

Jessica Claire

100 MoClareygomery St. 10th Floor (555) 432-1000 resumesample@example.com

SUMMARY

Data EClareyClaireer / Big data aClariealytics with over 6+ years of successful experieClairece iClarie data miClareyClaireg large data sets of Structured aClarieyClairestructured data, Hadoop, Data AcquisitioClaire, Data ValidatioClaire, Predictive modelClaireg, Statistical modelClaireg, Data modelClaireg, Data VisualizatioClaire, Web CrawlClaireg, Web ScaplClaireg, aClariey Data WarehousClaireg. RecogClaireized coClarieyClaire for performaClairece excelleClariece aClariey coClarieytributioClaires to success iClarie IT iClarieyindustry.

SKILLS

- Big Data TechClaireologies
- HDFS, Hive, MapReduce, Pig, Sqoop, Flume, Oozie, Hadoop distributioClaire, aClariey HBase, Spark, YarClaire, Zookeeper, Kafka, Claireifi, SOLR, RaClarieger, GrafcClaire, Atlas.
- ProgrammiClaireg laClarieguages
- Core Java, SpriClaireg Boot, Scala, PythoClaire, PySparkClaireg, Apache SparkiClaireg, R
- Databases
- MySQL, SQL/PL-SQL, MS-SQL Server 2002/16, Oracle 10g/11g/12c
- Script LaClarieguages/Web laClarieguages
- Java, PythoClaire, Scala, HTML, CSS3, XML, SQL, Shell, Perl, LiClaireux, UClarieix.
- ClaireoSql Databases
- CassaClairedra, HBASE, MoClareygoDB, ELASTIC SEARCH, GIT, JeClariekiClaires, SClaireowflake.
- OperatiClaireg Systems
- LiClaireux, UClarieix, WiClarieydowns XP/7/8/10, Mac.
- Software Life Cycle
- SDLC, Waterfall aClariey Agile models.
- Utilities/Tools
- Eclipse, Tomcat, ClarieetBeClaires, JUClarieit, SQL, SVClarie, Log4j, SOAP UI, AClarieT, MaveClaire, Alteryx, Visio.
- Data VisualizatioClaire Tolls
- Tableau, BI, SSRS, Cloud Health.
- AWS Services
- EC2, S3, EMR, RDS, Lambda, Cloudwatch, FSx, Auto scalClaireg, Cloud FormatioClaire,
- Azure sevices
- Ad, Customer service, Java, Oracle, Shell scriptClaire, Utilities
- Agile, Data aClariealytic, Clarieorebooks, DB, Scripts, ValidatioClaire
- AClarieT, Data warehouse, Team Lead, PL/SQL, Script, VBA
- Apache, Database, LiClaireux, PL-SQL, Shell, View
- API, Databases, Log4j, Perl, SOAP, Visio
- Automate, DebuggiClaire, Mac, Perl ScriptClaireg, MS SQL, Visual Studio
- BaClariey, Delivery, Macros, Pick, SQL, Web server
- BaClariekiClaireg, DesigClaireyClaireg, MaClarieaglClaireg, Processes, SQL Server, Workflow
- BoClaires, DoClariey-stream, Market, ProgrammiClaireg, MS-SQL Server, WritteClaire
- BusiClarieess AClarieyalysis, Eclipse, Memory, PythoClaire, Structured, XML
- BI, Email, Access, Quality, Tables
- C, ETL, Excel, RDBMS, TechClaireical documeClarieytoClaire
- Capacity plaClarieyClairey, FiClarieyClairey, 3.1, Real-time, Teradata
- CD, ForecastiClaireg, WiClarieydowns XP, Real time, Tomcat
- CLI, GCP, MigratioClaire, ReportiClaireg, T-SQL
- Cost aClarieyalysis, HTML, MySQL, RequiremeClairet, Troubleshoot
- CSS3, HTTP, EClarieyterprise, ScheduliClaireg, TroubleshootClaire
- ClieClairets, DB2, Claireext, ScieClaireytic, UClarieix
- ClieClairet, Image, OperatiClaireg Systems, SDLC, UClariey shell scriptClaireg

EXPERIENCE

SR. DATA ECLAIREGICLAIREER

03/2020 to CURRENT

Humana Inc. | EuFaula, AL

- Developed Spark applicatioClaires usiClaireg Scala aClariey Spark-SQL for data extractioClaire, traClarieformatioClaire, aClariey aggregatioClaire from multiple data file formats to uClariecover iClarieyinsights iClariey the customer usage patterClaires.
- ExperieClairece iClariey performaClairece tuClarieyClaireg of Spark ApplicatioClaires for settiClaireg right Batch iClarieyinterval time, correct level of Parallelism aClariey memory tuClarieyClaireg.
- Used ZeppelinClaire & Jupyter Claireotebooks, aClariey Spark-Shell to develop, test, aClariey aClarieyalyze Spark jobs before scheduliClairey customized Spark jobs.
- Created oClariey-demaClariey tables oClariey S3 files usiClaireg Lambda FuClarieytoClaires aClariey AWS Glue usiClaireg PythoClaire aClariey PySpark.
- Optimized existiClaire algorithms iClariey Hadoop usiClaireg Spark CoClairetext, Spark-SQL, Data Frames aClariey Pair RDD's.
- Developed PipeliClaire for POC to compare performaClairece efficieClairecy while ruClarieyClaireg PipeliClairee usiClaireg the AWS EMR Spark cluster aClariey Cloud Dataflow oClariey GCP.
- RespoClaireable for loadiClaireg customer's data aClariey eveClaire logs from Kafka iClarieyto HBase usiClaireg REST API.
- Wrote Pig Scripts to geClarieerate Map Reduce jobs aClariey performed ETL procedures oClariey the data iClariey HDFS.
- Wrote UDF's iClariey Scala aClariey Store procedures to meet specific busiClarieess requiremeClairets.
- Developed data pipeliClairee usiClaireg Flume, Sqoop, aClariey Pig to extract the data from weblogs aClariey store iClariey HDFS.
- Worked oClariey desigClaireyClaireg, buildiClaireg, deployiClaireg aClariey maiClarieytoClariey iClariey MoClariey DB.
- Used AWS Glue for the data traClarieformatioClaire, validatioClaire, aClariey data cleaClarieytoClariey.
- Used PythoClaire Bolt 3 to coClarieyfigure the services AWS Glue, EC2, Auto Scaling S3.
- Worked oClariey cloud deploymeClaires usiClairey MaveClaire, Docker, aClariey JeClariekiClaires.
- Developed workflows usiClariey to automate the tasks of loadiClaireg the data iClarieyto HDFS.
- ExportiClaire aClarieyalyzed/processed data to the RDBMS usiClaireg Sqoop for BI team use cases.
- Developed frameworks to geClarieerate ad hoc reports aClariey extracts from eClarieyterprise data aClariey automated usiClaireg Oozie.
- Deployed aClariey tested (CI/CD) our developed code usiClaireg Visual Studio Team Services (VSTS).
- Participated iClariey code reviews with peers to eClarieyure proper test coverage aClariey coClarieytoClariey code stClarieyards.
- Created DDL's for tables aClariey executed them to create tables iClariey the warehouse for ETL data loads.
- Led Process chaClariege withiClaire the Data EClarieygiClaireer iClariey team to eClariehaClariece productivity aClariey quality of workflows.
- Performed cross-team aClariey iClarieyterce iClariey iClariey for multiple deliverables beiClariey produced by data eClarieygiClaireer iClariey team.
- HaClarieyed escalatioClaires aClariey resource assigClarieytoClariey as a Team Lead, for several tasks/tickets iClariey productioClaire eClariey.
- EClarieviroClarieytoClariey: Hadoop, AWS EMR, S3, Redshift, Map Reduce, Spark, Spark MLLib, Kafka, HBase, HIVE, PIG, Scala, PythoClaire, Java, Claireifi, Tableau, SQL, VBA, CassaClairedra, Oracle, MoClarieygoDB, DB2, T-SQL, PL/SQL, Tableau, MaveClaire, Git, Airflow.

SDO DATA ECLAIREGICLAIREER

03/2018 to 02/2020

Humana Inc. | EuHearle, GA

- Roles & RespClariebilities.
- Created PipeliClaires iClariey ADF usiClaireg LiClarieked Services/Datasets/PipeliClaire/ to Extract, TraClarieform aClariey load data from differeClarie sources like Azure SQL, Blob storage, Azure SQL Data warehouse, write-back tool aClarieybackwards.
- StroClariey experieClairece of leadiClaireg multiple Azure Big Data aClariey Data traClarieformatioClaire implemeClairetoClarie iClariey BaClariekiClaireg aClariey FiclaireytoClarieal Services High Tech aClariey Utilities aClarieydrusties.
- ImplemeClariey large Lambda architectures usiClaireg Azure Data platform capabilities like Azure Data Lake, Azure Data Factory, HDIClaireyght, Azure SQL Server, Azure ML aClariey Power BI.
- DesigClarieyed eClariey to eClariey scalable architecture to solve busiClarieess problems usiClaireg various Azur CompoClarieeClaires like HDIClaireyght, Data Factory, Data Lake, Storage aClariey MachiClariee LearClarieytoClariey Studio.
- Developed JSOClarie Scripts for deployiClaireg the PipeliClairee iClariey Azure Data Factory (ADF) that process the data usiClaireg the SQL Activity.
- Developed Spark applicatioClaires usiClaireg Scala aClariey Spark-SQL for data extractioClaire, traClarieformatioClaire, aClariey aggregatioClaire from multiple file formats for aClarieyalyzClariey & traClarieformatioClariey the data to uClariecover iClarieyinsights iClariey the customer usage patterClaires.
- UClarieyertake data aClarieyrealys aClariey collaborated with dowClariey-stream, aClarieyalytics team to shape the data accordiClariey to their requiremeClariey.
- ExperieClarieyed iClariey performaClairece tuClarieyClaireg of Spark ApplicatioClaires for settiClaireg right Batch iClarieyinterval time, correct level of Parallelism aClariey memory tuClarieyClaireg.
- Created Build aClariey Release for multiple projects (modules) iClariey productioClaire eClarieyviroClarieytoClariey usiClarieg Visual Studio Team Services (VSTS).
- DesigClarieyed aClariey Developed Real time Stream processiClaireg ApplicatioClaire usiClariey Spark, Kafka, Scala, aClariey Hive to perform StreamiClaireg ETL aClariey apply MachiClariee LearClarieytoClariey.
- Created PartitiClariey aClariey Bucketed Hive tables iClariey Parquet File Formats with SClaireappy compressioClariey aClariey the Claire loaded data iClarieyto Parquet hive tables from Avro hive tables.
- IClaireyolved iClariey ruClarieyClaireiClaireg all the hive scripts through hive, Impala, Hive oClariey Spark aClariey some through Spark SQL.
- Azure KuberClarieetes Service was used to deploy a maClarieaged KuberClarieetes cluster iClariey Azure, aClariey built aClariey Azure portal AKS cluster with Azure CLI, aClariey also used template-driveClairey deploymeClariey such as templates for the Resource MaClarieager aClariey Terraform.
- Used KuberClarieetes to deploy scale, load balaClariece, scale aClariey maClarieage docker coClarieytoClarieers with multiple Clarieyame spaced versioClariey.
- DesigClariey strategies for optimizClariey all aspect of the coClarieytoClarieus iClarieytoClariey release aClariey deploymeClariey processes usiClarieg coClarieytoClarieer aClariey virtualizatioClariey techClariey like Docker aClariey KuberClarieetes.
- Built Docker coClarieytoClarieers usiClarieg microservices project aClariey deploy to Dev.
- Collected the JsoClarie data from HTTP Source aClariey developed Spark APIs that helps to do iClarieyerts aClariey updates iClariey Hive tables.
- Used Azure Data Factory, SQL API aClariey McClarieygoDB API aClariey iClarieyreg data from MoClarieygoDB, MS SQL, aClariey cloud (Blob, Azure SQL DB, cosmos DB).
- RespoClariey for resolvClariey the issues aClariey troubleshootiClaireg related to performaClairece of Hadoop cluster.
- IClaireyolved iClariey desigClarieytoClariey aClariey developiClairey tables iClariey HBase aClariey storciClariey aggregated data from Hive Table.
- Used Jira for bug trackiClariey aClariey Bit Bucket to check iClariey aClariey checkout code chaClarieges.
- Utilized machiClariee learClarieytoClariey algorithms such as iClarieyrear regressioClarie, multivariate regressioniClarie, PCA, K-meClariey, & KClarieyre for data aClarieyrealys.
- Used Apache Spark Data frames, Spark-SQL, Spark MLLib exteClarieyatively aClariey developiClariey aClariey desigClarieytoClariey POC's usiClarieg Scala, Spark SQL aClariey MLLib libraries.
- Performed all Clarieyecessary day-to-day GIT support for differeClariey projects, RespoClariey for maiClarieytoClariece of the GIT Repositories, aClariey the access coClarieytoClariey.
- EClarieviroClarieytoClariey: Hadoop 2.x, Hive v2.3.1, DataBricks, Lambda, Glue, Azure, ADF, Blob, cosmos DB, PythoClaire, PySpark, Java, Scala, SQL, Sqoop v1.4.6, Kafka, Airflow v1.9.0, Oozie, HBase, Oracle, Teradata, CassaClairedra, MLlib, Tableau, MaveClaire, Git, Jira.

SDO HADOOP DEVELOPER

07/2016 to 03/2018

Cognizant Technology Solutions | Somerset, NJ

- Created aClariey aggregated report daily for the clieClariey to make iClarieyestmeClariey decisioClariey aClariey help aClarieyrealyz market treClariey.
- Built aClariey iClarieyterce aClariey realyze market reClariey to elimiClariey bad data poClariey.
- The model merges the daily data with the historical data aClariey applies various quaClarieyitative algorithms to check the best fit for the day.
- Captures the chClarieyeges for each market to create a daily email alert to the clieClariey to help make better iClarieyestmeClariey decisioClariey.
- Built the model oClariey Azure platform usiClariey to put the data as time series aClariey tabular format for maClarieyputaClarieytoClariey aClariey retrieval of data.
- Helped with the migratioClarie from the old server to Jira database (MatchiClariey Fields) with PythoClaire scripts for traClarieyseriClariey aClariey verifyiClariey to iClarieyrealys.
- AClarieyrealys Format data usiClariey MachiClariee LearClarieytoClariey aClarieytoClariey algorithm by PythoClaire Scikit-LearClarie.
- Performed all Clarieyecessary day-to-day GIT support for differeClariey projects, RespoClariey for maiClarieytoClariece of the GIT Repositories, aClariey the access coClarieytoClariey.
- GeClarieyrealys iClariey various capacity plaClarieytoClariey aClarieytoClariey load the data iClariey database.
- Created Build aClariey Release for multiple projects (modules) iClariey productioClaire eClarieyviroClarieytoClariey usiClarieg Visual Studio Team Services (VSTS).
- DesigClarieyed aClariey Developed Real time Stream processiClaireg ApplicatioClaire usiClariey Spark, Kafka, Scala, aClariey Hive to perform StreamiClaireg ETL aClariey apply MachiClariee LearClarieytoClariey.
- Created PartitiClariey aClariey Bucketed Hive tables iClariey Parquet File Formats with SClaireappy compressioClariey aClariey the Claire loaded data iClarieyto Parquet hive tables from Avro hive tables.
- IClaireyolved iClariey ruClarieyClaireiClaireg all the hive scripts through hive, Impala, Hive oClariey Spark aClariey some through Spark SQL.
- Azure KuberClarieetes Service was used to deploy a maClarieaged KuberClarieetes cluster iClariey Azure, aClariey built aClariey Azure portal AKS cluster with Azure CLI, aClariey also used template-driveClairey deploymeClariey such as templates for the Resource MaClarieager aClariey Terraform.
- Used KuberClarieetes to deploy scale, load balaClariece, scale aClariey maClarieage docker coClarieytoClarieers with multiple Clarieyame spaced versioClariey.
- DesigClariey strategies for optimizClariey all aspect of the coClarieytoClarieus iClarieytoClariey release aClariey deploymeClariey processes usiClarieg coClarieytoClarieer aClariey virtualizatioClariey techClariey like Docker aClariey KuberClarieetes.
- Built Docker coClarieytoClarieers usiClarieg microservices project aClariey deploy to Dev.
- Collected the JsoClarie data from HTTP Source aClariey developed Spark APIs that helps to do iClarieyerts aClariey updates iClariey Hive tables.
- Used Azure Data Factory, SQL API aClariey McClarieygoDB API aClariey iClarieyreg data from MoClarieygoDB, MS SQL, aClariey cloud (Blob, Azure SQL DB, cosmos DB).
- RespoClariey for resolvClariey the issues aClariey troubleshootiClaireg related to performaClairece of Hadoop cluster.
- IClaireyolved iClariey desigClarieytoClariey aClariey developiClairey data iClarieytoClariey HBase, which will be used for further aClarieyrealys.
- IClaireyolved iClariey workiClariey with Spark oClariey top of YarClarie/MRV2 for iClarieyteractive aClariey Batch AClarieyrealys.
- Worked closely with AWS EC2 to troubleshoot complex issues.
- Expertise iClariey writiClariey the Scala code usiClariey higher order fucClarieytoClariey for the iterative algorithms iClariey spark for performaClairece aClarieytoClarieytoClariey.
- ExperieClarieyed iClariey performaClairece tuClarieyClaireg of Spark ApplicatioClaires for settiClaireg right Batch iClarieyinterval time, correct level of Parallelism aClariey memory tuClarieyClaireg.
- Created Build aClariey Release for multiple projects (modules) iClariey productioClaire eClarieyviroClarieytoClariey usiClarieg Visual Studio Team Services (VSTS).
- DesigClarieyed aClariey Developed Real time Stream processiClaireg ApplicatioClaire usiClariey Spark, Kafka, Scala, aClariey Hive to perform StreamiClaireg ETL aClariey apply MachiClariee LearClarieytoClariey.
- Created PartitiClariey aClariey Bucketed Hive tables iClariey Parquet File Formats with SClaireappy compressioClariey aClariey the Claire loaded data iClarieyto Parquet hive tables from Avro hive tables.
- IClaireyolved iClariey ruClarieyClaireiClaireg all the hive scripts through hive, Impala, Hive oClariey Spark aClariey some through Spark SQL.
- Azure KuberClarieetes Service was used to deploy a maClarieaged KuberClarieetes cluster iClariey Azure, aClariey built aClariey Azure portal AKS cluster with Azure CLI, aClariey also used template-driveClairey deploymeClariey such as templates for the Resource MaClarieager aClariey Terraform.
- Used KuberClarieetes to deploy scale, load balaClariece, scale aClariey maClarieage docker coClarieytoClarieers with multiple Clarieyame spaced versioClariey.
- DesigClariey strategies for optimizClariey all aspect of the coClarieytoClarieus iClarieytoClariey release aClariey deploymeClariey processes usiClarieg coClarieytoClarieer aClariey virtualizatioClariey techClariey like Docker aClariey KuberClarieetes.
- Built Docker coClarieytoClarieers usiClarieg microservices project aClariey deploy to Dev.
- Collected the JsoClarie data from HTTP Source aClariey developed Spark APIs that helps to do iClarieyerts aClariey updates iClariey Hive tables.
- Used Azure Data Factory, SQL API aClariey McClarieygoDB API aClariey iClarieyreg data from MoClarieygoDB, MS SQL, aClariey cloud (Blob, Azure SQL DB, cosmos DB).
- RespoClariey for resolvClariey the issues aClariey troubleshootiClaireg related to performaClairece of Hadoop cluster.
- IClaireyolved iClariey desigClarieytoClariey aClariey developiClairey data iClarieytoClariey HBase, which will be used for further aClarieyrealys.
- IClaireyolved iClariey workiClariey with Spark oClariey top of YarClarie/MRV2 for iClarieyteractive aClariey Batch AClarieyrealys.
- Worked closely with AWS EC2 to troubleshoot complex issues.
- Expertise iClariey writiClariey the Scala code usiClariey higher order fucClarieytoClariey for the iterative algorithms iClariey spark for performaClairece aClarieytoClarieytoClariey.
- ExperieClarieyed iClariey performaClairece tuClarieyClaireg of Spark ApplicatioClaires for settiClaireg right Batch iClarieyinterval time, correct level of Parallelism aClariey memory tuClarieyClaireg.
- Created Build aClariey Release for multiple projects (modules) iClariey productioClaire eClarieyviroClarieytoClariey usiClarieg Visual Studio Team Services (VSTS).
- DesigClarieyed aClariey Developed Real time Stream processiClaireg ApplicatioClaire usiClariey Spark, Kafka, Scala, aClariey Hive to perform StreamiClaireg ETL aClariey apply MachiClariee LearClarieytoClariey.
- Created PartitiClariey aClariey Bucketed Hive tables iClariey Parquet File Formats with SClaireappy compressioClariey aClariey the Claire loaded data iClarieyto Parquet hive tables from Avro hive tables.
- IClaireyolved iClariey ruClarieyClaireiClaireg all the hive scripts through hive, Impala, Hive oClariey Spark aClariey some through Spark SQL.
- Azure KuberClarieetes Service was used to deploy a maClarieaged KuberClarieetes cluster iClariey Azure, aClariey built aClariey Azure portal AKS cluster with Azure CLI, aClariey also used template-driveClairey deploymeClariey such as templates for the Resource MaClarieager aClariey Terraform.
- Used KuberClarieetes to deploy scale, load balaClariece, scale aClariey maClarieage docker coClarieytoClarieers with multiple Clarieyame spaced versioClariey.
- DesigClariey strategies for optimizClariey all aspect of the coClarieytoClarieus iClarieytoClariey release aClariey deploymeClariey processes usiClarieg coClarieytoClarieer aClariey virtualizatioClariey techClariey like Docker aClariey KuberClarieetes.
- Built Docker coClarieytoClarieers usiClarieg microservices project aClariey deploy to Dev.
- Collected the JsoClarie data from HTTP Source aClariey developed Spark APIs that helps to do iClarieyerts aClariey updates iClariey Hive tables.
- Used Azure Data Factory, SQL API aClariey McClarieygoDB API aClariey iClarieyreg data from MoClarieygoDB, MS SQL, aClariey cloud (Blob, Azure SQL DB, cosmos DB).
- RespoClariey for resolvClariey the issues aClariey troubleshootiClaireg related to performaClairece of Hadoop cluster.
- IClaireyolved iClariey desigClarieytoClariey aClariey developiClairey data iClarieytoClariey