Vaccination Data Analysis and Power BI Report
Project Overleaf Template

Prepared by: thedynasty

September 7, 2025

Abstract

This document is an editable Overleaf template for the Vaccination Project report. It preserves the original figures and content by including the original PDF and Power BI screenshots. Replace the sections below with your finalized text, or keep the placeholders for collaborative editing.

Contents

1   Executive summary

2   Overview of Project Status (Notebook review)

2.1   Files ingested

- coverage-data.xlsx
- incidence-rate-data.xlsx
- vaccine_introduction.xlsx
- vaccine_schedule.xlsx
- reported_cases.xlsx

2.2   Processing Steps and Status

- Data load and initial profiling (pandas): shapes, null counts, head.
- Cleaning: whitespace normalization, ISO3 mapping, NaN handling, negative value checks.
- Validation: `pandera` schemas applied for structure and type checks.
- Aggregations: coverage by schedule round, incidence aggregations, total case counts.
- Upload: cleaned dataframes written to PostgreSQL hosted on Neon using SQLAlchemy.
- Verification: sample queries (e.g., `SELECT COUNT(*) FROM coverage_data`).

3   Power BI Pages — Design and Functionality

This project includes two Power BI pages that were built on top of the Neon-hosted PostgreSQL tables. Both pages include slicers for interactive filtering (Year, Country/Name, Disease, Antigen). Below each page is documented in detail: the exact visual, why that visual was chosen, and what the visual tells (the specific analytical question it answers).

3.1   Page 1: Dose Drop-off & Booster Uptake

Summary of purpose: identify dose-wise drop-off and booster adoption trends by region and over time. Key components (replace text as needed):

- KPI cards: Dose Drop-off %, Booster Uptake %, Avg Schedule Rounds.
- Bar chart: Avg Coverage % by Schedule rounds.
- Horizontal bar: Avg Coverage % by WHO region.
- Line chart: Booster Uptake % by year and WHO region.
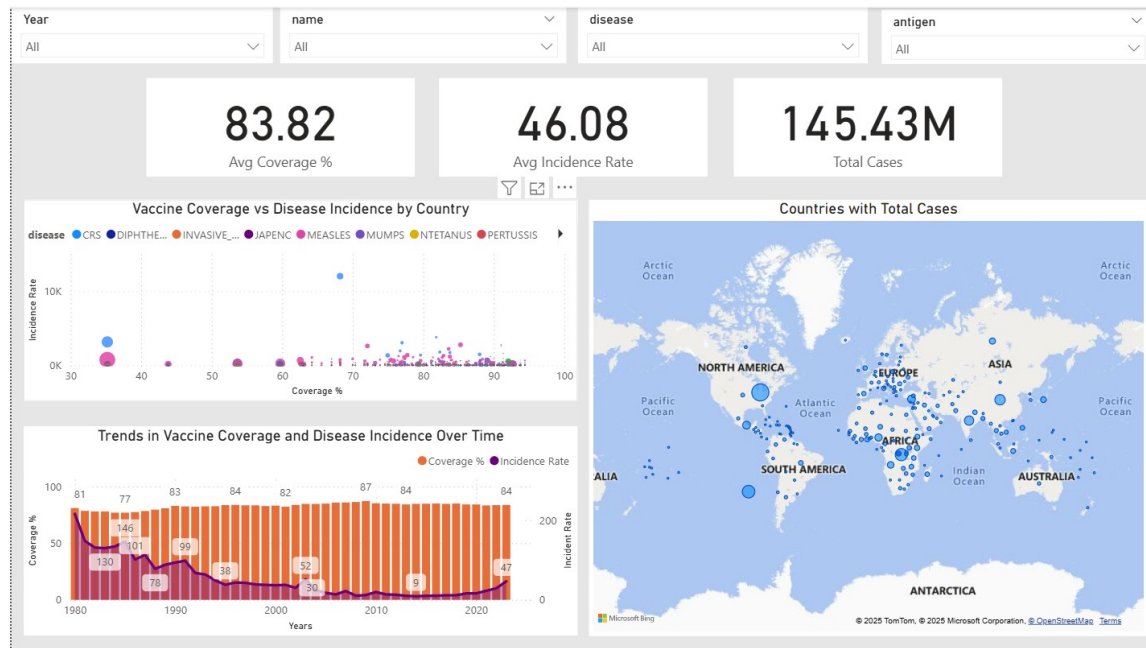- Heatmap: Disease vs WHO region (matrix view).

Figure 1: Power BI — Page 1: Dose Drop-off & Booster Uptake. If this box shows "Missing image", upload the screenshot with name page1$_p$owerbi.jpgor"page1powerbi.jpg".

3.2   Page 2: Coverage vs Incidence & Geographical Spread

Summary of purpose: compare coverage with incidence and show spatial distribution of cases.

- KPI cards: Avg Coverage %, Avg Incidence Rate, Total Cases.
- Scatter/Bubble: Vaccine Coverage vs Disease Incidence by Country (bubble size = total cases, color = disease).
- Map: Countries with Total Cases (geographic bubble map).
- Combo chart: Trends in Coverage and Incidence over time (bar + line with dual Y-axis).
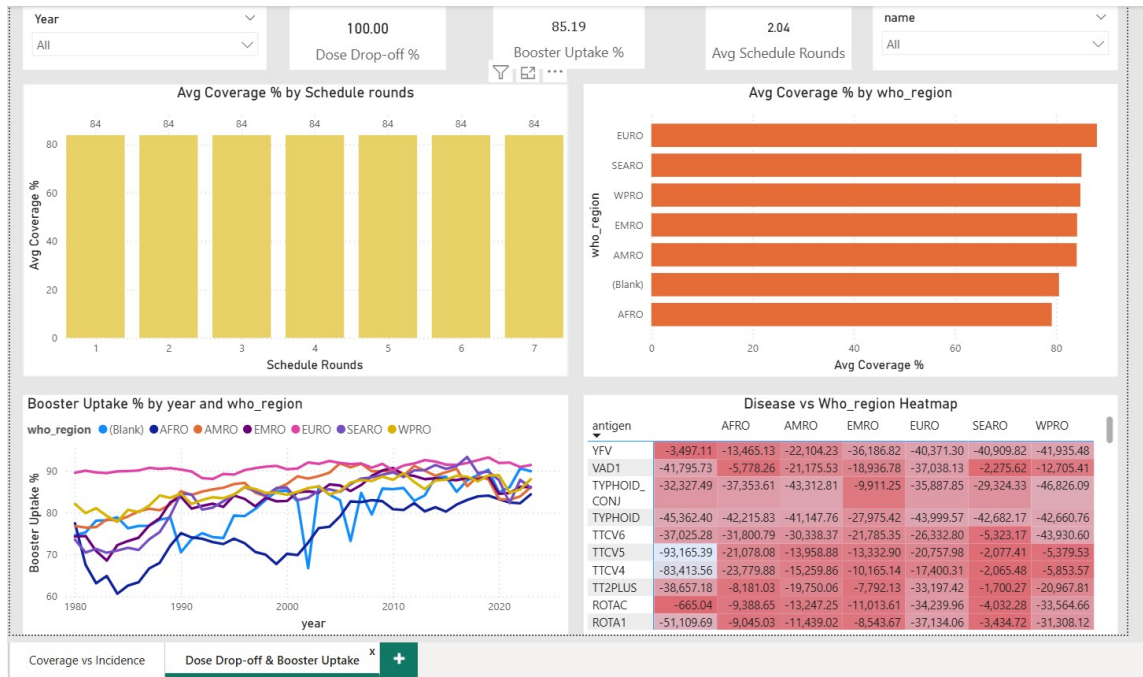
Figure 2: Power BI — Page 2: Coverage vs Incidence and Map. If missing, upload page2$_p$owerbi.jpgor"page2powerbi.jpg"

4   Neon (PostgreSQL) Role and Integration

4.1   How data were written to Neon

Example: cleaned DataFrames were written to Neon via SQLAlchemy from the notebook and then verified with simple COUNT queries. (Credentials should never be stored in the notebook; use environment variables or secrets management.)

4.2   How Power BI connects to Neon

Power BI uses the PostgreSQL connector. Typical connection settings:

- Server: ¡neon-hostname¿
- Database: ¡db-name¿
- Authentication: Basic (username/password) or managed identity depending on setup
- SSL: required (set to "Require")
- Mode: Import or DirectQuery (choose based on refresh needs)

5   Visual-to-Question Mapping

Use this table to map each visual to the question it answers and suggested follow-up actions.

Visual	Question answered	Suggested action
Avg Coverage by Schedule rounds	Where are dose drop-offs occurring?	Target interventions between rounds
Avg Coverage by WHO region	Which regions underperform?	Reallocate outreach/resources

Booster uptake by year/region	How is booster adoption changing over time?	Tailor booster campaigns
Coverage vs Incidence scatter	Is higher coverage linked to lower incidence?	Investigate outliers for surveillance issues
Map of total cases	Where is disease burden concentrated?	Geographically-targeted responses