Layoff Prevention

W205 Final Project Presentation

Aug. 10, 2016

By **by Madison Myers, Arvin Sahni and Shih Yu Chang**

Outlines

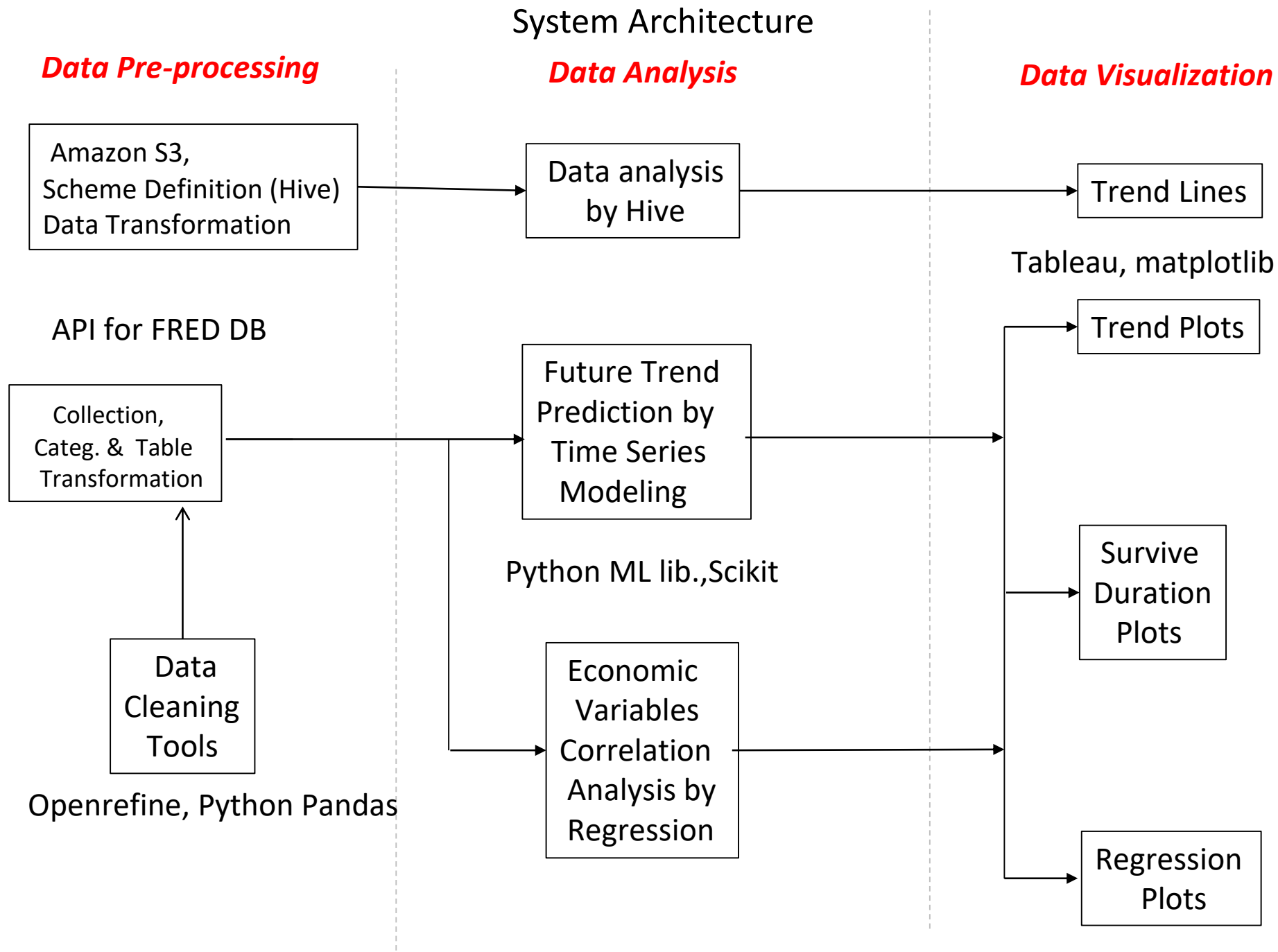
- Motivations and problems
- Data Sources
- System Architecture
- Data Acquisition API
- Prediction Model: ARIMA and Exponential Smooth
- The Effect of Minimum Wage to Unemployment
- Limitations and future extensions
- Conclusions

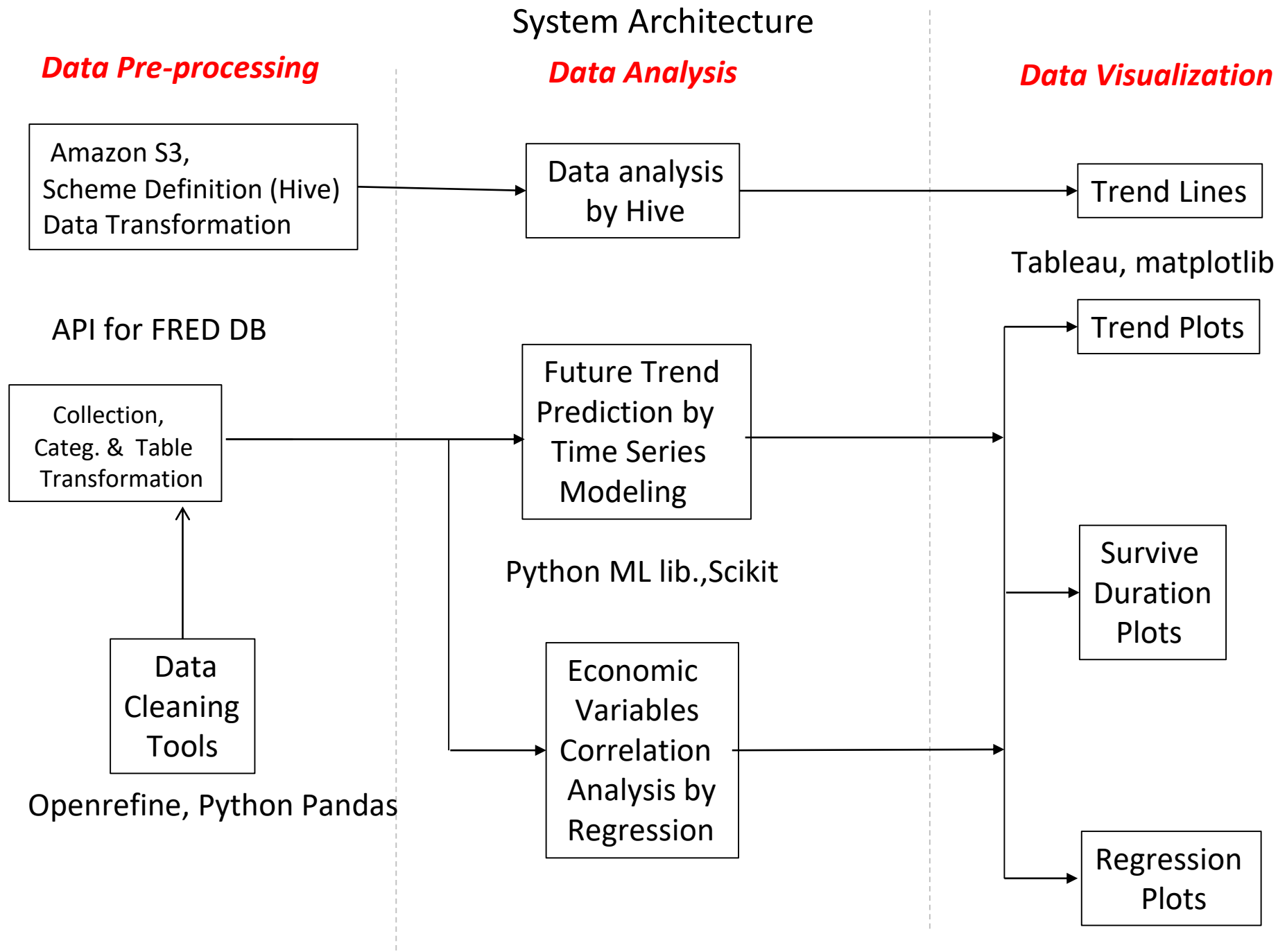
Motivations and problems

- Economic crisis makes many people got layoff.
- We wish to apply data science to help layoff people.
- How to estimation survival duration? By considering
  - Locations
  - Industries
  - Individual financial status

Data Sources

- Ministry of Manpower
- Department of Statistics Singapore
- FRED: http://research.stlouisfed.org/fred2/
  - A rich database of U.S. and international economic data maintained by the research department of the Federal Reserve Bank of St. Louis.
  - There are 390,000 time-series contributed by 79 various data sources, the data size is about 39 GB.
  - Moreover, this database will be updated day-by-day.

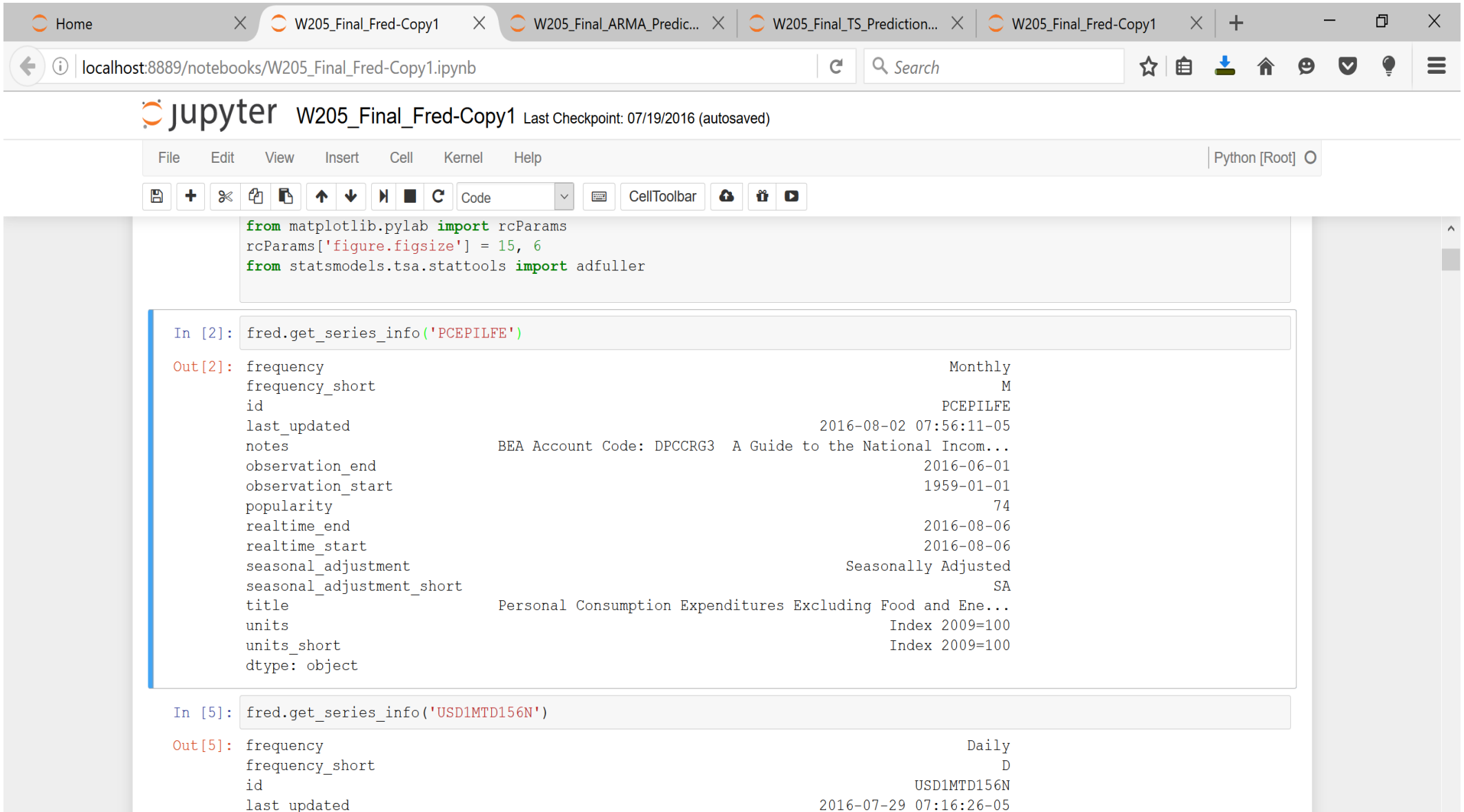

System Architecture
Data Pre-processing
Data Analysis
Data Visualization
Amazon S3,
Scheme Definition (Hive)
Data Transformation
Data analysis by Hive
Trend Lines
Tableau, matplotlib
API for FRED DB
Trend Plots
Future Trend Prediction by Time Series Modeling
Collection, Categ. & Table Transformation
Python ML lib.,Scikit
Survive Duration Plots
Data Cleaning Tools
Economic Variables Correlation Analysis by Regression
Openrefine, Python Pandas
Regression Plots


System Architecture
Data Pre-processing
Data Analysis
Data Visualization
Amazon S3,
Scheme Definition (Hive)
Data Transformation
Data analysis by Hive
Trend Lines
Tableau, matplotlib
API for FRED DB
Trend Plots
Future Trend Prediction by Time Series Modeling
Collection, Categ. & Table Transformation
Python ML lib.,Scikit
Survive Duration Plots
Data Cleaning Tools
Economic Variables Correlation Analysis by Regression
Openrefine, Python Pandas
Regression Plots

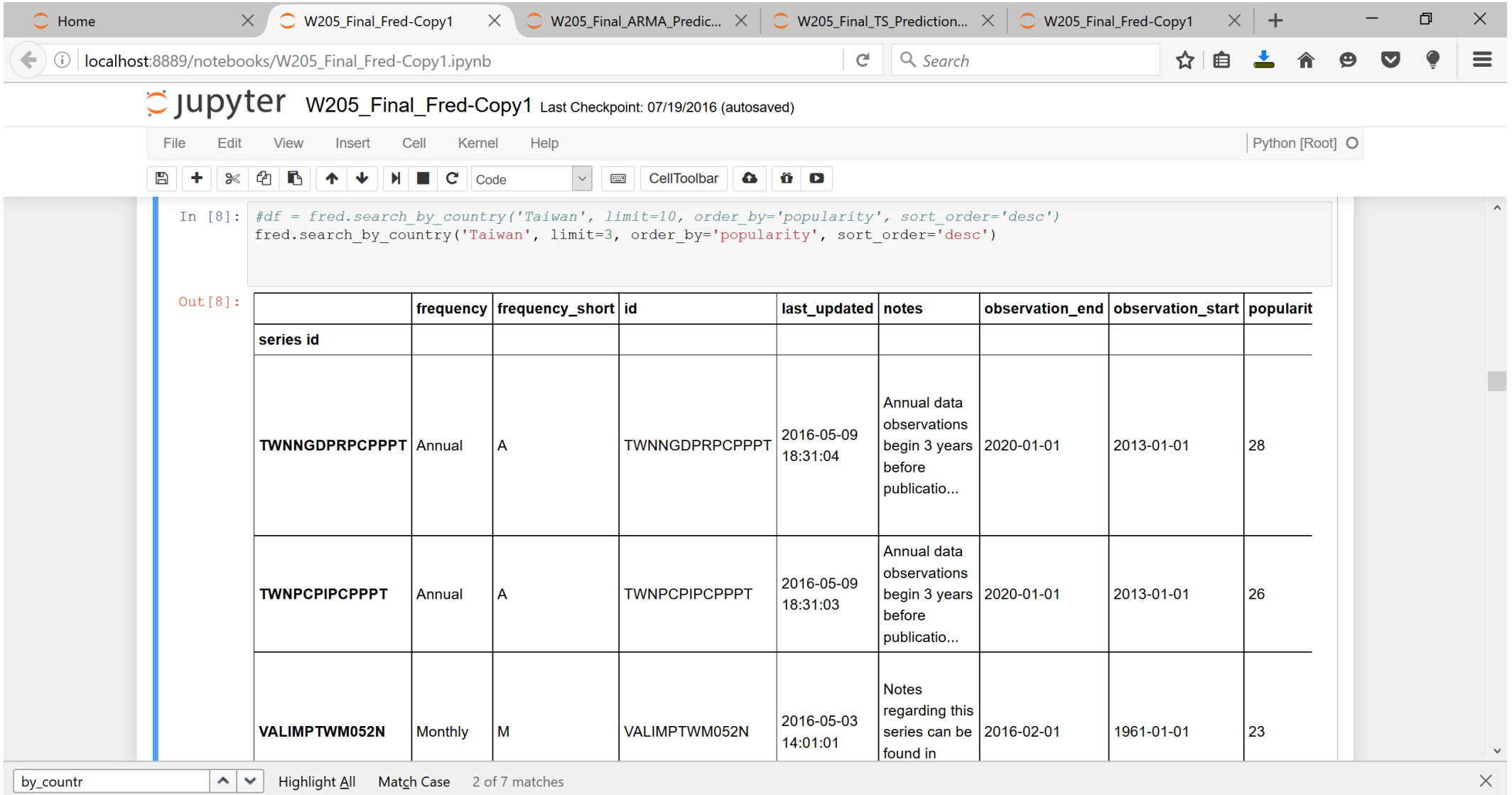
Data Acquisition API

- The Economic Research Division of the Federal Reserve Bank of St. Louis has created APIs (application programming interfaces) to allow developers to write programs and build applications that interact with content on its websites.
- Users may query our Federal Reserve Economic Data (FRED) databases to retrieve the specific data desired (according to source, release, category and series among other preferences).

FRED API Demo: Last Updated

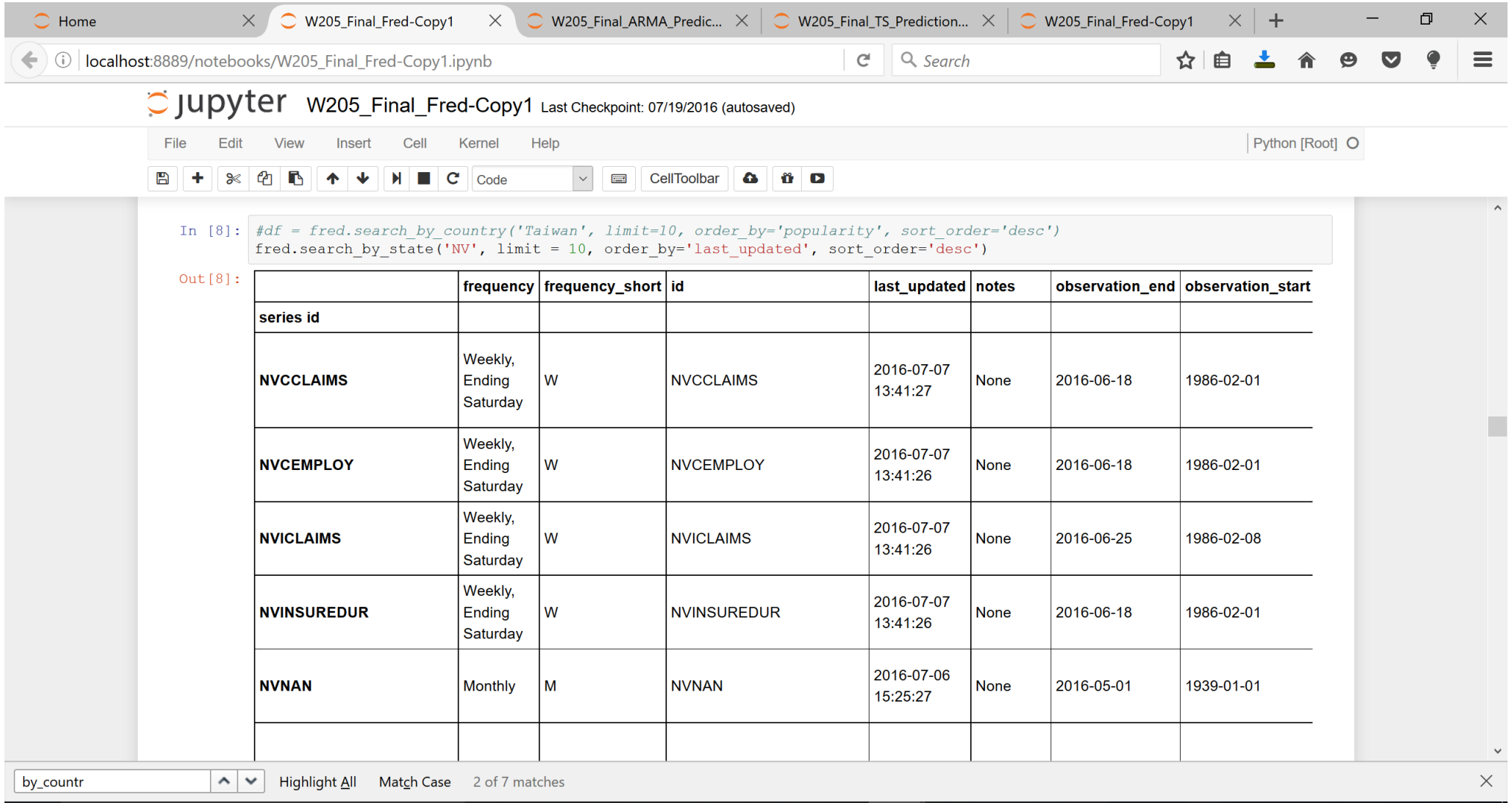


FRED API Demo: Search by Country



	frequency	frequency_short	id	last_updated	notes	observation_end	observation_start	popularit
series id								
TWNNGDPRPCPPPT	Annual	A	TWNNGDPRPCPPPT	2016-05-09 18:31:04	Annual data observations begin 3 years before publicatio...	2020-01-01	2013-01-01	28
TWNPCPIPCPPPT	Annual	A	TWNPCPIPCPPPT	2016-05-09 18:31:03	Annual data observations begin 3 years before publicatio...	2020-01-01	2013-01-01	26
VALIMPTWM052N	Monthly	M	VALIMPTWM052N	2016-05-03 14:01:01	Notes regarding this series can be found in	2016-02-01	1961-01-01	23

FRED API Demo: Search by State



series id	frequency	frequency_short	id	last_updated	notes	observation_end	observation_start
NVCCLAIMS	Weekly, Ending Saturday	W	NVCCLAIMS	2016-07-07 13:41:27	None	2016-06-18	1986-02-01
NVCEMPLOY	Weekly, Ending Saturday	W	NVCEMPLOY	2016-07-07 13:41:26	None	2016-06-18	1986-02-01
NVICLAIMS	Weekly, Ending Saturday	W	NVICLAIMS	2016-07-07 13:41:26	None	2016-06-25	1986-02-08
NVINSUREDUR	Weekly, Ending Saturday	W	NVINSUREDUR	2016-07-07 13:41:26	None	2016-06-18	1986-02-01
NVNAN	Monthly	M	NVNAN	2016-07-06 15:25:27	None	2016-05-01	1939-01-01

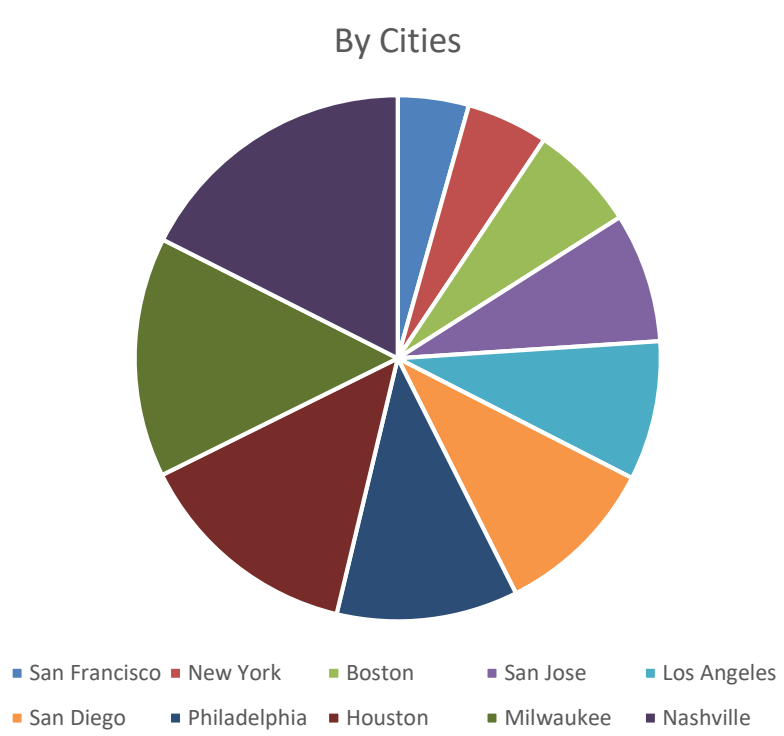
Prediction Model by ARIMA

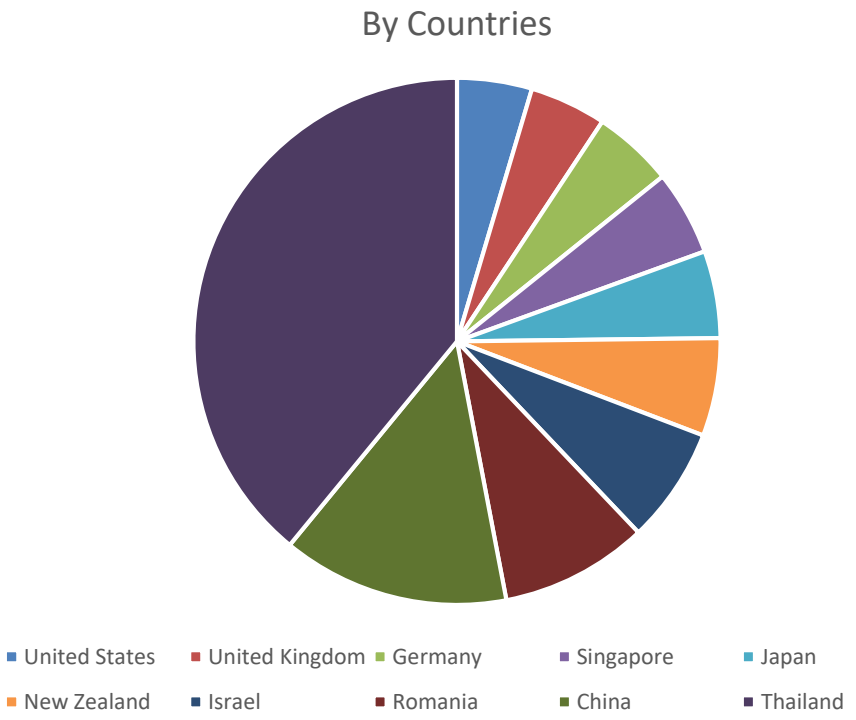
- In time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model.
- By applying ARIMA prediction mechanism, we can get range of survival time for each countries and states based on each individual all assets values at lay off time, say X.
- Then, survival time duration range can be determined as MIN: X / (MAX living cost in prediction), and MAX: X/(MIN living cost in prediction).

Prediction Model by ES

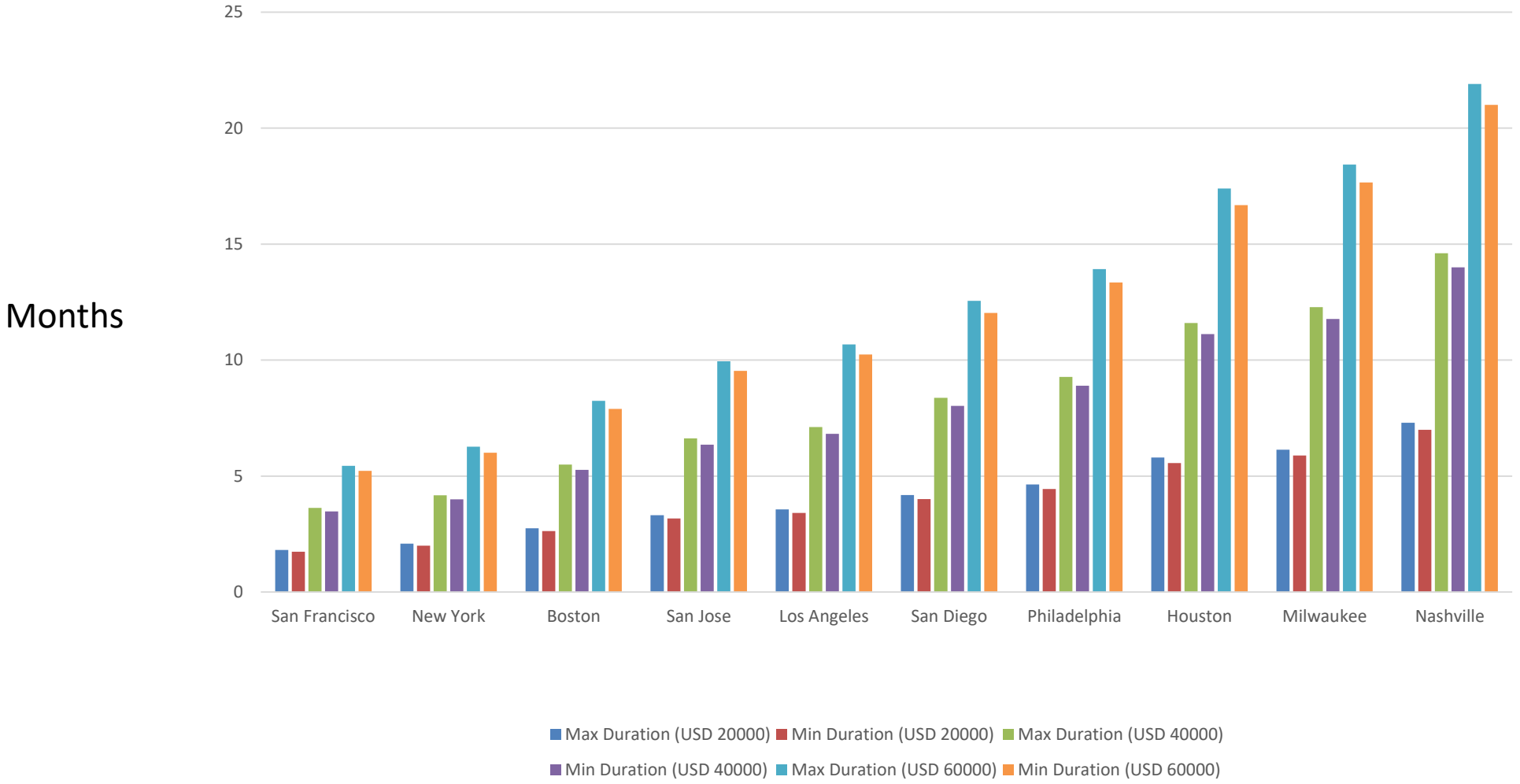
- Another prediction method used in time series analysis is Exponencial Smoothing (ES).
- Exponential Smoothing assigns exponentially decreasing weights as the observation get older. In other words, recent observations are given relatively more weight in forecasting than the older observations.

Living Cost Ranking

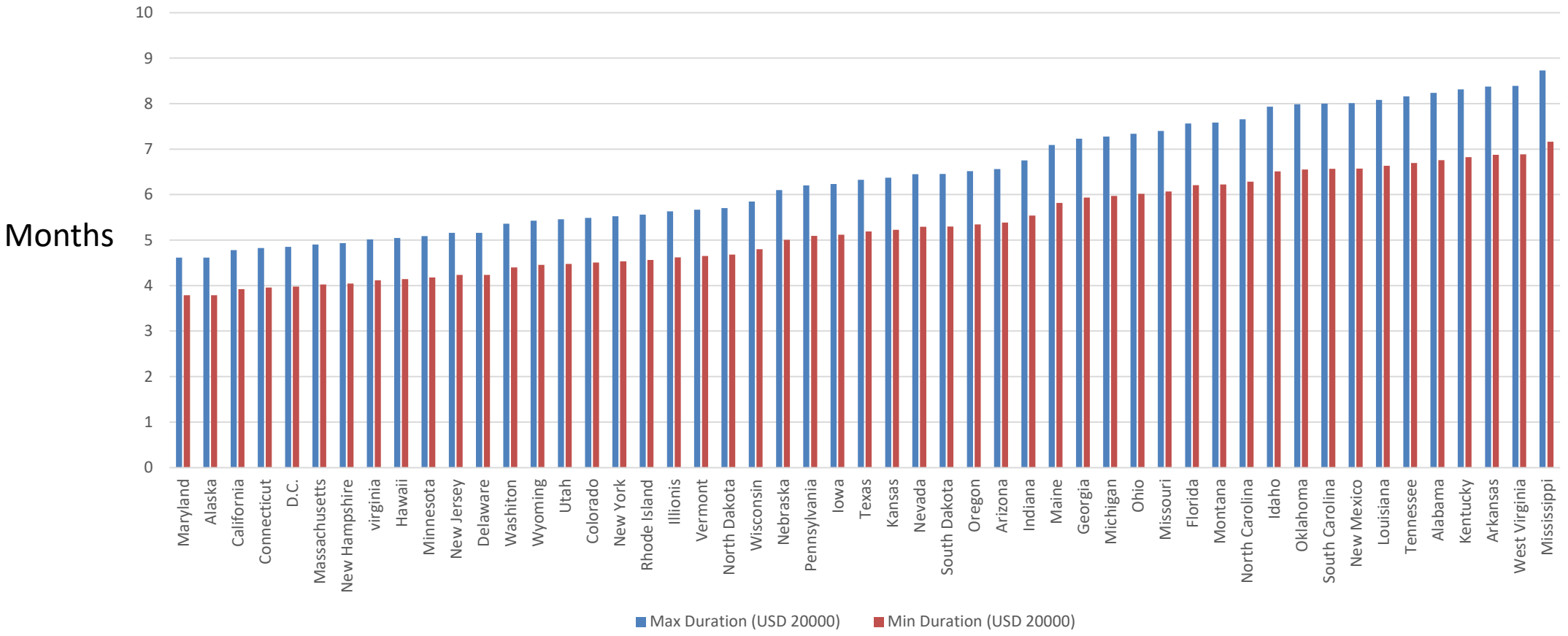




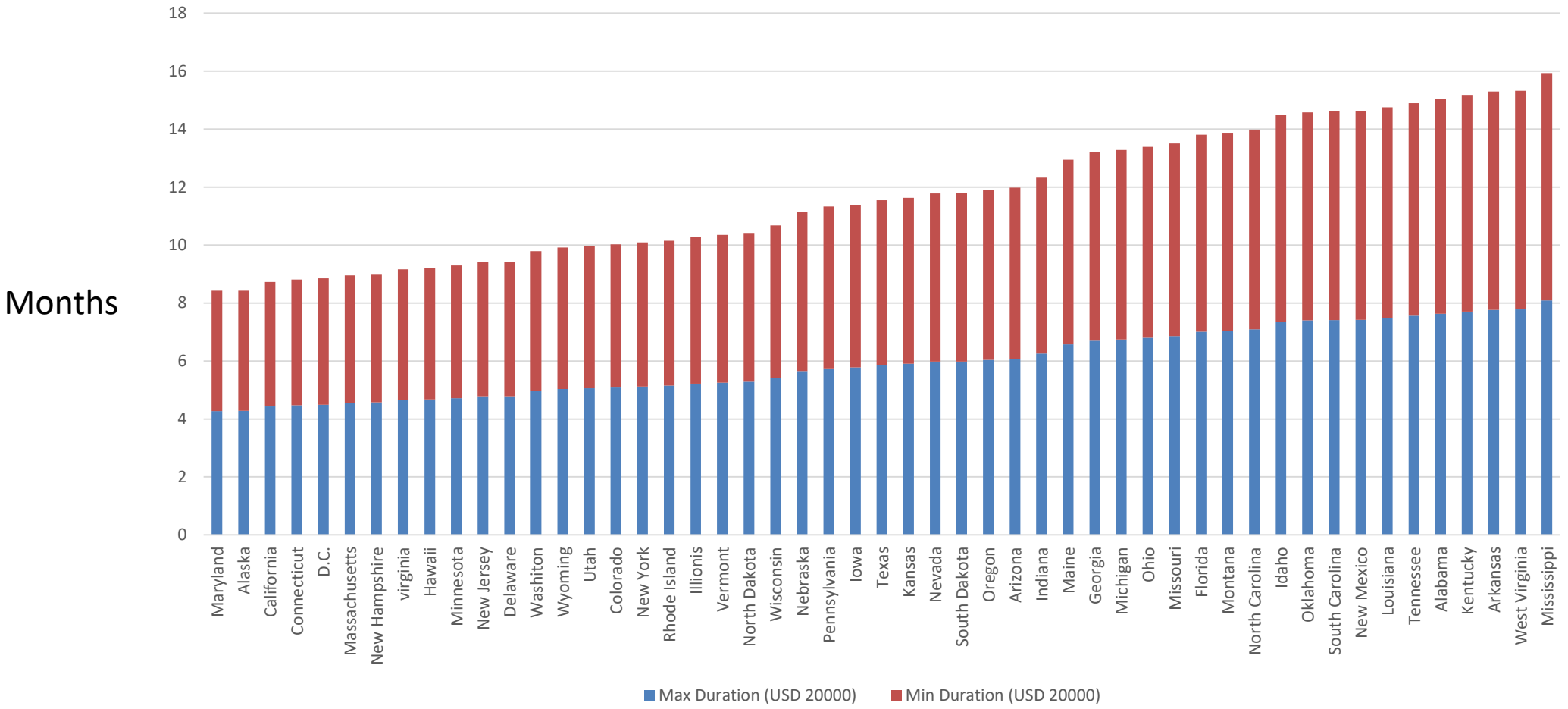
Survival Months by Cities Prediction (Smooth)



Survival Months by States Prediction (ARIMA)

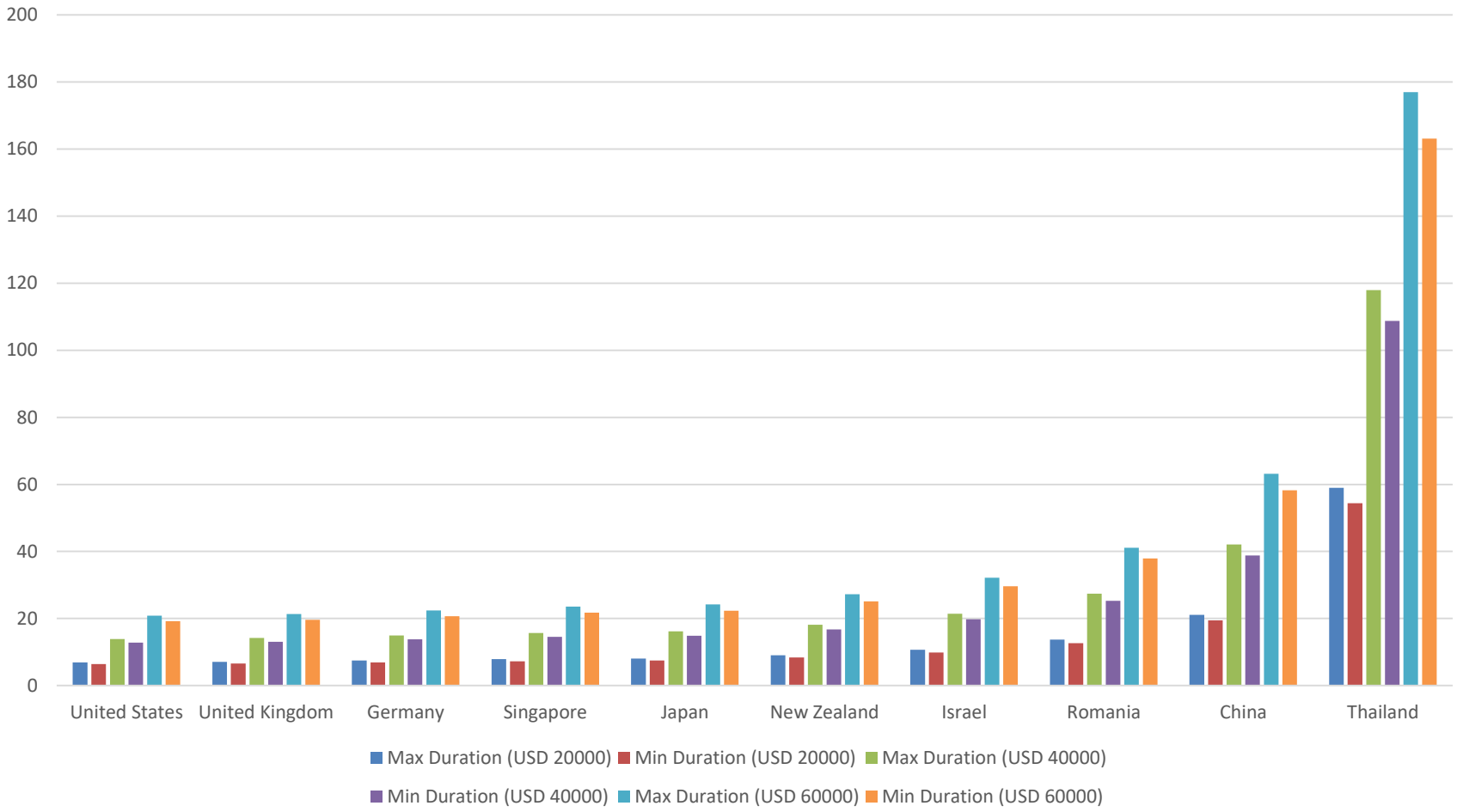


Survival Months by States Prediction (Smooth)



Survival Months by Countries Prediction (ARIMA)

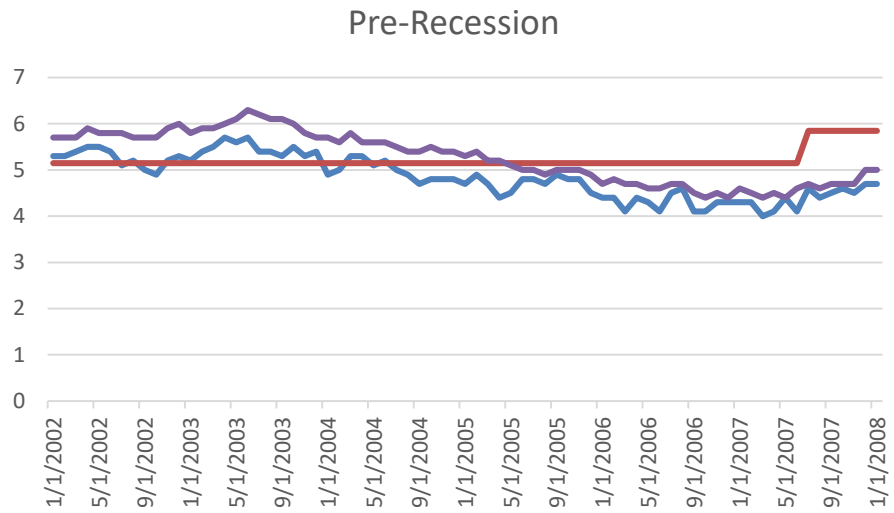
Months

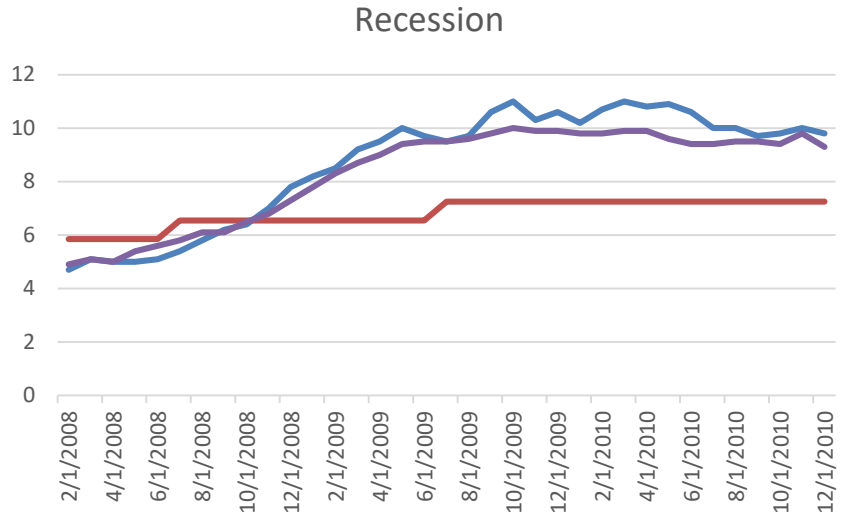


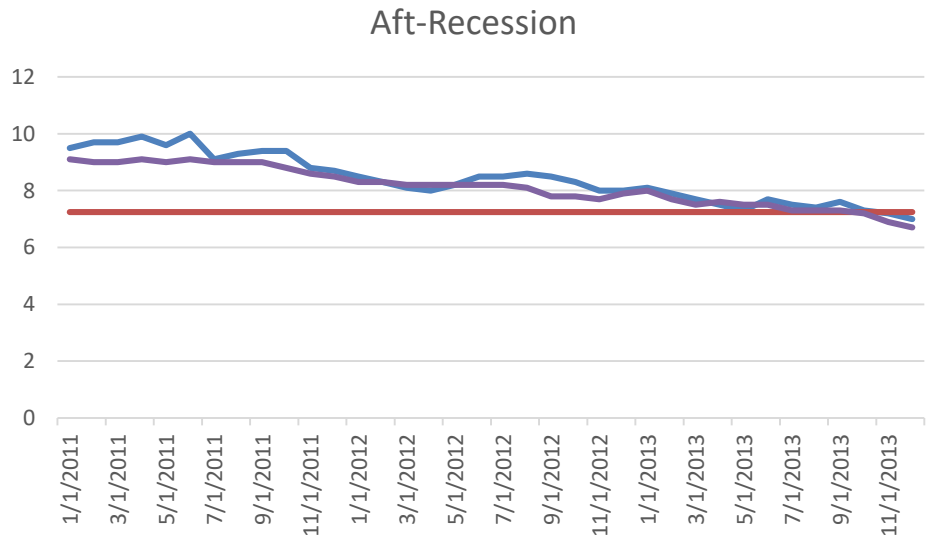
The Effect of Minimum Wage to Unemployment

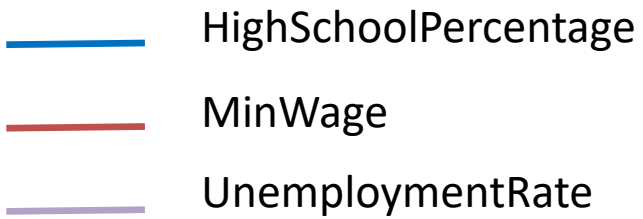
- This task aims to model and quantify the relationship between minimum wage and unemployment rate.
- For the multiple regression models, it was concluded that minimum wage had a significant effect on unemployment when the economy was during the recession.
- We also observed that higher percentage of high-school educated people in USA will have higher unemployment rate. This quite different from general sense.

Trends of Education Leve, Min_Wage and Unemployment









Perform Multi-Regression Analysis by R

Ind. Variables	Before-Recession	Recession	After-Recession
Min_Wage	NA	0.164	NA
Hgih_School_Percentage	0.647 ***	0.815 ***	0.442 ***
G_D_P	~0 ***	~0	~0 ***
Intercept	4.857 ***	2.496	13.872 ***
R-square	0.908	0.992	0.956

Limitations and future extensions

- Limitations:
  - FRED data source focus more at US only, not the whole economic world.
  - Other useful data sources, e.g., labor department, do not provide API to get data.
- Possible Extensions:
  - Other prediction models.
  - Preparing API to have more features.

Conclusions

- Developing API to get time-series from FRED website up-to-date.
  - Get time-series data from geographical query inputs.
- Estimating survival duration of layoff by ARIMA and ES model based on individual financial status.
- Discovering fact that minimum wage has a significant effect on unemployment when the economy condition is at regression by multiple regression analysis.