ANVESH AKULA

Senior Data Analyst

PH: 248-662-7020  
Email: [anveshakula27@gmail.com](mailto:anveshakula27@gmail.com)

**EXPERIENCE:**

* Strongprofessional experience with emphasis on **Analysis, design, development, testing, maintenance** and implementation of Data Mapping, Data Validation, and Requirement gathering in Data warehousing Environment.
* Experience in data-warehousing applications using ETL tools and programming languages like **python, SQL/PLSQL, Oracle and Sql server databases, Informatica, SSIS.**
* Experience in handling huge set of financial and participant 401k data using cloud clusters like **Amazon Web Services (AWS), Azure, Hadoop and Hive DB** and archiving the data. Enabled **CI CD Pipelines** using code pipeline, code build, code deploy, cloud formation, and Lambda for our DevOps pipelines. Used Cloud Formation templates to deploy these pipelines.
* Designed ETL pipelines, data extraction, data managing and scheduled ETL data jobs using **airflow.** Performed configuration, deployment and support of cloud services including **Amazon Web Services**. Established data pipeline using Redshift DB and scheduled ETLs using **Apache Airflow** in server less mode.
* Experience in providing custom solutions like Eligibility criteria, Match and Basic contribution calculations for major clients using **Informatica 10 and reports using Tableau**. Involved in Unit Testing and developed the unit test cases using PyUnit framework.
* Helped various projects **integrating Python with Tableau** and Informatica and provided customized solutions for participant data. Worked on Python advanced packages and modules like Requests, Beautiful Soup, Multithreading, Pandas, NumPy, SciPy, Matplotlib, SQL based database connections JDBC/ODBC and Py-Virtual Environment.
* Working experience with **U- SQL and T- SQL** in creating Azure data lakes and building a repository using HDFS and **Azure Data Factory** for server less data integration and data transformation. Worked on Micro Services deployments on **AWS ECS and EC2** instances. Refactored existing batch jobs and migrated existing legacy extracts **from Informatica to Python based micro services and deployed in AWS** with minimal downtime.
* Expertise in writing **Scripts to facilitate the extraction, transformation and loading of legacy data.**
* Experience in **MS SQL Server 2014/2012/2008 (BIDS)** with Business Intelligence in **SQL Server Reporting Service and SQL Server Integration Service, SQL Server Analysis.**
* Solid database experience on **Oracle 11g/10g/9i/8i/8.0, PL/SQL, SQL \*Loader**, Physical database design, Logical database design, Data Modeling (PowerBuilder), Data Mapping, Cleansing, Data Exports & Imports using **SAP MDM**, Inbound/Outbound Interfaces, Application Support, **Performance Tuning, Security Administration, Resource Management, Software Maintenance, User & Security Management, Oracle Tuning & Optimization.**
* Skilled in **Capital Markets** and **fixed income** market data with broad based background in regular markets. Seasoned in **bank** and **financial services** and exposed to utility, industrial and floating rate bonds. Experience in analysis of **Equities, fixed income, derivatives and wealth investments**, venture and growth capital.
* Proficient in Software Development Life Cycle (**SDLC**), Project Management methodologies and **Microsoft SQL Server** database management. Exposed to all phases of SDLC and in-depth knowledge of **agile methodology, Enterprise Reporting Life Cycle,** Problem management and escalation procedures.
* Expert in **Business Intelligence** solutions in Data warehousing / Decision Support Systems using ETL tools Informatica, OLAP, OLTP and Data Warehouse Reports using Business Objects, Crystal Reports.
* Experience in **Data cleansing and data mining using Python**. Hands on experience in using Python data modules like **NumPy, Pandas**. Exposure to Big Data Hadoop environment and its components **HDFS, Apache Pig, Hive, Map Reduce.**
* Experience in designing **star schema** (identification of facts, measures and dimensions), **Snowflake schema** for Data Warehouse, ODS architecture by using tools like **Erwin data modeler, E-R Studio.**
* Well versed in **Normalization (1NF, 2NF and 3NF)** and **De normalization** techniques for optimum performance in relational and dimensional database environments.
* Highly proficient in the use of **T-SQL** for developing **complex stored procedures, triggers, cursors, tables, user profiles,** relational database models and data integrity**,** SQL joins and query writing.
* Experience in creating **Sub Reports, Tabular Reports, Matrix Reports, Parameterized Reports, drill down Reports** and **Charts ad hoc reporting** using SQL Server Reporting Services (**SSRS**) and web reporting by customizing URL Access. Experience including analysis, modeling, design, and development of **Tableau** reports and dashboards for analytics and reporting applications.
* Expertise in **Power BI, Power BI Pro, Power BI Mobile**. Expert in creating and developing Power BI Dashboards in to rich look.
* Experience on **Cognos BI tools** – **versions 10.2, 8.4, & 8.2** (Report Studio, Workspace, Workspace Advanced, Analysis Studio, Query Studio and Cognos Connection).
* Awarded **certificate** for successful completion of **Google Analytics** Individual and Advanced courses. Successful completion of **data cleaning, data analysis and data visualization courses using Python** under Dataquest.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **Databases** | Oracle 11g, 10g, 9i, SQL Server 2012/2008R2, Hadoop, HIVE, MongoDB, and MySQL, Netezza, SAS, Hadoop, U-SQL, Snowflake, PostgreSQL |
| **Languages** | SQL, PL/SQL, TSQL, Python, HTML, XML, Nodejs, UNIX, JAVASCRIPT |
| **Analytics & Reporting** | Tableau, SSRS, Cognos, Power BI, Crystal Reports, Data bricks |
| **ETL & Cloud Platform** | Informatica Cloud, SQL Server DTS, Visual Studio 2012/2010, Azure Data Lake & Factory, Pandas, AWS S3, Lambda, NumPy |
| **BI Tools** | SSIS, BI Development Studio, Power BI, Visual Studio 2012, Performance Monitor, Power Pivot, Spark, Scala, Kafka, Devops, Asana, Jira |
| **Design Tools** | ERWIN, ER Studio, MS VISIO |
| **Operating Systems** | Windows 8/7/XP, UNIX |

**EDUCATION:**

* **Bachelor of Technology, Electronics and Communications Engineering, JNTU, India.**
* **Master’s in computer science and Information Systems, University of Central Missouri, Warrensburg, MO**

**PROFESSIONAL EXPERIENCE:**

**Business Data Analyst**

**L.A. Care, Los Angeles, CA Apr 2022 – Present**

* Manage projects to connect with client partners to receive and ingest Real-World clinical data from **electronic health records (EHR**), assist partners with the **deidentification** of this data, analyze and map data to the Holmusk normalized data model for use in the **NeuroBlu** platform and other internal applications.
* Conduct data field and code **mapping**, perform data validation and collaborate with cross-functional teams to generate data insights for business development.
* Conduct current and future state analysis for **data exchange** and systems integration using best practices.
* Develop and maintain system protocols and procedures, technical user guides and use cases supporting quality assurance activities.
* Leverage appropriate approaches and modeling techniques to gather requirements, document according to corporate standards and procedures and prioritize for assigned areas
* Utilized Informatica Cloud for designing ETL logical mapping to pull data from data partner sites and load the data periodically to the staging source of **Canonical Data model (CDM**).
* Write complex scripts for mapping to do analysis and create **OMOP concept-based tables** like Person, Visit occurrence, Condition Occurrence, Encounters, Drug related information etc.
* Setting up data quality metrics and reproducible analytics **pipelines embedded in cloud architecture**. Using Linux scripts and **Python commands**, worked on data transfer to **AWS S3** by working on serverless environments.
* Provide dashboard and reports utilizing the **Tableau** to the consumers with OMOP based data tables.

**Technologies**: Jira, Asana, Informatica Cloud, PostgreSQL, SQL, AWS S3, Tableau, Python, ELT data pipelines, GitHub, Docker

**Senior Data Analyst**

**Blue Cross Blue Shield of Kansas City, Kansas City, MO Aug 2020 – Apr 2022**

* Work on multiple projects, benefits, consulting services to clients, provide business solutions, deliver business models and reports, develop programs and applications.
* Analyze performance data for various departments and vendors by identifying potential trends and process errors from data, advise key elements on developing action plans and prepare reports on the findings.
* Perform Automation of test cases using **Q-Test** and helping QA team in Lab and Prod validations. Develop test cases and queries using **SAS, Hadoop, Hive and Netezza** and **Snowflake** for validations and create documentation of S2T, data model validations.
* Utilized **SQL, CDMA mapping, ADO, Data Stage, Azure Data Factory and Tableau** to pull, analyze and present data effectively for quality and ad-hoc analysis. Coordinated with team in building an **Azure Data Lake** storage to handle Facets structured unstructured and semi structured data.
* Automated CM and maintained a **CI/CD data pipeline**. Developed a monitoring application which captures the error related data and stores it in the database. Involved in a scrubbing project, which updates the existing data with hashed values. Reverse engineer and re-implement legacy back-end software into Python with minimal downtime.
* Involved in analyzing the problems using transaction breakdown, network monitoring and resource monitoring to find the performance bottleneck. Written complex SQL queries for data validation based on **ETL mapping** specifications.
* Created a database using **MySQL**, wrote several queries to extract data from the database. Setup automated con jobs to upload data into the database, generate graphs, bar charts, upload these charts to wiki, and backup the database.
* Wrote **Python modules** to extract/load asset data from the MySQL source database. Designed and implemented a dedicated MYSQL database server to drive the web apps and report on daily progress. Developed views and templates with **Python** & view controller and templating language to create a user-friendly website interface.
* Responsible for formalizing requirements and technical architecture documents.
* Performed trend and variance analyses, identified root cause of data & model issues, and lead resolution efforts. Developed **Power BI reports and dashboards** (storytelling) analyzing and tracking KPI’s performance such as data mapping, Authorization data, Patient claims, fraud claims etc.
* Provided accuracy by eliminating error procedures in Authorization of patients and provider data and prepared a business model for increase in revenue of hospital claims.

**Technologies**: SQL Server, Azure Devops, Azure Data Factory, Python, Django, ELT Pipeline, JSON, Tableau, Power BI, Jira, Service Now, Snowflake, Oracle, IBM Data Stage, AQT, SAS, Netezza, Hive, Salesforce CRM, Apex Data loader

**Principal Data Analyst**

**Fidelity Investments, Durham, NC Dec 2018 – July 2020**

* Worked on various client related projects from internal **kick offs until Project goes live**. Helped consultants and technical consultants in **gathering requirements and signing off the BRD**.
* Involved in cleansing raw participant **401k data and loading them to Hadoop clusters using Hive** and then archived the data in online (ODA) Cloud platforms like **AWS and Azure**.
* Achieved 400% growth in data quality check creation process by creating Rule automation tool in **Python**. Performed data analysis on Data Lake of Capital One using **PySpark** on data bricks platform.
* Experience with data quality issues through exploratory data analysis (EDA) using **SQL, Python and Pandas**. Worked on **AWS cloud services and** implemented the requirements on **Salesforce** platform and IDE using Eclipse.
* Involved in the code migration of quality monitoring tool from **AWS EC2 to AWS Lambda** and built logical datasets to administer quality data monitoring on **snowflake** warehouses.
* Contributed in helping to transform the department into Squads using **Agile methodology for projects using Jira, cloud platforms like Azure and AWS.**
* Worked in data extractions using Maria DB, SCF file formats and developed calculations and reports using **Informatica Power designer and workflow monitor**. Helped in **integrating Jira** to the team and developed documentation on using **Jira** for Informatica related tasks.
* Used **Toad** as central tool for connecting different sets of databases like **Maria DB, WIDE, Oracle, SQL and MySQL. Extensive experience in handling databases using Snowflake.** Create an **Azure data model** for the eligibility and Match program and Data security team using and **Azure SQL, U-SQL.**
* Participated in various stages of building Data Lake like **source identification, data ingestion, clean up, creating queries for stage data and visualization using BI tools and Tableau**.
* Developed programs using **Informatica and automated the schedule process using Control-M** for providing a formula-based **401k Eligibility criteria and Match contributions Calcs** per pay period.
* Helped in automating the **reports using Tableau and Python scripts**. Developed Python scripts for **integration with Tableau and Informatica**.
* Worked on providing client related support for automated process of eligibility and match calculations, helped in **Match true up calculations using Python scripts**.
* Worked on providing holistic approached benefits design, creating a customer centric **defined benefit plan data, delivered business models for personal benefits and investments**, scheduled multilayered job models for defined benefit (DB) and defined contributions (DC) data.
* Provided end to end solutions to the yearly services on personal benefits data of major clients like **Tenet health, Stanford Health care and Broadridge Financial solutions.**
* Developed personalized benefits access in **dashboard models using Tableau** and made separate templates for each benefit and contribution.

**Technologies:** SQL Server 2014/12, Oracle 11g, 12c, Python, Databricks, Spark, Snowflake, Eclipse, Pandas, AWS EC2, S3, Lambda, TOAD, Maria DB, Informatica 10, Power designer and workflow monitor, Hadoop, AWS, Hive, Azure, Jira, Tableau.

**Senior Data Analyst**

**Boardwalk Pipeline Partners, Owensboro, KY Jul 2016 – Nov 2018**

* Involved in **Data Migration** from Oracle Legacy to SQL using SQL Navigator and SQL Management Studio. Involved in **Analysis, profiling and cleansing** of legacy data and hence making the data available for further use.
* Created **STTM documentation** for data mart integrating different source systems. Involved in documentation of **Data Mapping & ETL specifications** for development from source to target and implemented the change Data Capture (CDC) while integrating the enterprise data sources.
* Assisted to build analytic tools to manage data and streamline data analyses **using Python and SQL.** Complex incoming **data cleaning and formatting** using Python. Ensured data accuracy through the creation and implementation of **data integrity queries**.
* Developed **queries** to build a new database in SQL using SSIS and **migrated the data from the legacy**. Created **batch files** that execute the ETL scripts which invokes SSIS packages. Created SSIS Packages using **Pivot Transformation, Fuzzy Lookup, Derived Columns, Condition Split, Term extraction, Aggregate, Execute SQL Task, Data Flow Task, and Execute Package Task etc.** to generate underlying data for the reports and to export cleaned data from Excel Spreadsheets, Text file, MS Access and CSV files to data ware house by using BIDS.
* Prepared **documentation** of Data Migration and Data Mapping with its specifications for development from source to target. Facilitated in **developing tests and procedures** and with **UAT testing**. Used **TOAD and SQL navigator** as a part of business requirement.
* **Prepared scripts** to ensure proper data access, manipulation, and reporting functions with Python and SQL. Involved in modifying the existing Reports based upon the change requested by the **users**. Involved in resolving **Performance** issues with various reports.
* Created reports in **Business Insight** Advanced and Report Studio. Manage Departmental Reporting systems, troubleshooting daily issues, and integrating existing. Access databases with numerous external data source including (SQL, Visual Studio 2010, Excel, & Access).
* Designed and implemented multiple dashboards using Power BI - Power Pivot & Power Query tools for in house metrics.

**Technologies:** SQL Server 2014/12, Oracle 11g, 12c, Python, TOAD for SQL, SQL developer, MS Office, SSIS, SSRS, Cognos, Hive, Hadoop, Windows 7.

**Senior Data Modeler/Analyst**

**TransAmerica, CedarRapids, IA Sep 2015 – Jun 2016**

**Responsibilities:**

* Involved in **Analysis, profiling and cleansing** of source data and understanding the business process. Involved in extensive **Data Validation** using **SQL** queries and back-end testing.
* Developed ER and Dimensional Models using **Erwin** advanced features. Created Conceptual, logical and physical data models using **Erwin**. Identified facts and dimensions based on KPI and identified the source tables. Worked with various process improvements, **normalization, de-normalization, data extraction, data cleansing, data manipulation.**
* In-depth analysis of **balance sheets, income statements, cash flow statements**, statement of schedule, capital call statements, distribution statements to abstract necessary entity performance data.
* **Forward Engineered** the physical data model and generated DDL script using Forward Engineering option in **Erwin.**
* Extract data from multiple systems, conduct detailed data analysis, and load data into **SAP-MDM**. Perform detailed data mining, analysis, and functional interpretation of Confidential Master Data.
* Create and map transactional script templates using SAP field names in **Win shuttle Studio**. Maintain the data integrity in **SAP** by tracking asset dispositions, performing inventories, and updating SAP for all specified sites.
* Facilitated in developing testing procedures, test cases and User Acceptance Testing (UAT). Developed **regression test scripts** for the application and Involved in **metrics gathering, analysis and reporting** to concerned team and tested the testing programs.
* Worked extensively **in Sybase for generating business reports**. Writing **Unix Batch jobs** to check for data integrity at scheduled time between database servers. Worked extensively on **Autosys** for overnight job monitoring and support.
* Mastered the ability to design and deploy rich **Graphic** visualizations using **Tableau**. Developed **Star** and **Snowflake** schemas based dimensional model to develop the data warehouse.

**Technologies:** Erwin 9.5, Oracle 12c, JDE 9.0, SQL Developer, Autosys, Toad for MySQL 6.3, MS Office, SAP MDM, Winshuttle, Crystal Reports, Tableau, UNIX.

**Data Modeler/Analyst**

**JP Morgan Chase, Dallas, TX Jan 2015 – Aug 2015**

**Responsibilities:**

* Worked with the Business Analyst and DBA for requirements gathering, business analysis, and testing and project coordination. Performed data manipulation, data cleansing and data profiling in the source systems that are required for Claims **Datamart**.
* **Analyzing** raw data, drawing conclusions & developing recommendations also writing **SQL** scripts to manipulate data for data loads and extracts. **Analyzed and evaluated** existing system for assessing how it supported functions.
* Implemented Referential Integrity using **primary key** and **foreign key relationships.** Involved in **Dimensional modeling** of the Data. Identified **the Facts and Dimensions**.
* Produced **functional decomposition and logical models** into an **enterprise data model**. Employed **process and data analysis** to model a Customer Information Business System.
* Consulted with client management and staff to **identify and document** business needs and objectives, current operational procedures for creating the logical data model. Developed **Logical data model** using **Erwin**.
* Tested the database to check field size validation, check constraints, stored procedures and cross verifying the field size defined within the application with **Metadata**.
* Used Power BI, Power Pivot to develop data analysis prototype, and used Power View and Power Map to visualize reports
* Published Power BI Reports in the required originations and Made Power BI Dashboards available in Web clients and mobile apps

**Technologies:** Oracle 10g, Toad, ERWIN, Crystal reports 9, Windows XP, Informatica Power Center, Teradata, DB2.

**Data Analyst**

**Nationwide Financial services, Columbus, OH Oct 2014 – Dec 2014**

**Responsibilities:**

* Designed and developed Extract, Transform, Load (**ETL**) processes using Base **SAS** and **Data Integration** Studio to populate **Data warehouse’s** and **Reporting** Data Marts.
* Modeled new tables and added them to the existing data model using **Power Designer** as part of data modeling.
* Worked in designing and optimizing various **T-SQL** database objects like tables, views, stored procedures, user defined functions, **Indexes** and **Triggers**. Worked on pre-existing macros for data validation by checking data distribution and comparison to standard data.
* Extracted data from different sources like claims data mart and text files using **SAS/Access, SAS SQL** procedures and created **SAS datasets**.
* Performed data cleaning by analyzing and elimination duplicate and inaccurate data using **PROC FREQ, PROC UNIVARIATE, PROC RANK PROC, TABULATION and macros in SAS.**
* Developed Strategies for **Data Analysis** and **Data Validation** with the help of **Pl/SQL**, **SQL and UNIX**.
* Analyze source systems and target data requirements and identify gaps in data, document issues and seek resolutions. Reviews mapping documents provided by Business Team and implement the business logic into the **UNIX and PL/SQL** scripts that load data structures in database.

**Technologies:** Oracle 10g, SQL, ERWIN, Crystal reports 9, SAS, Teradata, Sybase power designer, Windows 7, UNIX, Informatica Power Center