

Susmitha Arikatla

+1(281)-750-6048|susmitha.usk7@gmail.com |[GitHub](#) | [Linkedin](#) | Houston, Texas

EDUCATION

- **Masters in Data Science** | University of Houston | Houston, Texas | GPA: 3.6/4.0 May 2023
- **Relevant Coursework:** Machine Learning, Statistics, Data Analytics, Database Management, Big Data, Time Series Forecasting, Artificial Intelligence, Data Visualization, Time Series Analysis, Data Analysis

WORK EXPERIENCE

Data Visualization Teaching Assistant | University of Houston Aug 2022 – May 2023

- Showcased the utilization of various software tools including Excel, R, Tableau to demonstrate the implementation of data science techniques and machine learning models.
- Generated visualizations for professors to use in class, increasing student comprehension by 17% in 4 months.

Data Analyst | Freelance | Remote India June 2020 – Apr 2022

- Conducted in-depth statistical analysis utilizing advanced techniques to identify trends, patterns, and anomalies in client data, resulting in actionable insights for decision-making. Increased efficiency by 20%, identified cost-saving opportunities worth \$1M.
- Implemented innovative data collection methods using SQL queries and Python scripting, enabling extraction, transformation, and loading of diverse data sources with utmost integrity and accuracy. Reduced data processing time by 50%, improved data quality by 25%
- Spearheaded the development of customized Tableau and Excel dashboards and reports, effectively communicating critical insights to stakeholders and improving decision-making speed by providing real-time data-driven insights that increased executive visibility into KPIs by 30%.

SKILLS

- Tableau, Looker, Power BI, Qlik Sense, Excel, Python, R, Data Analytics, Data Mining
- Microsoft Office, DAX, PowerPoint, SharePoint, Hypothesis Testing, SQL, SAS, ETL, Data Modeling
- Pandas, Numpy, TensorFlow, Keras, Seaborn, Matplotlib, Scikit-learn, PyTorch.
- AWS S3, MYSQL, Snowflake, DBT, Microsoft Azure
- Logistic Regressions, Decision Trees, Clustering, Neural Networks, Random Forests

PROJECT EXPERIENCE

Construction Safety Analysis using OSHA Dataset

Tech Stack: Numpy, Google Collab, Power BI, Excel

- Extracted and analyzed 100k records from the OSHA Website using advanced web-scraping techniques, resulting in enhanced data quality and improved understanding of safety trends within the construction industry.
- Utilized Principal Component Analysis (PCA) to effectively reduce the dimensionality of the dataset, resulting in more efficient data representation and analysis.
- Employed K-means Clustering technique to determine the optimal number of clusters (K) and identify distinct safety profiles among construction companies, allowing for targeted safety interventions and improvements.

Visualization of Data Scientist Job Salaries

Tech Stack: Snowflake, S3Bucket, Power BI, SQL

- Performed extensive data analysis on job salary data for data scientist positions, identifying key industry trends and patterns, resulting in actionable insights for optimizing compensation strategies.
- Utilized advanced cloud-based technologies such as Snowflake and S3Bucket to efficiently store and retrieve large volumes of job salary data, streamlining the analytical process by 30%.
- Designed and developed interactive dashboards using Power BI to visually depict the distribution of Data Science job salaries based on experience levels and job titles for 2023, facilitating easy interpretation of complex data by stakeholders.

Seoul Bike Sharing Demand

Tech Stack: Regression, Google Collab, MySQL, Python, Tableau

- Collected and aggregated data from various sources to create a comprehensive dataset for analysis.
- Cleaned, normalized, and engineered features in the dataset to ensure data integrity and suitability for modeling.
- Developed predictive models, including regression and machine learning algorithms, to accurately forecast the demand for bikes at different stations. Achieved 80% accuracy rate, significantly improving bike sharing demand forecasting in Seoul.

CERTIFICATIONS

- Google Data Analytics Professional Certificate
- Tableau Desktop Specialist (Tableau Software, LLC)
- Machine Learning-HarvardX,
- Python & Machine Learning for Financial Analysis