

TP Final

“Calidad de datos en la información de producción de pozos de gas y petróleo”

Introducción

La secretaría de energía publica los datos referidos a la producción de gas y petróleo tanto de las fuentes convencionales como no convencionales.

Los objetivos de este trabajo práctico son:

1. Aplicar los conocimientos adquiridos en la materia al análisis de un conjunto particular de datos para efectuar un diagnóstico de su calidad
2. Proponer un mecanismo para detectar desvíos en la información referida a la producción de gas y petróleo.

Fuente de datos

Los conjuntos de datos están disponibles en

<http://datos.energia.gob.ar/dataset/produccion-de-petroleo-y-gas-por-pozo>.

- a) Capítulo IV: Pozos: es la información sobre los pozos
- b) Producción de pozos de gas y petróleo no convencional: contiene la información sobre la producción de los pozos no convencionales. Contiene información de varios períodos.
- c) Producción de pozos de gas y petróleo 2024 : contiene la información sobre la producción de los pozos durante el año 2024

Resultados Técnicos Esperados

- a. Analizar la unicidad de los datos en cada una de las fuentes.
- b. Analizar la consistencia de los datos dentro de cada fuente y entre las fuentes.
- c. El Instituto Geográfico Nacional cuenta con información sobre las unidades geográficas (<https://www.ign.gob.ar/NuestrasActividades/InformacionGeoespacial/CapasSIG>) , analice cómo puede usarse esa información para validar los datos del trabajo.
- d. Efectuar un análisis descriptivo de la producción de petróleo y gas tanto para los pozos convencionales como los no convencionales para el año 2024. Efectuar el análisis de petróleo y gas por separado.
- e. Analizar de qué forma la ubicación geográfica de los pozos puede contribuir a mejorar el análisis del punto d.
- f. ¿Qué mecanismos podrían utilizarse para detectar “casos anómalos” de producción?
- g. Identifique dos fuentes complementarias a la utilizada que podrían ser de utilidad para la detección de casos anómalos y utilícelas para enriquecer el resultado del punto f)
- h. Genere un informe donde se indiquen todos los casos anómalos con una justificación sobre los mismos.

- i. Definir 5 reglas de validación que se puedan aplicar sobre los datos, indicar los datos que no cumplen con cada una de las mismas.

Resultados no Técnicos Esperados

- a) Descripción funcional del conjunto de datos
- b) Informe sobre la calidad de datos