CELEBAL TECHNOLOGIES

INTERNSHIP REPORT

CELEBAL SUMMER INTERNSHIP- 2023
(27 MAY 2023 - 29 JULY2023)

NAME :   ADITYA RANJAN PRADHAN

INTERN ID : CT-CSI23/DE0331

DOMAIN :   DATA ENGINEERING

Introduction

The internship report provides an overview of my internship experience at Celebal Technology. During the internship period from 27 May 2023 to 29 July 2023, I had the opportunity to work in the Data Engineering Department and gained valuable insights into the operations and functioning of the company. This report aims to summarize the key tasks, achievements, challenges, and learning experiences encountered during the internship.

Company Overview

- Celebal Technologiesis a premier software services company specializing in Data Science, Big Data, and Enterprise Cloud.

- Solutions powered by Robotics, Artificial Intelligence, and Machine Learning algorithms.

- Focus on improving business efficiency in an interconnected world.

- Talented team of industry professionals with expertise in diverse technologies, including Blockchain, Analytics & Visualization, IoT, Chatbot, Data Science - AI & ML.

- Legacy and on-premise application modernization through modern cloud paradigms: Microservices, Advanced Analytics, API Management, and Cloud-native services.

Technologies Used :

I - Azure Data Factory

Azure Data Factory, a cloud-based data integration service, played a crucial role in the data engineering projects during the internship. ADF facilitated the creation of data-driven workflows to orchestrate data movement, data transformation, and data loading processes. Some key functionalities and features of Azure Data Factory utilized during the internship include:

- **Data Pipeline Creation:** Designing and implementing data pipelines to extract data from various sources, transforming it, and loading it into the target data storage.
- **Integration with Azure Services**: Leveraging Azure Data Factory's seamless integration with other Azure services like Azure Blob Storage, Azure SQL Database, and Azure Data Lake Store.
- **Data Transformation:** Applying data transformation activities using ADF's Data Flow feature to cleanse, enrich, and aggregate data.
- **Scheduling and Monitoring:** Configuring pipeline schedules and monitoring pipeline executions for efficient data processing

II - SQL Server Management Studio

SQL Server Management Studio was extensively used to interact with SQL Server databases for data warehousing purposes. Key activities performed with SSMS include:

- **Database Management:** Creating and managing databases, tables, and indexes to store the transformed data.
- **Query Development:** Writing and optimizing SQL queries to extract data for analysis and reporting purposes.
- **Data Validation:** Ensuring data integrity by performing data validation checks on loaded data.
- **Stored Procedures :** Creation of Stored Procedures to reduce the Complex code Structures.

III - Microsoft Power BI

Power BI, a powerful business analytics tool, was utilized to create interactive and insightful visualizations for data analysis and reporting. Key functionalities used in Power BI include:

- **Data Connection:** Connecting Power BI to the SQL Server database to retrieve real-time data.
- **Report Creation:** Designing engaging and dynamic reports and dashboards to convey data insights effectively.
- **Report Publishing:** Publishing reports to the Power BI service to share with stakeholders for collaborative analysis.

Projects Undertaken :

During the internship, I had the privilege to work on several data engineering projects that showcased the application of the above technologies. Some notable projects include:

Telecom and Communication Dashboard:

Objective: Develop a comprehensive telecom and communication dashboard to analyze key performance indicators (KPIs) and monitor network and customer-related metrics for a telecommunications company.

- Created an extensive telecom and communication dashboard using Power BI to track and assess critical KPIs for a telecommunications company.
- Data extracted from various data sources within the company, including network performance data, customer usage data, billing and revenue data, and customer feedback.
- Utilized Azure Data Factory to efficiently extract, transform, and load data from different sources into a centralized data repository.
- Applied real-time updates and visualizations to monitor network performance metrics such as call drop rate, data throughput, signal strength, and network availability.
- Utilized Power BI's interactive features, slicers, and filters to enable users to explore data and identify patterns or anomalies.
- Technologies used: Power BI, SQL Server Management Studio, Azure Data Factory, Telecom Network Performance Data, Customer Usage Data, Billing and Revenue Data.

Project Link:

https://github.com/aditya8725/CSI_23_Internship_Data_Engineering

Healthcare Dashboard:

Objective: Develop a comprehensive healthcare dashboard specifically focused on monitoring and analyzing COVID-19 data for healthcare organizations, public health authorities, and the general public.

- Created an extensive COVID-19 healthcare dashboard using Power BI to track and assess critical COVID-19 indicators.
- Data extracted from reliable sources, including government health departments, public health APIs, and global health organizations, through Azure Data Factory and API integrations.
- Utilized real-time updates and geospatial visualization to monitor COVID-19 infection rates, testing data, hospitalization rates, vaccination coverage, and other relevant pandemic-related metrics across different regions and demographics.
- Incorporated specific COVID-19 KPIs such as positivity rate, case fatality rate, and reproduction number (R0) to provide insights into the severity and transmission dynamics of the virus.
- Implemented Power BI's interactive features, slicers, and filters to allow users to explore the data in-depth, enabling them to identify trends, correlations, and patterns in the spread of the virus.
- Technologies used: Power BI, SQL Server Management Studio, Azure Data Factory, Public Health APIs, Global Health Databases.

Project Link:
https://github.com/aditya8725/CSI_23_Internship_Data_Engineering

Skills Gained :

Throughout the internship, I had the opportunity to develop and enhance various skills, including:

- **Power BI Proficiency**: I have developed expertise in creating dynamic and interactive visualizations using Power BI. This skill enables me to present data insights effectively and support data-driven decision-making.
- **Azure Data Factory:** Working with Azure Data Factory has provided me with hands-on experience in cloud-based data integration and ETL workflows. I can now efficiently orchestrate data pipelines and manage data movement between different sources and destinations.
- **SQL Server Management Studio (SSMS):** I have gained proficiency in using SSMS to manage and query relational databases. This skill allows me to perform various database administration tasks, ensuring smooth data operations.
- **Data Modelling and Warehousing:** Understanding data modeling principles and data warehousing concepts has equipped me to design efficient data structures and build data warehouses for analytical purposes.
- **Data Transformation and ETL Processes:** I have learned techniques to transform, cleanse, and enrich data during the ETL process, ensuring data accuracy and consistency.
- **Problem-Solving and Troubleshooting**: The internship has honed my problem solving abilities, enabling me to identify and resolve data-related issuesefficiently.
- **Collaboration and Teamwork:** Working on cross-functional projects has improved my collaboration and teamwork skills, fostering effective communication and coordination with colleagues.
- **Time Management:** Meeting project deadlines and prioritizing tasks effectively has been a crucial aspect of the internship, and I have learned to manage my time efficiently

Feedback and Learning Experience :

I am immensely grateful for the support and guidance provided by my mentors, **Ishita Agarwal Mam**, **Nikhil Agarwal Sir** and **Ravindra Bagadi Sir** and the HR team during my data engineering internship at Celebal Technology. Our Mentors played a crucial role in shaping my learning experience by sharing their extensive knowledge in data engineering and providing valuable insights on project execution. Their patience and willingness to answer my questions empowered me to explore new technologies like Power BI, data warehousing, and Azure Data Factory.

The HR team of **Piya Soni Mam**, **Priyanka Maharishi Mam**, **Prithvi Kumawat Sir** and **Sharthak Acharjee Sir** was exceptionally supportive throughout the internship, from the initial onboarding process to ensuring a smooth integration into the team. They made sure I had all the resources I needed to excel in my projects and regularly checked in to ensure my well-being and satisfaction. I, truly appreciate the collaborative and inclusive work culture at Celebal Technologies , which allowed me to work effectively with cross-functional teams and learn from experienced professionals. The mentorship and guidance from our mentors and the continuoussupport from the HR team have been invaluable in my professional development.

Working on the Telecom and Communication Dashboard and Healthcare Dashboard during my internship at Celebal Technology has been a truly rewarding experience. The project provided me with an excellent opportunity to apply my theoretical knowledge to real-world data engineering challenges. I gained valuable skills in Data Engineering, and I was able to develop a deeper understanding of tools like [e.g., Power BI, Azure Data Factory, SQL Server Management Studio]

Conclusion :

Data engineering internship in the data engineering department of Celebal Technology has been a transformative experience. Throughout the internship, I had the privilege to work with cutting-edge technologies such as Azure Data Factory, Power BI, and SQL Server Management Studio, gaining practical skills that will undoubtedly shape my future career.

The hands-on projects enabled me to apply theoretical knowledge to real world scenarios, strengthening my understanding of data engineering principles and best practices. Working on data integration, transformation, and visualization tasks has sharpened my problem-solving abilities and taught me the importance of data accuracy and quality.

The guidance and mentorship provided by the team, especially my mentors, have been instrumental in my professional growth. Their constant support, constructive feedback, and encouragement fuelled my passion for data engineering.

Overall, this internship has solidified my aspiration to pursue a career in data engineering. I am grateful to Celebal Technology for this invaluable opportunity and look forward to utilizing the skills acquired to contribute meaningfully to the field of data engineering