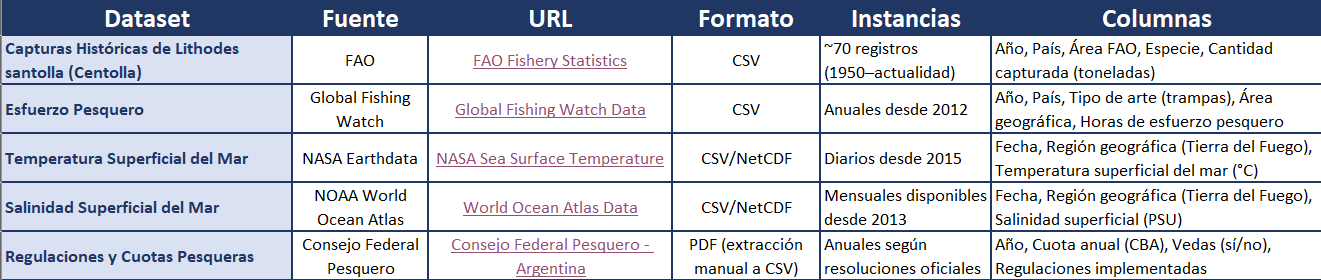
****

**Descripción de los data set**

Este proyecto de aprendizaje automático usara dataset multiples debido a que no se encontró una data set que contenga el conjunto de variables a analizar en un solo data set. Por ello El uso de múltiples datasets responde a la necesidad de obtener una visión holística y multidimensional para realizar una predicción precisa del volumen sostenible de captura de centolla. Cada dataset individual aporta variables claves como datos históricos de captura, esfuerzo pesquero, condiciones ambientales y regulaciones de gestión. Posteriormente, estos datasets separados se consolidarán en una base de datos unificada mediante la integración anual y geográfica, facilitando su análisis conjunto para identificar patrones complejos y relaciones entre las variables y así permitir predicciones robustas y fundamentadas.

**Descripción General de los Data Set**



**Proceso de Recopilación y Preprocesamiento Realizado**

1. **Descarga:** Todos los data sets se descargaron desde fuentes oficiales indicadas.
2. **Filtrado y Validación:** Cada dataset fue filtrado específicamente para la región de Tierra del Fuego y la especie centolla cuando corresponde.
3. **Preprocesamiento:** Los datos ambientales y de esfuerzo pesquero se promediaron o agregaron anualmente para coincidir con el data set principal de capturas.
4. **Limpieza:** Se revisaron datos inconsistentes, faltantes o anomalías, excluyendo registros inválidos.
5. **Integración:** Los datasets se estructuraron para permitir su integración sencilla utilizando Python y herramientas de análisis como pandas y scikit-learn.

Estos datasets ofrecen una perspectiva integral para desarrollar modelos predictivos precisos.

Todos los datasets seleccionados son de dominio público y cuentan con licencia abierta.

**Almacenamiento en Repositorio Git**

Los datasets completos y este informe se encuentran almacenados en el repositorio Git dentro del proyecto generado por Cookiecutter Data Science:

* **Datasets:** /data/raw/
  + Capturas: centolla\_capturas\_anuales\_fao.csv
  + Esfuerzo Pesquero: esfuerzo\_pesquero\_gfw.csv
  + Temperatura: sst\_tdf\_nasa.csv
  + Salinidad: sss\_tdf\_noaa.csv
  + Regulaciones: regulaciones\_centolla\_tdf.pdf
* **Informe:** /docs/

**Fuentes Consultadas**

* https://www.economia.gob.ar/catalogo-sspmi/nuevo\_dataset/pesca-y-puertos-pesqueros.csv – Dataset de capturas pesqueras.
* https://pescare.com.ar/centolla-el-consejo-federal-pesquero-determino-la-captura-biologicamente-aceptable-y-gestion-para-la-temporada-2024-2025 – Cuotas y sostenibilidad 2024-2025.
* https://www.argentina.gob.ar/noticias/la-pesqueria-de-centolla-en-argentina-recibio-la-certificacion-del-marine-stewardship-council – Certificación MSC.
* https://www.unp.edu.ar/centolla-sigue-explotandose-capturas-nunca-antes-registradas – Capturas intensivas.
* https://www.conicet.gov.ar/cientificos-del-conicet-contribuyen-a-la-sustentabilidad-de-la-pesca-de-centollas – Investigaciones sobre sostenibilidad.