ATTACHMENT 1-3. Explanation of Census of Agriculture Table Generation

See attached memorandum, Explanation of Census of Agriculture Table Generation, from the Biological and Economic Analysis Division.

OFFICE OF CHEMICAL SAFETY
AND POLLUTION PREVENTION

February 26, 2020

MEMORANDUM

SUBJECT: Explanation of Census of Agriculture Table Generation

FROM: Claire Paisley-Jones, Biologist
Science Information and Analysis Branch
Biological and Economic Analysis Division (7503-P)

THRU: Hope Johnson, Acting Chief
Science Information and Analysis Branch
Biological and Economic Analysis Division (7503-P)

TO: Kristina Garber, Senior Science Advisor
Environmental Risk Branch II
Environmental Fate and Effects Division (7507P)

Explanation of Census of Agriculture Table Generation

"The Census of Agriculture is a complete count of U.S. farms and ranches and the people who operate them. Even small plots of land - whether rural or urban - growing fruit, vegetables or some food animals count if $1,000 or more of such products were raised and sold, or normally would have been sold, during the Census year. The Census of Agriculture, taken once every five years, looks at land use and ownership, operator characteristics, production practices, income and expenditures. For America's farmers and ranchers, the Census of Agriculture is their voice, their future, and their opportunity."
https://www.nass.usda.gov/AgCensus/

In addition to the published narrative reports and tables, National Agricultural Statistics Service (NASS) also makes the underlying raw data available to the public on their quick stats developer page (ftp://ftp.nass.usda.gov/quickstats/). The full 2012 NASS Census of Agriculture raw data file is available at the following URL: ftp://ftp.nass.usda.gov/quickstats/qs.census2012.txt.gz. The file is very large and

contains a great deal of information that is not needed for ESA assessments. To obtain a smaller file with only the data that is needed, a series of filtering steps must be performed. The goal of the filtering process is to produce a file with only census data, which describes individual crops and production practices at an appropriate non-duplicative level of aggregation (with specific production and utilization practices, sub-classes of crop, and terms which aggregate multiple crops are all filtered out), that shows the use statistics for the available value closest to total acres grown for each crop. The final filtered file has non-duplicative data for **Acres** and **Number of Operations** at different geographic levels (national/state/county) for each available crop. The filtering steps to obtain this file are described below.

To filter for only Acres and Operations for all production practices or equivalent term:

ORDER of filtering Steps:	Column	Selection
1	SOURCE_DESC	CENSUS
2	SECTOR_DESC	CROPS
3	PRODN_PRACTICE_DESC	ALL PRODUCTION PRACTICES
		IN THE OPEN
4	UNIT_DESC	ACRES
		OPERATIONS
5	DOMAIN_DESC	TOTAL
6	DOMAINCAT_DESC	NOT SPECIFIED

To determine the appropriate level crop/commodity description for each individual crop:

STEP 7: There are four pieces of data that determine the full crop name: **COMMODITY_DESC, CLASS_DESC, PRODN_PRACTICE_DESC, and UTIL_PRACTICE_DESC**. To prevent double counting terms, all four of these terms must be considered to determine which terms are inclusive of other available terms.

The basic crop name is determined by the combination of two columns: **COMMODITY_DESC, CLASS_DESC**. To prevent double counting here, preferentially selected terms that list ALL CLASSES in the **CLASS_DESC** column. If no ALL CLASSES terms exist, then the other terms in that column will be used.

For example, for cotton, select ALL CLASSES rather than PIMA and UPLAND.

COMMODITY_DESC	COMMODITY_DESC
COTTON	ALL CLASSES
~~COTTON~~	~~PIMA~~
~~COTTON~~	~~UPLAND~~

Whereas for peas, no ALL CLASSES term exists, so the individual class terms are selected.

COMMODITY_DESC	COMMODITY_DESC
PEAS	AUSTRIAN WINTER
PEAS	DRY EDIBLE
PEAS	DRY, SOUTHERN (COWPEAS)

STEP 8: There are two additional columns **PRODN_PRACTICE_DESC** and **UTIL_PRACTICE_DESC** that modify the crop/class term. In general, to prevent double counting here, preferentially use ALL PRODUCTION PRACTICES and ALL UTILIZATION PRACTICES class production and utilization. If no, ALL terms exist, then the other terms in that column will be used.

For example, for corn, use ALL PRODUCTION PRACTICES and ALL UTILIZATION practices exist, so those terms are used rather than the grain and silage terms.

COMMODITY_ DESC	CLASS_ DESC	PRODN_PRACTICE_ DESC	UTIL_PRACTICE_ DESC
CORN	ALL CLASSES	**ALL PRODUCTION PRACTICES**	**ALL UTILIZATION PRACTICES**
~~CORN~~	~~ALL CLASSES~~	~~ALL PRODUCTION PRACTICES~~	~~GRAIN~~
~~CORN~~	~~ALL CLASSES~~	~~ALL PRODUCTION PRACTICES~~	~~SILAGE~~

Whereas for foliage plants there is no ALL PRODUCTION PRACTICES, so IN THE OPEN is used.

COMMODITY_ DESC	CLASS_ DESC	PRODN_PRACTICE_ DESC	UTIL_PRACTICE_ DESC
FOLIAGE PLANTS	INDOOR USE	IN THE OPEN	**ALL UTILIZATION PRACTICES**

Similarly, for sugarcane there is no ALL UTILIZATION PRACTICES term, so SEED and SUGAR are used.

COMMODITY_ DESC	CLASS_ DESC	PRODN_PRACTICE_ DESC	UTIL_PRACTICE_ DESC
SUGARCANE	ALL CLASSES	**ALL PRODUCTION PRACTICES**	SEED
SUGARCANE	ALL CLASSES	**ALL PRODUCTION PRACTICES**	SUGAR

After this is done, larger group totals that represent the aggregation of multiple crops also need to be removed (e.g. ORCHARDS and NURSERY TOTALS). As a part of this step, checks were done to ensure the individual crops retained added up to the aggregate terms that were removed.

STEP 9: Below is a list of terms that are returned by the filtering steps above, and must be excluded manually from the final run.

GROUP_DESC	COMMODITY_DESC	COMMODITY_DESC (CLASS_DESC)	UTIL_PRACTICE _DESC
field crops	delete anything with "TOTAL" in the commodity name, as well as "Hay & Haylage", and "Small Grains"	delete: Cotton (pima, and upland); Hay (alfalfa, small grain, tame, and wild); Haylage (alfalfa, and excl alfalfa); Mint (peppermint and spearmint); Sunflowers (non-oil type, and oil type); Wheat (spring excl durum, spring durum, and winter)	delete: Sugarbeet (seed); Corn (grain, and silage)
vegetables	delete anything with "TOTAL" in the commodity name	delete: Lettuce (head, leaf, and romaine); and Squash (summer, and winter)	--
fruit & tree nuts	delete anything with "TOTAL" in the commodity name	--	--
fruit & tree nuts	delete anything with "TOTAL" in the commodity name	delete: Oranges (mid & navel, and valencia); Pecans (improved, and native & seedling); Pears (bartlett and exl bartlett); Peaches (clingstone and freestone); Plums and Prunes.	--

The resulting terms were compared to the terms in the full census file, filtered only for Acres to ensure that all relevant **COMMODITY_DESC (CLASS_DESC)** combinations were captured in the final file.

To determine the available use statistic closest to total acres grown:
STEP 10: Finally, the STATISTIC CAT_DESC value for each GROUP_DESC and Crop to determine the available metric that is as close to acres grown as possible. This following selections should be made for each group:

AREA HARVESTED
- FIELD CROPS
- VEGETABLES
- HORTICULTURE - SOD

AREA GROWN
- FRUIT & TREE NUTS – *(all the berries)*

AREA BEARING & NON-BEARING
- FRUIT & TREE NUTS – *(all fruit and nut trees)*

AREA IN PRODUCTION
- HORTICULTURE – *(all crops except sod)*

Note: AREAS HARVESTED figures are available in multiple groups, but is not the closest to total area grown for all groups.

The file resulting from these steps contains non-duplicative crop information at various geographic levels (county/state/national). Use statistics (acres and number of operations) for crops are provided at the highest level available (i.e. the level closest to total acres grown).

Step 11: Following this filtering process, data should be split into two work sheets by UNIT_DESC: one for ACRES and one for OPERATIONS. Either of these tables can be used in the following step, as both produce the same values after transformation. Using either table, create a new pivot table with crops as the columns and geographic areas as the rows. Transform the data in this table using the following function:

> IF the value in the cell is > 0, report 1
> IF the value in the cell is 0, report 0

The resulting table provides presence/absence information for all crops in all geographic areas. If a crop is reported to be grown in the area according to the Census, it will be assigned a 1. If the crop is not reported, it will be assigned a 0.