# SUMAIYYA FAREED

sumaiyya.fareed@gmail.com | (346) 521-9198 | LinkedIn | Github

## EDUCATION

**University of Michigan** — **August 2022 - April 2024**
Master of Science - Data Science | GPA: 3.8/4.0 — Dearborn , MI

**Birla Institute of Technology & Science Pilani** — **August 2017 - June 2021**
Bachelor of Engineering - Computer Science — Dubai , UAE

## SKILLS

**Data & Analytics Tools**: Looker, Tableau, Power BI, Excel Suite, BigQuery, Spark, Kafka , ETL, Snowflake, Data Wrangling
**Programming Skills:** Python, R, SQL (MySQL, SQLite, PostgreSQL), ETL , Hadoop, Spark
**Analytics Expertise:** Predictive Analytics, Customer Life-time Value Analysis, Marketing Analytics, Customer Segmentation, A/B test
**Certifications:** AWS Solutions Architect- Associate (SAA – C03)

## WORK EXPERIENCE

**GKN Automotive (GKN Driveline North America Inc.)** — **Sept 2023 – Oct 2024**
Data Scientist – Knowledge based Engineering — Auburn Hills , MI

- Improved data accuracy and consistency by 30% through robust reconciliation processes across multiple systems
- Enhanced data integration capabilities by developing workflows with dbt and BigQuery, reducing processing errors by 18%
- Boosted database performance by 15% by optimizing relational databases (PostgreSQL, MySQL) and leveraging NoSQL for unstructured data.
- Built and deployed linear regression models to analyze key performance issues in products, identifying actionable product insights

**Chalhoub Group** — **Sept 2021 - Aug 2022**
Data Analyst — Dubai , UAE

- Increased customer engagement by 17% with targeted marketing campaigns derived from segmentation of 1.2M+ customer dataset
- Designed scalable ETL workflows in collaboration with data engineers, ensuring seamless data integration optimized & infrastructure
- Enhanced marketing ROI by 15% through detailed data modeling and analysis of large datasets
- Built interactive dashboards using Tableau, Looker (LookML) and AWS Quicksight, to provide actionable business insights, enabling 500+ stakeholders to monitor KPIs in real-time and make data-driven decisions efficiently.
- Improved user engagement by 12% by implementing and analyzing A/B testing frameworks for product feature optimization
- Delivered impactful presentations and executive summaries, driving informed decision-making across cross-functional teams
- Evaluated brand-specific KPIs and customer segmentation metrics, achieving a 15% uplift in campaign effectiveness and improving future targeting strategies

**Emirates Hospital Group** — **Jan 2021 - Aug 2021**
Intern — Dubai , UAE

- Ensured compliance with data privacy regulations by aligning data handling practices with company policies
- Conducted impact analysis to identify inefficiencies, implementing improvements that reduced teamwork inefficiencies by 10%
- Supported medical staff by troubleshooting and resolving data issues within the hospital and clinics internal software

## ACADEMIC PROJECTS

**Feature based Recommendation system using Ecommerce Data | *Python & SQL*** — **Jan 2024 – April 2024**

- Built a recommendation system on using machine learning to deliver personalized product suggestions.
- Analyzed user behavior & product features to provide tailored recommendations, enhancing customer satisfaction and engagement.
- Demonstrated 20% increase in click-through rates during testing, showcasing the system's potential to drive sales & optimize the user journey.

**Prediction Analysis on rise in COVID-19 cases | *Machine learning & Statistics*** — **Aug 2020 – Dec 2020**

- Performed a time-series analysis of COVID-19 cases globally using data from Johns Hopkins COVID-19 Statistics.
- Tested machine learning models such as XGBoost, SVM, ARIMA, Polynomial Regression, and Bayesian Ridge Regression achieving up to 90% accuracy in forecasting COVID-19 case progression and identifying the most effective algorithms for trend prediction.
- Visualized case progression using advanced graphical analysis, enabling insights into pandemic trends and future projections.

**Heart Disease Prediction Modelling | *Python , SQL & Tableau*** — **Aug 2023 – Dec 2023**

- Examined heart disease data of 2M patients with high probability of experiencing heart attacks in future based on existing data.
- Predicted patients at highest risk by developing machine learning models (Random Forest , SVM) with 92% accuracy.
- Reduced processing time for data preparation by 30% through optimized SQL queries and Python scripts, enabling faster iterations.