

Nitin Jain

Email: nitinj025@gmail.com | Phone: +1 (585) 733-6042 | LinkedIn: <https://www.linkedin.com/in/nitinj025/>

Technical Skills

- **Programming & Scripting Languages:** Python, Ruby, SQL, PL/SQL, Shell Script, Terraform, Ansible, helm chart.
- **Cloud & Data Technologies:** *GCP, AWS* (S3, DataPipeline, Lambda, etc.), *Azure* Data Factory, Databricks, *Snowflake, BigQuery*, Redshift, Starburst Galaxy, *Datadog, Monte Carlo*, Looker, Airbyte.
- **Data Management & Analytics:** Apache Airflow, Terraform, Kafka, Pub/Sub, Cloud Functions, Kinesis, Lambdas, SQS, Cloud Composer, Data modeling(Star Schema), Data Marts, Data lake, *data governance*, GDPR, CCPA, *Data mesh architecture*, GenAI, AI/ML models for predictive modeling, Bedrock, Vertex AI.

Work Experience

| | | |
|--|---|----------------------------|
| Sevenrooms, New york | Staff Data Engineer (Only DE in Team) | Aug 2022 - Present |
| <ul style="list-style-type: none">• Spearheaded the design and construction of Google Cloud Platform (GCP) data infrastructure at SevenRooms, optimizing data management and analytics processes that helped save company approx. \$600K/Year.• Implemented robust <i>data security and governance</i> policies ensuring compliance and data integrity, enhancing trust in data assets using Data plex and dynamic data masking.• Introduced <i>proactive data quality checks using Monte Carlo</i> simulations, fortifying data reliability and preempting potential issues.• Orchestrated efficient data workflows utilizing GCP services such as Google Cloud Composer, Google Dataflow, and Airbyte, alongside Kubernetes for scalable infrastructure deployment.• Collaborated closely with <i>AI/ML Engineers to leverage BigQuery AI and Vertex AI</i> for advanced analytics, offering MLOps support and enhancing operational efficiency in model deployment and management. | | |
| Appfolio Inc. Santa Barbara, CA | Sr. Data Engineer – > Data Manager (3 People) | Jan ‘20- Aug ‘22 |
| <ul style="list-style-type: none">• Led the transition to real-time data processing, utilizing <i>Kubernetes (K8s), Kafka, and Debezium</i>, enhancing agility and responsiveness in data analysis for production data and business services. This initiative involved <i>close collaboration</i> with various teams to ensure smooth integration and alignment with organizational goals.• Led the transformation of legacy and current systems into Infrastructure as Code (IaC) using <i>Terraform</i>, ensuring scalability and reproducibility while employing Ansible for streamlined system configuration.• Architected the establishment of a comprehensive data lake, enabling centralized and democratized data access across diverse business units.• Achieved a remarkable <i>50% reduction in Snowflake usage</i> costs through meticulous query optimization and the implementation of industry best practices, enhancing operational efficiency and cost-effectiveness.• Strengthened data security measures by implementing <i>Role-Based Access Control (RBAC) on Snowflake</i>, ensuring strict control and confidentiality over sensitive data assets. Additionally, pioneered the establishment of an infrastructure for efficient data sharing and monitoring, facilitating seamless collaboration and insights delivery with the company's clients and partners. | | |
| MindBody Inc. San luis Obispo, CA | Data Warehouse Engineer II | Nov 2017 - Jan 20 |
| <ul style="list-style-type: none">• Architecting and designing ETL processes to consolidate 100,000 databases into central databases hosted on an AWS RedShift cluster. This involved utilizing Python, SSIS, and Shell Scripting.• Working extensively with AWS services such as CloudFormation, S3, Data Pipeline, DynamoDB, RDS, RedShift, Lambda Function, SNS, and VPC to manage and optimize data infrastructure.• Integrating APIs from various data sources including SalesForce, Internal MB API, and Stripe, processing millions of data rows daily to ensure seamless data ingestion.• Conducting data analysis by writing complex Advanced SQL queries to perform data analytics, providing valuable insights to end-users.• Implementing dimensional modeling and 3NF on databases to enhance performance and efficiency for central databases and reporting purposes, utilizing Tableau for visualization.• Utilizing Python for QA purposes, writing complex ETL scripts and analytical functions to ensure data quality and integrity within the AWS RedShift database. | | |
| Bank of America, Infosys Limited, Los Angeles | Tech Lead (4 People) | Aug 2016 - Nov 2017 |
| <ul style="list-style-type: none">• ETL (Oracle/Data stage): Automated 80% of ETL code from scratch using PL/SQL. Design and Transforming Data from Legacy system into the Landing-Staging-Target. (Data Marts, Oracle PL/SQL, Data Stage)• Requirement gathering and reporting for LOB on a daily basis. - (BRD, Data Lineage, conceptual, Logical, physical model and creating interactive ad-hoc Tableau report)• Data modeling: Designing Conceptual, Logical and Physical Model on Mortgage Foreclosure domain (Power Designer). | | |
| Florida International University, Miami (ITPA- Informed Traveler program and applications) | Research Analyst | Apr 2015 - Apr 2016 |
| <ul style="list-style-type: none">• Bus arrival and Parking space optimization: Created dashboard for notification, ETA & parking garage occupancy to monitor real-time status• Implemented predictive algorithm for real-time Bus arrival & Parking space (Python)• Reduced API response time for bus location service by 85% via state memory caching SQL query optimization• Designed & built Database (RDBMS in MySQL), Backend Services (Python, SQL), RESTful API (Spring)• Developed mobile app for notification services in real-time (JQuery, HTML5, CSS, JS) | | |
| Infosys Limited, Hyderabad Area, India | System Engineer | Jun 2011 - Dec 2014 |
| <ul style="list-style-type: none">• RICE / Oracle Alerts / Register custom applications / Created concurrent programs / Define responsibilities / users. | | |

Education

- Masters of Science, Information System, May 2016 - Florida International University, Miami, FL (GPA-3.9/4)
- Bachelors of Technology, Computer Science, April 2011 - Narsee Monjee Institute of Management Studies, India