

# The Impact of COVID-19 on Employment



Yijin Zhu

## **Problem Introduction**

**The COVID-19 pandemic has triggered one of the worse jobs crises since the Great Depression. Huge damage has been done to the economy at the states, national, and occupational level.**

**This presentation will focus on my approach of processing the large-scale datasets related to CA COVID-19 cases and EDD Short-term Occupational Projections, then Visualizing analysis results.**

**Marco Data**

**Computer Data**

## Data Resource

**California -Coronavirus data:**  
<https://github.com/datadesk/california-coronavirus-data>

**California EDD-**  
**Employment Development Department:**  
**Coronavirus 2019**  
[https://edd.ca.gov/en/about\\_edd/coronavirus-2019](https://edd.ca.gov/en/about_edd/coronavirus-2019)

**Short-Term Occupational Employment Projections,**  
**2021 Second Quarter-2023 Second Quarter**  
<https://data.edd.ca.gov/Employment-Projections/Short-Term-Occupational-Employment-Projections/guh4-bakwrter>

**US BUREAU Stats: California**  
Occupational Employment and Wage Statistics  
<https://www.bls.gov/oes/tables.html>

## **Problem Introduction**

**The COVID-19 pandemic has triggered one of the worse jobs crises since the Great Depression. Huge damage has been done to the economy at the states, national, and occupational level.**

**This presentation will focus on my approach of processing the large-scale datasets related to CA COVID-19 cases and EDD Short-term Occupational Projections, then Visualizing analysis results.**

**Marco Data**

**Computer Data**

## **Computer Data**

**How to :**  
**Convert Marco Data to Computer data**

**Step 1. importing.** Checkout large-scale data from data source.

**Step 2. Partition.** Splitting large CSV data into small raw data files for future data processing.

**Step 3. Cleaning.** Remove non-readable or any non-relevant data.

**Step 4. Visualization.** Display data Information.

## **Problem Introduction**

**The COVID-19 pandemic has triggered one of the worse jobs crises since the Great Depression. Huge damage has been done to the economy at the states, national, and occupational level.**

**This presentation will focus on my approach of processing the large-scale datasets related to CA COVID-19 cases and EDD Short-term Occupational Projections, then Visualizing analysis results.**

**Marco Data**

**Computer Data**

# The Impact of COVID-19 on Employment



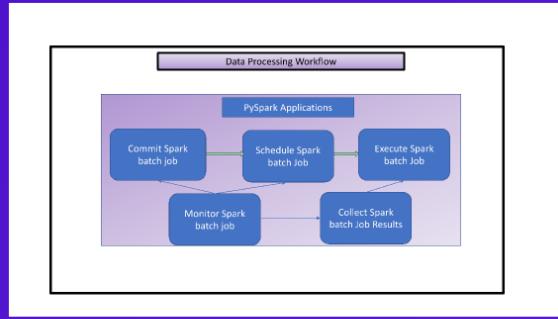
Yijin Zhu



# Research Development Proposal

## Major Components:

1. Data Collection.
2. Job Scheduling.
3. Data Cleaning and Data Visualization.



Data Collection

Job Scheduling

Data Cleaning  
And Data  
Visualization

## Data Collection Workflow Using NoSQL Big Data

**Data Collection Workflow** replies on Spark near-realtime data analysis.

### 1. Why NoSQL Database?



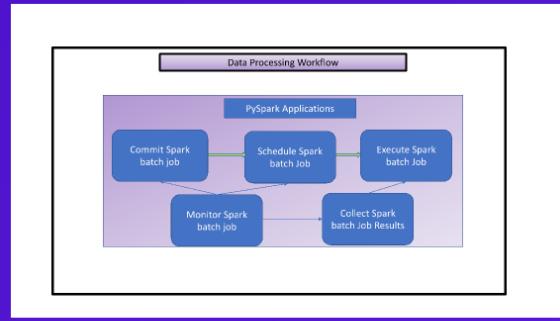
### 2. Why Spark?



# Research Development Proposal

## Major Components:

1. Data Collection.
2. Job Scheduling.
3. Data Cleaning and Data Visualization.



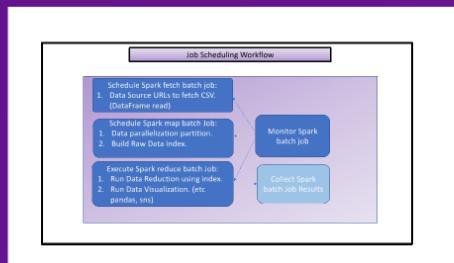
Data Collection

Job Scheduling

Data Cleaning  
And Data  
Visualization

## Job Scheduling Workflow

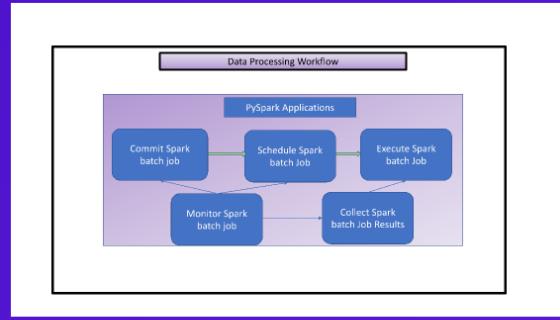
1. High level workflow in each level job scheduling.
2. Monitoring Watchdogs collector system health index to sync up communication.
3. Data collection module is to render final visual graphs.



# Research Development Proposal

## Major Components:

1. Data Collection.
2. Job Scheduling.
3. Data Cleaning and Data Visualization.



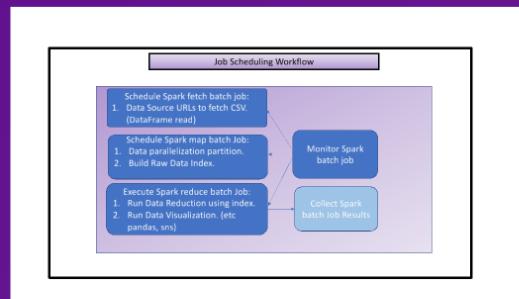
Data Collection

Job Scheduling

Data Cleaning  
And Data  
Visualization

## Data Cleaning and Data Visualization

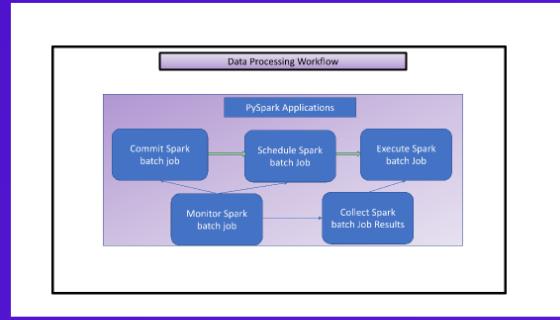
1. Data Cleaning work is being conducted in data parallelization partition.
2. Data Visualization work is being conducted in batch Job results collection module.



# Research Development Proposal

## Major Components:

1. Data Collection.
2. Job Scheduling.
3. Data Cleaning and Data Visualization.



Data Collection

Job Scheduling

Data Cleaning  
And Data  
Visualization

# The Impact of COVID-19 on Employment



Yijin Zhu



## Data Mapping

- 1. What data models can we use to present COVID-19 and CA Occupational Projections?**
- 2. What data correlations can we find between COVID-19 and CA Occupational Projections?**

Data Models

Data Correlations

## Data Models

**Network Model:**  
similar to hierarchical model, allowing one-to-many relationships using graph representation.

**Relational Model:**  
collection of predicates over finite set of predicate variables defined with constraints on the possible values and combination of values.

**Flat Model:**  
single array, or two-dimensional array of data elements.

Reference:  
<https://www.talend.com/resources/data-model-design-best-practices-part-1/>

## Data Mapping

---

- 1. What data models can we use to present COVID-19 and CA Occupational Projections?**
- 2. What data correlations can we find between COVID-19 and CA Occupational Projections?**

Data Models

Data Correlations

## Data Correlations

This will show what potential correlation between COVID-19 and CA Employment Data.

**COVID-19 test positive rate VS job openings.**

**COVID-19 test positive rate VS annual salary.**

**COVID-19 test positive rate VS job category.**

**COVID-19 test positive rate VS job requirement(eg. education level, experience).**

.....

## Data Mapping

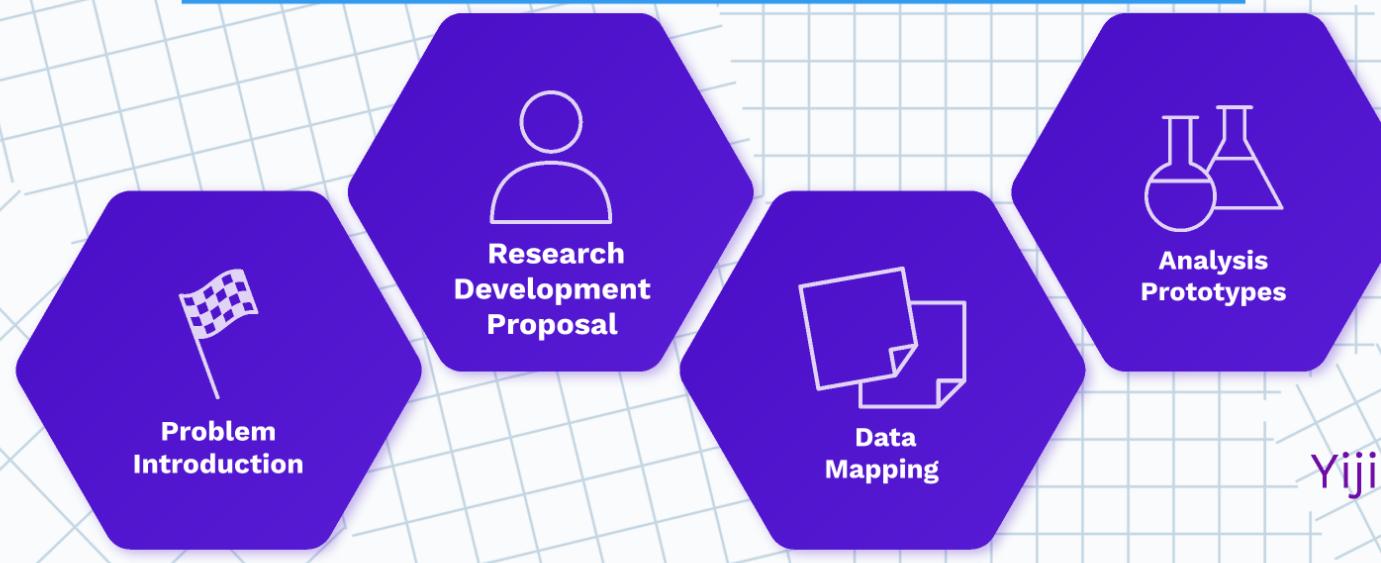
---

- 1. What data models can we use to present COVID-19 and CA Occupational Projections?**
- 2. What data correlations can we find between COVID-19 and CA Occupational Projections?**

Data Models

Data Correlations

# The Impact of COVID-19 on Employment

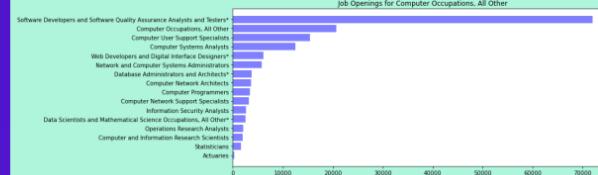
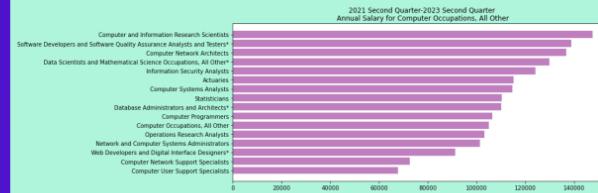


Yijin Zhu



# Data Analysis Prototypes

Data generated from my personal github repository:  
<https://github.com/judyma2294/informatics>



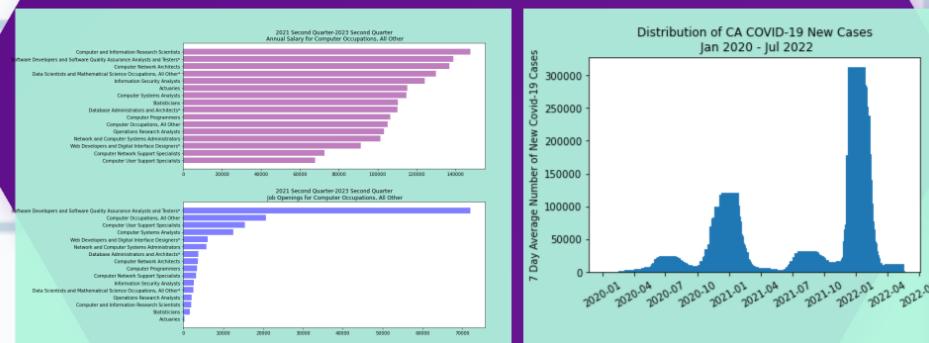
Data Science related work is promising!

Prototypes

Conclusion

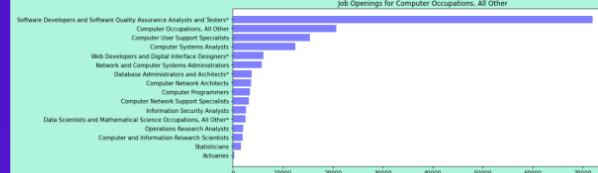
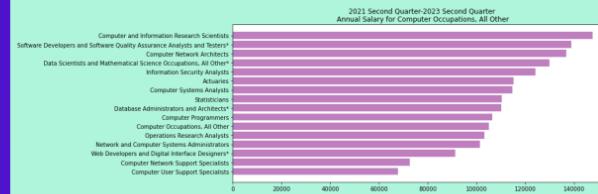
# Data Visualization Prototypes

Rendering DataFrame object using pandas,  
sns and matplotlib.pyplot.



# Data Analysis Prototypes

Data generated from my personal github repository:  
<https://github.com/judyma2294/informatics>



Data Science related work is promising!

Prototypes

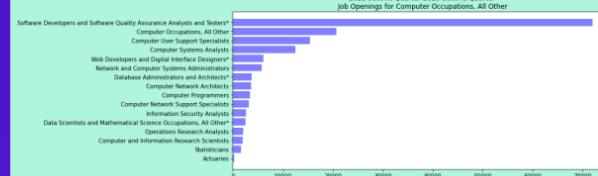
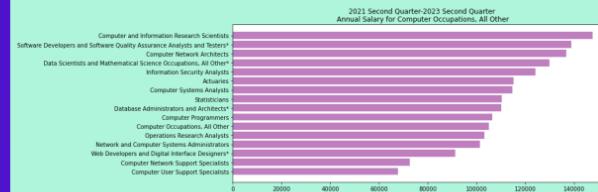
Conclusion

## Possible Conclusion

1. From year 2020 - 2021  
COVID-19 Test Positive Rate has significant impacts on job titles, which requires less working experiences or education degrees requirement.
  2. From year 2020 - 2021  
Data Science related annual salary and job openings both have been increasing.
- ...

# Data Analysis Prototypes

Data generated from my personal github repository:  
<https://github.com/judyma2294/informatics>



Data Science related work is promising!

Prototypes

Conclusion

# The Impact of COVID-19 on Employment



Yijin Zhu





As a future informatics professional, I believe I can contribute toward our society, by interpreting data into human language, which may help scientists and policymakers to identify patterns and find solutions to the many challenges raised by this pandemic.

**Thank you!**

## **References:**

1. The impact of COVID-19 on employment and jobs  
<https://www.oecd.org/employment/covid-19.htm>
2. The role of data in a world reshaped by Covid-19  
<https://penntoday.upenn.edu/news/role-data-world-reshaped-covid-19>
3. The COVID-19 Data Plan: 3 Innovative Ways Johnson & Johnson Is Using Data Science to Fight the Pandemic  
<https://www.jnj.com/innovation/how-johnson-johnson-uses-data-science-to-fight-covid-19-pandemic>

# The Impact of COVID-19 on Employment



Yijin Zhu

