**Name**

City, ST ZIP | Phone | Email | LinkedIn URL

***Lead Data Scientist***

Innovative IT professional with experience designing and developing complex big data solutions. Lead projects through all lifecycle development phases. Utilize relational and dimensional data modeling techniques. Design and implement data ingestion pipelines to read from multiple sources. Team player, leveraging excellent communication skills to serve as subject matter expert (SME) and handle escalated troubleshooting. Areas of expertise include:

Requirement Gathering | Code Development | Production Support | Team Leadership | Upgrade Implementation

Test Cases | Architecture & Design | Cloud Technologies | Data Mining & Analysis | Reporting/Documentation

***Technology Skills***

***Operating Systems:*** Windows NT/2000, UNIX

***Technologies:*** HDFS, Spark, Storm, Nifi, Kafka, Oozie, Elastic Search, Logstash, Google Cloud Platform (Dataproc, Dataflow, BigQuery, GCS), Java, J2EE (JSP, JDBC, EJB2.0, STRUTS1.1, JNDI), HTML, XML, Java Script, AJAX, SQL, PL/SQL and CSS

***Servers:*** Apache Tomcat, Web Logic

***Databases:*** Oracle 8i, 9i, 11g, DB2, MS Access, Mongo DB, Postgres

***Middleware:*** BEA Web Logic8.1/7.0, Apache Tomcat 4.0/5.0

***GUI:*** Java Server Pages (JSP), Java Swing

***IDE:*** Eclipse3.0/2.3, BEA Web Logic version 8.1, Rational Application Architect, Intellij

***Tools:*** TFS, VSS, CVS, GIT, Toad, Visio, IBM Rational Clear Case & Clear Quest, Advanced Query Tool, Camel, Hive, MicroStrategy, Lucid Chart, WebStorm

***Professional Experience***

Company, City, ST

**Lead Data Engineer, Data Sciences** Mo/Year – Mo/Year

Developed REDEX product, enabling big data platform users to configure data pipelines to read from high-velocity Kafka streams and ingest into big data ecosystem. Ensured strong performance and monitoring capabilities.

* Initiated reading from Kafka in Avro and persist data in Hive ORC. Utilized schema evolution and deduplication configuration. Accurately worded?
* Merged Avro Kafka messages from multiple topics based on uber schema. Wrote to Hive while supporting schema evolution.
* Provided data movement scheduling capability in REDEX.
* Modified Apache Spark open source code to read from non-consecutive offsets/compacted topics
* Delivered required capabilities for a REDEX UI and REDEX yarn application. Designed schema to persist user configurations.
* Developed Nifi-secure cluster and flows for clickstream data from website into Elastic Search.
* Enabled REDEX for PII ingestion (SecureRed) to deliver egress/ingress capabilities into restricted PII storage.
* Created alerts and monitored for REDEX jobs. Captured metadata and published execution metrics onto Kafka topic for user consumption.

*Environment: Spark, Nifi, Storm, Oozie, HDFS, Hive, Elastic Search, Logstash, Kafka*

**Senior Data Engineer, Digital Marketing & Guest Identification** Mo/Year – Mo/Year

Designed and developed technical solutions to deliver relevant information via identification and reporting.

* Interviewed business team members to gather information for identified story points, develop reporting capability roadmap, identify relevant keywords, and manage project scope/schedule.
* Developed detailed technical design for multiple products, such as reports and dashboards.
* Collaborated with Google, Facebook, Pinterest, and Twitter teams to determine appropriate APIs and pull in data to support guest identification.
* Defined architecture. Created design per approved patterns.
* Pulled activity, click, and impression log files from Google storage using DCM/DFA and trafficking APIs. Loaded data into Hadoop, BIG RED private cloud.

***Professional Experience***

Company, City, ST

**Senior Data Engineer, Digital Marketing & Guest Identification** *Continued*

* Mined and analyzed data to decipher logs and identify guests for audience targeting in paid search.
* Associated digital actions of guests on external/internal digital platforms to identify guests and missed opportunities. Relayed sales info to digital marketing team.
* Managed SDLC end-to-end, including analysis, design, development, testing, and prod activities.
* Developed Reviews Landing Page and site map generators.
* Created modifications and enhancements to the external landing page.
* Composed detail design document and developed required code.

*Environment: Intelli J,**GIT, Hadoop, Hive, Oozie, Spark, Solr, Tableau, Domo, Java, Groovy, Spring Boot, Junit, Spock, Oracle 11g, Teradata, Google Cloud Platform (Dataproc, Dataflow, BigQuery, GCS)*

**Senior Data Engineer, Enterprise Business Intelligence & Analytics**  Mo/Year – Mo/Year

Determined key performance indicators (KPIs) for website, applying business analytics/marketing strategies.

* Identified business needs, developed application/capability roadmap, and developed project timeline.
* Developed and implemented processes to scrape inventory and pricing statistics from competitor websites.
* Analyzed gathered requirements and documented technical requirements.
* Collaborated with project manager to plan roadmap execution.
* Defined application architecture and created detailed design per approved patterns.
* Created detail design document and developed the required code.
* Designed and implemented analytics data solution for gathered data from various web analytics platforms.
* Integrated and completed user acceptance testing (UAT) activities.
* Showcased products to partner teams. Served as subject matter expert (SME) on utilizing data and products.

*Environment:**Rational Application Architect****,*** *Java (J2EE), MS Visio, DB2, Oracle 11g,Teradata, Advanced Query Tool, Benthic Software, MicroStrategy, Team Foundation Server, GIT, Adobe Site Catalyst, Hadoop, Tableau*

Company, City, ST

**Senior Developer/Tech Lead** *(Client: Company)* Mo/Year – Mo/Year

Tested, implemented, documented, and supported new and existing web analytics requirements and procedures.

* Led requirements gathering meetings with internal clients.
* Analyzed gathered requirements to create functional specification document.
* Established application architecture and detailed design.
* Reviewed and documented current web analytics data collection.
* Created detailed design document and wrote required Java J2EE code.
* Led development team to standardize existing technical documentation, deliver code to match analytics code base, and implement necessary changes.

*Environment: Eclipse Indigo, Java (J2EE), MS Access, Microsoft Visio, Adobe Site Catalyst, Charles, WebStorm 3.0, Java Script*

*Early Roles*

**Java Developer** *(Client: Company)* **|** Company, City, ST Mo/Year – Mo/Year

**Senior Java Developer** *(Client: Company)* **|** Company, City, ST Mo/Year – Mo/Year

***Education & Certifications***

**Bachelor of Engineering in Computer Sciences** | University, City, Country

**Infosys PR 101 J2EE Foundation Certification**

**Google Certification**, DoubleClick Search Fundamentals

**Oracle Certified Associate (OCA)**, Introduction to Oracle: SQL AND PL/SQL (1Z0-001)