|  |
| --- |
| SUMMARY |
| Priya Upadhyay holds the position of Associate Data Engineer L2 at Publicis Sapient's remote Bengaluru office. She became a part of the Sapient team in 2022 and has since been involved in numerous projects, covering the entire spectrum of the Data Engineer lifecycle. Her expertise also extends to Big Data. |
|  |
| Priya Upadhyay is a distinguished professional who earned her B.Tech. degree from Dr. A.P.J Abdul Kalam University in 2017. Her career journey led her to a dedicated three-year tenure at Jocata, where she excelled from January 2019 to March 2022. During this period, she showcased exceptional proficiency in AWS Cloud technologies, demonstrating mastery in Glue, S3, Cloud Watch, Lambda, and SNS, as well as expertise in Spark. Her skills extended to crafting intricate AWS Glue Scripts, proficiently managing S3 buckets, and harnessing Spark SQL for precise data transformation. Of particular significance is her wealth of experience in seamlessly transferring data between on-premise systems and the AWS Cloud. At Jocata, she undertook the pivotal responsibility of managing data within S3 buckets and data lake accounts. She adeptly implemented AWS Glue scripts, leveraging the power of Spark, and formulated transformation logics tailored to project-specific requirements. Her contributions extended to the creation of Lambda functions, facilitating timely notifications for actions carried out on S3 buckets. As a devoted member of the migration testing team, she rigorously assessed results based on meticulously prepared mapping documents provided by clients. Noteworthy is Priya's dedication to fostering a collaborative work environment. She willingly offered guidance to her fellow team members, supporting them in accomplishing their tasks successfully. Her commitment to clear and efficient communication was evident in her active participation in Scrum Calls, as well as her regular updates on task progress provided to clients. Furthermore, Priya proactively took charge of documenting system processes and procedures, ensuring a valuable resource for future reference. |

|  |
| --- |
| WORK EXPERIENCE: |
| **Publicis Sapient April 2022 To Present** |
| **Hypothesis Bank, Data Engineer** |
| **Description:**The Goal of the project is to write transformation logic on the files provided in GCP bucket using PySpark, storing the transformed data in Big query Table and visualizing it by Looker. |
| **Roles and Responsibilities:**• Priya as a Data Engineer, she specialized in crafting and managing complex data workflows using PySpark in Airflow jobs. Her role involved orchestrating these jobs to ensure the smooth execution of intricate data operations. Additionally, She leveraged the power of Google Cloud Platform (GCP) services, including Big Query, Dataflow, and Storage components, to efficiently deploy and manage data pipelines. This enabled her team to achieve optimal performance and scalability in handling large datasets. |
|  |
| • One of her key contributions was the design and implementation of data models in Big Query. These models were carefully crafted to streamline the storage and analysis of extensive datasets, ensuring quick accessibility and efficient data management. In addition, Priya played a pivotal role in establishing a Looker Dashboard, providing a visual representation of critical business metrics. This tool proved instrumental in driving data-informed decision-making throughout the organization. |
|  |
| • Priya' capacity as a Data Engineer, She was responsible for the end-to-end management of data pipelines. This encompassed a range of operations, from data extraction and transformation to loading (ETL) and validation. By automating these processes through scheduled Airflow jobs, She significantly improved operational efficiency, allowing for more streamlined and reliable workflows. |
| **Skills/Domain:**GCP, Airflow, PySpark, Big Query |
| **TLD - ENI, Data Engineer** |
| **Description:**TLD is a Data warehouse which processes and store trade data and generating reports. With the help of ETL service we are generating reports for End user to perform Analytics. |
| • Priya' role as a Data Engineer, She spearheaded the creation of robust ETL pipelines aimed at extracting, transforming, and loading data from diverse sources into target databases. This meticulous process ensured the accuracy and reliability of the data, a critical aspect for any data-driven organization. Furthermore, She placed a strong emphasis on maintaining data integrity by constructing ETL pipelines that were specifically designed to safeguard the quality and consistency of datasets. |
|  |
| • An integral part of Priya' responsibilities involved executing queries for data manipulation, aggregation, and retrieval. She paid meticulous attention to optimizing these queries, a measure that significantly enhanced their efficiency, ultimately contributing to smoother data operations. Additionally, Priya' proficiently utilized IBM Infosphere DataStage for tasks ranging from job creation and development to scheduling and performance optimization. This utilization played a pivotal role in the facilitation of effective data integration workflows. |
|  |
| • A key focus of Priya' work revolved around the creation and management of data models, encompassing both logical and physical designs. These models served as the backbone for ETL processes, guaranteeing data uniformity across the organization. To further enhance data quality and reliability, She strategically integrated data validation protocols within ETL workflows. This practice not only identified irregularities or inconsistencies but also provided a streamlined approach to rectifying them, ensuring the high dependability of data. |
| **Skills/Domain:**ETL, DataStage, SQL |

|  |
| --- |
| EDUCATION: |
| B.Tech. (Bachelor of Technology), (Computer Science & Engineering): Dr. A.P.J Abdul Kalam University, 2017 |
|  |

|  |  |
| --- | --- |
| **SKILLS:** | |
| **Additional Skill(s):** | Data Engineering |
| **Other Skills:** | Spark, Pyspark, SparkSQL | Aws Glue, S3, CloudWatch, Lambda, SNS, Athena | GCP, Big Query, Airflow | ETL, DataStage, SQL |
|  |  |

|  |
| --- |
| ADDITIONAL DETAILS: |
| **Awards** |
| Willingness to Learn & Client Influence -- 2022 |