**Aziz Maksudov**

Tashkent maqsudovazizbek85@gmail.com

Uzbekistan

+998 97 342 01 63

Data Engineer with 2+ years of IT experience in querying, designing and maintaining databases on Postgres, Oracle and MS SQL Server. Strong knowledge in developing ETL packages using SSIS and Apache airflow to automate data migration from various sources. Great range of experience in building pipelines using python, Apache Airflow and Azure Data Factory. Responsible team member with excellent time management skills that fosters collaboration.

# Professional Summary

* Strong skills in creating and managing ETL and ELT pipelines that take data from various sources such as Microsoft SQL Server, Greenplum, Oracle, MongoDB, Excel Files, Parquet Files, and Azure Blob Storage. Proven experience with utilizing tools like SSIS, python (pandas and pyspark libraries), Azure Data Factory, and Databricks Asset Bundles to build and manage variety of pipelines.
* Strong knowledge in performance tuning, debugging and query optimization by analyzing the runtime, execution logs and statistics to identify the weaknesses of queries and provide the most efficient solution (be it a clustered, non-clustered index, partition, hashing, or completely different solution based on the time complexity of the query) to improve the speed of the query/stored procedure/data set query/view in question.
* Designed and implemented Delta Live Tables (DLT) pipelines for efficient data processing.
* Worked with append-only Delta tables and views, utilizing soft delete to optimize resource usage and maintain historical data efficiently.
* Improved performance and cost-efficiency by minimizing storage overhead while ensuring data integrity.
* Hands on experience in working with CDC pipelines created using Azure Data Factory that utilize Last Updated/ Modified Date of Control tables and monitor the transaction logs of the source tables to update the staging tables incrementally while also keeping the historical data in platforms like Azure Blob Storage.
* Strong knowledge in creating Star Schemas in 2NF and Snowflake Schemas in 3NF utilizing Fact and Dimension tables in OLAP data warehouses.
* Created ETL pipelines in Azure Synapse and Airflow. Excellent experience in working with DevOps tools like Github Actions.

# Skills

| Advanced T-SQL coding | Python(pymongo, pandas,  Pyspark) | Debugging/Troubleshooting |
| --- | --- | --- |
| ETL implementation | Data warehousing | Performance tuning |
| Cloud technologies  **Professional Experience** | Agile methodology | Business analysis |
| **Data Engineer** |  | **Apr 2023 – March 2024** |

*NTHRIVE(remote)*

* Collaborated with Business Intelligence and Executive teams to understand data gathering requirements.
* Involved in Designing, Building, and Optimizing OLAP data warehouses using ETL and ELT pipelines that take data from various sources including Oracle, Microsoft SQL Server, MySQL, excel files and more. Also involved in Data Mapping, Documentation, Debugging existing pipelines, stored procs, and queries that may need maintaining due to source changes, data type issues, etc.
* Parsed json files to pull data from no-sql databases like MongoDB.
* Designed table schemas and database schemas according to medallion structure and Software Development Lifecycle.
* Built parametrized pipelines using python’s functional programming and built fast queries using multiprocessing and multithreading techniques.
* Built parametrized stored procedures, efficient CTE’s, tables, views, functions, indexes, and DML triggers for different tasks.
* Filtered out the bad data using both pandas dataframes and complex sql statements.
* Staged data from various sources while building data warehouse, build incremental loads, fact, and dim tables.
* Constantly did data quality checks while extracting data from various sources.

**Data Engineer Nov 2022 – Apr 2023**  *GlobalTranz(remote), Texas, USA*

* Built efficient ETL pipelines in Azure Data Factory migrating data from various sources such as Azure blob storage, data lakes and data warehouses of different kinds.
* Built dynamic delta load tables and DLT pipelines using databricks cli extension of VScode and auto-deploy feature of GitHub actions. Parametrized sql scripts to take variable values from GitHub workflow configurations and adjust the code as the workspace the code is being deployed to different kinds of workspaces.
* Involved in maintaining data warehouses that are managed by ADF pipelines. Was responsible for controlling the data flow fixing and writing one time adjustments to pipelines/procedures/scripts when needed.
* Used sub-queries, CTEs and complex SQL structure while developing OTAs and analyzed complex stored procedures to maintain pipelines.
* Created indexes and views, stored procedures to enhance database/application performance.
* Applied reverse engineering to find the exact point the data become corrupted in various SSIS packages, stored procedures, and ADF pipelines.
* Created and managed append only delta tables/views that have soft delete columns to improve the efficiency and provide immediate data access for dashboards/reports.
* Worked with the BI team to validate the efficiency and quality of data and made adjustments to data flow when necessary per request of BI team.

# Education

Bachelor’s degree in Economics

UrDU, Urgench, Uzbekistan