

# Karan Gupta

Bryan, TX

☎ +1 (979)985-7541 ✉ [karan\\_mays@tamu.edu](mailto:karan_mays@tamu.edu) [in linkedin.com/in/karan-gupta-](https://www.linkedin.com/in/karan-gupta-) [🏠 people.tamu.edu/~karan\\_mays/](https://people.tamu.edu/~karan_mays/)

## EDUCATION

**Texas A&M University, Mays Business School**

**College Station, Texas**

*Master of Science in Management Information Systems* | GPA: 4.0/4.0

**May 2021**

**College of Engineering Pune**

**Mumbai, India**

*Bachelor of Technology in Electronics and Telecommunication* | GPA: 3.05/4.0

**May 2016**

## TECHNICAL SKILLS & CERTIFICATIONS

**Certifications:** Machine Learning (IBM), KPMG Data Analytics Consulting Virtual Internship, Tableau Data Scientist

**Languages:** Python, R, Java, SQL, HQL, HTML5, CSS3, JavaScript, Shell Scripting

**Tools:** SSMS, SSIS, SSAS & SSRS, QlikView, Tableau 10.5, Power BI, Advanced Excel (VBA), Elasticsearch and Kibana

**DB & Frameworks:** MSSQL, Oracle 12g, MySQL, PostgreSQL, MariaDB Galera Cluster, Google BigQuery, MongoDB, Hive, Spark, Hadoop, Kafka, Amazon Redshift, Amazon RDS, Amazon Kinesis Analytics

## EXPERIENCE

**HDFC Bank Ltd., Mumbai, Maharashtra** | *Data Analyst*

**July 2016 – June 2019**

- Designed optimized queries and procedures in MSSQL to improve the response time of digital loan webforms by 30%; elevated customer satisfaction index of loans to 97%
- Successfully interpreted key metrics for deriving transactional summary from digital loan data; the improved insights and reports contributed to raising the digital loan business of the bank by 90%
- Architected ETL pipelines to integrate transactional data coming from varied data sources using Python scripts
- Developed real-time Tableau dashboards for improved visibility and analysis of daily support ticket data
- Performed EDA on IIS web traffic log of the primary website of the bank using Pandas; provided actionable business insights through custom KPI's which helped in reducing technical declines by 80%

**Tata Consultancy Services, Pune, Maharashtra** | *Summer Analyst*

**May 2015 – July 2015**

- Engineered OLAP cubes that aggregated data received from multiple sensors for an autonomous car project
- Developed MATLAB code for fusing sensor data

## PROJECTS

**Anomaly detection in Machine Operation using Time Series Analysis -**

- Analyzed time-series data for 20 expensive machines operating in 3 distinct modes: normal, faulty & failed
- Developed an automated method to pinpoint the time of fault induction using rolling maxima for preventative action

**Dominick's Fine Foods – Data Warehousing -**

- Developed a Data Warehouse for a large enterprise dataset using Kimball's design of Conformed Data Marts
- Designed a dimensional model for the data and utilized SSIS, SSAS and SSRS for ETL, analysis and visualization

**COVID-19 Real-time interactive dashboard –**

- Designed a real-time dashboard in Tableau Public, utilizing the dataset provided by John Hopkins Univ.
- Developed a pipeline in Python to scrape the data from JHU's GitHub, transform the data & load the transformed data into Sheets using Google API; automated pipeline execution using Airflow

**Movie Recommendation System –**

- Designed a movie recommender system using Content-based filtering and Collaborative filtering approach
- Combined recommendations generated using metadata soup of movie cast, director(s), genre(s) and keywords with recommendations generated using plot description for a more wholistic content-based recommender

**Equipment Failure Prediction [TAMU 2019 Datathon]–**

- Trained a classification model using SVM, to predict the equipment failure using data from 172 sensors
- Ranked top 20 for achieving an F1 score of 98.54% with the model
- Weighed different algorithms for accuracy – KNN, Naïve Bayes, Logistic Regression and Decision Trees

**Flight Guru –**

- Developed a web app for suggesting flights and destinations to users, through analysis of 8 years flight records
- Created, configured and managed a cluster of 3 MariaDB databases on an AWS EC2 instance
- Migrated data from relational to a document database (MongoDB); reduced query execution time by 20%