

### 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

Dataset	Fuente	URL	Formato	Instancias	Columnas
Capturas Históricas de Lithodes santolla (Centolla)	FAO	FAO Fishery Statistics	CSV	~70 registros (1950–actualidad)	Año, País, Área FAO, Especie, Cantidad capturada (toneladas)
Esfuerzo Pesquero	Global Fishing Watch	Global Fishing Watch Data	CSV	Anuales desde 2012	Año, País, Tipo de arte (trampas), Área geográfica, Horas de esfuerzo pesquero
Temperatura Superficial del Mar	NASA Earthdata	NASA Sea Surface Temperature	CSV/NetCDF	Diarios desde 2015	Fecha, Región geográfica (Tierra del Fuego), Temperatura superficial del mar (°C)
Salinidad Superficial del Mar	NOAA World Ocean Atlas	World Ocean Atlas Data	CSV/NetCDF	Mensuales disponibles desde 2013	Fecha, Región geográfica (Tierra del Fuego), Salinidad superficial (PSU)
Regulaciones y Cuotas Pesqueras	Consejo Federal Pesquero	Consejo Federal Pesquero - Argentina	PDF (extracción manual a CSV)	Anuales según resoluciones oficiales	Año, Cuota anual (CBA), Vedas (sí/no), Regulaciones implementadas

# 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/
  - o Capturas: centolla\_capturas\_anuales\_fao.csv
  - o Esfuerzo Pesquero: esfuerzo pesquero gfw.csv
  - Temperatura: sst tdf nasa.csv
  - o Salinidad: sss tdf noaa.csv
  - o 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.