# **SWINIDHI MANCHIGANTI**

swinidhi100@gmail.com | 785-215-9057 | [LinkedIn](http://www.linkedin.com/in/swinidhi-m) | California

## PROFESSIONAL SUMMARY

* Passionate data-savvy professional with 7+ years of experience as a Data Analyst, skilled in advanced SQL, Python, Scala, R, and visualization tools like Tableau, PowerBI, D3 Library, and Advanced Excel.
* Expertise in Complex Data collection, Data Transformation, Data Validation, Data Modeling, Data Visualization and Reporting actionable insights, recommendations to support data-driven decisions.
* Experience in developing complex SQL queries and robust data models to optimize data storage and retrieval, providing actionable insights and metrics from high-volume datasets scalable to petabytes that informed key business decisions and improved overall operational efficiency and performance.
* Experience in leveraging big data tools like Spark, Hadoop for large-scale data processing, enhancing the analytical capabilities and the handling of complex datasets in terabytes to petabytes efficiently.
* Experience in cloud-based data warehouse systems like Amazon Redshift, Snowflake, GCP BigQuery to enhance data storage, processing, and retrieval ensuring improved scalability and performance.
* Implemented advanced statistical models and A/B tests to evaluate performance, resulting in measurable product enhancements and increased user engagement across various domain projects.
* Extensively worked on Data Migration, Data Integration in both RDBMS and Cloud infrastructure.
* Experience in designing and implementation of data warehouses and automated ETL/ELT pipelines, leveraging advanced SQL, Python, Scala and cloud technologies, resulting in scalable, efficient data processing solutions and significantly reduced manual data handling, cleansing and manipulations.
* Expertise in analyzing and interpreting daily metrics to identify patterns, trends, and areas for improvement, driving continuous enhancement, optimization of product features and performance.
* Designed and maintained interactive dashboards using Tableau, PowerBI and D3, facilitating real-time reporting and making data accessible to stakeholders, supporting informed strategic planning.
* Implemented data validation and quality assurance processes, maintaining a high level of data accuracy and reliability, ensuring all metrics and reports dependable for making business decisions.
* Collaborated closely with cross-functional teams, data engineers, scientists, product and strategy teams and business stakeholders, to define project requirements and metrics, ensuring that data solutions aligned with business objectives and met user needs effectively for informed decisions.
* Expertise in developing and presenting detailed analytical reports translating complex data into clear, actionable business insights, to both technical and non-technical users’ decision-making.
* Established seamless data connections between various data sources and visualization tools, enabling real-time reporting and facilitating accurate and timely dashboards for stakeholders.
* Expertise in implementing data retrieval optimization techniques to optimize storage costs, improve query performance, and reducing execution times from hours to seconds and enhancing efficiency.
* Experience in developing documentation for data pipelines, table schemas, data models and dashboards, ensuring clarity facilitating future maintenance and onboarding of new team members.
* Profound experience in Ad-hoc analysis and reporting, and in extracting, aggregating data by creating complex SQL Queries and reporting the metrics & insights to support quick decision making.
* Standardized and maintained various complex data sources, ensuring compliance with data governance policies and procedures, implementing encryption techniques, access control policies, and regular audits and maintaining confidentiality and security of sensitive business data.

## TECHNICAL SKILLS

Reporting & Visualization: Tableau, PowerBI, D3 Library, Looker, SQL Server Reporting Services SSRS, Google Analytics, Advanced MS Excel (Pivot Tables, Charts, VLOOKUPs, Complex Formulas, VBA Macros)

Databases and Cloud Datawarehouses: AWS S3, Redshift, PostgreSQL, Spark, Hadoop HDFS-Hive, Presto, Sqoop, Flume, Oracle, SQL Server, Visual Studio, MySQL, Snowflake, MS Access, Google Cloud BigQeury

Scripting Languages: Advanced SQL, Python (Jupyter Notebook, boto3, psycopg2, tableauserverclient, Pandas, Numpy, SciPy, Matplotlib, Seaborn, Scikitlearn), Scala, VBA-Macros, DAX, JavaScript-html and R

Data Modeling: Kimball’s Fact and Dimensional modeling, Star Schema, Snowflake Schema Designing, Slow Changing Dimensions, Entity-relationship designing, Data Marts designing, DBT build tool

Statistical models: Probability Distributions, Estimation Techniques, Confidence Intervals, Multivariate Analysis, A/B Experimentations- Hypothesis testing, A/B Testing (Z-test, t-Test, Chi-Square, ANOVA)

Automation: Python, AWS Glue, Lambda, AWS Cloud Watch Events, PowerApps, Power Automate, Airflow, Flume, Sqoop, SAS Enterprise Guide, SAS Data Integration, SAS Rapid Miner, Git, Apache Airflow.

Predictive models: Linear Regression, Logistic Regression, ANOVA, Chi-Square, Decision Trees, KNN, K-means, Random Forest clustering, Principal Component Analysis, text mining and sentiment analysis

Data Ingestion and Migration tools: AWS DMS, Google Cloud DTS, Informatica, Sqoop, SQL Server SSIS

## PROFESSIONAL EXPERIENCE

*Data Analyst, Amazon* *Dec 2022 – Present*

Project Overview:

The Alexa Data Analytics and reporting project at Amazon is aimed to increase customer engagement and satisfaction by improving Alexa Voice Search command’s feature. As a Data Analyst, I was responsible for designing, developing and maintaining data pipelines from defining the key metrics, extracting, collecting, transforming, modeling, analyzing & visualizing the key metrics reports, reporting dashboards to enhance the Alexa experience driving actionable insights & product optimizations.

Tools and Tech Stack: Amazon Web Services (AWS) for cloud infrastructure, Redshift and PostgreSQL for data warehousing and AWS S3 for raw data storage. Data processing and Automation with Advanced SQL, Python, AWS Glue, AWS Lambda, AWS Cloud Watch Events, and custom APIs. Designed and Maintained Tableau, Jupyter Notebooks, and D3 interactive dashboards

Roles and Responsibilities:

* Responsible for Collecting, transforming, aggregating, analyzing, interpreting and visualizing voice, remote, and touch data to identify patterns and trends affecting search success rates and user satisfaction, using advanced SQL, Python, statistical techniques and data visualization tools.
* Defined 15+ metrics and KPIS, built data models, processed high volumes of data, & reported daily, weekly, monthly metrics providing insights to track performance, achieving a 99% data accuracy.
* Developed advanced SQL queries in Redshift, PostgreSQL to collect, ingest, transform, manipulate, model, aggregate the data to calculate key performance metrics, to inform product improvements.
* Designed and implemented a AWS Redshift data warehouse with optimized table schemas, ensuring efficient storage and retrieval of large datasets and supporting complex query requirements.
* Involved in Creating ETL/ELT pipelines using AWS Glue, Python, AWS Lambda, cloud-watch events, to automate the ingestion, transformation, and loading from multiple sources into datawarehouses.
* Conducted statistical modeling and analyses, hypothesis testing, A/B tests to evaluate the impact of feature changes on user behavior and satisfaction, driving product enhancements and optimization.
* Implemented robust data validation checks and continuous monitoring processes to maintain data quality by 99% and identify anomalies or outliers, ensuring the reliability of the metrics and reports.
* Established a seamless data connection between Redshift, PostgreSQL, and Tableau, enabling real-time reporting and facilitating 99% accurate and timely dashboards reporting for stakeholders.
* Designed and developed user-friendly interactive dashboards using Tableau, Jupyter Notebooks and D3 to visualize key metrics, providing data-driven actionable insights, optimizing product features
* Monitored and analyzed daily metrics to identify patterns, trends and areas for improvement in the Alexa voice search feature, driving continuous enhancement and optimization of the product.
* Collaborated closely with data engineers, data scientists, strategy team and product managers to define project requirements and metrics, ensuring alignment with business objectives & user needs.
* Conducted regular stakeholder meetings to gather feedback, adjust project goals and updated key performance metrics, and ensure that the output aligned with user expectations & priorities.
* Implemented data partitioning and compression techniques in Redshift to optimize storage costs and improve query performance, enhanced efficiency & execution time from 3 hours to 2 seconds.
* Utilized AWS Cloud Infrastructure for storing & extracting raw data, implementing lifecycle policies to manage data archiving and optimize storage costs, ensuring cost-effective data management.
* Applied data enrichment techniques to integrate additional user demographics and device information into the modeling/analysis, providing a more comprehensive view of user interactions.
* Conducted detailed data profiling to understand data distribution and quality, aiding in the design of efficient data models and improving data processing efficiency, data quality & accuracy by 99% .
* Scheduled and automated regular data updates and ETL processes using AWS Data Pipeline, Lambda functions, Cloud watch events and custom APIs, ensuring that the Redshift and PostgreSQL tables are consistently updated with the latest data ensuring up to-date and timely metrics reporting.
* Ensured 100% compliance with data privacy regulations by implementing data encryption techniques, access control policies, and regular audits, protecting user data and maintaining trust.
* Performed data normalization and transformation to standardize data formats across multiple complex datasets, improving data consistency by 99%, facilitating accurate analysis and reporting.
* Developed and maintained comprehensive documentation for data pipelines, table schemas, SQL queries, APIs, Tableau, and D3 dashboards, ensuring 100% clarity facilitating future maintenance.
* Provided training and support to team members on using Jupyter Notebooks, Tableau, and D3 dashboards & interpreting metrics, enhancing the team's ability to leverage data insights effectively.
* Continuously iterated on data models & reporting processes, refining the accuracy & relevance of insights, adapting to evolving business requirements and user feedback improving efficiency by 99%.

*Data Analyst, GE Healthcare Jan 2021 – Dec 2022*

Project Overview:

GE Healthcare Medicaid Data Ecosystem Enhancement Project aimed to improve the data infrastructure and analytical capabilities. As a Data Analyst, I was responsible for designing and building data warehouses, analytical solutions, and software tools to meet the evolving needs of Medicaid services, such as integrating third-party healthcare data, creating data marts, and developing visualizations to provide actionable insights, ensuring compliance, and fostering innovation.

Tools and Tech Stack: AWS with Snowflake for cloud-based data warehousing, AWS S3, Postgres, RDBMS Oracle, Big data tools Spark and Hadoop. Leveraged advanced SQL, SAS, Python and Data visualization tools Tableau, Powe BI & advanced Excel. Git for version control and collaborative development

Roles and Responsibilities:

* Designed and built data warehouses and analytical solutions to support new and emerging healthcare demands, ensuring scalability and performance optimization using cloud technologies and big data tools such as SQL, Python, SAS, Hadoop, Oracle, Tableau, and Power BI.
* Developed and executed ETL/ELT processes using SQL, Python, and SAS to ensure 100% efficient data collection, cleaning, transformation, and ingestion into data warehouses and visualization tool.
* Responsible for acquiring, cleaning, aggregating, analyzing, and visualizing large volumes of healthcare data from complex sources, reporting extracted metrics and insights to drive decisions.
* Created data marts, designed, and developed various visualizations and reports to meet the specific needs of healthcare stakeholders, enhancing decision-making and operational efficiency by 20%.
* Monitored and optimized data flow, 20+ data marts, and metadata management through interconnected applications to support seamless data movement and transformation.
* Collaborated with business users to source and collect third-party healthcare data types, ensuring 99% data quality and integration. Handling Third-Party Data, Integrated SHA claims, IQVIA LAAD, Xponent, and SHA Payer prescriber datasets to support healthcare analytics and reporting.
* Developed comprehensive documentation, functional requirements documents, data mapping documents, and end-to-end requirements traceability, to ensure aligned and seamless execution.
* Conducted in-depth analysis with interpretive thinking to define issues and problem areas, develop innovative solutions, and ensure 100% effective problem resolution and continuous improvement.
* Streamlined the data ecosystem by researching and developing new technology solutions, including cloud enablement and big data tools, fostering innovation, and improving operational efficiency.
* Employed critical thinking to create innovative and interactive narrative reports, presentations, and visual mappings that clearly communicated analytical assessments to non-technical stakeholders.
* Implemented robust data validation and quality assurance processes to maintain 99% data accuracy, consistency, and reliability across all data sources and analytical outputs and dashboards.
* Escalated identified risks and sensitive areas, ensuring mitigation of potential issues and resolution.
* Ensured compliance with applicable healthcare regulations and policies, safeguarding Blue Cross Blue Shield, its clients, and assets through sound ethical judgment and adherence to policies.

*Data Analyst, Northern Illinois University* *Jan 2019 – Dec 2020*

Project Overview:

Institutional Research Data Analytics and Metrics project aimed to support the Institutional Research Office by collecting, manipulating, analyzing, visualizing and reporting complex datasets, providing actionable insights and recommendations to improve academic programs and student success metrics.

Tools and Tech Stack: Data Warehousing and Management with SQL, Oracle, MySQL, PostgreSQL, REDCap, Amazon Redshift and Data Processing and ETL with Python, SAS, Apache Airflow, Pandas, Numpy, with Data Visualization tools PowerBI, advanced excel and machine learning libraries.

Roles and Responsibilities:

* Responsible for Gathering and consolidating data from multiple sources, student information systems and admissions databases, ensuring comprehensive datasets for analytical use cases
* Applied data cleaning techniques to remove inconsistencies, errors, and missing values, ensuring data integrity and reliability for analysis & reporting driving 99% data quality, accuracy, & timeliness.
* Integrated data from various sources using SQL and Python, creating a unified dataset supporting comprehensive, accurate and timely analysis and reporting improving operational efficiency by 30%.
* Employed advanced SQL and Python scripting to transform raw data into structured formats suitable for detailed analysis, performed data mining to extract insights developing impactful visualizations.
* Utilized SAS for conducting statistical analyses, including hypothesis testing and regression analysis, to derive meaningful insights from student data leading to an increase in student retention by 15%.
* Developed interactive dashboards and visualizations in Tableau to present key metrics and trends related to student admissions, enrollments, retention, and graduation rates with increase by 10%
* Leveraged advanced Excel for complex data manipulations, including pivot tables, VLOOKUPs, and advanced formulas, to support detailed reporting and analysis for non-technical stakeholders.
* Conducted trend analysis on student enrollment, retention data to identify patterns, predict future trends, aiding in strategic planning & decision-making that led to metrics improvements by 21 %
* Created performance metrics to evaluate the effectiveness of academic programs and student support services, providing calculated insights to improve institutional effectiveness by 33%.
* Analyzed curriculum data and student retention data to assess course offerings, student performance, and instructional methods, and factors influencing dropout rates and recommended strategies to improve educational outcomes, student retention and success rate by 21%.
* Developed enrollment forecasting models using Python and statistical techniques to predict future student enrollment trends and support internal decision making and resource allocation planning.
* Monitored and analyzed graduation rates, identifying trends and potential barriers to student completion, and provided recommendations to improve graduation outcomes by 18%.
* Created comprehensive reports and presentations to communicate findings to stakeholders, and

collaborated with them to understand their data needs, providing customized analyses and insights.

* Provided training and support to university staff on data analysis tools and techniques, enhancing their ability to use data effectively in their roles contributing to effective data-driven decisions.

*Data Analyst, The Home Depot* *Dec 2016 – Dec 2018*

Project Overview:

The project aimed to support The Home Depot's strategic objectives by leveraging data analytics to translate business questions into actionable insights. Responsible for Designing, implementing data analysis, tools, and reporting capabilities, cleansing, validating, and maintaining data sources.

Tools and Tech Stack: Data Warehousing with Google Cloud BigQuery, SQL, MySQL, PostgreSQL, REDCap

Data Processing and ETL: Python, Apache Airflow with Data Visualization tools: Tableau, PowerBI with

Other Tools and Technologies: RESTful APIs, Git, Docker, Unix/Linux environments, Microsoft Office Suite

Roles and Responsibilities:

* Responsible for gathering, collecting, modeling, manipulating, analyzing, presenting data, ensuring relevant & understandable insights for stakeholders to support in their data driven decision making.
* Interpreted and translated results from complex analyses into impactful actionable business insights
* Collaborated with business stakeholders, data scientists, and business analysts to deliver high-quality analytics solutions tailored to address specific business needs, objectives and use cases.
* Evaluated, processed, analyzed, and interpreted statistical data to support various business functions and strategic initiatives using advanced SQL, Google Cloud BigQuery, Python and Tableau.
* Developed scripts in PYTHON to automate repeated data manipulation. Saved the data processing time from 15 hours per week to 20 minutes. Utilized Pandas in Python to perform advanced lookups, calculations, cleansing & transformation to make the data more suitable for analytical purposes.
* Responsible for Cleaning, validating and testing the collected data to ensure accuracy, reliability, and quality for analysis and reporting ensuring 99% of data quality, integrity and timeliness.
* Utilized RESTful APIs to integrate external data sources, enriching datasets for accurate insights.
* Developed interactive dashboards and visualizations in Tableau and Advanced Excel to present key metrics and trends, enabling stakeholders to gain valuable insights and recommendations.
* Conducted trend analysis on sales, customer behavior, and inventory data to identify patterns and predict future trends, aiding in strategic planning and overall operational efficiency by 28%.
* Created performance metrics to evaluate the effectiveness of marketing campaigns, product launches, and operational efficiencies and identified areas of improvement and recommendations
* Conducted hypothesis testing and A/B testing to evaluate the impact of different business strategies and interventions and contributed to improvement and optimization of the features implemented.

## EDUCATION

Northern Illinois University, Master’s in Management Information Systems Dec 2020

Related Coursework: Big Data Analytics, Business Statistics, Predictive Business Analytics, Business Systems Analysis & Design, Social media Analytics, Advanced Database Management System. GPA-3.8

JNTUH College of Engineering Kukatpally Hyderabad, Bachelor of Technology June 2016