

# Faisal Akbar

Queens, New York

(+1) xxx-xxx-xxxx | faisal.akbar.ny@gmail.com | dsfaisal.com | Github | LinkedIn | Tableau Public

## Summary

---

- Over 8 years of experiences as a **Data Scientist/ Data Analyst/ Senior Tableau Developer/ R Shiny Developer**.
- Experienced with business intelligence (BI) tools such as **Tableau, Microsoft Power BI, QlikView/ QlikSense, MicroStrategy** and **spreadsheets**.
- Experienced with Data Mining, Data Wrangling, Data Quality Assessment, Data Analysis, Machine Learning, Predictive Model.
- Strong analytical skills with ability to collect, organize, analyze, interpret, disseminate significant amounts of information with attention of detail and accuracy.
- Strong UX/UI design experience for desktop as well as small form-factor mobile devices such as tablets and mobile phones.
- Developed and automated Execution metrics, KPI, Mobile responsive and reporting dashboard.
- Collaborate with team and identify all the business requirements and take actions on priority.
- Experienced in employing **Tableau, Power BI, MATLAB, SSIS, R** and **Python** for data wrangling, data visualization, risk analysis and predictive analytics.
- Experienced on Big Data related technologies including but not limited to Hadoop, Hive, Pig, Sqoop, Cassandra, Oozie, Flume, Map-Reduce and Cloudera Manager for Data manipulation and data analytics.
- Hands on experienced with **Machine Learning** and deep learning algorithms.
- Experienced with other data repository including, but not limited to Oracle, XML, DB2, Teradata15/14, Netezza, SQL Developer.
- Extensively used Agile methodology as the Organization Standard to implement the data Models.
- Experienced with machine learning tools and libraries such as **Python** (Pandas, Numpy, SciPy, StatsModels, Scikit-learn, matplotlib, seaborn, plotly, nltk etc), **R** (Caret, dplyr, ggplot2, tidyverse, lubridate, DT, plotly, shiny, knitr, Rmarkdown etc).
- Hands-on experienced with NLP, mining of structured, semi-structured, and unstructured data.
- Experienced on Digital Ocean cloud hosting.
- Developed R shiny projects (<https://dsfaisal.com/shiny-apps/>) and deployed in Digital Ocean cloud.

## Work Experience

---

### Bank of America

New York, NY

#### SENIOR DATA ANALYTICS CONSULTANT WITH TABLEAU

Sep. 2020 - Present

- Created Dashboard using Tableau to communicate complex ideas to executive committee.
- Developed and automated daily, weekly tradable fee prediction monitoring dashboard with up/ down mover.
- Developed locate request and response dashboard to track client request, request by channel, request by region, requested quantity, approved quantity, requested market value, approved market value.
- Designed and developed market sentiment analysis dashboard.
- Applying cutting edge classification, regression, and clustering techniques, including deep learning, to handle high dimensional feature and outcome distributions.
- Building predictive models and liquidity intelligence tools based on numerous data sources.
- Wrote python script to check data quality, CRUD operation that helped speed up daily reporting.
- Performing ad-hoc analysis and present results in a clear manner.

### Healthfirst

New York, NY

#### DATA ANALYTICS CONSULTANT WITH TABLEAU

Dec. 2019 - Sep. 2020

- Worked with Appeals and Grievances management systems and processes like CCMS and Hyland Onbase.
- Developed aesthetically attractive visualization and following industry best practices to engage a large audience, and drive better business decision.
- Designed, developed, maintained, optimized, and automated the Appeals and Grievance tableau dashboard to identify the number of appeals and grievances case received and closed from both CCMS and Hyland, categorized by expedited Appeals, pre-service Appeals, post Service Appeals, standard grievances for different line of business, and reporting how many cases closed timely or untimely for each day, week and month.
- Developed, tuned and automated daily and weekly productivity metrics dashboard with guided features and performance quadrant metrics (high volume low quality, high volume high quality, low volume low quality, low volume high quality) for multiple teams' employee to track each individual and team lead performance based on cases closure rate, number of cases closed, productivity weight, worked hour, and non- productivity time.
- Developed and optimized modern UI design responsive dashboard for mobile and tablet devices to speed providing executives and leadership with condensed and timely analytics on their devices.
- Developed and automated Audit Risk Compliance Committee (ARCC) deck dashboard to track appeals and grievances case closed inventory, timeliness, processing accuracy and percentage metrics for different line of business.
- Created Urgent Care Dashboard and Clinical Partnership dashboard to find the number of visits occurring, top diagnosis, top family health center, and how old are the members based on start and end date.
- Created Health and Recovery Plan (HARP), Performance improvement project (PIP) quarterly and yearly dashboard to find the number of claims, members, members month and utilization per thousands.
- Healthcare Effectiveness Data and Information Set (HEDIS) and Quality Assurance Reporting Requirements (QARR) dashboard for different HEDIS measures.

- Designed and developed dashboard for West Midtown Medical Group to track number of members, claims, and paid amount with HEDIS measure, Substance used drug (SUD) for inpatient and outpatients for different categories like detox, rehab, behavioral health.
- Created and maintained Medicaid child carve dashboard number of members that received a Psychotropic medication, number of Prescribers written Psychotropic medication, number of Pharmacies dispensed Psychotropic medication, number of Psychotropic fills dispensed.
- Developed, designed and automated Talk time metrics KPI dashboard to track total number of calls, hours talked, hours worked, calls/ hour, survey/ call volume, total points, average points per call, completion rate, call disposition, completed vs failed call for each coordinator.
- Created Life Improvement Plan (HMO SNP) members dashboard based on demographic like gender, language, age groups.
- Created Health Outcomes Survey (HOS) dashboard to gather valid, reliable, and clinically meaningful health status from the Medicare Advantage (MA) program to use in quality improvement activities, performance, program oversight, and to improve health.
- Developed complex Audit Plan record entry form using excel vba and modern UI technique to improve data integrity and consistency and to minimize data error in Audit Plan dashboard for different responsible party including but not limited to compliance audit, internal audit, regulatory affairs.
- Extensively used tableau Level-of-Detail (LOD) calculations, nested LOD, context filters, table calculations, parameter actions, set actions, date functions, cross database join, wildcard union, modern UI design and similar.
- Produced insights and analyses using a combination of SQL, aginity workbench for redshift, python, R, alteryx, tableau prep and Tableau among other similar tools.
- Worked with building and navigating multiple data sources, leveraging clinical and administrative information for data integration, mapping, and analytics.
- Experience communicating technical results to technical and non-technical audience across all levels of an organization.
- Respond to ad-hoc requests from appeals and grievances, behavioral health, care management groups.

## CVS Health (e- Zest Solutions)

Irving, TX

DATA SCIENTIST & SENIOR TABLEAU DEVELOPER

Jun. 2018 - Dec. 2019

- Worked with Anthem pharmacy benefit management claims data-sets of varying degrees of size and complexity to identify and analyze the match mismatch scenarios for paid and reject claims and finding the root cause using Hadoop, Spark, HIVE, Machine learning and statistical modelling, and report analysis output using Tableau.
- Designed, developed, maintained, optimized, and deployed workflow tableau dashboard in tableau server, integrate the dashboard with web application via tableau JavaScript API and test in development and production environment and ensure tableau row level security.
- Developed smart triage tableau drill down dashboard for claims comparison and validation based on all the scenarios, Functionalities and reject codes
- Developed sophisticated tableau KPI dashboard and cost benefit analysis dashboard for claims execution, mismatch claims, reject and paid claims, mismatch by top features, functionality, top NDC, top drug type, copay comparison by state, city and pharmacy level.
- Automated and optimized all the daily, weekly, monthly reports including claim execution, mismatch, scenarios, reject codes, ESI claims versus
- Created CVS claims, mismatch by field, user activity and executive reports using tableau and saved more than 120 hours of work per week.
- Developed claims surveillance tableau dashboard to predict and flag potential fraudulent claims and anomalies detection, outliers, and trend changes.
- Developed completely responsive mobile application tableau KPI dashboard for executive.
- Extensively used tableau table calculations, Level-of-Detail (LOD) calculation, set actions, parameter actions, context filters, conditional filters, complex calculations, navigation, date functions, layout container and more.
- Analyze and select several potential appropriate modeling approaches for a given analytic problem (machine learning methods such as Ensemble and classification models, Decision Trees, Random Forest, AdaBoost, Logistic regression, Clustering, Principal Component Analysis (PCA), operations research; statistical modeling such as multivariate techniques).
- Data discovery and organization on SQL/ Teradata SQL/ Big data stack, cleansing and data prep to support development of new data science approaches and methodologies (e.g. neural nets, CART, Bayesian methods, etc.) to improve operations and business outcomes.
- Developed Machine learning and statistical analysis dashboard using RShiny for defect prediction, predictive analysis and claims to plan prediction.
- Designed machine learning projects to address specific business problems determined by consultation with business partners.
- Piping and processing massive data-streams in distributed computing environments such as Hadoop to facilitate analysis and implements batch and real-time model scoring to drive actions.
- Developed sophisticated visualization of analysis output for business users. Publish results and address constraints/limitations with business partners.
- Used R (RStudio), Python (PySpark), alteryx for data cleaning, data visualization, finding anomalies, statistical modeling and predictive analytics.
- Consulting with key internal and external stakeholders to determine how best to leverage machine learning and advanced analytic methods to support business objectives across CVS Health.
- Efficiently implement the models in a variety of modeling tools, achieving highly accurate models.
- Understands the underlying statistical concepts and computational approaches that enable efficient execution of models and may be able to design and implement modifications and enhancements to the computations.
- Developed sound analytic plans based on available data sources, business partner needs, and required timelines.

- Apply innovative approaches to understand and predict what will happen across the business.
- Extensively using agile methodology and JIRA as tracking tool.

## Capital One

New York, NY

### DATA SCIENTIST & TABLEAU DEVELOPER

Nov. 2016 - Jun. 2018

- Developed tableau dashboard for loan risk analysis, credit risk analysis, consumer spending and customer segmentation to improve targeted marketing strategies.
- Created Complex Tableau calculations, table calculations, Level-of-Detail (LOD), context filters, actions, time period functions, parameters based on requirements.
- Created visualizations including pie and bar charts, geographic maps, heat maps, treemaps, scatter plots, bubble chart, histograms, bullet charts, highlight tables, box charts, butterfly charts, and more.
- Conducted analysis on billions of customer transactions to identify sales and promotions scopes using Hadoop, Spark, HIVE and AWS suite and Tableau.
- Developed and designed interactive, functional and appealing web applications dashboard using R-Shiny for statistical, machine learning, and reporting.
- Wrote SQL and HQL to clean and investigate large, messy, data sets of numerical and textual data.
- Continuously built machine learning models from development through testing and validation to their 30+ million customers in production. ML used includes but not limited to Regression Analysis, Clustering, Boosting, XG Boost, SVM, and Random Forest, Bayesian HMM, Classification, Principal Component Analysis (PCA), MCA, Feed Forward Neural Network (FFNN).
- Investigated new technologies on the future of digital banking.
- Used R (RStudio) and Python (PySpark) programming for data cleaning, data visualization, risk analysis and predictive analytics.
- Worked with financial analysts to integrate algorithms and data into Return Path solutions.
- Extensively used Agile methodology (Scrum/Kanban).
- Built data pipelines for reporting, alerting, and data mining.
- Experienced with table design and data management using HDFS, Hive, Sqoop, MySQL, and Kafka.
- Mined geospatial data (GIS). Used R to create visualizations of different measures. Libraries used: ggmap, rgdal, gdal, rgeos, etc.
- Done NLP and text mining experience to create statistical predictive models with unstructured text data.
- Extracted data from web sources using we scrapping and web crawling in python and created text corpora.
- Worked on preprocessing the text data using pandas and nltk. Like, removing stop words, stemming, lemmatization etc.
- Used Convolution Neural Networks and toolkits including CAFFE, cuda-convnet, and others.
- Created predictive models with the text data using nltk and comparing these classification models (Naive Bayes, SVM, max entropy etc). Application involved: Sentimental analytics, sentence meaning extraction and topic modeling.
- Performed scoring and financial forecasting for collection priorities and predict fraud.
- Worked directly with executive stakeholders to define requirements of scoring models.
- Developed a model for predicting a debtor setting up a repayment rehabilitation program for student loan debt.
- Developed a model for predicting repayment of debt owed to small and medium enterprise (SME) businesses.
- Used JIRA as tracking tool.

## JP Morgan & Chase

Columbus, OH

### DATA SCIENTIST & TABLEAU DEVELOPER

Dec. 2015 - Oct. 2016

- Developed robust data and automated reporting tools for senior management (i.e. reduce manual reporting time by Tableau, Microsoft Power BI).
- Created complex tableau dashboard to analyze customer data in order to provide better customer banking experience.
- Developed dashboard to analyze customer data to track customer journey and preferences which helped marketing operations team in deciding website design, promotional materials and launching new products like Chase mobile app.
- Developed Logical and Physical Data model and organized data as per the business requirements using Sybase Power Designer, Erwin, ER Studio in both OLTP and OLAP applications
- Worked on different type of modules such as requests, boto, flake8, flask, mock and nose
- Done data cleaning and development of predictive model using RStudio and Python IDE (Anaconda)
- Utilized Hadoop based storage system, Cassandra, Hive, Pig and Sqoop to design business intelligence applications
- At the same time extracted and manipulated data from SQL and Oracle based transactional data base system
- Wear the hat of a data engineer and provided end to end support by extracting data, preparing analysis, interpreting data, making strategic recommendations and presenting before internal client team
- Wrote SQL scrips and used alteryx to extract & clean data and to join from different data sources
- Familiar with Crystal Reports, and SSRS - Query, Reporting, Analysis and Enterprise Information Management
- Drafted Business Intelligence level analytics on Pentaho and presented before clients
- Explored other databases including DB2, Teradata and MySQL.
- Worked with machine learning tools and libraries such as Scikit-learn, R, Spark.
- Reviewed suspicious activity and complex fraud cases to help identify and resolve fraud risk trends and issues.
- Analyzed of customer data to tune rules, exposes patterns, research anomalies, reduce false positives, and build executive and project-level reports
- Identified meaningful insights from chargeback data. Communicated analysis to engineers, product and stakeholders.
- Analyzed high-volume data to investigate, identify and report trends linked to fraudulent transactions
- Utilized Sqoop to ingest real-time data. Done Exploratory Data Analysis, trying to find trends and form clusters.
- Built models using techniques like Regression, Tree based ensemble methods, Time Series forecasting, KNN, Clustering and Isolation Forest methods, C4.5, SVM and others
- Worked both with structured and semi structured data
- Heavily utilized pandas, keras and other libraries and maintained RDD's using SparkSQL.
- Communicated and coordinated with other departments to collection business requirement.

## Medicare & Medicaid (CMS)

VA

DATA SCIENTIST (CONSULTANT) & TABLEAU DEVELOPER

Oct. 2013 - Nov. 2015

- Built sentiment analysis models that scored all customers interactions with the organization and a topic model that automated classification of text conversations.
- Worked on customer journey analytics, enhancing knowledge on various channels, hence, improved the overall customer satisfaction and reduce operational cost.
- Developed accurate and insightful visualizations and dashboards using Tableau.
- Implemented data mining and machine learning solutions to various business problems, hence, resolving issues.
- Applied LSTM networks to predict cost of healthcare services for individual customer.
- Used Tableau/ QlikView to design dashboards, reporting and to do high level BI.
- Analyzed and extracted relevant information from large amounts of data to help automate for self-monitoring, self-diagnosing, self-correcting solutions and optimize key processes.
- Built predictive/machine learning models using Regression, Boosting, GBM, NNs, HMMs, CRFs, MRFs, other deep learning, Decision Tree, Random Forest, Naïve Bayes, Correlation, Multivariate Analysis, Logistic Regression and Cluster Analysis.
- Done coding in R, Python, C, C++, SQL and UNIX.
- Strong Data Warehousing ETL experience of using Informatica.
- Done data mining in various data sources including but not limited to Oracle, MSSQL Server, DB2, ODS, and Hadoop based data storage systems.
- Done data mining, Factor Analysis, ANOVA, Hypothetical testing, normal distribution and other advanced statistical techniques.
- Excellent knowledge on Normalization (1NF, 2NF, 3NF and BCNF) and De-normalization techniques for improved database performance in OLTP, OLAP and Data Warehouse/Data Mart environments.
- Excellent track records of working with complex data sets and translating data into insights to drive key business decisions.

## Kaiser Permanente

MD

SENIOR DATA ANALYST

Feb. 2011 - Sep. 2013

- Built defection risk models to be put into production internally. These models were being used by those who interact with our members.
- Done cluster analysis to quickly understand the varied needs of our insurance members. These population groupings gave quick understanding on customer experience with kp.org.
- Wrote SQL queries for data extraction, manipulation and formation of Table.
- Maintained a Python/Django web application.
- Done coding in R and Python to form, test and tun predictive model.
- Wrote a python-based SQL generator that helped speed up a weekly reporting from several days.
- Assisted in tuning of the database, which includes indexes, and optimizing SQL statements, monitoring the server.
- Co-ordinated with various business users, stakeholders and SME to get Functional expertise, design and business test scenarios review, UAT participation and validation of financial data.

## Certifications

- Certified Tableau Desktop Specialist, data analyst, data scientist, designer.
- Data Science professional certification- Harvardx (9 Courses on Data Science).
- Data Science Specialization- Johns Hopkins University (9 courses on Data Science).
- Data Science and Data Analyst (R, Python, SQL, Reporting) Certifications courses from datacamp.

## Skills

<b>Machine Learning</b>	Linear Regression, Logistic Regression, Ridge Regression, Lasso Regression, Multicollinear regression Decisions Tree, Random Forest, AdaBoost, Artificial Neural Network, Bayesian Network/BBN, K-NN, SVM, K-means Clustering, and PCA, MCA, and other data mining, Statistical and ML algorithms
<b>Data Science and Analytics</b>	Tableau, Tableau Prep, Alteryx, Power BI, QlikView/ QlikSense, Google Analytics, R, R-Shiny, Python, SQL Server, Teradata SQL, SSIS, Jupyter notebook, Markdown, Github, Regex expression, familiar with PySpark, shell scripting, Gretl, Scala, spark, Sqoop, MapReduce, Amazon S3, Azure and Yarn.
<b>Web Development</b>	React JS, Firebase Authentication, JWT, Firebase Hosting, node JS, Express JS, React router, React hook, Redux, MongoDB, Heroku, Netlify, Vanilla JavaScript, ES6, Chrome Dev Tools, HTML, CSS, Bootstrap, Material UI, SEO, CDN
<b>Other Tools</b>	Excel (VLOOKUP, Pivot Table, IF Statement, excel vba etc), PowerPoint, Word, Visio, FireFTP, WinSCP, Firebug, Putty, VS Code
<b>Operating System</b>	Windows, UNIX, Linux, Kali Linux, Mac O

# Education

---

## **The City College of New York**

MASTER'S IN INFORMATION SYSTEMS, CGPA- 3.77/4.00

New York, NY

Aug. 2016 - Jun. 2018

## **Shahjalal University of Science and Technology**

BACHELOR IN CHEMICAL ENGINEERING, CGPA- 3.51/4.00

Sylhet, Bangladesh

Aug. 2005 - Dec. 2008

**Work Authorization:** US Citizen

**References:** Available upon request.