# Profile Summary

Strategic and detail-oriented Data Analyst with 2+ years of experience delivering data-backed insights that drive business performance and client success. Skilled at understanding complex business contexts, identifying opportunities through data, and communicating insights with clarity and confidence. I bring a consulting mindset to every project—thinking beyond the immediate ask, aligning with business goals, and building trust through accuracy, ownership, and thoughtful analysis. Passionate about solving real-world problems in fast-paced, impact-driven environments.

* GitHub : https://github.com/Balaji036
* Kaggle : https://[www.kaggle.com/balaji036](http://www.kaggle.com/balaji036)
* LinkedIn : https://[www.linkedin.com/in/balajiraja28/](http://www.linkedin.com/in/balajiraja28/)

# Professional Summary

# ZoomInfo Technologies June 2023 – Present

# Data Analyst 1 Chennai, India

* Collaborate with multiple clients on daily basis to understand their requirements, offer strategic feedback, and deliver data in requested formats
* Developed and implemented an automation to address a product limitation, enabling $5.7M+ in deal closures and increasing ZoomInfo’s competitive value.
* Deal with high-value Enterprise and Strategic accounts, contributing to renewals, up-sells, and critical new business deals
* Identified anomaly in Round Robin logic, led a team of 4 to redesign and successfully implement a scalable solution
* Appointed as Data Steward to handle critical requests, ensuring consistent delivery of accurate, high-quality data
* Solely handled end-to-end data and business analysis for Grafana, TeamViewer, and Employsure (Peninsula), directly supporting closed deals totaling $212K, $70K, and $194K in ACV

# KEY SKILLS

* **Programming Languages: Python, SQL, Java**
* **Big Data & Distributed Compute:**PySpark, Databricks, Snowflake, Cyberduck
* **Visualization Tools:** Tableau, Excel
* **Data Analysis:** Data Cleaning, EDA, Business Insights, Data Storytelling
* **Machine Learning:** Linear Regression, Logistic Regression, Lasso, Ridge, Hypothesis Testing
* **Python Libraries:** Pandas, NumPy, Matplotlib, Seaborn, Plotly, Scikit-learn, Statsmodels, Itertools
* **Data Querying:** DDL, DML, DQL, TCL, MySQL Workbench, PostgreSQL
* **Environments: Jupyter** Notebook, Visual Studio Code, Sublime Text
* **Enterprise Tools:** Confluence, JIRA, Salesforce, Docket AI
* **Professional & Soft Skills:** Client Communication, Requirement Gathering, Business Understanding, Data Interpretation, Stakeholder Collaboration, Analytical Thinking, Problem-Solving, Cross-Functional Coordination, Data-Driven Decision Making, Leadership, Adaptability, Ownership, Time Management, Attention to Detail, Agile Mindset

# ACADEMIC PROJECTS

## Project 1: Titanic Disaster: A logistic regression-based Survival analysis Source : Kaggle

**Goal** : Examine variables that contributed to survival and use predictive modelling to predict individuals who survived the Titanic disaster

## Synopsis:

* Conducted data profiling to assess structure, completeness, and suitability of the dataset for analysis.
* Developed modular EDA scripts to handle multiple file formats, detect duplicates, analyze missing values, and explore univariate to multivariate relationships.
* Built reusable data transformation modules for imputation, categorical encoding, outlier removal, feature scaling, and train-test splitting.
* Executed multiple model iterations and hyperparameter tuning; selected optimal configuration based on correlation analysis and performance metrics.
* Delivered a logistic regression model with 78% accuracy, demonstrating effective end-to-end implementation from raw data to predictive insight.

**Key skills**: Python , Exploratory Data Analysis(EDA) ,Logistic Regression, Pandas, NumPy, Matplotlib, Seaborn, Confusion Matrix, OneHot Encoder, Label Encoder, Zscore Outlier Handling, Standardization, Classification Report, OOPS(Object Oriented Programming)

## Project 2: Used Car Price Prediction – A Linear Regression based Model

## Source : Kaggle

## Goal : To build a predictive model that accurately estimates used car prices based on historical data, enabling data- driven pricing decisions.

## Synopsis:

* Conducted data profiling and Performed EDA to detect data inconsistencies including duplicates, encodable categorical variables, and outliers using pair plots and box plots.
* Removed duplicate records, applied appropriate encoding techniques, and eliminated outliers using Z-score with ±3 threshold.
* Split limited dataset into 60/40 train-test ratio and applied normalization to scale features for consistent model input.
* Conducted multiple iterations with different configurations; selected optimal model using MSE, adjusted R², and correlation-based feature selection.
* Achieved an R² of 82% and reduced Mean Squared Error (MSE) to 205, demonstrating strong model performance and predictive accuracy.

**Key skills:** Python,Exploratory Data Analysis(EDA) , Linear Regression, NumPy , Pandas , Matplotlib , Seaborn, MinMaxScaler , OneHotEncoder, Label Encoder, Zscore Outlier Handling, R-Square, Adjusted R-Square, OOPS(Object Oriented Programming)

# HACKATHONS/OTHER CERTIFICATIONS

* Participated in the GLCA Data Science Hackathon program conducted in 2023.
* Taken part in IEEE 24 Hours Coding Marathon Program twice(2019 & 2020)
* HackerRank SQL Gold badge
* Apache PySpark By Example
* Tableau Essential Training
* SQL For Data Analysis, Intermediate SQL For Data Scientist

# EDUCATION

|  |  |  |  |
| --- | --- | --- | --- |
| **Course** | **Institution** | **Year** | **Remarks** |
| Data Analytics | Great Lakes Institute of Management | 2023 | Completed |
| BE Computer Science Engineering | Panimalar Engineering College | 2022 | 89.60% |
| 12th Grade | St. Joseph of Cluny Matriculation Higher Secondary School | 2018 | 87.3% |
| 10th Grade | St. Joseph of Cluny Matriculation Higher Secondary School | 2016 | 96% |

**OTHER ACHIEVEMENTS**

* Dream Winner – Q2 2024, ZoomInfo (Top Performer of the Quarter)
* **5× New Business Star** – Honored in Revenue All Hands for consistent top-tier performance.