

tidyqwi: A Tidy Approach to Accessing The US Census Bureau's Quarterly Workforce Indicators

Michael E. DeWitt Jr¹, Mona Ahmadiani^{1, 2}, and Adam S. Hyde^{1, 2}

1 Wake Forest University 2 Eudaimonia Institute

DOI: [10.21105/joss.01462](https://doi.org/10.21105/joss.01462)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 16 May 2019

Published: 17 June 2019

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

The purpose of `tidyqwi` is to access the U.S. Census Bureau Quarterly Workforce Indicators (QWI) API and return a tidy data frame for further analysis. QWI is a longitudinal dataset across NAICS industry groups that contains quarterly measures of employment flows including employment, job creation/destruction, hires, separations, and turnover – as well as net employment growth among other local labor market indicators. The source data for the QWI is the Longitudinal Employer-Household Dynamics (LEHD) combined with input data from Unemployment Insurance Earnings Data (UI), Quarterly Census of Employment and Wages (QCEW), Business Dynamics Statistics (BDS), and demographic data sources (2000 and 2010 Census, American Community Survey, Social Security administrative records, and individual tax returns) (for more information: https://lehd.ces.census.gov/doc/QWI_101.pdf). The QWI has been currently employed to investigate the effects of minimum wages on employment flows (Dube, Lester, & Reich, 2016), the effects of NOx budget trading program (NBP) on labor markets in the manufacturing sector (Curtis, 2018), the effects of the housing market price on labor market flows (Abowd & Vilhuber, 2012), and the effects of creative destruction on the reported subjective well-being of individuals (Ahmadiani, Hyde, & Jackson, 2019).

The QWI API has specific requirements including three endpoints of “Sex/Age,” “Sex/Education,” and “Race/Ethnicity” for data request. Additionally, a cardinality limit of 400,000 cells has been placed on an API query requests which makes calls for many industries and associated cross-tabulations of different endpoints challenging. API calls allow for collecting data at smaller geographical levels (i.e., metropolitan/micropolitan areas and county levels) that could make the data collection tedious if it gets to 400,000 cells limit. These aspects make retrieving multi-state, multi-industry data difficult and time-consuming.

Taking inspiration from `tidycensus` (Walker, 2018) which has provided a robust method for accessing US Census data from the decennial census and American Community Survey, `tidyqwi` provides a friendly way to interface with the US Census' available API for Quarterly Workforce Indicators. `tidyqwi` provides a way to easily access the US Census' API with multi-state calls, over multiple years, variables and cross-tabulations. The `tidy_qwi` function allows the user to specify key fields for retrieval (years of interest, quarters of interest, variables of interest, industry levels (NAICS 2-digits, 3-digit, and 4-digit codes), specific states, county/CBSA level, and the cross tabulation among other fields). Relying heavily on Wickham (2017) and Ooms (2014) multiple API calls are constructed structured following the [API documentation](#) (Bureau, 2018). These calls are then submitted to the US Census Bureau using the user's API Key (which can be requested from the US Census Bureau [here](#)). Internal to the `get_qwi` function these multiple returned calls are aggregated into a single `tbl_df` object with potential missing data from the US Census represented as NA.

The `tbl_df` object that is returned allows for labels to be added with the `add_qwi_labels` function if desired for easier interpretation of the variable names. This allows researchers the

ability to retrieve data for further analysis without having to specify multiple queries. Tidyqwi thus provides an ergonomic and replicable way of collecting data for economics research on local labor markets.

References

- Abowd, J. M., & Vilhuber, L. (2012). Did the Housing Price Bubble Clobber Local Labor Market Job and Worker Flows When It Burst? *The American Economic Review*, *102*(3), 589–593. doi:[10.1257/aer.102.3.589](https://doi.org/10.1257/aer.102.3.589)
- Ahmadiani, M., Hyde, A. S., & Jackson, J. J. (2019). Creative destruction, job reallocation, and subjective well-being. *Working paper*.
- Bureau, U. C. (2018, September). Quarterly workforce indicators (qwi) (time series: 1990 - present). U.S. Department of Commerce. Retrieved from <https://www.census.gov/data/developers/data-sets/qwi.html>
- Curtis, E. M. (2018). Who Loses under Cap-and-Trade Programs? The Labor Market Effects of the NO_x Budget Trading Program. *The Review of Economics and Statistics*, *100*(1), 151–166. doi:[10.1162/REST_a_00680](https://doi.org/10.1162/REST_a_00680)
- Dube, A., Lester, T. W., & Reich, M. (2016). Minimum Wage Shocks, Employment Flows, and Labor Market Frictions. *Journal of Labor Economics*, *34*(3), 663–704. doi:[10.1086/685449](https://doi.org/10.1086/685449)
- Ooms, J. (2014). The jsonlite package: A practical and consistent mapping between json data and r objects. *arXiv:1403.2805 [stat.CO]*. Retrieved from <https://arxiv.org/abs/1403.2805>
- Walker, K. (2018). *Tidycensus: Load us census boundary and attribute data as 'tidyverse' and 'sf'-ready data frames*. Retrieved from <https://CRAN.R-project.org/package=tidycensus>
- Wickham, H. (2017). *Httr: Tools for working with urls and http*. Retrieved from <https://CRAN.R-project.org/package=httr>