

SAI BHAVAN GUBBALA

[GitHub](#) | saibhavan9@gmail.com | +91 93988 10746 | [LinkedIn](#)

TECHNICAL SKILLS

Technical Skills – Python, Jupyter Notebook (Pandas, NumPy, Seaborn, Matplotlib), R, Advanced in MySQL Server, MongoDB, Power BI, Tableau, Advanced Excel.

Communication Skills – Telugu (Native), Hindi, English (TOEFL CEFR level C1), Japanese (Learning)

Core Analytical Skills – Data Transformation, Data Manipulation Data Mining, Data Visualization, ETL, Report Generation, Business intelligence, Project Management, Statistics Analysis (Hypothesis testing, A/B Testing), Predictive Modeling. Snowflake schema and Star Schema.

WORK EXPERIENCE

HCLTech

Machine Learning Intern

JAN 2022-MAY 2022

HEART DISEASE PREDICTION

- Conducted Comparative analysis of 5 machine learning algorithms. Random Forest ML algorithm achieved the highest accuracy at 85.22%. Improved 7% accuracy after analyzing 15 random sample datasets.
- Conducted in-depth exploratory data analysis (EDA) to evaluate key features influencing model performance, resulting in improved prediction accuracy by 4%.
- Implemented robust data preprocessing techniques, including handling missing values and outliers (3%), to ensure data quality and model reliability.
- Deployed ML model into a production environment, enabling real-time predictions and automated decision-making.

PROJECT EXPERIENCE

Bike Sales Analysis

Jun 2024 - July 2024

- Performed in-depth analysis of bike sales data leading to \$24.9M revenue, \$10.5M profit, and 25.2k orders while maintaining a low 2.2% return rate.
- Optimized product mix and market strategy by identifying bikes as the primary revenue driver (92% of profit) and uncovering a 34% higher order volume in the US compared to other regions.

Maven Market Analysis

Apr 2024 - May 2024

- Developed and designed 3 dashboards visualizing over 0.5 million data points for actionable insights.
- Constructed KPIs to monitor monthly sales, profitability, and return rates, identifying underperforming products.
- Executed in-depth sales analysis across the USA, Mexico, and Canada, identifying a 51.08% sales advantage for the USA. Quantified the profitability of Plato brand products, demonstrating a substantial 63.55% profit margin.

Massachusetts General Hospital Analysis

May 2024 - July 2024

- Performed exploratory data analysis (EDA) to identify trends, patterns, and outliers in surgery types, costs, and insurance coverage. Total surgery costs amounted to \$105.52M, with an average cost of \$2.21K per surgery.
- Analyzed 48K surgeries, identifying electrical cardioversion being the most common, with Medicare covering approximately 41% of total surgeries.

Human Resources Analytics (HR)

Feb 2024 - Apr 2024

- Engineered a custom function to transform raw data, facilitating efficient automated data integration and normalization across multiple spreadsheets saving time by 30%.
- Enhanced data consistency and efficiency by 20%, facilitating rapid analysis and actionable insights.
- Identified a 15% increase in work-from-home requests and observed 3% rise in sick leave occurrences overall, with a slight concentration on Mondays. signaling potential issues with employee well-being issues.

USA Online Store Sales Analysis

Mar 2024 - Apr 2024

- Uncovered technology as the top-revenue-generating category, capturing 38% of total sales, eclipsing furniture (35%) and office supplies (32%).
- pinpointed New York City as a key market, generating a substantial \$103,852 USD profit over two years.
- Discovered Q1 and Q4 as peak revenue quarters, contributing to 36.14% of annual sales, highlighting significant sales opportunities during these periods.

EDUCATION

Sathyabama Institute of Science and Engineering

Jun 2019 - May 2023

Bachelors of Engineering in Computer Science & Engineering - Cumulative GPA - 9.15/10

CERTIFICATIONS

- Microsoft Power BI for Business Intelligence (Udemy)
- Google Data Analytics Professional Certificate (Coursera)
- Computer Science for Artificial Intelligence (Harvard)
- Data Analytics and Visualization (Accenture)
- Data Programming with R (Coursera)