

PRACTICA 1

ASIGNATURA: TIPOLOGÍA Y CICLO DE
VIDA DE LOS DATOS

MÁSTER DE CIENCIA DE DATOS



Universitat Oberta
de Catalunya

Alumnos: José Luis Rivas Calduch y Mariano Jiménez Barca
Octubre 2020

INDICE

0. Participantes del grupo	2
1. Contexto:	2
2. Título del dataset:	2
3. Descripción del dataset:	2
4. Representación gráfica:	2
5. Contenido:	3
6. Agradecimientos:	4
7. Inspiración:	4
8. Licencia:	5
9. Código:	5
10. Dataset:	6
11. Tabla de contribuciones:	6
BIBLIOGRAFÍA:	6

0. Participantes del grupo

Jose Luis Rivas Calduch y Mariano jimenez Barca

1. Contexto:

El proyecto consiste en analizar si existe alguna relación entre el coste del dinero y el mercado de la vivienda. Para ello se ha obtenido por un lado información de la Asociación Hipotecaria Española (AHE) que se encarga de recopilar información de diferentes fuentes sobre todo lo relacionado a la evolución del mercado hipotecario Español, y por otro lado se ha obtenido información del Instituto Nacional de Estadísticas Español (INE), del Ministerio de Fomento y de la página idealista.com sobre los principales indicadores del mercado de la vivienda.

Dado que existen muchos datos oficiales al respecto, hemos procurado obtener datos de diferentes fuentes y mediante distintos métodos de cara a poner en práctica varias técnicas.

2. Título del dataset:

historico_indices_hipotecario_vivienda_{fecha}.csv

3. Descripción del dataset:

El conjunto de datos es una serie histórica integrada por los índices de referencia del mercado hipotecario español (EURIBOR, MIBOR, ...) publicados por la AHE en su página web. Dicha asociación se encarga de recopilar esta la información de estas series publicadas por diferentes instituciones (BCE, BdE, ...). Por otro lado, se ha agregado al dataset información publicada por INE y el Ministerio de Vivienda relativas a la evolución del mercado de la vivienda (número de compraventa mensual de viviendas, precio medio de la vivienda, ...). También se han añadido el precio medio de venta y alquiler por metro cuadrado obtenido desde idealista.com

4. Representación gráfica:



5. Contenido:

El dataset comprende los campos descritos en la tabla inferior que comprenden el periodo temporal desde enero de 1986 a SEPTIEMBRE de 2020 siendo cada tupla de la tabla un mes de la serie.

Nombre campo	Ejemplo	Unidades	Fuente	Comentario
ano	2007	años		
mes	1	meses		
tentesp	4811	% 1/1000	AHE	Tipo medio préstamos hipotecarios España
zeuro	4780	% 1/1000	AHE	Tipo medio préstamos hipotecarios zona Euro
euribor	4064	% 1/1000	AHE	
mibor	4052	% 1/1000	AHE	
irs	5375	% 1/1000	AHE	Interest Rate Swap (IRS)
deuda	3727	% 1/1000	AHE	
bancos	4780	% 1/1000	AHE	Tipo medio préstamos hipotecarios Bancos
cajas	4837	% 1/1000	AHE	Tipo medio préstamos hipotecarios Cajas

ceca	5375	% 1/1000	AHE	Indicador CECA
ahe	4000	% 1/1000	AHE	
ced	3763	% 1/1000	AHE	Tipo cédulas hipotecarias
CP335	447846590	personas	INE	Población española
34010110.XLS	230755.0	viviendas	Fomento	Transacciones de viviendas
34020110.XLS	38906160.9	€	Fomento	Precio medio vivienda
ETDP1826	83713.0	viviendas	INE	Número de compraventas
HPT34618	124826.0	hipotecas	INE	Número de hipotecas
CNTR4805	23304	MM€	INE	Producto interior bruto (PIB)
IPV949	2.4	%	INE	Variación mensual PIB
precio_venta_m2	1750.0	Euros/m2	Idealista	Precio medio m2 venta
precio_alq_m2	11.5	Euros/m2	Idealista	Precio medio m2 alquiler

Los datos del AHE se han recogido de la página <http://www.ahe.es/bocms/sites/ahenew/estadisticas/indices-referencia/archivos/historico-de-indices.htm?version=106> mediante scraping

Los datos del INE se han recogido mediante acceso vía API del INE. <https://www.ine.es/dyngs/DataLab/es/manual.html?cid=45>

Los datos del ministerio de Fomento se han recogido desde <https://apps.fomento.gob.es/BoletinOnline2> cuya salida eran ficheros excel.

Los datos de Idealista se han obtenido desde <https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/alquiler/historico/> y <https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/venta/historico/> mediante scraping con selenium.

Los datos son extraídos a una fecha dada la cual es añadida al nombre del fichero.

6. Agradecimientos:

- Asociación Hipotecaria Española (AHE).
- Instituto Nacional de Estadística (INE).
- Ministerio de Fomento de España.

7. Inspiración:

La idea que nos ha impulsado elegir este tema para la práctica es analizar si existe algún tipo de correlación entre los índices de referencia del mercado hipotecario, también

denominados precio del dinero, y los principales indicadores del mercado de vivienda. Dichos índices reflejan el tipo de interés que paga el comprador de una vivienda que haya que recurrir a financiación bancaria, que se produce en la mayoría de las compraventas dado que se establecen hipotecas. En resumen, el objeto es estudiar si la oscilación del precio del dinero en el tiempo afecta al mercado de la vivienda.

8. Licencia:

Licencia del Software:

Copyright (c) 2020 José Luis Rivas Caldach & Mariano Jiménez Barca

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Licencia del dataset:

Creative Commons Attribution 4.0 International

9. Código:

Enlace github: <https://github.com/RivasCaldach/IndiceReferenciaMercadoHipotecario/tree/master/scr>

10. Dataset:

Enlace zenodo: <https://doi.org/10.5281/zenodo.4263379>

11. Tabla de contribuciones:

Contribuciones	Firma
Investigación Previa	JLRC MJB
Redacción de las respuestas	JLRC MJB
Desarrollo Código	JLRC MJB

BIBLIOGRAFÍA:

Subirats, L., Calvo, M. (2018). Web Scraping. Editorial UOC.

Masip, D. El lenguaje Python. Editorial UOC.

Lawson, R. (2015). Web Scraping with Python. Packt Publishing Ltd. Chapter 2. Scraping the Data.

Simon Munzert, Christian Rubba, Peter Meißner, Dominic Nyhuis. (2015). Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining. John Wiley & Sons.

Tutorial de Github: <https://guides.github.com/activities/hello-world>

