PM566 Final Project

Audrey Omidsalar

December 10th, 2021

Contents

Introduction

An issue that I have seen being constantly spoken about is the economy and the rise of inflation. Being a student in the science field, I have very little background on economics; frankly, the last time I took a course in economics was in high school. I was interested in examining wage disparities, more specifically throughout California, within the education sector and within colleges and universities, as this is the area I'm currently working in.

The question that I will be exploring is: How have wages throughout California, in the education sector, and in Colleges and Universities changed from 2004 until 2020? The dataset I used is California's Quarterly Census of Employment and Wages from data.gov (https://catalog.data.gov/dataset/quarterly-census-of-employment-and-wages-qcew). It contains 4351631 rows and 15 columns, categorized by the following variables:

- Area type (county, state-wide, US; character data)
- Area name (name of county; character data)
- Year (2004 - 2021; integer data)
- Quarter (1st - 4th Quarter; character data)
- Ownership (state government, federal government, private; character data)
- NAICS (North American Industry Classification System) level (integer data)
- NAICS code (character data)
- Industry name (character data)
- Number of Establishments (integer data)
- Average monthly employment (integer data)
- 1st, 2nd and 3rd month employment (if there is a quarter designation; integer data)

- Total wages for all workers (numerical data)
- Average weekly wages (numerical data)

The specifications I was interested in looking in for this project were the average weekly wages throughout various counties in California, within the Education as a whole, and within the *Colleges and Universities* industry.

Methods

The dataset was acquired from the data.gov database and can be downloaded using the following link: https://data.edd.ca.gov/api/views/fisq-v939/rows.csv?accessType=DOWNLOAD. It was downloaded as a csv (comma-separated values) file, and imported to Rstudio using the *data.table* package. This dataset is quite large (531.5 MB), so I cleaned the data using the following steps:

1. I filtered the dataset to only include average annual wages. I did by looking at the column named *quarter*, which subdivides the data per each quarter (1st through 4th) as well as the annual values. In order to control for fluctuations throughout the year, I used the annual values for the subsequent steps. As a result, I removed the *1st, 2nd and 3rd Month Employment* columns as those have NA values when looking at annual data.

2. I omitted any nationwide observations (where the *Area Type* column was set to *United States*) to focus on California-specific observations only.

3. A smaller dataset (*education*) was made by subsetting the dataset by NAICS code values that begin with 61, which is the code for Education Services[1].

4. The *education* dataset was subset to make a *colleges and universities* dataset that only contained information from the *Colleges and Universities* as the Industry Name. In order to make the last two plots in this report, this was separated by ownership type (*private* and *state government*), and the *average weekly wages* were adjusted for inflation (adjusted to equivalent prices in 2020) using the package *priceR*.

The graphs in this report were made using the ggplot2 package, and interactive plots were made with plotly.

Dependencies / Packages Used:

- data.table
- ggplot2
- dplyr
- tibble
- tidyverse
- plotly
- DT
- knitr
- priceR

Results

1. Which industries have the highest and lowest wages?

To do this, I first searched for the median average weekly wage for each unique industry overall per year. I chose to look at the median rather than the mean in order to control for any outliers. There are likely

industries that have seasonal changes, and I felt the median would be a fair estimation to compare. The top and bottom five industries are listed below.

The industries with the highest median wages include *all other information services* (this includes some big tech companies like Facebook, Linkedin and Yelp), *primary production of alumina and aluminum* (likely adjacent to the tech industry), *petrochemical manufacturing* (includes aerospace companies, making petroleum byproducts), *automobile manufacturing* (Tesla is likely a big contributor to this in California), and *credit card issuing*[2].

Table 1: 5 Industries with the Highest Median Weekly Wages per Year

Industry Name	Year	median_wage
All Other Information Services	2020	8496.0
Primary production of alumina and aluminum	2019	6633.0
Petrochemical Manufacturing	2017	6529.0
Automobile Manufacturing	2020	5807.5
Credit Card Issuing	2018	5384.0

The industries with the lowest median wages are *mens underwear & nightwear manufacturing* (it is commonly known that fast fashion does not give livable wages to many of its employees), *motion picture theaters, ex. Drive-Ins*, and *motion picture and video exhibition.* I would guess that these motion picture industries have a lot of fluctuations throughout the year and likely have high numbers of part-time employees, which would contribute to the low weekly wages. In addition, these observations are all from the two earliest years from this dataset; I would expect wages to increase yearly due to cost-of-living adjustments, so it makes sense why wages from 2004 and 2005 would be among the lowest.

Table 2: 5 Industries with the Lowest Median Weekly Wages per Year

Industry Name	Year	median_wage
Mens Underwear & Nightwear Manufacturing	2004	185.5
Motion Picture Theaters, ex. Drive-Ins	2004	193.0
Motion Picture Theaters, ex. Drive-Ins	2005	196.0
Motion Picture and Video Exhibition	2004	201.0
Motion Picture and Video Exhibition	2005	201.5

2. Which counties have the highest and lowest wages?

To do this, I found the median of the average weekly wage for each unique county per year. Again, I chose to look at the median rather than the mean in order to control for any outliers. The top and bottom five industries are listed below.

San Francisco County, San Mateo County, and Santa Clara County have the highest median weekly wages overall. This makes sense given that there is a high cost of living in these areas and there are many high-paying tech-related industries in these areas. It also makes sense that the observations among the top 5 are from recent years, as I would expect salaries to increase yearly due to cost-of-living adjustments.

Table 3: 5 Counties with the Highest Median Weekly Wages per Year

Area Name	Year	median_wage
San Francisco County	2020	1769.0
San Francisco County	2019	1603.0
San Francisco County	2018	1511.5
San Mateo County	2020	1461.0
Santa Clara County	2020	1432.0

The counties with the lowest median wages are Trinity County and Siskiyou County. These counties are located in more mountainous areas in Northern California near the Oregon border. The median wages being so low in these areas likely correspond to low population (Trinity County is the 4th least populated county in California) as well as seasonal jobs (likely more jobs in the winter months), and perhaps more part-time or volunteer workers since there are no major cities in these counties[3].

Table 4: 5 Counties with the Lowest Median Weekly Wages per Year

Area Name	Year	median_wage
Trinity County	2004	434
Siskiyou County	2004	441
Siskiyou County	2005	456
Trinity County	2006	458
Trinity County	2008	459

3. Which education sectors have the highest and lowest weekly wages?

I filtered the dataset to the top average weekly wages per industry. The top and bottom five are listed below.

The industries with the highest median wages are *Computer Training*, *Educational Support Services* and *Flight Training*. This makes sense to me given that the computer training category is tech-adjacent. I also know that flying is very costly, and educational support services (college counseling, standardized test preparation) can also be very costly.

Table 5: Top 5 Median Weekly Wages per Year

Area Name	Industry Name	Year	median_wage
Sacramento County	Computer Training	2017	2981
Sonoma County	Educational Support Services	2020	2890
San Mateo County	Computer Training	2016	2843
San Mateo County	Computer Training	2017	2784
Solano County	Flight Training	2020	2591

The industries with the lowest median wages are *Language Schools* and *Sports and Recreation Instruction.* Some contributing factors as to these low values are part-time and volunteer employees. The median wage of zero in Kings County means that either all the employees were volunteers (I do know that many of the sports leagues in communities have parents or older children volunteer to be coaches, so I didn't feel comfortable removing this value) or there were errors in inputting this data.

Table 6: Bottom 5 Median Weekly Wages per Year

Area Name	Industry Name	Year	median_wage
Kings County	Sports and Recreation Instruction	2004	0
Yolo County	Language Schools	2010	46
Yolo County	Language Schools	2011	49
Yolo County	Language Schools	2006	51
Yolo County	Language Schools	2012	52

The below table shows the highest weekly wages per industry in the *education* category, and is an interactive table where the dataset can be explored. One of the top industries mentioned earlier (*computer training*) is included here; however, the *Colleges and Universities* industry had the highest average weekly wage out of all the industries within the *education* sector. This makes sense to me given that this category likely encompasses universities with large endowments and high-salaried employees such as coaches for popular sports and the board of directors. Additionally, this industry will be the focus of the subsequent analysis; I was interested in looking at this in particular because I am a student worker at my university.

Table 7: 5 Education Sectors with the Highest Average Weekly Wages

Area Name	Industry Name	Year	Average Weekly Wages
California	Colleges and Universities	2019	3638
California	Colleges and Universities	2019	3638
California	Colleges and Universities	2019	3638
Sacramento County	Computer Training	2017	2981
Sacramento County	Computer Training	2017	2981

4. Within the *Colleges and Universities* sector, how have wages changed from 2004 to 2020?

In order for the comparisons on wages to be a little more fair, I adjusted these wages for inflation, using the `priceR` package. I corrected the values in the *Average Weekly Wages* column for inflation in the US to the year 2020 (this is the most recent year in this colleges and universities dataset). All subsequent data and graphs will be based on these inflation-corrected wages.

This first scatterplot shows the annual average weekly wages in local government, private, and state government colleges and universities throughout California. The outliers that pop out to me are private colleges/universities in San Mateo County (Stanford is located here), Santa Clara County (Santa Clara University is here), and Los Angeles County (USC, among others). These are all in areas with major cities and a high cost-of-living. Interestingly, Los Angeles also appears to have the only local government-owned colleges and universities in the state and only contains entries from 2006 to 2011; I believe these schools are the Los Angeles Community College District of schools[4]. However, since only five years' worth of data is shown for these colleges and they are still open to this day, it is very likely that the *ownership* and/or *Area Name* for these observations were input differently than in other years. Additionally, there seems to generally be a smaller range in average wages in the colleges and universities that are owned by the state (the points are more clustered together vertically), with San Joaquin County being an exception with comparatively lower wages. This is likely due to standards and policies that are maintained across these institutions. Something to note is that some counties do not have both private and state-government run colleges and universities, and some annual wages were not reported.

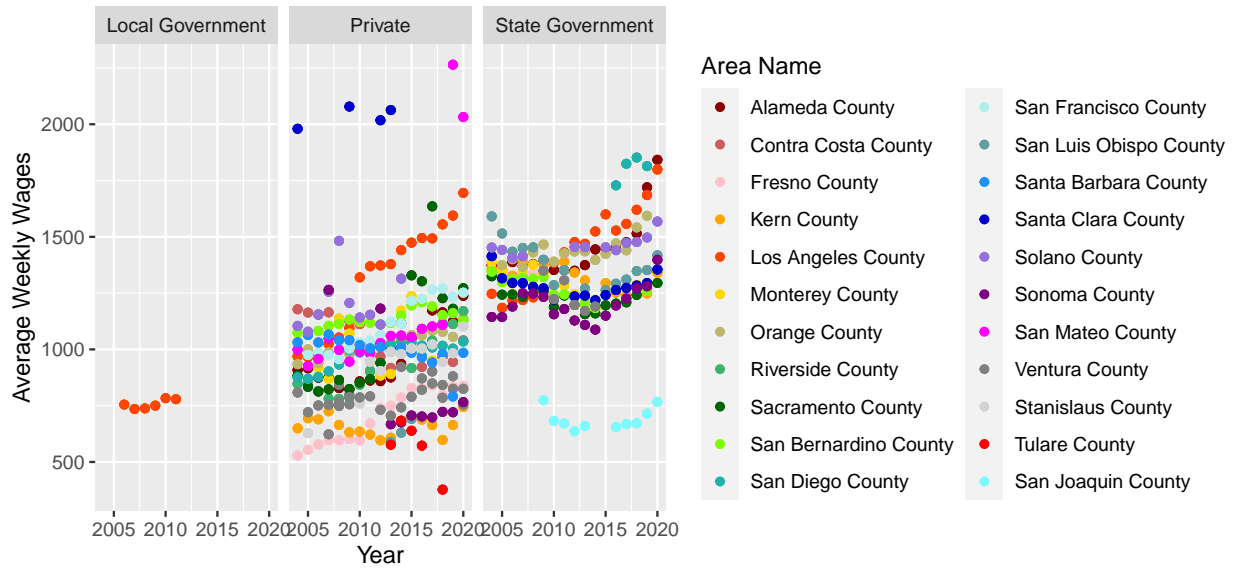# Figure 1: Average Weekly Wages at Colleges and Universities in California

Figure 1 *Average Weekly Wages at Colleges and Universities in California* This graph shows the annual average weekly wage in the Colleges and Universities industry, separated by ownership (local government owned, privately owned, state government owned) throughout California. It is colored according to county.

To look more closely at how salary has grown, I calculated the percent changes (compared to the previous year reported) for privately owned and state government owned colleges and universities.

The bar graph plotting these percent changes for colleges and universities owned by the state government is below. Some years with more negative percent changes that stand out to me are 2008-2010. I know that there there was a big recession during this time period. For the most part and especially in the last five years, there are positive percent changes in wages, which likely has to do with annual cost-of-living adjustments.

Figure 2: Percent Changes in Average Weekly Wages of State Government Owned Colleges and Universities in California
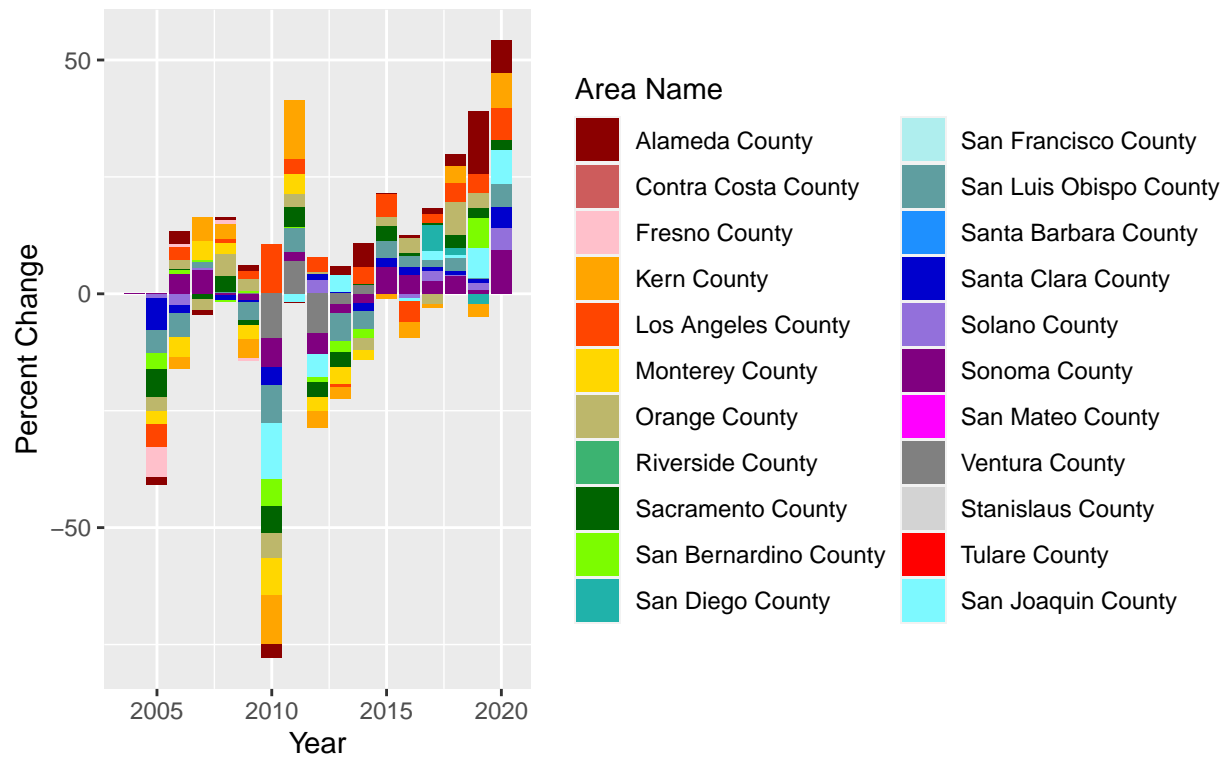
Figure 2 *Percent Changes in Average Weekly Wages of State Government Owned Colleges and Universities in California* This graph shows the yearly percent change in the average weekly wages of State Government owned Colleges and Universities throughout California. It is colored according to county.

Below is the bar graph showing the percent changes for privately owned colleges and universities in California. Overall, I see more negative percent changes here compared to the previous graph, and this is likely due to the nature of these universities not having government regulation and having more autonomy with how they manage themselves. What stands out to me the most here is the positive percent change from 2018 to 2019 in San Mateo County. I believe Stanford University is the major private university in this county, and I found an article stating that their football coach David Shaw received "over 8.9 million in pay" in 2019, so this was likely a big contributor to the jump in average weekly wage for that year[5].

Figure 3: Percent Changes in Average Weekly Wage of Privately Owned Colleges and Universities in California
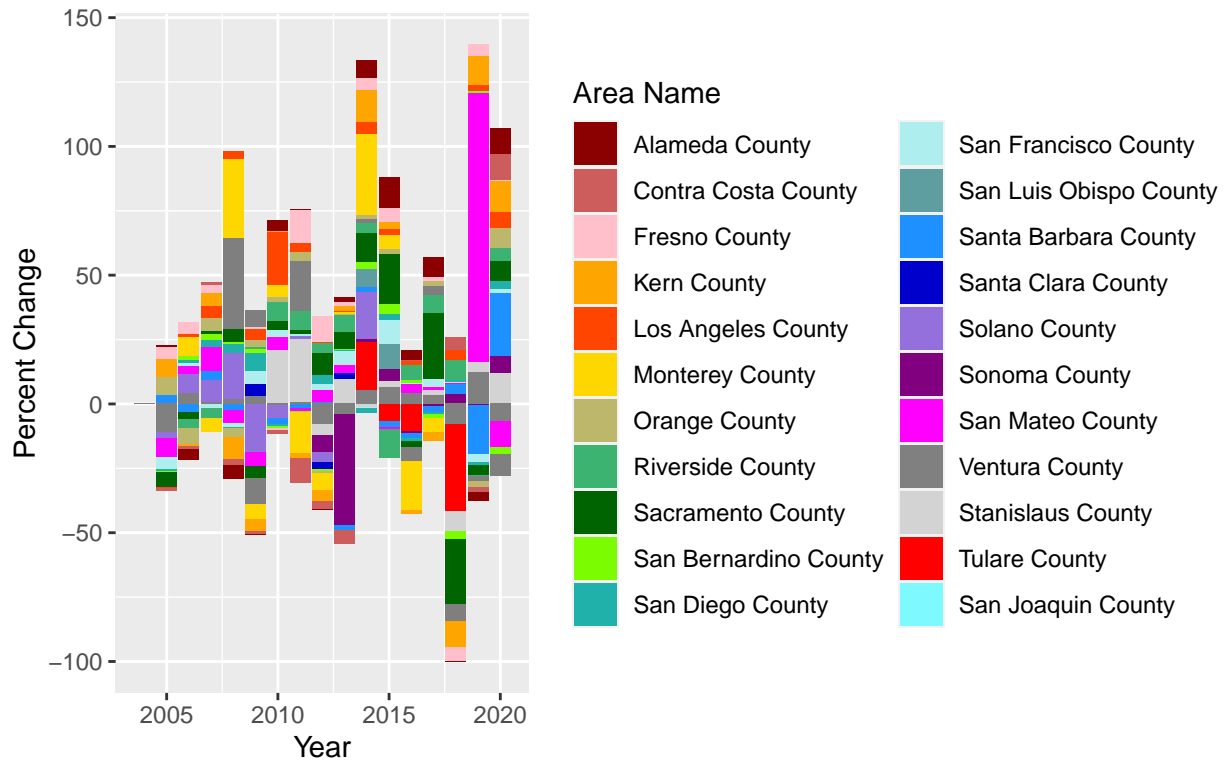

Figure 3 *Percent Changes in Average Weekly Wage of Privately Owned Colleges and Universities in California* This graph shows the yearly percent change in the average weekly wages of privately owned Colleges and Universities throughout California. It is colored according to county.

Conclusion

From this analysis, I have found that counties with major cities generally have higher wages than those that do not. Many of the top wages come from industries that are tech-related and are especially prevalent in the areas surrounding Silicon Valley. There is less county-wide variation in average weekly wages in colleges and universities owned by the state government compared to those that are privately owned. Colleges and universities that are owned by the state government also appear to be more likely to have positive percent changes in annual wage.

References

1. "NAICS & Sic Identification Tools." NAICS Association, https://www.naics.com/search/.
2. "NAICS Code: 519190 All Other Information Services." NAICS Association, https://www.naics.com/naics-code-description/?code=519190.
3. "California Counties by Population." California Outline, https://www.california-demographics.com/counties_by_population.
4. "Our Colleges." LACCD Colleges, https://www.laccd.edu/about/pages/our-colleges.aspx.
5. Berkowitz, Steve. "Stanford Football Coach David Shaw Credited with More than $8.9 Million in Pay for 2019." USA Today, Gannett Satellite Information Network, 4 Aug. 2021,

https://www.usatoday.com/story/sports/ncaaf/pac12/2021/08/04/stanford-football-coach-david-shaw-2019-pay/5488869001/.