

# SANKET JANOLKAR

Salt Lake City, UT | (385) 461-9772 | janolkarsanket98@gmail.com | [LinkedIn](#) | [GitHub](#)

## EDUCATION

<b>University of Utah, David Eccles School of Business</b>	Salt Lake City, UT
Master of Science, Business Analytics (GPA 3.7/4)	December 2024
<b>Centre for Development of Advanced Computing (ACTS)</b>	Pune, India
Diploma in Big Data Analytics (DBDA)	April 2021
<b>Savitribai Phule Pune University</b>	Pune, India
Bachelor of Engineering, Computer Science	May 2020

## SKILLS

**Programming:** R, Python, PySpark, SQL, HiveQL  
**Analytics/Visualization/CRM:** Power BI, Jira, Synapse Analytics, Tableau  
**Cloud Platform:** AWS Services, Azure Cloud, Snowflake, DBT  
**Big Data / ETL:** SQL, Hive, Ab Initio, Alteryx, Airflow, Kafka, MongoDB, Data Modeling, Data Warehousing

## EXPERIENCE

<b>Data Engineer</b>	Pune, India
<b>Hoonartek Pvt. Ltd</b>	May 2021 - July 2023
<ul style="list-style-type: none"><li>Reduced process disruptions by 30% through architecting <b>data migration</b> to AWS using AWS Glue, Kafka, and Airflow, enabling scalable data workflows and ensuring uninterrupted operation of business-critical applications with seamless <b>data ingestion</b>.</li><li>Achieved a 40% improvement in data processing efficiency by implementing scalable <b>data architecture</b> using AWS services and Ab Initio, enabling seamless scaling to handle Airtel's rapidly growing telecom data volumes.</li><li>Enhanced processing speed by 10% and minimized errors by designing and automating ETL processes to standardize diverse file formats, including JSON, ZIP, and CSV.</li><li>Increased data throughput by 5% and improved decision-making insights for telecom operations by developing and maintaining comprehensive <b>ETL pipelines</b>, ensuring accurate and timely data processing for billing systems.</li><li>Improved analytical reporting and query performance by 20% for financial dashboards by developing robust <b>data models</b> using SQL and HiveQL, streamlining data handling and enhancing operational efficiency.</li><li>Optimized <b>data insertion</b> time by 15% in credit data pipelines at Experian, enhancing analytics and reporting efficiency. Integrated multiple sources with PySpark and <b>SQL</b>, ensuring data accuracy and completeness, and implemented robust <b>data validation</b> processes using Regex and <b>Python</b>.</li><li>Enhanced <b>data consistency</b> with <b>Ab Initio GDE</b> by generating daily summaries (e.g., Null Records, Invalid Values) from UNIX-based extractions on Experian's servers. Leveraged <b>Metadata Hub</b> for data governance and Control Center for efficient monitoring.</li><li>Increased automated testing coverage by 12% through <b>data quality</b> frameworks with longitude and latitude sanity checks, data cleansing, and validation routines. Designed test cases to ensure processes met requirements.</li></ul>	

## PROJECTS

<b>Swire Coca-Cola Predictive Maintenance</b>	Dec 2024
<ul style="list-style-type: none"><li>Reduced maintenance losses by 45% through ETL pipelines for machine data across 6 plants and forecasting models. Created real-time dashboards with PowerBI and Streamlit for monitoring and alerts, and proposed IoT integration to boost data ingestion.</li></ul>	
<b>Environmental Sensor Data Analysis</b>	April 2024
<ul style="list-style-type: none"><li>Streamlined environmental sensor data analysis by building an end-to-end data pipeline to process and analyze over 1 million data points daily, enhancing weather trend visualization and machine learning predictions by 30%. Leveraged tools like Azure Cloud, Airflow, and Synapse Analytics for seamless ingestion, transformation, and real-time insights.</li></ul>	
<b>Home Credit Default Risk Prediction</b>	March 2024
<ul style="list-style-type: none"><li>Achieved top-tier loan default prediction with a LightGBM model, attaining an AUC of 0.75 and F1 score of 0.92. Conducted advanced feature engineering and hyperparameter tuning to optimize precision on a dataset of 100,000 records.</li></ul>	
<b>Low-light Image Enhancement Model</b>	April 2021
<ul style="list-style-type: none"><li>Developed and implemented a GAN-based image enhancement tool to transform low-light images into clear day-light images, designed for integration with advanced surveillance software.</li></ul>	

## ACTIVITIES

- Won a sponsored case competition for Swire Coca-Cola by developing a predictive machine maintenance model, analyzing machine downtimes, and delivering solutions to optimize maintenance processes and reduce operational costs.
- Analyzed 50,000+ Super Bowl ad tweets during the 2024 Game Day Analytics Challenge, uncovering key insights and creating impactful infographics to highlight top advertising trends and audience engagement metrics.
- Led operations at the University of Utah Stadium, Arena, and Campus Event Services, overseeing 160+ volunteers, workflows and vendors to ensure efficient service during large-scale events, including Monster Jam, concerts, and college Football.