Anshuman Jindal
Senior Software Engineer

+91-98153-67444 | ajanshu.1996@gmail.com | www.linkedin.com/in/ajanshu

Professional Experience

➢ **Data Engineer @ GroupM Media India Pvt. Ltd, Gurugram** **Aug 2023 - Present**

Project- Advertisement & Marketing Web Analytics: Extensible solution for integrating the web analytics data like clicks, impressions, campaigns, actual costs, engagements, conversions, etc. related to advertisement & campaign marketing from multiple web platforms into snowflake and providing with the capacity to visualize the data on Power BI. Working with Global and European client's team.

- Engineered and integrated web analytics data from multiple platforms like Google Adwords, Google Analytics, Meta, Bing, LinkedIn, Pinterest, TikTok, etc. into Snowflake using batch processing and visualized it on Power BI to help business for useful insights and to take decision.
- Developed robust data pipelines using APIs, platform custom reports, Python, AWS S3, Step Functions and Cloud Watch, leading to efficient data processing in terms of performance & data correctness for European clients by providing 98 % data accuracy and completeness.
- Implemented strategic planning, data profiling and backend development by successfully incorporating new advertising accounts into the ETL pipeline and providing business and media team with the transformed data which is used to decide future spends and commit budget.
- Resolved data inconsistencies and performed RCA, code fix for existing implementation, providing high standards of data quality and integrity.
- Developed and automated various customized reports using Python, SQL and Excel providing insightful data to business development team.

➢ **ETL and Big Data Developer @ Amdocs Development Centre India LLP, Pune** **Feb 2019 - July 2023**

Project - Amdocs Data Hub: Built to handle Big Data, ADH provides a carrier-grade solution that accumulates, manages, and transforms massive volumes of data from multiple different sources like oracle, kafka topics, flat files, etc. to deliver a real-time and robust data management foundation for analytics. Worked with offshore clients from Jakarta, Indonesia, and Brazil.

- Planned and developed landing-source, staging and target entities as per product specification & data modeling to create the required Kafka topics, HBase tables and hive tables for the real-time data processing with 60 seconds latency as per SLA.
- Conducted end to end testing by ingesting & transforming data from multiple sources, like kafka-topics, oracle, third party flat files into hive.
- Assisted in creating custom spark jobs for Data Quality & Reconciliation Framework, using this recon engine the recon running time efficiency was improved by 66% as compared to SQL engine for billions of source & target data.
- Contributed to development of ongoing production job Monitoring Framework for L1, L2, L3 support using Unix, PySpark, MySQL as operational DB and Grafana dashboard for front end which enabled operational and performance monitoring for ops team.
- Implemented jenkins-based auto deployment workflows, improving continuous integration/continuous deployment (CI/CD) process.
- Re-designed, developed and improved existing ongoing process with code fixes and enhancements decreasing incident count by 20 %.
- Inspected and tuned the complex SQL queries which provide useful insights and customized reports to business in real-time.
- Analyzed and stabilized the production environment, after successfully managing the data migration for 22 million subscribers and 7 billion rows, by providing root cause analysis, appropriate code fix and re-processing of data for ongoing issues.

Project - GALAXY ODS: ODS is a operational data store designed to integrate data from OLTP system which as a source for the enterprise data warehouse and reporting to provide useful insights to client. Worked with offshore clients from Peru, Chile, and Argentina.

- Modeled CDC mapping for extraction of DML changes from the redo logs on production data base and storing it in landing data base.
- Developed landing, staging, target schema & entities in oracle along with ETL jobs in IBM datastage and SQL scripts for data processing.
- Established control M chains for automated job scheduling along with monitoring of current production batch jobs for continuous data flow.
- Created an Android app for monitoring production logs on mobile phones using MIT App Inventor for real-time job monitoring.
- Introduced purging policy in ODS by removing historical data by understanding customer requirements thus reducing operational cost by 4%.
- Automated many routine tasks like daily monitoring, sanity, and data synch. using linux scripting, saving 30 hours of monthly human effort.
- Programmed new enhancements, CRs as per customer request for integrating useful data required for reporting purpose.
- Maintained migration activities to deploy the new enhancements in UAT, BUAT and production environments for user access.

➢ **Programmer Analyst Trainee @ Cognizant Technology Solutions India Private Ltd, Pune** **Oct 2018 - Jan 2019**

- Completed training modules in SQL, Core Java and performance engineering as a top performer demonstrating strong foundational tech skills.

Education

Bachelor of Technology – Computer Science and Engineering **Aug 2014 - Jul 2018**

Department of Computer Engineering, Punjabi University, Patiala CGPA: 7.42

Skills and Tools

Programming Languages: Python, Linux, Core Java, C/C++
Databases & DWH: Snowflake, Oracle, MySQL, SQLite, HBase, Hive, Teradata
Frameworks & Tools: AWS, JIRA, Hadoop (HDFS, YARN), Spark(PySpark), Visual Studio, Power BI, Grafana, SVN, Perforce, Git, Control-M, CDC, Attunity, IBM Infosphere Datastage

Achievements

- Represented Data Quality and Reconciliation framework in Amdocs's Innovation Festival 2021 and Won Apple watch.
- Went onsite in 2020 for business meetings and supporting production migration activities.
- Performed at many events as a guitarist with International flute maestro Ustad Dr. Mujtaba Hussain.