

# SHANTHAN KUMAR CHAKUNTA

☎ (667)-345-4196 ✉ [shanthankumarchakunta@gmail.com](mailto:shanthankumarchakunta@gmail.com) in [LinkedIn](#)

## EDUCATION

- University of Maryland, Baltimore County (UMBC), Baltimore May 2023  
*M.S Information Systems* GPA: 3.58/4.0
- Jawaharlal Nehru Technological University, Hyderabad, India. May 2017  
*B. Tech, Mechanical Engineering*

## SKILLS

**Programming Languages:** Python, Java, SQL, Oracle PL-SQL, PostgreSQL, Shell Scripting.  
**Databases:** MySQL, Oracle, SQL Server.  
**Big Data:** PySpark  
**Packages & other:** Scikit-Learn, NumPy, Pandas, TensorFlow, Matplotlib, Statistics, Git.  
**Data Visualization Tools:** Power BI, Tableau, Advanced MS Excel.  
**Cloud Services:** Amazon Web Services (S3, EC2, Lambda, EBS, Route53, IAM, Cloud Formation, Beanstalk).  
**ETL & Integration Tools:** Snowflake, SQL Server Integration Services, Snaplogic  
**Operating Systems:** Windows, Linux, MacOS

## WORK EXPERIENCE

- Data Engineer Intern | Multiplan Inc | Maryland, USA** Jun 2022 – Dec 2022
  - Spearheaded data ingestion enhancements in a claims processing system, adhering to Multiplan's best practices. Transformed healthcare data of 90M+ rows from source to target in Oracle database by testing through SQL queries.
  - Engineered ETL pipeline optimizations, achieving a 40% acceleration in data flow and seamless data build mutations. Utilized SQL for data retrieval and analysis, effectively supporting decision-making processes.
  - Consolidated Python scripts to facilitate data ingestion into a Hadoop-based data lake. Segregated data into raw, trusted, and analytical zones, enabling 50% more data accommodation and robust data infrastructure.
  - Developed rules-based exception handling using custom shell scripts, automating fail-safes. Streamlined script deployment, bridging QA to preproduction phases, and elevating operational efficiency.
  - Fostered stakeholder collaborations, focusing on business product refinement through the construction of inter-zone data pipelines, data modeling, and streaming data to analytical zones.
  - Successfully conceived and executed a Snowflake proof-of-concept, validating its feasibility and alignment with Multiplan's data infrastructure development goals. Managed high-volume data, handling millions of rows in both batch and streaming formats.
  - Maintained historical data of 20M rows in dimensional, fact tables using Slowly Changing Dimension.
- Data Analyst | Apple India Pvt.Ltd | Hyderabad, India** Nov 2019 – Aug 2021
  - Collaborated with cross-functional teams within the Siri Organization to elevate user experience. Employed Python, SQL, and a Grading Tool to conduct rigorous data evaluation and analysis on AI and ML technologies.
  - Engineered ETL and data migration pipelines to transfer disparate data sources into a centralized data warehouse. Ensured seamless data analytics by solving complex data problems through scalable pipeline ingestion.
  - Accomplished 25% improvement in data retrieval efficiency by optimizing SQL queries on an Apple cloud environment, managing 100 million rows of data. Addressed the challenges of transforming and validating voluminous datasets.
  - Navigated data conversion challenges as a core team member, analyzing 500 to 700 transcriptions and false triggers daily. Achieved a direct boost in Siri's capabilities and led innovative technology solutions for data complexity.
  - Achieved a 25% improvement in data retrieval on datasets comprising 100 million rows in an Apple cloud environment.
  - Successfully analyzed 500 to 700 transcriptions and false triggers daily, enhancing Siri's functionality.
- Junior Data Analyst | Amazon Development Centre India Pvt.Ltd | Hyderabad, India** Oct 2018 – Jan 2019
  - Orchestrated cross-functional collaborations using Power BI and Tableau to optimize North American logistics, adeptly handling high-volume data inflows during peak seasons for accurate forecasting.
  - Deployed SQL queries on Amazon Redshift for real-time transportation analysis, resolving 100 logistical issues daily and implementing strategies that notably boosted sales during peak periods.
  - Developed business-focused Power BI dashboards to identify lost deal causes and mitigation steps, while partnering with Relay Operations Outbound teams to maintain real-time dashboards and stakeholder information flow.
  - Handled both batch and real-time streaming data to monitor sales, on-time delivery, and customer satisfaction, implementing key performance indicators (KPIs) that directly influenced business strategy and outcomes.
  - Adapted to the complexities of predicting logistical bidding, leveraging Power BI and Tableau visualizations to make data-driven decisions from millions of data points, ensuring timely delivery and heightened customer satisfaction.
  - Utilized dashboards to track monthly net sales trends, offering valuable insights for data-driven decision-making. Achieved precise real-time reflection of strategic decisions on operational metrics, including sales and product delivery.
  - Developed analytical dashboards on the churn rate of customers in various quarters of the year using Tableau.

## CERTIFICATIONS

- AWS Certified Cloud Practitioner Certification