

## Summary

- Data Scientist & Analyst** with over **5 years** of experience in **Data Analytics** and Modeling, specializing in **Machine Learning, Statistical Analysis, and Predictive Modeling** to enhance patient outcomes and streamline operational efficiency.
- Proficient in **SQL, Python, and AWS**, with a proven track record of processing **large-scale datasets** and delivering scalable solutions.
- Skilled in **ETL processes** and proficient in managing **SQL databases (MySQL, PostgreSQL)**, with expertise in **AWS** for scalable data integration and processing of large healthcare datasets.
- Skilled in crafting **interactive dashboards and visualizations** using **Power BI and Tableau**, driving actionable insights for healthcare stakeholders and decision-makers.
- In-depth knowledge of medical records systems (**EHR, EMR**), **HL7 standards**, and **HIPAA compliance**, ensuring secure and compliant handling of sensitive healthcare data.
- Advanced knowledge of **statistical techniques** (regression, hypothesis testing, forecasting) and **machine learning**, optimizing clinical and operational workflows.
- Collaborative communicator, experienced in working with **cross-functional teams, healthcare providers**, and IT specialists to deliver tailored analytical solutions and support continuous improvement initiatives.

## Skills

Programming Languages	Python, R, SQL
Packages	Pandas, NumPy, Matplotlib, Seaborn, ggplot, Jupyter Notebook
Data Visualization Tools	Tableau, Power BI, QuickSight
Machine Learning & Data Science	Scikit-learn, Random Forest, Logistic Regression, Decision Trees, SVM, KNN, Gradient Boosting, Statistical Modeling (t-test, Chi-squared, Regression, Forecasting)
Databases	SQL Server, MySQL, PostgreSQL, MongoDB (NoSql), Redshift (explored), Oracle (basic)
ETL Tools	Informatica, Talend
Statistical Tools	SAS, R
Cloud	AWS, Azure
Data Management	Data Wrangling, Data Integration, Data Warehousing, Data Mining
Healthcare Knowledge	HIPAA, Electronic Health Records (EHR), (HL7) Standards
Methodologies	Agile, Scrum, Kanban
Reporting	Advance Excel (VBA, Pivot Tables)
Collaboration Tools	Jira, Confluence, GitHub

## Education

Masters in Information Systems – Business Analytics

California State University, Fullerton, CA

| Jan 2024

## Experience

Cigna Healthcare Group, USA

Data Scientist

| Feb 2024 – Present

- Utilized Python and SQL to develop and automate ETL pipelines, integrating diverse healthcare data sources into a centralized SQL Server database, ensuring data consistency and synchronization.
- Engineered key features (e.g., claim types, patient demographics, provider information) from integrated datasets to enhance model performance, deploying predictive models within the claims processing workflow for actionable insights.
- Applied statistical tools like SAS and R to develop machine learning models (e.g., regression, gradient boosting machines), predicting claim processing times and identifying error-prone claims, contributing to a 30% reduction in processing times and a 20% decrease in errors.
- Analyzed healthcare data from diverse sources, including EHR systems and claims data, to identify trends, patterns, and opportunities for improvement, enhancing decision-making processes for patient care and clinical outcomes.
- Conducted exploratory data analysis (EDA) using Python with Pandas, NumPy and Seaborn to uncover insights into operational efficiencies and device performance, informing data-driven optimization strategies.
- Designed and optimized SQL Server to streamline data collection, querying, and reporting processes, ensuring data integrity

and supporting the development of economic models and cost calculators for health economics analysis.

- Implemented ETL processes to validate and transform data for Business Intelligence objectives, improving data accuracy, consistency, and the efficiency of report generation.
- Developed and maintained interactive dashboards and visualizations using Power BI to effectively communicate complex healthcare information, track key performance indicators (KPIs), and monitor healthcare quality metrics.
- Ensured compliance with healthcare regulations (HIPAA, HL7) during the implementation and integration of EMR software, maintaining data privacy and security standards.
- Reviewed and analyzed medical codes and claims to ensure accuracy and alignment with the care provided, facilitating proper reimbursement and compliance.

## **ValueLabs, India**

*Healthcare Data Analyst (Lead on Machine Learning Projects)*

*| Mar 2019 – Dec 2021*

- Applied Agile methodology within a cross-functional team to ensure efficient project execution, collaboration, and iterative delivery of healthcare data solutions.
- Optimized Python scripts to automate data extraction, cleaning, and transformation, reducing manual effort by 25% and increasing operational efficiency.
- Experienced in using Jupyter notebooks for reproducible data analysis and modeling, conducting exploratory data analysis (EDA) to identify trends and refine predictive models.
- Conducted detailed exploratory data analysis (EDA) using Python libraries such as Pandas, Matplotlib, and Seaborn to identify trends and patterns in product data, supporting improved operational efficiency and clinical decision-making.
- Utilized SAS for advanced statistical analysis and predictive modeling of patient outcomes, providing accurate forecasts to inform strategic decisions related to patient care and resource allocation.
- Performed data mapping and logical data modeling, created diagrams, and utilized SQL queries to filter and analyze data, enhancing data accuracy and operational efficiency.
- Collaborated with regional leaders to gather data on global reimbursement codes and payment amounts by procedure, country, and site of care, presenting findings through interactive Power BI dashboards.
- Delivered Proof of Concepts (POCs) using AWS services to demonstrate innovative visualization solutions tailored to business requirements and objectives.
- Processed charge postings in EHR systems, handling billing for prescriptions and medical visits using ICD-10 and CPT codes for cancer patients.
- Communicated with providers and payers to process claims efficiently and gained a thorough understanding of HIPAA regulations for data privacy and security.

## **Hexaware, India**

*Data Visualization Analyst*

*| Jan 2018 – Feb 2019*

- Gathered business requirements by collaborating with end users, Project Managers, and SMEs to understand and document business processes, ensuring visualization solutions aligned with stakeholder needs.
- Created detailed Source-to-Target Mapping (STM) documents outlining data transformation logic, ensuring accuracy and consistency in visualization outputs.
- Installed, configured, and maintained Tableau Server across multiple environments to support scalable, enterprise-level visualization deployments.
- Integrated Python and NumPy with Tableau to perform advanced statistical analysis and enhance the depth of data insights presented in visualizations.
- Conducted data analysis and profiling using SQL on source systems such as Oracle and Teradata, ensuring data quality and integrity for visualization purposes.
- Developed interactive and customized Tableau Dashboards, incorporating Custom SQL from sources like Teradata, Oracle, and Hadoop, and performed data blending to provide comprehensive, cross-platform insights.
- Authored complex SQL scripts to create BI layers within the Teradata data warehouse, ensuring optimized data retrieval for Tableau reporting.
- Designed user-friendly dashboards and visualizations in Tableau, presenting data insights on multiple platforms to enhance user engagement and decision-making efficiency.

## **Certifications**

- **Salesforce Administrator**
  - **Salesforce AI Associate**
-