

Vishwas Arunkumar

Data Engineer

An engineer with over 14 years of experience in designing and building data platforms and a MSc in Data Science.

I enjoy working with data and building simple scalable solutions



Dubai, UAE

vishwas.ma@gmail.com

SUMMARY

- Building self serve data platforms
- Building data lakes and data marts
- Batch and real time ELT
- Data modeling (OLAP and OLTP)
- Productionizing ML models and services
- Collaborating with data scientist and business users
- Defining a MVP and organizing tasks into sprints
- Managing and mentoring teams
- Retail, Finance, Insurance Domains
- **Languages:** Python, SQL, Golang
- **Cloud Platforms:** AWS
- **Databases:** Snowflake, Oracle, MS SQL Server
- **ETL:** Airflow, SSIS
- **Data Modelling:** DBT, Data Vault, Star Schema
- **DevOps:** Docker, Terraform
- **Education:** Msc in Data Science (University of London)

EXPERIENCE

River Island

Senior Data Engineer
London, UK

MAR 2022- Present
(> 1 year 3 months)

Role: Building a self serve data platform on AWS

Tech Stack: Airflow(2.4) on ECS, AWS Batch, DBT, Snowflake, Terraform, Python

- Developed a repeatable pattern to productionize ML inference jobs
- Defined standards (compute, security, access) for resources on Snowflake and developed Terraform modules using the Snowflake provider to help teams spin up infra on Snowflake as per set guidelines
- Used DBT to build a Data Vault warehouse (hubs, satellites and links) with incremental models that were used as sources for start schemas serving the presentation layer
- Developed DAGs in Airflow for ELT into snowflake with E2E tests with fake data that ran in the CI
- Mentored other product teams to onboard them onto the data platform and enabled them to self serve

LYST

Senior Data Engineer

London, UK

MAY 2021 - FEB 2022

(9 months)

Role: Working on content personalization to drive user engagement

Tech Stack: AWS Batch, Step Functions, Terraform, Snowflake, Python

- Productionized user-product recommendation models using AWS Batch and AWS Step Functions
- Developed a service in Python to expose user-product recommendations as an API to the app
- Set up A/B tests for different versions of the recommendation models
- Helped define the overall strategy and direction of the newly formed content personalization product team
- Transformed business requirements into technical deliverables that were scoped, estimated, prioritized and broken down into sprints

River Island

Senior Data Engineer

London, UK

FEB 2018 - MAY 2021

(3 years 3 months)

Role: Building a self serve data platform on AWS and worked on solving problems around stock replenishment and allocation

Tech Stack: Airflow on ECS, Lambdas, Snowflake, SQS, Terraform, Python, Golang

- Collaborated with architects and product owners to develop the overall vision of the data platform
- Developed a framework for validation of data ingested into the data lake using AWS Lambda functions
- Developed a pipeline to expose real time events (SQS, Lambda) from a legacy oracle OLTP to the data lake (S3) in AWS
- Developed Airflow DAGS for batch ELT of data
- Collaborated with data scientists to productionised demand forecast models to drive allocation and replenishment of products to stores.

Tata Consultancy Services

(6 years 10 months)

Lead Data Engineer

Birmingham, UK

JULY 2016 - SEP 2017

(1 year 3 months)

Role: Working with actuaries to build new insurance products

Tech Stack: Oracle/SQL

- Designed and developed Oracle PL/SQL modules for new annuity products based on statistical models
- Developed python scripts to process and load semi-structured statistical data into the database
- Collaborated with actuaries and technical architects
- Documented technical specifications and performed code and design reviews

Lead Data Engineer

Bangalore, India

JUL 2015 - JUN 2016

(1 year)

Role: Building a new data platform for the reporting of hedge fund data

Tech Stack: Oracle/SQL

- Designed and developed a scalable framework for reporting of hedge fund investments
- Designed and developed PL/SQL modules for bulk processing of data for financial statements

Senior Data Engineer

DEC 2010 - JUL 2015

Bangalore, India/Melbourne, Australia

(4 year 7 months)

Role: Building a platform for data analysis

Tech Stack: Oracle/SQL

- Developed data marts and ETL scripts in PL/SQL to aid downstream analysis
- Collated data from multiple sources into a data mart with partitioned tables and materialized views.
- Implemented a framework for parallel extraction and processing of Terabytes of data

Senior Software Engineer

DEC 2010 - SEP 2013

(2 years 10 months)

Role: Building a windows client for managing superannuation

Tech Stack: Microsoft SQL Server|SSIS

- Designed a framework for ETL of structured and semi-structured data using SSIS
- Developed batch processes in TSQL to perform financial transactions and reconciliations

Accenture

JUN 2006-NOV 2010

Software Engineer/Senior Software Engineer

(4 years)

Bangalore, India

Role: Backend development for a web app

Tech Stack: Microsoft SQL Server|SSIS