**1. Project Title & Overview**

* **Project Title:** Analysing Ghana Ministry of Health's 2020 COVID-19 Response & Healthcare Infrastructure
* **Purpose:** This project analyses key data from the Ghana Ministry of Health's 2020 Annual Report to assess the country's COVID-19 response, healthcare infrastructure, and human resources performance during the year.
* **Key Questions Addressed (or intended to be addressed):**
  + What were the key COVID-19 case metrics (total cases, recoveries, deaths) as of early 2021?
  + How was the staff distributed across categories, age, and sex?
  + What was the capacity of healthcare infrastructure (labs, treatment centres, beds, ventilators)?
  + How many tests were conducted and what were the positivity rates by surveillance type?

**2. Data Source**

* **Source Document:** Ghana Ministry of Health 2020 Annual Report.
* **Date of Report:** Report covers data up to December 31, 2020, with some COVID-19 figures up to January 18, 2021.
* **Link/Location:** Ministry of Health Website

**3. Data Extraction Process**

* **Method:** How was the data extracted from the PDF?
  + Manual extraction by copying and pasting text/tables.
* **Challenges (if any):** tables were not easily copy-pasted, required careful manual transcription.

**4. Data Organization (Excel)**

* **Software Used:** Microsoft Excel
* **File Name:** 2020 COVID 19 RESPONSE & INFRASTRUCTURE ANALYSIS.xlsx
* **Sheet Structure:** Describe each sheet and its purpose. For each sheet, list its column headers.
  + **Overall Staff & Trends Sheet:**
    - Purpose: Contains high-level staff numbers and percentage change.
    - Columns: Metric, Value
  + **COVID-19 Cases Overview Sheet:**
    - Purpose: Contains key COVID-19 metrics and regional active case distribution.
    - Columns: COVID-19 Metric (As of Jan 18, 2021), Value (for metrics); Active Cases by Region (Majority), Percentage (%) (for regional data).
    - *Note: Original raw data was in two distinct blocks within this sheet.*
  + **COVID-19 Testing Data Sheet:**
    - Purpose: Details on testing by surveillance type.
    - Columns: Surveillance Type, Total no. Tested, Total No. positive, Positivity Rate (%)
  + **Infrastructure Inventory Sheet:**
    - Purpose: Information on labs, treatment centres, beds, ambulances, and drone centres.
    - Columns: Testing Infrastructure | Detail, Case Management & Facility Metrics | Count, Procured resources (2020) | Count and Zipline drone delivery centres | Location
  + **Staff Categories Sheet:**
    - Purpose: Breakdown of staff by Junior/Senior categories.
    - Columns: Staff Category, Number, Percentage (%)
  + **Staff Class Sheet:**
    - Purpose: Breakdown of staff by Administrative/Accounting class.
    - Columns: Staff class, Percentage (%)
  + **Age Category Sheet:**
    - Purpose: Breakdown of staff by age groups.
    - Columns: Age category, Number, Percentage (%)
  + **Sex Distribution Sheet:**
    - Purpose: Breakdown of staff by sex.
    - Columns: Sex distribution, Number, Percentage (%)

**5. Data Transformation & Cleaning (Power BI / Power Query)**

* **Software Used:** Power BI Desktop (Power Query Editor)
  + Initially, the Staff Demographics Excel sheet contained multiple data blocks in one sheet.
  + To address this, the original Staff Demographics query was duplicated four times in Power Query Editor.
  + For each of the resulting queries (Staff Categories, Staff Class, Age Category, Sex Distribution), detail the applied steps:
    - Removed Top Rows (e.g., 3 rows) to get to the correct header.
    - Used First Row as Headers.
    - Kept Top Rows (e.g., 2 rows) to isolate the specific block.
    - Changed Data Types for 'Number' to Whole Number and 'Percentage (%)' to Percentage/Number.
    - Renamed columns as necessary (e.g., from 'Column1' to 'Number').
  + **Custom Sort Order for COVID-19 Metrics:**
    - Added a new column Metric Order in the Excel COVID-19 Cases Overview sheet to define the custom sort order for the 'COVID-19 Metric' column.
    - In Power BI, used 'Sort by Column' feature on COVID-19 Metric to sort by Metric Order.

**6. Data Model (Power BI)**

* **Description:** Briefly explain the tables in your Power BI model.
* The Power BI data model consists of the following loaded tables, each corresponding to a clean sheet/query from Excel:
  + COVID-19 Cases Overview
  + COVID-19 Testing Data
  + Infrastructure Inventory
  + Staff Categories
  + Sex Distribution
* **Relationships:** Currently, no explicit relationships are defined between these tables as they represent distinct statistical snapshots rather than related transactional data. (This is usually true for this type of summary report).

**7. Visualizations & Analysis**

* **Dashboard Page 1: Overview**
  + Visual 1: (Multi-row Card: Key COVID-19 Metrics) shows (Total cases, Recoveries, Deaths, etc.)
  + Visual 2: (Clustered Column Chart: COVID-19 Testing by Surveillance Type) Explains comparison of tested vs. positive by routine/contact tracing.
* **Dashboard Page 2: Human Resources**
  + Visual 1: (Donut Chart: MoH Staff Category Distribution) Shows Junior vs. Senior staff.
  + Visual 2: (Donut Chart: MoH Staff Sex Distribution) Shows Male vs. Female staff.

**8. Key Findings / Insights**

* Summarize 3-5 of the most important things you learned from the data.
  + The Ministry of Health saw a marginal staff increase of 2.34% from 2019 to 2020, with senior staff comprising the majority.
  + As of Jan 18, 2021, Ghana had a recovery rate of 95.7% and a CFR of 0.61%, with Greater Accra contributing over half of active cases.
  + Ghana established 98 treatment/isolation centres with a total bed capacity of 2,002 beds as of mid-July 2020.
  + Contact tracing accounted for the majority of COVID-19 tests conducted (254,342), with a positivity rate of 9.03%.

**9. Tools Used**

* PDF Reader (Google chrome)
* Microsoft Excel
* Power BI Desktop

**10. Limitations and Future Work**

* **Limitations:**
  + Data is static (snapshot as of specific dates), not live or time-series data.
  + Data was extracted manually, which introduces potential for transcription errors.
  + The report provided summary statistics, not granular raw data.
  + Detailed information on some categories (e.g., full age breakdown, specific locations of infrastructure beyond regions) was not available in the source.
* **Future Work:**
  + Integrate additional years of data for trend analysis.
  + Incorporate more granular data if available (e.g., daily case counts, district-level breakdowns).
  + Explore more advanced Power BI features like DAX calculations for deeper insights.
  + Develop a more interactive dashboard with filters for regions, dates, etc.

**11. Date and Version Control**

* **Date Last Updated:** July 4, 2025
* **Version:** 1.0 (Initial Analysis)