

Using SPSS for CBM001 Reports

TAIR 2015 – Session G5
Presented by Jane Haas
Del Mar College

Step 1:

Import raw data from certified
CBM001 text file into SPSS

TAIR 2015 – G5 - Importing CBM001 txt into SPSS

** Del Mar College * Office of Strategic Planning and Institutional Research*

** Last updated 09/08/14 by Jane Haas*

** This program reads Coordinating Board CBM001 student term enrollment data.*

** File location R:\IRE-Limited\Data Warehouse\CBM 01 Files\CBM001 Layouts-Fall 2013 CURRENT.sps*

** CURRENT "CBM Reporting Manual" on THECB website is dated "Fall 2013" (Checked on 09/05/14 by JH)*

****NOTE: This layout below will ONLY work starting with FALL 2013 since the report record length changed from 170 to 182 characters!!!***

****The following variables changed in length and format in Fall 2013:***

** sch_naf (was F2.0 now F 4.2 = changed from numeric total length 2 with zero decimals to 4 in length, which includes 2 decimals)*

** excs_dev (was F3.0 now F4.2)*

** dual_hrs (was F2.0 now F4.2)*

** uglimit (was F2.0 now F4.2)*

** remhrs (was F2.0 now F4.2)*

** techhrs (was F2.0 now F4.2)*

** scht_nsf (was F2.0 now F4.2)*

** scha_nsf (was F2.0 now F4.2)*

** To indicate that the number of SCH need to have a decimal point inserted, add **(2)** after the position & length*

** such as: sch_naf61-64 **(2)** and it will import "300" as "3.00"*

* **Read the text data file from the Coordinating Board:**

* *THECB data files get downloaded here: 'R:\IRE-Limited\CBM Files\Downloaded Raw Data Files\CBM 01*

* **replace the name of the FOLDER and FILE NAME in location of the file for each semester**

* **data list file** = 'R:\IRE-Limited\CBM Files\Downloaded Raw Data Files\CBM 01**FOLDER**\01_NYYYY.txt' **fixed**
records=1

* *The most current data is census for Fall 2014:*

* *R:\IRE-Limited\CBM Files\Downloaded Raw Data Files\CBM 01\2014-15\FA14\01_12014.txt*

NEW FILE.

data list file= 'R:\IRE-Limited\CBM Files\Downloaded Raw Data Files\CBM 01\2014-15\FA14\01_12014.txt' **fixed**
records=1

<u>/1 ssn</u>	8-16 (a)
<u>gender</u>	17 (a)
<u>class</u>	18
<u>b_year</u>	19-22
<u>b_month</u>	23-24
<u>b_day</u>	25-26
<u>tuition</u>	27 (A)
<u>reside</u>	28-30
<u>ftic</u>	31-36
<u>cha_ind</u>	37-40
<u>cha_od</u>	41-44

<u>chv_ind</u>	45-48
<u>chv_od</u>	49-52
<u>major</u>	53-60
<u>sch_naf</u>	61-64 (2)
<u>tuit_exm</u>	65-66
<u>remote</u>	67
<u>typmjr</u>	68
<u>name_f</u>	69-78 (a)
<u>name_m_i</u>	79 (a)
<u>semester</u>	80
<u>year</u>	81-84
<u>flex_ent</u>	85
<u>dis_acad</u>	86
<u>dis_econ</u>	87
<u>dis_dis</u>	88
<u>dis_lep</u>	89
<u>dis_sex</u>	90
<u>dis_hmkr</u>	91
<u>dis_spar</u>	92
<u>excs_dev</u>	93-96 (2)
<u>iinst_ac</u>	97-99
<u>iinst_tc</u>	100-102

<u>dual_hrs</u>	103-106 (2)
<u>uglimit</u>	107-110 (2)
<u>remhrs</u>	111-114 (2)
<u>techhrs</u>	115-118 (2)
<u>rexcss</u>	119-121
<u>sintent</u>	122
<u>nondiscl</u>	123
<u>cht_nsf</u>	124-126
<u>cha_nsf</u>	127-129
<u>scht_nsf</u>	130-133 (2)
<u>scha_nsf</u>	134-137 (2)
<u>name_l</u>	138-157 (A)
<u>restradm</u>	158-159 (A)
<u>HS_code</u>	160-165 (A)
<u>PEIMS</u>	166-174 (A)
<u>ethnic</u>	175 (A)
<u>white</u>	176 (A)
<u>black</u>	177 (A)
<u>asian</u>	178 (A)
<u>natamer</u>	179 (A)
<u>intl</u>	180 (A)
<u>unkn</u>	181 (A)
<u>nathawn</u>	182 (A)
.	

EXECUTE.

** In case you are reading in a multiple-semester data file (not uncommon),
be sure to select the records from one semester at a time and save them to the appropriate .SAV file.

** SELECT IF (semester=1).*

** EXECUTE.*

VARIABLE LABELS

- ssn 'Student ID'/
- gender 'Gender - M or F'/
- class 'Classification'/
- b_month 'Birth Month'/
- b_year 'Birth Year'/
- b_day 'Birth Day'/
- tuition 'Tuition Status'/
- reside 'Residence'/
- ftic 'First Time Transfer/In College'/
- cha_ind 'Contact Hours Academic - In District'/
- cha_od 'Contact Hours Academic - Out of District'/
- chv_ind 'Contact Hours Voc-Tech - In District'/
- chv_od 'Contact Hours Voc-Tech - Out of District'/
- major 'Major'/
- sch_naf 'SCH Not Affected'/

Intl 'International' /
unkn 'Unkknown' /
nathawn 'Nat Hawaiian Pacific Isl' /

•
EXECUTE.

** If using old data file where these variables are present, add them:*

** unused1 ''/*
** unused2 ''/*
** unused_1 ''/*
** unused_2 ''/*
** unused_3 ''/.*

VALUE LABELS

class 1 'Freshman' 2 'Sophomore' 3 'Unclassified' 4 'Associate Degree' 5 'Baccalaureate or above' /

tuition 1 'In-District' 2 'Out-of-District' 3 'Nonresident' 5 'Tuition Exemption/Waivers'

'A' 'Intl Classif as a Resident and pays In-Distr Tuition'

'B' 'Intl Classif as a Resident and pays Out-of-Distr Tuition'

'C' 'Intl Pending Resident Status and pays In-Distr Tuition'

'D' 'Intl Pending Resident Status and pays Out-of-Distr Tuition'

'E' 'Good Neighbor Waiver'

'N' 'Nonresident' /

typmjr 1 'Academic' 2 'Technical' 3 'Tech-Prep' /

*** Syntax for IPEDS Ethnicity Categories:**

STRING race (A01).
COMPUTE race='0'.

*** Step 1: * Figure out the "two or more races" FIRST, regardless of Ethnic Origin**

***** Note: If the person is "white" and also "Hispanic", the race will still be "white"**

IF (white='1') **AND** (black<>'2') **AND** (asian<>'4') **AND** (natamer<>'5') **AND** (intl<>'6') **AND** (nathawn<>'8') race='1' .
IF (black='2') **AND** (white<>'1') **AND** (asian<>'4') **AND** (natamer<>'5') **AND** (intl<>'6') **AND** (nathawn<>'8') race='2' .
IF (asian='4') **AND** (white<>'1') **AND** (black<>'2') **AND** (natamer<>'5') **AND** (intl<>'6') **AND** (nathawn<>'8') race='4' .
IF (natamer='5') **AND** (white<>'1') **AND** (black<>'2') **AND** (asian<>'4') **AND** (intl<>'6') **AND** (nathawn<>'8') race='5' .
IF (intl='6') **AND** (white<>'1') **AND** (black<>'2') **AND** (asian<>'4') **AND** (natamer<>'5') **AND** (nathawn<>'8') race='6' .
IF (unkn='7') race='7' .
IF (nathawn='8') **AND** (white<>'1') **AND** (black<>'2') **AND** (asian<>'4') **AND** (natamer<>'5') **AND** (intl<>'6') race='8' .
EXECUTE.

**Note: Since all "single response" races have been assigned a value (1-8), the remaining zeros are "two or more races", so need to recode them "9" not to confuse with system missing:*

RECODE

```
race ('0'='9') (ELSE=COPY)
```

```
INTO race.
```

```
EXECUTE .
```

```
VARIABLE LABELS race 'Race'.
```

```
VALUE LABELS race '1' 'White'
```

```
          '2' 'Black'
```

```
          '4' 'Asian'
```

```
          '5' 'Amer Indian Alaskan'
```

```
          '6' 'International'
```

```
          '7' 'Unreported'
```

```
          '8' 'Hawaiian Pacific Isl'
```

```
          '9' 'Two or more races' /.
```

```
EXECUTE.
```

*** Step 2: IPEDS ethnicity category with Hispanic overwrite:**

COMPUTE ipeth1=0.

***Note: "ethnic" and "race" variables were STRING, the "ipeth1" variable is NUMERIC**

{

****** Assign IPEDS ethnicity to "Non-Hispanic" first**

IF (race='1') **AND** (ethnic<>'1') ipeth1=6.

IF (race='2') **AND** (ethnic<>'1') ipeth1=5.

IF (race='4') **AND** (ethnic<>'1') ipeth1=4.

IF (race='5') **AND** (ethnic<>'1') ipeth1=3.

IF (race='6') **AND** (ethnic<>'1') ipeth1=1.

IF (race='7') **AND** (ethnic<>'1') ipeth1=7.

IF (race='8') **AND** (ethnic<>'1') ipeth1=8.

IF (race='9') **AND** (ethnic<>'1') ipeth1=9.

EXECUTE.

****** Then assign "Hispanic" Ethnic Origin (it should replace all zeros with 2)**

IF (ethnic='1') ipeth1=2.

EXECUTE.

**Assign labels and you will see that Non-Hispanic & White = white, non-Hispanic*

** Non-Hispanic & White + other race = two or more races*

** Hispanic & White (+ other race) = Hispanic of any race*

VARIABLE LABELS ipeth1 'IPEDS Ethnicity'.

VALUE LABELS ipeth1_1 'Non-resident alien'

2 'Hispanic of any race'

3 'American Indian'

4 'Asian, non-Hispanic'

5 'Black, non-Hispanic'

6 'White, non-Hispanic'

7 'Unknown'

8 'Native Hawaiian'

9 'Two or more races' /.

EXECUTE.

**Note that the race/ethnicity numbering is changed from CBM001 order (white=1) to IPEDS order (white=6)*

*** Look for blank records.**

SORT CASES BY

ssn (A) .

*** Deleting blank records based on classification being system missing.**

FILTER OFF.

USE ALL.

SELECT IF(SYSMIS(class) ~ = 1).

EXECUTE .

*****:

*** RUN TILL HERE AND SAVE**

*** At this point, stop to look for duplicates before proceeding.**

*** Flex Entry students do have duplicate records -- duplicates in this case**

*** or in the cases where students enter in multiple terms within a term like**

*** rapid track and mini sessions should not be deleted either.**

```
COMPUTE dup=0.
SORT CASES BY
  ssn (A).
IF ssn=lag(ssn) dup=1.
FREQUENCIES
  VARIABLES=dup
  /ORDER ANALYSIS.
EXECUTE.
```

```
*****
SELECT IF (dup=0).
EXECUTE.
```

```
*****
* Evaluate data for possible problems.
```

```
FREQUENCIES
  VARIABLES=gender class b_month b_year tuition reside ftic cha_ind cha_od
  chv_ind chv_od major tuit_exm remote ethnic semester year flex_ent
  dis_acad dis_econ dis_dis dis_lep dis_sex dis_hmkr dis_spar excs_dev iinst_ac
  iinst_tc dual_hrs uglimit remhrs techhrs rexcexs sintent residenc nondiscl restradm
  /ORDER ANALYSIS.
EXECUTE.
```

```
*****
* If the datafile looks clean you should save it at this point
* Save the file using the dialogs just to ensure that you do not overwrite a previous file that you need.
```

```
*****
```

Step 2:

Add semester to “Main” Database

- 1) In Access – to “CBM001.mdb”
- 2) In SPSS – to “CBM001.sav”

Step 3:

Use “CBM001 MASTER Syntax”
to produce various tables

TAIR2015-G5-CBM001 MASTER Syntax

**R:\IRE-Limited\Data Warehouse\CBM 01 Files\CBM001 MASTER Syntax.sps

* Created by Jane Haas 09-04-2014 from bits and pieces of various syntax found throughout.

* **Purpose: To calculate student demographic and academic statistics for Program Review tables.**

* Also used for Statistical Profile|

* Also used for IPEDS Enrollment Survey

* Also used for Title III/V IDUES MSEIP program annual data report.

* Also used for CAFR and Budget Statistical Supplements.

* Also used for College Board Survey

* **Create SPSS data file from CBM text file using the "CBM001 Layouts-Fall 2013 CURRENT.sps"**

* (from Fall 2013 forward. Still applicable to Spring 2015). For prior semesters use prior "Layouts".

* **Some of the variables have already been created by the "Layouts syntax". Check what is missing**

** Compute total contact hours.*

COMPUTE ch=0.

COMPUTE ch=cha_ind+cha_od+chv_ind+chv_od+iinst_ac+iinst_tc+rexcess+cht_nsf+cha_nsf.

EXECUTE.

VARIABLE LABELS ch 'Contact Hours'.

** Compute total non-funded contact hours:*

COMPUTE ch_nsf=cht_nsf+cha_nsf.

EXECUTE.

FREQUENCIES

VARIABLES=ch

/STATISTICS=SUM

/ORDER= ANALYSIS .

** Compute contact hours by student tuition status.*

TABLES

/OBSERVATION ch xtot

/FORMAT BLANK MISSING('.

/FTOT TOTAL 'Total'

/TABLES ch > tuition + TOTAL BY (STATISTICS)

/TITLE 'Term Census Student Contact Hours by Tuition Status'

/STATISTICS SUM(ch(COMMA5.0) 'Contact Hours')

*** To get selected students only (by term). Make sure change semester and term to current:**

```
SELECT IF (sem=20141 & term =20139).  
EXECUTE.
```

USE ALL.

```
COMPUTE filter_$=(PrimaryLast = 1).  
VARIABLE LABEL filter_$ 'PrimaryLast = 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMAT filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE .
```

*** To get FTICs only.**

```
SELECT IF (ftic=1).  
EXECUTE.
```

***To get Full-time only.**

```
SELECT IF (fullpart>=1).  
EXECUTE.
```

|

```
IF MISSING(fullpart) fullpart=fpgrad.  
EXECUTE.
```

**** END of CBM001 MASTER Syntax**

Note:

The “CBM001 MASTER Syntax.docx” is **23** pages long and contains syntax used for various tables from IPEDS to Statistical Profile. I would be glad to share it with you.

I prefer to save a copy of SPSS syntax as a Word document and make changes in Word (so I can highlight items and insert screen prints) before pasting them to SPSS syntax.

Thank You!

Jane Haas

Del Mar College

Office of Strategic Planning and Institutional Research

101 Baldwin Blvd., HA-157

Corpus Christi, TX 78404

Phone (361)698-1552, Fax (361)698-1090

E-mail jhaas@delmar.edu