***Assignment***

***Case Study:*** *Scalable and Secure Client Reporting for a Growing Patient Advocacy Organization*

***Context***

*We are a small but fast-growing* ***Patient Advocacy Group (PAG)*** *supporting patients with various medical conditions. We work with clients like* ***United Healthcare*** *and* ***Humana****, who rely on our data to understand patient treatment trends and optimize healthcare outcomes. At the same time, we serve* ***internal stakeholders*** *like program managers and senior leadership who rely on timely, reliable insights for decision-making.*

*Our data ecosystem is undergoing modernization:*

*• A* ***PostgreSQL reporting data warehouse***

*• ETL pipelines managed via* ***Airflow***

*• A transition from legacy SQL scripts to* ***DBT*** *for transformation*

*• Visualizations built using* ***Tableau****, with plans for deeper* ***role-based access controls***

*• Programming is mainly in* ***SQL***

*• Data privacy and regulatory compliance (e.g., HIPAA) is critical to everything we do*

***Your Task***

*You are stepping into the role of* ***Senior Data & Reporting Analyst****, helping us build and scale our internal and external reporting capabilities. We want to evaluate how you’d approach real-world, messy data and create thoughtful, stakeholder-driven solutions that scale with our mission.*

***Assignment***

*You’ll receive a* ***sample dataset*** *containing mock patient procedures, insurance info, and other anonymized metadata. Assume it's already been ingested into our reporting warehouse.*

*Deliver your responses as:*

*• A brief write-up or slide deck (PDF or Markdown is fine)*

***Task 1: Build a Client-Facing Report (United Healthcare)***

*United Healthcare wants to:*

*•* ***Standardize insurance coverage for treatments*** *(group similar procedures)*

*•* ***Identify high-volume procedures*** *and* ***patients missed*** *(especially uninsured)*

*•* ***Export the data*** *sorted by procedure and latest procedure date*

*• Receive data visualizations via Tableau, without exposing patient PII*

***Your Deliverables:***

*• Describe how you would model and transform this data (DBT or SQL logic is welcome)*

*• Design a Tableau (or any other visualization tool) dashboard mockup*

*• Define filters or permissions needed for secure client access (SQL level or Tableau)*

*• Include an explanation of your approach to anonymization, grouping, and export logic*

How to do this:

* Set Up PostgreSQL Database (Locally or on cloud postgressql service, for ex. Google Cloud SQL?)
* Load csv files into PostgreSQL tables.
* Install and configure dbt to connect to PostgreSQL (create a profiles.yml file)
* Write dbt models, transformations to raw csv files to create clean, analysis ready tables.

Notes/Summary:

Build everything locally (PostgreSQL + dbt + Tableau) for now

Create Tableau extracts and publish to Tableau Public for stakeholder access

Later I can document how I’d set this up in cloud infrastructure for production

***Task 2: Internal Report for Multi-Specialty Hospital Team (Humana)***

*Humana needs:*

*• A dashboard that allows team members to* ***filter by specialty*** *(procedure that they know everything of)*

*• A mechanism to* ***restrict visibility*** *of* ***"hospice treatments"*** *unless authorized*

*• A way to track* ***patients helped*** *by* ***time frame and procedure***

***Your Deliverables:***

*• Describe the technical setup (think about how access and transformation might be handled)*

*• Propose a dashboard layout that supports specialty filtering + sensitive data hiding*

*• Explain how you would integrate these permissions and privacy constraints*

***Task 3: Senior Management Deep-Dive Reporting***

*Our leadership wants:•* ***Detailed operational insights*** *with as much granularity as possible*

*• Extracts available to all staff, but* ***visualizations only for execs***

*• Ongoing visibility into* ***volume, cost drivers, missed patients, and trends***

***Your Deliverables:***

*• Describe the structure of the “executive dashboard” vs. the raw extracts*

*• Explain how you’d use Tableau permissions or data-layer strategies to separate access*

*• Propose 3–5 key metrics or KPIs that would matter most to leadership*

***Task 4: Data Pipeline & Governance Proposal***

*Our team has been using Airflow and DBT but is still early in the process.*

*We'd like to see how you'd ensure:*

*•* ***Clean, well-modeled data***

*•* ***Repeatable, scalable pipelines***

*•* ***Governance and security best practices***

*•* ***Smooth Tableau reporting layer integration***