

SUMMARY

- Accomplished Data Scientist boasts 7 years of background in system engineering, Big Data migration, and advanced data analysis. Combined with a Master of Professional Studies in Analytics from Northeastern University, specializing in Applied Machine Intelligence.
- Adept in spearheading a data migration strategy using Apache Sqoop, boosting smart meter installation tracking by 40%, and developing a payment defaulter prediction model that increased revenue by \$900K.
- Proficient in various tools and languages, including Python, R, MySQL, Tableau, Power BI, and Alteryx complemented by certifications like Tableau Desktop Specialist.
- Expertise in creating efficient data repositories and implementing Change Data Capture (CDC) methodologies at NCR Corporation, reducing data latency by 75%.
- Demonstrates knowledge of cutting-edge business intelligence and data warehouse solutions for top-tier clients like Uber Technologies and Meta, significantly enhancing decision-making and compliance processes.
- Volunteered for Non-Profit organization working at the intersection of clean air, climate and health to transform policies, educate communities, and advance solutions to clean our air and protect our climate.
- Expertise in corporate sustainability through OPF Academy, focusing on GHG accounting, LCA modeling, and decarbonization strategies. Enthusiastic about using data science skills to address climate change.

TECHNICAL SKILLS

Python | R | MySQL | Tableau | Google Colab | Power BI | Apache Hive | GitHub | Unix | Databricks | MLOps | Alteryx | Sqoop | Scikit-learn | PyTorch | NLTK | PySpark | spaCy | SciPy | Matplotlib | Seaborn | XGBoost | HuggingFace | Snowflake | Microsoft Azure data factory

AREAS OF EXPERTISE

- Data Science and Analytics
- Data Cleaning and Mining
- Big Data and Statistics
- GHG Accounting
- LCA Modeling
- Sustainability Reporting
- Decarbonization Strategy
- Data Visualization
- Machine Learning
- Feature Extraction
- Clustering
- Time series analysis
- Agile Methodology
- Brainstorming
- Business Recommendations
- Product Recommendations
- Storyboarding
- Natural Language Processing (NLP)
- Artificial Neural Networks (ANN)
- Convolutional Neural Networks (CNN)
- Data engineering

CAREER CONTOUR

Deloitte Risk and Financial Advisory, Gurgaon, India

Oct 2021 – Nov 2022

Lead Solution Advisor

- Designed and executed advanced Business Intelligence (BI) and Data Warehouse (DW) solutions tailored to meet the needs of high-profile clients such as Uber Technologies and Meta.
- Managed the maintenance and development of the Snowflake data warehouse, leading the migration from a legacy Oracle system.
- Worked extensively with Type 1 and Type 2 Slowly Changing Dimensions (SCD) within Snowflake to support accurate data tracking and analysis.
- Utilized specialized SQL scripts to identify and rectify missing invoices, ensuring meticulous compliance with reporting standards.
- Employed Tableau as the primary tool for generating detailed reports that informed strategic decision-making processes.
- Enhanced overall data visualization capabilities, improving client organization decision-making processes.

Tata Power Delhi Distribution Limited, Delhi, India

Aug 2019 – Oct 2021

Data Analyst - IT Business Analytics

- Spearheaded a strategic data migration initiative using Apache Sqoop, enhancing smart meter tracking efficiency by 40%.

- Designed and implemented a predictive analytics model for payment defaulters, increasing revenue by \$900K and cutting costs by 10%.
- Crafted detailed monthly usage reports for over 10,000 smart meters using Tableau, delivering critical insights into consumption patterns.
- Streamlined data analysis processes, boosting operational efficiency and enhancing decision-making capabilities.

NCR Corporation, Gurgaon, India

May 2018 – July 2019

Big Data Engineer - Supply Chain Analytics

- Engineered and set up structured data repositories in Hive, significantly enhancing data organization and accessibility for advanced reporting.
- Streamlined data storage designs, focusing on efficient data organization, facilitating quicker and more reliable data retrieval for analytical purposes.
- Implemented Change Data Capture (CDC) methodology to efficiently track and capture data modifications, reducing data latency by 75%.
- Boosted the efficiency of real-time analytics by optimizing data processing workflows, ensuring timely data availability and decision-making.

PREVIOUS EXPERIENCE

- System Engineer at Tata Consultancy Services, India | Sep 2014 – Aug 2017

PROJECTS

Advanced Customer Segmentation Strategy for Straumann LLC USA | June 2024

- Led the development of an advanced customer segmentation framework by applying cohort analysis, RFM models, and K-means clustering, resulting in more targeted and effective marketing strategies.
- Analyzed a comprehensive four-year transaction dataset to categorize customer segments, streamlining targeting efforts and improving customer engagement.
- Delivered actionable insights and strategic recommendations that enhanced digital marketing campaigns, expanding the customer base and driving overall business growth.

VOLUNTEER

Breathe Mongolia – Clean Air Coalition

Sep 2024 – Present

Air Quality Data Analyst

- Analyzing historical air quality data from multiple sensor stations across Mongolia.
- Development of interactive dashboard and reports with spatial air quality data analysis.
- Creating an air quality monitoring plan based on identified data gaps.

CERTIFICATIONS

- Certified as Tableau Desktop specialist | Aug 2020
- Certified as OnePointFive- Fellow - Sustainability Consulting Accelerator | July 2024 – Aug 2024
- Alteryx Designer Core Certification | Oct 2024

EDUCATION

- Master of Professional Studies in Analytics, Concentration in Applied Machine Intelligence from Northeastern University, Boston, MA | 2024
- Bachelor of Technology in Electronics and Electrical Engineering from Kalinga Institute of Industrial Technology, India | 2014