**Saiteja Kandula | Healthcare Data Analyst**

**CONTACT:**  +1 (203)545-3245 | **Email ID:** [Kandulasaiteja07@gmail.com](mailto:Kandulasaiteja07@gmail.com)

**LINKEDIN:** [*https://*](https://www.linkedin.com/in/sai-teja-kandula-96b8372b2/)*www.linkedin.com/in/kandulasaiteja*

**SUMMARY:**

* **Over 9 years** of Professional Experience in different environments and platforms including **Data Analysis, data warehousing, data modeling, data mapping, data lineage, data profiling, Informatica Data Quality (IDQ), Informatica Data Analyst (IDA), Informatica Power Center** and Client - Server applications in Healthcare Industry.
* Played a pivotal role in streamlining data mapping procedures, aiding in **ETL** processes, and facilitating the transition of data from **OLTP to OLAP** using **ETL** tools like **Talend** and **Informatica PowerCenter**.
* As a Data Analyst responsible for Data Modelling, Enterprise Data Management, Data Presentation, Visualization, Optimization, Risk assessment, predictive analysis, advance Data Analysis, Business Research and Quantitative Analysis.
* Extensive familiarity with healthcare regulations including **HIPAA, HITECH, PPACA, CLIA,** and **GDPR** guidelines.
* Skilled in utilizing medical coding systems such as **ICD-10, HCPCS**, and **LOINC**, proficient in integrating and aligning various coding standards.
* Proficient in employing messaging standards like **HL7** and **FHIR** to ensure interoperability in healthcare systems.
* Expertise in healthcare quality metrics like **HEDIS** and **CMS** measures, facilitating thorough evaluation of clinical data.
* Extensive experience in implementing health information systems, such as E**lectronic health records (EHR)** or H**ealth information exchanges (HIE).**
* Demonstrated expertise in data privacy and data protection regulations, such as **GDPR** and **CCPA**, and assisted in ensuring compliance through data governance practices.
* Proficient in claims processing, **ICD-9** and **ICD-10** coding systems, with expertise in analyzing and interpreting complex medical coding data.
* Proficient in leveraging industry leading **EHR** platforms like Cerner Citrix and Epic to extract actionable insights and optimize healthcare outcomes.
* Familiarity with the **Medicare** and **Medicaid** sectors within the healthcare industry, encompassing aspects related to inpatient and outpatient care as well as reimbursement methodologies.
* Proficient in leveraging comprehensive Healthcare management systems like **QNXT** and **Facets** for streamlined operations, member enrollment, provider management, and financial functionalities.
* Collaborated with cross-functional teams, including clinicians, developers, and testers, employing Agile practices to create rapid prototypes and mock ups of data visualizations.
* Facilitated meticulous requirements gathering sessions with stakeholders, developed user stories in **JIRA** establishing clear project specifications and acceptance criteria for streamlined project execution.
* Proficient in **JIRA** project management software, with a strong understanding of **Agile methodologies** for sprint planning and user story mapping.
* Employed advanced **SQL** queries, encompassing joins, views, pivots, and stored procedures, to examine extensive datasets while ensuring sensitive patient information remained masked.
* Leveraged **PowerBI** for thorough analysis of clinical data, patient demographics, resource allocation, and patient flow within healthcare establishments, generating insights to enhance quality and operational efficiency.
* Utilized **Tableau** for comprehensive analysis of **Electronic health records (EHR**), claims processing data, and patient journeys, fostering insights for refining patient care and devising strategies for infection control.
* Hands-on experience in application development utilizing **Java**, relational databases, **TALEND**, and **Linux shell scripting**, with proficiency in **DB2**.
* Proficient in utilizing the **Spark** ecosystem, employing **Spark SQL** and **Scala** queries to handle various data formats such as Text files, Avro, and Parquet files. Skilled in working with diverse databases including **MySQL, SQL Server, DB2, Oracle, as well as NoSQL databases like MongoDB, Cassandra, DynamoDB, Redshift, and HBase**.
* Experienced in using Sqoop for data migration between relational databases**, NoSQL** database and **HDFS.** Competent in managing Hadoop clusters and services through **Cloudera** Manager.
* Possess expert knowledge of the **Software Development Life Cycle (SDLC)** and have practical experience across all phases.
* Proficient in **Python** programming with the ability to craft and automate data analysis workflows, construct predictive models, and generate visualizations using widely used libraries such as **Pandas**, **NumPy**, and **matplotlib**.
* Skilled in utilizing the **R** language for statistical analysis and data visualization, enhancing the toolkit for data exploration and modeling.
* Managed midsize teams and developed various business solutions on **patient analytics**, including physician database standardization and sales and **claims data analysis**.
* Developed **Python** scripts with Lambda functions to extract data from **AWS** **DynamoDB**, transforming it into **JSON** and inserting it into an **SQL** server.
* Built automated visualization dashboards for testing voluminous incremental data feeds, analyzing marketing data, and delivering **analytic solutions** to clients.
* Created **PowerBI** dashboards using **DAX** and power queries, engaging with clients to understand business requirements, and managing **Ad-Hoc** requests.
* Implemented streamlined workflows using high-performance **API** services and developed Spark jobs in **Python** for **ETL** transformations, eliminating the need for **ETL** tools.
* Executed **Hadoop** and **Spark** jobs on **AWS** **EMR**, utilizing data stored in **AmazonS3**, and analyzed data using **Alteryx** workflows and **ETL** processes.

**EDUCATION**:

* **Master’s** in Health Informatics, Sacred Heart University, Fairfield, CT – 2017.

**TECHNICAL EXPERIENCE:**

|  |  |
| --- | --- |
| **Big Data/Hadoop Technologies** | Spark, SparkSQL, Azure, Spark Streaming, Kafka, PySpark, Hive, HBase |
| **Languages** | HTML5, JSON, Java, Scala, Python (NumPy, SciPy, Pandas), Java Script, Shell Scripting |
| **Vizuvalization Tools** | Tableau.PowerBI |
| **Healthcare Tools and Modules** | FACETS, McKesson, QNXT configuration, Billing, Claims processing and submission, |
| **Healthcare Standards and Security** | HIPAA 4010,5010 HL7, FHIR, ICD9,10, HL7 v2.x,  HL7 C-CDA, |
| **EDI transactions** | EDI 835, EDI837, EDI834, EDI276, EDI277, EDI270, EDI271 |
| **Healthcare Claims and Payers** | FACETS , QNXT. |
| **NO SQL Databases** | Cassandra, HBase, MongoDB, |
| **Web Design Tools** | HTML, CSS, JavaScript, JSP, JQuery, XML |
| **Public Cloud** | EC2, IAM, S3, Auto scaling, Cloud Watch, Lambda, Glue, Kinesis, S3, EMR, Redshift, snowflake. |
| **Development Methodologies** | Agile/Scrum, Waterfall |
| **Reporting Tools** | MS Office (Word/Excel/Power Point/ Visio/Outlook), SSRS |
| **Operating Systems** | All versions of Windows, UNIX, LINUX. |

**PROFESSIONAL EXPERIENCE**:

**Banner Health, Remote Jan/2021-Current**

**ROLE: Healthcare Data Analyst**

**RESPONSIBILITIES:**

* Experience in building multiple **Data pipelines**, end to end **ETL** and **ELT** process for **Data ingestion** and transformation in **AWS** and coordinate task among the team.
* Worked extensively on **Kythera data** (voluminous **raw medical claims data**) by building several schemas to break it into various dimensions and facts.
* Point of contact for various Payer data feeds like **MMIT, DRG & CMS Medicare** and Claims datasets like **SHA** & **IQVIA LAAD**. Implemented various business solutions using these datasets right from summarizing the raw data.
* **Incentive Compensation (IC**) implementation using Payer data for the **Salesforce** team.
* Led initiatives to integrate new data sources into existing frameworks, ensuring seamless data flow and consistency across platforms.
* Expertise in summarizing and understanding the copay data from various feeds like **Sonexus** & **McKesson**.
* Built reports and data flows to automate copay tracking of various client copay programs on Monthly basis like Voucher Access programs etc. Created insights using copay and **claims data** for various clients.
* Outlined table structures, established relationships between entities, and implemented foreign key constraints in **relational databases.**
* Participated in cross-functional meetings to **align software** engineering efforts with broader business objectives and client needs.
* Managed the **EHR** improvement project following an **Agile methodology**, employing iterative development cycles to deliver incremental enhancements.
* Applied **Waterfall methodology** to develop detailed data analysis plans for clinical trials, ensuring comprehensive understanding of requirements and adherence to timelines.
* Conducted thorough analysis of organization's data architecture and interoperability frameworks, ensuring alignment with strategic business objectives.
* Implemented **ETL** processes using Informatica and **MS Azure** Synapse Analytics to efficiently transform and analyze trends and patterns in vast amounts of healthcare data.
* Collaborated closely with IT teams and vendors to analyze system capabilities, specifically focusing on **Medicare** and **Medicaid** claims processing requirements.
* Coordinated **user acceptance testing (UAT)** to ensure that the improved **EHR** system met quality standards and fulfilled the needs of end users before deployment.
* Created tailored integration roadmaps for **HL7** messaging standards, considering specific data exchange needs and regulatory requirements.
* Leveraged Pandas and NumPy libraries for data profiling and quality assessment before integration.
* Utilized **Python**, **R**, and **PySpark** to analyze large datasets, identifying trends and providing actionable insights, resulting in a **20%** increase in revenue.
* The Python **PySpark** package offers a potent distributed data processing architecture that enables effective analysis of big datasets.
* Created **Python Merge** tasks for data extraction and loading into MySQL databases.
* Developed custom **Tableau** dashboards presenting key performance indicators such as patient wait times and claim processing turnaround times.
* Conducted **COPQ** (Cost of Poor Quality) and **DPPM** (Defective Parts Per Million) analysis in **Power BI**, aiding Global Quality Leaders in making business-driven decisions.
* Trained junior associates on **PowerBI** and **SQL** Star schema concepts using **SHA** **claims** **datasets**.
* Implemented **row-level security (RLS)** in **Power BI** to monitor access to sensitive patient data.
* Applied **DAX** functions and **FHIR** to create measures and calculated fields for analyzing data, enabling logical visualization.
* Utilized **PostgreSQL** for data storage and management, optimizing performance and scalability.
* Developed interactive dashboards in Salesforce Einstein Analytics Studio using **SQL queries**, facilitating analysis of **CRM** data.

**TRACEGAINS, REMOTE Jan/2017 – Dec/2020**

**ROLE: Healthcare Data Analyst**

**RESPONSIBILITIES**:

* Worked on implementing scalable infrastructure and platform for large amounts of **data ingestion, aggregation, integration, analytics** in **Hadoop** using **Spark**, **Snowflake** and **Hive**.
* Supported data and application backbone for **Life Sciences** market access and commercialization.
* Provided strategic support in areas of growth strategy and drug launch excellence to senior leadership of **biopharmaceutical** clients.
* Developed and documented several **Request for Proposals (RFP)** and **Statement of Work (SOW)** for various client engagements.
* Carried out early adopter analysis for one of the drugs in commercial space which was in initial clinical trials phase.
* **HCP** Targeting & Segmentation: Co-managed a project to find the potential Physicians for our sales team to target to increase the sales of the drug.
* Managed Midsize team and developed various business solutions on **Patient Analytics**.
* Created a Model to standardize our Physician database which is coming from different Sources.
* Extensively worked on data related to sales, **claims**, calls, **patients**, **physicians & payers**.
* Derived Projection factors and calculated Projected claims based on Company sales and claims data.
* Led the **Data Steward** Team by taking complete responsibility of data loading as well as data governance activities.
* Designed automated scripts to identify data discrepancy as well as to perform data quality checks such as validating numbers shared by clients against production code output and sending alerts via email.
* Developed intricate **Python** scripts, incorporating **Lambda** functions to extract data from **AWS** **DynamoDB** in **NoSQL** format, transforming it into **JSON**, and inserting it into an **SQL** server.
* Built Visualization dashboards which are automated to test various voluminous incremental data feeds.
* Worked extensively on Marketing data and analyzed various sources of Marketing data like samples, emails, calls, sales, **NPP**, etc.
* Delivered Analytic solutions like Source of Business, Patient Persistency, Monthly Business Review, Call Adherence, Payer Contract Evaluation Tool**, Integrated Delivery Network (IDN)** Report, Non- Personal Promotion Tracker.
* Provided various insights to clients based on patient weekly claims, their statuses, and their respective payer analytics.
* Built several dashboards in **PowerBI** using **DAX** and power queries.
* Actively engaged in conversations with multiple clients to understand their business requirements and support them in achieving goals.
* Managed all the **Ad-Hoc** requests from the client.
* Responsible for end-to-end automation of reports.
* Worked on developing streamlined workflows using high-performance **API** services and dealing with large amounts of structured and unstructured data.
* Developed Spark jobs in **Python** to perform data transformation, creating Data Frames and **Spark** **SQL**, **Snowflake**.
* Developed **Spark** applications using spark libraries to perform **ETL** transformations and thereby eliminating the need for utilizing **ETL** tools. Developed the end-to-end data pipeline in spark using **python** to ingest, transform and analyses data.
* Created **Hive** tables using **HiveQL**, then loaded the data into b tables and analyzed the data by developing **Hive queries**.
* Executed Hadoop and Spark jobs on **AWS** **EMR** using data stored in **AmazonS3**. Implemented Spark RDD transformations to map business analysis and apply actions on top of transformations.
* Used **ALTERYX** to analyze the data using repeatable workflow. Experience in using **Alteryx** workflows and **ETL** processes.
* Provided SQL and stored procedure support to end-users and developers, assisting with query optimization and problem resolution.
* Implemented and managed database security protocols within stored procedures to control user access and protect sensitive information.
* Utilized dynamic SQL within stored procedures to create flexible and reusable code that can adapt to varying input parameters and conditions.
* Optimized stored procedures for performance, ensuring fast execution times and minimal resource usage through indexing, query rewriting, and efficient use of joins and subqueries.
* Created stored procedures for batch processing tasks, allowing efficient handling of large data volumes and scheduled data processing operations.

**CLIENT: Sun Pharmaceutical Industries Ltd, India Jan/2015– Jul/2016**

**ROLE: Healthcare Data Analyst**

**RESPONSIBILITIES**:

* Created comprehensive **BRDs** detailing stakeholder requirements, objectives, and scope for clinical trials data analysis projects, ensuring alignment with business goals.
* Developed **FRS** documents outlining the specific functionalities, features, and interactions required for data analysis tools and platforms within clinical trials contexts.
* Defined **NFRS** criteria, including data security, privacy compliance, performance benchmarks, and user

experience standards, ensuring high-quality outcomes.

* Integrated diverse clinical data streams such as patient records, lab results, and adverse event reports, from EHRs and flat files using data mapping techniques to load data to **Google Cloud Platform (GCP)** repository.
* Developed **data-driven strategies** for optimizing inventory levels, ensuring medical supplies are available when needed without overstocking.
* Analyzed supplier data to assess performance metrics such as delivery times, quality, and pricing, contributing to informed supplier selection and negotiation.
* Analyzed lead times for different supplies and identified opportunities to reduce lead times, improving **supply chain responsiveness.**
* Collaborated with suppliers to implement **data-sharing practices**, improving supply chain visibility and enabling better coordination.
* Ensured compliance with healthcare regulations and standards to maintain **supply chain integrity** and mitigate risks through thorough risk assessments.
* Conducted Exploratory Data Analysis using **Python** to uncover insights into patient characteristics, treatment modalities, and survival outcomes across different cancer types and stages.
* Developed predictive models using **Python (scikit-learn)** to identify prognostic factors and predict patient outcomes, such as overall survival, progression-free survival, and treatment response.
* Excellent in Data Extraction, Transforming and Loading (ETL) using **SQL Server Integration Services (SSIS)**, DTS, Bulk Insert, **BCP**.
* Strong working experience in the Data Analysis, Data Verification and Validation of an **ETL Applications** using Backend/Database Testing
* Strong experience in SQL queries on Database such as Oracle, SQL Server
* Possess good knowledge in working with Linux / Unix Environment using Putty (UI) for checking error log files.
* Expert in design and development of **SQL** Server Integration Services (SSIS) packages and using various control flow tasks and data flow transformation.
* Created Database Maintenance Plans for the performance of **SQL Server** which covers Database Integrity checks, update database Statistics and Re-indexing.
* Stored Procedures, User Defined Functions, Views, **TSQL** scripting for complex business logic. Extensively used Joins and **sub-queries** for complex queries involving multiple tables from different databases.
* Good understanding of Data Marts, Data warehousing, **Operational Data Store (ODS)**, **OLAP**, Star Schema Modeling, Snow-Flake Modeling, Fact and Dimensions Tables using MS Analysis Services.
* Implemented Tableau for visualizations and views including **scatter plots**, **box plots**, heatmaps, tree maps, donut charts, highlight tables, word clouds, reference lines, etc.
* Created dynamic dashboards and reports using Power BI and **Tableau** to provide real-time financial insights to senior management, enhancing **decision-making processes**.
* Generated complex reports in **Tableau** using multiple functionalities including context filters, hierarchies, parameters, **LOD** Expressions, time intelligence, etc.
* Built interacting and storytelling dashboards in Tableau workbooks using **dashboard** actions.
* Involved in migrating the Excel Power View Reports to POWER BI Designer Reports.
* Responsible for creating the reports based on user requirement in POWER BI Designer (POWER BI Desktop).
* Created KPI&#39;s, Scorecard&#39;s and Dashboard&#39;s using the POWERBI.
* Used **Power BI**, Power **Pivot** to develop data analysis prototype, and used Power View and **Power Map** to visualize reports.
* Published Power BI Reports in the required originations and Made **Power BI Dashboards** available in Web clients and mobile apps.
* Developed a comprehensive **data dictionary** to document supply chain data elements, facilitating a common understanding among **stakeholders** and defining clear guidelines for data collection, validation, and transformation to enhance overall data quality.
* Developed **Oracle Data Integrator** (**ODI**) mappings to extract inventory data from **Oracle SQL** database and other local files, applying **data transformation** rules to standardize and cleanse the data for analysis.