild the census database.

Census Separations with No Prime Separation Action

svn: ap\portal\trunk\db\verify-2013-census**census-separated-no-prime-separation-action.sql**

This answers the question: Is there anyone who **would** show up on the separation report, if only they had been separated in Prime?

Possible resolutions:

1) Separate the person in Prime so they show on the Separations Report.
2) Notify me that they should be included on the separations report, but should not have a separation action in Prime.
3) Have unit fix appointment data in HEPPS, and also let me know what the correct data is.
4) Do nothing, the person won't show in the Separations report.

How do we resolve issues? Determine if any of the people on this audit should be included in the End of Year Separations report. If so, refer to "Forcing someone to be included in Separations" in this document.

Some of the people on this report may have accurate data in both systems and still not be considered Separations for reporting purposes.

VRI people not in Census Separations

svn: ap\portal\trunk\db\verify-2013-census**in-vri-not-in-separations.sql**

This answers the question: Is there anyone who has accepted VRI with a retirement date in the reporting period, who are not in the census2013separated table?

How do we resolve issues? If someone is listed in the audit and we want them to be included in the End of Year Separations report, first check to make sure that they are not in the census (census2013_jcc will be null). Refer to "Removing people from the Census" in this document; if they have a Separation personnel action in Prime that should fix it, otherwise refer to "Forcing someone to be included in Separations."

Retirement_date is from the VRI manager.

census2012_jcc and census2013_jcc will be null if they were not in the census population during that year.

separation_type_name, separation_reason_name, separation_date fields are the separation information from Prime. If these are null it means that person has no separation reasons.

Prime New Positions not in Census New Hires

svn: ap\portal\trunk\db\verify-2013-census**prime-new-positions-not-in-census2013new.sql**

This answers the question: Is there anyone who has had a New Position Personnel Action entered in Prime, but who are not showing up in the New Hires Report?

Possible resolutions:

1) Notify me that they should be included on the New Hires Report, but that data as entered is correct (I will add them to the New Hires Report manually)
2) Have unit/ahr fix appointment data in HEPPS/Prime, and also let me know what the correct data is.
3) Do nothing, the person won't show in the New Hires Report.

How do we resolve issues? If someone is listed in the audit and you want them to show up on the End of Year New Hires report, then first make sure they are in the census correctly (the census2013_jcc will be null or will not match the prime_jcc column). If null, refer to "Adding people to the Census" section of this document, if the census data is incorrect, refer to the "Changing peoples Census data" section of this document.

Some of the people on this report may have accurate data in both systems and still not be considered New Hires for reporting purposes.

In Census New Hires not in Prime New Positions

svn: ap\portal\trunk\db\verify-2013-census**census2013new-not-in-prime-new-positions.sql**

This answers the question: Is there anyone who will show up on the New Hires report, but do not have a New Position Personnel Action in Prime?

Possible resolutions:

1) Notify me that they should not be included on the New Hires Report, but that data as entered is correct (I will remove them from the New Hires Report manually)
2) Have unit/ahr fix appointment data in HEPPS/Prime, and also let me know what the correct data is.
3) Do nothing, the person will show in the New Hires Report.

How do we resolve issues? If someone is listed on this audit and you do not want them to show up on the End of Year New Hires report, then first make sure they are in the census correctly (the census2013_jcc will be null or will not match the prime_jcc column). If null, refer to "Adding people to the Census" section of this document, if the census data is incorrect, refer to the "Changing peoples Census data" section of this document. If the data in the current census is correct they may be showing up as a new hire because they were not in the last census. In this case, make sure they are not reported as a new hire by following the instructions in the "Removing someone from New Hires" section of this document.

Some of the people on this report may have accurate data in both systems and still not be considered New Hires for reporting purposes.

Age on Separation is missing

svn: ap\portal\trunk\db\verify-2013-census**age_on_separation_is_null.sql**

This answers the question: Is there anyone who will show up on the portions of the Separation report which show age, but for whom Age_on_separation is not filled in?

How do we resolve issues? Re-run the Microsoft Access-based process for calculating the Age_on_separation data, check it into SVN and rebuild the census.

Configuring the Census

Due to data entry errors in HEPPS or Prime, you may need to force people to be included or excluded from certain datasets.

Removing people from the Census

Add the EID's you wish to remove to the file **fill-census2013-removefrom2013.sql**

Adding people to the Census

This is trickier because we need data to insert.

First, edit **download-edw-data-2013-supplemental.sql** to include the EID's you wish to search for. This file allows for appointment marked inactive in EDW, and also doesn't check appointment start and end dates in EDW. If this doesn't catch everyone, try changing the record effective dates at the top of the query to be later than the census date.

Note, if you can't get **download-edw-data-2013-supplemental.sql** to download the data you see in Opus, then you should either wait for a paycalc to update HumanResources, or use **download-edw-data-2013-supplemental-fallback1.sql** with a connection to the ODS views (which are updated daily): ucssqlc1db1.admin.washington.edu,4358. This will not catch PII data (fields which are denoted by returning the field name). You may find running **download-edw-data-2013-supplemental-fallback2.sql** against the normal HumanResources database useful for pulling those fields, then combine the two outputs in Excel to get the most accurate data.

Once you have at least one appointment for everyone, load that data into an Excel spreadsheet. In Excel, delete extra appointments so that the file is unique by EID, then edit the appointment to closely match the correct data as you can.

When this is done, save the file to tab delimited, and use **tabdelim-to-sqlinsert.php** to create the **fill-census2013-addto2013.sql** file.

Changing peoples Census data

If you want to change someones Census data, then add them both to the **fill-census2013-removefrom2013.sql** file and also get the correct data into the **fill-census2013-addto2013.sql.**

Forcing someone to be included in Separations

If you've ensured that someone has an appointment last year and no appointment this year, and they are not in the census2012separated table, the reason is because they have not been separated in Prime.

Force them to be included by adding the EID to **fill-census2013separated.sql**. Then, edit **fill-census2013separated-addsepreasons.sql** to create separation reason data for those people. Separation Date should be the end-date of the appointment.

Forcing someone to be included in Promotions

If you have a situation where someone has been promoted and then later separated, they will only show up in the separations report. In this case, you do want the promoted person to be absent from the census population, but also show up in the census2013changes table so they are also counted as a promotion.

Force them to be included by filling in the appropriate information in **fill-census2013changes-addto2013.sql**. This is trickier because we need data to insert.

First, edit **download-edw-data-2013-supplemental.sql** to include the EID's you wish to search for. This file allows for appointment marked inactive in EDW, and also doesn't check appointment start and end dates in EDW. If this doesn't catch everyone, try changing the record effective dates at the top of the query to be later than the census date.

Once you have at least one appointment for everyone, load that data into an Excel spreadsheet. In Excel, delete extra appointments so that the file is unique by EID, then edit the appointment to closely match the correct data as you can. Use the spreadsheet **fill-census2013changes-addto2013.xlsx** because this file has more collumns in it than the other tables. You need to be sure to line up the data you pull with the 2013 collumns, leaving the other ones empty. Then, fill in the 2012 collumns by cutting and pasting from the census2012 table.

When this is done, save the file to tab delimited, and use **tabdelim-to-sqlinsert.php** to create the **fill-census2013-addto2013.sql** file.

Now, if you are including them in Promotions because they were also separated, you should update the separation data to be the appointment that they were separated from (ie, the appointment which they were promoted into). Do this by adding their eid to **update-separated-from-changes.sql**.

Removing someone from New Hires

Add the EID's you wish to remove to the delete statement at the bottom of the file **fill-census2013new.sql**

Configuration Tables For Portal Reports

DeptOrgLookup and CollegeOrgLookup

DeptOrgLookup and CollegeOrgLookup are used to modify the Org structure to be something which is reportable.

Edit CollegeOrgNameMod or DepartmentOrgNameMod to be the appropriate reporting name, if the name is wrong in Prime or HEPPS.

To group multiple org codes together, change the CollegeOrgCodeMod and CollegeOrgNameMod to be the same, desired, org code and name..

jccs

The jccs table is used to group JCC's together into larger reporting groups and determines if JCC's are included in certain reports. It also contains labels used in the Roster and Factsheet.

JCC

Key value, used to join to the census dataset.

Rank

Used in reporting query views.

Category

Category is used to build the New Hires and Separations report. If this data element is changed you must follow the instructions in the Applying Configuration section of this document.

Track

Used in reporting query views.

Roster

Determines inclusion in the Roster. Allowed values are 'Yes', 'No'

RosterGroup

Groups JCC's into special groups for the Roster. This is also used as the label in the report.

See roster-groupings-2013.xlsx for more information.

RosterGroupOrder

Controls the ordering of the JCC's on the report. Default is to set this to be the same for all JCC's in a Group (which then orders the Groups), After that JCC is used for ordering.

See roster-groupings-2013.xlsx for more information.

RosterLabel

This is how the title shows up on the Roster. If one sets two JCC's to have the exact same value for the RosterLabel and the RosterGroupOrder, then those two JCC's will be combined on the report.

See roster-groupings-2013.xlsx for more information.

FactSheet

Determines inclusion in the Fact Sheet. Allowed values are 'Yes', 'No'

FactSheet is used to build the New Hires and Separations report. If this data element is changed you must follow the instructions in the Applying Configuration section of this document.

FactSheetLabel

Can be changed if we want to display Title differently on the FactSheet.

Applying Configuration

You need to rebuild the census. Follow the instructions in the "Build the Census Dataset" section of this document.

Tweaking Reports

You may need to tweak a report in the portal front end if the "truth" is more complex than the database can represent. For example, if someone changes jobs twice during the year, and you want to represent that as a promotion in each category.

use the census.specialcases table!

At time of writing only these two reports have specialcases implemented in the query. These are used to include or exclude people from the *Number Of Promotions Awarded in 2012* report on the portal:

NumberOfClinicalAffiliatePromotions
NumberOfProfessorialInstructionalPromotions

Table definitions

report	year	eid	action

this column can be anything, as long as it's the same string that your report uses to select the data	corresponds to the reporting year	eid to force either for inclusion or exclusion	include or exclude from a certain report.

Sample SQL

To utilize this table in other reports in the future, include them into the sql statement in a fashion similar to this:

WHERE IdNbr IN (SELECT eid FROM census.specialcases WHERE report = 'NumberOfProfessorialInstructionalPromotions' AND YEAR = '2012' AND ACTION = 'include')
OR (
IdNbr NOT IN (SELECT eid FROM census.specialcases WHERE report = 'NumberOfProfessorialInstructionalPromotions' AND YEAR = '2012' AND ACTION = 'exclude')
AND (column1 = 'example1' AND column2 = 'example2' ... etc))

Deliver Raw Data to Rhonda Forman

Use instructions located here: Exporting 10/31 census for delivery to Rhonda Foreman