## línea horizontal



EDA sobre el potencial económico del teletrabajo

28/07/2021

**─**

Juan Manuel Cornejo

# Visión general

El teletrabajo antes del covid se situaba en una media europea de casi el 15% en europa, con una clara ventaja en los paises nórdicos,con una media del 30%, destacando Países Bajos con el 45% .

En el caso de España la media precovid se situaba en el 7%, no obstante según el paper del Banco de España: Artículos analiticos 2/2020 El Teletrabajo en España, tenemos un potencial para teletrabajar del 30% del mercado laboral y un 50% en profesiones intelectuales

# Objetivos

Verificar el potencial económico y la posiblidad de elegir tu lugar de residencia en base a preferencias personales y no laborales, lo que nos va a llevar a analizar los siguientes puntos

1. Análisis genérico de la situación del mercado laboral en función de la tasa de paro y los afiliados a la seguridad social.
2. Rendimiento del salario en las diferentes provincias, pudiendo valorar el cambio de poder adquisitivo según en qué provincia estemos.
3. Contrastar los precios del mercado inmobiliario para poder valorar donde residir.

# Fuentes

* <https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/>
* <https://prensa.fotocasa.es/fotocasa-fomenta-el-open-data-en-su-nueva-pagina-interactiva-del-indice-inmobiliario/>
* [https://www.ine.es/](https://www.ine.es/jaxiT3/Tabla.htm?t=2184)
* <https://iermb.uab.cat/wp-content/uploads/2020/11/IERMB-WORKING-PAPER-E-20.01.pdf>
* <https://repositorio.bde.es/handle/123456789/12361>

## Uso de fuentes

* Fotocasa solo fue usado para contrastar datos de idealista.
* Idealista se han obtenido datos de m2 alquiler y compra
* INE:

- Valor escriturado total de compra.

- Número total de escrituras.

- Tasas de paro.

- Afiliados a la seguridad social.

- Otras de mera consulta que no han sido relevantes.

* Paper sobre el coste de vida en comunidades, provincias y municipios, para el coste de vida
* Paper del Banco de España, para datos sobre el teletrabajo precovid y su potencial en el país.

# 

# Técnicas utilizadas y desarrollo del proyecto

## Librerías utilizadas

El proyecto ha sido realizado en Python, usando las siguientes librerías:

* Streamlit
* Pandas
* Numpy
* Matplotlib
* Seaborn
* PLotly

## Técnicas utilizadas

1. Búsqueda de fuentes fiables. “ El primer EDA iba orientado para la predicción del cálculo del precio perfecto en base a los competidores por calidad y situación de alquiler turístico y coliving, pero no fue posible en el tiempo disponible encontar una fuente fiable que no tuviera los datos contaminados por el covid, dada las repercusiones económicas que ha tenido para dicho sector. Esto me llevó a cambiar el tema y buscar fuentes de fiabilidad.
2. Tratamiento de datos: exploración inicial de datos en excel descartando datos no relevantes, para su posterior tratamiento mediante lectura de csv, y uso de Dataframes.
3. Generar datos que no existan originalmente pero que mediante relaciones se podían obtener como son:

* Valor medio escriturado > Obtenido a raíz de la suma total de valor escriturado y el número total de escrituraciones.
* Tamaño medio escriturado> Obtenido a raíz del valor medio y del precio medio del m2.
* Poder adquisitivo en función del coste de vida> obteniendo un modificador a raíz del coste de vida en relación al coste de vida en españa y aplicándolo al salario medio.

1. Uso de Seaborn y PLotly para visualización de datos como soporte a la presentación.
2. Uso de Streamlit para presentación dinámica y poder añadir una herramienta financiera como utilidad extra para el usuario.

## Retos y uso del tiempo

La mayor inversión del tiempo ha sido den encontrar los datos, depurarlos y tratarlos, superando las 20h.

Posteriormente en generar las visualizaciones se ha invertido cerca de 8h.

Presentación en streamlit e implementación de la herramienta financiera mas de 16h

En los mayores retos y dificultades encontradas, fue encontrar una manera que el funcionamiento normal de streamlit no afectase negativamente a la experiencia de un posible usuario. Además de tras seleccionar un tema, encontrar datos limpiarlos y trabajar descubrir que esas horas (no contadas) no eran productivas.