# Comentarios Generales:

El siguiente archivo recopila la ejecución de algunos comandos ejecutados sobre el site objeto de la práctica:

https://datosmacro.expansion.com

Así como la reproducción de algunos aspectos de interés recopilados durante la inspección inicial de la página, que sin duda orientan al lector en relación al trabajo realizado y las posibilidades que ofrece la página.

El contenido aquí expuesto, no sigue ningún orden en particular, y son solo notas de trabajo que soportan algunas de las decisiones de diseño incorporadas en el producto desarrollado, y puestas a la mano del lector que requiera conocer mayores detalles del ejemplo.

Se someten a scraping las páginas siguientes, y en el caso de la url .../pib, las referencias que hace a páginas auxiliares:

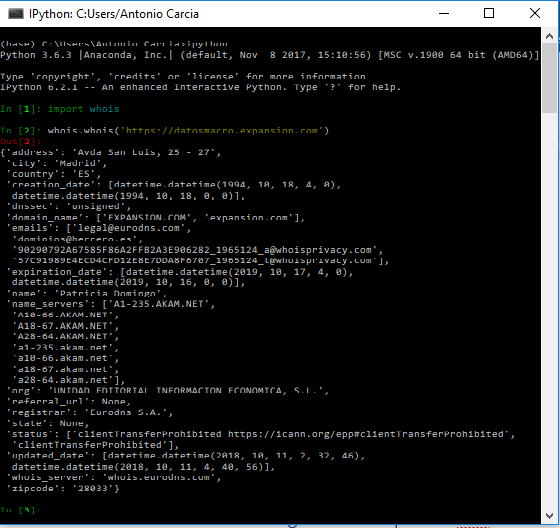
<https://datosmacro.expansion.com/smi>

<https://datosmacro.expansion.com/mercado-laboral/salario-medio>

<https://datosmacro.expansion.com/pib> (Pestaña PIB per cápita)

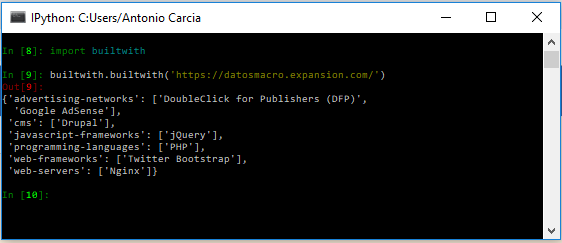
# Información de Propietario del Site

A partir de la URL de inicio del site que provee la información, se determina la entidad a quien corresponde la propiedad del mismo:



# Información acerca de la Tecnología utilizada para el Site

En la figura se reporta la tecnología utilizada por la confección del site



El hecho de que se trate de una página PHP simplifica el trabajo, por cuanto el contenido se despliega en la página, tan pronto esta es requerida para descarga.

# Información Legal publicada en el Sitio

En las url siguientes se muestra la información de condiciones de uso y otras consideraciones legales en relación al site, que, para efectos de simplificar su consideración, se reproducen a continuación:

https://datosmacro.expansion.com/legal/terminos

“El titular de este Sitio Web es Alldatanow, S.L. (en adelante, “Alldatanow”), con C.I.F.: B-87328910 y Registro Mercantil de Madrid Tomo 33.655, Sección 8ª, Folio 60, Hoja M-605792, e-mail: info@datosmacro.com”

https://datosmacro.expansion.com/legal/acerca-de

“Datosmacro nace con la intención de aglutinar las principales variables económicas, tanto de España como del resto del mundo, para ofrecer una visión global de la situación económica en cada momento y en cada país.

En datosmacro encontrarás las series históricas de los principales datos de macroeconomía publicados en cada país. Empezamos publicando datos del paro ([desempleo](https://datosmacro.expansion.com/paro) y datos trimestrales de la encuesta de población activa,[EPA](https://datosmacro.expansion.com/paro-epa/espana)), [PIB](https://datosmacro.expansion.com/pib), [IPC](https://datosmacro.expansion.com/ipc), [inflación armonizada (IPCA)](https://datosmacro.expansion.com/ipca), [SMI](https://datosmacro.expansion.com/smi), [tipos de interés de los bancos centrales](https://datosmacro.expansion.com/tipo-interes), [tipos de referencia de las hipotecas](https://datosmacro.expansion.com/hipotecas), [deuda pública](https://datosmacro.expansion.com/deuda), [déficit](https://datosmacro.expansion.com/deficit), [prima de riesgo](https://datosmacro.expansion.com/prima-riesgo), [divisas](https://datosmacro.expansion.com/divisas), [materias primas](https://datosmacro.expansion.com/materias-primas), pero nuestra vocación es crecer, añadiendo más datos y más países.

Ofrecemos también datos de [población](https://datosmacro.expansion.com/demografia/poblacion), [indice de desarrollo](https://datosmacro.expansion.com/idh) [natalidad](https://datosmacro.expansion.com/demografia/natalidad) , [homicidios](https://datosmacro.expansion.com/demografia/homicidios), [Esperanza de vida](https://datosmacro.expansion.com/demografia/esperanza-vida), etc de más de 180 países.

Puedes ver los datos que mostramos en la página de: [Indicadores](https://datosmacro.expansion.com/indicadores%20econ%C3%B3micos) y acceder desde los siguientes links a los [países](https://datosmacro.expansion.com/paises) y las [comunidades autónomas de España](https://datosmacro.expansion.com/ccaa).

Nuestras fuentes son los organismos oficiales de los distintos países y zonas a las que corresponden los datos. Si encuentras algún error o tienes alguna sugerencia por favor escríbenos a info@datosmacro.com”

# Restricciones de scraping especificadas sobre robots.txt

A través del browser, se accesa a la dirección:

[https://datosmacro.expansion.com/robots.txt](https://ec.europa.eu/eurostat/robots.txt)

obteniendo la siguiente especificación, que da fe del libre acceso a cualquier agente sobre el site.

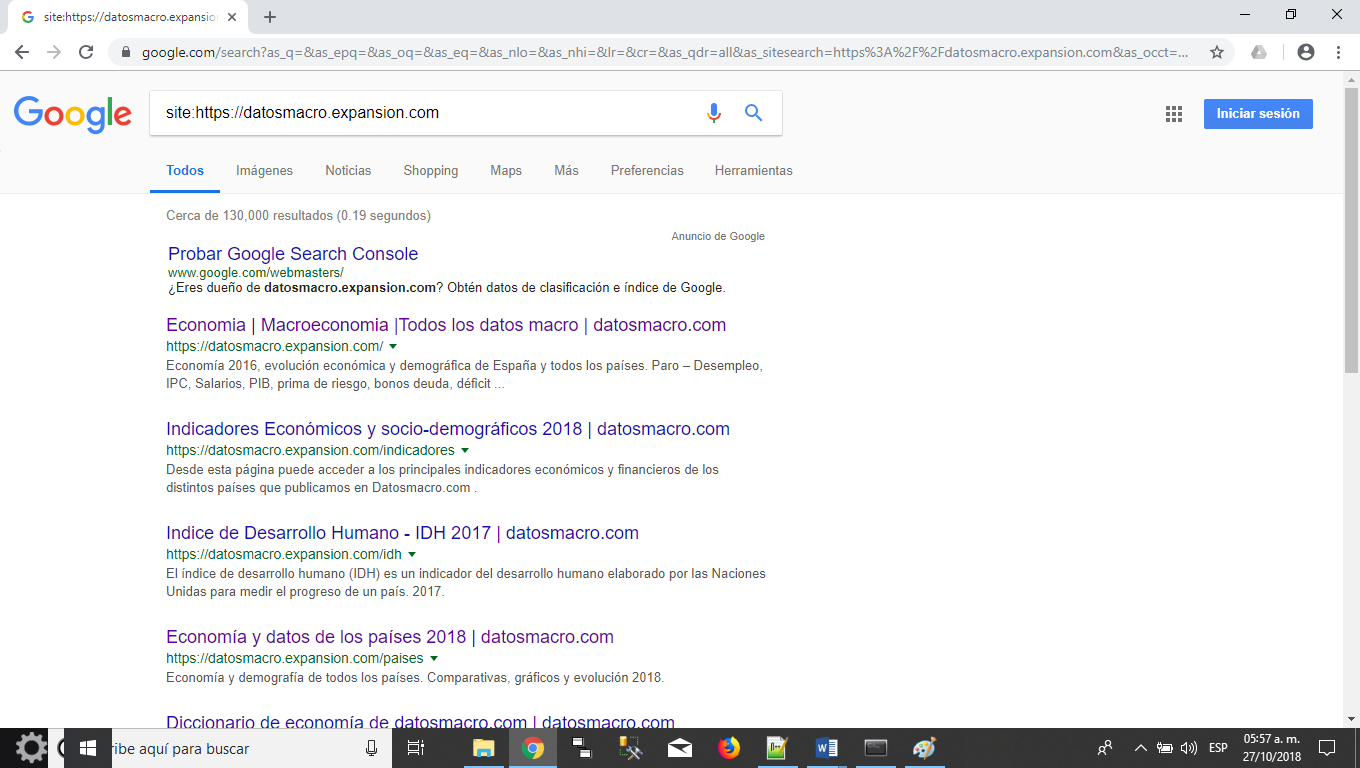
|  |
| --- |
| Sitemap: https://datosmacro.expansion.com/sitemap.xml.gz  Sitemap: <https://datosmacro.expansion.com/sitemap1.xml.gz>  User-agent: \*  Crawl-delay: 10  Allow: /misc/\*.css$  Allow: /misc/\*.css?  Allow: /misc/\*.js$  Allow: /misc/\*.js?  Allow: /misc/\*.gif  Allow: /misc/\*.jpg  Allow: /misc/\*.jpeg  Allow: /misc/\*.png  Allow: /modules/\*.css$  Allow: /modules/\*.css?  Allow: /modules/\*.js$  Allow: /modules/\*.js?  Allow: /modules/\*.gif  Allow: /modules/\*.jpg  Allow: /modules/\*.jpeg  Allow: /modules/\*.png  Allow: /profiles/\*.css$  Allow: /profiles/\*.css?  Allow: /profiles/\*.js$  Allow: /profiles/\*.js?  Allow: /profiles/\*.gif  Allow: /profiles/\*.jpg  Allow: /profiles/\*.jpeg  Allow: /profiles/\*.png  Allow: /themes/\*.css$  Allow: /themes/\*.css?  Allow: /themes/\*.js$  Allow: /themes/\*.js?  Allow: /themes/\*.gif  Allow: /themes/\*.jpg  Allow: /themes/\*.jpeg  Allow: /themes/\*.png  # Directories  Disallow: /user/  Disallow: /includes/  Disallow: /misc/  Disallow: /modules/  Disallow: /profiles/  Disallow: /scripts/  Disallow: /themes/  # Files  Disallow: /CHANGELOG.txt  Disallow: /cron.php  Disallow: /INSTALL.mysql.txt  Disallow: /INSTALL.pgsql.txt  Disallow: /INSTALL.sqlite.txt  Disallow: /install.php  Disallow: /INSTALL.txt  Disallow: /LICENSE.txt  Disallow: /MAINTAINERS.txt  Disallow: /update.php  Disallow: /UPGRADE.txt  Disallow: /xmlrpc.php  # Paths (clean URLs)  Disallow: /node/  Disallow: /en/node/  Disallow: /admin/  Disallow: /en/admin/  Disallow: /comment/reply/  Disallow: /filter/tips/  Disallow: /node/add/  Disallow: /search/  Disallow: /user/register/  Disallow: /user/password/  Disallow: /user/login/  Disallow: /user/logout/  # Paths (no clean URLs)  Disallow: /?q=admin/  Disallow: /?q=en/admin/  Disallow: /?q=comment/reply/  Disallow: /?q=filter/tips/  Disallow: /?q=node/  Disallow: /?q=node/add/  Disallow: /?q=search/  Disallow: /?q=user/password/  Disallow: /?q=user/register/  Disallow: /?q=user/login/  Disallow: /?q=user/logout/ |

# Inspeccionando Tamaño

Utilizando las facilidades de búsqueda avanzada de Google ([www.google.com/advanced\_search](http://www.google.com/advanced_search)), para el sitio

https://datosmacro.expansion.com

Encontramos la cantidad (abrumadora) de 130.000 de páginas, de las que tan solo estamos interesados en algunas



# Base de Datos de Interés

Del site puede extraerse la siguiente nota, que inspira el trabajo que se propone y proporciona las bases para la idea de descargar datos de diversas, y alinearla por año para la confección de un dataset de interés:

“El salario mínimo, junto con el salario medio, la renta per cápita y el IDH de un buen indicador del nivel de vida de un país”.

En consecuencia, se procede a examinar y preparar la extracción de los contenidos de las páginas que de seguido de enumeran. Algunos comentarios asociados a cada una, se reproduce durante la enumeración, para efectos de simplificar la comprensión del material.

## Salario Mínimo Interprofesional

<https://datosmacro.expansion.com/smi>

**“El salario mínimo interprofesional o SMI** fija la cuantía retributiva mínima, o el valor del salario mínimo, que debe recibir un trabajador por la jornada legal de trabajo, independientemente del tipo de contrato que tenga, en un país determinado.

Los datos del SMI que publicamos son mensuales, suponiendo un salario de doce mensualidades cada año

 último salario mínimo establecido en cada uno de los países y su variación respecto al inmediatamente anterior”

## Salario Medio

<https://datosmacro.expansion.com/mercado-laboral/salario-medio>

El **salario medio** es la cuantía retributiva media que recibe al año un trabajador de un país determinado por la jornada legal de trabajo. La ganancia media anual bruta de los **empleados a tiempo completo** en las empresas con 10 o más

## Producto Interno Bruto

## Per Cápita

<https://datosmacro.expansion.com/pib>

“El **Producto Interior Bruto** **(PIB)** mide el valor monetario de la producción de bienes y servicios finales de un país durante un año. El PIB también se define como el conjunto de todos los bienes y servicios finales producidos en un país durante un año. El **PIB** se utiliza habitualmente como medida del grado de bienestar de la población de un país.

El **PIB per cápita** (o PIB por habitante), se obtiene dividiendo el PIB entre el número de habitantes.”

En este caso en particular, debe navegarse por las referencias a cada país ubicadas en la tabla inicial, para luego accesar el detalle de la información tanto del PIB del país como para el PIB per cápita, para el año de interés.

Probado que no hay recursión en los links que deben seguirse