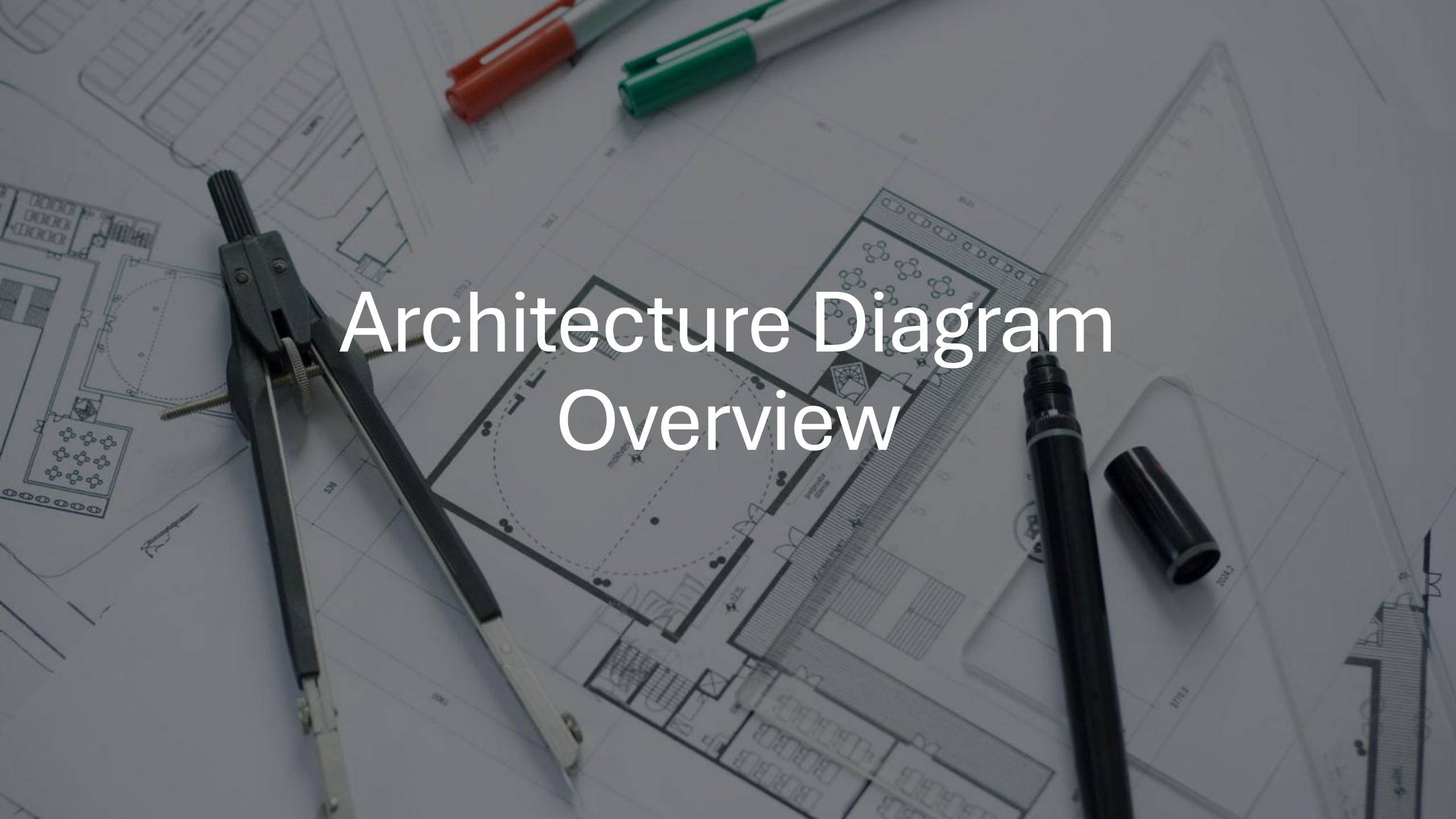
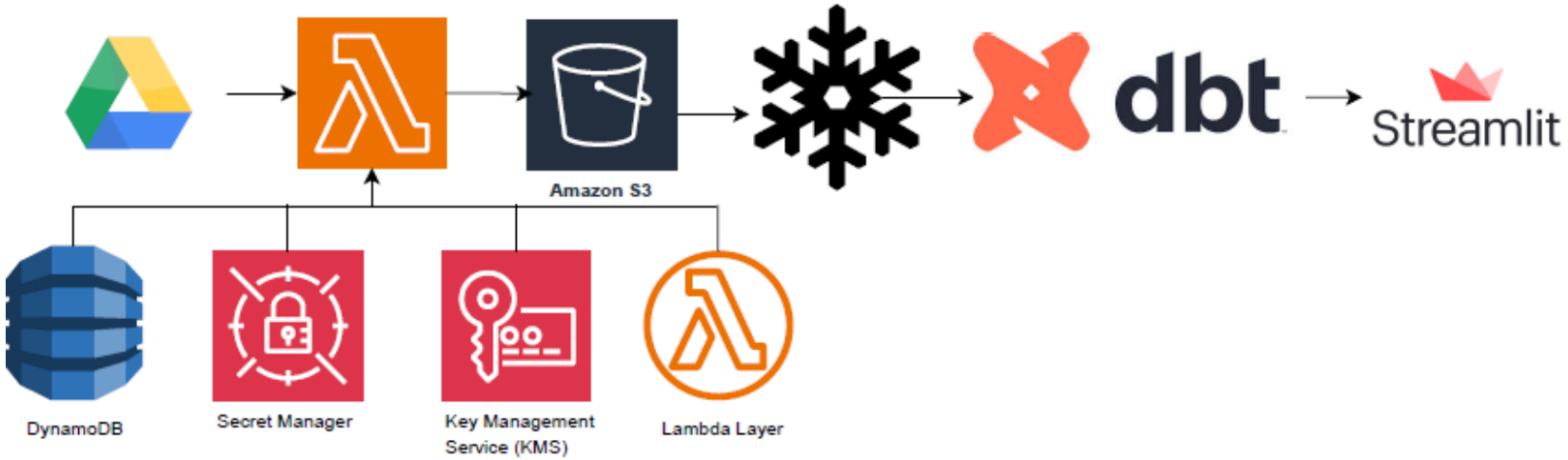


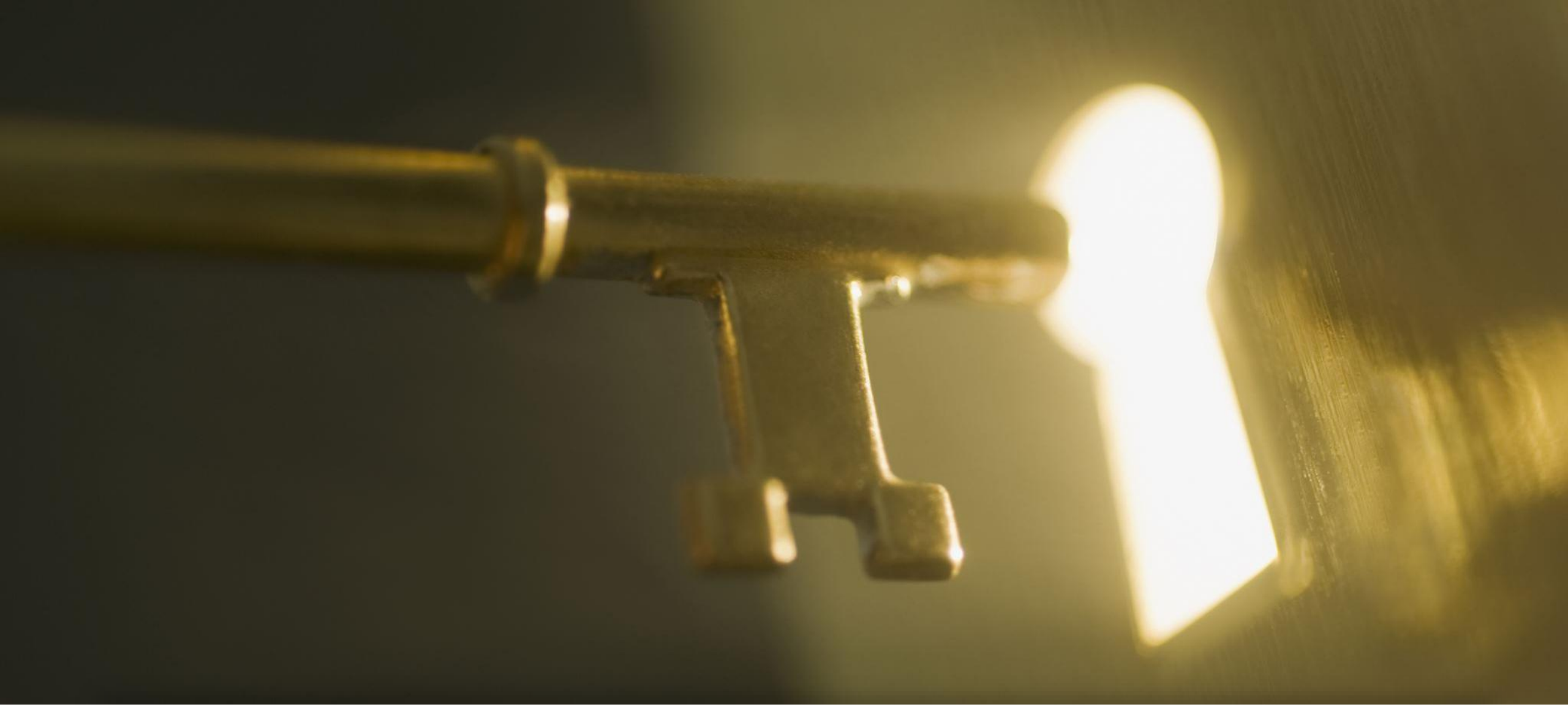
Architecture Diagram Overview

Healthcare Metrics Project



- Google Drive: Serves as the initial source of truth, where raw CSV files are stored and maintained.
- **AWS Lambda:** A serverless function that securely fetches only the required raw CSV files from Google Drive and uploads them to an encrypted S3 bucket. The Lambda function integrates with:
    - **Secrets Manager** to retrieve the Google service account credentials
    - **AWS KMS** to decrypt credentials and encrypt S3 uploads
    - **Lambda Layer** to include external Python packages (e.g., `google-api-python-client`)
    - **DynamoDB** to track ingested file IDs and ensure incremental ingestion

    This step ensures secure, controlled ingestion while avoiding redundant file transfers.
- **Amazon S3:** Functions as the raw data landing zone. Files are stored in an organized directory structure to support external staging in Snowflake.
- **Snowflake:** Acts as the centralized data warehouse. External stages are created to access raw files in S3, and raw tables are populated from those files. This layer is the entry point for structured ingestion.
- **dbt:** Powers the transformation pipeline using a multi-layered approach:
    - **Bronze Layer**: Uses a custom macro (triggered via pre-hooks) to ingest raw CSVs from S3 into Snowflake raw tables.
    - **Silver Layer**: Implements modular dbt models to clean, transform, and enrich the raw data. These transformed datasets form the foundation for analysis and reporting and are materialized in Snowflake.
- **Streamlit:** Serves as the Gold Layer by providing an interactive front-end for exploring and visualizing insights derived from the Silver layer in Snowflake. It uses the `snowflake-connector-python` package to securely connect to Snowflake and execute parameterized SQL queries. Query results are loaded into pandas DataFrames for analysis and visualization using Streamlit's built-in UI components. The application features interactive filters (e.g., dropdowns, sliders, checkboxes) to enable dynamic exploration by facility, ownership type, state, or time range.

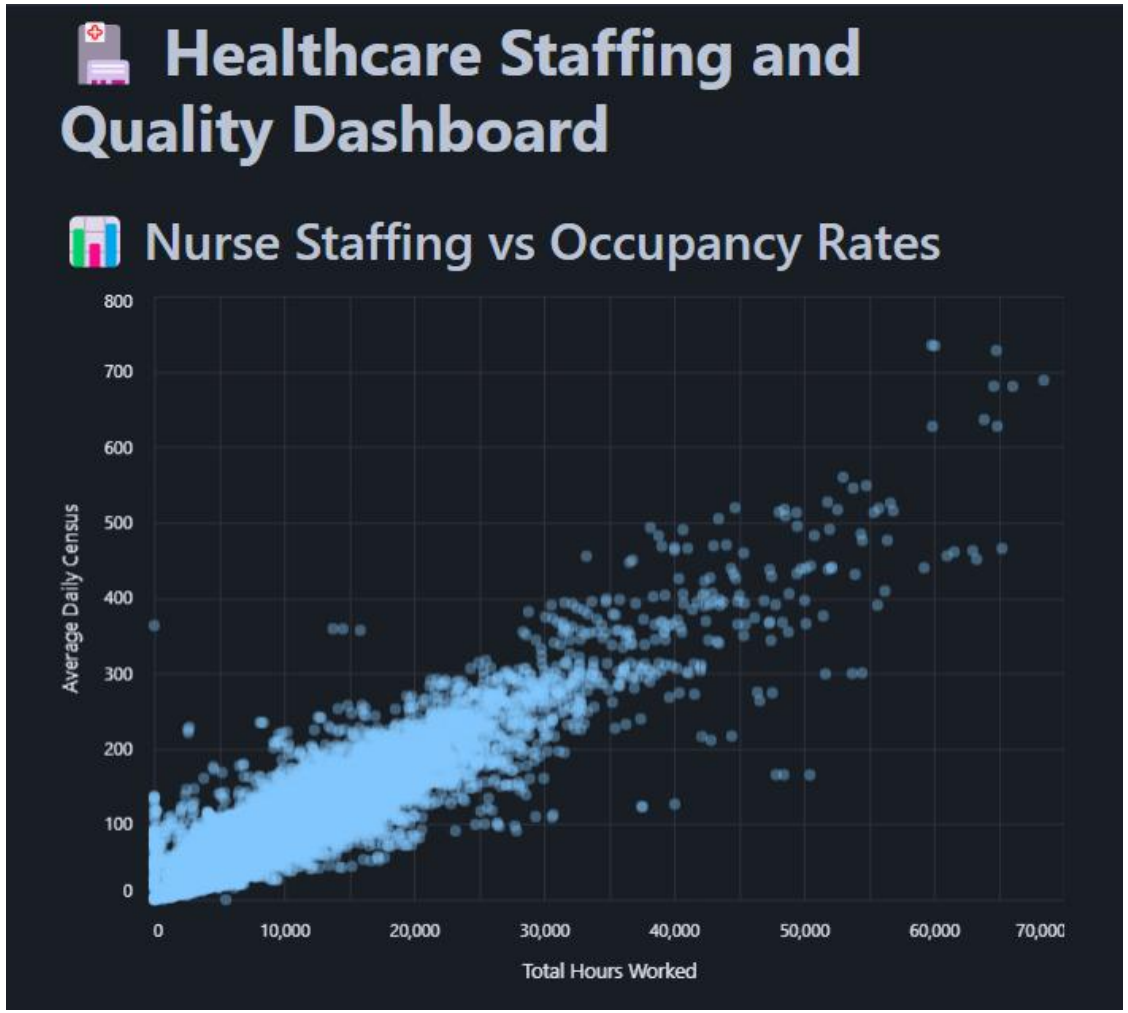


Metric Key Insights & Recommendations (Q2 2024 Data)

Metric Key Insights & Recommendations

- **Key Insights**
  - Strong positive correlation between staffing hours and resident census.
  - Outliers detected: Some overstaffed or understaffed facilities need review.
  - Wide variation at mid-staffing levels shows inconsistent staffing practices.

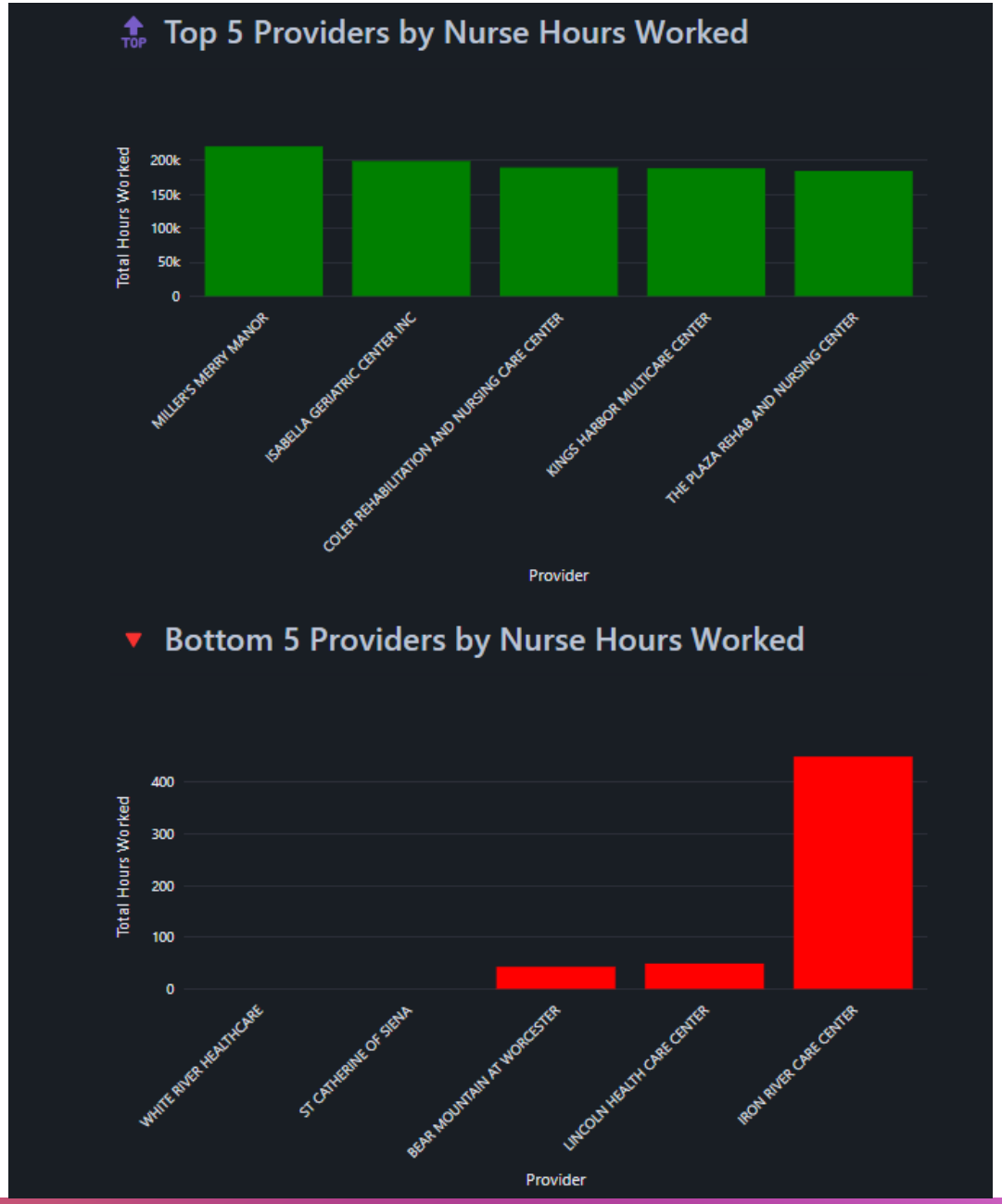
- **Recommendations**
  - Audit outliers for efficiency or quality issues.
  - Set benchmarks using well-aligned facilities.
  - Use census forecasting to optimize staffing levels.



Metric Key Insights & Recommendations

- **Key Insights**
  - Top providers exceed 180k+ nurse hours, reflecting strong staffing capacity.
  - Bottom 5 providers have critical staffing gaps, with some below 50 total hours.
  - Suggests possible reporting issues or severe understaffing.

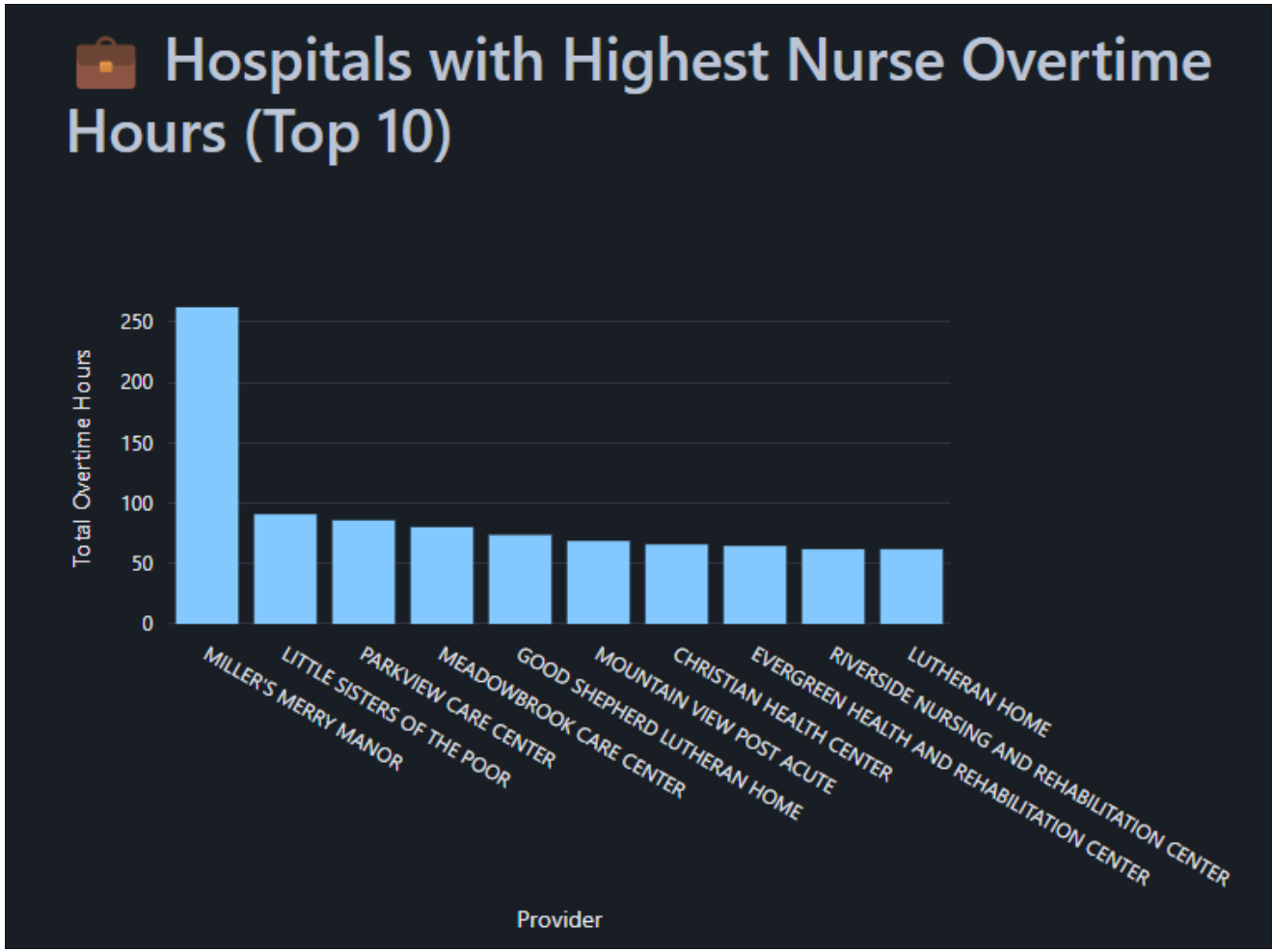
- **Recommendations**
  - Investigate bottom performers for data integrity or operational concerns.
  - Use top performers as staffing benchmarks.
  - Flag extreme outliers for compliance or quality audits.



Metric Key Insights & Recommendations

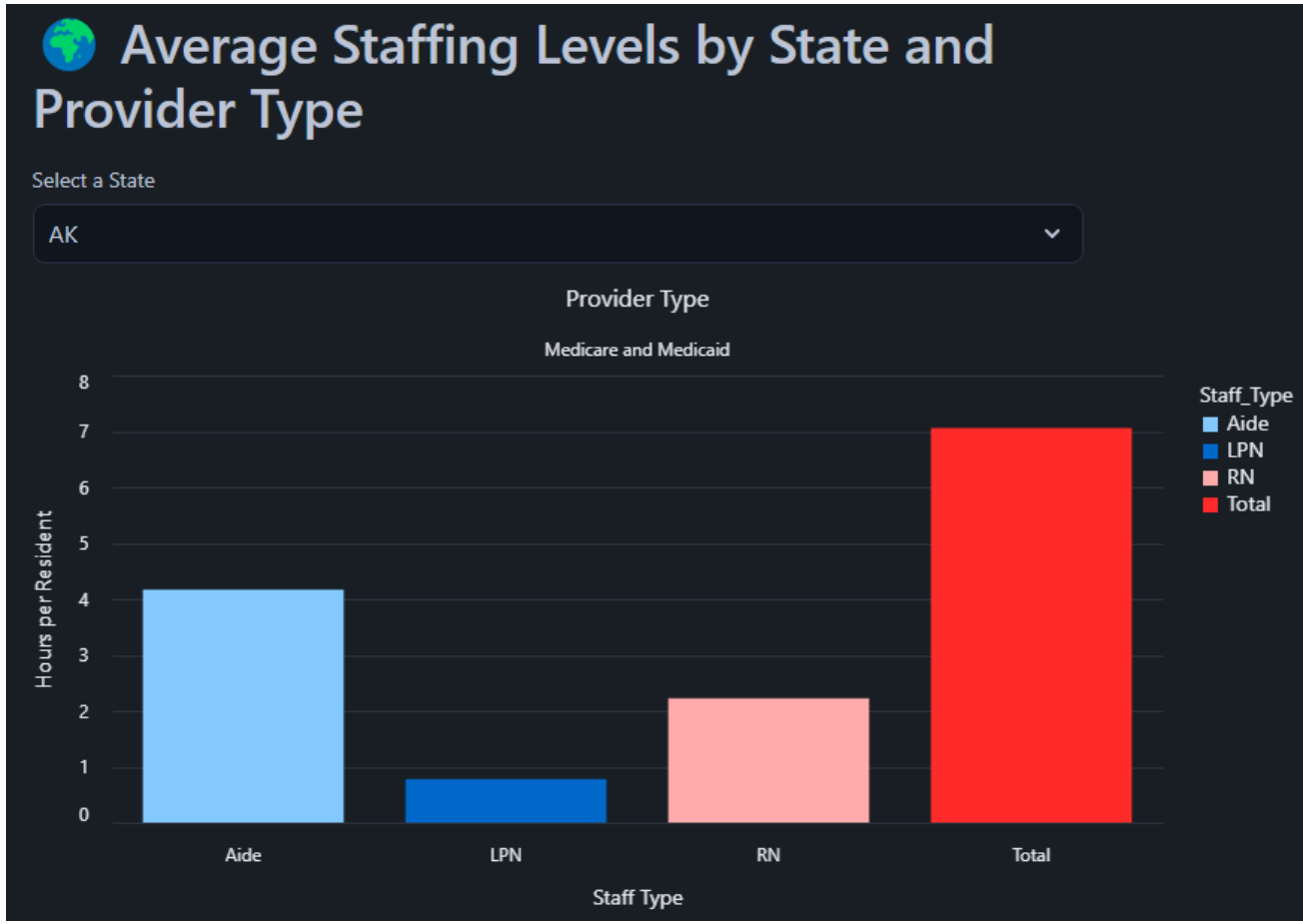
- **Key Insights**
  - Miller's Merry Manor has more than 2.5x the overtime hours of any other hospital.
  - Remaining providers have consistent but elevated overtime levels (~60–90 hours).
  - High overtime may indicate staffing shortages, burnout risks, or scheduling inefficiencies.

- **Recommendations**
  - Investigate bottom performers for data integrity or operational concerns.
  - Use top performers as staffing benchmarks.
  - Flag extreme outliers for compliance or quality audits.



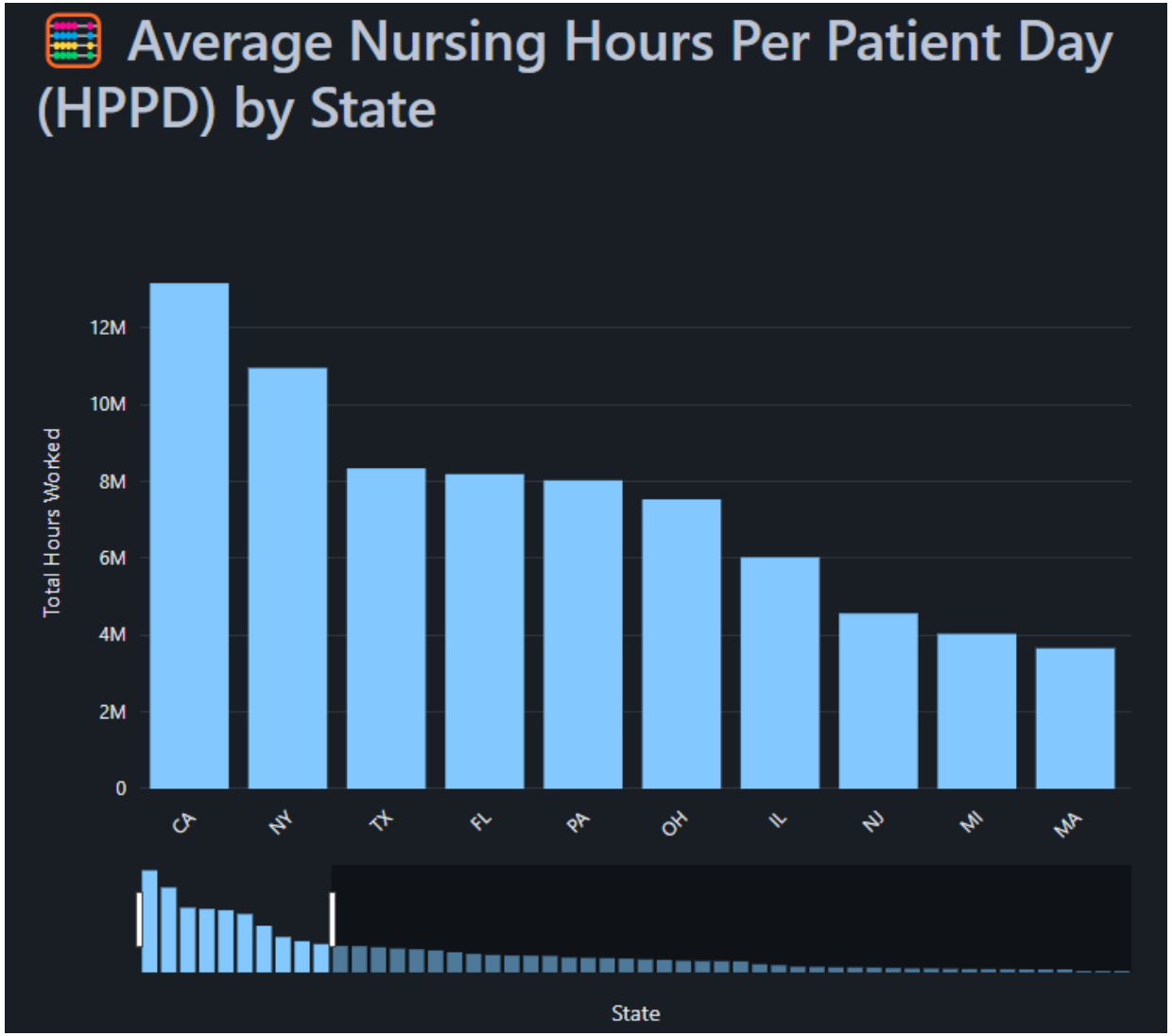
Metric Key Insights & Recommendations (AK)

- **Key Insights**
  - Aides deliver most of the hands-on care, contributing 4+ hours per resident in Alaska.
  - Registered Nurses (RNs) average just over 2 hours per resident, supporting skilled care.
  - Licensed Practical Nurses (LPNs) contribute the fewest hours across staff types.
  - Total combined staffing exceeds 7 hours per resident, suggesting high overall coverage for this state.
- **Recommendations**
  - Maintain strong aide staffing while monitoring for burnout due to high load.
  - Explore opportunities to rebalance responsibilities among RNs and LPNs for efficiency.
  - Use this benchmark to identify underperforming states or facilities with lower staffing hours.



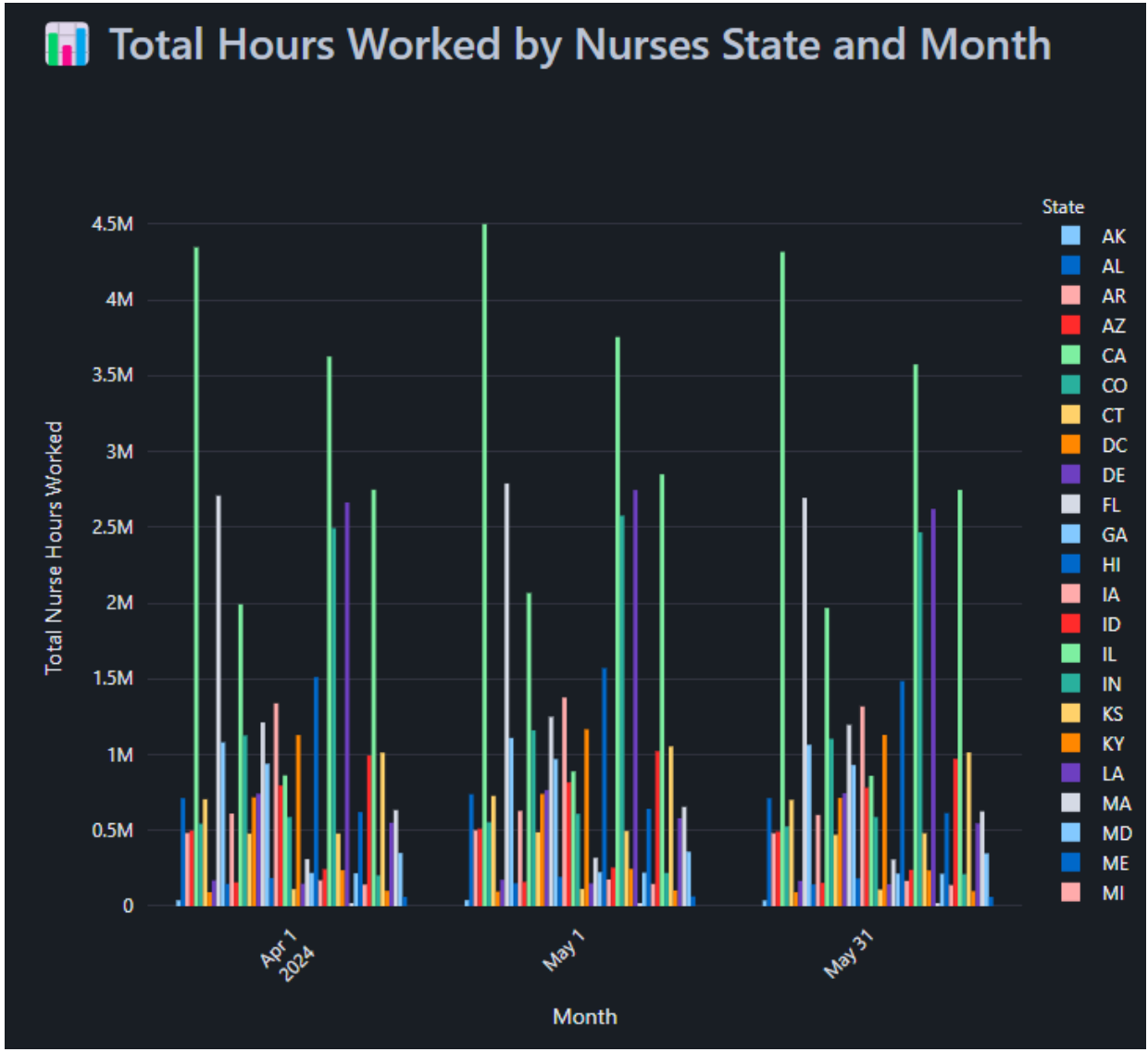
Metric Key Insights & Recommendations

- **Key Insights**
  - This chart displays the top 10 states by total nursing hours worked, offering a snapshot of where the most staff time is being delivered.
  - California (CA) and New York (NY) lead by a wide margin, with 13M+ and 11M+ hours, respectively.
  - Texas (TX), Florida (FL), and Pennsylvania (PA) round out the top 5, all showing similarly high staffing levels.
  - Despite smaller differences, the bottom half of the top 10 (OH, IL, NJ, MI, MA) still show significant volume—indicating robust care infrastructure.
- **Recommendations**
  - Use these top 10 states as benchmarks for staffing capacity and explore operational practices that contribute to their high performance.
  - Investigate whether higher hours correlate with better patient outcomes, particularly in states like CA and NY.
  - Consider scaling resource allocation models from these states to underperforming regions for improved care equity.



Metric Key Insights & Recommendations

- **Key Insights**
  - California (CA) consistently leads in total nurse hours worked across all months shown—exceeding 4 million hours monthly.
  - New York (NY) and Texas (TX) follow closely, maintaining stable contributions above 2.5M hours.
  - States like Florida (FL), Illinois (IL), and Pennsylvania (PA) also show strong staffing levels month over month.
  - Most other states hover below 1M hours, showing consistent but lower relative activity.
- **Recommendations**
  - Benchmark staffing in CA, NY, and TX to understand the systems supporting high utilization and potentially better patient coverage.
  - Drill down by month in lower-staffed states to explore seasonal fluctuations or resource gaps.
  - Consider highlighting a few underperforming or rapidly improving states for more focused intervention or support.



Metric Key Insights & Recommendations

- **Key Insights**
  - 21.94% of short-stay residents are hospitalized after a nursing home (NH) admission — a significant rate indicating potential care transition issues.
  - 12.68% of short-stay residents had an ED (Emergency Department) visit, reflecting common acute events post-discharge.
  - Long-stay residents experience:
    - 1.57 hospitalizations and
    - 1.67 ED visits per 1,000 resident days, suggesting stable but notable levels of acute care utilization.
- **Recommendations**
  - Targeted care coordination for short-stay residents may reduce preventable readmissions.
  - Review discharge planning protocols across facilities with high rehospitalization/ED visit rates.
  - Use predictive modeling to flag high-risk residents for proactive interventions before discharge or transfer.