

# SOHAM AGARWAL

West Lafayette, IN | (765) 327-6218 | sohamagarwal00@gmail.com | linkedin.com/in/soham-agarwal/ | www.sohamagarwal.com/

## PROFILE

- A **proactive data analyst and leader** with 1.5 years of client-facing experience and 2 years of startup experience. Led over 87 teams by discovering insights and creating strategies for problems in technology, consulting, and healthcare industries.
- **Programming Languages:** SQL, Python, R | Others: Tableau, Excel, Data Analysis, Data Mining, Statistics, Cloud (GCP/AWS), Scikit-learn, Machine Learning, ETL, Git, Minitab, SAS EM, Google Analytics, Digital Advertising.
- **Certifications:** Tableau Desktop Specialist (Tableau, 2023), Certified Analytics Professional (INFORMS), AWS Cloud Practitioner (Amazon, 2023), Azure Fundamentals (Microsoft, 2023), Operations Research with SAS (SAS, 2023).

## EDUCATION

**Purdue University, Daniels School of Business (GPA 3.76/4.0, Beta Gamma Sigma Honor)**

*Master of Science in Business Analytics and Information Management - President*

**West Lafayette, IN**

**August 2024**

**Mahindra University**

*Bachelor of Technology, Electrical Engineering*

**Hyderabad, India**

**June 2022**

## PROFESSIONAL EXPERIENCE

**Accenture**

**West Lafayette, IN**

**Student Data Analyst Lead (Contract via Krenicki Centre, Purdue)**

**November 2023 – August 2024**

- Led a team of 7 through strong communication in identifying problem areas across 800+ processes in SAP Signavio. Examined large process-level data to mine business-oriented insights and translate them for non-technical stakeholders.
- Built Python and SQL-based ETL pipelines to discover process blockers and risks for generating descriptive reports. Significantly reduced manual analysis from **4 hours to under 2 minutes** to meet Accenture's clients' KPIs efficiently.

**FAANG Company**

**West Lafayette, IN**

**Data Science Intern (Capstone Course, Purdue)**

**January 2024 – May 2024**

- Analyzed hundreds of SEC 10-K filings to translate complex financial data into investor-friendly information, enhancing accessibility for diverse experience levels. Communicated insights to head of AI services via reporting and visualizations.
- Developed and deployed a Gen AI chatbot with GCP on top of the 10-K documents. Displayed 60% higher accuracy than the base Text Bison 32k model, **potentially saving 13,300 man-hours annually** and improving decision-making efficiency.

**Capgemini**

**Bangalore, India**

**Senior Analyst (Full Time)**

**September 2022 – April 2023**

- Reviewed large datasets extensively using SQL for trends to inform the development of AI chatbots. Dramatically revamped customer support and attained a **36.4% increase in customer satisfaction** while cutting down manpower dependency.
- Engineered dashboard automation adopting Python and APIs, enabling data-driven decision-making for over 10 international stakeholders. Tested pipelines through sophisticated data analysis and communicated with storytelling through Tableau.

**Sailotech**

**Hyderabad, India**

**Data Analyst Intern**

**January 2022 – May 2022**

- Utilized Python algorithms and employed NLP for custom entity extraction, contributing to the delivery of the **'iCapture OCR software'** 5 weeks ahead of schedule. Leveraged SQL for querying and to create a database from extracted entities.
- Extracted KPIs and manipulated warranty data from Sailotech's client. Through experimentations constructed statistical models and classification algorithms to predict fraudulent claims, achieving over 96% accuracy.

**Save The Young Heart**

**Hyderabad, India**

**Marketing Analyst and Founder (Self Startup)**

**January 2020 – January 2022**

- Spearheaded a 9-person team, including doctors, for an online healthcare initiative with a budget of \$8,500+. Received 31,000 unique website visitors and 44,000 session hits, **reaching over 800,000 people** via Google and Facebook Ads.
- Used SQL and Tableau to create data visualizations from website data sources identifying profitable advertising channels.
- Rolled out an **'Interheart Risk Score'** Python model based on classification algorithms to recognize major combinations of factors of heart disease. In turn, drove thousands of high-risk patients for further investigation on the platform.
- Decreased email Cost-Per-Clicks through A/B tests from **\$2.58 to \$0.29** boosting company-wide revenue by over 104%.

## LEADERSHIP, PROJECTS, AND COMPETITIONS

- **President** of the MS-BAIM Program at Purdue University. Leading a cohort of 110 graduate students since 2023.
- Placed **2nd nationally, and 1st on Kaggle**, for the NCAA Crossroads competition comprising 67 teams and 280+ students. Presented key insights through Tableau dashboards to analyze ticket sales for the 2024 women's basketball tournament.
- Secured **1st place** in the CYC Case Competition (Purdue, 2023) by developing solutions to combat Purdue's housing crisis.
- Achieved **1st in the Purdue Data Mining Competition** and 2nd in Kaggle. Presented an ensemble pipeline in SAS using neural networks and gradient boosting to predict firm bankruptcies with a high AUC of 0.96.
- Led a team to **1st place at Purdue** and 4th out of 290 U.S. teams in the Data4Good competition. Utilized open-orca 7B Large Language Models to review 2,001 transcripts; reducing man-hours by 96% for every medical transcript parsed.