# Hamza Aziz Khan

hamzazizkhan@gmail.com

Kenyan national with an undergraduate degree from Malaysia, currently pursuing a master's degree at Uppsala University, Sweden.

## EDUCATION

**Uppsala University** **Expected graduation date: December 2025**
*Master's in Computational Science specializing in Data Engineering*

**Sunway University**
*Bachelor's in Actuarial Science*

## KEY COURSES TAKEN

**High Performance Programming (10hp)** **Jan. 2024 – Mar. 2024**
- Learned serial code optimisation techniques: compiler optimisation, arithmetical expressions, efficient memory usage and optimisations related to instruction-level parallelism (ILP).
- Project (C): Implemented a bucket sort algorithm in serial and parallel (using OpenMP) with a focus on performance using optimisation techniques.

**Statistical Machine Learning (5 hp)** **Jan. 2024 – Mar. 2024**
- Learned the mathematical models used in machine learning; Neural networks; convolutional neural networks.
- Project (Python, Scikit-learn): Used classification methods to determine whether the Hesa Fredrik siren could be heard based on various factors. The models were tuned using grid search and performance was evaluated using cross validation.

**Data Mining (7.5 hp)** **Aug. 2023 – Oct. 2023**
- Learned how to pre-process large complex databases and conduct data analysis on it.
- Project (Python, Pandas): Identified genes associated with Alzheimer's disease using K-Means clustering (from sklearn) on a large dataset of sick patients from the GEO NCBI database browser. Data pre-processing was done using Pandas and Matplotlib was used for visualization.

## PROJECTS

**Solving Sokoban puzzles using Artificial Intelligence (R)**
- Created a greedy search algorithm combined with breadth-first-search (BFS) to solve Sokoban puzzles in R.
- Solved 17 levels of the Sokoban puzzle set - Microban (by David W. Skinner) in under 5 minutes.
- Learned how to tune greedy search algorithms to achieve a specific goal for complex game environments.
- Developed an animated visualization in R that dynamically showcases the Sokoban puzzle solutions, replaying the moves like a movie for an intuitive and engaging experience.

**Focused Web Scraper (Python, SQL)**
- Developed an ETL pipeline using Python and BeautifulSoup to extract, transform, and load extensive MMA fighter data from Wikipedia into a SQLite database.
- Designed a focused crawling algorithm using BFS to selectively follow hyperlinks, ensuring the scraper only visits relevant fighter pages.
- Optimized HTML parsing techniques to efficiently extract and process structured data as part of the ETL workflow.
- Performed data analysis to investigate the effect of the seed link on the distribution of fighters (by country and employing organisation) selected by the spidering algorithm.

## SKILLS, LANGUAGES AND CERTIFICATIONS

- **Programming languages (proficient):** Python (with expertise in NumPy for numerical computing, Pandas for data manipulation, Matplotlib for data visualization).
- **Programming languages (basic):** R, React, C, SQL.

- **Software:** Microsoft Office Suite.
- **Languages:** English (Native).
- **Certifications:** Python for everybody specialization (5 courses – University of Michigan).