

UTILIZING MACROS TO ANALYZE DATA FROM MULTIPLE DATASETS

Diana Tersteeg
ADF&G

The Task

What is requested

- We want to list the Harvest per unit of effort (HPUE) by Site and species sampled by individual year and five-year average to post on our website for public information.

What is available

- Datasets are output by port and by year with species as variables.
- To note: We redesigned our program in 2011 so data prior to this is no longer consistent, so is currently a 4-year average.

What was done...

- At the end of each year a program is run to output the HPUE for each port for that year with the HPUE for six different species.
- We had nine SAS programs to grab these files (one for each Site) with code for each Species written as one data step after the other that had to be updated each year and run individually.
- Nine sites with six species each amounted to updating file names and years for 54 individual data steps and sitting in front of the computer to wait for each program to finish so you could run the next one in line.

The Data

- We have nine different Sites:

- Ketchikan
- Prince of Wales
- Petersburg
- Wrangell
- Sitka
- Elfin Cove
- Gustavus
- Juneau
- Yakutat

And we want the mean HPUE by week from six different species for each of the Sites:

- King salmon
- Coho salmon
- Pink salmon
- Chum salmon
- Halibut
- Rockfish (all species)

Site	Year	Week	King	Coho
Ketchikan	2011	9	0.0125	0.0001

The Output

- For each Site sampled – want a spreadsheet to display by week the HPUE per year and the mean HPUE over years requested by species.

Site	Week	KS_2011_ HPUE	KS_2012_ HPUE	KS_2013_ HPUE	KS_2014_ HPUE	KS_mean _HPUE
Juneau	12	0.00485	0.0854	0.1887	0.8746	0.2884

Procedure

- **Outer Macro (CAT_GO):** by **Site** – Runs program for each of the Sites called and pulls in all requested years SAS datasets for that Site.
 - % Let statements – assign macro variables
 - Initial ods tagsets.ExcelXP file listing
 - **Inner Macro (NAME):** by **Species** – Within a Site, runs the macro for each species called.
 - Each species output to the ods tagsets.ExcelXP
 - **Mend Inner Macro**
 - ods tagsets.ExcelXP close;
 - **Mend Outer Macro**

Outer Macro

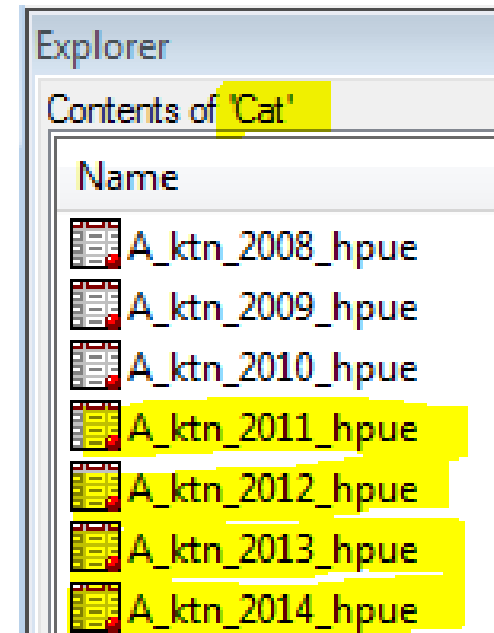
```
%MACRO CAT_GO ( INLIB=CAT ,  
                SITE= ,  
                NUMYR=4 ,  
                YEAR1=2011 ,  
                YEAR2=2012 ,  
                YEAR3=2013 ,  
                YEAR4=2014 ,  
                YEAR5= ,  
                REDCT=Y ,  
                SAVELIB=CAT ) ;
```

```
%mend CAT_GO ;
```

```
%CAT_GO ( SITE=KETCHIKAN ) ; **total of 9 site calls **
```

```
%CAT_GO ( SITE=POW ) ;
```

```
%CAT_GO ( SITE=PETERSBURG ) ;
```



Macro variables for input files (V1)

```
/* Create Prefix strings for naming the input files */
```

```
%LET SITE=%UPCASE(&SITE);  
%if      "&SITE" EQ "KETCHIKAN"      %then %let N=A_KTN_;  
%else %if "&SITE" EQ "POW"            %then %let N=B_POW_;  
%else %if "&SITE" EQ "PETERSBURG"    %then %let N=C_PTG_;  
%else %if "&SITE" EQ "WRANGELL"      %then %let N=C_WRG_;  
%else %if "&SITE" EQ "SITKA"        %then %let N=D_SIT_;  
%else %if "&SITE" EQ "JUNEAU"        %then %let N=E_JNU_;  
%else %if "&SITE" EQ "GUSTAVUS"     %then %let N=G_GUS_;  
%else %if "&SITE" EQ "ELFIN COVE"   %then %let N=G_ELF_;  
%else %if "&SITE" EQ "YAKUTAT"     %then %let N=H_YAK_;
```


Or Use a Format lookup (V2)

```
proc format;  
  value $sitecode  
    "KETCHIKAN"    = 'A_KTN_'  
    "POW"          = 'B_POW_'  
    "PETERSBURG"   = 'C_PTG_'  
    "WRANGELL"     = 'C_WRG_'  
    "SITKA"        = 'D_SIT_'  
    "JUNEAU"       = 'E_JNU_'  
    "GUSTAVUS"     = 'G_GUS_'  
    "ELFIN COVE"   = 'G_ELF_'  
    "YAKUTAT"      = 'H_YAK_'  
    other          = 'UNK  ' ;  
run;
```

```
%let n = %sysfunc(putc(&site,$sitecode.));  
%put &n;
```

n=A_KTN_

```
%let P=&n&YEAR1._&YEAR4.;
```

P=A_KTN_2011_2014

Set Statement – what we need per site

```
DATA A_KTN_ALL;  
    set  
    CAT.A_ktn_2011_hpue(keep=site year week tking tcoho tpink  
                        tchum thalibut trockfish)  
    CAT.A_ktn_2012_hpue(keep=site year week tking tcoho tpink  
                        tchum thalibut trockfish)  
    CAT.A_ktn_2013_hpue(keep=site year week tking tcoho tpink  
                        tchum thalibut trockfish)  
    CAT.A_ktn_2014_hpue(keep=site year week tking tcoho tpink  
                        tchum thalibut trockfish)  
  
    ;  
run;
```

Macro variables for dataset names (V1)

* Create the infile for each site and year;

```
%let dsn1=&INLIB..&N.&YEAR1._hpue;
```

```
%let dsn2=&INLIB..&N.&YEAR2._hpue;
```

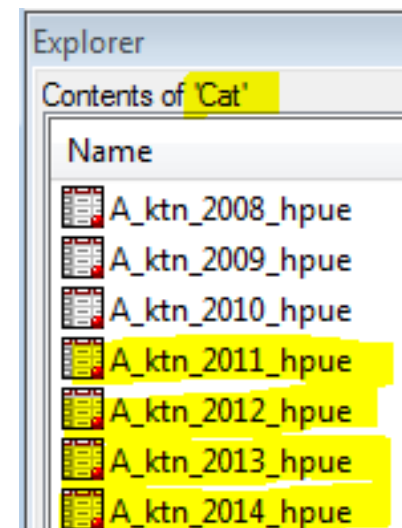
```
%let dsn3=&INLIB..&N.&YEAR3._hpue;
```

```
%let dsn4=&INLIB..&N.&YEAR4._hpue;
```

```
%let dsn5=&INLIB..&N.&YEAR5._hpue;
```

• INLIB=CAT

• %let N=A_KTN_;



* create outfile for each site and year;

```
%let dsn6=&savelib..&P.;
```

• SAVELIB=CAT

• P=A_KTN_&YEAR1._&YEAR4.

Set Statement (V1)

Do loop within the set statement allows you to bring in multiple datasets:

```
DATA &N.ALL;  
    set  
        %do I=1 %to &NUMYR. %by 1;  
            &&dsn&I (keep=site year week tking tcoho  
tpink tchum thalibut trockfish)  
  
        %end;  
    ;  
run;
```

Set Statement (V1)

Do loop within the set statement allows you to bring in multiple datasets:

```
DATA &N.ALL;
```

```
  set
```

```
    %do I=1 %to &NUMYR. %by 1;
```

```
      &&dsn&I (keep=site year week tking tcoho  
tpink tchum thalibut trockfish)
```

```
    %end;
```

```
  ;
```

```
run;
```

- %let N=A_KTN_;

- NUMYR=4,

- %let dsn1=&INLIB..&N.&YEAR1._hpue;

Set Statement (V1)

```
DATA A_KTN_ALL;
```

```
set
```

```
  %do I=1 %to 4 %by 1;
```

```
    &dsn1 (keep=site year week tking tcoho tpink  
tchum thalibut trockfish)
```

```
  %end;
```

```
;
```

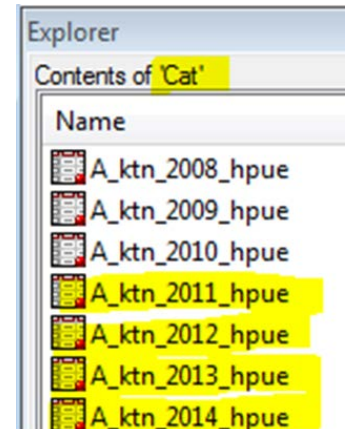
```
run;
```

- `%let N=A_KTN_;`

- `NUMYR=4,`

```
%let dsn1=&INLIB..&N.&YEAR1._hpue;
```

```
Dsn1=CAT.A_KTN_2011_hpue;
```



Explorer	
Contents of 'Cat'	
Name	
A_ktn_2008_hpue	
A_ktn_2009_hpue	
A_ktn_2010_hpue	
A_ktn_2011_hpue	Highlighted
A_ktn_2012_hpue	
A_ktn_2013_hpue	
A_ktn_2014_hpue	

Or alternate Set Statement (V2)

```
DATA &N.ALL;
```

```
set
```

```
%do I=&YEAR1. %to &YEAR4. %by 1;
```

```
    &INLIB..&N.&I._hpue (keep=site year week  
tking tcoho tpink tchum thalibut trockfish)
```

```
%end;
```

```
;
```

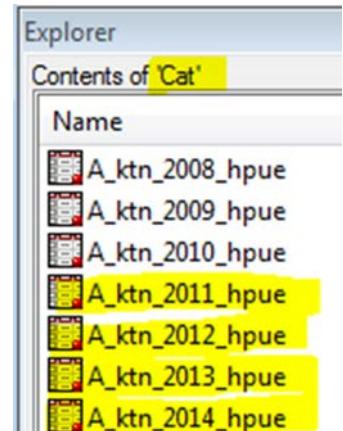
```
run;
```

```
CAT.A_KTN_2011_hpue;
```

- %let N=A_KTN_;

- INLIB=CAT

Now Do not need to define dsn1-4



Or alternate Set Statement (V2)

```
DATA A_KTN_ALL;
```

```
set
```

```
%do I=2011 %to 2014 %by 1;
```

```
    CAT.A_KTN_2011_hpue (keep=site year week  
tking tcoho tpink tchum thalibut trockfish)
```

```
%end;
```

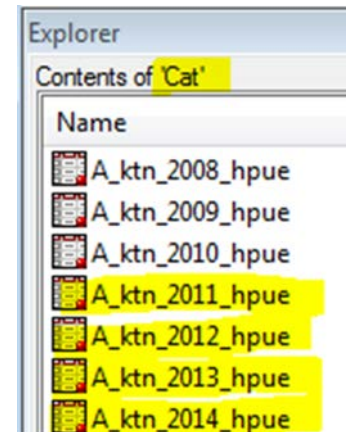
```
;
```

```
run;
```

- `%let N=A_KTN_;`

- `INLIB=CAT`

```
CAT.A_KTN_2011_hpue;
```



Set Statement – what we will get

```
DATA A_KTN_ALL;
```

```
    set
```

```
CAT.A_ktn_2011_hpue(keep=site year week tking tcoho tpink  
                    tchum thalibut trockfish)
```

```
CAT.A_ktn_2012_hpue(keep=site year week tking tcoho tpink  
                    tchum thalibut trockfish)
```

```
CAT.A_ktn_2013_hpue(keep=site year week tking tcoho tpink  
                    tchum thalibut trockfish)
```

```
CAT.A_ktn_2014_hpue(keep=site year week tking tcoho tpink  
                    tchum thalibut trockfish);
```

```
run;
```

Set up excel output sheet

- Locate within outer macro, before starting inner macro.
- We want one excel output file per site.

```
ods tagsets.ExcelXP file="S:\creel\Catch_Rates\Multi Year\  
&SITE._HPUE_&YEAR1._&YEAR4._&sysdate..xml"  
style=meadow options(embedded_titles='yes'  
                      embedded_footnotes='yes');
```

Inner Macro

```
%Macro Name ( spec ) ;
```

```
Data &N&spec ;
```

```
    set &N.All ;
```

```
keep site year week t&spec ;
```

```
run ;
```

```
proc sort data=&N&spec ;
```

```
    by SITE WEEK ;
```

```
run ;
```

```
%mend NAME ;
```

```
%NAME ( KING )
```

```
%NAME ( COHO )
```

```
%NAME ( PINK )
```

```
%NAME ( CHUM )
```

```
%NAME ( HALIBUT )
```

```
%NAME ( ROCKFISH )
```

Inner Macro

```
%Macro Name (KING) ;
```

```
Data A_KTN_KING ;
```

```
    set A_KTN_All ;
```

```
keep site year week tKING ;
```

```
run ;
```

```
proc sort data=A_KTN_KING ;
```

```
    by SITE WEEK ;
```

```
run ;
```

```
%mend NAME ;
```

```
%NAME (KING)
```

```
%NAME (COHO)
```

```
%NAME (PINK)
```

```
%NAME (CHUM)
```

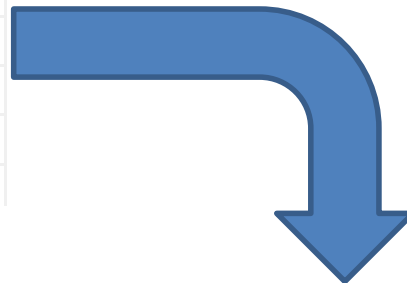
```
%NAME (HALIBUT)
```

```
%NAME (ROCKFISH)
```

Now Transpose Data

```
proc transpose data=&N&spec  
                out=&N&spec._out(drop=_NAME_) prefix=T&spec._;  
  by SITE week;  
  id YEAR;  
run;
```

year	WEEK	TCOHO
2013	8	0.0471612322
2014	8	0
2011	9	0.0013943883
2012	9	0.03155885
2013	9	0.0430751174

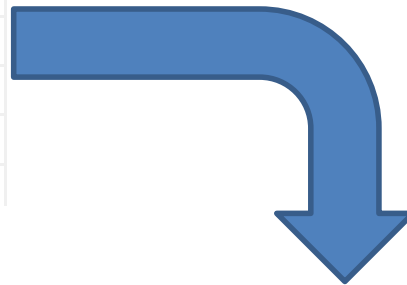


WEEK	TCOHO_2012	TCOHO_2011	TCOHO_2013	TCOHO_2014
9	0.03155885	0.0013943883	0.0430751174	0.0076233964
10	0.0134246372	0.0150518926	0.1310942326	0.0101991758
11	0.0520222833	0.0959246557	0.2245366322	0.05994144
12	0.1126584225	0.0812614658	0.3080735562	0.1209284623

Now Transpose Data

```
proc transpose data=A_KTN_KING  
                out=A_KTN_KING_out(drop=_NAME_) prefix=TKING_  
  by SITE week;  
  id YEAR;  
run;
```

year	WEEK	TCOHO
2013	8	0.0471612322
2014	8	0
2011	9	0.0013943883
2012	9	0.03155885
2013	9	0.0430751174



WEEK	TCOHO_2012	TCOHO_2011	TCOHO_2013	TCOHO_2014
9	0.03155885	0.0013943883	0.0430751174	0.0076233964
10	0.0134246372	0.0150518926	0.1310942326	0.0101991758
11	0.0520222833	0.0959246557	0.2245366322	0.05994144
12	0.1126584225	0.0812614658	0.3080735562	0.1209284623

Proc Means & Merge Data

- Some surveys start later than others and the way weeks are calculated, there may be some weeks with no survey for many of the years surveyed.
- Utilize Proc Means as it analyses the non-missing values
- Merge yearly data with proc means data

Set missing data to 'N'

- For the visual report, we want to denote between
 - HPUE of 0: survey but no harvest and
 - HPUE of missing i.e. 'N': No survey conducted, no data

```
If t&spec._&year1=. then t&spec._&year1=.N;
```



```
If tKING_2011=. then tKING_2011 =.N;
```

Calculate Hours per fish

- Most anglers do not understand catching $.025468$ of a fish per 1 rod hour of effort so we report the reciprocal which roughly gives the numbers of hours you will need to fish to get one fish.

Output to Site specific excel file

- Output 1 sheet/tab per species into the same excel file – why the initial ods tagset.ExcelXP file= is outside the inner species macro call
- But the sheet call is within the inner macro
 - `ods tagsets.ExcelXP options(sheet_name="&SPEC");`
- Sets up the excel file for the port, then as works through each species within the macro, outputs that species information into that species tab.

Finish off

- Once all %NAME calls (by species) are complete
- Close ods tagsets.ExcelXP for that Site
- Print log to file if requested (coding to print prior to first data statement) or
- Print user macro variables to the temp log
- Close outer macro - %mend CAT_GO;
- Proceed to the next %CAT_GO macro call for the next Site and do it again.

QUESTIONS?

Diana Tersteeg
Alaska Dept. of Fish & Game
Sport Fish Division, Juneau
diana.tersteeg@alaska.gov