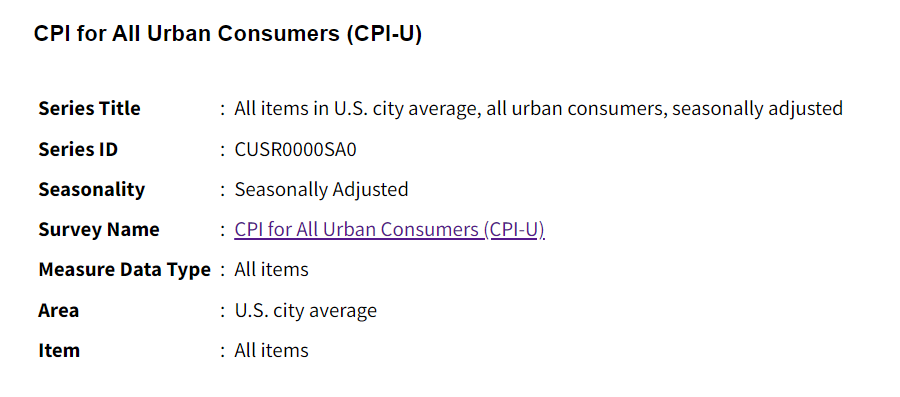
**Consumer Price Index for All Urban Consumers (CPI-U)**

*\*data is seasonally adjusted*

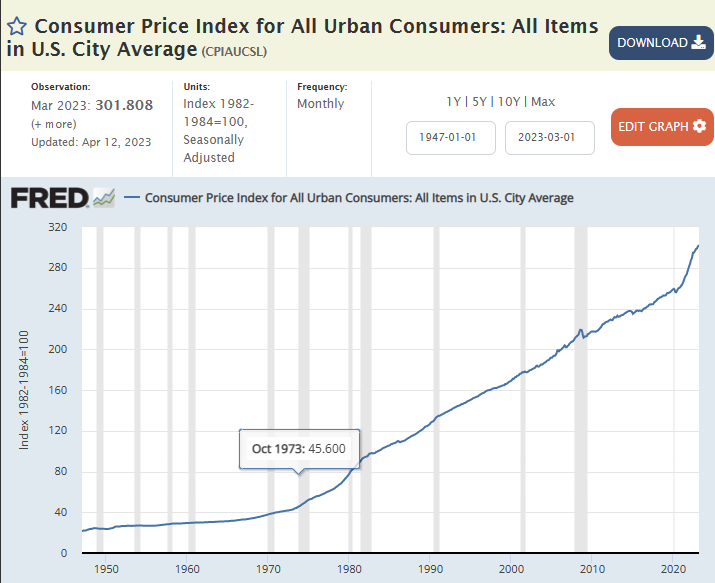
**

# Objectives

Our a plan is to create a time series analysis to predict CPI values (not rate of change). To be specific, since a prediction model normally can predict one label, predicting CPI values allows us to calculate both monthly and yearly inflation.

# CPI Information

* According to U.S. Bureau of Labor Statistics, CPI is ameasure of the average change over time in the prices paid by urban consumer for a market basket of consumer goods and services.
* There are two common versions of the CPI, which are the CSP for all urban consumer (CPI-U) and the CPI for urban wage earners and clerical workers (CPI-W). The CPI-U is designed to measure price changes faced by urban consumers, who represents 93% of the U.S. population. This is also the direct figure used to measure inflation. The CPI-W is used to adjust Social Security benefits. ([Source](https://www.bls.gov/cpi/overview.htm))
* The rate is calculate by determining the current value of a particular basket of goods and services, then dividing it by the value of those same goods and services from a year or month prior. The result is then multiplied by 100.
* Sources of data: Prices of goods and services used to calculate the CPI are collected in 75 urban areas throughout the country and from about 23,000 retail and service establishments. The weight for an item is derived from reported expenditures on that item as estimated by the Consumer Expenditure Survey.
* Without context or comparison to other periods, the CPI value alone does not provide information about the rate of inflation, the direction of price changes, or the trends in price movements. To gain meaningful insights about inflation or price trends, you need to compare CPI values across different time periods, such as month-to-month or year to year.



* Uses
* CPI can be used as an economic indicator. It is the most widely used measure of inflation and deflation. Changes in this figure affect nearly every market and person in America. Therefore, it is an indicator of the effectiveness of government policy.
* Business executives, labor leaders and other private citizens use the index as a guide in making economic decisions.
* …

# Outlined thoughts:

## Preprocessing

* Let’s gather CPI data from Fred first. I can retrieve it quickly from R. The data would be monthly.
* Please find other relevant economic indicators and variables that could influence the CPI: unemployment rate, Change in 1 year treasury bonds, Change in 10 years treasury bonds, gap between these two, etc (We need to find more here)
* Clean

## Explore the data

* CPI: Analyze trends, seasonality, patterns
* Relationship potential predictors

## Feature Engineering

* Explore lags, moving average.
* Autocorrelation plot, Partial Autocorrelation

## Model Selection

* ARIMA, SARIMA, XGBoost, **Gets**
* Experimenting Long Short-Term Memory (LSTM), or hybrid model that combine traditional time series models with machine learning approaches
* Baseline model for comparison: Naive, random walk?

## Training and Validation

* Find a reasonable cut-off point. The covid time may severely affect the model performance. We may need to cut them into 3 different periods: train - validation - test
* Metrics used for validation: RMSE, MSE, MAE.
* For each model, make sure we satisfy the model's assumption by checking residual plots - autocorrelation, partial autocorrelation.

# A few question to ask Professor:

* Should I normalize the all variables, both label, lags, and other predictors?
* What would be the best model I should use as baseline? Naive?

# Data Dictionary

CPI

* Units: Growth rate previous period,Not Seasonally Adjusted
* Frequency: Monthly
* Time range:
* <https://fred.stlouisfed.org/series/CPALTT01USM657N>
* <https://beta.bls.gov/dataViewer/view/timeseries/SUUR0000SA0>

PCE

* Units: Percent Change,Seasonally Adjusted Annual Rate
* Frequency: Monthly
* Time range:
* <https://fred.stlouisfed.org/series/PCE>

Unemployment Rate

* Units: Percent Change,Seasonally Adjusted Annual Rate
* Frequency: Monthly
* Time range:
* https://fred.stlouisfed.org/series/UNRATE

…