Nilay Soni

Bengaluru| 7999505321 | Nilaysoni25@gmail.com |**linkedin.com/in/nilay-soni-619724388**| https://github.com/nilaysoni25

Professional Summary

Data Analyst skilled in **Python (Pandas, NumPy), SQL, Advanced Excel, Power BI (DAX), and Tableau** for data cleaning, transformation, and dashboard development. Experienced in **Exploratory Data Analysis (EDA), A/B Testing, Segmentation, KPI Tracking, and Data Visualization** to drive data-driven decision-making. Hands-on exposure to **MySQL, Oracle, and Microsoft Azure (Azure SQL)** for cloud-based data management and analytics.

Technical Skills

- Programming : Python (Pandas, NumPy), SQL
- Data Visualization : Power BI (DAX), Tableau, Advanced Excel
- Databases : MySQL, Oracle
- Cloud : Microsoft Azure (Azure SQL)
- Analytics : Exploratory Data Analysis (EDA), A/B Testing, Segmentation, KPI Tracking
- Tools : Git, Jupyter Notebook, VS Code

Projects

Customer Analytics & Data Modeling (SQL)

- Analyzed 12+ months of acquisition & transaction data for 10K+ customers using advanced SQL (CTEs, joins, window functions) to evaluate engagement, spend behavior, and Customer Lifetime Value (CLV).
- Built data-driven customer segmentation and cohort models, identifying ~25% high-value customers contributing ~60% revenue, while reducing manual reporting effort by ~75% through reusable SQL datasets.

Burning Train Data Analysis (Python)

- Performed end-to-end data cleaning, transformation, and Exploratory Data Analysis (EDA) on train incident dataset using Python (Pandas, NumPy, Matplotlib, Seaborn) to identify casualty trends, cause patterns, and high-risk factors.
- Applied feature engineering, aggregation, and statistical analysis to uncover key safety insights, improving incident pattern detection by 30%+ and enabling data-driven risk assessment.

Banking Analytics & Business Intelligence (Python & Power BI)

- Analyzed 5,000+ banking customer records using Python (Pandas, NumPy) and SQL-based logic to build acquisition & engagement funnels, identify top 20% customers driving 60%+ revenue, and improve forecasting accuracy by 30% through KPI modeling.
- Developed interactive Power BI dashboards integrating 10+ datasets, enabling performance tracking across acquisition, retention and revenue metrics; adopted by 50+ stakeholders, reducing reporting time by 40%.

Azure Data Engineering End-to-End Project | Azure Data Factory | Databricks | Azure Synapse

- Designed and implemented an end-to-end data pipeline using Azure Data Factory, Azure Data Lake, Databricks (PySpark), and Azure Synapse Analytics, enabling scalable data ingestion, transformation, and warehousing for big data processing.
- Performed distributed data transformation using Apache Spark (PySpark) and built optimized analytics-ready datasets, improving pipeline efficiency and supporting real-time reporting and cloud-based data engineering workflows.

Work Experience

PySpiders | 6 Months

- Analyzed structured datasets using Python (Pandas, NumPy) and SQL, performing EDA, data cleaning, and KPI analysis to improve reporting accuracy by 30%+.
- Developed interactive Power BI dashboards (DAX) for performance tracking, customer segmentation, and trend analysis, reducing manual reporting time by 40%.
- Built optimized SQL queries using CTEs, joins, and aggregations, and supported cloud-based data storage using Microsoft Azure (Azure SQL) for streamlined data access and stakeholder reporting.

Education

MCA (Specialization: AI & ML) LNCT University, Bhopal – 77.7% / CGPA (2023–2025)

Certifications

- IBM Data Analytics – SQL & Excel, issued by IBM via Coursera
- SQL (Advanced) Certification – Hacker Rank
- Python Programming Certification – Infosys