**Vanathi Sakthivelu**

[vanusakthi@gmail.com](mailto:vanusakthi@gmail.com)

07985373168

# Professional Profile

I am recognized for my strong work ethic, reliability, and a relentless passion for acquiring new knowledge. Whether collaborating within a team or working independently, I thrive on adapting to new challenges. Transitioning to Data Engineering in the IT industry excites me. I seek a role that challenges me to grow, expand my skills, and excel in this dynamic field.

# Education

* MSc (Computer Science) 2013-14

Coventry University, Coventry, UK

# Key Skills

# SQL

# Pyspark

# Azure Data Factory, Azure Databricks, Azure Synapse, Data Lake

# Python

* CI/CD pipelines
* Microsoft Power BI

# Projects

**Azure Data Engineering Project**

Steps:

* Created azure blob storage
* Designed and implemented an Azure Data Factory pipeline for seamless data transfer. Successfully transformed JSON data from Azure Blob Storage to an Azure Blob Container, converting it into CSV format for further processing.
* Integrated Azure Databricks with Azure Blob Storage by mounting it, ensuring easy access to data.
* Utilized PySpark within Databricks to conduct data analysis, uncovering valuable insights and patterns
* Leveraged Power BI to create engaging and informative data visualizations. Transformed processed data into meaningful reports and dashboards, making it accessible and actionable for stakeholders.

**Skills: Azure Databricks · Azure Data Factory · PySpark · Microsoft Power BI**

**Analyse Hospital Dataset (MIMIC) on Azure Databricks**

Steps:

* Played a key role in a data engineering project focused on healthcare data, utilizing Spark and Azure Databricks. Collaborated with the team to process and analyze hospital datasets, honing my data engineering skills.
* Downloaded complex hospital datasets from the MIMIC website and comprehensively understood the dataset's structure and challenges. Designed and implemented data ingestion, cleansing, and transformation workflows, preparing data for analysis on Azure Databricks.
* A subscription is set up for using Microsoft Azure and categorisation of resources is done into a resource group.
* Set up a secure Azure storage account to store all the essential data for our hospital dataset analysis.
* Creating and managing containers and uploading the dataset were a crucial part of ensuring data availability.
* Created Azure Data Factory to automate data movement and transformation.
* With a data copy pipeline, Moved data from Azure Storage to Azure Data Lake Storage and transformed it into formats that worked best for analysis.
* Performed partitioning, repartitioning and coalesce on the dataset in Databricks. Finally performed data analysis on the repartitioned dataset.

**Skills: Python3, Azure Data factory, Azure Databricks, ADLS**

# Work Summary

# Java Developer, GE Solutions, England [Jan 2021-Oct 2023]

* The project involved the development of an API Integration Platform that connects two Enterprise Applications and do data exchange between them by using webhooks and Polling Technique and to automate the data by every second.
* Contributed to all stages of software development lifecycle. Design, implement, test and maintain Java-based applications that can be high-volume and low-latency.

# Information Analyst , Palace Surgery (NHS primary care), Fulham,England [Jan 2014 – Dec 2020]

* Extract, analyse the **Systmone Searches data** to monitor high risk medicine patients to ensure 100% CQC compliance.
* Analyse and report the systmone data to improve the Cancer two week wait referral process.
* Create and Modify referral templates to extract essential data and update read codes for referrals on Systmone to monitor referral process and data return for GPs.
* Manage and monitor QOF updates on Systmone based on Practice Manager’s feedback to achieve the target every year.
* Update discharge summaries, test results from secondary care and specialist providers into the clinical system based on GP advice and ensure accuracy on patient profile.
* Maintain long term disease register data and facilitate CCG Locality meetings where areas of improvement and practice income were identified and agreed.
* Created Pop up **Dashboard** on new patient health check to find out any missing field to improve fast track payment.

**Public Health Analyst (Voluntary Placement), London Borough of Lewisham [Mar 2015-Apr 2015]**

* Helped in the design and development of GP Neighbourhood Profiles.
* Analysed PBR SUS data to find areas for emergency admission reduction
* Assisted in preparing core public health datasets from various national and local data sources for the profiles.
* Applied Advanced Statistical techniques (DSRs, ISR, CIs) and charts to summarise key findings on population health.
* Developed good understanding of JSNA process, Public Health Outcomes Framework and Performance monitoring dashboard and its role in health and social care integration.

**References**

Available upon request