**Sivaji Ganesh Konka**

**Data Analyst**

**Halethorpe, MD | Mobile: (667) 292-5858 | Email:** konkasivajiganesh@gmail.com

**SUMMARY**

* 4 years of proven expertise in leveraging Data Analytics to enhance operational efficiency and drive informed decision-making in healthcare and finance sectors.
* Proficient in conducting detailed data analysis, leveraging skills in SQL, Python, and Excel to extract actionable insights from complex datasets.
* Adept at utilizing advanced tools such as Tableau, Power BI, and Excel (including Pivot Tables, VLOOKUP) for creating interactive dashboards and visualizations.
* Experienced in designing and maintaining ETL pipelines using SSIS and SQL to ensure a seamless flow of data for analysis and reporting.
* Expertise in managing diverse databases, including MySQL, MongoDB, Microsoft SQL Server, and Teradata, ensuring efficient data storage and retrieval.
* Proficient in integrating cloud technologies, including Microsoft Azure, AWS, and Google Cloud Platform, for scalable and flexible data analysis solutions.
* Experienced in project management using Jira, ensuring efficient task assignment, tracking, and collaboration for timely project deliveries.
* Strong communication and presentation skills, enabling the translation of complex analytical findings into actionable insights for both technical and non-technical stakeholders.

**SKILLS**

|  |  |
| --- | --- |
| **Methodologies:** | SDLC, Agile, Waterfall |
| **Programming Language:** | Python, Java, R, SQL |
| **Packages:** | NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, TensorFlow, Seaborn, Plotty,  TensorFlow, NLTK |
| **Visualization Tools:** | Tableau, Power BI, IBM Cognos, Advanced Excel (Pivot Tables, VLOOKUP) |
| **IDEs:** | Visual Studio Code, PyCharm, Jupyter Notebook, IntelliJ |
| **Database:** | MySQL, Redis, MongoDB, DynamoDB, PostgreSQL, Teradata |
| **Cloud Technologies:** | Microsoft Azure (ADF, SSMS, Blob Storage), AWS (EC2, RDS, S3), Google Cloud  Platform |
| **Other Technical Skills:** | Google Analytics, DAX, Power Query, Alteryx, SAS, JIRA, SAP, SSIS, SSRS, Machine Learning Algorithms, Probability distributions, Confidence Intervals, ANOVA, Hypothesis Testing, Regression Analysis, Linear Algebra, Advance Analytics, Data Mining, Data Visualization, Data warehousing, Data transformation, Data  Storytelling, Association rules, Clustering, Classification, Regression, A/B Testing, Forecasting & Modelling, Data Cleaning, Data Wrangling. |
| **Big Data Tools:** | Apache Hadoop, HDFS, Spark, Hive, Pig, Zookeeper |
| **Version Control Tools:** | Git, GitHub |
| **Operating Systems:** | Windows, Linux, Mac iOS |

**EDUCATION**

**Master of Science in Data Science** – University of Maryland Baltimore County, Baltimore, Maryland

**Bachelor in Electronics and Communication Engineering** – GITAM University, Vishakhapatnam, India

**EXPERIENCE**

**Data Analyst | CDC Foundation, Atlanta, GA Jan 2022 - Present**

* Spearheading a healthcare analytics initiative at CDC Foundation, leveraging advanced statistical techniques, and assisting machine learning algorithms to enhance claims processing accuracy by 15%.
* Conducting in-depth analysis of extensive healthcare datasets using SQL and Python, achieving a noteworthy 20% reduction in data processing time.
* Developing and implementing predictive models utilizing R, resulting in a significant improvement in forecasting accuracy for patient admission rates.
* Employing Tableau for data visualization, facilitating intuitive interpretation of complex healthcare metrics to support informed decision-making.
* Demonstrate a commitment to data security and patient privacy by ensuring strict adherence to HIPAA regulations in all aspects of healthcare data analysis. Implemented robust data handling protocols to safeguard patient data and maintain compliance with industry standards.
* Conducting data cleaning procedures to address inconsistencies, missing values, and outliers in EHR datasets.
* Applying advanced statistical techniques to identify patterns and trends within millions of patient records, enhancing overall data-driven insights.
* Conducting A/B testing methodologies to assess the impact of implemented changes, providing valuable insights into the effectiveness of optimization strategies.
* Developing and maintaining automated ETL (Extract, Transform, Load) pipelines, ensuring a consistent flow of updated and accurate data for ongoing analytics efforts.
* Integrating cloud-based solutions, leveraging technologies such as AWS, to scale and handle large datasets efficiently, optimizing storage and computational resources.
* Integrating Matplotlib and Seaborn for data visualization, creating insightful charts and graphs to communicate complex healthcare metrics effectively.
* Employing SQLAlchemy, an SQL toolkit for Python, to interact with databases and optimize database operations, enhancing the overall performance of data retrieval.
* Knowledge of safety data standards like CDISC, medical terminologies such as MedDRA and regulatory guidelines.
* Proven ability to collaborate cross-functionally with healthcare professionals, administrators, and technical teams, fostering effective communication and synergy between data analytics and healthcare domains.

**Data Integration Engineer| TCL CSOT, India Feb 2019 – Dec2021**

* Developed ETL pipelines to integrate structured data from ERP and MES systems with unstructured sensor data in JSON and XML formats to create a comprehensive manufacturing data lake on AWS.
* Built analytics to optimize manufacturing performance, including OEE dashboards tracking downtime, cycle times, and quality metrics to identify issues and improve processes.
* Created data models leveraging dimensionality reduction techniques to track KPIs including production counts, defect rates, equipment effectiveness, and material waste analytics.
* Proficiently Designed optimized star schema data models and snowflake architectures on cloud platforms like Azure to enable scalable and performant data analytics.
* Architected a scalable cloud data lake on AWS, leveraging Redshift, EMR, and S3 to combine structured ERP and MES data with unstructured JSON and XML data from sensors. Designed data models and schema to enable efficient analytics.
* Created data visualization dashboards and reports in Power BI to communicate key insights to executives and senior leadership. Developed KPIs to monitor manufacturing performance.
* Collaborated cross-functionally with business stakeholders, operations, quality, and data science teams to understand analytics requirements and provide technical expertise to design end-to-end data pipelines. Partnered with stakeholders to optimize infrastructure for advanced analytics.
* Developed metrics as code to track data pipeline performance and data quality, providing visibility into analytics systems for data-driven decision making.
* Successfully migrated from on-prem data warehouses to cloud data lakes, enhancing scalability while reducing costs by 40%.
* Built Developed SQL queries and Python scripts to efficiently ingest, transform, and process structured and unstructured data (JSON, XML) from multiple sources, including ERP, MES systems, and equipment sensors.
* Managed projects and tracked tasks using Jira. Participated in Agile ceremonies with cross-functional analytics, operations, and quality teams to deliver high-impact analytics projects.

**CERTIFICATION**

**Azure Data Engineer, Associate Level, Microsoft Certification**

**SQL Advanced, Hacker rank**