**Soundarya Parvatam**

**+1(573) 494-2667**

[**soundaryabala57@gmail.com**](mailto:soundaryabala57@gmail.com)

**Data Analyst Profile :** [**https://public.tableau.com/app/profile/soundarya.bala.parvatam**](https://public.tableau.com/app/profile/soundarya.bala.parvatam)

# CareerSummary:

* **Around 6+ years of professional experience in Machine learning, Deeplearning and Data science technologies with extensive experience working with Business Intelligence/data visualization tools with specialization in Tableau.**
* Experienced in employing Python, R Programming, Tableau and SQL for data cleaning, data visualization and predictive analytics
* Experience in Installation, Configuration and administration of Tableau Server in a multi-server and multi-tier environment including AWS.
* Introduced Tableau to Clients to produce different data visualizations and presented dashboards on web and desktop platforms to End-user to help them make effective business decisions.
* Administered users, user groups, and scheduled instances for reports in Tableau and documented upgrade plan from Tableau 10.2.2 to 10.5.2
* Expertise in Software Design, Development, Integration, Implementation and Maintenance of Business Intelligence and Database Platforms.
* Knowledge of Data warehouse concept, Data modeling principles - star schema, snowflake schema, normalization/de-normalization techniques.
* Well-versed in analytics to find meaningful insights fromdata.
* ProfessionalexperiencewithstrongcodingskillsinPythonandR.
* ManualdatavalidationinexcelusingVLOOKUPandpivottables.Thishelpsustovalidatewhetherthe sales are correct in the system and the file weget.
* CreatedadatavalidationsystemusingSQLqueriesandWebreportsdemonstratingtheresultsinbusiness terms.
* Experience as a Data Analyst and in converting user requirement documents to business requirement documents.
* Used a blend of Azure Data Factory, T-SQL, Spark SQL, and U-SQL Azure Data Lake Analytics, gather, convert, and load the data from source systems to Azure Data Storage service
* Strong experience with ETL Alteryx.
* Installed Tableau Server in a Distributed Environment and worked on Worker Nodes and added additional worker nodes.
* Connected to different databases such as My SQL, MS SQL Server, Snowflake, Oracle, etc.,
* Experience in Publishing and Sharing Workbooks, Data Sources with User Filter to Tableau Server, Setting Schedules, Subscriptions, Authorization and Permissions, incremental refresh extracts and updating server data that are using extracts.
* Used Configuration utility to work on setting up of worker nodes by configuring and updating Primary nodes
* Expert in Data Extraction from multiple sources and creating visulalization using reporting tools like Power BI
* Extensively worked on setting up of Sites, Projects, groups, Users, Schedule refreshes & subscriptions
* Monitored server health for configuration alerts and subscriptions and optimized workbooks using Performance Optimization options.
* Experience with statistical methods, and versioncontrol(git).
* Created Custom Visualization of data usingTableau.
* Excellent team worker, leader and mentor, thrives in pressure and challenging working environment

# Experience:

# Designation: Data Analyst

**Company: Huntington Bank**

**Duration: Feb 2022 – Present**

* Requirement gathering and convert business language to technical language.
* Built charts, graphs and formuals using Excel.
* Responsible for developing , supporting and maintenance for the ETL processes using Informatica PowerCenter.
* Extensive SQL querying on Staging , Datawarehouse and DataMarts.
* Develop Mappings and Workflows to generate staging files.
* Worked in various transformations like lookup, joiner, sorter , Aggregator.
* Developed SQL quesries for querying data against data bases for data verification process.
* Prepared the test plan and testing strategies for Data Warehousing Application.
* Tested number of complex ETL mappings, mapplets and rsusable transformations for daily data loads.
* Manipulating, cleansing & processing data using Excel, Access and SQL.
* Used Data mapping to match the fields from SQL and Snowflake.
* Extensively worked with pivot tables, v-lookup, graphs and charts to analyse large sets of data using Excel.
* Created dynamic repots with pivot tables using Excel.
* Created tabular model reports to slice and dice huge volumes of data.
* Experience in combining difference sources by using power query.
* Experience in data transformations such as adding calculated columns, manage relationships, create different measures, remove rows, replace values, split column, date and time column.
* Experience in combining data from different sources by using power query.
* Troubleshoot and support reports in PowerBI workspace whenever there are issues in production.
* Utilized (Power BI, Pivot/View) to design multiple scorecards and dahsboards to display information required by different departments and upper level management.
* Extensive use of DAX functions for the reports and for the tabular models.
* Created Power BI reports by using joins in multiplr tables from multiple data bases using complex SQL queries.
* Experience in scheduling refresh of Power BI reports on- demand.
* Interact with different teams and analyze joins and required data using Salesforce, Excel, Power BI

and Microsoft SQL.

* Developed visual reports, dashboards, Cubes and KPI scorecards using Power BI. publishing and scheduling Power BI reports as per the business requirements.
* Implemented row level security on data and have an understanding of application security layer models in Power BI.
* Used Python to understand the relationship between different attributes of dataset and causal relationship between them.

# Designation: ETL DATA ANALYST

# Maximus, Falls Church, VA.

# April 2021 – Jan 2022

* Analyzed healthcare data and fraud claims.
* Involved in writing complex quesries using SQL for data integrity checks.
* Validating the data in various stages of data movement from Source to Staging and from Staging to Data Warehouse tables.
* Created Test Cases using SDLC procedures and reviewed them with the Test Manager.
* Validating the load process of ETL to make sure the target tables are populated according to the data mapping provided that satifies the trasnformations rules.
* Built and published customized interactive reports, dashboards and created report schedules using Tableau server.
* Created data blending, hierarchies, calculated fields, filters, parameters, groupings, relationships, actions, in both Tableau and excel.
* Created different reports using line charts, text tables, tree maps, pie, heat maps, pivot tables, bar, combination charts in excel and tableau.
* Took place in creating stored procedures, views, and custom SQL queries to import data from SQL server to Tableau.
* Created views in Tableau Desktop that were published to internal team for review and further data analysis and customization using filters and actions.
* Involved in creating interactive, parameterized dashboards by extracting data from different sources using data blending and applying actions - filter, highlight and URL.
* Created dashboard from scratch that included drill-through, filters, Heat/Geo/Tree Maps, Pie/Bar Charts, Circle Views, Line/Area Charts, Scatter Plots and Bullet Graphs.
* Defined and documented the best practices for Tableau report development and effectively used data blending feature in tableau.
* Processed data using Python pandas to examine the claims data , identify outliers and inconsistencies.
* Wrote complex custom SQL to gather data from multiple databases.
* Developed Complex Stored Procedures and views to extract multi source data for Tableau reports.
* Experience in redesigning as well as end to end development of complex Tableau visualization for business users.
* Conducted exploratory data analysis using python NumPy and Seaborn to see the insights of data and validate each feature through different charts and graphs.
* Analyzed data from various databases and formatted it into various visuals using Tableau Desktop.
* Designed Technical design document and provided Training plans for the business users

# Designation: DataAnalyst

**Company: Accenture Private Limited Duration: June 2016 – September 2019**

**Data Analyst** *(Project: Confidential, Client: Leading Global Pharma Company)*

* Responsible for analyzing pharmaceutical data from multiple sources.
* Collected, cleaned, and furnished structured data analysis for sound decision-making.
* Performed Exploratory Data Analysis to understand and analyze the data.
* Used machine learning algorithms such as random forest, k-nearest neighbors, and support vector machine to predict the company sales with an accuracy rate of 91%.
* Presented data and conclusions to the team to improve business strategies.
* Performed ad-hoc analysis requests for stakeholders.
* Created visually impactful dashboards for data reporting.
* Experience in working with GitHub and ETL pipeline.
* Experience in Agile environment and familiar with JIRA

**SQL Developer** *(Project: Confidential, Client: Leading Global Pharma Company)*

* Responsible for Enhancement Releases from Planning, Design, and Testing to Deployment stages.
* Excellent grasp of Data warehousing and ETL Process.
* Performed Data extraction and Analysis from various data sources.
* Performance Enhancement of database by optimizing the queries.
* Created Custom Visualization of data upon requests for users across 26 European countries.
* Gathered extensive requirements from clients for various change requests.
* Extensively investigated different data issues and communicated with the clients to resolve the issues.
* Maintained comprehensive documentation of Estimates, Deployment, and Design documents.

**Academic Projects**

# Project #1:

**Disease Prediction: -**

**Algorithms:** [Random Forest](https://en.wikipedia.org/wiki/Random_forest), XGBoost, Convolutional neural network.

**Technologies:** TensorFlow, Keras, scikit-learn,

**Data:** Breast Cancer, Diabetes, Heart Disease, Covid-19, Pneumonia, Alzheimer's from Kaggle.

* Breast Cancer Detection: Used [Random Forest](https://en.wikipedia.org/wiki/Random_forest) for this use case. The accuracy achieved was around 91.81%.
* Diabetes Prediction: Used Random Forest for this use case. The accuracy achieved was around 66.8%.
* Heart Disease Prediction: Used XGBoost for this use case. The accuracy achieved was around 86.96%.
* Covid-19 Prediction: Used custom-made CNN architecture for this detection. The accuracy achieved was around 93%.
* Pneumonia Prediction: Used custom CNN architecture for this use case. The accuracy achieved was around 83.17%.
* Alzheimer Prediction: Used Random Forest for this use case. The accuracy achieved was around 73.54%.
* I have created an app using FLASK in which it will take input from the users and predict if there is any chance of these diseases.

# Project #2:

**Climate change impact on Agriculture: -**

**Algorithms:** principal component analysis, k-means clustering, fixed effect model.

**Data:** Collected from world bank data.

**Implementation:** The hypothesis is that climate change variables like greenhouse gases impact on agriculture production. In this I have used principal component analysis to reduce the four greenhouse gases into 2 dimensions.BasedonthisIhavedoneclusteranalysis,togetoptimalnumberofclustersIhaveusedsilhouetteand gap statistic method and used these results in statistical analysis like panel data methods and I have found that agriculture is impacted by greenhousegases.

# Project #3:

**Amazon Fashion Recommendation: - ML Problem:** [recommended systems](https://en.wikipedia.org/wiki/Recommender_system)

**Algorithms:** dimensionality reduction, collaborative filtering, classification algorithms.

**Technologies:** sci-kit-learn, tensor flow, Keras and NLTK.

**Data:** Amazon fashion items from amazon organization Implementation: web application.

**Role:** preprocessing of data using good feature engineering techniques and developed model using machine learning and neural networks algorithms.

# Project #4:

**Traffic signs Recognition:**

**ML Problem:** classification problem **Algorithms:** convolutional neural network **Technologies:** sci-kit-learn, tensor flow, Keras

**Role:** classified the traffic signs classifier with 95% accuracy and visualized how our accuracy and loss change with time, which is pretty good from a simple CNN model.

# Project #5:

**Amazon product review : -**

**ML Problem:** sentiment analysis

**Technologies:** sci-kit-learn, pandas

**Role:** preprocessing of data using good feature engineering techniques and developed model using machine learning techniques.

# TechnicalSkills:

* ProgrammingLanguages : Python andR.
* WebDesigning : HTML, CSS,JAVASCRIPT.
* Database : ORACLE, MYSQL,SQL.
* **Supervised machine learning**: Linear models, Decision Trees, Neural networks, Random Forest, Support Vector Machines, clustering, causal inference, Natural Language Processing, anomaly detection, Gradient boosting, Adaboost, XGBoost,KNN.
* **Unsupervisedmachinelearning**:k-means,DBscan,Hierarchical,Principalcomponentanalysis.
* **Statistical techniques**: hypothesis testing, ANOVA, correlation, regression analysis, Bayesian method, Inferentialstatistics,Descriptivestatistics,Hypothesistests,t-test,ANOVA,chi-squaretest,type1error, type 2 error, one tail test, two tail test,etc.
* Exploratory data analysis, Feature Engineering, Featureselection.
* **Deeplearning:**Artificialneuralnetworks,Convolutionalneuralnetworks,Optimizer,activation functions.
* **NaturalLanguageProcessing:**Tokenization,Lemmatization,Stopwords,Bagofwords,TFIDF, Unigrams, Bigrams, Word Embedding, word2vec,Avgword2vec.
* **Reporting Tools:**Tableau and Power BI.

# EducationalQualifications:

* *Master of Science* in Applied Statistics andDecisionAnalytics 2021

WesternIllinoisUniversity GPA3.9

* *Bachelor of Science* in ComputerscienceEngineering 2016

Gandhi Institute of ManagementandTechnology GPA7.4

# Achievements:

* VolunteeredfromAccentureinrescue&reliefoperationsheldduringChennaifloodsinIndia.
* VolunteeredforUnitedAnantapurwelfareorganizationbyteachingkidsinruralareasbySupport Budding Knowledge ’14program.
* Volunteered for blood donation camp in my university atGITAM.