**HASTI**

Email: hastipateldataanalyst0208@gmail.com

732-344-4392

**GitHub**: https://github.com/hastipatel14395

**Education:**

Professional certification in Data Science Rutgers University (2020)

Bachelor of Computer Science & engineering BITS edu campus (2016)

**Summary:**

* Around 4 years of IT experience with Business and Data Analysis, ETL Development, Data Modeling, Data Profiling, Data Migration, Data Integration and Metadata Management Services.
* Worked on various domains like Banking, Finance, Health Care and Manufacturing.
* Strong experience on gathering business requirements from business/user, source to target mapping, creating source to target mapping document, creating process flows, and DFD.
* Excellent understanding of Data Warehousing Concepts - Star and Snowflake schema, SCD Type1/Type2, Normalization/De-Normalization, Dimension & Fact tables.
* Proficient experience as a Data Analyst in gathering data from various sources, data profiling, data definition, and loaded the data to business warehouse.
* Strong experience in Data Analysis, Data Migration, Data Cleansing, Transformation, Integration, Data Import, and Data Export using multiple ETL tools such as Informatica Power Center and Data Stage.
* Develop project management plans, identify key milestones, track metrics and managed stakeholders throughout the lifecycle of multiple projects to ensure adherence to project schedules & budgets
* Worked with heterogeneous relational databases such as Teradata, Oracle, MS Access and SQL Server.
* Experienced in using IDQ tool for applying rules and develop mappings to move data from source to target systems
* Applied the rules and profiled the source and target table's data using IDQ
* Technical expertise in ETL methodologies, Informatica 6.x/7.x/8.x/9.x - Power Center, Power Mart, Client tools – Mapping Designer, Workflow Manager/Monitor and Server tools – Informatica Server Manager, Repository Server Manager, and Power Exchange.
* Experience with Relational Databases like Oracle 10g/9i/8i, DB2 ESE, MS SQL Server 2005/2010, DTS, Sybase, Teradata V2R5, MS Access and formats like flat-files, CSV files, COBOL files and XML files
* Design project plans and teams to deliver on-time, on-budget and on-scope results within risk tolerances without undue project management overhead.
* Expertise in broad range of technologies, including business process tools such as Microsoft Project, MS Excel, MS Access, MS Visio, technical assessment tools, MicroStrategy, Data Warehouse, Data Modeling and Design
* Ability to handle and interact with Business users, collect and analyze business / functional requirements.
* Excellent communication, documentation and presentation skills with clear understanding of business process flow.

**Technical Skills:**

Frontend: HTML, JavaScript, CSS, jQuery, Bootstrap, SharePoint

Backend: MongoDB, MySQL, AWS, C++, Java, Python

**Non-Technical Skills:**

Communication, Analytics and Reporting, Data Mining, Data analysis and Design, Documentation, Relationship management, Data Virtualization, Problem solving

**Work Experience:**

**Client:** EMD Serono, Inc – Rockland, MA **Aug’19 –Current**

**Role:** Data Analyst

**Responsibilities:**

* Participated in user meetings, gathered Business requirements & specifications for the Data-warehouse design. Translated the user inputs into ETL design docs.
* Defined, and documented the technical architecture of the Data Warehouse, including the physical components and their functionality.
* Created Star schema dimension modelling for the data mart using Visio and created dimension and fact tables based on the business requirement.
* Analyze and gather user requirements and create necessary documentation of their data migration
* Designed ETL architecture to process large number of files and created High-level design, low-level design documents.
* Work alongside clients to develop strategies for migration of their business data across platforms utilizing Microsoft SQL Server
* Creating Data Dictionary for the model.
* Maintain the database and data model for any further requirements and changes.
* Redesign and extend the data model for further enhancements.
* Involving in writing code to extract, clean and validate Data from tables.
* Estimate schedules for data modeling activities and complete them on time, adhering to predetermined specifications and quality standards.
* Worked with OLAP tools such as ETL, Data warehousing and Modeling.
* Created new strategies for preventing fraud.
* Used Informatica to extract, transform & load data from SQL Server to Oracle databases.
* Involved in the creation of Informatica mappings to extracting data from oracle, Flat Files to load in to Stage area.
* Worked data mapping, data cleansing, program development for loads, and data verification of converted data to legacy data.
* Worked on Master Data Management (MDM) for maintaining the customer information and also for the ETL rules to be applied
* Skilled at implementing project management processes
* Created MS Excel reports for metadata extracts using MS Excel pivot tables.
* Developed SQL Server views for generating metadata reports.
* Building, publishing customized interactive reports and dashboards, report scheduling using Tableau server.
* Created action filters, parameters and calculated sets for preparing dashboards and worksheets in Tableau.
* Designed and Created data cleansing, data conversion, validation and External loading scripts like MLOAD and FLOAD for Teradata warehouse using Informatica ETL tool.
* Involved in error handling, performance tuning of mappings, testing of Stored Procedures and Functions, Testing of Informatica Sessions, and the Target Data.
* Experienced in loading data into Data Warehouse/Data Marts using Informatica, Teradata Multiload, Fast load and BTEQ utilities.
* Worked with DBA in making enhancements to physical DB schema. Also, coordinated with DBA in creating and managing table, indexes, table spaces, triggers, DB links and privileges.
* Possess expertise with relational database concepts, stored procedures, functions, triggers and scalability analysis. Optimized the SQL and PL/SQL queries by using different tuning techniques like using hints, parallel processing and optimization rules.
* Supported application development timelines by implementing designs as well as incremental changes to database definition in a timely manner into production and non-production Teradata systems.
* Responsible for migrations of the code from Development environment to QA and QA to Production.
* Documented the existing mappings as per the design standards followed in the project.
* Prepared the validation scripts to compare the new data with legacy systems.
* Carry out Defect Analysis and fixing of bugs raised by the Users
* Highly motivated with the ability to work effectively in teams as well as independently

**Environment:** Informatica Power Center 9.x, Oracle 11g, SQL Server 2008, SPSS, PL/SQL, Teradata, UNIX Shell Scripting, Visual Studio, Erwin, Business Objects, Tableau, MS Excel, MS Access, Windows XP.

**Client:** Service Channel, Pleasanton, CA **July’18- July’19**

**Role:** Data Analyst

**Responsibilities:**

* Processed large Excel source data feeds for Global Function Allocations and loaded the CSV files into Oracle Database with SQL Loader utility.
* Performed code Inspection and moved the code into Production Release.
* Documented all the Relative activities in Quality Centre and coordinated with QA team.
* Performed Data filtering, Dissemination activities, trouble shooting of database activities, diagnosed the bugs and logged them in version control tool.
* Project Management, Requirements Analysis Software development lifecycle, full lifecycle documentation, database
* Performed source data analysis and captured metadata, reviewed results with business. Corrected data anomalies as per business recommendation.
* Involved in migration of the mapping from IDQ to Power center
* Created adhoc reports to users in Tableau by connecting various data sources.
* Created source to target mappings for multiple source from SQL server to oracle.
* Used excel sheet, flat files, CSV files to generated Tableau adhoc reports.
* Performed the batch processing of data, designed the SQL scripts, control files, batch file for data loading.
* Performed the Data Accuracy, Data Analysis, Data Quality checks before and after loading the data.
* Designed and developed the Database objects (Tables, Materialized Views, Stored procedures, Indexes), SQL statements for executing the Allocation Methodology and creating the OP table, CSV, Text files for business.
* Maintained a metadata repository which included administration of Data Stewards, and ETL data.
* Performed the physical database design, normalized the tables, worked with Denormalized tables to load the data into fact tables of Data warehouse.
* Designed and developed Dimensional Model and ETL process using informatica.
* Developed SQL Joins, SQL queries, tuned SQL, views, test tables, scripts in development environment.
* Used SQL\*Loader to load data from external system and developed PL/SQL programs to dump the data from staging tables into base tables.
* Developed data migration and data validation scripts from old to the new system
* Extensively wrote SQL Queries (Sub queries, correlated sub queries and Join conditions) for Data Accuracy, Data Analysis and Data Extraction needs.
* Created mapping documents using the Metadata extracted from the Metadata repositories.
* Resolved the data type inconsistencies between the source systems and the target system using the Mapping Documents.
* Designed star schema with dimensional modeling, created fact tables and dimensional tables.
* Involved in data analysis, data discrepancy reduction in the source and target schemas.
* Designed and developed the star schema Data model, Fact Tables to load the Data into Data Warehouse.
* Implemented one-many, many-many Entity relationships in the data modeling of Data warehouse.
* Analyzed the tables once the imports were done.
* Edited the Database SQL files with Vi Editor, used Unix Commands like Job scheduling, Executing files, Process, Background, Grep etc.
* Performed the column analysis using Information Analyzer jobs.
* Written PERL Scripts to extract the data from Oracle Database, checking the database connections, regular expressions etc.
* Developed the VBA Integration with Excel feeds and SQL database, SQL Extraction, Transformations of Excel Data into SQL database.
* Involved in writing Oracle PL/SQL procedures, functions, Korn Shell scripts that were used for staging, transformation and loading of the data into base tables.
* Involved in debugging and troubleshooting of database objects.
* Involved in Data loading and data migration – Used SQL \* Loader to load data from excel file into staging table, Data Cubes and developed PL/SQL procedures to load data from staging table into base tables of Data warehouse.
* TOAD, SQL Developer tools were used to develop programs for executing the queries.
* Worked with SQL Query performance issues.
* Created the XML control files to upload the data into Data warehousing system.

**Environment:** Oracle 11g, Tableau, Data Warehouse, OLAP, SQL Navigator, Visual Studio 2010, SQL Developer, Erwin 4.0, XML, OLTP, MS-Excel 2000, MS-office 2000, Microsoft XP Professional.

**Client:** AETINS Software PVT LTD, India **July’17 – March’18**

**Role:** Data Analyst

**Responsibilities:**

* Responsibilities include gathering business requirements, developing strategy for data cleansing and data migration, writing functional and technical specifications, creating source to target mapping, designing data profiling and data validation jobs in Informatica, and creating ETL jobs in Informatica.
* Designed the data marts using the Ralph Kimball’s Dimensional Data Mart modeling techniques
* Gathered all the Sales Analysis report prototypes from the business analysts belonging to different Business units
* Involved in Data Extraction, Transformation and Loading (ETL) from source systems
* Responsible with ETL design (identifying the source systems, designing source to target relationships, data cleansing, data quality, creating source specifications, ETL design documents, ETL development (following Velocity best practices).
* The data received from Legacy Systems of customer information were cleansed and then Transformed into staging tables and target tables in DB2
* Used External Tables to Transform and load data from Legacy systems into Target tables
* Use of data transformation tools such as DTS, SSIS, Informatica or Data Stage.
* Conducted Design reviews with the business analysts, content developers and DBAs
* Designed, developed, and maintained Enterprise Data Architecture for enterprise data management including business intelligence systems, data governance, data quality, enterprise metadata tools, data modeling, data integration, operational data stores, data marts, data warehouses, and data standards.
* Incremental loading of Fact table from the source system to Staging Table done on daily basis
* Coding SQL stored procedures and triggers.
* Designed different types of STAR schemas like detailed data marts and Plan data marts
* Monthly Summary data marts, Inventory data marts using Erwin (with various Dimensions Like Time, Services, Customers, Sales Hierarchy, Orders Snowflake Dimensions and various FACT Tables)
* Worked on data modelling and produced data mapping and data definition documentation
* Worked on SQL Server 2005 concepts SSIS (SQL Server Integration Services), SSAS (Analysis Services) and SSRS (Reporting Services).
* Designed a reporting table approach to get better performance of complex queries and Views.
* Data maintenance activities like REORG, RUNSTATS, LOAD, UNLOAD, IMPORT and EXPORT.
* Performance & tuning and installation of DB2 on Linux.
* Performance tuning was an ongoing process, tuning queries.
* Monitoring and providing 24x7 supports for the production database.
* **Environment:** Linux, IBM Web sphere 6.0, Erwin, SQL Server 2000/2005, Crystal Reports 9.0, HTML, Data Stage Version 7.0, Oracle, Toad, MS Excel, PowerPoint.

**Client:** MACROSERVE Infotech Pvt Ltd - Vadodara, India **July16 – June’17**

**Role:** Technical Support Specialist

* Provide thorough support and problem resolution or customers, build and maintain successful relationships with the dealers, consumers, and service providers. Supported customers with an online account and billing issues.
* Researched issues various computer systems and databases to resolve the complaints, outline the solutions, and answer the inquires. And informed the customers about issues resolution progress. Maintained composure and patience and face of difficult consumer situations.
* Received customer calls and solves the queries and data connectivity issues.
* Assisted customers in IT-related issues such as the installation of software programs, proper registration of PC or mobile device in the software’s network, operational matters, and problems with the functioning of hardware.
* Communicated effectively and collaborated with the company’s sales teams; shared workload, took inbound calls, and performed other appropriate support tasks.

**Projects:**

**Project Name: Uber in Greater Metropolitan Area (Research) Rutgers University (2020)**

* Collected all the csv files needed for our project and I analyzed and interpreted them for potential visualization.
* We used the matplotlib and pandas frame for this project.
* Designed workflow and plots like Bar charts, Pie charts, Heat map and created data frame with API keys.
* **Tools/Technologies:** MS Office, MS SQL server, HTML 5, CSS, JavaScript, jQuery, Python, Pandas, Matplotlib

**Project Name: NJ State Legislators (ETL Process)**  **Rutgers University (2020)**

* The purpose of this short project is to link the State legislators with local townships that enabled to find out our state and local elected leaders are. Information gathered about NJ State legislators and Mayor - Head of local Township and used in csv, excel format. We used Pandas to read HTML data from NJ state website, csv and xls files from local directory.
* Jupyter notebook is used to read and transform raw data (csv, excel and webpage) into final data frames - equivalent to the database tables. Used lambda functions to cleanup data.
* **Tools/Technologies:** MS Office, MS SQL server, HTML 5, CSS, JavaScript, jQuery, Python, Pandas, Matplotlib, PostgreSQL

**Project Name: Fatal Police Encounter Rutgers University (2020)**

* The scope of the project is to highlight the number of police encounters that have resulted in fatalities.
* Used PANDAS for Data Wrangling and PANDAS and PostgreSQL for Data Loading.
* Created flask app to be served as an API for data visualization and used lambda functions to cleanup up data. Also created a dashboard.
* HTML and CSS, and JS libraries like D3, Leaflet, and Chart.js to create the visualizations.
* **Tools/Technologies:** MS Office, MS SQL server, HTML, CSS, Python, Pandas, Matplotlib, PostgreSQL(pgAdmin), JavaScript (D3, Leaflet and Chart.js)

**Project Name: Sentiment Analysis on twitter Rutgers University (2020)**

* The scope of the project is to explore what is the social media trend on presidential election, compare the presidential candidate’s sentiment standing.
* Twitter API and Twitter data streaming methods were used extract tweet data. We used an hour window every day for seven days to extract tweet data.
* From our raw data inputs, we cleaned and normalized the data with the goal of extracting useful keywords and predict the tweet sentiment.
* Normalized data was then loaded into a Postgres development database prior to being converted into a Heroku database.
* Visualizations were built using variety of JS libraries.
* The dashboard function of the website was built around the interactive visualizations.
* **Tools/Technologies:** NLTK, textblob, Sklearn, tweepy, Pandas, NumPy, TensorFlow, matplotlib, word cloud, seaborn, MS SQL server, HTML, CSS, Python, PostgreSQL(pgAdmin), JavaScript (D3, Leaflet and Chart.js), Heroku, Bootstrap, PyCharm, Visual Studio

**Project Name: Corporate Staff Performance Evaluation System with Hierarchical Approvals Reviews and Ratings.** **BITS edu campus (2016)**

* I have studied the module of our project, design forms and implementation which includes searching for any employee of any department. Designed all E-R diagrams in our project like Use-case diagram, Sequence diagram, Activity diagram, Class diagram.
* Successfully designed all diagrams and canvases like AEIOU summary, Empathy summary, Ideation summary and BMC (Business Model Canvas) report.
* **Tools/Technologies:** Visual Studio, asp.net, MS SQL server, C# language, HTML 5, CSS, JavaScript, jQuery