

# Workforce Data Quality Initiative | Indiana

## Assigning standard occupation codes using Public Licensing Agency data

The Indiana Business Research Center (IBRC) incorporated public licensing data from Indiana into the Workforce Data Quality Initiative (WDQI) Occupation Assignment project. The Indiana Department of Workforce Development made the de-identified data from 2005 – present available, and provided a randomly generated universal ID field that allowed these data to be matched to their wage records, unemployment insurance claims, and public higher education records from the Indiana Research and Analysis Warehouse data system.

IBRC incorporated these data to develop an approach that allowed the active license(s) of an individual to serve as a proxy for their occupation. To acknowledge that a license holder may be employed in an occupation not related to their license or certification, we developed procedures to check an individual's salary and industry of employment against the occupation(s) associated with each active license.

### Public licensing data

The Indiana Professional Licensing Agency (PLA) maintains a public database and online portal of individuals that have been granted a license or certificate required for an occupation within the state of Indiana. The public is allowed to access the following information for each license:

- Name
- License number
- Address
- Issue date
- Expiration date
- License status (Active, Cancelled, Expired, Pending, Probation, Superseded, Suspended, etc.)
- Date the license status was updated

Currently, records from more than 50 licensing boards are compiled in the PLA database:

- Accountancy Board
- Acupuncture Committee
- Alcoholic Beverage
- Appraiser Board
- Architect Board
- Asbestos
- Athletic Division
- Athletic Trainer Board
- Auctioneer Commission
- Behavioral Health Board
- Chiropractic Board
- Cosmetology & Barber Board
- Dentistry Board
- Dietician Certification Board
- Drinking Water
- Engineer Board
- Environmental Health Board
- Funeral Board
- Gaming
- Health Facility Administration Board
- Hearing Aid Dealer Committee
- Home Inspectors Board
- Hypnotist Committee
- Interior Design Registry

- Land Surveyor Board
- Landfill
- Lead
- Librarian
- Manufactured Home Installers
- Massage Therapy Board
- Medical Licensing Board
- Nurse Aides
- Nursery
- Nursing Board
- Occupational Therapy Committee
- Optometry Board
- Pharmacy Board
- Physical Therapy Committee
- Physician Assistant Committee
- Plumbing Commission
- Podiatric Medicine Board
- Private Investigation & Security Guard Board
- Psychology Board
- Radioactive Material
- Radiology
- Radon
- Real Estate Commission
- Speech Language Pathology & Audio Board
- Timber
- Tobacco
- Veterinary Board
- Wastewater

Two notable absences from the PLA database are records of licensed attorneys and records of licensed educators.

## Methodology

### Determining the dates a license is active

The process to use an individual's license as a proxy for their occupation began with cleaning the PLA database. Many licensing boards that supply data to PLA have their own license status categories, therefore we developed a condensed and standard list of license status values based upon information available from the licensing boards. The standard list of status values was ranked in order of importance to indicate which license should be given priority in the occupation assignment process if there were several licenses held by the same individual. For example, the priority ranking will consider a Home Inspector License with an active status rank to be of a higher priority than a Real Estate Agent License with an expired status rank from the same individual.

<b>Standardized License Status</b>	<b>Priority Rank</b>	<b>Non-Standard Status Values Assigned to Category</b>
<b>Active</b>	1	Active
<b>Conditionally Active</b>	2	Conditional, Probation, Probation/Referral, Valid to Practice While Reviewed
<b>Superseded</b>	3	Superseded
<b>Not Practicing</b>	4	Current/Not Practicing, Retired
<b>Expired</b>	5	Dead, Expired, Expired/Holding Application, Expired/Non-renewable, Inactive, Inactive/Expired, Inactive/Probation, Null and Void, Probation/Expired, Referral/Expired, Registry History, Storage, Unassigned/Expired, Voluntary Surrender
<b>Suspended</b>	6	Emergency Suspension, Suspended
<b>Revoked</b>	7	Finding, Rescinded, Revoked
<b>Not Awarded</b>	8	Abandoned Application, Application Denied, Cancelled, Deleted Application, Failed Exam, Renewal Denied, Withdrawn Application

<b>Pending Application</b>	9	Pending Application, Reinstatement Pending
<b>Unassigned</b>	99	Unassigned

For each license, we tabulated the number of months it was active each quarter of each year. Conditionally active licenses are considered to be active until the status is updated in the database. Not practicing, expired, suspended, and revoked licenses are considered to be active until the date at which the license status was updated to its current status. Not awarded and pending application licenses are considered null and inactive. Superseded licenses are considered active until their expiration date. However, these licenses are typically temporary in nature and are therefore given a lower priority in the occupation assignment process than their more stable, active counterparts. For example, an individual may have a pharmacy technician in training license that expires on December 1, 2016. On September 1, 2016, the in training license was superseded by an active pharmacy technician license. The training license will still be considered active until December 1, 2016, but it will be given a lower priority than the active pharmacy technician license from September 1 to December 1 of 2016.

### Linking licenses, occupations, and industries of employment

The Indiana Department of Workforce Development supplied IBRC with a crosswalk that listed the standard occupation codes (SOC) associated with each type of license or certificate that is tracked in the PLA database. This crosswalk was then joined to the Occupation Employment Statistics (OES) program annual staffing patterns data for Indiana, which list the possible industries of employment for each standard occupation code. The join resulted in a crosswalk for each year that contained one record per license and SOC combination with the full list of possible industries of employment for each record. The staffing patterns were used for the most detailed level of NAICS industry codes available for the year in question. For data prior to 2010, the License to SOC crosswalk was changed to use 2000 SOC codes rather than 2010 SOC codes.

Next, we returned to the PLA database and dropped all licenses that were not active for the year in question. We performed a pairwise join to merge the license-SOC-NAICS crosswalk with the active licenses for the year. These data were joined to the wage records for the year using a one (wage record) to many (license-SOC-NAICS combinations) merge. Non-matched records were dropped, since we are only interested in individuals with one or more active licenses and a valid wage record for the year.

### Assigning a Standard Occupation Code

Every firm in Indiana that pays Unemployment Insurance (UI) on behalf of its employees is required to submit quarterly wage records for each UI-covered employee. The wage records list the 6-digit NAICS industry of employment and the wage per individual and quarter in a calendar year. Records of up to three employers per quarter are maintained in the database, and the model attempts to assign a SOC code for each available simultaneous job in the database of wage records.

The SOC assignment was conducted on a quarterly basis to account for licenses that were not active for an entire calendar year. If a license was active for at least one month in a quarter, all of the possible NAICS industries of employment for each license and SOC combination are evaluated to determine whether there is a match to the dominant NAICS from the individual's wage record. The number of NAICS matches for each individual and active license-SOC combination were counted quarterly. If an individual only had one NAICS match for a quarter, they were assigned the 6-digit SOC code from the license and SOC combination associated with the matching NAICS industry. If there were multiple matches, a person had multiple active licenses for the quarter in question, and the highest priority

license had a single NAICS match, then the SOC code associated with that match was used. Individuals with only a single possible SOC assignment due to one NAICS match were given an assignment description as 6-digit direct.

When a 6-digit direct assignment was not possible, and there were multiple potential SOC codes based upon NAICS matches, less detailed SOC codes were used when possible. For example, an individual with an active Physician License will have a list of seven potential SOC codes associated with that license:

- 29-1061 Anesthesiologists
- 29-1062 Family and General Practitioners
- 29-1063 Internists, General
- 29-1064 Obstetricians and Gynecologists
- 29-1065 Pediatricians, General
- 29-1066 Psychiatrists
- 29-1069 Physicians and Surgeons, All Other

If that same individual works in the General Medical and Surgical Hospitals Industry, they will have a NAICS match to all seven of the potential SOC codes. While it is not possible to assign this individual a 6-digit SOC code, it is possible to assign them a less detailed 5-digit SOC code (29-1060: Physicians and Surgeons) that includes all of the potential 6-digit SOC codes. This assignment would be given a description of 5-digit direct. 4-digit direct and 3-digit direct are also valid assignment descriptions. We considered a 2-digit assignment to be too generic for a license assignment, therefore it was not used.

For the remaining unassigned license holders, we incorporated annualized wage data to eliminate occupations with non-likely wages. National OES staffing patterns with wage data for industry and occupation combinations was joined to the wage/license/SOC/NAICS data. We chose to use national data for industry-specific occupation wages rather than state data because the state data does not have industry-specific wage ranges by occupation. Potential occupations associated with each license and industry of employment were checked to see whether the annualize wage from the wage record fell within the 10<sup>th</sup>-90<sup>th</sup> percentile for the occupation and industry combination. If the annualized wage of the individual fell within the percentile ranges of the national salary range for only one of the possible occupations, then that SOC code was assigned to the individual. If not, then the occupations with wage matches were counted and the process began for determining whether all the potential matches fell into the same 5-digit, 4-digit, or 3-digit, less detailed SOC code. These assignments were given a description of 10-90 wage match.

If there were still multiple potential SOC matches for an individual based upon their active licenses and industry of employment, then the wage match range was narrowed to the 25<sup>th</sup>- 75<sup>th</sup> percentile of the national salary for the occupation and industry in question. The occupations with wage matches within the narrower range were then evaluated to determine whether the matches resulted in a possible 6-digit, 5-digit, 4-digit, or 3-digit SOC assignment. These assignments were given a description of 25-50 wage match.

After following all of these procedures, we were still left with many unassigned license holders, the vast majority of which were individuals with Alcoholic Beverage licenses. For those individuals, the SOC code closest to the median quarterly wage of the person was assigned.

Engineer Board license-holders also required additional measures to assign SOC codes. For potential engineering SOC codes, the median wage for the national industry and occupation combination was compared to the annualized wage of the individual license holder. The potential occupations were ranked by how closely the annualized wage was to the national median. If the Engineering Manager Occupation was ranked first, then we assigned that SOC code to the individual. If the Engineering Manager Occupation was ranked last, then that SOC code was dropped as a potential match. A 6-digit, 5-digit, 4-digit, or 3-digit SOC code was assigned as appropriate and these records were given the description of engineering wage tie break. This method produced an issue with Engineers being over assigned the 17-2171: Petroleum Engineer occupation due to its high salary range, which is similar to the salary range of 11-9041: Architectural and Engineering Managers. Individuals holding Professional Engineering licenses with salaries that match the ranges associated with both Petroleum Engineers and Architectural and Engineering Managers, will be assigned the Architectural and Engineering Managers occupation code unless they work in one of the following industries:

- NAICS 211110: Oil and Gas Extraction
- NAICS 213112: Support Activities for Oil and Gas Operations
- NAICS 486000: Pipeline Transportation
- NAICS 221200: Natural Gas Distribution
- NAICS 324000: Petroleum and Coal Products Manufacturing

There were also a group of unassigned individuals with active asbestos workers licenses. If any of these individuals were employed in the remediation or other waste management services industry, then they were assigned the hazardous materials removal workers SOC code, and the assignment description was listed as asbestos tiebreak.

For the remaining unassigned records, the model identified individuals that held multiple licenses, but one or more of the licenses were not active for all three months in the quarter. Occupations related to a license that was active for less than the full quarter were suppressed, and the model assigned a SOC code based upon the other active license(s) for the quarter in question.

Indiana does not have a specific license for Nurse Practitioners, but does have two prescriptive authority licenses that indicate a nurse is more likely to be a nurse practitioner than a registered nurse. Therefore, if an individual held an active prescriptive authority license and an active registered nurse license, the nurse practitioner license was given priority and assigned to the individual as long as they were employed by a firm in an expected industry.

### Public license module results

	<b>2013</b>
<b>Individuals with an active license for at least one month during the year</b>	461,109
<b>Percent of individuals with more than one active license during the year</b>	26.7%
<b>Licensed individuals on an Indiana payroll for at least one quarter</b>	373,614
<b>Licensed individuals on an Indiana payroll and working in an industry associated with their license for at least one quarter</b>	308,130

Licensed individuals with a SOC assignment for a least one quarter

307,079

			Q2 2013 SOC Assignment Counts		
SOC Code Assignment	SOC Title		Highest paying job for the quarter	2nd Highest paying job	3rd Highest paying job
<b>Management Occupations</b>					
11-9000	Other Management Occupations		24	2	
11-9021	Construction Managers		2	1	
11-9041	Architectural and Engineering Managers		270		
11-9051	Food Service Managers		2,841	16	
11-9061	Funeral Service Managers		436	21	
11-9111	Medical and Health Services Managers		677	36	2
<b>Business and Financial Operations Occupations</b>					
13-2000	Financial Specialists		1		
13-2011	Accountants and Auditors		4,204	309	21
13-2021	Appraisers and Assessors of Real Estate		197	4	
<b>Architecture and Engineering Occupations</b>					
17-1000	Architects, Surveyors, and Cartographers		3		
17-1010	Architects, Except Naval		2		
17-1011	Architects, Except Landscape and Naval		325	10	2
17-1012	Landscape Architects		81	2	
17-1022	Surveyors		222	13	
17-2000	Engineers		4,141	133	4
17-2011	Aerospace Engineers		18	4	1
17-2021	Agricultural Engineers		4		
17-2041	Chemical Engineers		14		
17-2051	Civil Engineers		25	1	
17-2061	Computer Hardware Engineers		7		
17-2071	Electrical Engineers		32	2	
17-2072	Electronics Engineers, Except Computer		22		
17-2111	Health and Safety Engineers, Except Mining Safety Engineers and Inspectors		15		1
17-2112	Industrial Engineers		52	6	
17-2131	Materials Engineers		12		
17-2141	Mechanical Engineers		23	2	
17-2151	Mining and Geological Engineers, Including Mining Safety Engineers		9		
17-2161	Nuclear Engineers		4		
17-2171	Petroleum Engineers		12		
17-2199	Engineers, All Other		52	5	
<b>Life, Physical, and Social Science Occupations</b>					
19-1032	Foresters		125	6	

	19-3030	Psychologists	412	64	2
	19-3031	Clinical, Counseling, and School Psychologists	18	6	1
	19-3032	Industrial-Organizational Psychologists	1		
	19-4091	Environmental Science and Protection Technicians, Including Health	6	1	
<b>Community and Social Service Occupations</b>					
	21-1000	Counselors, Social Workers, and Other Community and Social Service Specialists	605	94	8
	21-1010	Counselors	854	120	5
	21-1011	Substance Abuse and Behavioral Disorder Counselors	266	33	3
	21-1013	Marriage and Family Therapists	158	15	4
	21-1014	Mental Health Counselors	6		
	21-1015	Rehabilitation Counselors	3	1	
	21-1020	Social Workers	2,940	332	32
	21-1021	Child, Family, and School Social Workers	5	7	1
	21-1022	Healthcare Social Workers	6	4	
<b>Education, Training, and Library Occupations</b>					
	25-1071	Health Specialties Teachers, Postsecondary	3		
	25-1194	Vocational Education Teachers, Postsecondary	156	30	1
	25-4021	Librarians	6		
<b>Arts, Design, Entertainment, Sports, and Media Occupations</b>					
	27-1025	Interior Designers	104	6	
<b>Healthcare Practitioners and Technical Occupations</b>					
	29-1000	Health Diagnosing and Treating Practitioners	97	9	1
	29-1011	Chiropractors	215	6	1
	29-1020	Dentists	964	25	1
	29-1021	Dentists, General	61	47	1
	29-1031	Dietitians and Nutritionists	693	97	15
	29-1041	Optometrists	383	18	
	29-1051	Pharmacists	3,328	327	21
	29-1060	Physicians and Surgeons	6,608	1,028	30
	29-1065	Pediatricians, General	1		
	29-1066	Psychiatrists	4	2	
	29-1067	Surgeons	1		
	29-1069	Physicians and Surgeons, All Other	584	61	2
	29-1071	Physician Assistants	674	64	1
	29-1081	Podiatrists	93	7	
	29-1000	Health Diagnosing and Treating Practitioners	103	9	1
	29-1120	Therapists	4	1	
	29-1122	Occupational Therapists	3,148	680	91
	29-1123	Physical Therapists	3,047	596	87
	29-1124	Radiation Therapists	299	18	1
	29-1126	Respiratory Therapists	3,029	302	23
	29-1127	Speech-Language Pathologists	1,974	414	55
	29-1131	Veterinarians	749	19	
	29-1141	Registered Nurses	53,586	5,326	367

	29-1171	Nurse Practitioners	2,938	656	82
	29-1181	Audiologists	255	18	1
	29-1190	Miscellaneous Health Diagnosing and Treating Practitioners	1		
	29-1199	Health Diagnosing and Treating Practitioners, All Other	14	3	
	29-2000	Health Technologists and Technicians	75	21	3
	29-2021	Dental Hygienists	2,907	206	14
	29-2030	Diagnostic Related Technologists and Technicians	20	2	
	29-2031	Cardiovascular Technologists and Technicians	65	3	
	29-2033	Nuclear Medicine Technologists	311	24	1
	29-2034	Radiologic Technologists	2,362	254	20
	29-2050	Health Practitioner Support Technologists and Technicians	1	2	
	29-2052	Pharmacy Technicians	8,280	961	64
	29-2056	Veterinary Technologists and Technicians	674	32	2
	29-2061	Licensed Practical and Licensed Vocational Nurses	17,030	1,880	159
	29-2092	Hearing Aid Specialists	26	5	
	29-9091	Athletic Trainers	638	136	18
	29-9092	Genetic Counselors	26	1	
<b>Healthcare Support Occupations</b>					
	31-1010	Nursing, Psychiatric, and Home Health Aides	6,962	1,950	279
	31-1011	Home Health Aides	3,866	1,034	143
	31-1014	Nursing Assistants	29,335	3,817	345
	31-2021	Physical Therapist Assistants	1,883	354	41
	31-9000	Other Healthcare Support Occupations	1		
	31-9010	Massage Therapists	1		
	31-9011	Massage Therapists	1,076	154	10
	31-9091	Dental Assistants	1,995	86	8
<b>Food Preparation and Serving Related Occupations</b>					
	35-1012	First-Line Supervisors of Food Preparation and Serving Workers	6,255	47	
	35-3011	Bartenders	38,329	11,705	1,224
<b>Personal Care and Service Occupations</b>					
	39-5000	Personal Appearance Workers	122	8	
	39-5010	Barbers, Hairdressers, Hairstylists and Cosmetologists	5		
	39-5011	Barbers	68	11	1
	39-5012	Hairdressers, Hairstylists, and Cosmetologists	7,648	926	62
	39-5090	Miscellaneous Personal Appearance Workers	25	4	
	39-5092	Manicurists and Pedicurists	234	33	3
	39-5094	Skincare Specialists	504	93	4
<b>Sales and Related Occupations</b>					
	41-2031	Retail Salespersons	2,148	181	21
	41-9000	Other Sales and Related Workers	40	2	
	41-9020	Real Estate Brokers and Sales Agents	162	11	
	41-9021	Real Estate Brokers	1,274	72	6



	41-9022	Real Estate Sales Agents	1,768	167	17
	41-9099	Sales and Related Workers, All Other	434	60	4
<b>Construction and Extraction Occupations</b>					
	47-1011	First-Line Supervisors of Construction Trades and Extraction Workers	228	20	2
	47-2130	Insulation Workers	43	6	3
	47-2131	Insulation Workers, Floor, Ceiling, and Wall	35	2	
	47-2132	Insulation Workers, Mechanical	21	4	1
	47-2152	Plumbers, Pipefitters, and Steamfitters	3,897	315	26
	47-4000	Other Construction and Related Workers	10		
	47-4011	Construction and Building Inspectors	265	18	
	47-4040	Hazardous Materials Removal Workers	16		
	47-4041	Hazardous Materials Removal Workers	294	79	15
<b>Installation, Maintenance, and Repair Occupations</b>					
	49-9095	Manufactured Building and Mobile Home Installers	14	2	
<b>Production Occupations</b>					
	51-8031	Water and Wastewater Treatment Plant and System Operators	19	3	
	<b>Total</b>		<b>243,646</b>	<b>35,719</b>	<b>3,368</b>