**SIDDHARTHA KOMAKULA**

**+1(980) 428-6225 | skomakula09@gmail.com |** [**Linkedin**](https://www.linkedin.com/in/siddhartha-komakula/) **|** [**Github**](https://github.com/isidkoma/final3.git) **|** [**Portfolio**](https://isidkoma.github.io/)

**Professional Summary:**

* Data Analyst with over 7 years of experience delivering impactful insights through data analysis, visualization, and reporting across diverse industries including healthcare, finance, and retail.  
  Proven expertise in writing complex SQL queries, creating stored procedures, and optimizing large relational databases (MySQL, PostgreSQL, MS SQL Server) to support scalable data operations.
* Highly proficient in Excel with hands-on experience in advanced functions such as PivotTables, Power Query, INDEX-MATCH, and VBA macros for automating repetitive reporting tasks.
* Strong programming skills in Python for data wrangling, statistical modeling, and automation, using libraries like Pandas, NumPy, Matplotlib, and Seaborn to analyze large and complex datasets.
* Designed and deployed self-service dashboards using Power BI, Tableau, and Looker Studio, enabling business stakeholders to monitor KPIs, track performance, and make real-time decisions.
* Deep understanding of data modeling techniques, including normalization and dimensional modeling (star and snowflake schemas), to ensure data consistency and reporting efficiency.
* Experienced in building and maintaining ETL pipelines using tools like Informatica, Talend, Azure Data Factory, and Python scripts to automate data ingestion and transformation from multiple sources.
* Integrated structured and semi-structured data from APIs, flat files, spreadsheets, and cloud storage systems (AWS S3, GCP BigQuery) to create unified and trustworthy datasets for analytics.
* Familiar with applying machine learning techniques such as linear regression, logistic regression, clustering, and decision trees using Scikit-learn to build predictive and classification models.
* Implemented data governance practices by maintaining data dictionaries, defining standardized metrics, and executing validation rules to ensure high data quality and trustworthiness.
* Effectively collaborated in Agile/Scrum environments, contributing to sprint planning, user story refinement, and cross-functional communication with engineering, product, and business teams.
* Provided mentorship to junior analysts, conducted training sessions on BI tools and SQL best practices, and championed analytics best practices across departments.
* Continuously upskilling through certifications in Google Data Analytics and Advanced Microsoft Excel, and currently exploring dbt, Databricks, and modern data stack tools to stay ahead in the evolving analytics landscape.

**Technical Skills**

Methodologies : Agile, Scrum, Waterfall, SDLC, CRISP-DM, Data Governance, KPI

Frameworks, Data Storytelling

Programming Languages : SQL (Advanced), Python (Pandas, NumPy, matplotlib, Seaborn), R,

DAX, VBA, Shell Scripting, Regex

BI & Visualization Tools : Power BI, Tableau, Looker Studio, Qlik Sense, SSRS, Excel

(PivotTables, Power Query, VBA), matplotlib, Plotly

Data Management & ETL : SSIS, Alteryx, Apache Airflow, Informatica, dbt, Talend, Excel

Power Query, Fivetran, Airbyte

Databases : SQL Server, PostgreSQL, MySQL, Oracle, MongoDB, Google

BigQuery, Snowflake, Amazon Redshift, Teradata.

Cloud Platforms : AWS (EC2, S3, RDS, Lambda, Glue, Athena), GCP (BigQuery,

Cloud Storage, Dataflow), Azure Synapse Analytics

Analytics & Modeling : EDA, Time Series Forecasting, Regression, Classification,Clustering,

Hypothesis Testing, A/B Testing, Predictive Modeling

Tools & Collaboration : Jupyter Notebook, Git, GitHub, Bitbucket, JIRA, Confluence,

ServiceNow, SharePoint, Teams, Slack

Advanced Tools : Databricks, Docker (Basic), SAP BO, DataRobot, MLflow

(Introductory), AutoML, Google Colab

Other Tools : VS Code, Anaconda, Notion, Miro, PowerPoint (Advanced

Visualizations), Google Sheets

**Professional Experience**

**Client : Wells Fargo, USA Feb 2025 - Present**

**Role : Data Analyst**

* Analyzed large-scale financial datasets from internal banking systems to identify patterns in customer behavior, loan performance, and fraud detection, driving data-backed decision-making.
* Developed interactive dashboards in Power BI and Tableau to track KPIs such as loan default rates, customer churn, and risk exposure, enabling faster executive reporting cycles.
* Automated end-to-end reporting workflows using SQL, Python, and Excel VBA, reducing manual effort by over 40% and improving accuracy across monthly reporting tasks.
* Wrote optimized SQL queries and built complex stored procedures across SQL Server and Oracle databases to support real-time analytics and regulatory reporting.
* Built ETL pipelines using Informatica and Python to integrate data from APIs, Excel files, and cloud storage, ensuring accurate ingestion of financial and transactional records.
* Collaborated with cross-functional teams including finance, compliance, and risk to translate business needs into data models, ensuring alignment with regulatory requirements.
* Applied statistical models (logistic regression, clustering) using Scikit-learn to support risk scoring and segmentation strategies for loan applicants and credit card users.
* Conducted data validation and enforced data governance standards by maintaining metadata, building data dictionaries, and defining standardized business metrics.
* Participated in Agile sprints, provided estimation for analytics tasks, and contributed to user story grooming and backlog prioritization with product owners and tech leads.
* Mentored junior analysts on SQL optimization and Power BI best practices, promoting a culture of continuous learning and data literacy within the analytics team.

**Client : Citi Bank ,USA Aug 2023 - Dec 2024**

**Role : Data Analyst**

* Developed and maintained robust SQL-based reporting solutions to track credit risk, loan performance, and customer segmentation across retail and commercial banking units.
* Created executive dashboards using Power BI and Tableau to visualize portfolio health, delinquency trends, and operational efficiency, enabling leadership to make faster, data-driven decisions.
* Led the migration of legacy reports from Excel to modern BI platforms, incorporating Power Query, DAX measures, and row-level security to support self-service analytics across departments.
* Automated monthly financial reporting processes using Python (Pandas, NumPy) and VBA, reducing report turnaround time by 60% and minimizing manual errors.
* Designed and implemented ETL pipelines with SSIS and Informatica to consolidate data from internal systems, cloud platforms, and external vendors (e.g., Bloomberg APIs).
* Applied predictive models (logistic regression, decision trees) to forecast default probabilities and customer churn, integrating results into BI dashboards for actionable insights.
* Partnered with risk, compliance, and product teams to define KPIs and ensure alignment of analytics efforts with Basel III and CCAR regulatory reporting standards.
* Ensured data integrity and audit readiness by applying data governance principles, including the development of data dictionaries, validation rules, and metric standardization.
* Participated in Agile ceremonies (sprint planning, retrospectives, daily stand-ups), collaborating closely with developers, product owners, and data stewards to deliver scalable solutions.
* Provided knowledge sharing and mentorship to team members on SQL optimization, BI visualization best practices, and financial data modeling techniques.

**Client: UnitedHealth Group, Minneapolis, MN Apr 2021 – Jul 2023**

**Role: Senior Data Analyst**

**Responsibilities:**

* Managed the extraction, transformation, and analysis of large healthcare datasets including medical claims, eligibility records, and provider networks using SQL (SQL Server, Teradata) to support compliance reporting, quality improvement, and population health initiatives.
* Built and maintained interactive dashboards using Power BI to track vital healthcare KPIs such as hospital readmission rates, claims denial patterns, HEDIS scores, and preventive service utilization, enhancing visibility for operational and clinical decision-makers.
* Conducted exploratory and statistical analysis with Python (Pandas, NumPy, Seaborn) to identify high-cost utilization patterns, detect care gaps, and flag anomalies in patient outcomes across commercial and Medicare Advantage populations.
* Partnered with care management and clinical quality teams to gather reporting requirements, interpret business questions, and deliver tailored data solutions aligned with value-based care strategies and regulatory mandates.
* Developed automated Excel-based tools using Power Query and PivotTables to streamline production of monthly and quarterly compliance reports, reducing manual effort and error rates.
* Actively contributed in Agile ceremonies, including sprint planning, backlog grooming, and retrospectives, ensuring analytics tasks aligned with evolving business needs and delivery timelines.
* Applied dimensional modeling techniques to design efficient and scalable reporting datasets, significantly improving dashboard responsiveness and query performance.
* Supported the development of a 30-day hospital readmission risk model by preparing and transforming claims and discharge data, engineering relevant features, and evaluating model accuracy using classification algorithms in Scikit-learn with an achieved ROC-AUC score of 0.81.
* Visualized model outputs and risk categories in Power BI dashboards used by clinical care managers to prioritize outreach and interventions for high-risk patients, improving care coordination and resource planning.
* Led root cause analysis of anomalies in quality metric submissions and implemented solutions to improve data pipeline reliability.
* Conducted hypothesis testing and statistical validation of population health interventions, presenting findings to leadership and suggesting data-driven optimizations.

**Client: Accenture, Hyderabad, India Nov 2019 – Mar 2021  
Role: Data Analyst**

**Responsibilities:**

* Assisted in extracting and transforming large datasets from internal systems using SQL, supporting recurring client and operations reports.
* Created and updated Excel-based reports and dashboards using PivotTables, Power Query, and VBA macros to automate repetitive reporting tasks, reducing manual effort by 25%.
* Supported senior analysts in developing visual dashboards using Power BI and Tableau to track project KPIs, resource allocation, and operational metrics.
* Collaborated with business stakeholders to understand reporting requirements and documented report logic, filters, and data sources for internal knowledge sharing.
* Helped clean and structure raw business data using Python (Pandas) for quick ad-hoc analysis and visual summaries.
* Participated in data quality checks and validation activities to ensure data integrity in client deliverables.
* Engaged in weekly project meetings and sprint reviews, using tools like Excel and OneNote to take notes, track action items, and support project documentation.
* Explored basic process automation opportunities by identifying repetitive data handling tasks and suggesting improvements with VBA and Power Query.
* Received hands-on exposure to client-focused data analytics projects, learning how insights are translated into business decisions in a consulting environment.

**Client: Equinix, Bengaluru, India Mar 2019 – Oct 2019  
Role: Data Analyst Intern**

**Responsibilities:**

* Assisted in gathering, cleaning, and organizing structured data from internal sources using SQL and Google BigQuery to support reporting needs.
* Helped design and update interactive dashboards in Tableau and Power BI, allowing stakeholders to monitor KPIs and operational metrics.
* Supported basic data analysis tasks using Python (Pandas, NumPy) to identify patterns, summarize trends, and highlight anomalies in business data.
* Participated in project planning sessions with product and business teams to understand data requirements and translate them into reporting tasks.
* Documented data pipelines and dashboard logic clearly for internal knowledge sharing and onboarding new team members.
* Used Jira to track task progress, report blockers, and update status during agile sprint meetings.
* Conducted quality checks on reports and dashboards to ensure data accuracy and consistency across sources.
* Collaborated with cross-functional teams on business analysis tasks, supporting research on customer behavior and market trends.
* Received hands-on exposure to Google Cloud Platform (GCP) tools for data storage, querying (BigQuery), and basic cloud data workflows.
* Contributed to product lifecycle documentation by maintaining data tracking sheets and supporting project coordination tasks.
* Created Excel reports using PivotTables, VLOOKUP, and Power Query to automate routine reporting needs for the analytics team.

**Education:**

Master of Science, Information Technology | University of North Carolina at Charlotte, USA