

# Yash Tandel MSc - Data Engineer

[yashtandel.11@outlook.com](mailto:yashtandel.11@outlook.com) | [LinkedIn](#)

Building data pipelines and providing data solutions are a great passion of mine. Data Engineer with over 2 years of experience delivering high throughput projects in Data space, very keen to work in the sports, automobile, or healthcare sectors.

## Professional Experience

**Jr Data Engineer (November 2020 - June 2022) - Infosys.Ltd, Pune**

***Operated the hadoop tech stack to contribute to a Big Data project, facilitating rapid data analysis and delivering effective solutions.***

- **Main Contribution/Achievement:** Collaborated with a 3 member team providing consultation and client support in **Big Data analytics** to British Telecom.
- Designed HiveQL scripts to extract the data from staging tables and executed impala queries for data analysis to meet the business requirement and speed up the data analysis process by 15%.
- Written spark transformations and actions for the required input data and performed the data transformations using spark core.
- Good understanding of Apache Spark api and other Big Data tools.
- Performed the loading of data from different sources through SQOOP and storing the data in HDFS and hive from existing SQL server.
- Regularly analysed all the workflows and coordinators and was involved in daily processing activity.
- Managed the container by performing housekeeping activity to remove the old data and improve storage management for the new incoming data.
- **Skills & Stack:** HDFS, Apache sqoop, Apache hive, Impala, SQL, Linux, Java, pySpark, Apache oozie, GCP.

**MSc, Data Science and Analytics (2022 - 2023) - University of Westminster, London**

In a highly diverse environment with international students and a good learning space, I graduated with a First Class Master's Degree in Data Science and Analytics.

- **Modules undertaken:** Data mining and machine learning, Big Data theory and practise, Web and social media analytics, Data visualisation and dashboarding, Data warehousing and business intelligence, Business analytics.
- As part of dissertation, Implemented a real time data pipeline and conducted sentiment analysis on the streaming data.
- **Tools:** Databricks, Python, Spark, Apache Kafka, google cloud platform.

## Projects:

Dashboard on road casualties in the UK.

- Developed a Powerbi dashboard, producing creative visuals to draw valuable insights on different features based on the collected data and further providing business solutions.
- **Tools:** Powerbi.

## Yash Tandel MSc - Data Engineer

[yashtandel.11@outlook.com](mailto:yashtandel.11@outlook.com) | [LinkedIn](#)

Connected vehicles.

- Designed a Azure ELT pipeline to extract the data from AWS S3 bucket, validate the JSON data and store it in SQL database which can be used by the data science team in the future.
- **Tools and technologies: Azure, Data Factory, S3 bucket, Azure function, Azure SQL server.**

### Technical Skills

#### Programming Languages

Java, Python, SQL, Scala(basic level).

#### Big Data Technologies

Apache hive, Apache Sqoop, Apache Spark, Hadoop, Apache kafka, Databricks.

#### Data warehouse and cloud platforms:

Azure, GCP.