Automating Occupational Profiles Using R Markdown

GASP! Workshop

McLeod Brown
U.S. Bureau of Labor Statistics
September 23, 2019

Occupational Requirements Survey (ORS)

- The ORS is a survey conducted by the Bureau of Labor Statistics' (BLS) National Compensation Survey (NCS) program.
  - Provides job-related information regarding physical demands; environmental conditions; education, training, and experience; as well as cognitive and mental requirements for jobs in the U.S. economy.
  - **Annual** survey and produces data at the **national** level
  - Newest survey at the BLS, having released its first set of estimates in November of 2016
  - Full estimate list via Excel is published at www.bls.gov/ors

ORS Occupational Profiles

- Released in conjunction with the news release and dataset
  - ORS program also produces high-level occupational profiles of different occupations using ORS data and continues to create and update these throughout the year.
- Currently, we have 40 ORS profiles available on our website
- Standardized layout

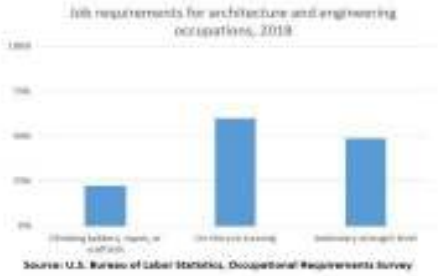
ORS Profile Example

Occupational Requirements Survey: Occupational Profiles

Based on the 2018 Occupational Requirements Survey

- All workers
- Architecture and engineering occupations

This occupational group includes architects, surveyors, cartographers, photogrammetrists, engineers, drafters, engineering technicians, electro-mechanical technicians, and mapping technicians.



Job requirements for architecture and engineering occupations, 2018

Requirements	Estimate
Climbing ladders, ropes, or scaffolds required	22.5%
On-the-job training required	59.0%
Sedentary strength	48.9%

Physical requirements

A sedentary strength level is required for 48.9 percent of workers, a light strength level is required for 21.9 percent of workers, and a medium strength level is required for 22.9 percent of workers. Reaching at or below the shoulders is required for 65.8 percent of workers and reaching overhead is required for 26.0 percent of workers. All (100.0 percent) workers are required to perform fine manipulation and 99.7 percent of workers are required to perform gross manipulation.

Environmental conditions

Working in high and exposed places is a requirement for 17.5 percent of workers. The use of personal protective equipment to mitigate risk associated with this exposure is required for 9.2 percent of workers. Proximity to moving mechanical parts is required for 17.1 percent of workers. A moderate noise intensity level is experienced by 64.5 percent of workers, a quiet noise intensity level is experienced by 30.1 percent of workers, and a loud noise intensity level is experienced by 5.4 percent of workers.

Most workers are not exposed to extreme cold (98.5 percent) or extreme heat (96.4 percent). Climbing ladders, ropes, or scaffolds is a requirement for 22.5 percent of workers and climbing ramps or stairs for work-related tasks is a requirement for 21.6 percent of workers.

Education, training, and experience requirements

Prior work experience is required for 74.4 percent of workers and on-the-job training is required for 59.0 percent of workers. The preparation time necessary for workers is greater than four years but less than or equal to 10 years for 42.8 percent of workers.

Current Process and Reason to Automate

- Very repetitive
  - Staff spends hours reviewing estimates, choosing which are "interesting," and creating profiles
  - Drains resources from other programs within Office of Compensation and Work Conditions (OCWC)
  - Creates conflict with other duties related to Employment Cost Index (ECI), Employer Costs for Employee Compensation (ECEC), Benefits, and other programs especially during production cycles
- With latest release of ORS data, we now have estimates for 420 occupations
  - 420 – 40 = 380 occupations just waiting to be written about!

Why R Markdown

- First used Python
  - Cleaning data beforehand, parsing column text
  - Nested for-loops
  - Output not as neat and compact as we'd like

Why R Markdown, cont.

- R Markdown is used to create fully reproducible documents
  - Cross-language meaning you are able to incorporate your language of choice into the document - R, Python, SQL
  - Supports static and dynamic outputs with HTML, CSS, JavaScript
  - AWESOME documentation and tutorials

R Markdown Starter Code

```
---
title: "ORS Profiles"
author: "U.S. Bureau of Labor Statistics"
date: "19 September, 2019"
output: html_document
runtime: shiny
---
```

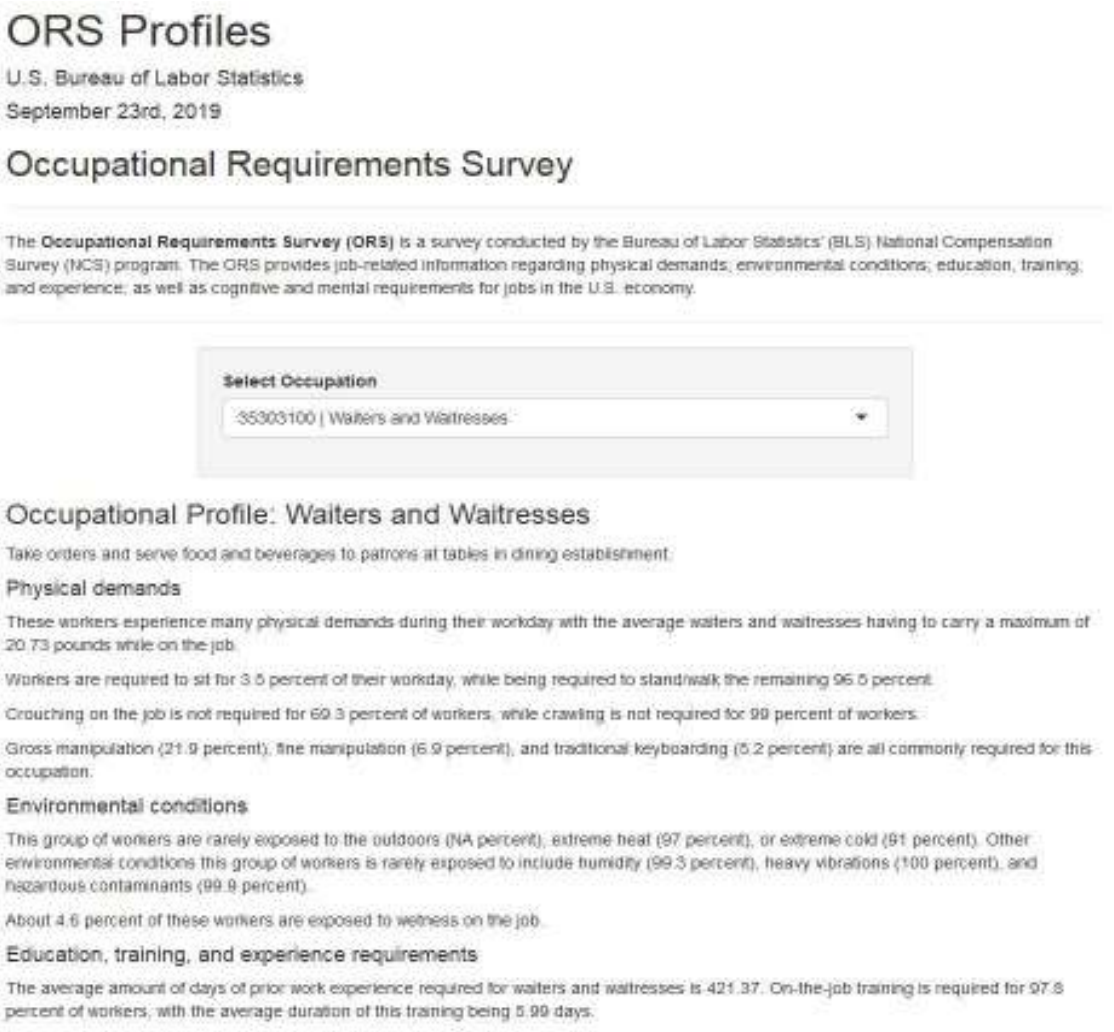
R Markdown Starter Code, cont.

- YAML header
  - "Yet Another Markup Language"
  - Set title, author, date, output, css, figure height and weight options
  - R Markdown cheat sheet
  - Run document or Knit function depending on incorporation of Shiny

Current R Markdown Product

- Our current R Markdown tool is functional and creates a profile for each occupation in our dataset
- Reactive function allows us to make our tool dynamic and... **reactive**
  - Pulls estimates according to occupation chosen
  - Essentially performing a VLOOKUP

Current R Markdown Product, cont.



ORS Profiles

U.S. Bureau of Labor Statistics

September 23rd, 2019

Occupational Requirements Survey

The **Occupational Requirements Survey (ORS)** is a survey conducted by the Bureau of Labor Statistics' (BLS) National Compensation Survey (NCS) program. The ORS provides job-related information regarding physical demands; environmental conditions; education, training, and experience; as well as cognitive and mental requirements for jobs in the U.S. economy.

Select Occupation

35303100 | Waiters and Waitresses

Occupational Profile: Waiters and Waitresses

Take orders and serve food and beverages to patrons at tables in dining establishment.

Physical demands

These workers experience many physical demands during their workday with the average waiters and waitresses having to carry a maximum of 20.73 pounds while on the job.

Workers are required to sit for 3.5 percent of their workday, while being required to stand/walk the remaining 96.5 percent.

Crouching on the job is not required for 69.3 percent of workers, while crawling is not required for 99 percent of workers.

Gross manipulation (21.9 percent), fine manipulation (6.9 percent), and traditional keyboarding (5.2 percent) are all commonly required for this occupation.

Environmental conditions

This group of workers are rarely exposed to the outdoors (NA percent), extreme heat (97 percent), or extreme cold (91 percent). Other environmental conditions this group of workers is rarely exposed to include humidity (99.3 percent), heavy vibrations (100 percent), and hazardous contaminants (99.9 percent).

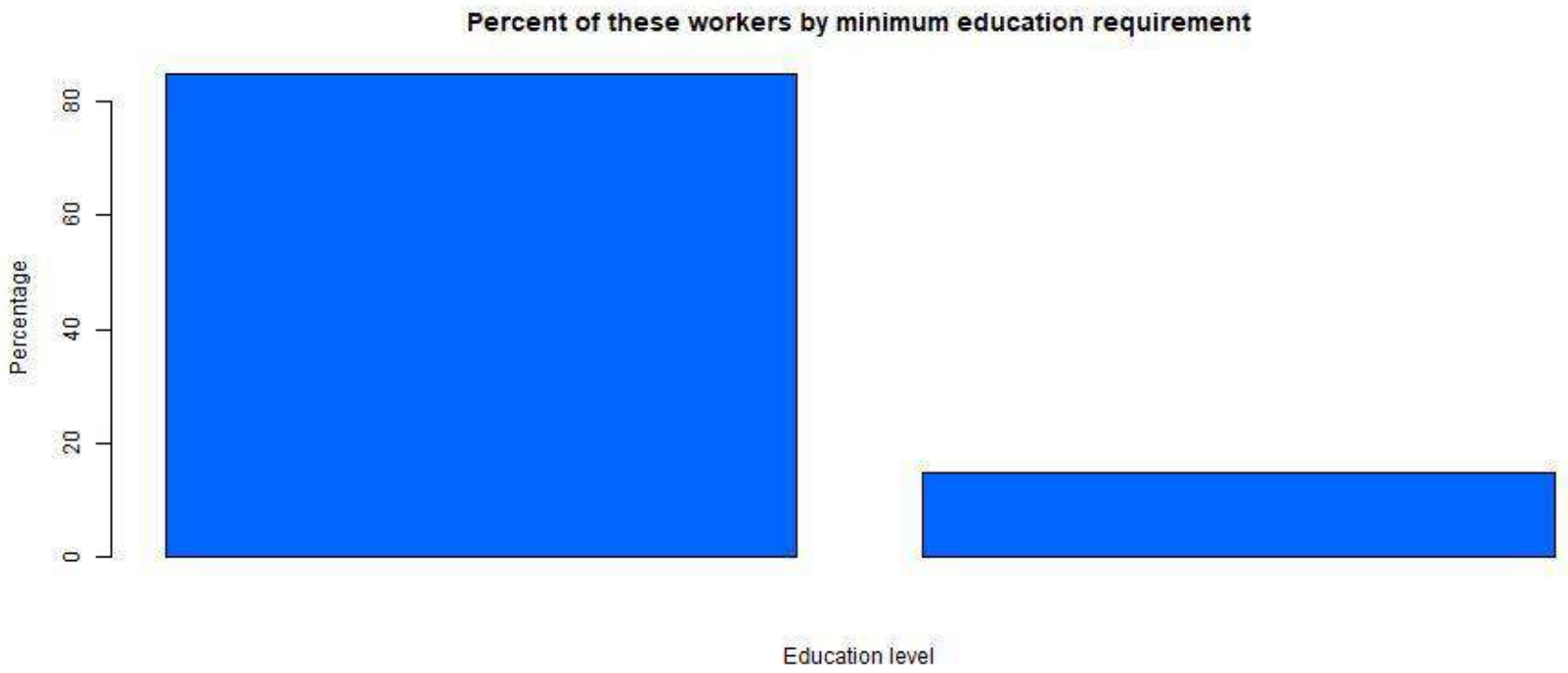
About 4.6 percent of these workers are exposed to wetness on the job.

Education, training, and experience requirements

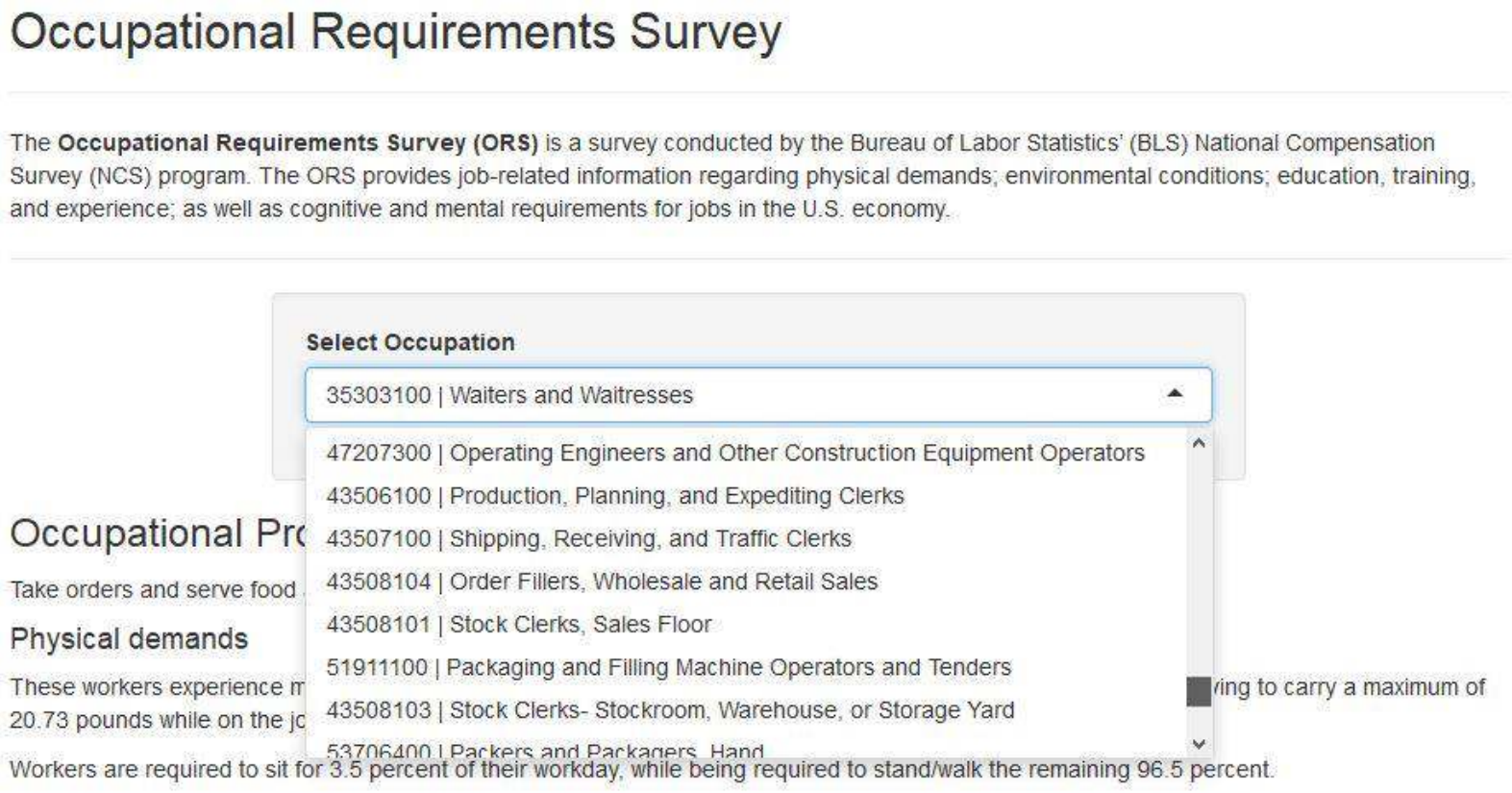
The average amount of days of prior work experience required for waiters and waitresses is 421.37. On-the-job training is required for 97.8 percent of workers, with the average duration of this training being 5.99 days.

Current R Markdown Product, cont.

Select characteristics for Waiters and Waitresses



Current R Markdown Product, cont.



Cons of Current R Markdown Product

- Dynamic, but not as dynamic as we'd like
  - Shows same types of estimates for every occupation
  - Not all estimates are created equal aka some are missing or just not interesting
- So, how do we fix that?

Improving Our R Markdown Product

- We've identified three main ways to improve our tool and have ordered them by most time-consuming to least
  - 1) **Machine Learning**
  - 2) Grouping by major occupation group
  - 3) The params function
- All viable options, but how feasible is each?

Improvements, cont.

- Short-term options are grouping by two-digit SOC or the params function
  - Major SOCs *should be* similar enough
  - Params function offers great flexibility with our program
    - Parameters are declared using “params” field within YAML header
    - (i.e., results for specific geo location or time period)

Improvements, cont.

- Long-term goal is machine learning
  - Gathering training data
  - Perk of the ORS being such a young survey
- Reticulate package
  - Incorporate Python code we've already written

Pros of Current R Markdown Product

- If we rolled this out tomorrow, it would already save our staff hours of data review and profile creation
  - Reduces time from scanning data and pulling estimates
  - Creates charts for staff according to types of estimates desired
  - Responsibilities would revolve around copying and pasting and fact-checking

Pros, cont.

- The ORS is a very diverse dataset, with many different types of estimates so if we want to use this product in other programs...
  - It should be *fairly* easy to implement
  - Modeled Wage Estimates (MWE) profiles
  - Standardized fact sheets
  - Similar reproducible documents

Contact Information

Questions? Comments? Advice?

McLeod Brown
Occupational Requirements Survey
www.bls.gov/ors
202-691-5670
brown.mcleod@bls.gov