Dimple Francis

Data Engineer | Data Analyst| Data Science Enthusiast | Deputy Scrum Master

University of Stirling, March2022

Contact +447769491554, Edinburgh

I am a seasoned Information Technology professional with over a decade of experience in the field. In March 2022, I graduated with an MSc in Big Data from the University of Stirling, expanding my knowledge and skills in the ever-evolving world of data and technology. I am currently upskilling myself in Data Science.

ETL Development Expertise: I have specialized in ETL development, with a primary focus on providing solutions within Data Warehousing. My expertise lies in utilizing tools such as Informatica and SQL to transform complex data into valuable insights.

Data Engineering Proficiency: Since March 2022, I have diversified my skill set, working with various technologies such as Snowflake, Streamsets, MongoDB, Hadoop, Tableau, dbt, Python, Pyspark. This expansion has enabled me to seamlessly integrate diverse data sources and construct robust workflows and pipelines to optimize data management processes.

Leadership and Project Management: I have also played the lead role, successfully managing teams to deliver projects. Additionally, I have provided technology solutions to clients across different categories, showcasing my adaptability and problem-solving skills.

Scrum Master Experience: As the deputy scrum master for my team, I have played a pivotal role in ensuring agile project management, improving efficiency, and maintaining a collaborative team environment.

Full Project Lifecycle Involvement: I excel in handling projects from inception to implementation. This includes requirement analysis, application design and development, application maintenance, reviews, and testing.

Tool Proficiency: SQL, Snowflake, Spark, Python, Jupiter notebook, Mongo dB, Git, PyCharm, Autosys, SQL, Scala, Informatica, Hadoop, Tableau, Machine Learning. I have extensive experience working with ETL tools, including Informatica Power Center 8.X/7.X/9.X. I am adept at developing workflows, sessions, and tasks to effectively manage large data loads. I have also excelled in source system/data analysis to identify and develop data requirement documents.

Client-Centric Approach: My strong communication and interpersonal skills allow me to effectively engage with customers, clients, and end-users. I have demonstrated my ability to work independently and handle peaks in demand.

Soft Skills: I am known for my ability to quickly adapt to new and challenging technological environments, staying at the forefront of industry trends. I also possess strong verbal and written communication skills, which are crucial in conveying complex technical information to both technical and non-technical stakeholders.

Technical Skills	
Hardware / Platforms	Windows 98/XP/2000, Unix
Technology	Data Engineering, Data warehousing – Informatica ETL Tool, Streamsets, Databricks, Machine Learning, dbt, Python, PySpark, Spark, SQL, Unix basics, Autosys, Python, Scala, Git (Bitbucket), Hadoop, AWS
Databases	Oracle 10g, Mongo dB basics, Snowflake
ETL Tools	Informatica 7.1, Informatica 8.1, Informatica 9.1

Education		
❑ **Degree**	❑ **College/University**	❑ **Year of Passing**
❑ Msc in Big Data	❑ University of Stirling	❑ 2022
❑ B.Tech in Information Technology	❑ Kerala University	❑ 2006
❑ HSC	❑ CBSE – Indian School Salalah	❑ 2002
❑ SSLC	❑ CBSE – Indian School Salalah	❑ 2000

Work Experience

Company	Role	Year	Location
❑ bigspark	❑ Data Engineer	❑ Mar 2022 – Oct 2023	❑ Edinburgh
❑ Business Stream	❑ Integrations Developer	❑ Oct 2019 – Oct 2020	❑ Edinburgh
❑ Cognizant	❑ Senior Associate	❑ Apr 2011 – Dec 2015	❑ India
❑ IBM	❑ ETL Specialist	❑ Mar 2010 – Mar 2011	❑ India
❑ Wipro	❑ Senior Software Engineer	❑ Nov 2006 – Feb 2010	❑ India

Certifications/Papers/Case Study/Awards

o dbt fundamentals certificate – November 2023

o Employee of the month - August 2023 – for delivering the ERPNext HR module efficiently.

o Snowflake course – Hands on training – August 2023

o Snowflake – Hands on Essentials -Datawarehouse – 2022

o Informatica Certified Developer -2011.

o Provided Informatica training for 3 batches in CTS and got good ratings.

o Prepared Case Study for Informatica Upgrade project in Cognizant.

o Abbott SNOP Team won the Project of the year –Runner up in CTS Excellence Awards. Received CTS SPOT certificate and cash award for the same.

o Cognizant Certified Professional – Informatica Practitioner -2014

o Published approach note on "Delete logic – performance tuning" in Wipro.

Relevant Project Experience

Organization: Bigspark

Project #1

Title   Fincrime Project
Tools   Streamsets, SQL, Oracle, Mongo dB, Snowflake, Python, Scala, Git (Bitbucket), Autosys, Jira, Hadoop, Tableau
Team Size   12
Role   Data Engineer
Period   March 8th, 2022 – May 2023

Project Objective

The project focused on automating the current reporting tools by delivering the data for a few heterogenous sources into snowflake where tableau was used for reporting this data to the stakeholders.

Role and Responsibilities

- Handled data analysis and requirements gathering.
- Responsible for creating streamsets pipelines to load the data from respective sources to AWS S3 bucket and further loaded to Snowflake.
- Responsible for the end-to-end release of Mongo dB source till Snowflake.
- Responsible for debugging all errors related to these pipelines.
- Assisted in fetching results for Tableau and handled data transformation rules.
- Create the needed jil scripts and respective yaml files.
- Created the SDK scripts and respective yaml files.
- Pushed the code to Bitbucket using Git/Mobax/PyCharm.
- Monitored and validated the data after release.
- Handled snowflake deployment.
- Actively participated in daily Agile scrums.

Project #2

Title	bigspark World(internal)
Tools	ERPnext, Frappe framework, SQL, Javascript, GIT
Team Size	15
Role	HR Module Lead
Period	May 2023 – September 2023

Project Objective

The project focused on building HR, Project, Recruitment, CRM Modules on ERPNext platform which would help everything to be stored in a centralized area. It would also bring enhanced decision making.

Role and Responsibilities

- Responsible for successfully spearheading the integration of ERPNext's HR module into the company's operations, optimizing human resource management processes and enhancing overall organizational efficiency.
- Collaborated with HR team to identify requirements, configure the module.
- Migrated huge employee related data and leave applications of 2020 till date from BrightHR portal to ERPnext.
- Responsible for overall delivery and quality of HR Module.
- Debugged and resolved issues.
- Actively participated in daily Agile scrums and acted as the deputy scrum master for 3 weeks in the absence of the scrum master.

Organization: Business Stream

Project #1

Title	Integrations Support and enhancement
Client	Business Stream
Tools	Informatica 9.5, Oracle 10g, SQL, batch script
Team Size	5
Role	Integrations Developer
Period	Oct 2019 – Oct 2020

Project Objective

This is primarily an enhancement project. It would handle all enhancements/automation for Integrations team.

Role and Responsibilities

- Responsible for automating the manual tasks.
- Involved in data analysis and requirements gathering meetings.
- Responsible for ensuring delivery of 'small continuous development projects' on time.
- Co - ordinate with the team, help them out in technical aspects/issues.
- Hands on informatica mappings- development/enhancement.
- Responsible for handling incidents raised by internal business to fix issues relating to data/code.
- Responsible for fixing failures in production.
- Review the code.
- Responsible for overall delivery of metrics for support.

Organization: Cognizant Technology Solutions

Project #1

Title   IM – CLDW-CDF
Client   Hartford
Tools   Informatica 9.5,Oracle 10g, Unix
Team Size   5
Role   ETL Offshore Lead
Period   Jan 2014 –Dec 2015

Project Objective

It is primarily a support cum enhancement project. It would handle all
tickets/issues/enhancements for Commercial Lines Data warehouse – Catastrophe Data Facility.

Role and Responsibilities
- Allocation of ETL upgrade tasks to team members.
- Mentoring new joiners.
- Responsible for reviewing maintaining the understanding docs.
- Handle status calls and share the tasks with the team members.
- Co - ordinate with the team, help them out in technical aspects/issues.
- Hands on informatica mappings- development/enhancement
- Review the code
- Responsible for overall delivery of project and project metrics.

Project #2

Title   PPD Informatica Upgrade Project
Client   ABBOTT
Tools   Informatica 8.5, Informatica 8.6, Informatica 9.1, Oracle 10g, Unix
Team Size   7
Role   ETL Lead
Period   October 2012 – March 2014

Project Objective

This project focused on the Informatica upgrade of all projects in PPD division of ABBOTT to
version:9.1.

There were 2 sets of upgrades involved:

1)Informatica8.5 to Informatica9.1
2)Informatica8.6 to Informatica9.1

The first phase of the project involved dry run of all workflows in dev. The second phase involved
the data testing of few selected workflows in each folder in QA. The project also optimized the
performance of few workflows.

Informatica8.5 to Informatica9.1 Upgrade Snapshot:

The upgrade handled 8 folders which constituted of 188 workflows.

Informatica8.6 to Informatica9.1 Upgrade Snapshot:
The upgrade handled 21 folders which constituted of 1483 workflows.

Role and Responsibilities
- Allocation of ETL upgrade tasks to team members.
- Mentoring new joiners.
- Prepared CASE STUDY for the project.
- Responsible for reviewing maintaining the Upgrade Plan and Test docs.
- Handled status calls and shared the tasks with the team members.
- Co - ordinated with the team, helped them out in technical aspects.
- Provided support during pre-UAT and UAT phases.
- Was effective in meeting the given deadlines.

Project #3

Title	ADC S&OP Implementation Project
Client	ABBOTT
Tools	Informatica 8.1, Oracle 10g, Unix
Team Size	20
Role	ETL Developer
Period	August 2011 – October 2012

Project Objective

Abbott Diabetes Care (ADC) has a manual process for maintaining global supply & demand data and report generation. ADC SNOP project will focus on backend automation of demand & supply data integration and report generation by reducing the manual interventions.

Business rules are distributed in macros, queries, and multiple Microsoft Access databases. A centralized architecture does not exist. The current solution is not flexible to support the continued growth of the ADC business.

ADC S&OP project focused on: Analyzing and developing harmonized business rules globally, Design and develop a scalable robust integration framework, Design, and develop a centralized database design for data integration and reporting, implement a data integration and reporting solution based upon the ADC BI solution.

Role and Responsibilities
- Involved in design phase and preparation of low level and high-level documents
- Responsible for development, testing of ETL code.
- Allocation of ETL development to team members.
- Mentoring new joiners.
- Responsible for creation and maintaining the SOP and UAT documents.
- Co - ordinated with Onsite team during the status calls and shared the same with the team members.
- Co - ordinated with the team, helped them out in technical aspects.
- Provided support during pre-UAT and UAT phases.

Project #3

Title	PepsiCo People Master
Client	PepsiCo
Tools	Informatica 8.1, Oracle 10g, Unix
Team Size	6
Role	ETL Developer
Period	May 2011 – Aug 2011

Project Objective

The People Master is a data repository that contains information about employees, contractors, and third parties. It is also an authoritative system for master and operational data required to provision entitlements. Provisioning reference data includes roles and entitlements. The People Master will synchronize people-related attributes across various applications, data stores, and directories.

Role and Responsibilities
- Creating ETL Specifications, ETL code
- Unit Testing
- Support during UAT and Implementation phases

Organization: IBM India Pvt Limited

Project #1

Title	Pfizer - PSS Remediation Migration
Client	PFIZER
Tools	Informatica 8.1, Oracle 10g, UNIX, Teradata
Team Size	2
Role	ETL Developer
Period	March 2010 to March 2011

Project Objective
The objective of this project is to migrate US Commercial Pharmaceuticals Data Warehouse (DW), i.e., US Data Management (DM) environment, and its data marts, associated processes (i.e., ETL, reporting, etc.) from the current technology/platform (i.e., Solaris - Oracle) to the Teradata technology/platform. The technical focus of this project is to 'Lift and Shift' the existing infrastructure components of the US Commercial Data Warehouse to the new Teradata technology platform without any new business requirement or enhancement to existing functionality.

Role and Responsibilities
- Understanding of the Data flow and data mappings rules and implementations.
- Key responsible resource for Design/Build of various complex interfaces.
- Handling challenging timelines and motivating the team.
- Integration design of the different components of the projects.

Organization: Wipro Technologies

Project #1

Title	SAP One on One
Tools	Informatica 7, informatica 8.1, Oracle 9i
Team Size	10
Role	ETL Developer
Period	July 2007 – February 2010

Project Objective

     The client is migrating its applications from Legacy Systems to SAP. So, the project includes creating ETL mappings to populate tables for generating BI reports as well as to generate flat files to support the already existing BI applications.

    Role and Responsibilities
- Creation of Informatica mappings using ETL CoE standards.
- Review of mappings created by other team members.
- Unit Testing and data load
- Implementing Change Requests.
- Allocation of work to team members

Project #2

Title	Informatica ETL Production Support
Client	Hanes brand Inc, USA
Tools	Informatica 8.1, Oracle 10g, Unix
Team Size	10
Role	ETL Developer
Period	July 2007 – February 2010

Project Objective

Production Support was started on a trial-and-error basis and has now matured into a well-defined process with a resolution window of 3 hours in a 12hr on call environment.

    Role and Responsibilities
- Understanding the already existing applications.
- Handling on Call support and providing resolution in the stipulated period of 3 hours for high priority tickets.
- Documentation of ticket details into logs.
- Analyzing the logs for repetitive tickets and finding permanent solutions to prevent their reoccurrence.
- Proactive interaction with the client and the onsite operations team to improve and enhance the support process.
- Guiding the new team members in the support team.