Using SAS In Generating Multi-Jurisdictional Reports for U.S. Retail Industries' Employment Trends

Abdullah Khan

Associate Professor of Economics,
School of Business, Claflin University
akhan@claflin.edu

Abstract

This paper demonstrates how the SAS ODA web application can be used to combine historical data files into a large data set so that multi-jurisdictional trends of interest can be calculated programmatically. The U.S. retail industry as a whole serve as the second largest source of private-sector employment. The spatial analysis of employment trends in this sector is therefore important for both academic and industry insights. The prime objective of this paper is to demonstrate how SAS ODA can be used to generate tabular reports, maps, and multivariate scatterplots in depicting employment trends for various U.S. retail industries at various jurisdictional levels using the County Business Patterns data of the Census Bureau for the years 2001 through 2021. This paper presents some snippets of SAS codes and output depicting the employment trends for the U.S. retail sector at the county level, state level, census division level, census region level, as well as national level. SAS is a versatile program which can be used for high quality industrial data analytics and visualization.

Key words: *SAS, Retail industry, employment trends, data analytics, data visualization*

INTRODUCTION

According to the U.S. Census Bureau[1] and the Bureau of Labor Statistics (BLS), the retail industry currently is the second largest employment source in the private sector preceded only by the service sector. In terms of employment, the U.S. retail industry surpassed the manufacturing industry several years ago. According to the BLS data, as of August 2023, total employment in the retail sector was 15.54 million whereas the manufacturing sector's total employment was less than thirteen million. Currently the retail sector provides 10.1 percent of all non-farm employment in the USA[2]. The employment landscapes for the retail industries are changing rapidly due to the exponential rise of e-commerce facilitated by expanding coverage internet access (Khan, 2023a) and the rapid emergence of multi-channel and omni-channel retail business practices (Khan, 2023b; Hanninen, Kwan, and Mitronen, 2021; Liu, Lobschat and Verhoef, 2018). Besides, the recent COVID-19 pandemic-related exigencies also prompted businesses and consumers to utilize digital transactions with an increased sense of urgency. Despite the historic prominence of the U.S. manufacturing industries as one of the most important non-farm employment-generating sectors, this sector's employment growth rates in recent years have been lackluster (Acemoglu et al., 2016; Pierce and Schott, 2016; Rider and Khan, 2010). Thus, the study of employment trends in the retail industry as a whole and sub-sectors is of renewed interest to both academics and industrial practitioners.

[1] U.S. Census Bureau, County Business Pattern Database (https://www.census.gov/programs-surveys/cbp/data.html)
[2] U.S. Bureau of Labor Statistics report.

SAS CODES

Creation of multi-jurisdictional reports using total retail employment count and retail establishment count (NAICS codes 44-45):

Sample codes for step 1:
The following sample codes are used in importing comma-delimited data files into SAS ODA library for the years 2001 through 2021. Please note that the .txt files were downloaded from the county business pattern database and were uploaded on the SAS ODA. For space-saving purposes, a subset of SAS codes is shared below as a sample:

```
PROC IMPORT DATAFILE='/home/u1341415/test/cbp01st.txt'
OUT=SESUG23.cbp01st
DBMS=dlm REPLACE;
DELIMITER= ",";
GETNAMES=YES;
RUN;
;
;
```

*NOW READ 2021 DATASET;

```
PROC IMPORT DATAFILE='/home/u1341415/test/cbp21st.txt'
OUT=SESUG23.cbp21st
DBMS=dlm REPLACE;
DELIMITER= ",";
GETNAMES=YES;
RUN;
```

Sample SAS codes for step 2:
Add "year" variables and enter corresponding year values into these imported SAS files for the years 2001 through 2021 as follows:

```
*ENTER YEAR 2001;

DATA SESUG23.CBP01ST;
SET SESUG23.CBP01ST;
LENGTH YEAR 4.;
YEAR=2001;
RUN;
;
;

DATA SESUG23.CBP01ST;
SET SESUG23.CBP01ST;
LENGTH YEAR 4.;
YEAR=2021;
RUN;
```

Sample SAS codes for Step 3:
Now I take a subset of only retail sector data from all sector data for the years 2001 through 2021 as follows:

```
DATA SESUG23.RETAILTOT_01ST;
SET SESUG23.cbp01st;
WHERE NAICS="44----";
KEEP YEAR FIPSTATE EMP EST NAICS;
RUN;
;
;
DATA SESUG23.RETAILTOT_10ST;
SET SESUG23.cbp10st;
WHERE NAICS="44----" & LFO = “ -”;
KEEP YEAR FIPSTATE EMP EST NAICS;
RUN;
;
;

DATA SESUG23.RETAILTOT_21ST;
SET SESUG23.cbp21st;
WHERE NAICS="44----" & LFO= "-" ;
KEEP YEAR FIPSTATE EMP EST NAICS;
RUN;
```

[Special note: When we are dealing with historical data over several years or decades, it is imperative that we pay special attention to the data dictionary for each of these years. For example, in the above SAS code, one can notice that total retail employment and establishment data extraction codes between 2010 and 2021 were different compared to those between 2001 and 2009. [3]]

Sample SAS codes for Step 4:

In this step, we concatenate all these files containing total employment and establishment count as follows:

```
DATA SESUG23.RETAILTOT_Y0121ST;
SET SESUG23.RETAILTOT_Y01ST
;
;
SESUG23.RETAILTOT_Y21ST
RUN;
```

[3] Since year 2010, county business pattern data started reporting the data into various Legal Forms of Ownership (LFO). Therefore, we need to filter data appropriately so that we get total employment count and business establishment count across all types of LFOs.

[Special note: Sometimes concatenation of large files for several years may require taking a ‘gradual approach.’ For example, to concatenate twenty years of data into one file, one could first concatenate year 1 through year 7 in a file, then add another seven years of data, and then can adjoin the remainder of the years to get all the twenty years of data in to one mega concatenated file.]

Sample SAS codes for Step 5:

In this step the group variable variables ‘census division’ and ‘census region’ are created.

Sample SAS code for this step:

```
DATA SESUG23.RETAILTOT_Y0121ST;
SET SESUG23.RETAILTOT_Y0121ST;
LENGTH CENREGIONS4 $50 NINECENSDIV$30;

IF FIPSTATE = '09' OR FIPSTATE = '23'
OR FIPSTATE = '25' OR FIPSTATE= '33' OR FIPSTATE = '44' OR FIPSTATE = '50'
THEN NINECENSDIV = 'NEW ENGLAN;
;
;
IF NINECENSDIV = 'MOUNTAIN' OR NINECENSDIV= 'PACIFIC' THEN CENREGIONS4=
'R4WEST';
RUN;
```

Sample SAS codes for Step 6:

In this step, we use the PROC REPORT command to generate reports of aggregate retail employment.

```
*NOW CREATE PROC REPORT and proc rank FILE BY YEAR NINECENSDIV AND THEN BY
CENREGIONS4 SEQUENTIALLY;

PROC REPORT DATA=SESUG23.RETAILTOT_Y0121ST
OUT=SESUG23.RETAILTOT_Y0121NINECENSDIVRPRT NOWD;
COL YEAR NINECENSDIV EMP EST;
DEFINE YEAR / ANALYSIS GROUP;
DEFINE NINECENSDIV / ANALYSIS GROUP;
DEFINE EMP / ANALYSIS SUM;
DEFINE EST / ANALYSIS SUM;
RUN;

*now create proc report and proc rank by 4 census regions;

PROC REPORT DATA=SESUG23.RETAILTOT_Y0121ST
OUT=SESUG23.RETAILTOT_Y0121FOURCENREGIONS NOWD;
COL YEAR CENREGIONS4 EMP EST;
```

```
DEFINE YEAR / ANALYSIS GROUP;
DEFINE CENREGIONS4 / ANALYSIS GROUP;
DEFINE EMP / ANALYSIS SUM;
DEFINE EST / ANALYSIS SUM;
RUN;

*PROC REPORT AND PROC RANK FOR NATIONWIDE (US LEVEL);

PROC REPORT DATA=SESUG23.RETAILTOT_Y0121ST OUT=SESUG23.RETAILTOT_Y0121USTOT
NOWD;
COL YEAR EMP EST ;
DEFINE YEAR / ANALYSIS GROUP;
DEFINE EMP / ANALYSIS SUM;
DEFINE EST / ANALYSIS SUM;
RUN;
```

Creation of multi-jurisdictional reports using retail employment count and retail establishment count at 3-digit NAICS code levels (NAICS codes 441, 442, 443, 444, 445, 446, 447, 448, 451, 452, 453, and 454):

Sample SAS codes for Step 1: Removal of special character for the data set, extraction of NAICS 3-digit level data extraction for the years 2001 through 2021 as follows:

*Remove special characters from 2001 state level data set;

```
DATA SESUG23.cbp01st;
SET SESUG23.cbp01st;
NAICS2= INPUT(COMPRESS(NAICS, "/"), COMMA6.);
RUN;
```

*Now extract 2001 3-digit retail data from the state level data set;

```
DATA SESUG23.RETAILTOT3DGT_01ST;
SET SESUG23.cbp01st;
WHERE NAICS2> 439 AND NAICS2<460;
KEEP YEAR FIPSTATE EMP EST NAICS NAICS2;
RUN;
;
;

*NOW REMOVE SPECIAL CHARACTERS FOR 2021;
DATA SESUG23.cbp21st;
SET SESUG23.cbp21st;
NAICS2= INPUT(COMPRESS(NAICS, "/"), COMMA6.);
RUN;

*NOW EXTRACT 3-dgt retail emp est data FILE FOR 2021;
DATA SESUG23.RETAILTOT3DGT_21ST;
```

```
SET SESUG23.cbp21st;
WHERE NAICS2> 439 AND NAICS2<460 AND LFO = "-";
KEEP YEAR FIPSTATE EMP EST NAICS NAICS2;
RUN;
```

Sample SAS codes for Step 2: Now let us concatenate all these yearly files using a data step as follows.

```
DATA SESUG23.RETAILTOT3DGT_Y0104ST;
SET SESUG23.RETAILTOT3DGT_01ST
SESUG23.RETAILTOT3DGT_02ST
;
;
SESUG23.RETAILTOT3DGT_21ST;
RUN;
```

Sample SAS codes for Step 3: Create census division variable and Census region variable in the concatenated file as follows:

```
DATA SESUG23.RETAILTOT3DGT_Y0121ST; SET SESUG23.RETAILTOT3DGT_Y0121ST;
LENGTH CENREGIONS4 $50 NINECENSDIV$30;
 IF FIPSTATE = '09' OR FIPSTATE = '23' OR FIPSTATE = '25' OR FIPSTATE= '33'
OR FIPSTATE = '44' OR FIPSTATE = '50' THEN NINECENSDIV = 'NEW ENGLAND';
IF FIPSTATE = '34' OR FIPSTATE = '36' OR FIPSTATE = '42' THEN NINECENSDIV =
'MIDATLANTIC'; IF FIPSTATE = '18' OR FIPSTATE = '17' OR FIPSTATE = '26' OR
FIPSTATE = '39' OR FIPSTATE = '55' THEN NINECENSDIV = 'ENORTHCENTRAL';
IF FIPSTATE = '19' OR FIPSTATE = '20' OR FIPSTATE = '27' OR FIPSTATE= '29' OR
FIPSTATE = '31' OR FIPSTATE = '38' OR FIPSTATE = '46' THEN NINECENSDIV =
'WNORTHCENTRAL';
;
;
;
IF NINECENSDIV = 'ENORTHCENTRAL' OR NINECENSDIV= 'WNORTHCENTRAL' THEN
CENREGIONS4= 'R2MIDWEST';
IF NINECENSDIV = 'SATLANTIC' OR NINECENSDIV= 'ESOUTHCENTRAL' OR NINECENSDIV=
'WSOUTHCENTRAL' THEN CENREGIONS4='R3SOUTH';
IF NINECENSDIV = 'MOUNTAIN' OR NINECENSDIV= 'PACIFIC' THEN CENREGIONS4=
'R4WEST';
RUN;
```

Sample SAS code for step 4: Now create report of retail employment count and retail business establishment count across nine census divisions, four census regions and at the national level (USA-level aggregate) as follows:

*Now create report using PROC REPORT for nine census divisions;

```
PROC REPORT DATA=SESUG23.RETAILTOT3DGT_Y0121ST
OUT=SESUG23.RETAILTOT3DGT_Y0121NINECENSDIVS NOWD;
COL YEAR NAICS3 NINECENSDIV EMP EST;
DEFINE YEAR / ANALYSIS GROUP;
```

```
DEFINE NAICS3 / ANALYSIS GROUP;
DEFINE NINECENSDIV / ANALYSIS GROUP;
DEFINE EMP / ANALYSIS SUM;
DEFINE EST / ANALYSIS SUM;
RUN;
```

*Now create proc report for four census regions;

```
PROC REPORT DATA=SESUG23.RETAILTOT3DGT_Y0121ST
OUT=SESUG23.RETAILTOT3DGT_Y0121FOURCENREGION NOWD;
COL YEAR NAICS3 CENREGIONS4 EMP EST;
DEFINE YEAR / ANALYSIS GROUP;
DEFINE NAICS3/ANALIYSIS GROUP;
DEFINE CENREGIONS4 / ANALYSIS GROUP;
DEFINE EMP / ANALYSIS SUM;
DEFINE EST / ANALYSIS SUM;
RUN;
```

*Now create proc report for U.S. national total retail employment across subsectors;

```
PROC REPORT DATA=SESUG23.RETAILTOT3DGT_Y0121ST
OUT=SESUG23.RETAILTOT3DGT_Y0121USTOTAL NOWD;
COL YEAR NAICS3 EMP EST;
DEFINE YEAR / ANALYSIS GROUP;
DEFINE NAICS3/ANALIYSIS GROUP;
DEFINE EMP / ANALYSIS SUM;
DEFINE EST / ANALYSIS SUM;
RUN;
```

Sample SAS code for county level report: county level report for the year 2021;

Step 1: extract county level retail employment data for the year 2021 and then aggregate NAICS 4-digit level data into NAICS 3-digit level data as follows:

*First remove special characters for 2021 county level data file;

```
 DATA SESUG23.cbp21co;
SET SESUG23.cbp21co;
NAICS4= INPUT(COMPRESS(NAICS, "/"), COMMA6.);
YEAR='2021';
RUN;
```

*NOW EXTRACT 4-dgt retail emp est data file for 2021;

```
DATA SESUG23.RETAIL4DGT_21COUNTYLEVEL;
SET SESUG23.cbp21co;
WHERE NAICS4> 4399 AND NAICS4<4600;
```

```
KEEP YEAR FIPSTATE FIPSCTY EMP EST NAICS NAICS4;
RUN;
```

*Now first truncate the last digit of the four-digit NAICS-coded data and then generate proc report at county level totals for NAICS 3-digit level;

```
DATA SESUG23.RETAIL4DGTTRUNC3DGT_21CNTY3DGT;
SET SESUG23.RETAIL4DGT_21COUNTYLEVEL;
NAICS4CHAR= PUT(NAICS4, 4.);
NAICS3CHAR= SUBSTR(NAICS4CHAR,1,3);
RUN;
```

*SAS code for creation of map using U.S. State-level total retail employment data for the year 2001.

```
PROC GMAP MAP=MAPS.US DATA= SESUG23.RETAILTOT_01XST ALL;
ID STATE;
CHORO EMP / DISCRETE;
FORMAT EMP EMP_.;
NOTE ' STATE-LEVEL RETAIL EMPLOYMENT IN THE UNITED STATES IN 2001';
RUN;
QUIT;
```

Now crate maps showing total retail employment in every U.S. states for the year 2001, 2021 (the output maps are shown in Figure 1 and Figure 2 in the result section). Also, I created a map showing the percent change in employment between 2001 and 2021 across U.S. states (this map is shown in Figure 3 in the result section). Besides, Figure 4 contains a map showing percentage changes in retail sector employment across states grouped by magnitude of percentage change in retail employment. The SAS codes used to generate these three maps are given here below and the state level retail sector total employment data for 2001 and 2021 are presented in Table A2 in the Appendix section.

```
*NOW CREATE THE STATE LEVEL MAP AS FOLLOWS;

PROC GMAP MAP=MAPS.US DATA= SESUG23.RETAILTOT_01XST ALL;
ID STATE;
CHORO EMP / DISCRETE;
FORMAT EMP EMP_.;
NOTE ' STATE-LEVEL RETAIL EMPLOYMENT IN THE UNITED STATES IN 2001';
RUN;
QUIT;

*NOW CREATE MAP FOR EMP 2021;

PROC GMAP MAP=MAPS.US DATA= SESUG23.RETAILTOT_21XST ALL;
ID STATE;
CHORO EMP / DISCRETE;
FORMAT EMP EMP_.;
NOTE ' STATE-LEVEL RETAIL EMPLOYMENT IN THE UNITED STATES IN 2021';
RUN;
QUIT;
```

```
*Now create a map showing the percent change;
DATA SESUG23.RETAILEMPTOT2001_2021_4PCCHNAGE;
SET SESUG23.RETAILEMPTOT2001_2021_4PCCHNAGE;
PC_CHANGE_EMPX = INPUT(PC_CHANGE_EMP, 10.1);
PC_CHANGE_RETAIL_EMP = INPUT(PC_CHANGE_EMPX, 4.);
run;

*NOW CREATE MAP showing percent change in retail employment between 2001 and
2021;

PROC GMAP MAP=MAPS.US DATA= SESUG23.RETAILEMPTOT2001_2021_4PCCHNAGE ALL;
ID STATE;
CHORO PC_CHANGE_RETAIL_EMP / DISCRETE;
FORMAT PC_CHANGE_RETAIL_EMP PC_CHANGE_RETAIL_EMP_.;
NOTE ' PERCENT CHANGE IN STATE-LEVEL RETAIL EMPLOYMENT BETWEEN 2001 AND
2021';
RUN;

*Recognizing the need for grouping the states by magnitude of change in
retail employment, now I created another map in which the individual states
are banded into five grouped based on range of magnitudes of changes in the
retail sector employment over two decades between 2001 and 2021. The codes
for generating this map are given below and the map is presented in Figure
no. 4.
SAS code used to generate the map presented in Figure no. 4:

PROC GMAP MAP = MAPS.US DATA= SESUG23.RETAILEMPTOT2001_2021_4PCCHNAGE;
ID STATE;
CHORO PC_CHANGE_RETAIL_EMP / levels=5 legend=legend1 coutline=gray33;
NOTE j=1 ' PERCENT CHANGE IN RETAIL EMPLOYMENT BETWEEN 2001 AND 2021'
j=1 'ACROSS INDIVIDUAL STATES GROUPED BY MAGNITUDE OF CHANGE';
RUN;
QUIT;

*Now I create a scatter plot for retail sector employment across four U.S.
census regions as follows (result is shown in Figure 5 in the results
section).
/* Creating a multivariate scatter plot over time */
proc sgplot data=SESUG23.REGION_RETAILEMP0121_4SASGRAPH;
title 'Retail employment across U.S. Census regions';
scatter x=year y=northeast_region_retail_emp /
markerattrs=(symbol=circlefilled) datalabel=northeast_region_retail_emp;
scatter x=year y=midwest_region_retail_emp /
markerattrs=(symbol=squarefilled) datalabel=midwest_region_retail_emp;
scatter x=year y=south_region_retail_emp /
markerattrs=(symbol=trianglefilled) datalabel=south_region_retail_emp;
scatter x=year y=west_region_retail_emp / markerattrs=(symbol=trianglefilled)
datalabel=west_region_retail_emp;
xaxis label='Year';
```

```
yaxis label='Retal employment across census regions';
run;
PROC PRINT;
RUN;
```

RESULTS

Snippets of SAS output regarding the multi-jurisdictional reports using total retail employment count and retail establishment count (NAICS codes 44-45):

While complete tables for selected years are provided in the appendix section of the paper, in this section I share a few screenshots of the multijurisdictional reports.

TABLE 1 A partial snippet of SAS output for all retail industries' employment and establishment counts for the nine census regions.

5-883b-401fa5fb88b0/results

YEAR	NINECENSDIV	emp	est
2001	ENORTHCENTRAL	2457090	171420
	ESOUTHCENTRAL	898262	73492
	MIDATLANTIC	1945947	159578
	MOUNTAIN	1007650	72653
	NEW ENGLAND	806123	61479
	PACIFIC	2141349	154070
	SATLANTIC	2882177	219825
	WNORTHCENTRAL	1168182	87400
	WSOUTHCENTRAL	1583509	120033
2002	ENORTHCENTRAL	2399889	170784
	ESOUTHCENTRAL	885256	73518
	MIDATLANTIC	1964599	161386
	MOUNTAIN	1013855	73440
	NEW ENGLAND	828810	61895
	PACIFIC	2119433	155224
	SATLANTIC	2878226	221890
	WNORTHCENTRAL	1148238	86744
	WSOUTHCENTRAL	1581598	120812

TABLE 2 partial snippets of SAS output for all retail industries employment and establishment count for the four census regions

da4aa-998a-44a8-8e62-658f7cf342cf/results

YEAR	CENREGIONS4	emp	est
2001	R1NORTHEAST	2752070	221057
	R2MIDWEST	3625272	258820
	R3SOUTH	5363948	413350
	R4WEST	3148999	226723
2002	R1NORTHEAST	2793409	223281
	R2MIDWEST	3548127	257528
	R3SOUTH	5345080	416220
	R4WEST	3133288	228664
2003	R1NORTHEAST	2800983	220517
	R2MIDWEST	3532380	255319
	R3SOUTH	5329536	411467
	R4WEST	3204926	228603

TABLE 3 A partial snippet of all retail industries employment and establishment count for the Nine census divisions at NAICS 3-digit level retail industries

odamid-usw2.oda.sas.com/SASStudio/sasexec/submissions/8b6cd6b9-5d64-4c95-8e3d-05308119ac91/results

Obs	YEAR	NAICS3	NINECENSDIV	emp	est	_BREAK_	NAICS3DGT_DESCRIPTION
1	2001	441	ENORTHCENTRAL	299528	19616		Motor vehicle and auto parts dealers
2	2001	441	ESOUTHCENTRAL	115190	9693		Motor vehicle and auto parts dealers
3	2001	441	MIDATLANTIC	206402	14047		Motor vehicle and auto parts dealers
4	2001	441	MOUNTAIN	139108	8532		Motor vehicle and auto parts dealers
5	2001	441	NEW ENGLAND	91301	6231		Motor vehicle and auto parts dealers
6	2001	441	PACIFIC	280340	16588		Motor vehicle and auto parts dealers
7	2001	441	SATLANTIC	362016	25083		Motor vehicle and auto parts dealers
8	2001	441	WNORTHCENTRAL	140947	11119		Motor vehicle and auto parts dealers
9	2001	441	WSOUTHCENTRAL	215386	14819		Motor vehicle and auto parts dealers
10	2001	442	ENORTHCENTRAL	88639	9608		Furniture and home furnishing stores

When I run the SAS code generate reports at the 3-digit level for the four census regions, the resulting output table look like as follows:

TABLE 4 A partial snippet of SAS output for all retail industries employment and establishment count for the four census regions at NAICS 3-digit level retail industries

odamid-usw2.oda.sas.com/SASStudio/sasexec/submissions/ce11306f-92f1-4e48-b729-e0df7854f2e2/results

Obs	YEAR	NAICS3	CENREGIONS4	emp	est	_BREAK_	NAICS3DGT_DESCRIPTION
1	2001	441	R1NORTHEAST	297703	20278		Motor vehicle and auto parts dealers
2	2001	441	R2MIDWEST	440475	30735		Motor vehicle and auto parts dealers
3	2001	441	R3SOUTH	692592	49595		Motor vehicle and auto parts dealers
4	2001	441	R4WEST	419448	25120		Motor vehicle and auto parts dealers
5	2001	442	R1NORTHEAST	105479	12186		Furniture and home furnishing stores
6	2001	442	R2MIDWEST	128314	14199		Furniture and home furnishing stores
7	2001	442	R3SOUTH	207239	24782		Furniture and home furnishing stores
8	2001	442	R4WEST	126286	14599		Furniture and home furnishing stores
9	2001	443	R1NORTHEAST	73526	8967		Electronics and appliance stores
10	2001	443	R2MIDWEST	103885	11077		Electronics and appliance stores
11	2001	443	R3SOUTH	138478	16510		Electronics and appliance stores
12	2001	443	R4WEST	109847	11193		Electronics and appliance stores
13	2001	444	R1NORTHEAST	209910	16665		Building materials, garden equipment and supplies dealers
14	2001	444	R2MIDWEST	333117	26341		Building materials, garden equipment and supplies dealers
15	2001	444	R3SOUTH	444818	33181		Building materials, garden equipment and supplies dealers
16	2001	444	R4WEST	261281	17948		Building materials, garden equipment and supplies dealers
17	2001	445	R1NORTHEAST	625936	39191		Food and beverages stores
18	2001	445	R2MIDWEST	663672	33494		Food and beverages stores
19	2001	445	R3SOUTH	1063520	52828		Food and beverages stores

Please note that by the U.S. census region's definition, U.S. census region 1 (Northeast region) is composed of the following two census divisions: New England division and Middle Atlantic division. As shown in Table 3, for the year 2001, the total retail employment for the retail industry NAICS 441 for the Northeast region and Middle Atlantic regions were 91,301 and 206,402, respectively. Therefore, total retail employment for 2001 for the census region 1 (Northeast) was found to be 297,703 which has been accurately calculated by the SAS ODA (see Table 4).

TABLE 5 A partial snippet of SAS output for U.S. total retail employment across twelve subsectors (NAICS 3-digit level)

odamid-usw2.oda.sas.com/SASStudio/sasexec/submissions/293b6fc3-f28b-40fa-8f86-e5feb418f4d3/results

Obs	YEAR	NAICS3	emp	est	_BREAK_	NAICS3DGT_DESCRIPTION	NA
1	2001	441	1850218	125728		Motor vehicle and auto parts dealers	
2	2001	442	567318	65766		Furniture and home furnishing stores	
3	2001	443	425736	47747		Electronics and appliance stores	
4	2001	444	1249126	94135		Building materials, garden equipment and supplies dealers	
5	2001	445	2963801	156261		Food and beverages stores	
6	2001	446	958072	81898		Health and personal care stores	
7	2001	447	927284	118892		Gasoline stations and convenience stores	
8	2001	448	1392626	151668		Clothing and clothing accessories stores	
9	2001	451	622261	64453		Sporting goods, hobby, book and music stores	
10	2001	452	2525974	41169		General merchandise stores	
11	2001	453	841594	129311		Miscellaneous retail stores	
12	2001	454	566279	42922		Non-store retailers	
13	2002	441	1890916	126644		Motor vehicle and auto parts dealers	
14	2002	442	551567	66360		Furniture and home furnishing stores	
15	2002	443	418725	49600		Electronics and appliance stores	
16	2002	444	1270736	94109		Building materials, garden equipment and supplies dealers	
17	2002	445	2883997	155677		Food and beverages stores	
18	2002	446	988347	82574		Health and personal care stores	
19	2002	447	895983	117100		Gasoline stations and convenience stores	
20	2002	448	1408948	151895		Clothing and clothing accessories stores	
21	2002	451	617726	65933		Sporting goods, hobby, book and music stores	
22	2002	452	2546094	41069		General merchandise stores	

TABLE 6 A partial snippet of SAS output for U.S. County level employment and establishment counts for retail industries (at the 3-digit NAICS codes level)

odamid-usw2.oda.sas.com/SASStudio/sasexec/submissions/299e8136-fcc0-4598-9fb7-147a378cd11c/results

YEAR	fipscty	NAICS3CHAR	emp	est
2021	001	441	43991	2557
		442	11234	1128
		443	5758	492
		444	32275	1737
		445	86446	4012
		446	23538	2054
		447	21261	2251
		448	33706	3056
		451	11181	1132
		452	56670	988
		453	17872	2573
		454	23789	2006
	003	441	58438	3272
		442	16033	1359
		443	7833	707
		444	39235	1876
		445	105370	4724
		446	33337	2984
		447	27396	2857
		448	57655	4839
		451	14364	1354

FIGURE 1A **STATE LEVEL RETAIL EMPLOYMENT IN 2001**

damid-usw2.oda.sas.com/SASStudio/sasexec/submissions/54453ea0-4fbf-4fde-9253-20be5771572a/results

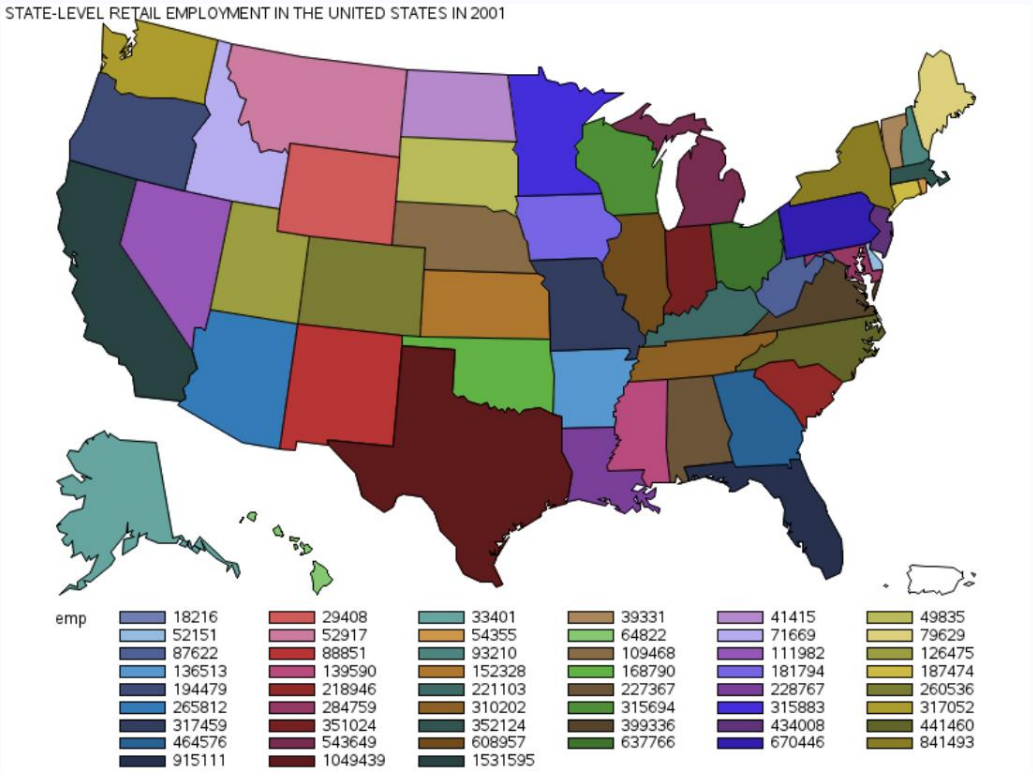


FIGURE 2 **STATE LEVEL RETAIL EMPLOYMENT IN 2021**

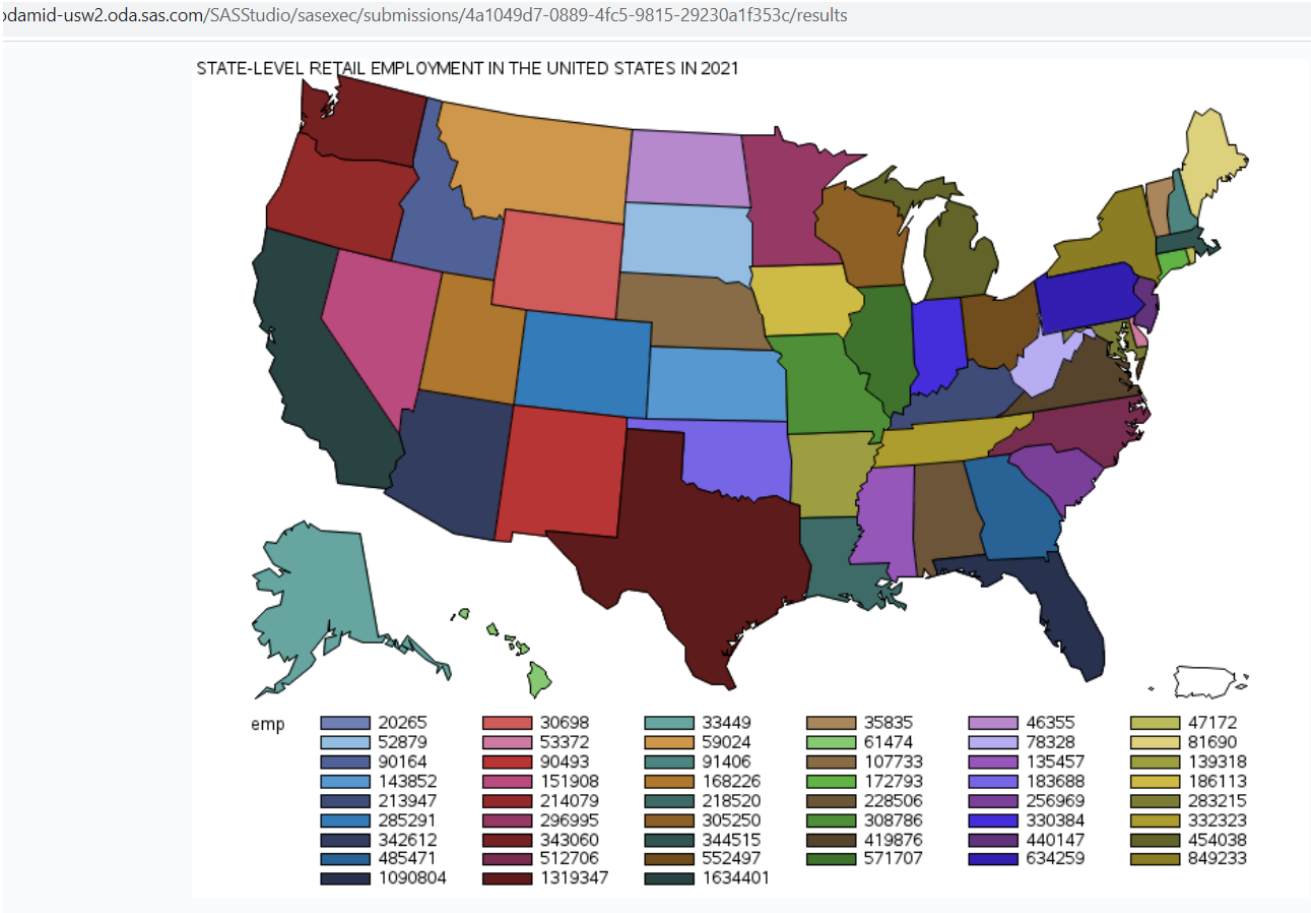


FIGURE 3 **MAP SHOWING STATE LEVEL PERCENTAGE CHANGES IN RETAIL SECTOR EMPLOYMENT BETWEEN 2001 AND 2021**

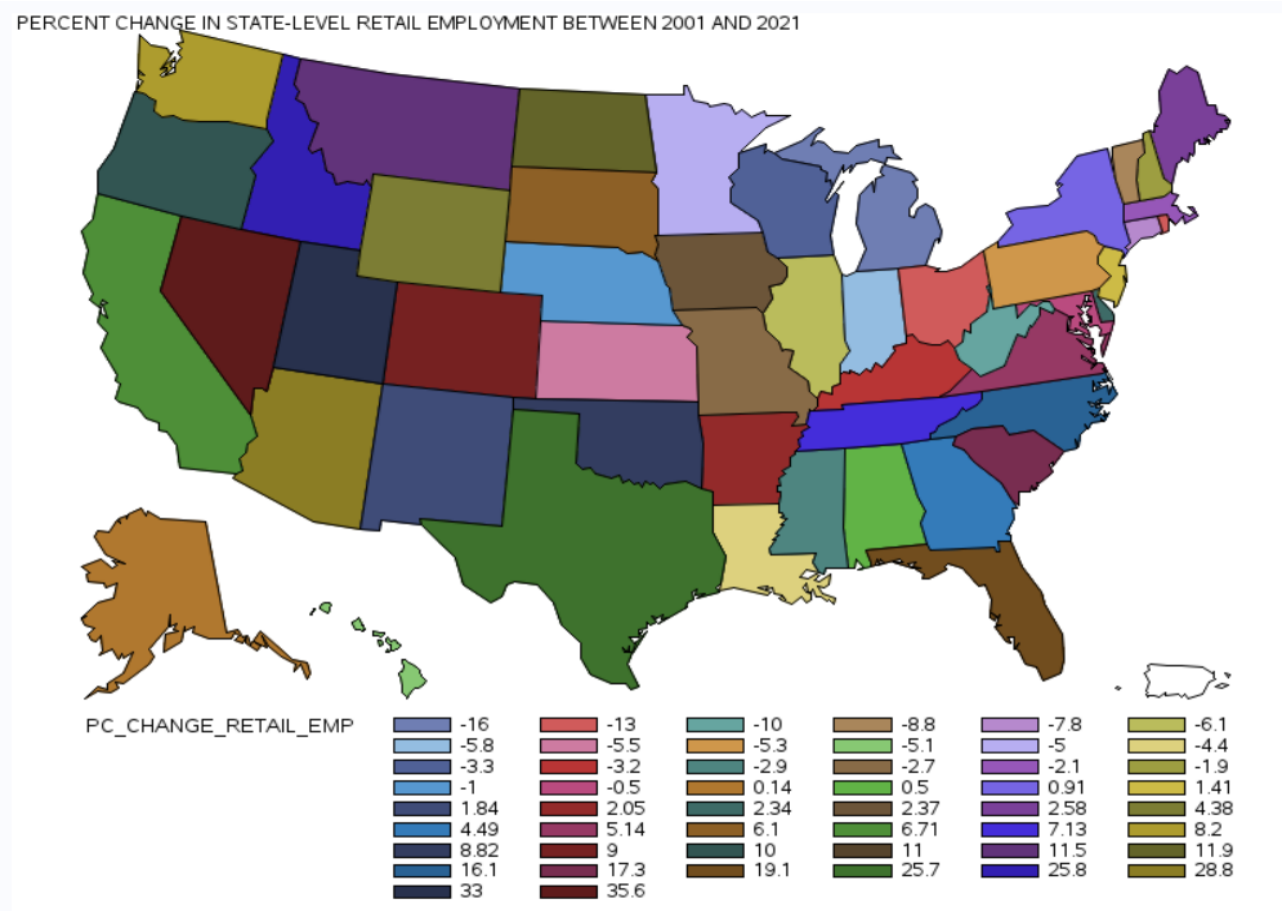


FIGURE 4 **MAP SHOWING PERCENTAGE CHANGES IN RETAIL SECTOR EMPLOYMENT BETWEEN 2001 AND 2021 ACROSS INDIVIDUAL STATES GROUPED BY MAGNITUDE OF CHANGE**

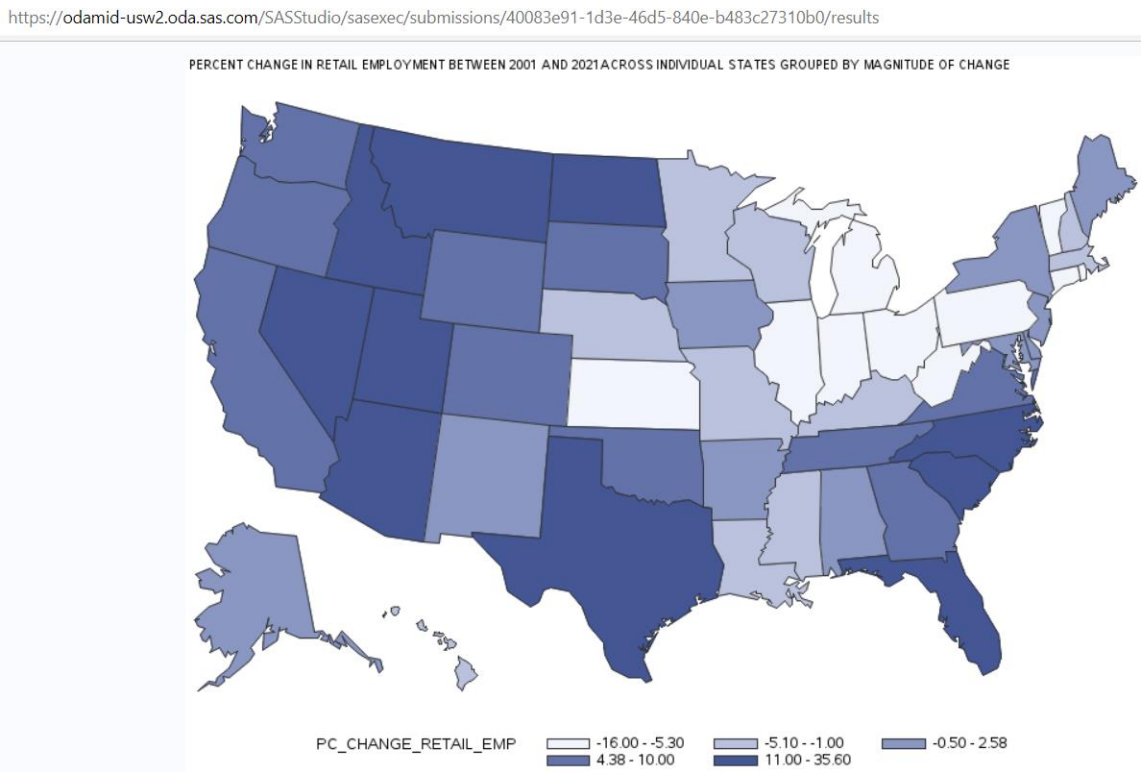
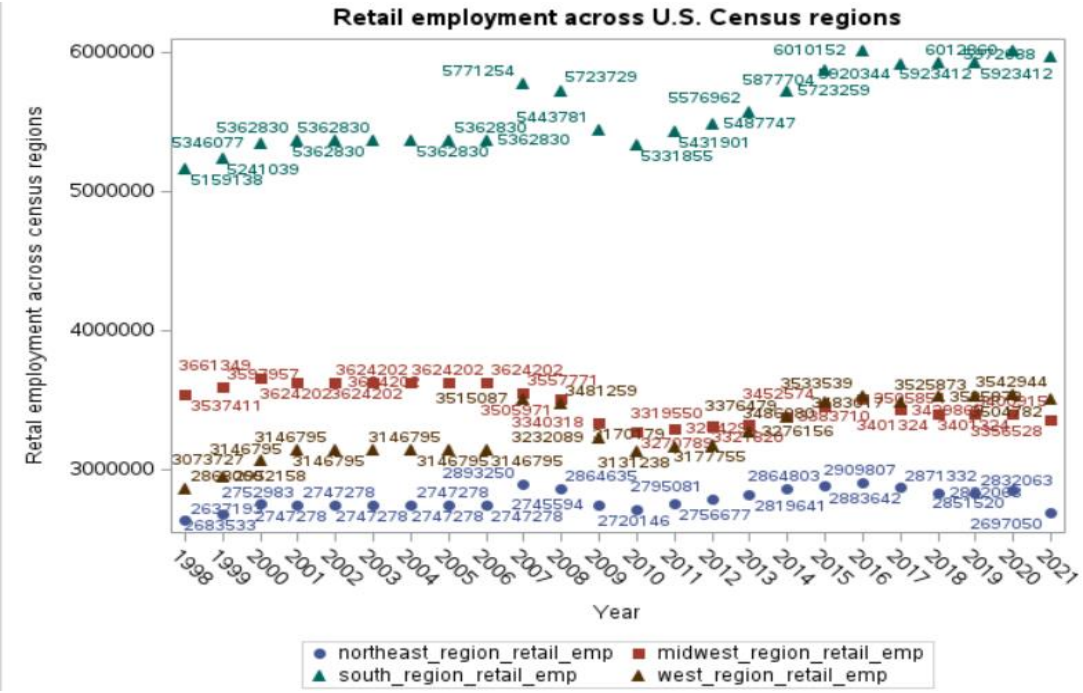


FIGURE 5 **RETAIL EMPLOYMENT ACROSS CENSUS REGIONS, 2001-2021**



Per the color-coded map (Figure 4) of the states grouped into categories based on the magnitude of change in retail employment and the jurisdictional definition of census divisions and regions (appendix Table A1) along with the state-level employment summary table (Appendix Table A2) one can delineate several trends. One but all states (except for Hawaii) and all the census divisions included in the West region experienced growth in total employment in the retail industry. Eleven states - one out of twelve states in this region experienced growth in retail sector employment. In this region the top three states that experienced the highest percentile growth in employment between 2001 and 2021 are: Nevada (35.65%), Utah (33.01%) and Arizona (28.89%).

In terms of employment growth, the South region also did well. Twelve states out of seventeen states of the region experienced positive employment growth in the retail sector during the indicated period. In terms of employment gain, the top three states in the region are Texas (25.72%), Florida (19.20%), and South Carolina (17.37%). West Virginia is the state that saw the most percentile decline (-10.61%) in retail employment in the region

The Midwest region contains the highest number of states that experienced a decline in the total retail sector employment. Nine out of twelve states saw a decrease in retail employment in 2021 relative to 2001. The three states in this region that experienced an increase in retail sector employment over the indicated period are North Dakota (11.93%), South Dakota (6.11%), and Iowa (2.38%). In this region Michigan experienced the largest percentile decline in retail sector employment between 2001 and 2021 (-16.48%).

In the Northeast region, seventy five percent of the states (six states out of nine states) experienced a decline in retail sector employment. The three states that had a gain in retail sector employment for the mentioned period are: Maine (2.59%), New Jersey (1.41%), and New York (0.92%). Rhodes Island lost the most in terms of retail employment in percentage for the indicated period (-13.21%). The retail employment trends from 2001 through 2021 across the U.S. census region are shown in scatterplot presented in Figure 5.

CONCLUSIONS

This paper discusses and demonstrates how to manipulate industrial employment data using SAS. SAS provides a rich set of analytical functionalities (such as regression analysis, correlation studies etc.), customized report generation features, and map creation and other insightful data visualization tools (Tavakoli, Jones and Raynor, 2022; Okerson, 2010; Zdeb, 2004) which can be explored using SAS ODA, as well as SAS Viya AI and analytics platform. Data visualization helps us see the granulated trends in data that are useful for follow up research and to add value to industrial insights and policy perspectives. While overall retail employment growth is observed, the national level (Table A1 in the appendix section), we have also visualized in the maps (specially, in Figure 3 and Figure 4) that the employment growth in the retail sector employment is skewed across the census regions. While the South region and the West region experienced the overall growth in the retail sector employment (11.34% and 11.37% respectively), the Midwest region and Northeast region saw a decline (-7.41% and -2.07% respectively) in the retail sector employment.

REFERENCES

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H., & Price, B. (2016). Import competition and the great US employment sag of the 2000s. Journal of Labor Economics, 34(S1), S141-S198.

Hänninen, M., Kwan, S. K., & Mitronen, L. (2021). From the store to omnichannel retail: looking back over three decades of research. The International Review of Retail, Distribution and Consumer Research, 31(1), 1- 35.

Khan, A. (2023a). Influence of Residential Internet Access and Patented Innovation on U.S. Non-store Retail Industries' Employment, Advances in Social Sciences Research Journal, 10(80):107-124. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4540393).

Khan, A. (2023b). Employment Trends and Relative Influence of U.S. Retail Subsector Employments on GDP. Archives of Business Research, 11(7). 189-205. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4552655)

Liu, H., Lobschat, L., & Verhoef, P. C. (2018). Multichannel retailing: A review and research agenda. Foundations and Trends in Marketing, 12(1), 1-79.

Okerson, B. (2010). SAS® Maps as Tools to Display and Clarify Healthcare Outcomes, SAS Global Forum 2010, Conference proceedings. (https://support.sas.com/resources/papers/proceedings10/211-2010.pdf)

Pierce, J. R., & Schott, P. K. (2016). The surprisingly swift decline of US manufacturing employment. American Economic Review, 106(7), 1632-1662.

Rider, M., & Khan, A. M. (2010). The Impact of Globalization on Agglomeration: The Case of US Manufacturing Employment from 1988 to 2003. Andrew Young School of Policy Studies Research Paper Series, (10-07). (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1944995)

U.S. Bureau of Labor Statistics (2023). Employment By Industry. Retrieved September 10, 2023, from: Weblink: https://www.bls.gov/charts/employment-situation/employment-levels-by-industry.htm#:~:text=Thousands%20Employment%20levels%20by%20industry%2C%20seasonally%20adjusted%20Click,150%2C000%20160%2C000%20Hover%20over%20chart%20to%20view%20data.

U.S. Bureau of Labor Statistics. Industries at a glance. Webpage accessed on 12th August 2023. (https://www.bls.gov/iag/).

U.S. Census Bureau. County Business Pattern Data.

Zdeb, M., & Rensselaer, N. (2004). The Basics of Map Creation with SAS/GRAPH. *SUGI 29*, 251-29. (https://support.sas.com/resources/papers/proceedings/proceedings/sugi29/251-29.pdf)

TRADEMARK NOTICE

Contact Information

Abdullah Khan
Associate Professor of economics
School of Business
Claflin University
400 Magnolia St.,
Orangeburg, SC 29115
Phone: 803-535-5479 (office)
Email: akhan@claflin.edu

(APPENDICES) **Table A1 Census Divisions, Census Regions and Sates Therein**

Census Regions	**Census Divisions therein**	**States therein**
Census region 1 (Northeast Region) .	*Division 1 (New England division)*	Division 1 (New England) includes following U.S. States: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont.
	Division 2 (Middle Atlantic division)	Division 2 (Middle Atlantic) includes the following states: New Jersey, New York, and Pennsylvania.
Census region 2 (Midwest Region)	*Division 3 (East North Central division)*	The Division 3 includes the following states: Indiana, Illinois, Michigan, Ohio, and Wisconsin.
	Division 4 (West North Central division)	The Division 4 (West North Central) includes following states: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota
Census region 3 (South Region)	*Division 5 (South Atlantic division)*	The division 5 (South Atlantic) includes District of Columbia and the following eight states Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia.
	Division 6 (East South-Central division)	The division 6 (East South Central) includes following states: Alabama, Kentucky, Mississippi, and Tennessee.
	Division 7 (West South-Central division)	Division 7 (West South Central) includes the following states: Arkansas, Louisiana, Oklahoma, and Texas.
Census region 4 (West Region)	*Division 8 (Mountain division)*	The division 8 (Mountain) includes following states: Arizona, Colorado, Idaho, New Mexico, Mountain, Utah, Nevada, and Wyoming.
	Division 9 (Pacific division)	The division 9 (Pacific) is comprised of following states: Alaska, California, Hawaii, Oregon, and Washington.

Note: U.S. Census Bureau's definition of Census Regions and Census divisions can also be accessed from the Bureau's webpage[4].

[4] https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

(APPENDICES) TABLE A2 Retail Employment by States (2001-2021)

State Name	2001	2021	% change between 1988-
Alabama	227,367	228,506	0.50%
Alaska	33,401	33,449	0.14%
Arizona	265,812	342,612	28.89%
Arkansas	136,513	139,318	2.05%
California	1,531,595	1,634,401	6.71%
Colorado	260,536	285,291	9.50%
Connecticut	187,474	172,793	-7.83%
Delaware	52,151	53,372	2.34%
District of Columbia	18,216	20,265	11.25%
Florida	915,111	1,090,804	19.20%
Georgia	464,576	485,471	4.50%
Hawaii	64,822	61,474	-5.16%
Idaho	71,669	90,164	25.81%
Illinois	608,957	571,707	-6.12%
Indiana	351,024	330,384	-5.88%
Iowa	181,794	186,113	2.38%
Kansas	152,328	143,852	-5.56%
Kentucky	221,103	213,947	-3.24%
Louisiana	228,767	218,520	-4.48%
Maine	79,629	81,690	2.59%
Maryland	284,759	283,215	-0.54%
Massachusetts	352,124	344,515	-2.16%
Michigan	543,649	454,038	-16.48%
Minnesota	315,883	296,995	-5.98%
Mississippi	139,590	135,457	-2.96%
Missouri	317,459	308,786	-2.73%
Montana	52,917	59,024	11.54%
Nebraska	109,468	107,733	-1.58%
Nevada	111,982	151,908	35.65%
New Hampshire	93,210	91,406	-1.94%
New Jersey	434,008	440,147	1.41%
New Mexico	88,851	90,493	1.85%
New York	841,493	849,233	0.92%
North Carolina	441,460	512,706	16.14%
North Dakota	41,415	46,355	11.93%
Ohio	637,766	552,497	-13.37%
Oklahoma	168,790	183,688	8.83%
Oregon	194,479	214,079	10.08%
Pennsylvania	670,446	634,259	-5.40%
Rhode Island	54,355	47,172	-13.21%
South Carolina	218,946	256,969	17.37%
South Dakota	49,835	52,879	6.11%
Tennessee	310,202	332,323	7.13%
Texas	1,049,439	1,319,347	25.72%
Utah	126,475	168,226	33.01%
Vermont	39,331	35,835	-8.89%
Virginia	399,336	419,876	5.14%
Washington	317,052	343,060	8.20%
West Virginia	87,622	78,328	-10.61%
Wisconsin	315,694	305,250	-3.31%
U.S. Total	14,890,289	15,530,630	4.30%