Arun kumar Jegarkal

arunjeg@gmail.com

(806) 620–8017   **https://arunjegarkal.github.io/arunjegarkal/**

Summary: Developer with 6+ years of extensive experience in Python, Amazon Web Services, Microsoft Azure, Spark, Tableau, Power BI, Data Analytics and ETL technologies; worked for various clients and projects in every phase of Software Development Life Cycle, from business requirement gathering to project delivery. Excellent design and integration problem solving skills.

Professional Experience

Sr Data Engineer, Peterbilt Motors Company(Paccar Inc.), Denton, TX (2021 Dec - Present)
Data Engineer (2019 Nov –2021 Dec)
Responsibilities
- As a Data Engineer I focus on making possible fast, accurate, and reliable access to data.
- Build data pipelines, manage a data warehouse, and support the production use of our data.
- I interact with business users and data scientists to gather the requirements and develop the AWS infrastructure and deployment pipelines.
- Also, I will involve in data migration process to move data from OnPrem Data sources to cloud platform using ETL tools and Apache spark.
- As part of responsibilities, I develop tableau dashboards, snowflake views, schemas, and python modules.
- I also build a continuous integration and continuous development pipelines for machine learning models using Jenkins and Git Actions.
- Experienced in developing web applications using python flask to easy end user interaction with backend snowflake data and ML model initiation.
- Use the knowledge of Agile / Scrum methodology as a development process.

Business Systems Analyst, Netrush, Vancouver, WA (2018 Oct – 2019 Nov)
Primary Duties
- Developing Azure ETL Pipelines to extract data from Azure storage and transform into dimensional data and load into analytical database.
- Scheduling ETL pipelines using Azure Automation.
- Developing Power BI reports and dashboards to transfer insights in data to visuals.
- Writing Runbook scripts in Azure Portal to Auto scale Power BI premium Tier and trigger reports dataset refresh.
- Dimensional Data Modeling using star schema. And writing SQL Quires and Procedures to pull reports.
- Peer review reports and dashboards before publishing to production.
- Interacting with Stakeholders and End users to collect User Requirements.
- Translating User Stories into development features.
- Testing the Netrush internal developed tools, Reports and Dashboards.
- Writing release, user documents as a guidance of the tools for End Users and Stakeholders.
- Experience in agile methodology.

Reporting Analyst, Modis, Beaverton, OR (2018 Oct – 2018 Oct)
Primary Duties
- I was responsible for developing SQL Server Reporting Services (SSRS) reports based on requirement specification document.
- Experienced in writing complex SQL quires to pull data from source database.

Graduate Assistant at Athletics Department, Texas Tech University, Lubbock, TX (2016 Dec – 2018 May)
Responsibilities

- I developed tableau dashboards and stories to get insights and trends of expenditures and athletes' performance in each game and in their academics.
- I created charts, reports and motion charts using tableau.
- Deployed dashboards to the Tableau public and used tableau bridge to keep the data UpToDate at the server side.
- Scheduled Tableau server to send reports via email.
- Involved in writing SQL queries, PL/SQL procedures and functions to get data from various data sources.
- Also performed logical and physical Data Modeling using normalization Techniques.
- Performed pivoting, clustering, created calculated and parameter fields to model the data according to the business needs.

Software Engineer S2-Data Warehouse Analyst at EVRY India Pvt Ltd, Bangalore, India (2014 Aug – 2016 Aug)
Primary Duties
- I was responsible for the Extraction Transformation and Loading (ETL) data from multiple sources into Data warehouse.
- Extensively worked on designer to designing and developing mappings from varied transformation logic using Expression, Source Qualifier, Filter, Router, Joiner, Lookup, Sorter, Aggregator and Update Strategy etc.
- Designed ETL logic for Dimension Tables using SCD – Type1 and Type2.
- Worked with heterogeneous data sources like DB2, Flat Files, and Mainframe Files.

- Experienced in identifying and solving performance bottlenecks in various levels like source, target, mappings and session levels.
- Experienced in SQL and Informatica Cloud services and working experience on PL/SQL and UNIX commands.

Projects

- **Engine Reliability Analysis –** This project was initiated as a directive from PACCAR's executive team. The MX13 EPA17 engine family is experiencing multiple catastrophic failure modes which have resulted in a significant increase in warranty costs. These issues also have a profound impact on the customer downtime, and the customer's perception of PACCAR's quality. I was part of Data Modeling, AWS Infrastructure setup and developing deployment pipelines.
- **Truck Traffic –** This project is to understand the level of truck activity in a given dealership location's territory. This insight produces more precise recommendations for the need of more (or possibly less) of the following in a given dealership's area/market: Number of dealerships, Mobile Service units, Outside Parts Salespeople, Service operation hours etc.
- **Materials Parts Summary –** Project to identify the number of trucks with part shortages by build day and allows the plant to adjust build rates to reduce the impact of part shortages.
- **Offline no shortage priority –** Management spend time using EWI data to lookup defects and shortages on trucks to create priority lists to guide the offline flow to park trucks, the process of creating these lists is time consuming and the lists are not up to date by the time they reach the operators. I built a dashboard that will help the management to quickly see the number of defects on the trucks. Identify and prioritize low defect trucks and work on those so that those are ready to deliver. See the truck parking location in the lot so that drivers can bring to plant for the work.
- **Netrush Portal –** This is a Netrush Analytical and reporting project, focused on developing reports and dashboards to get insights of sales, inventory, compliance, violations, pricing etc.,
- **Netrush Platform –** This is a Netrush in-house retail domain project, focused on developing e-commerce platform. I am responsible for gathering user requirement and translating into development features. Also involve in testing tools.
- **National Eligibility Database (NEDB Carrier Loads) –** This is an Insurance domain project for the client HMS. The goal is to use proprietary logics to avoid cost from the government payers and recovery from the commercial or third-party insurance vendors. For which we received files on timely from all the vendors and performed ETL operations to uniform the data collected, before applying match processing.
- **Geospatial Visualization of Twitter Data –** This is a National Science Foundation funded research, it is focused on building a model for researchers to analyze trends in Twitter using machine learning algorithms for detecting the topics discussed and obtaining people attitude towards the discussions. Moreover, to find co-occurring topics and its evolution over a period.

Technical Skills

Programming languages	:	Python, Java, C, SQL, PL/SQL, Shell scripting, Data Analytic Expressions (DAX)
Web Designing Skills	:	HTML5, HTML, CSS, JavaScript, PHP, XML, JSON
Database	:	Oracle 9i, SQL Server , Neptune GraphDB
Cloud Experience	:	**Microsoft Azure :** Data Factory, Azure Storage, Azure Automation **Amazon Web Services :** S3, Lambda, SQS, SES, SNS, Step functions, Cloud Front, SSO, Cognito, EC2, EKS, ECR, Glue Etc. Snowflake, Docker, Spark, Jenkins, Terraform, Gremlin, Git Actions
BI Tools	:	Power BI Pro, Power BI Desktop, Tableau, Tableau Desktop, Tableau Bridge, Tableau Server and Tableau Public, Google Analytics, Looker, QuickBase, D3.js
Tools & Utilities	:	Git, VSTS, Attunity, Apache Spark, Informatica PowerCenter Client 9.5.1, Informatica Data Explorer (IDE), Informatica Cloud, SQL Developer and Toad, Power Exchange

Education

Master of Science in Computer Science	**Graduation: May 2018**

Texas Tech University, Lubbock, TX

Bachelor of Engineering in Computer Science	**Graduation: June 2014**

Visvesvaraya Technological University, Belgaum, India

Diploma in Computer Science	**Graduation: June 2011**

R N Shetty Polytechnic, Belgaum, India

Honors, Awards and Achievements

- Part of SQL steering committee at Peterbilt objective to spread SQL knowledge across the company.
- Geospatial Visualization of Twitter Data - Poster presentation at national workshop for NSF grant support.
- Awarded competitive scholarship: 75% fee waiver by Texas Tech University.
- WebMaster at Texas Tech India student Association 2017-18.