

Hadoop & Kafka Administrator Hadoop & Kafka Administrator Austin, TX Strong, self-motivated, Hadoop-certified Software Developer with primary experience in Big Data, ETL, and NoSQL focused environments. Strives for following good design patterns and maintainable code. Proficient in diverse languages including Java and Python. Experienced with Hadoop, Kafka, Spark, Terraform, Ansible, and Docker. Able to effectively self-manage during independent projects, as well as collaborate as part of a productive team Authorized to work in the US for any employer Work Experience Hadoop & Kafka Administrator Austin, TX March 2017 to Present Administering and maintaining a hybrid cloud environment between a 100+ node on-premise Hadoop cluster, multiple 5-20 node Kafka clusters, and several cloud environments for a vacation rental company. Providing business hour support through company Slack channels with regards to general questions, maintenance notifications, and debugging assistance. Tools involved include Kafka (Confluent Platform), Avro, Hortonworks HDP, Hue, Hive, Qubole, PrestoDB, Spark 1.6+, Elastic ELK stack, AWS EC2, S3, EMR, Docker on Mesos/Marathon, Consul, Vault, Terraform, Jenkins, and Splunk. Staff Consultant Avalon Consulting, LLC - Austin, TX June 2015 to Present Became familiar with creating repeatable builds of software solutions through the combinations of Bash, Ansible, Vagrant, and Docker. Developed multiple demos using Elasticsearch, Logstash, Kibana, and Beats (the ELK stack) for textual and visual analysis on structured datasets. Hadoop EDW Offloader Dallas, TX February 2017 to March 2017 Offloaded a multinational Arts and Crafts retailer's Oracle database to the Cloudera CDH Platform in AWS, and reduced the runtime of an Oracle Pro*C ETL process from 5 hours to within a 30 minute window using a combination of Apache Impala and the Parquet data format. Apache Spark Streaming Developer Austin, TX September 2016 to November 2016 Built a Spark Streaming ETL pipeline for payment transaction documents made available on HDFS and processed through to Kafka and into MemSQL, Hive, and ElasticSearch. Utilized custom Hadoop InputFormats with regex support to handle parsing semi-structured document files into Spark RDD Row events. Graph Database Developer Austin, TX August 2016 to September 2016 Researched differences in Neo4j and OrientDB graph databases while translating relational, tabular data into a format more suitable for a graph. Used both

technologies to query and identify connected components of an online retailer's web traffic metadata. Apache Spark MLlib Developer Austin, TX May 2016 to August 2016 Used Apache Spark MLlib to refine an iterative Random Forest algorithm to categorize issues reported through a help-desk system. Each iteration of development and manual parameter tuning resulted in above 80% of generated model accuracy against the provided dataset. Tasks included manual classification of training data, word stemming, removal of stop words, and feature vectorization.

Education Bachelor's in Software Engineering And Computer Science Rose-Hulman Institute of Technology - Terre Haute, IN August 2011 to June 2015 Skills Apache Kafka (2 years), Hadoop (3 years), Docker (2 years), Java (5 years), Python (4 years), Ansible (2 years), Git (4 years)

Assessments Data Analysis Expert August 2019 Measures a candidate's skill in interpreting and producing graphs, identifying trends, and drawing justifiable conclusions from data. Full results: https://share.indeedassessments.com/share_assignment/w8mgl-ngwwjgw16q Indeed Assessments provides skills tests that are not indicative of a license or certification, or continued development in any professional field.

Name: Jeffery Ortiz

Email: linda36@example.net

Phone: (707)315-4818x2030