

TEAM PROJECT : DATA SOURCE ANALYSIS & PLANNING

Project Information

Project Name: Analisis Keterkaitan Temperature (Suhu) Permukaan dengan Produksi Perikanan Air Tawar Tahunan

Created By: Data Engineering Kelompok 1

Date: 13 April 2025

Version: 1.0

1. Executive Summary

1.1 Project Overview

- **Tujuan Project:** Menganalisis korelasi antara pola suhu dengan hasil produksi perikanan air tawar untuk analisa korelasi.
- **Scope Project:** Integrasi Temperature Permukaan dan Produksi Ikan Air Tawar per Tahun.
- **Expected Outcomes:**
Laporan Analisa yang mengidentifikasi korelasi antara Temperature Permukaan tercover Jumlah Produksi Ikan Air Tawar per Tahun, serta rekomendasi untuk peningkatan efektivitas penggunaan atau penyesuaian temperature yang ada.
- **Timeline:** 3 bulan (Maret – Juni 2025)

1.2 Stakeholders

- **Project Owner:** Kementerian Kelautan dan Perikanan
- **Team Members:**
 - **Data Engineer:** Fajar Hakiki
 - **Data Analyst:** Lyan Fairus, Alif Rahmatul J
 - **Project Manager:** Frezy Ananta
- **End Users:**
 - Google Earth Engine
 - BPS

2. Data Source Analysis

2.1 Produksi Perikanan Budidaya Menurut Komoditas Utama

Source Details

- **Dataset Name:** Produksi Perikanan Budidaya Menurut Komoditas Utama
- **URL/Access Point:** <https://www.bps.go.id/id/statistics-table/2/MTUxMyMy/produksi-perikanan-budidaya-menurut-komoditas-utama>
- **Data Owner:** BPS
- **Update Frequency:** Tahunan

Data Analysis

- **Format Data :** CSV
- **Volume Data :** 15 kb

- **Time Coverage :** 2023
- **Data Quality :**
 - Completeness : Lengkap untuk coverage provinsi tahun 2023 (berdasarkan BPS)
 - Accuracy : Tinggi (langsung oleh BPS sendiri)
 - Consistency : Baik (format data BPS)
 - Timelines : diperbarui (seharusnya) setiap tahun.

2.2 CFSR: Climate Forecast System Reanalysis

Source Details

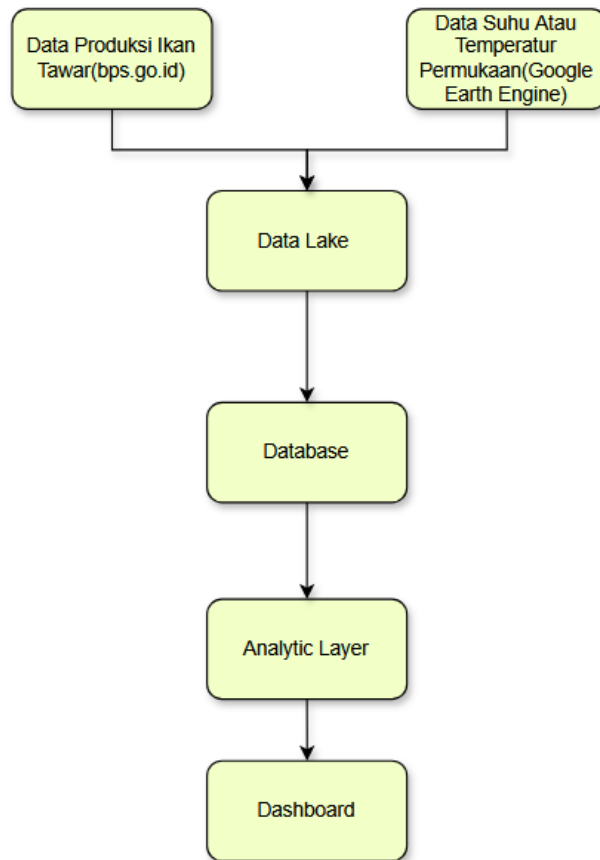
- **Dataset Name:** CFSR: Climate Forecast System Reanalysis
- **URL/Access Point:** https://developers.google.com/earth-engine/datasets/catalog/NOAA_CFSR
- **Data Owner:** Google Earth Engine
- **Update Frequency:** Perbulan
- **Last Update:** 27 April 2025

Data Analysis

- **Format Data :** CSV
- **Volume Data :** 15 kb
- **Time Coverage :** 2023
- **Data Quality :**
 - Completeness : Lengkap untuk coverage provinsi tahun 2023 (berdasarkan Google Earth Engine)
 - Accuracy : Tinggi (langsung diverifikasi menggunakan alat ukur temperature dari google satellite)
 - Timelines : diperbarui setiap bulan

3. Data Flow Mapping

3.1 Data Integration Architecture



3.2 ETL Process Design

- **Extraction Methods**
 - Data Produksi Ikan diekstrak/ scraped dari website BPS yaitu badan pusat statistic yang berupa tabel, disini kode menggunakan identifikasi tabel yang ada dalam website dan diekstrak data tersebut ditulis ulang menjadi data CSV.
 - Data Suhu/Temperature dari GEE diekstrak dalam format yang sama yaitu CSV dari website Google Earth Engine.
- **Transformation Rules**
 - Standarisasi Format Data : Menyesuaikan format agar seragam antara semua dataset.
 - Membuang beberapa ikan yang tidak digunakan sebagai suguhan data utama yang akan diketahui akibat suhu yang ada.
 - Pengelompokkan data berdasarkan tahun dan juga yang dibutuhkan.
 - Memastikan agar tidak ada data yang duplikat atau tidak valid.
- **Loading Procedures**
 - Data yang telah dibersihkan dan diolah dimasukkan kedalam Data Warehouse.

- Data yang telah siap digunakan akan disalurkan ke Analytics Layer untuk visualisasi data.
- **Scheduling**
 - Frekuensi pembaruan setiap tahun mengikuti pengadaan produksi ikan tawar dalam data BPS
 - Menggunakan pipeline ETL yang telah dan memperbarui pipeline untuk setiap data yang akan dipanggil secara berkala.