Soumya Shah

+ 1 (520) 5897446 • soumyashah14@gmail.com • linkedin.com/Soumya

SUMMARY

Analyst with 2+ years of experience in Business Intelligence and Data Analytics. Skilled in translating complex data findings into business solutions while collaborating across teams to deliver measurable business value

EDUCATION

M.S in Management Information Systems - University of Arizona, Tucson, AZ   **Aug 2023 - Dec 2024**
Coursework: System Analysis, Business Consulting, Data Mining, Business Intelligence   **GPA: 3.8**

B.Tech in Computer Science - University of Mumbai, India   **Aug 2019 - May 2023**
Coursework: Data Structures, Cloud Computing, Big Data Analytics, Data Warehouse Management   **GPA: 3.6**

PROFESSIONAL EXPERIENCE

Business Intelligence Analyst | Global Health and Body Composition Institute   **Jan 2024 - May 2024**
- Conducted comprehensive market research and analyzed 1000+ surveys to gather relevant data, identifying key customer segments and understand market trends for Go-to-Market strategy refinement
- Performed exploratory data analysis (EDA) on collected data, uncovering insights to guide marketing decisions
- Deployed ETL pipelines using Microsoft SSIS, integrating data from multiple sources into a centralized data warehouse in MS SQL Server to ensure accurate and timely data analysis
- Automated reports and dashboards using Power BI, tracking key performance indicators (KPIs) such as Cost Per Click, customer acquisition cost, and ROI across marketing channels
- Presented actionable insights and a detailed marketing plan to stakeholders, focusing on data-driven improvements for marketing strategies and ROI optimization

Data Analyst | Growth Ability Services   **Sept 2023 - Dec 2023**
- Led the design and implementation of a relational database using Oracle RDBMS and MySQL, streamlining data management across business operations and improving decision-making efficiency
- Conducted data modeling and transformation using MS-Visio/AWS Glue, ensuring data accuracy and consistency, leading to a 10% reduction in data discrepancies
- Deployed scalable ETL pipelines using SQL and Apache Airflow, automating data extraction and loading to improve integration and boost operational efficiency
- Implemented a real-time data interface on AWS using Lambda functions and S3 storage, reducing manual data handling by 25% and improving access to real-time business metrics

Data Science Intern | TechShots   **Jan 2023 - May 2023**
- Built and optimized data pipelines to preprocess and transform large datasets for a text summarization application, solving information overload and redundancy for business users
- Trained a Bi-directional Auto-Regressive Transformers (BART) model to accurately summarize key insights from large datasets, improving information extraction speed and accuracy
- Monitored engagement metrics by developing a Power BI dashboard with backend SQL queries to track retention, conversions and traffic sources enabling data driven evaluation of the application

Data Science Intern | Benchmark Computer Solutions   **Aug 2021 - Sep 2021**
- Developed an AI/ML-based Resume Parser with a focus on building scalable data extraction pipelines, automating resume data processing, and improving HR workflow efficiency
- Designed & deployed custom ETL processes via REST APIs to parse and display resume data, ensuring reliable performance through rigorous validation and integration checks
- Supported ad-hoc client tech requests by troubleshooting system issues, refining data extraction logic, and optimizing model performance to meet evolving business needs

PROJECTS

Customer Attrition Analysis | CitiBank
- Conducted data analysis on customer attrition using Excel, leveraging Pivot Tables, VLOOKUPs, and advanced functions to identify trends and key churn indicators
- Developed an ETL pipeline to clean, transform, and load structured customer data into Tableau, ensuring seamless visualization and reporting
- Designed interactive Tableau dashboards to track churn rates, uncover customer behaviour patterns, and provide data-driven recommendations to enhance retention strategies

TECHNICAL SKILLS

- **Core Skill:** ETL, Data Modeling, Data Warehousing, Data Integration, Data Analysis, Machine Learning
- **Programming:** Python, SQL, R, PL/SQL,
- **Data Visualization:** Tableau, Power BI, Google Analytics, Looker, Google BigQuery
- **ETL & Data Management:** SSIS, dbt, Apache Airflow, Snowflake, Oracle, SQL Server, Salesforce, AWS
- **Project Management:** Jira, Confluence, Asana, Monday.com, Slack, Azure Devops, Agile/Scrum Methodologies