HEMALATHA M

Bangalore, Karnataka, India hemalatham236@gmail.com, 7975025711, Github, https://www.linkedin.com/in/hemalatha-m-84b561233/

SUMMARY

Data Analyst with hands-on experience in **Python, Excel, Tableau, and Machine Learning**, specializing in **data extraction, preprocessing, statistical modeling, and visualization**. Proven ability to derive actionable business insights through **time series forecasting, predictive analytics, and dashboard reporting**. Seeking to contribute analytical expertise to innovative team.

SKILLS

- Data Analysis & Visualization: SQL, Excel, Tableau
- Programming: Python,
- Database: MySQL
- Statistical Modeling & Forecasting: EDA, Hypothesis Testing, Regression, ARIMA, XGBoost, Prophet
- Data Quality & Validation: Cleaning, Aggregation, Feature Selection
- Tools & Platforms: Jupyter Notebook, PyCharm, Visual Studio Code, GitHub
- MLOps & Deployment: CI/CD Pipeline, MLflow, AWS, Amazon SageMaker, Flask, Streamlit

EXPERIENCE

DATA SCIENCE INTERN(Time Series Forecast)

C5i Company- Bangalore **June 2024 - Dec 2024**

Time Series Forecasting of Temperature Trends Using AirQualityUCI:

- Extracted and transformed large datasets using Python, ensuring data readiness for analysis.
- Conducted data cleaning, aggregation, and resampling on AirQualityUCI data (9357 records to 391 daily points).
- Applied time series decomposition, stationarity tests, and feature selection using Granger Causality, correlation, and Lasso Regression.
- Built and fine-tuned models (ARIMA, SARIMA, XGBoost, Prophet, Holt-Winters) for temperature forecasting.
- Evaluated model performance using **MAE.**
- Delivered actionable insights for optimizing greenhouse heating, cooling, and irrigation systems.
- Collaborated in dashboard development for data visualization and reporting.

EDA on Ad Performance Data.

- Conducted in-depth EDA to analyse the relationship between Customer Engagement Score(CES) and Ad performance metrics(Amount spend Impression and click).
- Evaluated CES components including Recall, Cognitive Load, Ad Copy Effectiveness(Persuasiveness, text readability, attention).
- Investigated industry-wise CES distribution and spending patterns to identify correlations with ad performance outcomes.

Application Development using Flask in Python with CSS, HTML and JavaScript | Link

- Designed a data upload and visualization interface supporting XLSX, CSV, and Parquet files with error handling for unsupported formats.
- Implemented functionality to display the top 10 rows of the dataset upon file submission.
- Developed a visualization page enabling scatter plot creation by selecting two distinct numerical variables.
- Added dynamic dropdowns that update to prevent duplicate variable selection

PROJECTS

1. **Netflix Business Case Study |** Github Link SEP 2025

- Analyzed 8,800+ Netflix titles to uncover global content trends, audience preferences, and regional production patterns.
- Analyzed the distribution of movies and TV shows on Netflix to understand any trends or changes over time.
- Identified popular genres and evaluated the trending genres to receive the highest ratings and user engagement.
- Discovered that on Netflix, movies are Approximately 128.94% greater than Tv shows & In 2017-2018 more than 700 movies were released, with TV-MA receiving highest rating of 3207.

2. Walmart Business Case Study | Github Link **OCT 2025**

- Analyzed 550,000+ Walmart customer records using Python and the Central Limit Theorem (CLT) to ensure reliable population inferences and 95% confidence intervals on mean purchases (~₹9,260).
- Identified key demographic trends - 26–35 age group and City Category B showed the highest average purchases (₹10,000–₹11,000), while marital status showed no significant difference (overlapping confidence
- Discovered gender-based spending gap (~₹700) and high-value occupations (codes 4, 7, 10), delivering actionable insights that guided targeted marketing and regional expansion strategies for revenue growth.

3. Yulu Business Case Study: **Github Link** **OCT 2025**

- Analyzed bike-sharing demand data using Python (Pandas, NumPy, SciPy) to identify key factors influencing ride countsuch as season, weather,and working days.
- Performed ANOVA, t-tests, and correlation analysis to test hypotheses, revealing significant seasonal variation ($p < 0.05$) with higher demand during summer and clear weather condition.
- Delivered data-driven insights and visualizations that supported resource optimization and operational planning, improving demand forecasting accuracy

4. Target Business Case Study using SQL:

- Analyzed customer order patterns using SQL, identifying peak times and cost inefficiencies.
- Achieved 136% increase in cost efficiency through data-driven strategies.

EDUCATION

- Pursuing Data Science specialization program from Scaler Academy - **12/2023-Present.**
- Post Graduate Programme in Data Science. from Careerera. - **07/2023 - 08/2024 .**
- Bachelor of Science(Agricultural Marketing & Co-operation)Gandhi Krishi Vignana Kendra (GKVK) - Bangalore - **2012-2016**

CERTIFICATES

SQL Certificate from Scaler Link
Python Libraries Certificate from Scaler Link
EDAFundamentals Certificate from Scaler Link
Data Science Certificate from Careerera- Link