# Hiten Pragji

**London — 07450 217119 — hkpragji@gmail.com**

## Summary

Highly numerate statistical data analyst and physics graduate currently working as a data specialist. Confident using SQL, R, and Power BI; interested in applying mathematical methods and analytics and visualisation tools to drive successful data-driven decision making. Looking to apply and further develop my knowledge and skills in analytics and data science.

## Skills

- $\geq$ 5 years: Advanced Mathematics, Statistics, Excel, LaTeX.
- $\leq$ 2 years: SQL, R, Power BI.

## Training

- Advanced Statistics for Data Science (Coursera) - In progress
- Data Analysis with R Programming (Coursera) - July 2023
- Intermediate SQL for Data Scientists (LinkedIn Learning) - March 2023
- Querying Microsoft SQL Server 2022 (LinkedIn Learning) - Feb. 2023
- Microsoft SQL Server 2022 Essential Training (LinkedIn Learning) - Dec. 2022
- Power BI Essential Training (LinkedIn Learning) - Nov. 2022

## Education

**MSc Medical Physics - Distinction - University of Surrey, 2019-21**

- Distinction (82%) in the Research Project and Dissertation, "Understanding SARS-CoV-2 and the impacts of associated variants of concern on the COVID-19 pandemic". Awarded the 2021/2 IPEM Student Prize for outstanding performance in the MSc Medical Physics programme dissertation.

**BSc (Hons) Theoretical Physics - 2:2 - Queen Mary University of London, 2014-18**

- Distinction (72%) in Extended Independent Project, "An investigation into the spectral separation of semiconductor quantum dots for super-resolution microscopy."

## Experience

**Data Specialist - UK Health Security Agency, Apr. 2023 - Present**

- Led the extract-transform-load (ETL) process for the development of Power BI reports; worked with the senior software developer to write advanced SQL scripts to migrate data between databases.
- Transformed data in Power BI, derived statistical metrics, created advanced visualisations in R using the tidyverse package, and designed data tables for reports using DAX.
- Prepared scientific reports and analysed them in Excel and R; communicated key findings, including statistical inaccuracies, to ensure scientific validity and completeness of data.
- Developed R code to perform statistical hypothesis testing (specifically, the Kolmogorov-Smirnov test) on univariate datasets using various statistical packages.
- Automated several administrative processes such as splitting, merging, and bulk renaming files using file manipulation packages in R.

**Data Quality Analyst - IQVIA, Nov. 2022 - Mar. 2023**

- Reinforced and built on the existing data quality framework by supporting the automation of the quality assurance process for management information (MI) deliverables (Power BI dashboards and reports); this involved writing SQL queries and collaborating with software developers to integrate SQL Server data into Power BI to track fluctuations in key metrics.
- Supported agile methodology by pro-actively participating in the work-stream planning process which included daily scrum calls with the MI team, following each task through its life cycle, developing strong connections with colleagues.

**Data Officer - NHS Digital, Jun. 2022 - Sep. 2022**

- Managed the extract-transform-load (ETL) process for five national cancer datasets, ensuring the availability of high-quality (clean and complete) data for disease registration, in line with standard operating procedures and strict deadlines.
- Performed quality assurance and validation for five national cancer datasets; communicated with internal stakeholders and NHS Trusts to resolve data quality issues. Ensured data was received in agreed formats, in line with service level agreements.
- Analysed data transformation mappings to diagnose and solve compatibility issues between the source data and the data warehouse, ensuring that different data formats were integrated correctly into the ETL system.

**Data Administrator - UK Health Security Agency, Feb. 2019 - May 2022**

- Increased the effectiveness of the COVID-19 pandemic response and accelerated turnaround times for test reports by assisting in testing and deploying a VBA application to import COVID-19 patient data and diagnostic data into the laboratory information management system (LIMS).
- Efficiently and accurately updated the LIMS and resolved data validation issues through effective communication with internal and external stakeholders.