

# Reema Dass

Data Scientist | Data Engineer | Data Analyst



Email: [reema.g.dass@gmail.com](mailto:reema.g.dass@gmail.com)

Ph: +436767317181

Location: Vienna, Austria

[LinkedIn](#)

[GitHub](#)

[Website](#)

---

I am a Master of Data Science student with four years of expertise in data engineering and cloud computing and looking for a Full-time or part time opportunities. I enjoy leveraging machine learning and sophisticated analytics to convert data into useful insights that help organizations develop. My work has focused on developing clever, scalable solutions that have a real-world impact. I'm enthusiastic in finding innovative solutions to complicated challenges and making data work for everyone.

---

## Education:

**Masters in Data Science at Technical University of Vienna, Vienna**

**Mar 2022 – Dec 2025**

- Majoring in Machine Learning and Visual Analytics.

**Bachelors in Computer Science at VTU, India**

**Sep 2013 - Dec 2017**

- Thesis Topic: Image processing for a medical university for disease diagnosis.
- 

## Experience:

**Data Science Research Internship (FFG-Femtech) at SBA Research, Vienna**

**Sep 2023 - Feb 2024**

- Investigated security upgrades for edge-deployed Deep Neural Networks (DNNs), and proposed an alternative obfuscation method that reduced vulnerability risk by 10% through ethical attacks on the deployed model.
- Evaluated Trusted Execution Environments (TEEs) and encryption methods to validate and improve data security.
- Implemented improved obfuscation methods, reducing the chance of reverse engineering by 25% and considerably increasing the security of DNN model deployment.

**Software Developer at Nextpart, Linz**

**Nov 2022 - Jul 2023**

- Integrated three new features into a malicious analysis tool, improving threat detection accuracy by 20%. Developed Python logic to categorize threats as Malicious, Benign, Suspicious, or Unknown.
- Added external software connectors to the tool's architecture, enhancing its ability to identify and evaluate security issues by 30%, with internal decisions influenced by external integrations.
- Conducted vulnerability testing to uncover and address logic flaws, resulting in increased overall system resilience.

**Cloud Consultant at Deloitte, India**

**Aug 2021 - Feb 2022**

- Developed and executed data redaction and anonymization solutions for Personally Identifiable Information (PII), reducing exposure by 85% and enhancing compliance with data privacy regulations.

- Engineered and optimized data processing pipelines using AWS Fargate, achieving a 30% increase in processing speed and a 20% reduction in operational costs.
- Collaborated with over five cross-functional teams to ensure adherence to data privacy standards, resulting in a 95% compliance rate and mitigating legal risks.

Data Engineer at Stats Perform, India

Mar 2020 – Jul 2021

- Led the cleaning, processing, and analysis of extensive datasets for sports analytics, managing over 10 terabytes of data and improving accuracy by 50% using AWS Glue, AWS Crawler, Spark scripting, Redshift, and DynamoDB.
- Integrated AWS platforms with external APIs to enhance data pipeline scalability by 60% and reduce data retrieval times by 35%. Applied data storage solutions on S3 Glacier for auditing.
- Developed a self-sustaining Cloud ETL pipeline capable of handling dynamic loads, resulting in a 45% increase in system resilience and a 25% reduction in downtime.

Full stack Developer at Infosys Private Ltd, India

Dec 2018 - Feb 2020

- Led full-stack development for Belgium Post, focusing on enhancing data storage and retrieval operations. Achieved a 50% reduction in query response times and a 40% improvement in storage efficiency using backend technologies such as Azure Blob Storage, SQL for data tracing, C#, and RESTful APIs.
- Directed Angular UI development, incorporating features like language flexibility and dynamic functionality to enhance user experience and accessibility.

## Skills:

- **Programming Languages:** Python (Pandas, NumPy), R, SQL, Scala.
- **Machine Learning & Statistical Analysis:** TensorFlow, scikit-learn, PyTorch, Keras, Random Forests, SVM, Gradient Boosting, Time Series Analysis, Hypothesis Testing.
- **Data Management & Processing:** Hadoop, Spark, Kafka, MongoDB, Redshift, S3.
- **Data Visualization & BI Tools:** Tableau, Power BI, Excel, ggplot2, Matplotlib, Plotly.
- **Cloud Platforms:** AWS (S3, Glue, EMR, SageMaker), Azure (Data Factory, Machine Learning).
- **Containerization & Orchestration:** Docker, Kubernetes.

## Certificates:

AWS Cloud Practitioner.	English: C1
Microsoft Certified: Azure Developer Associate.	German: B1

## Projects:

Testing Identity theft:

- Tools: Transfer Learning, Record Linkage, Differential Privacy, ARX Data Anonymization.
- Ethical Record Linkage Attack.

Reinforcement Learning-Based Breakout Game

- Tools: Tuning rewards and penalty, Pygame for Game design, Evaluated over thousands of iterations.

Sentiment Analysis based on Twitter feeds:

- Tools: NLP Models, BERT, GPT, Tableau, NLTK.

Influenza Prediction:

- Tools: Pandas, numpy, Scikit-learn, Matplotlib, Seaborn.