## Creación de un *Dataset* a partir del Registro General de la Flota Pesquera

## Julio Úbeda Quesada y Lucas Zamora Vera

## 1. Contexto

El **Registro General de la Flota Pesquera (RGFP)**, gestionado por el **Ministerio de Agricultura, Pesca y Alimentación de España (MAPA)**, proporciona información detallada sobre la flota pesquera nacional. Se trata de una fuente confiable, ya que los datos provienen de registros oficiales del gobierno (pej.: diarios de pesca, notas de venta, inspecciones, etc.) y son actualizados periódicamente.

|  |  |
| --- | --- |
| **Recurso** | **Fuente** |
| Sitio web | <https://servicio.pesca.mapama.es/censo/ConsultaBuqueRegistro/Buques/Search> |
| Zenodo | <https://doi.org/10.5281/zenodo.15132458> |
| Repositorio con el código | <https://github.com/LucasZV/repo_tipologia.git> |
| Video de presentación |  |

**2. Título**

**"**Registro de la Flota Pesquera Española: Características y Distribución**"**

**3. Descripción del *Dataset***

El *dataset* contiene información sobre las embarcaciones pesqueras registradas en España. Incluye datos sobre su nombre, puerto base, arqueo bruto, potencia, modalidad de pesca y zona de operación.

**4. Representación Gráfica**

(Se incluirá un diagrama representando la estructura de los datos y su relación con las variables)

**5. Contenido**

Los campos incluidos en el *dataset* son:

* **Nombre de la embarcación:** Nombre del buque.
* **Código Buque:** Últimos 5 dígitos del CFR.
* **Código CFR:** Código de identificación único del buque en la Unión Europea.
* **Matrícula:** 4 o 5 dígitos, en el cual viene expresado el Puerto base al que pertenece el buque.
* **IRCS:** Sistema de identificación utilizado para asignar un identificador único a las embarcaciones a nivel mundial.
* **IMO**: Identificador de 3 letras, seguidas de un número de siete dígitos asignado cuando se construye el buque.
* **Modalidad de Pesca:** es el tipo de pesca para el que está autorizado el buque en ese momento.
* **Puerto base:** nombre del puerto donde está inscrito y facultado por la autoridad pesquera competente, de forma previa al inicio de la actividad, para la realización de sus operaciones pesqueras y comerciales.
* **Eslora Total (m):** longitud máxima del barco, definida como la distancia medida en línea recta desde el extremo anterior de la proa hasta el extremo posterior de la popa.
* **Arqueo bruto (GT):** es la medición de la capacidad de carga que tiene el buque expresada en toneladas brutas.
* **Potencia del motor (kW):** es la potencia del motor del buque expresada en Kilovatios.
* **Material del casco:** Tipo de material utilizado para la construcción de la estructura exterior de una embarcación
* **Fecha de Alta:** Fecha en la que el buque queda incluido en el registro.
* **Fecha de Baja:** Fecha en la que el buque sale del Registro (por retirada voluntaria de la actividad, desguace, exportación, siniestro, etc.).
* **Estado:** Situación actual en la que se encuentra el buque, siendo éstas el alta, la baja temporal y la baja definitiva.

El *dataset* abarca registros desde **XXXX hasta la fecha actual**.

**6. Propietario**

El **Ministerio de Agricultura, Pesca y Alimentación** es el propietario de los datos. Para garantizar el cumplimiento ético y legal, se han revisado las políticas de acceso a la información y se ha respetado el uso de datos abiertos.

Citas de análisis anteriores serían:

* [Informe anual de la actividad de la flota pesquera española 2024 (datos 2022).](chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https:/www.mapa.gob.es/en/pesca/temas/registro-flota/informeanualflota2024conpa_tcm38-687759.pdf)
* [Modernización de la flota pesquera española](https://dialnet.unirioja.es/descarga/articulo/2593248.pdf)
* Otros análisis similares.

**7. Inspiración**

Este conjunto de datos es valioso para:

* Analizar el envejecimiento de la flota.
* Identificar la distribución geográfica de los buques activos.
* Evaluar la potencia y capacidad de los buques entre modalidades de pesca.

En comparación con análisis anteriores el presente informe:

* Complementa….
* Se distingue….

**8. Licencia**

Se optará por la ***Open Database License* (ODbL)**, permitiendo el uso, modificación y redistribución de los datos siempre que se reconozca la fuente.

O

Released Under CC0: Public Domain License

**9. Código**

Se ha implementado un script en **Python** para extraer los datos del sitio web y formatearlos en CSV. Las librerías utilizadas incluyen:

* requests
* pandas
* beautifulsoup4

El código se encuentra en la carpeta /source del repositorio.

*En la memoria en PDF, se deben comentar los aspectos más relevantes sobre cómo el código realiza el proceso de recolección de datos, qué dificultades presenta el sitio web elegido, y cómo se han resuelto.*

*El script.py buscara en el código, nombre, matrícula, IRCS, CFR e IMO.*Principio del formulario

Final del formulario

**10. *Dataset***

El *dataset* se ha publicado en **Zenodo** con el DOI: <https://doi.org/10.5281/zenodo.15132458>

*Breve descripción del mismo.*

El archivo CSV también está en la carpeta /dataset del repositorio.

**11. Video**

El video explicativo está disponible en el siguiente enlace: <https://drive.google.com/XXXXXX>

## 12. Contribuciones

|  |  |
| --- | --- |
| Contribuciones | Firma |
| Investigación previa | Julio UQ, Lucas ZV |
| Redacción de las respuestas | Julio UQ, Lucas ZV |
| Desarrollo del código | Julio UQ, Lucas ZV |
| Participación en el vídeo | Julio UQ, Lucas ZV |