Aniket Belsare

Data Engineer

8390486326 aniketbelsare348@gmail.com *House No 28, Pagnis Paga, Near Water Tank, Indore, 452001*

github.com/abelsare348 linkedin.com/in/aniket-belsare-56557a217

PROFILE SUMMARY

[Dynamic, Microsoft Certified [DP-203, COF-02] Data Engineer with 3+ years in designing and implementing scalable data solutions on the Azure platform. Demonstrated expertise in data pipeline orchestration, ETL processes, and data warehouse modeling. Skilled in Azure Data Factory, Azure Databricks, Snowflake, and data lake management. Known for collaborating effectively with cross-functional teams to translate business requirements into impactful data-driven insights.]

Certifications

• **Azure DP-203 [Microsoft]** • **Databricks Data Engineer Associate** • **Snowflake (Snow Procore) COF02**

Experience

- ***47Billion, Indore***
  Position: Data Engineer
  [Sep 2024 – *Present*]
- ***Nucleus-Teq, Indore***
  Position: SDE-I (Data Engineer)
  [Dec 2021 – Sep 2024]
- ***QuantumSoft Technologies Pvt ltd, Amravati***
  Position: Project Engineer (Intern)
  [Nov 2020 – Nov 2021]

Education

SGBAU, Amravati, Maharashtra

BE (Computer Science):
Grade/CGPA: **9.19**/10
12th Science (General): **70.62%**
10th Board: **86.80%**

Skills

Languages:
SQL, Python, Py-spark, Scala-Spark
Tools:
• Databricks • Azure Data factory
• Azure Data Lake • Azure Synapse
• Postgres/DBMS • Microsoft Purview
• Snowflake • Git-GitHub • Shell-Script
• Tidal • Active Directory • Kafka
• Hadoop • Key-Vault • Logic Apps
Cloud: Azure, AWS, GCP
Specialty: Big-data, Data- Analytics, Data-Engineering, Data-Warehouse.
OS:
• *Windows* • *Ubuntu-24.04* • *Mac-OS*

Projects

HealthCare Data (Data Integration and Transformation of Electronic Health Records (EHR) for Insurance Companies Using Azure Data Services)
Sep, 2024 – Present **(*Data Engineer, 47Billion, Indore*)**
Tech Stack: - Azure Synapse, Azure Data Factory, Databricks, MS-Purview

- Data dumps from SAP and WEBTMA to the input layer of Azure Data Lake Gen2 using Azure Data Factory, improving data availability and reducing data transfer time by 30%.
- Established raw, refine, and enterprise layers in Azure Data Lake, implementing data partitioning strategies that optimized storage costs by and reduced data retrieval time
- Utilized Azure Databricks for data transformation processes, ensuring scalability and performance. Decreased overall processing time to 40% through parallel processing and optimized algorithms, resulting in faster insights generation.
- Implemented Microsoft Purview for data governance, enhancing data discoverability and compliance, and integrated Azure Functions and Logic Apps for custom event alert generation to improve operational efficiency

Sales Data (Data Modernization & Efficiency Enhancement)
Dec, 2021 - Sep, 2024 **(*Software Engineer/SDE-I, Nucleus-Teq, Indore*)**
Tech Stack: - Azure Databricks, Snowflake, Netezza, Informatica, Tidal

- Spearheaded migration project from Netezza IBM's on-premises system to Snowflake, a modern data warehouse solution, while transitioning from Informatica to Databricks for ETL processes, resulting in a 20% increase in data processing efficiency
- Oversaw the end-to-end migration process, encompassing planning, execution, and validation stages, ensuring seamless integration and minimal disruption to operations. Completed the migration project three weeks ahead of schedule, resulting in significant cost savings.
- Collaborated closely with stakeholders to assess requirements, design data migration strategies, and implement necessary changes to optimize performance and scalability in Snowflake, reducing data storage costs by 15% annually.
- Directed the conversion of ETL workflows from Informatica to Databricks, resulted into 30% improvement in processing speed; the new architecture supported project in efficient way that processed over 500,000 records in under two minutes.

Vendor Data (Azure Data Transformation Initiative)
Nov, 2020 - Nov, 2021 **(Project Engineer-Intern, QuantumSoft Technologies)**
Tech Stack: - Azure Data Factory, Postgres, Azure Synapse

- Migrated data from on-premises PostgreSQL to Azure Data Lake Gen2 using Azure Data Factory, reducing transfer time by 30% and improving data accessibility
- Established raw, refine, and enterprise layers in Azure Data Lake, optimizing storage costs by 20% and reducing data retrieval time by 25%.
- Used Azure Databricks for data transformation, cutting processing time by 40% through parallel processing and optimized algorithms.
- Implemented data governance measures with Active Directory and Key Vaults to accomplish security and data governance.