

SANKET JANOLKAR

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

EDUCATION

University of Utah, David Eccles School of Business

Master of Science, Business Analytics (GPA 3.66/4)

Salt Lake City, UT

December 2024

Centre for Development of Advanced Computing (ACTS)

Diploma in Big Data Analytics (DBDA)

Pune, India

April 2021

Savitribai Phule Pune University

Bachelor of Engineering, Computer Science

Pune, India

May 2020

SKILLS

Programming: R, Python, PySpark, SQL, HiveQL

Analytics/Visualization/CRM: Power BI, Jira, Synapse Analytics, Tableau

Cloud Platform: AWS S3 bucket, Glue, GCP, Azure Databricks, Snowflake, Azure Data Lake, DBT

ETL Tool: Ab Initio, Alteryx

Databases/Big Data: MySQL, Hive, Excel, Apache Airflow, Apache Kafka, MongoDB, Firebase, Data Modeling, Data Warehousing

EXPERIENCE

Associate Data Engineer / Data Consultant

Pune, India

Hoonartek Pvt. Ltd

May 2021 - July 2023

- Designed and implemented innovative scalable **data architecture** to streamline data ingestion, transformation, and storage across AWS S3, Glue, and Ab Initio.
- Designed and automated ETL processes to standardize diverse file formats (JSON,CSV,DOCX ,ZIP,TXT,XLSX), enhancing processing speed by 10% and minimizing errors.
- Developed and maintained comprehensive **ETL pipelines** for data extraction, transformation, processing, and loading, resulting in a 5% increase in data throughput and improved decision making insights.
- Developed robust **data models** to support analytical reporting and improve query performance using SQL and HiveQL, resulting in a 20% improvement in overall data processing efficiency.
- Led **data migration** to cloud platforms using AWS Glue, ensuring data scalability and minimal disruptions, with batch and real-time data processing using Apache Kafka and Airflow to support business-critical applications.
- Facilitated seamless **data integration** across multiple sources, leveraging PySpark and **SQL** to ensure data accuracy and completeness, reducing insertion time by 15% by creating new **data validation** processes using Regex and **Python**.
- Implemented enterprise-wide **data warehousing** solutions using Ab Initio GDE and Metadata Hub, enhancing reporting capabilities and ensuring data consistency, and job monitoring via Control Center.
- Established **data quality assurance** frameworks to enhance reliability, designing test cases and plans for sanity check **with data cleansing** and validation routines, which increased the automated testing process by 12%.

PROJECTS

1st Place - Swire Coca-Cola Capstone Case Study Competition

Dec 2024

- Designed ETL pipelines for maintenance data across 6 plants, built models to forecast activities, reduced loss by 45%, developed real-time dashboards (using PowerBI and Streamlit web app) with monitoring and alerts, and recommended IoT integration to improve data ingestion, operational efficiency and dynamic feedback.

Environmental Sensor Data Analysis

April 2024

- Constructed a comprehensive data pipeline for environmental sensor data analysis, leveraging GitHub, Visual Crossing API, Azure Data Lake, Apache Airflow, Google Colab, Databricks, and Azure Synapse Analytics for efficient ingestion, processing, machine learning, and visualizing weather trends.

Home Credit Default Risk Prediction

March 2024

- Developed a predictive model to assess loan default risk using supervised learning. Conducted feature engineering and implemented the LightGBM algorithm, optimizing hyperparameters for enhanced accuracy.

Low-light Image Enhancement Model

April 2021

- Developed and implemented an algorithm utilizing Generative Adversarial Network (GAN), a Deep learning model to enhance the low-light image to clear day-light image.

ACTIVITIES

- Won a sponsored case competition for Swire Coca-Cola by developing a predictive maintenance model, analyzing machine downtimes, and delivering solutions to optimize maintenance processes and reduce operational costs.
- Analyzed Super Bowl ad tweets in the 2024 Game Day Analytics Challenge to create insightful infographics.
- Led operations at the University of Utah Stadium, Arena, and Campus Event Services, overseeing workflows, volunteers and vendors to ensure efficient service during large-scale events.