Public Employee Jobs Analysis

A Processing Manual for the Assignment of Standard Occupation Classification (SOC) Codes to Public Employee Data

A joint project of the
Indiana Department of Workforce Development &
Indiana Business Research Center at the
Indiana University Kelley School of Business

The Indiana Department of Workforce Development (DWD) in collaboration with the Indiana Business Research Center (IBRC) have developed these instructions as a guide to analyze publicly available government employee data. This methodology overcomes the greatest challenges with this data – inconsistent job titles and descriptions. In this document, we provide step-by-step instructions on assigning Standard Occupational Classification (SOC) codes to raw employment data. We also provide background on job descriptions, SOC codes, recommendations for data usage, as well as a breakdown of Indiana public employee data.

Authors: Kellie McGiverin-Bohan

Contributors: Timothy Slaper, Sunjung Yoon, Thea Evans, Carol Rogers

Acknowledgments

This manual was prepared as a part of an ongoing collaboration with the Indiana Department of Workforce Development and Indiana Business Research Center to provide insights on how to best approach public employee data. We are thankful to have this opportunity from the U.S. Employment and Training Administration as part of the Workforce Data Quality Initiative.

Suggested Citation:

McGiverin-Bohan, K. et al. Indiana Business Research Center (2019). *Public Employee Jobs Analysis: A Processing Manual for the Assignment of Standard Occupation Classification (SOC) Codes to Public Employee Data*. Bloomington, IN: Hoosiers by the Numbers, Indiana Business Research Center, the Indiana Department of Workforce Development.

Contents

Executive Summary 1
Background 2
Public Employee Coding Steps 6
Data Analysis Examples 20
References 29

Tables

Table 1: SOC Major groups 4
Table 2: Example of hierarchical occupation code system (preschool through high school teachers) 5
Table 3: Formatted copy of raw public employee files 8
Table 4: Examples of common job title abbreviations 9
Table 5: Examples of common job title misspellings 9
Table 6: Examples of Uncodable Records 19
Table 7: Broad distribution of jobs (%) 20
Table 8: Breakdown of education employment (25-000) 22
Table 9: Percent jobs in each type government unit 22
Table 10: School Unit occupation breakdown 23
Table 11: Percent public jobs by county 24
Table 12: Marion County occupation breakdown 25
Table 13: Percentage of Occupations by MSA (split across two pages) 27

Figures

Figure 1: Public employee occupation analysis workflow 6
Figure 2: Employee file ready for coding 10
Figure 3: Sample of IBRC Crosswalk 11
Figure 4: Applying the SOC crosswalk 12
Figure 5: Example SQL Query - Applying SOC crosswalk 12
Figure 6: Teaching assistant record examples 14
Figure 7: Batch Code Example Step 1 & 2 15
Figure 8: Results from Batch Filters 15
Figure 9: Example O*Net Return for "Teacher Assistant" Query 16
Figure 10: Example of O*Net Job Description 17
Figure 11: Related Occupations List 18
Figure 12: Example of Inapplicable Search Results 18
Figure 13: Indiana MSAs 26

Executive Summary

The Indiana Public Employee Analysis (2019) explores both state and local governmental jobs. To conduct this analysis, the Indiana Department of Workforce Development (IN DWD) and the Indiana Business Research Center (IBRC) designed a crosswalk database. This crosswalk categorizes a wide variety of job titles into Department of Labor Standard Occupational Codes (SOC) to facilitate analysis. This document provides instructions on how to use Indiana's crosswalk, as well as develop crosswalk databases in other states.

Applying the crosswalk to Indiana, we found information such as:

- Cities account for only 8 percent of total government jobs, compared to public schools which account for 42 percent.
- Office and administrative support jobs comprise 13 percent of all government employment. At the county level, they make up 23 percent.
- College/university employment accounts for 20 percent of all public education jobs.
- In schools, bus drivers represent 7 percent of occupations.
- Marion has the most government jobs compared to other counties, comprising 8 percent.

Policy makers, government officials, and researchers can use these facts to:

- **Identify areas with the most government employees.** The rate of government employment can affect local economies by providing extra jobs, but can also potentially crowd out other industries.
- **Examine economic conditions and development.** Regions with few government employees can communicate with underdeveloped services to improve conditions.
- **Be more transparent.** Job data helps inform citizens about budget allocations.
- **Have policy effects.** Occupation data can serve discussions about the quality of education, incarceration, environmental protection, public safety, city services, etc.
- **Recruit talents.** Compare the average salaries for positions to other regions and offer citizens right opportunities.

The first half of this document explains the methodology underlying the Employment crosswalk. The second half demonstrates some ways to use occupational data.

Background

Different government units often use a variety of job titles for the same occupation which impedes certain aggregations on government employments. For example, teacher assistants may have different job titles such as school paraprofessional, teacher aide, special education teacher assistant, para-educator, and so on.

To combat these challenges, the Indiana Business Research Center (IBRC) and the Indiana Department of Workforce Development (IN DWD) created a crosswalk between public employee occupation records and the Standard Occupation Classification (SOC) codes. Managers and policy makers can use this system to organize their public employee records into an easy-to-manage database for reviews and analysis.

Job Titles

Organizations and departments assign job titles in a variety of ways. In a researcher's ideal world, job titles would be based on: the employees' specialized responsibilities and duties; the work design and workflow; job required equipment and materials; skills; paygrades; benefits and a variety of other characteristics. The standard acronym for a worker's job-related characteristic is KSAP – knowledge, skills, abilities, and personality. But organizations and departments face many different challenges, and they often adjust job titles to fit the needs of the unit (Strang & Baron 1990). Titles also change as job roles evolve. Therefore, not all job titles match precisely with SOC codes.

Public Employee Records

Public employee records have been publicly available in Indiana since 1943 (100R data under the IC 5-11-13-1 and IC 5-14-3.8-7). Beginning in 2013, government employers must submit public employee records to the Indiana State Board of Accounts. These data sets include employee names, business addresses, job titles, compensation, county, and type of government unit. When working with the data, there are a few things to keep in mind:

Salary does not mean a complete annual salary. It is the compensation for one person, at one job, for that one year. The data set includes all people that were employed at the end of the year, no matter how many days they worked. For example, an employee hired on Dec 1 has an "annual" salary equivalent to what they are paid in December. In another example, if someone quits in April, their total annual salary only reflects earnings from January to April.

The total number of employees does not equal the total number of jobs. A unit reports all the employees that worked that year. This includes employees who quit a job and new employees taking the job. If Pam quits her job, and Joe is hired as her replacement, the office reports both employees, even though they shared the same job that year. This also applies to employees who hold more than one job. For example, Anne is both a teacher and a debate coach. She would appear twice in the data set: one record for her teaching job and another time for her coaching job.

The definition of compensation varies across units. Organizations/units report compensation, but the definition of compensation varies. For example, a job may come with a meal plan or retirement contributions in addition to salary in one unit but not another.

The titles for the same type of job vary depending on the organization/department. Each unit uploads its own data. Not all units use the same job titles for a particular occupation. For example, one school has "teacher assistants," while other schools use titles like "assistant instructor" or "teaching assistant." The Crosswalk is specifically designed to overcome this challenge.

While the data set is available to the public, still double check with your data clearinghouse about confidentiality. It's better to check and be sure than unintentionally violate a person's confidentiality. The Indiana 100R data set – the one used for the crosswalk – is public domain. In Indiana, there are several websites that host searchable databases mostly used for information on highest paid employees. The IBRC redacts the name data, so that it is not a distraction.

For more information about Indiana 100R data, see documentation from Indiana Gateway.

Standard Occupational Classifications

Standard Occupational Classification codes (SOC) are a federal standard that can be used to classify employees into a standardized list of occupations. The U.S. Bureau of Labor Statistics (BLS) defines occupations based on job duties, responsibilities, education, skills, and others. There are 840 distinct classifications are grouped into a hierarchical/tiered system—23 major groups, 97 minor groups, and 461 broad occupations. Table 1 lists all major groups, and Table 2 provides an example of the hierarchical classification system.

SOC codes are a six-digit coding system that covers public, private, and military occupations. Applying SOC codes is useful for collecting, calculating, and sharing employee data. For example, the IBRC uses the data to compare the number of employees by occupation in various counties. The data can give insights into employment and workforce trends.

Table 1: SOC Major groups

11-0000	**Management Occupations**
13-0000	Business and Financial Operations Occupations
15-0000	Computer and Mathematical Occupations
17-0000	Architecture and Engineering Occupations
19-0000	Life, Physical, and Social Science Occupations
21-0000	Community and Social Service Occupations
23-0000	Legal Occupations
25-0000	Education, Training, and Library Occupations
27-0000	Arts, Design, Entertainment, Sports, and Media Occupations
29-0000	Healthcare Practitioners and Technical Occupations
31-0000	Healthcare Support Occupations
33-0000	Protective Service Occupations
35-0000	Food Preparation and Serving Related Occupations
37-0000	Building and Grounds Cleaning and Maintenance Occupations
39-0000	Personal Care and Service Occupations
41-0000	Sales and Related Occupations
43-0000	Office and Administrative Support Occupations
45-0000	Farming, Fishing, and Forestry Occupations
47-0000	Construction and Extraction Occupations
49-0000	Installation, Maintenance, and Repair Occupations
51-0000	Production Occupations
53-0000	Transportation and Material Moving Occupations
55-0000	Military Specific Occupations

Table 2: Example of hierarchical occupation code system (preschool through high school teachers)

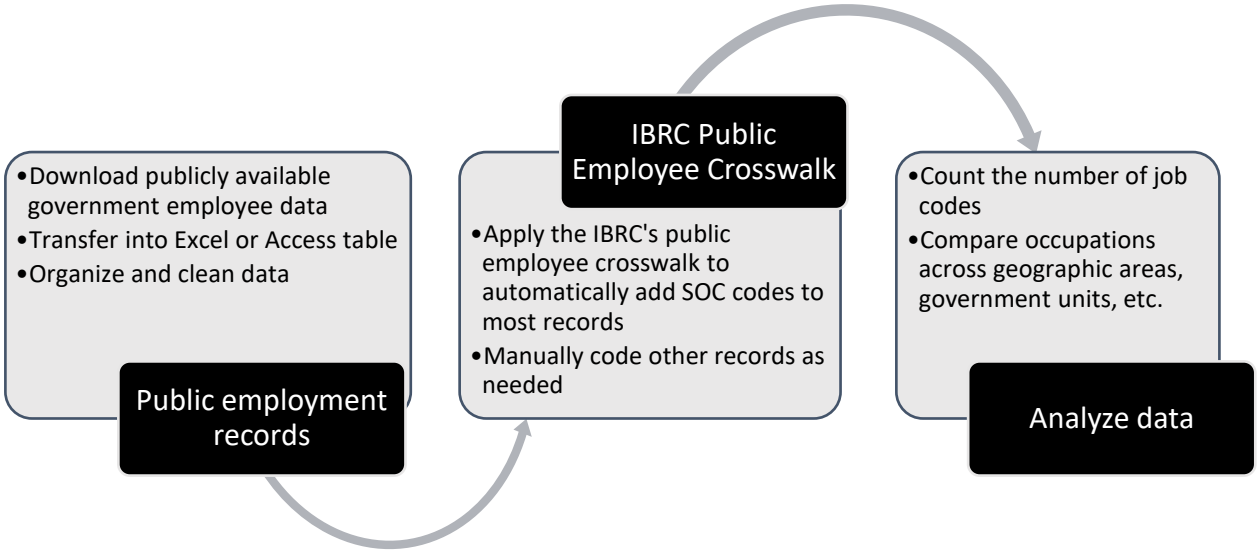
25-2000			**Preschool, Primary, Secondary, and Special Education School Teachers**
	25-2010		Preschool and Kindergarten Teachers
		25-2011	Preschool Teachers, Except Special Education
		25-2012	Kindergarten Teachers, Except Special Education
	25-2020		Elementary and Middle School Teachers
		25-2021	Elementary School Teachers, Except Special Education
		25-2022	Middle School Teachers, Except Special and Career/Technical Education
		25-2023	Career/Technical Education Teachers, Middle School
	25-2030		Secondary School Teachers
		25-2031	Education
		25-2032	Career/Technical Education Teachers, Secondary School
	25-2050		Special Education Teachers
		25-2051	Special Education Teachers, Preschool
		25-2052	Special Education Teachers, Kindergarten and Elementary School
		25-2053	Special Education Teachers, Middle School
		25-2054	Special Education Teachers, Secondary School
		25-2059	Special Education Teachers, All Other

Public Employee Coding Steps

The IBRC's approach to analyzing public employee data has three basic steps (Figure 1):

- Acquiring publicly available government employment records (as described on page 2).
- Preparing and coding the data using the IBRC SOC crosswalk and other sources.
- Conducting descriptive and other analyses of governmental employee records.

Figure 1: Public employee occupation analysis workflow



This section will focus on the steps for organizing and coding records. Possible analysis options are explored starting on page 20.

Raw file modification

Choose the program you want to use to modify the publicly available data file.

Raw files come in a variety of formats (e.g., text, csv, xml) that are impractical for analysis and need to be converted into accounting or database formats. There are tradeoffs in using MS Access, Excel, or other software programs.

- MS Access: The IBRC chose MS Access to proceed with coding. All examples shown in this document are based on MS Access. The program handles filtering, record matching, and

other useful coding activities, such as producing specialized tables. Also, when working off a server, multiple users can work in the database at the same time without interfering with each other.

- Excel: Due to the amount of data available, a user may find Excel rather slow in creating tables.
- Other programs: For any extensive analysis, a user may want to use other statistical programs, such as RStudio, Stata, or SAS.

Preparations and recommendations for modifying the working file

- Create a duplicate copy of the original file for convenience.
- Create a record ID. This helps with sorting, filtering, reference records, and other coding activities.
- Create copies of columns to convert text to codes (e.g., a copy of Unit_Type with the codes: City=1, County=2, etc.). Normalizing the data using numeric codes speeds up the filtering process.
- Add columns for coders, such as SOC code, coder name, and coding rationale. In the rationale column, you can enter information if job titles are ambiguous or can't be coded.
- We hide the following information: personal names, salary, address (except for city name and ZIP code). Such information does not help with coding and can distract coders (e.g., looking up your coworker's salary).

At the end of these steps, your file will look similar to Table 3.

Table 3: Formatted copy of raw public employee files

ID	unit_name	unit_type_cc	unit_type	cnty_cd	cnty_name	year	first_name	middle_initia	last_name	department	business_addres	k
30	ADAMS COUN	1	COUNTY	1	Adams	2014				CIRCUIT COURT		D
229668	INDIANAPOLI	2	CITY	49	Marion	2014				Dept of Parks & Recreation		I
229888	LAWRENCE C	2	CITY	49	Marion	2014				Fire Department		L
445	DECATUR CIV	2	CITY	1	Adams	2014				P & R		D
375073	WARRICK CO	5	SCHOOL	87	Warrick	2014				WARRICK COUNTY SCHOOL CORP.		E
323606	MISHAWAKA	5	SCHOOL	71	St. Joseph	2014				Administration Center		M
2773	BERNE PUBLI	6	LIBRARY	1	Adams	2014						E
161854	GOOD SAMAR	4	HOSPITAL	42	Knox	2014				Nursery		V
262929	MARION CO S	19	SOIL & WATI	49	Marion	2014				Marion County Soil & Water Conser		I
263338	Indianapolis A	25	AIRPORT AU	49	Marion	2014				Parking Mgmt/Cashiers		I
353586	TIPTON COUN	1	COUNTY	80	Tipton	2014				0801 Health		T
410195	State of India	40	OTHER STAT	93	State	2014				Natural Resources, Department of		I
439350	Indiana State	41	UNIVERSITY	93	State	2014				ATHLETICS-BASEBALL		T
454989	Indiana Unive	41	UNIVERSITY	93	State	2014				DENTAL EDUCATION		C
63371	Inspire Acade	32	CHARTER SC	18	Delaware	2014				INSPIRE ACADEMY		M
149999	CITY OF MADI	15	PORT AUTHO	39	Jefferson	2014				Madison Railroad		M
363039	EVANSVILLE L	29	FLOOD CON	82	Vanderburgh	2014				Levee		E
41627	JEFFERSONVI	23	HOUSING AU	10	Clark	2014				MAINTENANCE		J

Review your file for inconsistencies

All files have inconsistencies in job titles attributable to abbreviations (Table 4) and misspellings (Table 5). Going through each of these can be rather time consuming, but the Crosswalk has been designed to address many of these issues for you. For example, it will assign a 11-3031 code to both "Deputy Treasurer" and "Deputy Tresurer."

Table 4: Examples of common job title abbreviations

PART-TIME COMMUNITY CORRECTNS	OPERATING ENGR
PUBLIC INFO OFCR	CH MECH ENGR
GEN OP MGR/CC	MNTCE SPVR IV
ADMIN ASST III	CHLD SUPPT FILE ROOM SUPERVISR
DEP CORONER/PHYS ASST	ASST PUB DEFENDER
ENVIRONMENTAL HLTH SPCLST I	DEPUTY PROSC.
MNTCE WORKER I	ASST CODE OFFIC
HEALTH SERV NURSING SUPV	SYS ADM
PART-TIME PARKING-LOT-ATTENDNT	OUTSIDE MAINT
P/T Em.Dis. Instl Asst	F/T Dir Food Srvcs

Table 5: Examples of common job title misspellings

Assisstant	Township Assisstance Investigator
FOOR SERVICE SUPRVISOR	Recucling Specialist
Laundry Personn	PROPERTY TAX ADMINSTRATOR
Deputy Tresurer	SYSTEM & TECHNOLOGY COORDINATO
superintendant	Asst Directr
OCCUPTNAL THERAPIST	Cemetery Maintentance
PYSICAL THERAPIST	summer cleark
OCCUPT THERPY ASSIST	F/T Computer Network Techncn

Apply the IBRC crosswalk

The figures below demonstrate how we implemented the IBRC crosswalk to a public employees file using MS Access. Figure 2 is an example of the Indiana public employee records. At a minimum, the table must include: unit_name, department, job_title, SOC_code, SOC_logic, and Coder columns. Including a Coder column helps resolve inconsistencies that naturally arise when classifying inconsistent information. The Coder column is useful in any interrater reliability analysis.

Figure 2: Employee file ready for coding

YourRecords

unit_type	cnty_name	year	unit_name	department	job_title	soccode	soc_title	soclogic	Coder
COUNTY	Adams	2014	ADAMS COUNTY	MAINTENANCE	MAINTENANCE WORK				
SCHOOL	Allen	2014	FORT WAYNE COMMUNITY SCI	JEFFERSON	SCHOOL ASSISTANT				
TOWNSHIP	Allen	2014	CEDAR CREEK TOWNSHIP	township	Board				
HOSPITAL	Bartholomew	2014	COLUMBUS REGIONAL HOSPIT/	7880	286				
CITY	Floyd	2014	NEW ALBANY CIVIL CITY	CITY COUNCIL	COUNCIL MEMBER				
HOSPITAL	Floyd	2014	FLOYD MEMORIAL HOSPITAL/H	7020	136				
SCHOOL	Hamilton	2014	HAMILTON HEIGHTS SCHOOL C	HAMILTON HEI(	TRANSPORATATION DI				
TOWNSHIP	Hendricks	2014	GUILFORD TOWNSHIP		P/T Seasonal Facilities				
SCHOOL	Jasper	2014	RENSSELAER CENTRAL SCHOOL	RENSSELAER CE	SUB TCHRS				
CITY	Lake	2014	CROWN POINT CIVIL CITY	Parks Dept.	Parks Assistant				
TOWN	Lake	2014	MUNSTER CIVIL TOWN	PARKS	GOLF COURSE LABORE				
SCHOOL	Lake	2014	GRIFFITH PUBLIC SCHOOL CORI	READY ELEMEN	Teacher				
SCHOOL	Lake	2014	HAMMOND CITY SCHOOL CORI	0013 Gavit High	No Job Description				
SCHOOL	LaPorte	2014	LAPORTE COMMUNITY SCHOO	LaPorte Commu	Health Assistant				
SCHOOL	Marion	2014	M.S.D. WAYNE TOWNSHIP SCH	Substitute/Part	COACH				
TOWN	Morgan	2014	MORGANTOWN CIVIL TOWN						
SCHOOL	Porter	2014	DUNELAND SCHOOL CORPORA	CHESTERTON H	CUSTODIAN				
SCHOOL	Pulaski	2014	WEST CENTRAL SCHOOL CORP(		Bus Driver				
HOSPITAL	Rush	2014	RUSH MEMORIAL HOSPITAL	RMH	PT ACCESS REP				
SCHOOL	St. Joseph	2014	SOUTH BEND COMMUNITY SCI	Madison Prima	First Grade Teacher				
TOWN	White	2014	MONON CIVIL TOWN		PART TIME				
OTHER STATE UNIT	State	2014	State of Indiana	State Police, Inc	Secretary 4				
OTHER STATE UNIT	State	2014	State of Indiana	State Police, Inc	Motorcarrier Inspecto				
UNIVERSITY	State	2014	Indiana University	PUBLIC HEALTH	Dir, Communication &				
UNIVERSITY	State	2014	Ivy Tech Community College	Liberal Arts	Adjunct Faculty				

The IBRC crosswalk (Figure 3) includes unit_type, department, job_title, and SOC_code columns.

Figure 3: Sample of IBRC Crosswalk

IBRC_SOC_Crosswalk

unit_type	department	job_title	soccode
SCHOOL	Jane Ball Elementary	Teacher	25-2021
UNIVERSITY	WL - Curriculum and Instr	Faculty	25-1000
UNIVERSITY	Information Technology Services	Grad Stipend (100100)	43-9199
UNIVERSITY	GENERAL SURGERY	Faculty/Instructor	25-1071
UNIVERSITY	ACADEMIC AFFAIRS OFFICE	Acad Aff-Contract & Grants	25-9031
UNIVERSITY	PHILOSOPHY	Faculty/Instructor	25-1126
SCHOOL	KV Intermediate School	Teachers	25-2022
SCHOOL	K.V.H.S.	Tech Personnel	15-1151
UNIVERSITY	Academic Skills Advancement	Adjunct Faculty	25-1000
SCHOOL		Elem Adm	11-9032
SCHOOL		Bus Driver	53-3022
OTHER STAT	Department of Child Services	Family Case Manager 2	21-1021
SCHOOL	Sub/Mail	Substitute Teacher	25-3099
UNIVERSITY	WL - Mathematics	Graduate Student	25-1191
LIBRARY		Clerk	43-4121
SCHOOL	Kingsford Heights School	Elementary School Teacher	25-2021
UNIVERSITY	Elementary Education	Associate Professor	25-1081
SCHOOL	01 Administration Center	Substitute Library Aide	25-4031
SCHOOL		RETIRED TEACHER	25-2031
SCHOOL	099 Sub Teachers	Substitute Teacher	25-3099
CITY	PARKS	DAY CAMP COUNSELOR	39-9032
UNIVERSITY	Generalist	Adjunct Faculty	25-1000
SCHOOL		TEACHER	25-2031
LIBRARY	Administration	Director	25-4021
UNIVERSITY	INFORMATICS	Student Temporary Staff	43-9199
COUNTY	108 Prosecutor	Deputy Prosecutor	23-1011
SCHOOL	CLARK-PLEASANT CSC	ELEMENTARY TEACHER	25-2021
UNIVERSITY	Hospitality Administration	Instructor/ PC	25-1194
UNIVERSITY	CHEMISTRY	Student Temporary Staff	43-9199
SCHOOL	PIKE HIGH SCHOOL	TEACHER	25-2031
SCHOOL	Ben Davis High School	SPC ED SEC PARA	25-9041
UNIVERSITY	PSYCHOLOGICAL & BRAIN SCIENCES	Student Temporary Staff	43-9061
SCHOOL	J. EVERETT LIGHT CAREER	ASST DIRECTOR	11-9032
UNIVERSITY	WL - Environmental & Ecol	Admin/Prof	.

Before coding, use the IBRC crosswalk to automatically code job records to save time. This can be done quickly via an update query in MS Access. Figure 4 is an example of the query. Alternatively, modify the SQL code (Figure 5). We match on unit_type, department, and job_title variables. Since departments vary by state, a user may drop department from the match.

Figure 4: Applying the SOC crosswalk

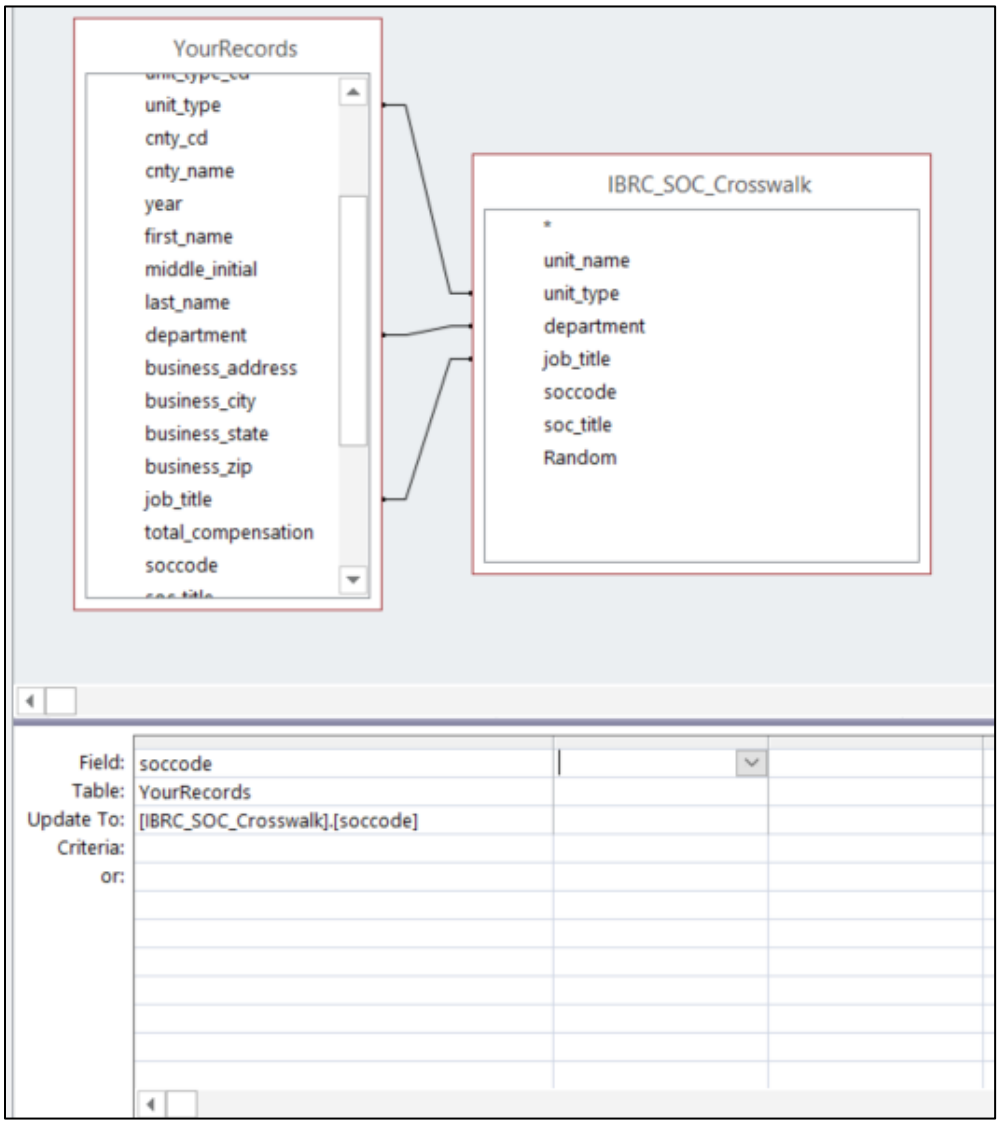


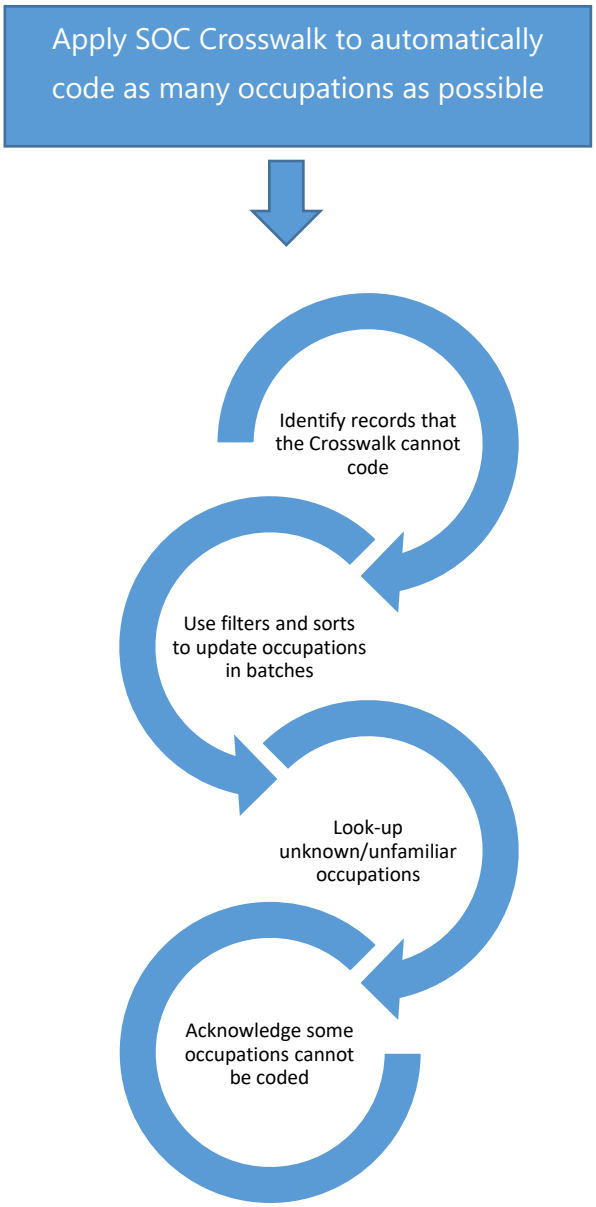
Figure 5: Example SQL Query - Applying SOC crosswalk

```
UPDATE YourRecords
INNER JOIN
IBRC_SOC_Crosswalk ON
(YourRecords.job_title = IBRC_SOC_Crosswalk.job_title) AND
(YourRecords.department = IBRC_SOC_Crosswalk.department) AND
(YourRecords.unit_type = IBRC_SOC_Crosswalk.unit_type)
SET YourRecords.soccode = [IBRC_SOC_Crosswalk].[soccode];
```

Coding Records Example – Teaching Assistants

This section demonstrates examples of the coding process. This set of instructions uses SOC code 25-9041 records that represent 'Teaching Assistant' jobs. One caveat is that while the IBRC SOC crosswalk may reduce coding time, hard-coding may be needed.

This section includes the steps and strategies the IBRC used to hardcode occupations. Using these techniques will speed up hardcoding. We use "teaching assistant" occupations as an example.



1. Review records for patterns

For example, different types of educational institutions, such as charter schools, and universities, employ teaching assistants, but job titles vary significantly. Below are screenshots of records that fall under the SOC code 25-9041, the "Teaching Assistant" occupation. Under the job_title column, note the following:

- Some job titles are spelled out as "Teacher Temp" or "Teaching Assistant," while others include recognizable and common abbreviations such as "Instr Asst" or "IA."
- Some include additional details, like the name of the school (unit_name), department, or type of classroom (unit_type).
- Some have department information that clarifies the job title, while at other times it is missing or obscure (e.g., department named "12").

Figure 6: Teaching assistant record examples

Query1 | employee_compensation_2014_full

unit_type	unit_name	department	job_title	soccode
SCHOOL	GREATER CLARK COUNTY SCHOOL CORPORATIC	CHARLESTOWN MS	TEACHER TEMP	25-9041
SCHOOL	DEKALB COUNTY CENTRAL UNITED SCHOOL COI	J R Watson Elementary Parapros	PARAPRO - JRW - TITLE 1	25-9041
SCHOOL	PIKE COUNTY SCHOOL CORPORATION	PIKE COUNTY SCHOOL CORPORATIOI	AIDE - G/T	25-9041
SCHOOL	FORT WAYNE COMMUNITY SCHOOL CORPORAT	LAKESIDE	SCHOOL ASST SPECIAL EDUC	25-9041
SCHOOL	Northwest Indiana Special Education Cooperat	Merrillville Corp	Ld - Para	25-9041
SCHOOL	M.S.D. WASHINGTON TOWNSHIP SCHOOL CORP	NORTH CENTRAL HIGH SCHOOL	INSTR ASST	25-9041
SCHOOL	MICHIGAN CITY AREA SCHOOL CORPORATION	High School	Paraprofessional	25-9041
SCHOOL	SWITZERLAND COUNTY SCHOOL CORPORATION	SWITZERLAND COUNTY SCHOOL COF	INSTR ASST/ELEM	25-9041
SCHOOL	NORTHWEST ALLEN COUNTY SCHOOL CORPORA	NORTHWEST ALLEN COUNTY SCHOO	SCHOOL ASSISTANT	25-9041
SCHOOL	LEBANON COMMUNITY SCHOOL CORPORATION	LEBANON COMMUNITY SCHOOL COF	INSTRUCTIONAL ASST	25-9041
SCHOOL	GREATER CLARK COUNTY SCHOOL CORPORATIC	NEW WASH ELEM	PARA EDUCATOR	25-9041
SCHOOL	HUNTINGTON COUNTY COMMUNITY SCHOOL C	Substitutes	Paraprofessional - SS - Sub	25-9041
SCHOOL	CARMEL-CLAY SCHOOL CORPORATION	WB Support Staff	Kindergarten IA	25-9041
SCHOOL	SALEM COMMUNITY SCHOOL CORPORATION	SALEM COMMUNITY SCHOOLS	AIDE	25-9041
SCHOOL	BROWNSBURG COMMUNITY SCHOOL CORPORA	EAGLE ELEMENTARY SCHOOL	INSTRUCTIONAL ASST	25-9041
SCHOOL	LINTON-STOCKTON SCHOOL CORPORATION	LINTON STOCTON SCHOOL	INST & SUPP PUP SER	25-9041
SCHOOL	M.S.D. WARREN TOWNSHIP SCHOOL CORPORAT	WARREN CENTRAL HIGH SCHOOL	INSTRUC ASST	25-9041
SCHOOL	Kokomo School Corporation	KOKOMO SCHOOL CORPORATION	PARAPROFESSIONAL/HR	25-9041
SCHOOL	CARMEL-CLAY SCHOOL CORPORATION	TM Support Staff	Kindergarten IA	25-9041
SCHOOL	MICHIGAN CITY AREA SCHOOL CORPORATION	Marsh Elementary	Instructional Aide	25-9041
SCHOOL	Central Indiana Education Service Center	CIESC	PROGRAM ASSISTANT	25-9041
SCHOOL	DELAWARE COMMUNITY SCHOOL CORPORATIO	DELTA MIDDLE SCHOOL	INSTRUCTIONAL ASST	25-9041
SCHOOL	M.S.D. PIKE TOWNSHIP SCHOOL CORPORATION	SNACKS CROSSING ELEM	IA MEDIA	25-9041
SCHOOL	BARTHOLOMEW CONSOLIDATED SCHOOL CORP	Administration	TCH ASSISTANT	25-9041
SCHOOL	WA-NEE COMMUNITY SCHOOL CORPORATION	005 Wakarusa Elementary	Assistant	25-9041
SCHOOL	ZIONSVILLE COMMUNITY SCHOOL CORPORATIO	PLEASANT VIEW ELEMENTARY	Paraprofessional	25-9041
SCHOOL	MONROE COUNTY COMMUNITY SCHOOL CORP	LAKEVIEW	TA PREV	25-9041
SCHOOL	SOUTHWEST DUBOIS COUNTY SCHOOL CORPOR	SOUTHWEST DUBOIS CO SCHOOLS	ASSIST-SP.NEEDS	25-9041
SCHOOL	WARRICK COUNTY SCHOOL CORPORATION	WARRICK COUNTY SCHOOL CORP.	PROGRAM ASSISTANT	25-9041
SCHOOL	JENNINGS COUNTY SCHOOL CORPORATION	JENNINGS COUNTY SCHOOL CORP	TEACHERS ASSISTANT	25-9041
SCHOOL	M.S.D. WARREN TOWNSHIP SCHOOL CORPORAT	LAKESIDE ELEMENTARY SCHOOL	INSTRUC ASST	25-9041

2. Update frequently occurring records in batches

Some records repeat themselves. For instance, a school may have many teaching assistants with the same job title (Figure 6). These records can be found through filtering the data in the table or in a query. Figure 7 and Figure 8 demonstrate how to filter records. Step 1 uses filters to select all records where unit_type is "school," and Step 2 filters for all records where job_title is "teaching assistant."

These steps can be repeated for all standard SOC jobs including: teacher assistant, teaching assistant, instructional assistant, etc. Job titles can be found on the Department of Labor's website.

Figure 7: Batch Code Example Step 1 & 2

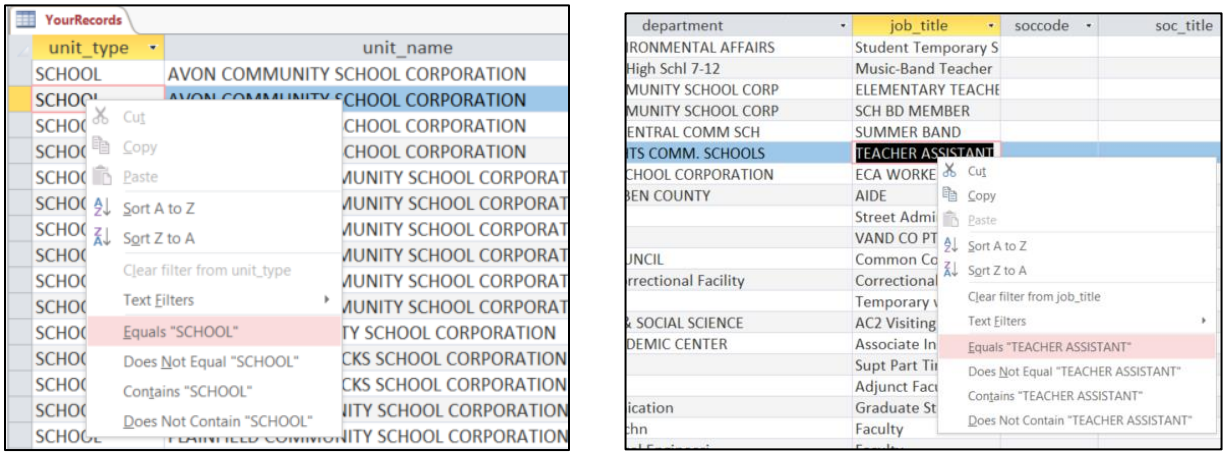
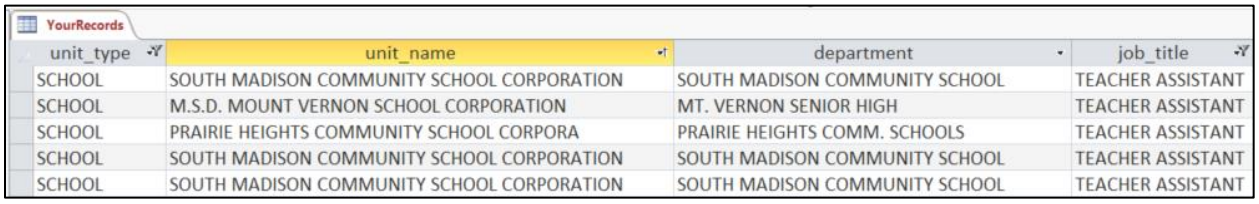


Figure 8: Results from Batch Filters

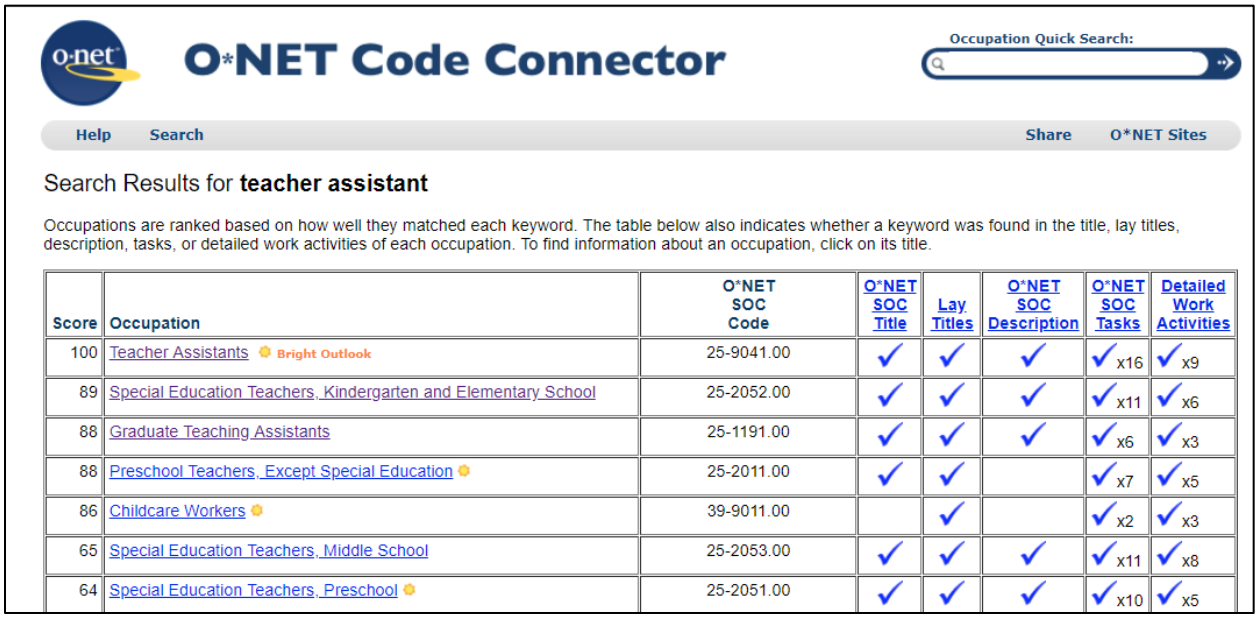


unit_type	unit_name	department	job_title
SCHOOL	SOUTH MADISON COMMUNITY SCHOOL CORPORATION	SOUTH MADISON COMMUNITY SCHOOL	TEACHER ASSISTANT
SCHOOL	M.S.D. MOUNT VERNON SCHOOL CORPORATION	MT. VERNON SENIOR HIGH	TEACHER ASSISTANT
SCHOOL	PRAIRIE HEIGHTS COMMUNITY SCHOOL CORPORA	PRAIRIE HEIGHTS COMM. SCHOOLS	TEACHER ASSISTANT
SCHOOL	SOUTH MADISON COMMUNITY SCHOOL CORPORATION	SOUTH MADISON COMMUNITY SCHOOL	TEACHER ASSISTANT
SCHOOL	SOUTH MADISON COMMUNITY SCHOOL CORPORATION	SOUTH MADISON COMMUNITY SCHOOL	TEACHER ASSISTANT

3. Use O*Net Code Connector

When job descriptions are ambiguous, use O*Net Code Connector to find the appropriate SOC code (Figure 9). A search for "Teacher Assistant" will also return a variety of similar professions like "Graduate Teaching Assistant." If the "Teacher Assistant" works for a University, the 25-1191 code is more appropriate.

Figure 9: Example O*Net Return for "Teacher Assistant" Query



O*NET Code Connector

Occupation Quick Search:

Help Search Share O*NET Sites

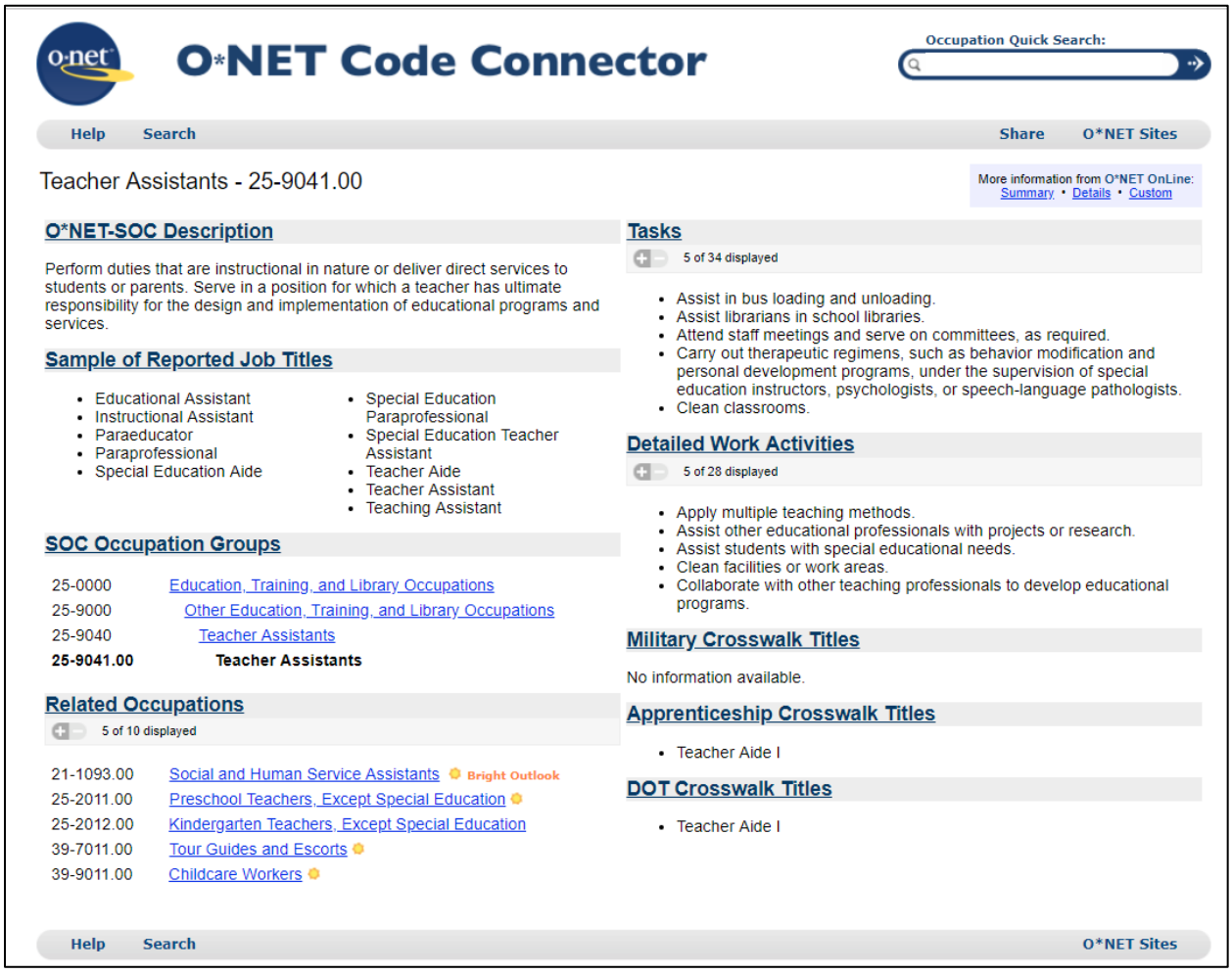
Search Results for **teacher assistant**

Occupations are ranked based on how well they matched each keyword. The table below also indicates whether a keyword was found in the title, lay titles, description, tasks, or detailed work activities of each occupation. To find information about an occupation, click on its title.

Score	Occupation	O*NET SOC Code	O*NET SOC Title	Lay Titles	O*NET SOC Description	O*NET SOC Tasks	Detailed Work Activities
100	Teacher Assistants Bright Outlook	25-9041.00	✓	✓	✓	✓ x16	✓ x9
89	Special Education Teachers, Kindergarten and Elementary School	25-2052.00	✓	✓	✓	✓ x11	✓ x6
88	Graduate Teaching Assistants	25-1191.00	✓	✓	✓	✓ x6	✓ x3
88	Preschool Teachers, Except Special Education	25-2011.00	✓	✓		✓ x7	✓ x5
86	Childcare Workers	39-9011.00		✓		✓ x2	✓ x3
65	Special Education Teachers, Middle School	25-2053.00	✓	✓	✓	✓ x11	✓ x8
64	Special Education Teachers, Preschool	25-2051.00	✓	✓	✓	✓ x10	✓ x5

If you choose "Teacher Assistant," the search returns details on the occupation, including "Sample of Reported Job Titles" that fall under this classification (e.g., teacher aid, para-educator, paraprofessional, etc.) (Figure 10).

Figure 10: Example of O*Net Job Description



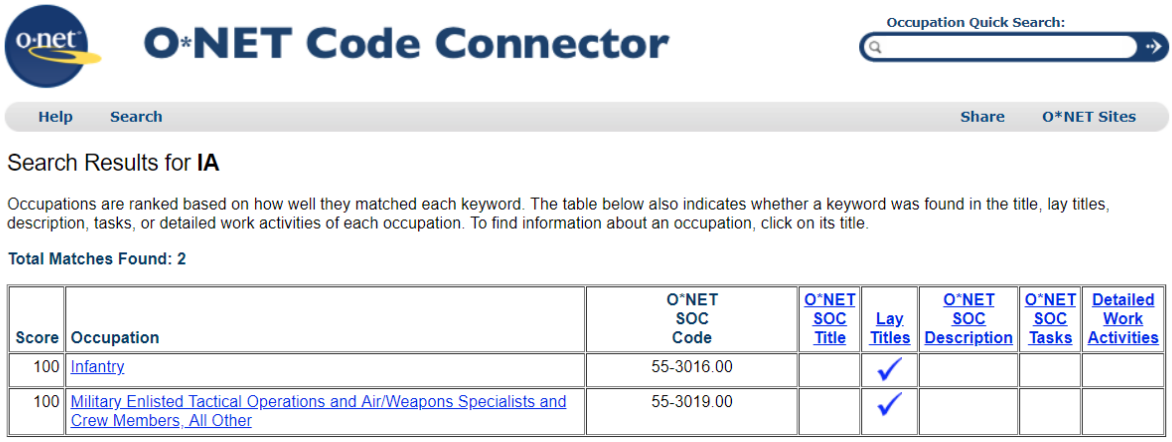
Looking up any of these titles will return the information for teacher assistants. In cases where finding a code is impossible, the "SOC Occupation Groups" can indicate a broader group, which might be more appropriate. For example, the 25-9000 category also includes other "Education, Training, and Library Workers" such as "Instructional Coordinators" (Figure 11).

Figure 11: Related Occupations List

25-4011.00	Archivists
25-4012.00	Curators
25-4013.00	Museum Technicians and Conservators
25-4021.00	Librarians
25-4031.00	Library Technicians
25-9011.00	Audio-Visual and Multimedia Collections Specialists
25-9021.00	Farm and Home Management Advisors Green
25-9031.00	Instructional Coordinators
25-9031.01	Instructional Designers and Technologists
25-9041.00	Teacher Assistants
25-9099.00	Education, Training, and Library Workers, All Other

If the job title is an abbreviation, look-ups can be more difficult. For example, looking up AI returns irrelevant codes (Figure 12). In these cases, department information can be particularly helpful.

Figure 12: Example of Inapplicable Search Results



Even with the most diligent research, some occupation titles are not codable (Table 6). We could not assign codes to 8 percent of records, and these are unevenly distributed. For example, hospitals tend to use numeric occupation titles (e.g., 1456), so 30 percent of those records are not codable. In other cases:

- Occupations such as "sheegoat" and "668 L A II" do not convey useful information.
- Some job lines may be vacant (e.g., retired, sick leave, etc.).

- Sometimes the titles are departments (e.g., food sanitation, information technology, athletic, etc.). In this example, titles like "bookworm" and "broad band executive" have no connection with the SOC naming system.
- Some occupation titles represent budget lines/items, such as retirement, training, firefighter clothing, and others. These records might indicate a constrained budgeting process that requires creative recordkeeping. For example, the university budget process might not have allowances to include awards, so human resource managers include it as a job budget line.

Table 6: Examples of Uncodable Records

Unit Type	Department	Job Title
LIBRARY	Information Technology	Information Technology
TOWN		Town Employee
CITY		Pay Per Run
COUNTY	Engineering	Error Correction
SCHOOL	BCHS CERT	INACTIVE
SCHOOL	1001 Direct Deposit-Aes	Miscellaneous Training
SCHOOL	TERMINATED	RETIRED
SCHOOL	SICK LEAVE	SICK LEAVE
UNIVERSITY	Geology and Physics	FA Bingham Award 2014
UNIVERSITY	WL - Libraries	Student
CITY	Mayors Office	Standard
CITY	unknown	unknown
COUNTY	170 Superior Court 6	668 L A II
COUNTY	GENERAL & UNDISTRIBUTED EXP	GENERAL & &
HOSPITAL		50235
OTHER STATE UNIT	SL	Sheegoat
SCHOOL	DEPARTMENT	DIRECTOR
UNIVERSITY	WL - 4ITIS	Admin/Prof
UNIVERSITY	Special Events & Scheduling Servs	SOA SGA AttnGen 090614
COUNTY	GENERAL & UNDISTRIB/EXPENSES	Garge Forman
UNIVERSITY	DEPARTMENT	DIRECTOR
SCHOOL	Central Office	Bookworm
COUNTY	GEN & UNDISTRIBUTED EXPENSES	
LIBRARY	Information Technology	Information Technology
SCHOOL	Carmel High Cafeteria	HS Media
OTHER STATE UNIT	Family and Social Services	Broad Band Executive
UNIVERSITY	VP & CHIEF FINANCIAL OFFICER	Senior VP, CFO & Treasurer

Data Analysis Examples

SOC coding sets up public employee data for various analyses. Researchers can use these data as a part of larger statistical studies for economic development, government budgeting, and employment trends. The data can also be used for descriptive analysis. This section includes data analysis examples based on Indiana public employee data.

Occupations and Units

Education, training and library occupations account for more than 33 percent of all Indiana public jobs (see Table 7). Considering most state schools are public, this high percentage is plausible. The second largest occupation is Office and Administrative Support, which accounts for around 13 percent of jobs.

Looking at subsectors of 'Education' jobs are informative (Table 8). Out of total education employees in Indiana (n=190,196), 35 percent work in preschools, primary/secondary, and special education occupations. About 22 percent work in a postsecondary (post high school, university, college, vocational) area. Another 23 percent of workers specialize in other education professions, such as adult basic education and self-enrichment teachers (e.g., parks and rec teachers specializing in art, music, etc.). Other education workers—teaching assistants, instructional coordinators, and farm and home management advisers—make-up about 18 percent of all educators. Less than five percent focus on libraries, curating, and archiving materials. Of other education workers (25-9000), over 90 percent (22,180) are teaching assistants.

Table 7: Broad distribution of jobs (%)

Group	Name	% jobs
25-0000	Education, Training, and Library Occupations	34.4%
43-0000	Office and Administrative Support Occupations	13.4%
33-0000	Protective Service Occupations	8.4%
11-0000	Management Occupations	7.2%
53-0000	Transportation and Material Moving Occupations	4.1%
37-0000	Building and Grounds Cleaning and Maintenance Occupations	3.4%
35-0000	Food Preparation and Serving Related Occupations	3.2%
29-0000	Healthcare Practitioners and Technical Occupations	3.0%
21-0000	Community and Social Service Occupations	2.4%
27-0000	Arts, Design, Entertainment, Sports, and Media Occupations	2.1%

13-0000	Business and Financial Operations Occupations	1.7%
39-0000	Personal Care and Service Occupations	1.4%
47-0000	Construction and Extraction Occupations	1.4%
49-0000	Installation, Maintenance, and Repair Occupations	1.3%
19-0000	Life, Physical, and Social Science Occupations	1.2%
15-0000	Computer and Mathematical Occupations	0.9%
23-0000	Legal Occupations	0.9%
51-0000	Production Occupations	0.6%
31-0000	Healthcare Support Occupations	0.6%
17-0000	Architecture and Engineering Occupations	0.4%
41-0000	Sales and Related Occupations	0.2%
45-0000	Farming, Fishing, and Forestry Occupations	0.1%
55-0000	Military Specific Occupations	0.0%
00-0000	Cannot be coded	7.5%

n=552,145

Table 8: Breakdown of education employment (25-000)

Group	Description	Jobs %
25-1000	Postsecondary Teachers	22%
25-2000	Preschool, Primary, Secondary, and Special Education School Teachers	35%
25-3000	Other Teachers and Instructors	23%
25-4000	Librarians, Curators, and Archivists	3%
25-9000	Other Education, Training, and Library Occupations	18%

n=190,196

Type of government unit

Indiana has 30 different types of government units, including schools, port authorities, hospitals, and others (Table 9). Usually, school employees dominate the total number of public jobs (approximately 40 percent in Indiana), because public schools mainly employ teachers, teaching assistants, secretaries, security, and cafeteria workers. In contrast, most job units account for less than 1 percent of the state's 552,145 public employees.

Table 9: Percent jobs in each type government unit

Unit	Job %	Unit	Job %	Unit	Job %
School	42%	Other local unit	<1%	Conservancy district utility	<1%
University	23%	Public transportation corporation	<1%	Soil & water conservation district	<1%
City	8%	Housing authority	<1%	Building authority	<1%
County	8%	Airport authority	<1%	Regional planning commission	<1%
Other state unit	7%	Other special district	<1%	Regional water/ wastewater district	<1%
Hospital	4%	Fire protection district	<1%	Flood control district	<1%
Town	2%	Solid waste district	<1%	Regional water district	<1%
Library	1%	Turnaround school	<1%	Military reuse authority	<1%
Township	1%	Regular waste-water (sewer) district	<1%	Port authority	<1%
Charter school	<1%	Conservancy district non-utility	<1%	Redevelopment authority	<1%

n=552,145

Since schools comprise the largest percentage of government jobs, we broke it further into subsectors (Table 10). Within schools, teachers (and other educators) hold 61 percent of jobs. The rest include food preparation, building maintenance, administrative support, etc.

Table 10: School Unit occupation breakdown

Group	Occupation Description	Examples	Jobs %
25-0000	Education, Training, and Library	Teacher, Teaching Assistants, Substitutes	61%
53-0000	Transportation and Material Moving	Bus drivers	7%
35-0000	Food Preparation and Serving Related	Cafeteria staff	7%
37-0000	Building and Grounds Cleaning and Maintenance	Custodians	5%
27-0000	Arts, Design, Entertainment, Sports, and Media	Coaches, Interpreters	4%
11-0000	Management	Principals, Instructional Coordinators	4%
43-0000	Office and Administrative Support	Clerks, Secretaries	4%
00-0000	Other	Security, Lawyers, Computer Support	9%

N=229,246

Geographic

Counties and State (Indiana)

Public employment data comprise of individual job information on the county level. From the county level, we found that public jobs in Indiana accounts for nearly 30 percent. Marion County, the capital and the most populated county in Indiana, has the highest public jobs compared to other counties in the state—over 8 percent. In Marion, more than 36 percent of jobs are in education compared to 42 percent statewide (a statistically significant difference, likely due to the number of state-level jobs in the area).

Table 11: Percent public jobs by county

County	% Jobs	County	% Jobs	County	% Jobs	County	% Jobs
Adams	0.52%	Franklin	0.21%	Lawrence	0.50%	Rush	0.27%
Allen	3.04%	Fulton	0.30%	Madison	1.18%	Scott	0.25%
Bartholomew	0.97%	Gibson	0.36%	**Marion**	**8.57%**	Shelby	0.54%
Benton	0.17%	Grant	0.76%	Marshall	0.53%	Spencer	0.26%
Blackford	0.16%	Greene	0.44%	Martin	0.12%	St. Joseph	2.71%
Boone	0.75%	Hamilton	2.98%	Miami	0.37%	Starke	0.28%
Brown	0.16%	Hancock	0.91%	Monroe	1.14%	Steuben	0.36%
Carroll	0.26%	Harrison	0.44%	Montgomery	0.47%	Sullivan	0.33%
Cass	0.58%	Hendricks	1.52%	Morgan	0.70%	Switzerland	0.12%
Clark	0.94%	Henry	0.66%	Newton	0.22%	Tippecanoe	1.43%
Clay	0.31%	Howard	0.95%	Noble	0.47%	Tipton	0.23%
Clinton	0.40%	Huntington	0.35%	Ohio	0.09%	Union	0.11%
Crawford	0.16%	Jackson	0.60%	Orange	0.32%	Vanderburgh	1.43%
Daviess	0.32%	Jasper	0.35%	Owen	0.17%	Vermillion	0.19%
Dearborn	0.80%	Jay	0.28%	Parke	0.18%	Vigo	1.00%
Decatur	0.38%	Jefferson	0.34%	Perry	0.30%	Wabash	0.48%
DeKalb	0.57%	Jennings	0.26%	Pike	0.17%	Warren	0.11%
Delaware	0.96%	Johnson	1.60%	Porter	1.88%	Warrick	0.53%
Dubois	0.52%	Knox	0.68%	Posey	0.28%	Washington	0.29%
Elkhart	2.11%	Kosciusko	0.88%	Pulaski	0.25%	Wayne	0.74%
Fayette	0.29%	LaGrange	0.36%	Putnam	0.49%	Wells	0.34%
Floyd	1.08%	Lake	5.96%	Randolph	0.38%	White	0.39%
Fountain	0.23%	LaPorte	1.40%	Ripley	0.37%	Whitley	0.31%
						State	29.86%

N=552,145

Table 12: Marion County occupation breakdown

Group	Occupation	% Jobs
25-0000	Education, Training, and Library	36%
33-0000	Protective Service	11%
43-0000	Office and Administrative Support	9%
29-0000	Healthcare Practitioners and Technical	6%
11-0000	Management	6%
53-0000	Transportation and Material Moving	6%
21-0000	Community and Social Service	3%
35-0000	Food Preparation and Serving Related	3%
37-0000	Building and Grounds Cleaning and Maintenance	3%
27-0000	Arts, Design, Entertainment, Sports, and Media	3%
00-0000	Other	15%

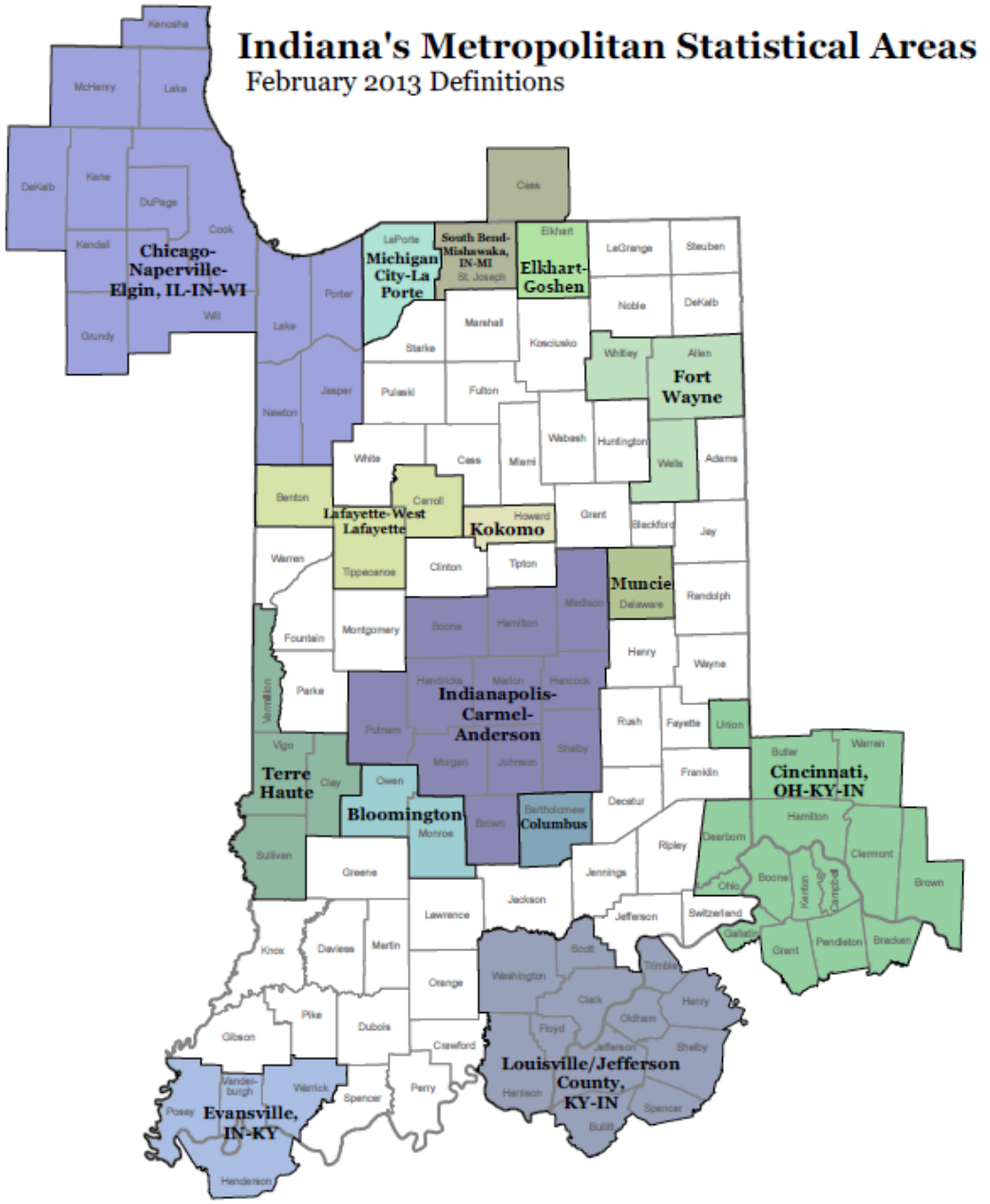
n=47,338

Economic Areas

Looking at Indiana MSAs (Figure 13):

- Indianapolis has nearly 19 percent of all public employees (Table 13).
- Almost all of Muncie's records were codable.
- Education accounts for around 45 percent in Lafayette and Elkhart-Goshen MSAs, compared to only 30 percent in the Cincinnati MSA.
- The Fort Wayne and Louisville MSAs have the most records unable to be coded.
- In most MSAs, healthcare practitioners and technical jobs make up 1-2 percent of total jobs, compared to 10 percent in the Cincinnati MSA and 5 percent in the Indianapolis MSA.

Figure 13: Indiana MSAs



Note: The four Indiana counties in the Chicago metro area form the Gary Metropolitan Division

Source: Indiana Business Research Center
www.stats.indiana.edu

Table 13: Percentage of Occupations by MSA (split across two pages)

	All State	Bloomington	Cincinnati	Elkhart-Goshen	Evansville-Henderson	Fort Wayne	Gary	Indianapolis	Kokomo	Lafayette	Louisville	Muncie	South Bend	Terre Haute	Min	Max
Education, Training, and Library	38%	33%	30%	44%	38%	38%	40%	38%	40%	45%	32%	38%	41%	39%	**30%**	**45%**
Protective Service	10%	10%	7%	8%	11%	10%	11%	10%	10%	10%	8%	12%	13%	12%	**7%**	**13%**
Office and Administrative Support	9%	14%	10%	7%	9%	8%	11%	8%	9%	7%	7%	8%	7%	9%	**7%**	**14%**
Management	7%	8%	8%	5%	6%	7%	5%	6%	6%	6%	6%	7%	5%	8%	**5%**	**8%**
Transportation and Material Moving	6%	7%	3%	6%	5%	5%	6%	6%	7%	6%	5%	5%	6%	8%	**3%**	**8%**
Building and Grounds Cleaning and Maintenance	4%	4%	3%	5%	5%	3%	6%	4%	5%	4%	4%	5%	3%	5%	**3%**	**6%**
Food Preparation and Serving Related	4%	3%	6%	5%	5%	4%	4%	4%	6%	4%	5%	5%	3%	3%	**3%**	**6%**
Healthcare Practitioners and Technical	3%	1%	10%	2%	2%	1%	1%	5%	1%	1%	1%	2%	1%	1%	**1%**	**10%**
Arts, Design, Entertainment, Sports, and Media	3%	2%	3%	4%	2%	3%	2%	3%	2%	3%	2%	5%	2%	1%	**1%**	**5%**
Community and Social Service	2%	1%	2%	2%	2%	2%	2%	2%	2%	1%	1%	2%	2%	2%	**1%**	**2%**
Installation, Maintenance, and Repair	2%	2%	2%	1%	1%	1%	3%	1%	2%	1%	1%	1%	2%	2%	**1%**	**3%**
Personal Care and Service	1%	6%	1%	2%	4%	2%	1%	1%	0%	2%	1%	2%	3%	0%	**0%**	**6%**
Construction and Extraction	1%	1%	1%	1%	2%	1%	1%	1%	3%	1%	2%	3%	2%	1%	**1%**	**3%**

	All State	Bloomington	Cincinnati	Elkhart-Goshen	Evansville-Henderson	Fort Wayne	Gary	Indianapolis	Kokomo	Lafayette	Louisville	Muncie	South Bend	Terre Haute	Min	Max
Business and Financial Operations	1%	2%	1%	1%	1%	1%	1%	2%	1%	1%	1%	1%	1%	1%	**1%**	**2%**
Legal	1%	2%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	**1%**	**2%**
Production	1%	1%	1%	0%	1%	1%	1%	1%	1%	1%	1%	1%	0%	1%	**0%**	**1%**
Life, Physical, and Social Science	1%	1%	1%	1%	1%	1%	0%	1%	1%	0%	1%	1%	1%	1%	**0%**	**1%**
Healthcare Support	1%	0%	3%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	0%	**0%**	**3%**
Computer and Mathematical	1%	0%	0%	1%	0%	1%	0%	1%	1%	1%	0%	0%	1%	0%	**0%**	**1%**
Architecture and Engineering	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	**0%**	**1%**
Sales and Related	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	**0%**	**0%**
Farming, Fishing, and Forestry	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	**0%**	**0%**
Uncodable	5%	1%	7%	4%	3%	8%	4%	4%	1%	3%	21%	1%	5%	4%	**1%**	**21%**

References

2010 SOC User Guide - https://www.bls.gov/soc/soc_2010_user_guide.pdf

Berman, E. M., Rabin, J., & Bartle, J. R. (2007). *Encyclopedia of public administration and public policy*. Marcel Dekker.

Baron, J. N., & Bielby, W. T. (1986). The proliferation of job titles in organizations. *Administrative Science Quarterly*, 561-586.

Nakao, K., & Treas, J. (1994). Updating occupational prestige and socioeconomic scores: How the new measures measure up. *Sociological Methodology*, 1-72.

Strang, D., & Baron, J. N. (1990). Categorical imperatives: The structure of job titles in California state agencies. *American Sociological Review*, 479-495.