YUVA KRISHNA THANNERU

+1 (945)-208-0433 | t.yuvakrishna@gmail.com | [LinkedIn](#) | [Github](#) | [Portfolio](#)

SUMMARY

Dynamic and results-oriented Data Analyst with proven expertise in data engineering, statistical analysis, and business intelligence. Proficient in leveraging tools like Python, SQL, Excel, Tableau, and Power BI to drive business outcomes. Adept at optimizing ETL workflows and managing large-scale datasets on cloud platforms such as AWS and Azure, while applying Six Sigma and Lean methodologies for process optimization. Experienced in predictive modeling, A/B testing, and building impactful dashboards, ensuring data-driven decisions.

TOOLS & TECHNOLOGIES

Languages: SQL, R, Python, C/C++, HTML/CSS, Java
Cloud Technologies: Google Cloud Platform, AWS, Microsoft Azure
Databases: PostgreSQL, MySQL, MS SQL server, NoSQL Datastores (Cassandra, MongoDB)
Statistical Techniques: Regression Analysis, A/B Testing, Hypothesis Testing, Time Series Analysis
Tools: MS Excel (VBA, pivot tables, macros, array functions), Tableau, Power BI, Snowflake, SAS, Alteryx
Certifications: Google Data Analytics Professional Certificate, AWS Solutions Architect Associate, Data Analysis and Visualization with Power BI, Snowflake Snowpro Core Certification

EXPERIENCE

Data Research Analyst | Northern Illinois University   **Aug 2023 - Present | Chicago, IL**

- Designed and implemented a real-time data pipeline using **AWS Lambda**, S3, and **Apache Airflow**, processing over **10,000+** posts daily from Reddit to extract mental health insights, improving automation and scalability by 50%.
- Analyzed user behavior and language patterns using **NLP** libraries like **spaCy** and **NLTK**, applying techniques such as **sentiment analysis** and **topic modeling** to identify early signs of mental health concerns in social media users.
- Automated data ingestion and preprocessing workflows by implementing ETL pipelines in **AWS Glue**, ensuring seamless data integration across structured and unstructured datasets reducing querying time by 30%.
- Built and deployed machine learning models with **TensorFlow** and **Scikit-learn**, analyzing language patterns and predicting user engagement with a 30% improvement in prediction accuracy.
- Created dynamic interactive dashboards levaraging **Power BI**, visualizing key performance metrics **KPIs** which improved data interpretation by 36%, enabling stakeholders to make data-driven decisions.

Data Analyst | Wipro, Client: Citi Bank   **June 2020 - June 2022 | India**

- Engineered a data warehousing solution in **Snowflake**, managing over **10TB** of structured and semi-structured data, reducing query times by 40% and enhancing business intelligence capabilities.
- Designed and maintained advanced **MS Excel** dashboards with automated calculations, and features like **PivotTables**, **Power Query** and **VBA macros** to analyze and visualize financial metrics, improving reporting efficiency by 45%.
- Automated repetitive data preparation tasks using **Alteryx**, reducing manual effort by 12+ hours per week and ensuring consistent reporting accuracy. Executed **A/B testing** on financial products, yielding a 20% increase in customer retention and improving revenue predictability.
- Performed Exploratory data analysis (EDA) over **1 million+ transactions** of financial data using SQL and Python to generate actionable insights and and minimize payment-related defects by 15%.
- Conducted process mapping and applied **Lean Six Sigma principles** to identify inefficiencies in financial workflows, reducing transactional errors by 18% and increasing process throughput by 12%. Mentored 10+ junior engineers on best practices for API design, query optimization, and data modeling, improving team performance by 36%.
- Developed **Tableau** dashboards to visualize financial **KPIs** and provide critical insights into transaction workflows, enabling informed decision-making for 25+ financial products.

PROJECTS

YouTube Data Analysis | AWS, Python, ETL (Extract, Transform, Load) Pipeline

- Designed and implemented a comprehensive data engineering pipeline for YouTube ad campaigns: Ingested **500,000** video records into Amazon S3, performed ETL processes with AWS Glue, and used AWS Lambda for data normalization. Reduced query times by 70% with AWS Athena. Developed an interactive dashboard in **QuickSight** to visualize YouTube metrics, significantly enhancing ad campaign effectiveness.

E-commerce Data Analysis | *Python, faker, K-Means, PCA, Machine Learning*

- Generated a realistic dataset of **5,000 transactions** with 100 customers and 150 products for robust e-commerce analysis. Applied the **Apriori** algorithm to uncover product associations for targeted marketing, and implemented K-Means clustering for customer segmentation, enabling personalized recommendations.

EDUCATION

Masters in Computer Science, Merit Scholarship Student   **Jan 2023 - May 2024**

Northern Illinois University, DeKalb, IL   *CGPA - 3.71/4*