

Shubham Mahalank

linkedin.com/in/shubhammahalank
Portfolio: shubhammahalank.github.io

Data Scientist

shubhammahalank@gmail.com
+91 90199 77888

EXPERIENCE

Caterpillar Inc. | Data Scientist II

Oct 2022 - Present

(Python, Tableau, PowerBI, Statistics, ML, Azure, Snowflake)

Bangalore, KA, India

- Created & optimized Power BI dashboards using advanced DAX calculations & custom semantic models to streamline data flows, improve report accuracy, & enhance decision-making for global teams
- Developed data-driven solutions in Power BI, leveraging advanced DAX & data modeling techniques to identify inefficiencies, recommend improvements, & automate reporting processes, boosting overall team productivity
- Collaborated with cross-functional teams to migrate & integrate diverse data sources into Power BI, streamlining reporting processes & improving data accessibility for global stakeholders, resulting in a 20% reduction in report generation time & enhanced decision-making capabilities
- Gained proficiency in Power Apps & Power Automate, collaborated with the team to develop & implement the Target Setting Dashboard in Power BI
- Leveraged Snowflake for data management & optimization, contributing to the consolidation of multiple data pipelines using DBT (Data Build Tool), improving data accessibility, performance, and reporting efficiency across projects
- Applied statistics (clustering & anomaly detection) to enhance the Automated Fault Code (AFC) PSE (Prioritized Service Events) development, improving recommendation accuracy & driving \$15M in Closed Won revenue
- Developing ML Pipeline (Regression, Neural Network) to derive insights for Condition Monitoring & perform CVA (customer value agreement) incremental analysis
- Developed custom functions using the Tableau API to strengthen data governance practices within Tableau, ensuring standardized data access, consistency, & security. This initiative improved operational efficiency, & reinforced compliance with governance standards, supporting long-term data reliability across reporting workflows

Caterpillar Inc. | Data Scientist

Mar 2021 – Jun 2022

(Python, Tableau, Azure, Snowflake)

Peoria, IL, USA

- Leveraged Python (Pandas, SciPy) to streamline data collection processes, perform advanced analytics, thus reducing analysis time & enhancing data quality for better decision-making in the form of Tableau dashboard
- Applied statistical analysis (ANOVA, regression) & ML models to enhance the RAM Premium offering, streamlining data visualization & event tracking
- Engineered data-driven solutions, to refine templates & optimize server load, improving efficiency & quality in OSI PI AF
- Collaborated with MWM to integrate advanced analytics like anomaly detection into the RAM Premium platform, enhancing asset monitoring & fault prediction accuracy

EDUCATION

New York University, Tandon School of Engineering

Sep 2017 – Dec 2019

MS in Computer Engineering

New York, NY, USA

B V B College of Engineering & Technology

Sep 2012 – May 2016

BE in Electronics & Communication

Hubli, KA, India

SKILLS

Programming Languages:

- Python
- SQL (MySQL, PostgreSQL, Hive)
- NoSQL (MongoDB, Cassandra)

Libraries, Tools & Frameworks:

- Analysis (Numpy, Pandas)
- Machine Learning (Scikit-Learn)
- Deep Learning (Tensorflow, PyTorch)
- Spark (PySpark, SparkML)

- Visualization (Matplotlib, Seaborn, Plotly)
- AWS, Azure
- MS Power Platform (Apps, Automate, BI)
- Tableau, Snowflake, Docker, Git, Jira, Linux

Statistics: A/B Test, Hypothesis Test, Descriptive & Inferential Statistics, PCA

Machine Learning: Regression, Regularization (Ridge & Lasso Regression), Classification, Bagging (Random Forest), Boosting (AdaBoost, XgBoost), Clustering, Recommendation Systems, Anomaly Detection, Forecasting (ARIMA, FB Prophet)