

Shubham Dinesh Parulekar

shubhamdparulekar1@gmail.com | +1-217-480-2983 | [Linkedin](#) | [Github](#) | [Portfolio](#)

SUMMARY

Data & AI Professional with 3+ years of experience transforming complex data into actionable insights across finance and insurance. Skilled in SQL, Python, and BI tools with expertise in ETL automation, predictive modeling, and data analytics

WORK EXPERIENCE

Research Assistant | *School of Information Science, University of Illinois* August 2025 - Present

- Analyzed historical and cultural patterns across time periods by fine-tuning large language models, improving consistency and contextual accuracy by 25%.
- Designed and implemented a scalable data preprocessing pipeline to clean, normalize, and structure multi-source historical and cultural data totaling over 2 million documents.
- Presented model evaluation findings and progress to stakeholders, translating complex LLM behavior into actionable insights that shaped project direction.

Business Data Analyst | *Quantiphi Inc* July 2021 - August 2024

- Improved decision-making by analyzing insurance and claims datasets with over 1 million records using SQL and conducting exploratory and root cause analyses to surface key business drivers.
- Reduced reporting turnaround time from multiple days to 3 hours by automating ETL pipelines, monitoring workflows, and creating reusable analysis scripts, improving decision accuracy.
- Increased customer retention by more than 10% by analyzing transaction and CRM data, conducting A/B testing, and building dashboards to optimize win-back campaigns.
- Reduced incident resolution time by 40% by building proactive monitoring and alerting dashboards using Power BI, improving visibility into production systems and service health.
- Drove solution adoption for 50+ clients by translating complex data findings into executive-ready presentations and workshops, directly influencing project scopes and client decisions.

Business Data Analyst Intern | *Quantiphi Inc* January 2021 - July 2021

- Increased product relevance by conducting market research through 10+ discovery sessions with senior stakeholders and translating findings into product roadmap improvements.
- Enabled faster data-driven decisions by partnering with cross-functional teams to translate business questions into structured analytical solutions and ad hoc deep dives.

Data Analyst Intern | *BitGenie Technologies* April 2020 - July 2020

- Improved product prioritization, communicated key insights to non-technical leadership through visual storytelling, helping leadership identify the top features driving user engagement.
- Reduced reliance on engineering teams by designing reusable SQL queries and standardized reports, cutting turnaround time for recurring business questions by more than 50%.

EDUCATION

University of Illinois at Urbana-Champaign | *MS, Information Management* Aug 2024 - May 2026

Coursework: Statistics, Data Modeling, Data Science, Machine Learning, Data Warehousing, Business Intelligence, LLM

University of Mumbai | *Bachelor's in Electronics and Telecommunications Engineering* Aug 2017 - May 2021

TECHNICAL SKILLS

Languages: Python, R, SQL, JavaScript | **Tools:** PostgreSQL, MongoDB, Power BI, Tableau, Spark, Airflow, Docker, Git

Cloud Platform: AWS S3, Redshift, SageMaker, DynamoDB, Azure Data Factory, Databricks, GCP BigQuery, Snowflake

Certification: Microsoft Certified Power BI Data Analyst, Microsoft Certified Azure Developer, Professional Scrum Master.

PROJECTS

Strategic Drug Demand Forecast (*Python, PySpark, Scikit-Learn, PostgreSQL, Power BI, Excel*):

- Improved national-level drug demand planning by developing time series forecasting models on 10+ years of public data, capturing multi-year trends, and improving forecasts by approximately 20%.
- Increased usability of analysis for policy planning by collaborating with the Pan American Health Organization and the UIUC Center for Health Informatics to translate model results into actionable allocation scenarios.

Intelligent Research Assistant (*SQL, Streamlit, Pandas, Pytorch, Pyspark, Keras*):

- Developed an NLP-powered assistant that ingests research papers and automatically generates concise summaries, explanations, related work mappings, and citation lists.
- Enhanced literature review and research by automating multi-step reading workflows for students and researchers

AI-Powered Breast Cancer Detection (*TensorFlow, JavaScript, Django, HTML, CSS, MongoDB, Scikit-Learn, OpenCV*):

- Developed a deep learning model that achieves 92% accuracy in early breast cancer detection by detecting and analyzing cancer cells from mammogram images.
- Created a secure, responsive web application for uploading mammogram images and generating automated diagnostic reports, improving accessibility of predictive healthcare technology. Published the research in IJRASET