cirlib 5

1967 CENSUS ESTABLISHMENT SIZE DISTRIBUTION DATA

ABSTRACT: This data set contains the 1967 Census statistics by establishment employment size classes for the 2- and 4-digit Census manufacturing industries. Data is also included for "all industries" under the 1-digit code "1". The data is often aggregated to avoid revealing confidential firm information.

JCL:
```
//GO.FT08F001 DD DSN=ESTSIZ67,DISP=OLD,UNIT=TAPE6250,
// DCB=(RECFM=FB,LRECL=80,BLKSIZE=3520),
// LABEL=(05,SL,,IN),VOL=SER=DATABK DTAD63
```
41

DATA SET ORGANIZATION: Each record is uniquely associated with a particular size class in a particular industry. Records are sorted according to the following major to minor sequence:

1. Census SIC
2. Size class code

The sort is in ascending sequence, with the following qualifications:

1. The 2-digit codes precede the 4-digit codes. The data set begins with an industry with SIC=1 that represents all manufacturing industries.
2. SIC 19 follows all the 2-digit codes; all SIC's in the 1900's follow the 4-digit codes.

The number of records per industry is variable because empty size classes are not given. However, industry totals and class 11 are always present. The size class codes are as follows:

Size class	Code
Industry totals	0
1-4	1
5-9	2
10-19	3
20-49	4
50-99	5
100-249	6
250-499	7
500-999	8
1,000-2,499	9
2,500 and over	10
establishments covered by administrative records	11

OF RECORDS: 4719 total records. 441 industries, but the number of records per industry is variable

FORMAT: (I4,2X,I2,1X,I1,F6.0,F6.1,F7.1,3F6.1,3F7.1,2F6.1)

VARIABLES:

1.	Census industrial code (SIC)	
2.	Size class code	
3.	Aggregation code (see note 1 below)	
4.	Number of establishments	
5.	Number of employees	(thousands)
6.	Payroll	($ millions)
7.	Number of production workers	(thousands)
8.	Man-hours	(millions)
9.	Wages	($ millions)
10.	Value added	" "
11.	Cost of materials	" "
12.	Value of shipments	" "
13.	Capital expenditures	" "
14.	End-of-year inventories	" "

SOURCE: U.S. Bureau of the Census. Census of Manufactures: 1967 Subject Statistics: Size of Establishments, MC67(1)-2. U.S. Government Printing Office: Washington, D.C., 1970.

This report is a preprint of a corresponding chapter of U.S. Census of Manufactures: 1967, Vol. 1, Summary and Subject Statistics

NOTES: 1. Aggregation--As noted earlier, some size classes are aggregated with others to prevent disclosure of confidential firm information. In the Census report, such aggregated data is recorded in the column for the smallest size class of the group and is indicated by underscoring. The subsegent size classes in the grouping are coded by the letter "D". Our corresponding aggregation codes are:

1. All values of the variables 5 through 9 of this size class are aggregates of the data of this and one or more subsequent size classes. (Indicated in the original source by a column of five underscored figures.)

3. The data for variables 5 through 9 of this size class is included i in a previous aggregation. (Indicated in the original source by a

column of five D's.)

2,4. Analogous to codes 1 and 3 for variable 9 (capital expenditures) only. (Indicated in the report by an underscoring and a subsequent D for capital expenditures.)

5. All values in variables 5 through 9 of this size class are aggregates of the data of this and one or more subseqent size classes. In addition, the data for variable 9 (capital expenditures) is included in a previous aggregation. (Indicated in the report by a column of four underscored figures, and a D for capital expenditures.)

6. All values for size class code 11 are suppressed. The data are included in figures for the two smallest size classes.

If both codes 3 and 4 are valid, code 3 is recorded.

2. For some size classes there are no data observations. There are no records for empty size classes. Suppressed values, indicated by the "D" code, are recorded as zeroes.

3. The size distribution data was summed and checked against the industry totals. A number of errors were found. The Bureau of the Census was able to correct most of these errors including all the major ones.

4. Report forms were not sent to companies that operated only one establishment and which Social Security payrolls indicated had fewer than 10 employees. Actual payrolls (and sales) for 1967 were obtained from administrative records of the Federal Government. The other statistics from these establishments were estimated from industry averages.

5. The Z code (value of variable is greater than zero but rounds to zero at the unit presented) is represented by a -1.

6. The 1967 data sets contain all the variables that the 1963 data sets contain, plus more, but in different units of measurement. Variables numbered 5, 6, 10, 12, and 13 do not have the same units in both 1963 and 1967.

7. The data is often aggregated to avoid revealing confidential firm information. We have developed a technique for disaggregating the aggregated data--the technique is described in separate documentation. The JCL for the data set containing our estimates of the true values of the aggregated data is given below:

```
//GO.FT08F001 DD DSN=SIZ67MCD,DISP=OLD,UNIT=TAPE6250,
// DCB=(RECFM=FB,LRECL=90,BLKSIZE=7290),
// LABEL=(06,SL,,IN),VOL=SER=DATABK
```

cirlin 6

1967 CENSUS ESTABLISHMENT SIZE DISTRIBUTION DATA, DISAGGREGATED

ABSTRACT: This data set contains the 1967 Census statistics by establishment employment size classes for the 2- and 4-digit Census manufacturing industries, but differs from the original data in one important respect. In the original data (available in another data set), the data is often aggregated to avoid revealing confidential firm information. We have developed a technique for disaggregating the aggregated data--the technique is described in separate documentation--and this data set contains our estimates of the true values of the aggregated variables.

JCL:
```
//GO.FT08f001 DD DSN=SIZ67MCD,DISP=OLD,UNIT=TAPE6250,
// DCB=(RECFM=FB,LRECL=90,BLKSIZE=7290),
// LABEL=(06,SL,,IN),VOL=SER=DATABK
```
42 DTA063

DATA SET ORGANIZATION: Each record is uniquely associated with a particular size class in a particular industry. Records are sorted according to the following major to minor sequence:

1. Census SIC
2. Size class code

The sort is in ascending sequence, with the following qualifications:

1. The 2-digit codes precede the 4-digit codes. The data set begins with an industry with SIC=1 that represents all manufacturing industries.
2. SIC 19 follows all the 2-digit codes; all SIC's in the 1900's follow the 4-digit codes.

For each industry, there are 11 records--one record for each of 10 size classes, plus an additional record for the industry totals. The size class codes are as follows:

Size class	Code
Industry totals	0
1-4	1
5-9	2
10-19	3
20-49	4
50-99	5
100-249	6
250-499	7
500-999	8
1,000-2,499	9

2,500 and over 10

 # OF RECORDS: 11 records per industry, 441 industries, and hence 4851 total records

FORMAT: (I4,1X,I2,1X,I1,F7.0,F7.1,F8.1,3F7.1,3F8.1,2F7.1)

VARIABLES:

1.	Census industrial code (SIC)	
2.	Size class code	
3.	Aggregation code (see note 1 below)	
4.	Number of establishments	
5.	Number of employees	(thousands)
6.	Payroll	($ millions)
7.	Number of production workers	(thousands)
8.	Man-hours	(millions)
9.	Wages	($ millions)
10.	Value added	" "
11.	Cost of materials	" "
12.	Value of shipments	" "
13.	Capital expenditures	" "
14.	End-of-year inventories	" "

SOURCE: U.S. Bureau of the Census. Census of Manufactures: 1967. Subject Statistics: Size of Establishments, MC67(1)-2. U.S. Government Printing Office: Washington, D.C., 1970.

This report is a preprint of a corresponding chapter of U.S. Census of Manufactures: 1967, Vol. 1, Summary and Subject Statistics. A copy of the original data is available on magnetic tape; the JCL is given below:

```
//GO.FT08F001 DD DSN=ESTSIZ67,DISP=OLD,UNIT=TAPE6250,
// DCB=(RECFM=FB,LRECL=80,BLKSIZE=3520),
// LABEL=(05,SL,,IN),VOL=SER=DATABK
```

NOTES: 1. Aggregation--As noted earlier, all aggregated data has been disaggregated. This data set retains the aggregation coding used in the original data set in order to indicate where the disaggregation took place.

1. All values of the variables 5 through 9 of this size class are aggregates of the data of this and one or more subsequent size classes. (Indicated in the original source by a column of five underscored figures.)

3. The data for variables 5 through 9 of this size class is included i in a previous aggregation. (Indicated in the original source by a column of five D's.)

2,4. Analogous to codes 1 and 3 for variable 9 (capital expenditures) only. (Indicated in the report by an underscoring and a subsequent D for capital expenditures.)

5. All values in variables 5 through 9 of this size class are aggregates of the data of this and one or more subsegent size classes. In addition, the data for variable 9 (capital expenditures) is included in a previous aggregation. (Indicated in the report by a column of four underscored figures, and a D for capital expenditures.)

If both codes 3 and 4 are valid, code 3 is recorded.

2. For some size classes there are no data observations. Variables 3 through 9 in the corresponding records contain zeroes. Missing values are recorded as zeroes. (In the original data set, empty size classes are not recorded.)

3. The original size distribution data was summed and checked against the industry totals. A number of errors were found. The Bureau of the Census was able to correct most of these errors including all the major ones.

4. Report forms were not sent to companies that operated only one establishment and which Social Security payrolls indicated had fewer than 10 employees. Actual payrolls (and sales) for 1967 were obtained from administrative records of the Federal Government. The other statistics from these establishments were estimated from industry averages.

5. The Z code (value of variable is greater than zero but rounds to zero at the unit presented) is represented by a -1. Included in this category are variables that have been made too small (less than 0.05) as a result of the disaggregation procedure.

6. There are three differences between the organization of this data and the original 1967 data:

first, data that was included in the original 1967 Census data for "establishments covered by administrative records" has been omitted in this data set; second, size classes with no data were omitted in the preparation of the original 1967 Census data, but in this data set, empty size classes are represented by records containing the SIC and size class codes and zeroes in all other variables. Thus, the number of records is not the same for the two 1967 Census data sets. And third, the format of the variables has been changed.

7. The 1967 data sets contain all the variables that the 1963 data sets contain, plus more, but in different units of measurement. Variables numbered 5, 6, 10, 12, and 13 do not have the same units in both 1963 and 1967.

8. Persons desiring to know the exact procedure used to disaggregate the aggregated variables should see the document: "The Census Size Distribution Disaggregation Procedure for Undisclosed Data."