**Python Project 3**

**We will do two things:**

* **Track the changes in house prices**
* **Examine the impact of interest rate on mortgage origination. Usually as interest rates goes up, housing market will cool down a bit.**
* We intend to use mortgage data from 2018 to Q1 2021 to examine how mortgage origination is impacted by interest rate.
* 30-year fixed interest rate will be used to represent market rate.

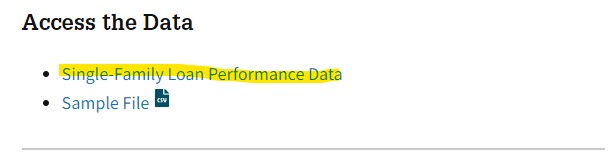
**1) Data Source 1 – mortgage data**

Take the following steps to extract the data (very similar to Project 2):

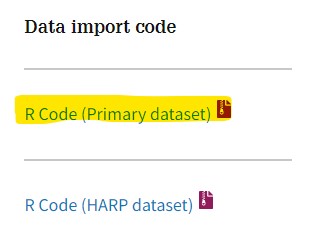
Step 1: Copy the following to Google and search

<https://capitalmarkets.fanniemae.com/credit-risk-transfer/single-family-credit-risk-transfer/fannie-mae-single-family-loan-performance-data>

Step 2: After opening the above web page, click on the following link highlighted in yellow:



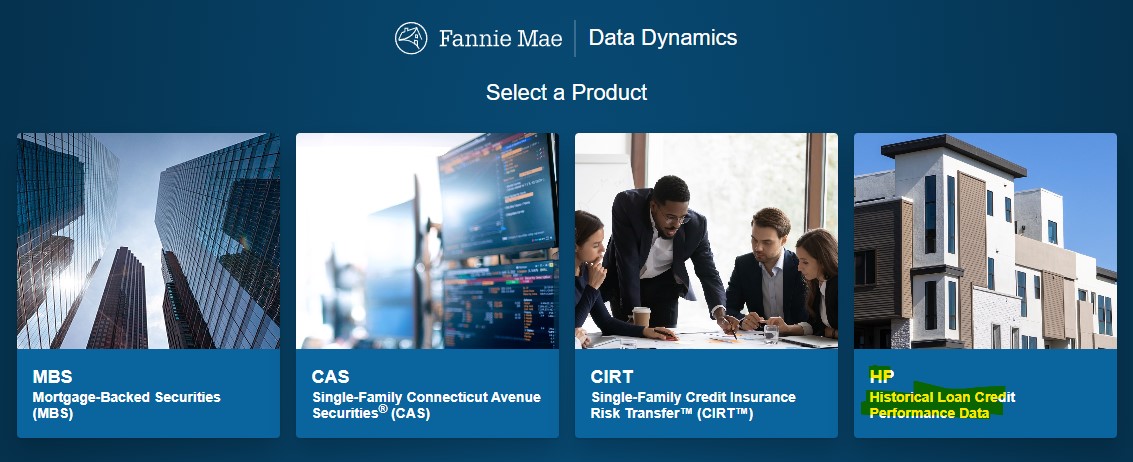
On the same page, download the following:



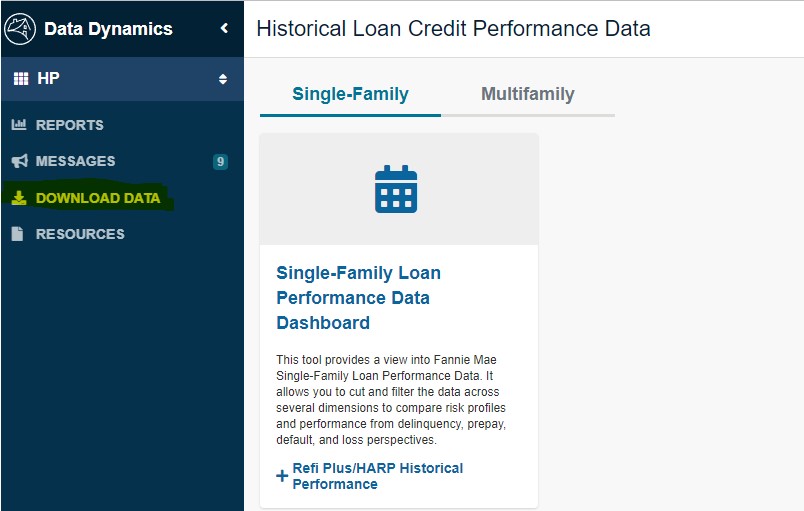
Open the R code in notepad, and you will get the list of column names for the data we are going to download.

Step 3: Register for a free account for login by using your email.

Step 4: After you have logged in, click on the icon on the very right (in yellow highlight):



Step 5: On the top left, click on DOWNLOAD DATA:



Step 6: Download all files for 2018 to 2022.

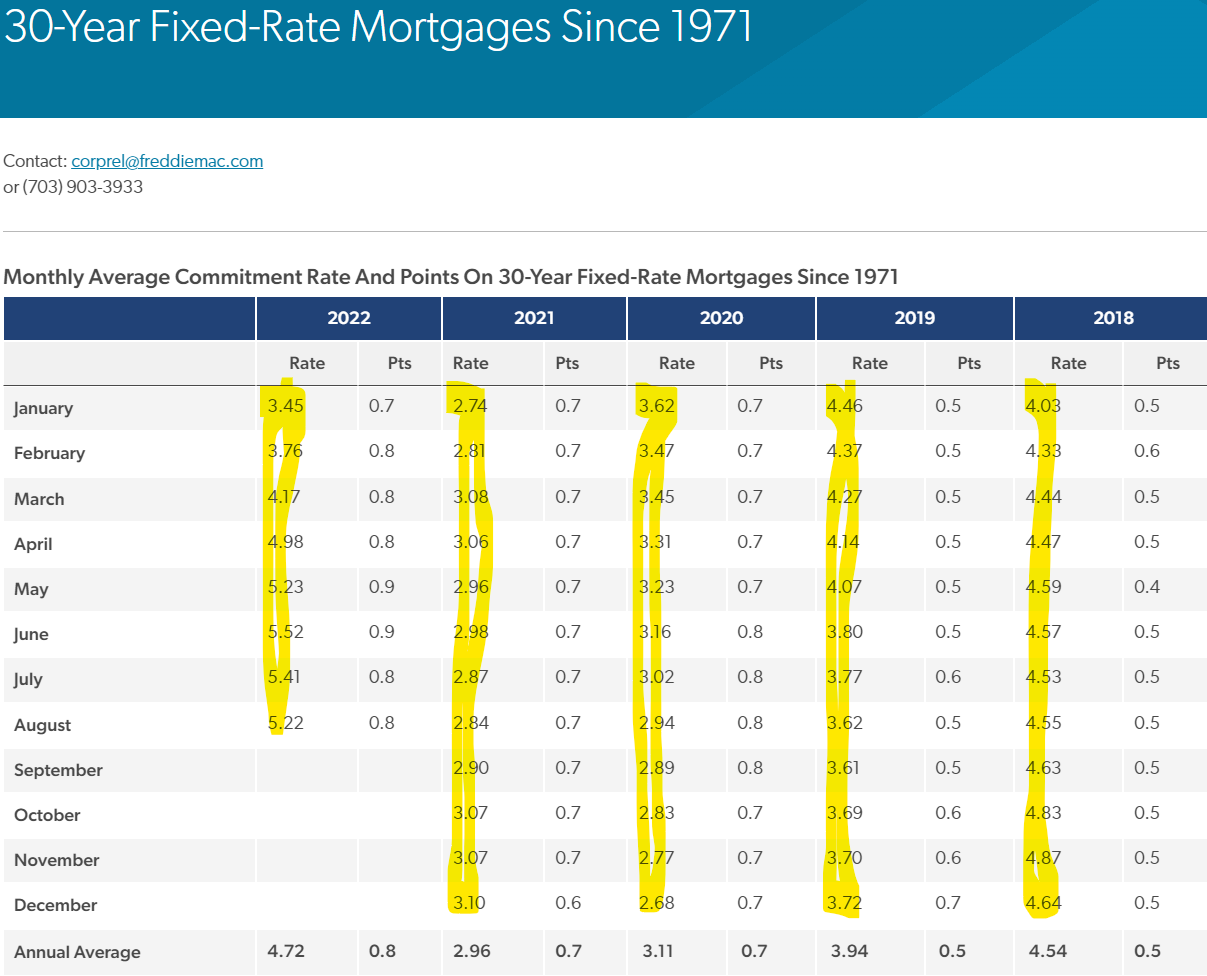


**2) Data Source 2 – 30-year fixed mortgage rates**

step 1: Click on the following page:

<https://www.freddiemac.com/pmms/pmms30>

Step 2: Copy the column for rate from the table below to Excel (no need to copy the last row for Annual Average):



Step 3: Save the data for 2018 to 2021 to Excel as follows:



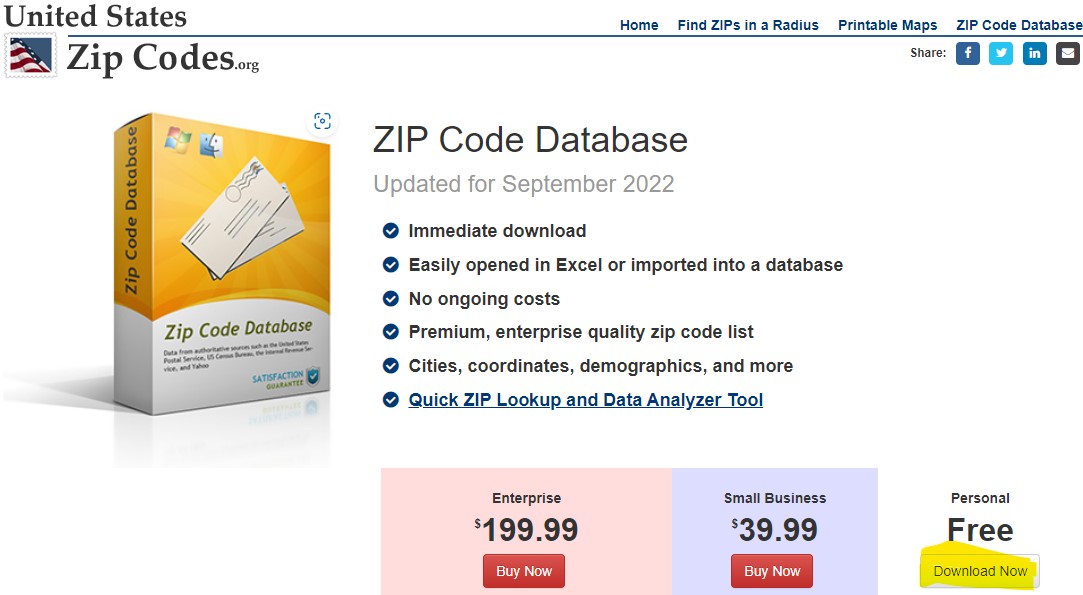
**3) Data Source 3 – list of US zip codes**

Because data for entire US are too big to handle. To keep the analysis within a manageable

scale, we will focus on southern California only.

Step 1: Open following link <https://www.unitedstateszipcodes.org/zip-code-database/>

Step 2: Click the Download Now button on the right bottom:



We take the following steps to scale the zip codes down to southern California (as used in the python code):

* Select state='CA'
* Find online all counties in southern California
* Retain those zip codes for southern California only by selecting SoCal counties.

**How to calculate change in home prices?**

Ideally if some homes change owners every month or every year, we can track these homes to calculate home price changes, but this is not the reality.

Case and Shiller thought out the following method:

* If an area (such as a city, a zip code, etc.) has 61 homes that changes owners in period 1, find the median price (中值) of these 61 homes.
* If the same area has 100 homes that changes owners in period 2, find the median price of these 100 home.
* Calculate the difference in home price between the two. Bear in mind that the two homes from two periods are usually different houses.