

# UC San Diego

## **California Wage Data Analysis** **11/23/2022**

Yuhao Huang, Zhiyan Zhu, Xi Yang, Yusuf Morsi, Siddharth Satyam

# Contents

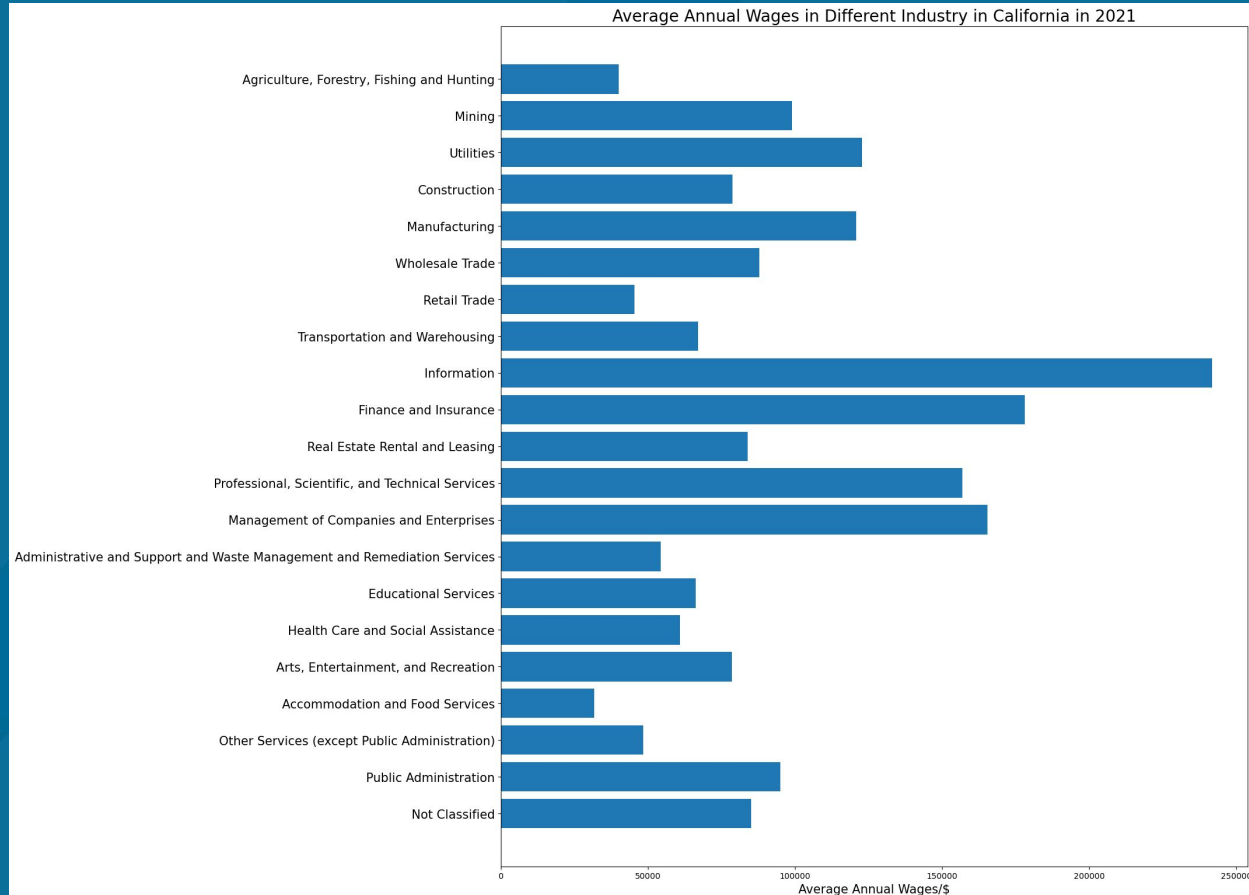
- **Overview**
- **Industry Type**
- **Ownerships**
- **Area Type Comparison**
- **Establishments**
- **Time Series Analysis and prediction for Employment Data**
- **Conclusion**

# Overview

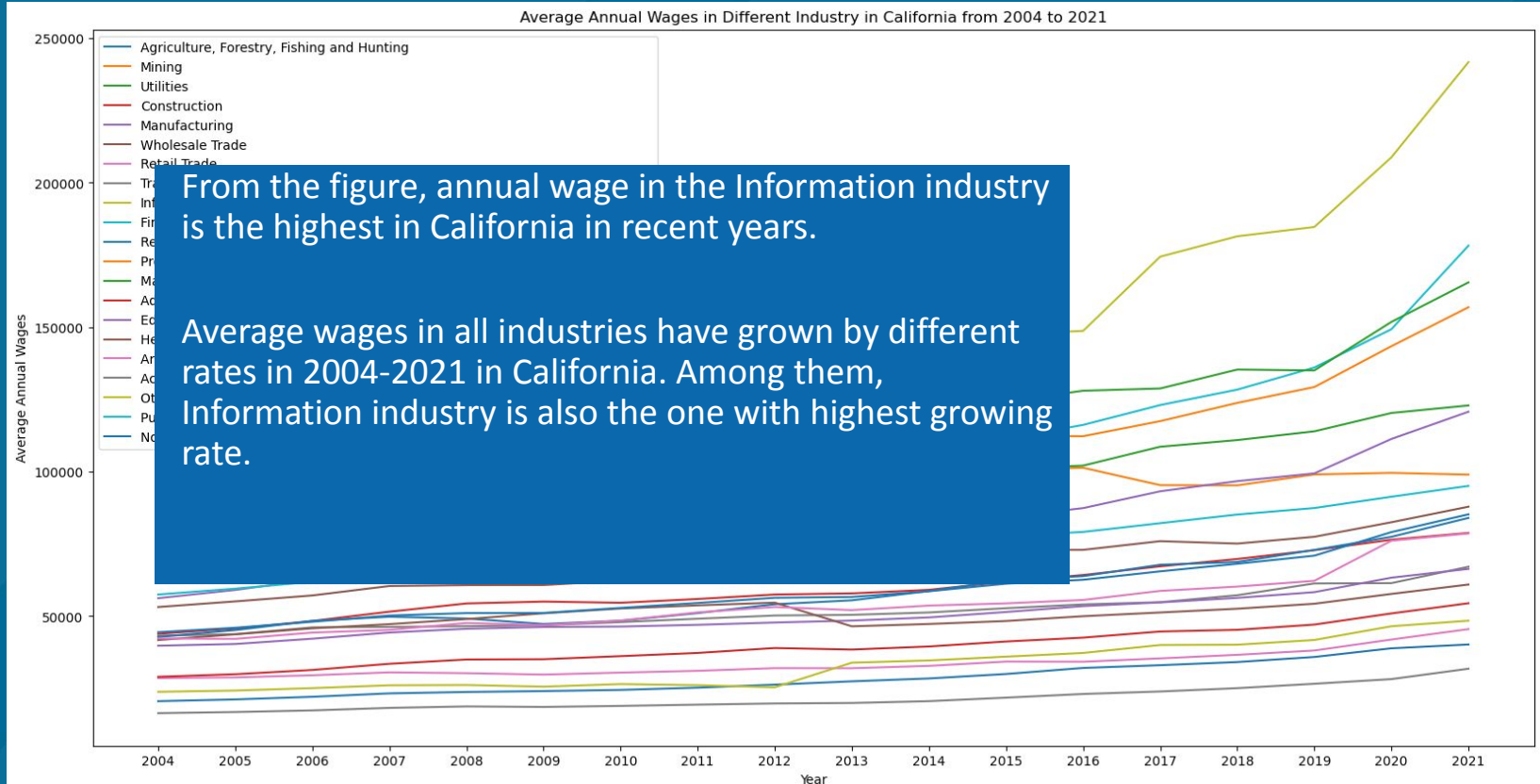
- Data Source: Quarterly Census of Employment and Wages (QCEW)
- Questions:
  1. What is the major industry types of California?
  2. What can we learn from the changing trend of employment of different ownership?
  3. What are the correlations between the trend of establishments, wages and monthly employment?
  4. Can we find out which industries are seasonal in nature?

# Industry Type

# Average Annual Wages in Different Industry in California in 2021



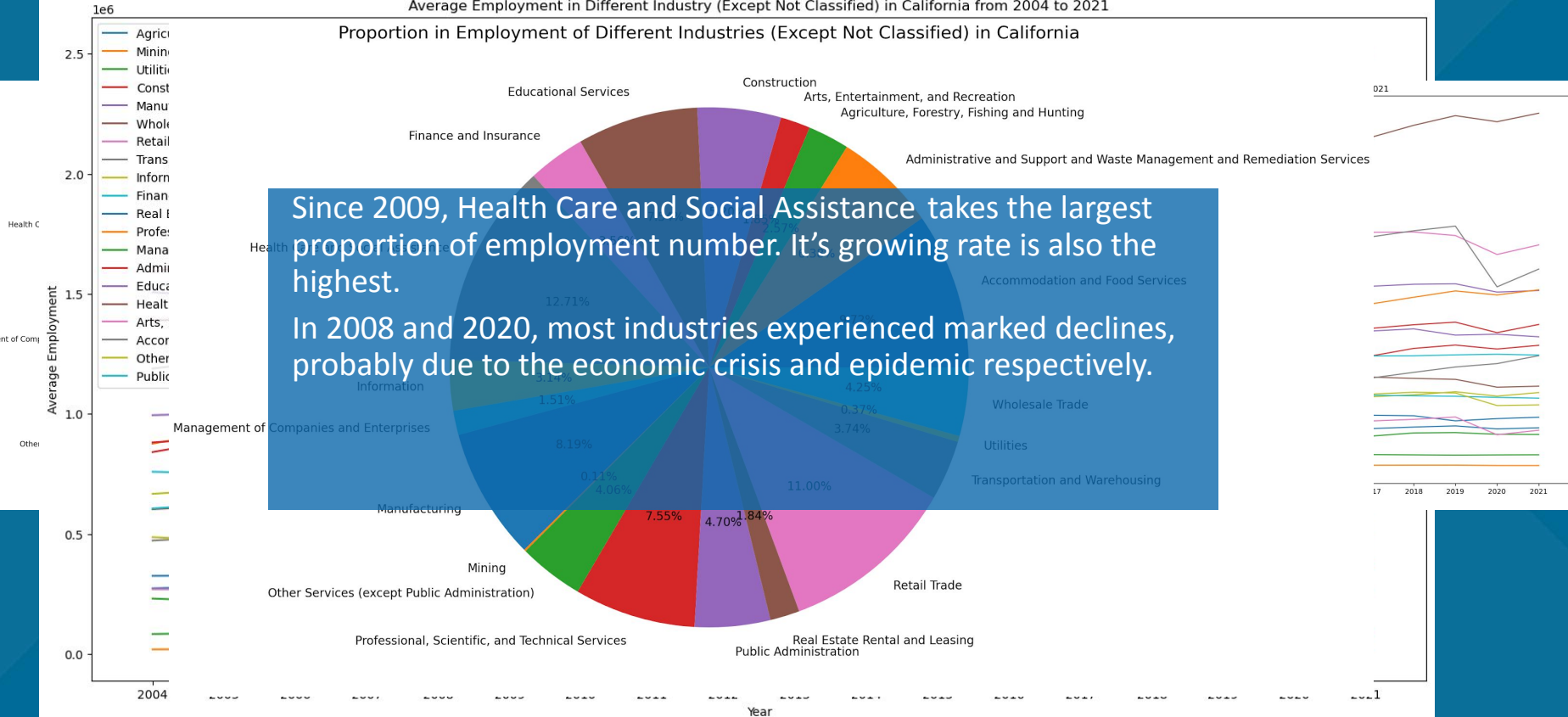
# Average Annual Wages in Different Industry in California from 2004 to 2021



# Employment of Different Industries

Average Employment in Different Industry (Except Not Classified) in California from 2004 to 2021

Proportion in Employment of Different Industries (Except Not Classified) in California

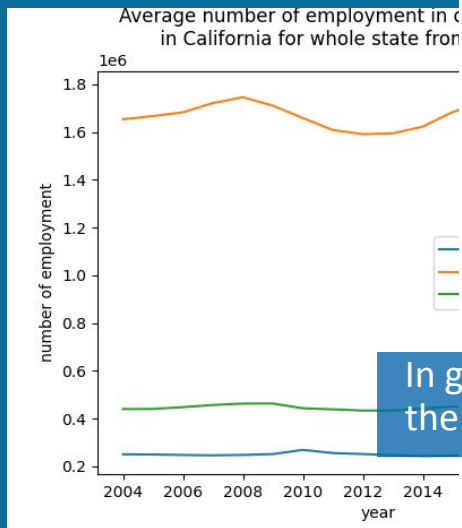


# Ownership



Number of employees with different ownership in California and the United States from 2004 to 2021

Private enterprises have the largest number of employees. We found that employment declined in 2008 and 2019. The reason is as the last student said.

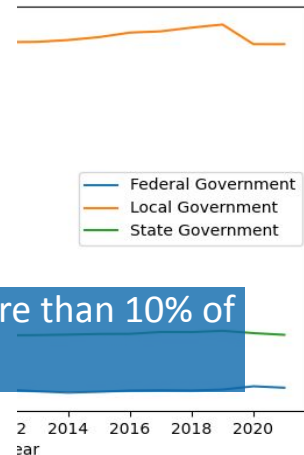


California's share of employment of the United States



In general, California's employment accounts for more than 10% of the U.S., which sounds great.

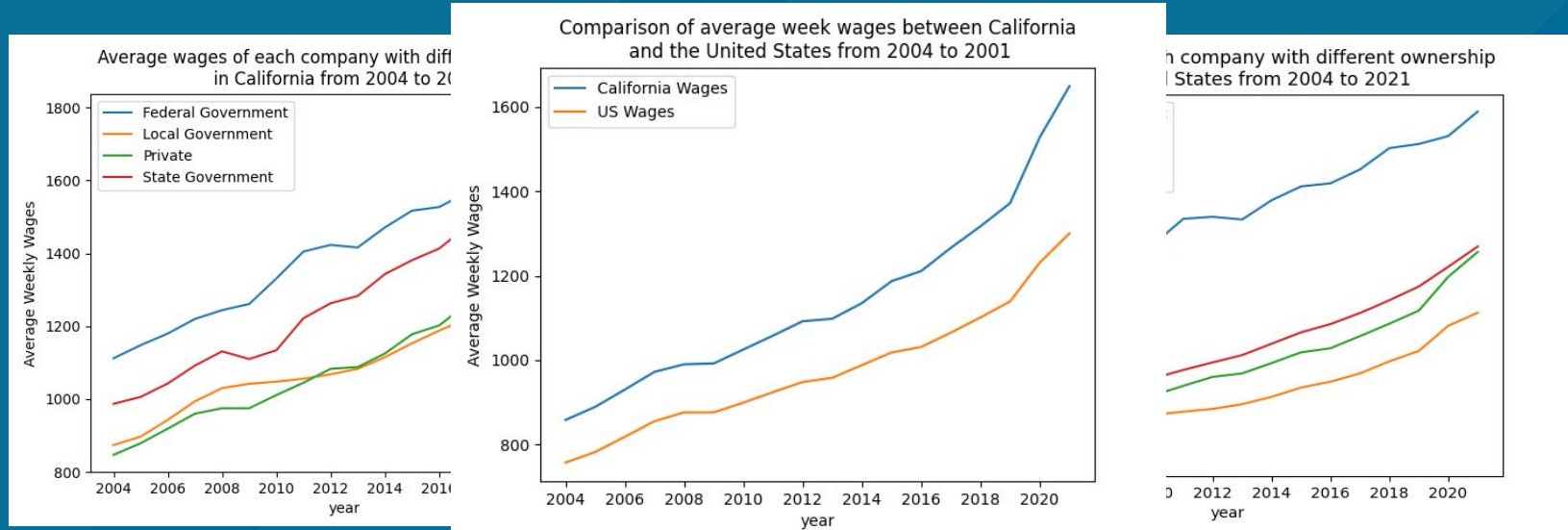
Employment in different ownership 2004 to 2021



In comparison with the United States, there is a certain decline, the overall employment remains stable.

there is a

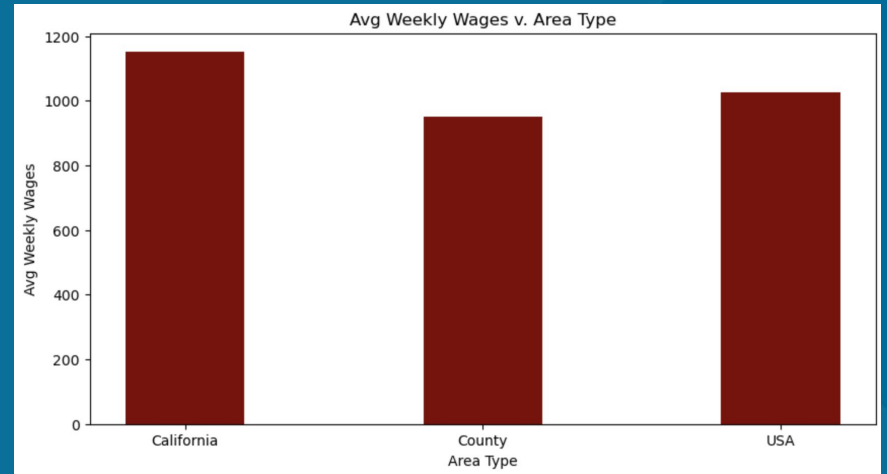
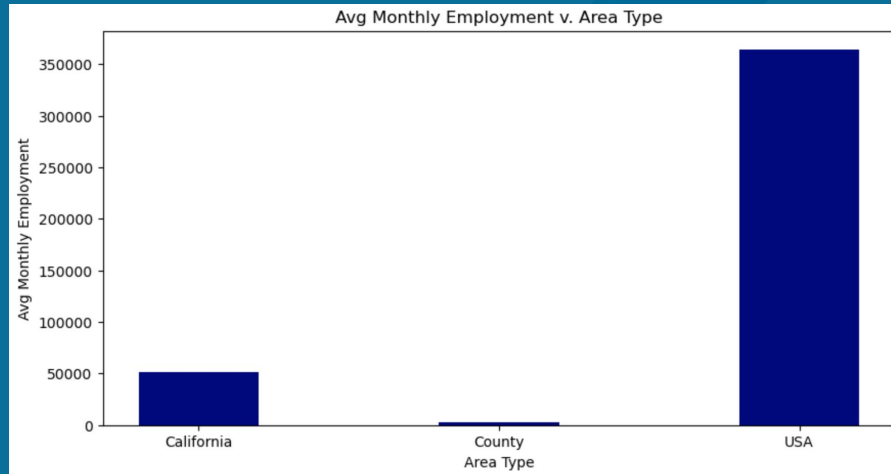
# Average weekly wages of each company with different ownership in California and the United States from 2004 to 2021



It looks like that California's enterprises not only have more employees, but also their average wages are higher than the average of the whole United States. This is really a good thing!

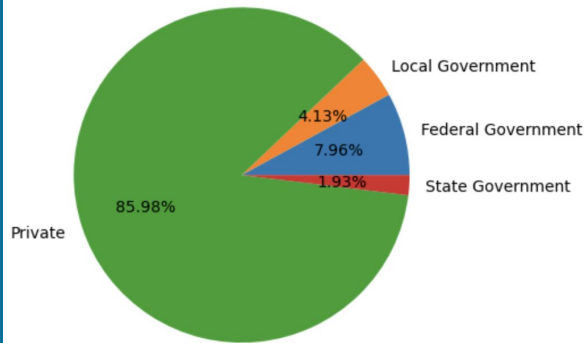
# Area Type Comparison

# Avg Monthly Employment, Weekly Wages Based on Area Type (2004-2021)

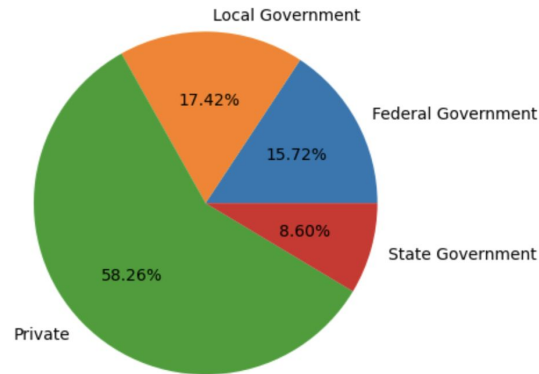


# Average Ownership Based on Area Type from 2004 to 2021

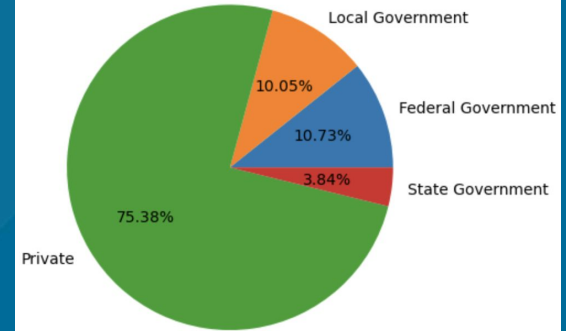
County Ownership Comparison



United States Ownership Comparison



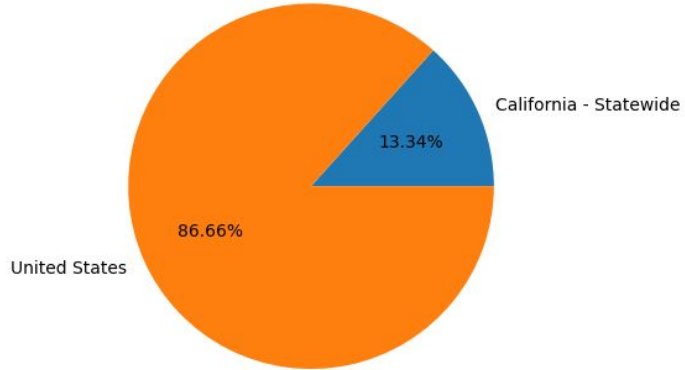
California Ownership Comparison



# Establishments

# Establishments in different areas

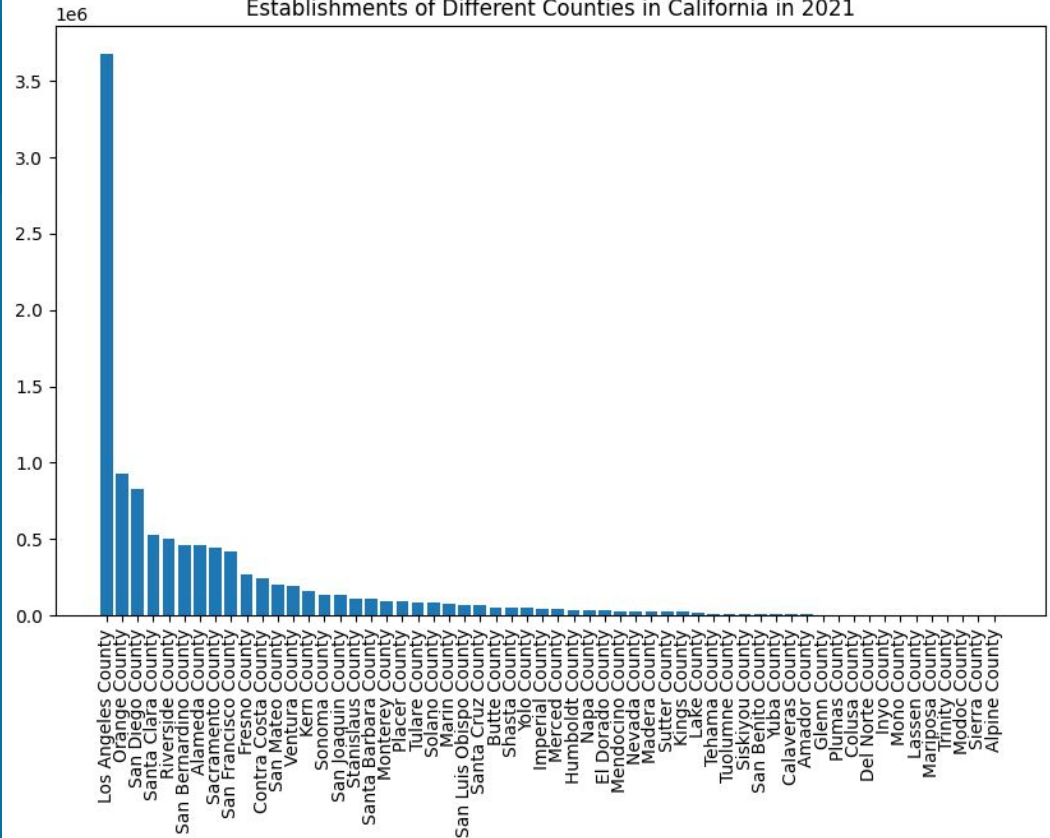
Establishments in California Statewide vs US Nationwide in 2021



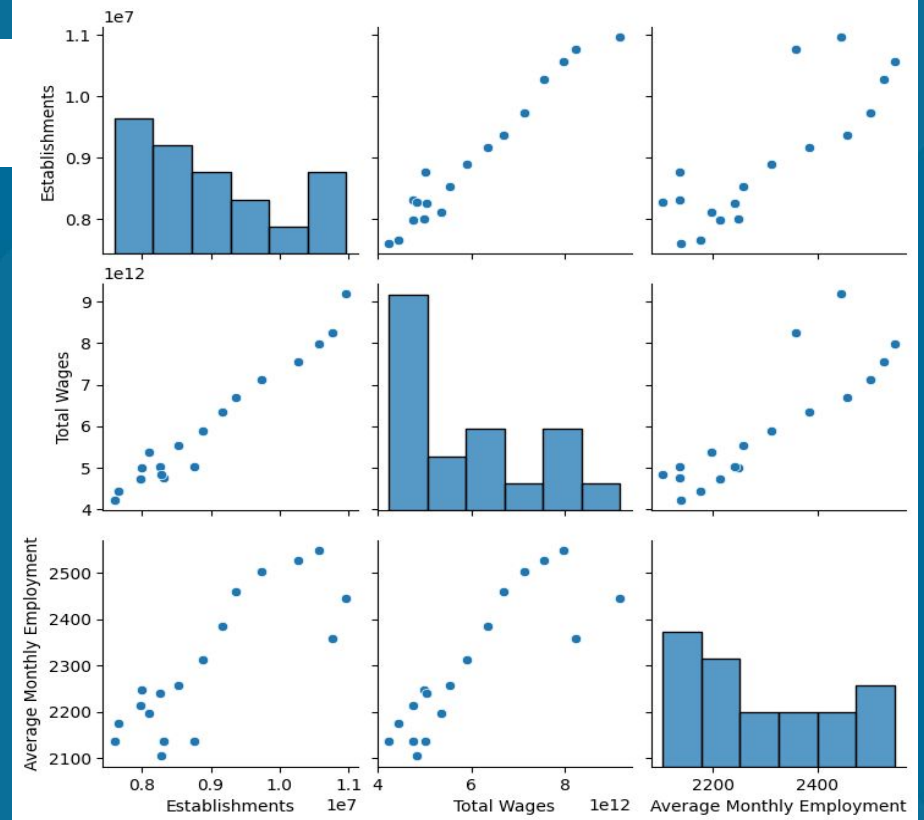
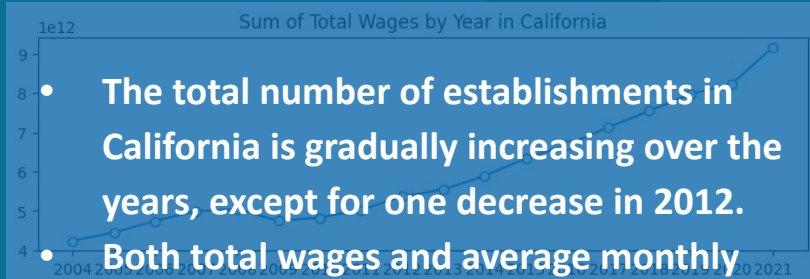
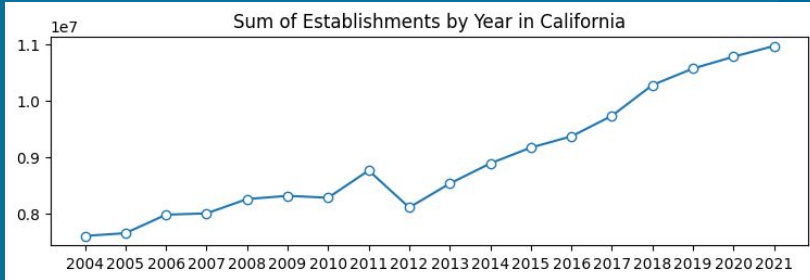
## Focusing on the data in 2021.

- The total number of establishments in California is about 1/7 of the total numbers in the US.
- In California different counties, Los Angeles has the highest number, which is 4 times bigger than the second largest number.

Establishments of Different Counties in California in 2021



# Correlations Analysis



Correlation	Establishments	Total Wages	Avg Monthly Employment
Establishments	1.000000	0.979327	0.829266
Total Wages	/	1.000000	0.859342
Avg Monthly Employment	/	/	1.000000



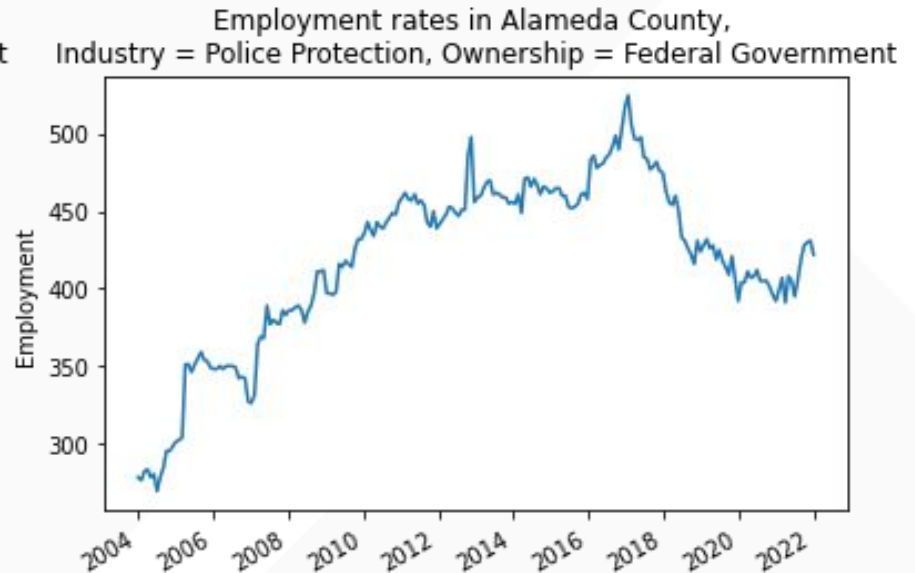
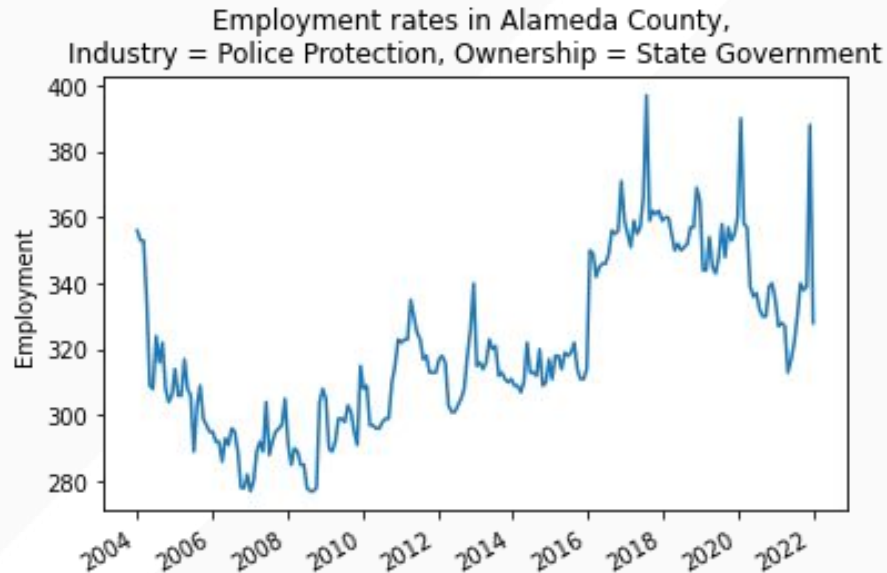
# **Time Series Analysis and prediction for Employment Data**

**Alameda County**

# Time Series Analysis of Employment Data

## Data overview:

- Time series data for employment numbers in each month from 2004 to the end of 2021.
- Data was extracted using the database management tool SQLite.



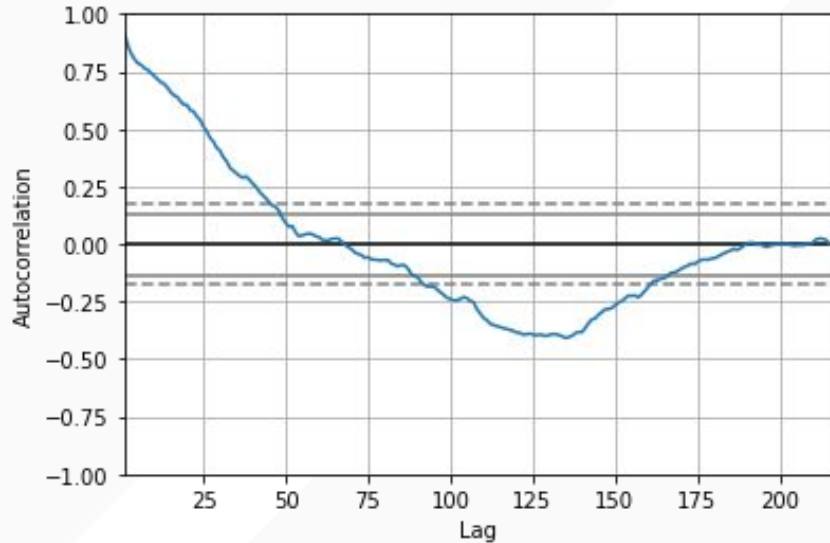
# Time Series Analysis of Employment Data

## Checking Stationarity, Trend and Seasonality using Statistical Tests

- Stationary: Constant Mean and Variance at each time interval.
- Augmented Dickey-Fuller Test:
  - Null Hypothesis:  $H_0$ : The time series is non-stationary.
  - P - value of Data 1: 0.14. ( $>0.05$ )  $\Rightarrow$  Non-Stationary
- Mann Kendall Trend Test:
  - Null Hypothesis:  $H_0$ : The time series has no trend.
  - P - value of Data 1: 0 ( $<0.05$ )  $\Rightarrow$  The series has a trend.
  - Found to have an overall positive trend.

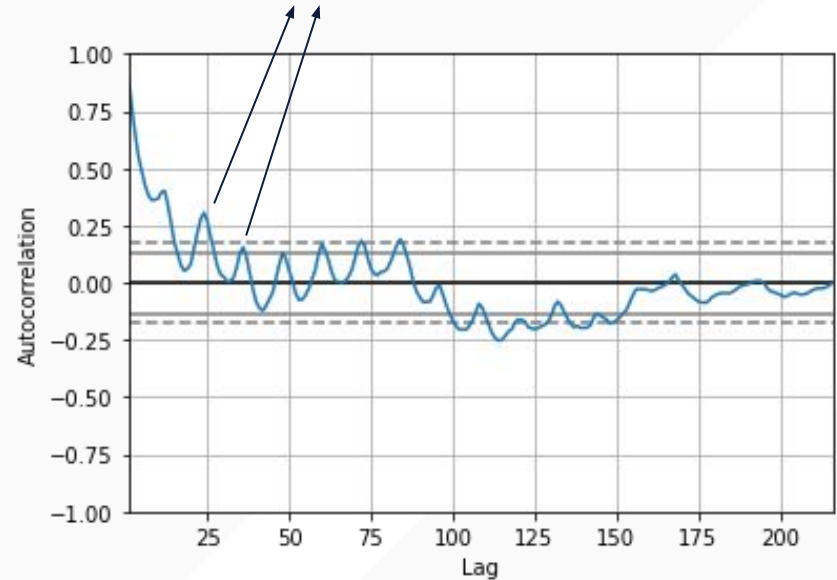
# Time Series Analysis of Employment Data

## Detection of Seasonality using Autocorrelation



Industry: Police Protection

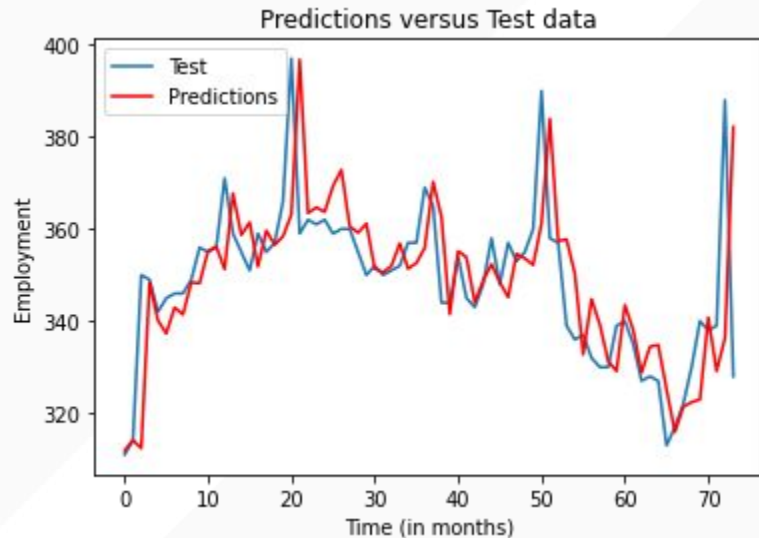
There are peaks in the plot at 12,24,36....



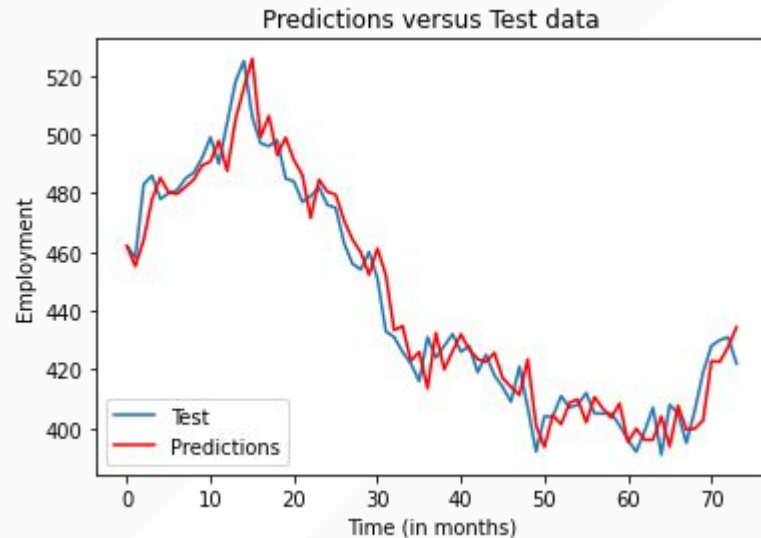
Industry: Nursery, Garden & Farm Supply Stores

# Time Series Prediction of Employment Data using ARIMA

- ARIMA(p,d,q): Autoregressive Integrated Moving Average Model
- P -> Number of lag observations to include =25, found from the autocorrelation plot



Police Protection (State Govt.), RMSE = 14.23



Police Protection (Federal Govt.), RMSE = 8.703

## Conclusion: Key Takeaways

- Major industry types in California are Information and Finance & Insurance
- Health Care and Social Assistance is the largest industry by employment number and Information has the highest average wage in California.
- Since 2009, healthcare and social assistance has had the largest proportion of employment
- The total number of establishments in California is about 1/7 of the total numbers in the US.
- California's employment accounts for over 10% of the United States' Has higher average wages than the entire U.S.

UC San Diego