****Role: Data Scientist / Big-Data Analyst

Kiran Achanta | Mob: 8095512333|Email: [akiran.tech123@gmail.com](mailto:akiran.tech123@gmail.com) |[www.linkedin.com/in/KiranA333](http://www.linkedin.com/in/KiranA333)

Summary

Having 7 years of extensive experience in the field of Data Analytics, experience in working with geographically divided global teams in Agile methodology. My core competency is in Data Analytics/Data Mining, Database Programming, Data Warehousing (ETL+BI Visualization), Statistics and Machine Learning.

Skill Stack

|  |  |
| --- | --- |
| **BI/Visualization** | Tableau, Seaborn, Advance Excel |
| **Database Query Language** | SQL Stored Proc and Stored Functions |
| **Programming Language** | Python 3x, R |
| **Machine Learning** | Supervised and Unsupervised models. |
| **Scripting Language** | Linux Shell Scripting |
| **RDBMS** | Oracle 10g Certified, MySQL, SQL Server 2012 |
| **ETL Tool** | DataStage 8.5 Certified |
| **Hadoop Eco System** | HiveQL Certified, PySpark, HDFS, CDH5 |
| **Cloud Environment** | AWS Certified, OpenStack, |
| **Agile Tools** | Jira, Git, Bitbucket, Confluence |
| **Domain Experience** | BFSI, Retail & Airlines |
| **Methodology** | Agile & Waterfall |

Certifications

* Oracle Database SQL Certified Expert.
* Certified on Hive from IBM-Big Data University.
* Certified ETL Developer - IBM Web Sphere Certified for Data stage V8.5.
* AWS Certified Solutions Architect - Associate
* Course Completion from Udemy -Statistics for Data Science and Business Analysis
* Course Completion from Udemy - Python for Data Science and Machine Learning Bootcamp
* Certification - Information Security Awareness (Amadeus Internal).

Professional Profile

* Cognate experience in extracting, interpreting and analyzing data to find the hidden insights to drive the business solutions.
* Experience comprises of Data Extraction, Cleaning, Preparation, Validation, and Analysis.
* Experience in performing in-depth statistical analysis on the business-related data using Python, Excel and Tableau.
* Having significant experience in RDBMS concepts. Writing, debugging complex SQL queries and fine tuning long running queries.
* Experience in writing Python code for data wrangling/ data munging, web scraping, transforming, visualizing and automating.
* Extensively used Python (PyData Libraries) - Numpy, Pandas, Matplotlib, Seaborn, sklearn, MySQL, SYS, OS, CSV, JSON etc. Modules.
* Experience in working with Machine Learning models for Regression & Classification problems using – (Linear Regression, Logistic Regression, Decision Trees, Random Forest & KNN Algorithms)
* Experience in writing queries using HiveQL to query against HDFS.
* Hands-on experience in Linux scripting to automate the manual tasks and schedule to run on daily/weekly basis.
* Experience in working with ETL Tool – Data stage 8.5 to extract data from various source systems, cleansing, transforming and loading into staging area and DWH.
* Have worked on excel advanced functions, pivot tables and charts for visual representation.
* Experience in working with applications hosted on cloud environment- AWS and Openstack.
* Hands on experience in working with AWS core services like VPC, Sub-Nets, EC2, EBS, S3, Redshift and RDS.

Achievements

* Spot Award – Received Extra Mile award for automation of reports in Santander Project.
* Received Great Team Member Card (Best performer of the team) for the project Target Corp.
* Received solid performer feedbacks from various clients.

Work Experience

* Working as a **Technical Member Staff** in Amadeus, Bangalore (Dec/2017- Till Date)
* Worked as **Associate IT Consultant** in ITC InfoTech, Bangalore (Oct/2013- Nov/2017)
* Worked as **Consultant** in Capgemini, Bangalore (Sep/2011-June/2013)

Key Projects Executed

Amadeus IT Group: is a major IT Provider for the global [travel](https://en.wikipedia.org/wiki/Travel) and [tourism](https://en.wikipedia.org/wiki/Tourism) industry.

Role: Analyst

Duration: Dec’ 17 – Till date.

Environment:Tableau Desktop, Python 3.6, Excel, SQL, Linux, Machine Learning.

Description: The aim of the project is to analyze Customer Experience Scorecard, Frequent Flyer Loyalty program to identify and engage customers with offers and predicting Air Line Delays.

* Responsible for interaction with business stake holders, gathering requirements and managing the delivery.
* Conduct data quality assessment and cleansing activities.
* Worked on Data Preprocessing activities like handling missing data, detecting outliers, feature scaling and feature selection techniques using python Libraries.
* Applying machine learning models like Linear Regression, Logistic Regression, Decision Trees, Random Forest to do prediction on Flight delays and analyze customer express.
* Fine tune the model performance by checking various regression and classification metrics.
* Produce comprehensive, precise and meaningful metrics, reports, and analyses.
* Created Rich dashboards using Tableau Dashboard and prepared user stories to create compelling dashboards to deliver actionable insights.
* Generate KPI for customer satisfaction survey results.

Santander UK Santander UK plc is a [British](https://en.wikipedia.org/wiki/United_Kingdom) bank, wholly owned by the [Spanish](https://en.wikipedia.org/wiki/Spain) [Santander Group](https://en.wikipedia.org/wiki/Santander_Group).

Role: Data Analyst

Duration: Feb’ 15 – Oct’17.

Environment:HiveQL, SQL, Tableau, Excel, HDFS, Oracle 11g, Linux, Python Pandas Framework, AWS.

Description: The existing DWH system was facing problem to store and analyze large structured and semi-structured historical data using RDBMS. The solution is based on the open source Big Data Hadoop Framework. The data stored in Hadoop file system and processed using Hive and python pandas frame work and visualized in Excel and Tableau Desktop.

* Involved in several client meetings to understand the business requirements.
* Developed and build HIVE tables for structured data using Hive QL
* Created Hive table schemas, written hive queries as per the business rules and generating CSV files.
* Developed the batch jobs and automated hive queries by writing Linux scripts to generate CSV files.
* Extensively used python pandas to cleanse, process semi structured data.
* Created dashboards for various metrics, KPIs to replace manual reports and achieved an increase in time efficiency.
* Have performed ad-hoc queries, weekly, monthly reports like customer reports, inventory reports, etc.
* Created reports, dash boards, user stories using Tableau and used excel for visualization.

NCB Bank: The National Commercial Bank (NCB) also known as AlAhli Bank is the first Saudi Arabian bank, the second largest bank by asset in the Arab World.

Role: DWH Developer.

Duration: Oct’ 13 – Jan’ 14

Environment: Data stage 8.5, Oracle 11g, Sql Server 2012, Linux, Cognos.

Description: NCB Data warehouse (DWH) is used for reporting and analysis. There are over 20 different source systems (BANCS-14, AFS, CCR, Cortex etc.) for which the data extracts happen at NCB. The source systems contain databases as Oracle, SQL Server and flat files.

The data extracted from these systems are transformed and loaded into the USA DB and DWH DB for reporting & Analytics.

* Involved in status meetings, and interacted with the Business Analyst to get the business rules and participated in the review of Technical, Business Transformation Requirements Document
* Developed complex extraction, transformation, and load (ETL) processes for data conversions and construction of data warehouse using IBM Web sphere Data stage8.5
* Extracted data from various source systems like Flat files, MS Excel, Sequential files, Oracle DB and SQL server DB.
* Extensively worked with various Parallel stages like Lookup, Peek, Transformer, Merge, Aggregator, Row generator, Column generator, Join, Filter.
* Created error files and log tables containing data with discrepancies to analyze and re-process the data.
* Worked on ETL enhancements and bug fixes as required.
* Supporting SIT, UAT users and helping release management team for deployment issues in test & production environments.

Target Corp:Target is a leading corporation of retail stores in the United States and Canada.

Role: ETL Developer

Duration: Sep’ 11 – Jun’ 13

Environment:Data stage 7.5, IBM DB2, Oracle 10g, Micro Strategy, Linux Scripting.

Description: This project involved in extracting data from different sources and loading into a Data Marts. The major job involved in cleansing the data and transforming the data to the staging area then loading into Data Marts. The Data Marts is an integrated Data Mine that provides feed for extensive reporting.

* Participated in Requirement gathering and preparation of Technical Design Document.
* Developed Data stage jobs to extract data from various source systems, transform and load the data to Landing DB and Core tables based on the TSD document.
* Written complex SQL queries to extract the data from Source Systems like Oracle, SQL Server.
* Developed ETL jobs for Delta Load and Incremental using methodologies SCD Type-1 and SCD Type-2
* Developed and executed the Unit Test Cases.
* Tuned Data stage jobs and SQL queries for better performance.
* Prepare the Linux scripts for processing the source files and automating the batch jobs.
* Worked closely with SIT, UAT teams to support the testing till the release has been signed off.

Education B-Tech in Computer Science Engineering from JNT University, Hyderabad.