THASNI M A

Thrissur, Kerala, India

+918590793528

GitHub

thasnihakeem2017@gmail.com

LinkedIn

Kaggle

💼 WORK EXPERIENCE

Junior Data Analyst, *Lulu Group India - Head Office*
Oct 2023 – present

- **Set up a secured MongoDB database** and implemented a backup system for storing daily sales POS data. Developed scripts for converting XML data and uploading it via daily scheduled tasks. Additionally, created indexes and views within the database and configured it for various purposes, including API integration for application use.
- **Automated data extraction** from the database by developing Python scripts for efficient querying and processing of information.
- **Developed a web application** for XML analytics of sales data using Flask and MongoDB as part of the Carnival XML log analysis project.
- **Worked extensively on sales and stock analytics**, creating several analytical dashboards using Power BI.

Data Science Intern, *Hashwave Technologies Pvt Ltd*
Dec 2022 – Oct 2023

- **Gained experience in web scraping and data analysis** using Python, with tools and libraries including **Pandas, NumPy, Seaborn, Matplotlib, Beautiful Soup, Playwright, and Selenium.**
- **Conducted extensive data analysis and visualization** with *Tableau*.
- **Authored technical blogs** on web scraping techniques, which were published on the **Datahut website** to share insights and knowledge with a wider audience.

Data Science Intern, *SEQATO Software Solutions*
Aug 2022 – Oct 2022

- **Developed an efficient food recommendation system** using both content-based and collaborative filtering techniques. Conducted evaluations and analysis of real-world datasets to derive quantitative and fundamental insights.
- **Acquired valuable experience** in data wrangling and feature engineering, with a focus on deriving actionable insights and optimizing recommendation algorithms.

🎓 EDUCATION

MSc Data Science,
Department of Futures Studies, University of Kerala
2020 – 2022 | Thiruvanathapuram, India
CGPA - 8.8(Out of 10)

BSc. Mathematics,
Sree Narayana College, Calicut University
2017 – 2020 | Nattika
CGPA - 3.47(Out of 6)

💻 SKILLS

Programming Languages:
- Python, SQL, MQL

Database
- MongoDB, MySQL

Data Science Skills
- Machine Learning
- Web Scraping
- Exploratory Data Analysis (EDA) & Visualization
- Natural Language Processing (NLP)
- Recommendation System Model Building
- Building Web Applications for Data Insights

Visualization Tools:
- Power BI
- Tableau

Web Development & Automation:
- Django, Flask (Backend Development), React, HTML, CSS, Bootstrap (Frontend Development)
- Beautiful Soup, Selenium, Playwright (Automation & Web Scraping)

Development Tools & Productivity Tools::
- VS Code (Code Editor)
- PyCharm (Python IDE)
- Jupyter Notebook

Soft Skills
Communication, Presentation, Report Making

📁 PROJECTS

PRODUCT MATCHING USING BERT EMBEDDINGS AND FUZZY MATCHING FOR E-COMMERCE PLATFORMS ↗

- Led a project focusing on matching products across two major e-commerce platforms, Amazon and Flipkart, utilizing advanced natural language processing techniques.
- Executed data preprocessing and **vectorization** using **CountVectorizer**, followed by **cosine similarity** calculations to assess product name similarity.
- Leveraging libraries like **pandas**, the project involved rigorous model training to identify optimal matches based on product name similarity.
- Implemented a multi-stage matching process, including setting thresholds for product name, brand, color, capacity, and model matching.
- Demonstrated expertise in natural language processing and model evaluation methodologies throughout the project's lifecycle.

WEB SCRAPING A DYNAMIC WEBSITES WITH PYTHON AND A COMPETITIVE ANALYSIS OF THE BEST-SELLING WOMEN FRAGRANCES ↗

- Extracted large volumes of women's fragrance data from Ulta Beauty's website using **web scraping** techniques with **BeautifulSoup, Selenium, and Playwright**.
- Utilized **lxml** and **HTML parsers** to parse and process the extracted data, ensuring its accuracy and completeness.
- Cleaned and prepared the data for analysis using **NumPy and Pandas**.
- Conducted a comprehensive analysis of women's fragrances sold at Ulta Beauty, with visualizations created using **Seaborn, Matplotlib, and Tableau**.

  - https://www.blog.datahut.co/post/scrape-a-dynamic-website-using-python
  - https://www.blog.datahut.co/post/sephora-vs-ulta-beauty-competitive-analysis

FOOD RECOMMENDATION SYSTEM USING COLLABORATIVE FILTERING APPROACH

- Developed an effective food recommendation system by implementing collaborative filtering approaches, including both **memory-based** and **model-based** methods, using the **Nearest Neighbour Algorithm**.
- Evaluated and analyzed real-world datasets from food tech firms to generate quantitative and fundamental insights, leveraging Python libraries such as **NumPy** and **Pandas**.

📖 PUBLICATIONS

STRUCTURAL ANALYSIS OF AIR TRANSPORT NETWORK USING NETWORK INDICATORS ↗

📖 CERTIFICATES

Deep Learning with PyTorch: Zero to GANs ↗
Jovian

Machine Learning with Python: Zero to GBMs,
Jovian ↗

Machine Learning through Python, *ICFOSS*

🧩 ACHIEVEMENTS

A 3X Kaggle Expert.

Earned Golden Badge for Python and Earned Bronze Badge for SQL in Hackerrank

Best Volunteer for the cyber crime awareness programme 'KOOTU 2022' for state level inauguration

🌐 LANGUAGES

- English
- Malayalam

📇 PERSONAL INFORMATION

Nationality
Indian

Date of Birth
31 August 1999

Gender
Female

Father's Name
M K Abdhul Hakeem

Marital Status
Single

✏️ DECLARATION

I declare that all the statements made in this C.V are true, complete and correct to the best of my knowledgeand belief and that in the event of any information being found false or incorrect, action can be taken against me.

THASNI M A

Sale Posting Data

Date	Store	Store Name	Total Payment	Expected Payment	Actual Payment	Cashier Short/Excess	ZDSR	ZRead	ZDSRZRead	ZPMC	Difference	Total Article Sale	Exception	Remarks
2025-02-08	9201	Lulu Hypermarket, Kochi	9603711.29	9603711.29	9603882.92	171.63	9603711.29	9603711.29	0.00	0.00	0.00	8070364.41	1533346.88	

Daily Checklist

Date: 2025-02-28

Date	Store Code	Store Name	Server Status	POS Status	Scale Status	PDT Status	Ind Store Backup	Sales Status	Acronics Backup	POS DB Backup	UPS Status	Remarks
2025-02-28	9201	Kochi	Not Ok	Not Ok	Not Ok	Not Ok	Not Ok	Not Available	Not Ok	Not Ok	Not Ok	Server - Server Name: C, Warranty Date: 2025-02-28, Status: Not Ok, Remarks: No remarks provided. POS - POS Number: 1, Status: Not Ok, Remarks: No remarks provided. Scale - Scale Number: 4, Status: Not Ok, Remarks: No remarks provided PDT - PDT Number: 2, Status: Not Ok, Remarks: No remarks provided Ind Store Backup - Server Name: 3, TPCentralDB: Not Ok, Remarks: No remarks provided. Ind Store Backup - Server Name: 5, TPCentralDB: Not Ok, Remarks: No remarks provided. Acronics Backup - Server Name: 3, TPCMDB: Not Ok, Remarks: No remarks provided. POS DB Backup - Date backup mismatch: File Name: 2, Size: 2, Backup Date: 2025-02-27, Remarks: No remarks provided. UPS - UPS Name: w, Warranty Date: None, AMC Period: None to None, Status: Not Ok, Remarks: No remarks provided.
2025-02-28	9401	Lucknow	Not Available	Not Available	Not Available	Not Available	Not Available	Not Available	Not Available	Not Available	Not Available	No remarks
2025-02-28	9211	Trivandrum	Not Available	Not Ok	Not Available	Not Available	Not Available	Not Available	Not Available	Not Available	Not Available	POS - POS Number: 2, Status: Not Ok, Remarks: No remarks

Date	Store Code	Store Name	Server Status	POS Status	Scale Status	PDT Status	Ind Store Backup	Sales Status	Acronics Backup	POS DB Backup	UPS Status	Remarks
												provided. POS - POS Number: 4, Status: Not Ok, Remarks: No remarks provided.