

# Sanjith Venkatesh

[sanjith.venkatesh@gmail.com](mailto:sanjith.venkatesh@gmail.com)

(714) 496-0804

[sanjithv.com](http://sanjithv.com)

American Citizen

## WORK EXPERIENCE

### Veeva Systems

Pleasanton, CA

#### Data Engineer

January 2022 to October 2023

- Constructed data pipelines to allow clients to ingest data from their databases and other company products to provide tailored solutions to various problems
- Maintained and extended a Python SDK to allow developers to access and modify data stored on S3 and Redshift
- Developed AWS Redshift stored procedure to migrate multiple tables from one database to another while maintaining security parameters, moving
- Converted Pandas DataFrames into AWS Redshift tables by interpreting data types in Python and constructing appropriate SQL query to execute using psycopg2
- Enhanced data platform in Java to allow for multiple SQL statements to generate custom table metadata, decreasing generation times by 85%
- Optimized existing customer databases by writing internal Python logic to detect where columns can be trimmed, decreasing table sizes by 24%
- Wrote help documentation for Python SDK through Jupyter Notebooks to demonstrate common use cases as well as how to get started regarding AWS infrastructure

#### Software Engineer

July 2020 to December 2021

- Automated manual Git workflows, cutting processing times by 80% with Jenkins pipelines
- Generated new pipelines for the automatic creation of isolated environments automatically through Jenkins using Terraform and Python
- Created JSON metadata modification scripts in Python for dev-integration environments, maintaining the sanity of each file to ensure no duplicate files
- Eliminated discrepancies in VPC pairing for testing environments by writing Python unit tests to identify duplicate VPCs and invalid pairings
- Developed database migration scripts in bash to assist in migrating 50+ customer databases to new product versions during production deployments
- Assisted with weekly production deployment preparation and work with IT and development teams to ensure no customer disruptions

### Boston Consulting Group Platinion

New York City, NY

#### Data Engineering Intern

June 2019 to September 2019

- Designed and implemented name similarity models, using Pyspark and PostgreSQL, that would identify potential unique customers across 5 independent databases and ranging from 1 to 7 million records.
- Contributed to Testaton source code by adding additional tests to allow corporations to run them with just a JSON file as well as a validator to check for syntax and anomalies

---

## SKILLS

Experience: **Python** **Java** **Jenkins** **Groovy** **Git** **MySQL** **AWS EC2, S3, SQS, Redshift** **R**

Exposure: **HTML/CSS** **NodeJS** **Express JS** **Neo4J** **C++** **React JS**

---

## EDUCATION

### University of California, Irvine

2016 - 2020

B.S. Software Engineering, B.S. Data Science , Irvine, CA

Co-Founder - Blockchain @ UCI

---

## PROJECTS

YayNay - Simplify Voting [yaynay.avosapps.us](http://yaynay.avosapps.us)

Vue.js app powered by Leancloud and Express JS designed to help organizations run meetings more efficiently by organizing votes.