

YASH PANDYA

☎ 07563041197 ✉ yashpandya0411@gmail.com [in linkedin.com/in/yash-pandya-data](https://www.linkedin.com/in/yash-pandya-data) github.com/yash041197

Summary

- Data Scientist with 5+ years of experience in developing and iteratively improving research datasets and machine learning models. Skilled in statistical analysis, data transformation, and cross-team collaboration to enhance data quality and meet rigorous program standards. Proven track record of delivering data solutions with business and research impact on cloud platforms such as Azure Databricks.

Technical Skills

Programming Languages: Python, SQL, Java, JavaScript, MATLAB, Scala

CRM & Platforms: SFDC, ServiceNow

Data Science & Machine Learning: Pandas, NumPy, Scikit-learn, PySpark,

Data Visualization: PowerBI, Tableau, QlikSense

Data Analysis: Data Validation, Data Cleaning, Data Preparation, Data Transformation

Other Tools: Jupyter Notebooks, Git

Automation & Testing: Selenium, Git

Methodologies: ITIL (Information Technology Infrastructure Library), Agile, Scrum

Soft Skills: Research, Communication, Attention to Detail, Report Writing

Professional Experience

Acorn Fertility

March 2023 – Present

IT Technical Consultant (Contract Work)

London, England

- Conducted data testing, analysis, and evaluation to meet evolving program requirements, supporting insights for improved decision-making.
- Helped the team during Sprints by ensuring tasks were completed on time and aligned with user stories.
- Processed large datasets to create reliable insights, aligned with stakeholder needs and quality benchmarks.
- Developed automation pipelines for data processing, improving operational efficiency by 25%.

The Hospital Fertility Group (iTrust Fertility)

Sep 2021 – Oct 2022

IT Data Analyst

London, England

- Conducted advanced statistical analysis on healthcare datasets using Python and SQL to improve business strategies.
- Created standardized reporting frameworks and dashboards in Power BI, enhancing decision-making processes.
- Led analysis and validation of patient datasets, focusing on data quality and iterative improvement for enhanced treatment predictions.
- Led cross-team collaborations to deliver data-driven insights, improving process efficiency by 20%.

Tata Consultancy Service

Jan 2019 – Aug 2021

Developer

Gujrat, India

- Designed and optimized predictive models to meet business impact objectives, improving data reliability and achieving a 25% increase in pricing accuracy.
- Led cloud migration projects, utilizing Azure and AWS to enhance application responsiveness and data accessibility for cross-functional teams.
- Automated testing processes using Selenium and ensured data validation post-deployment.
- Implemented and maintained data workflows on Databricks, focusing on data engineering and quality checks to align with program standards.
- Automated data pipelines using PySpark on Azure Databricks, improving data processing speed by 30%.
- Authored research reports on data findings and presented insights to stakeholders, demonstrating strong communication skills.
- Demonstrated regular use of version control systems such as Git to manage code and ensure consistency across projects.

Education

Bournemouth University

Sep. 2021 – Jan 2023

Masters of Data Science in Artificial Intelligence

Bournemouth, England

Ganpat University

Sep. 2015 – Sep. 2018

Bachelor of Technology in Computer Engineering

Gujrat, India

Certifications

Dataiku Advanced Designer Certified	Mar 2024
Dataiku Developer Certified	Mar 2024
Dataiku ML Practitioner Certified	Mar 2024
Microsoft Certified Data Science Associate	Dec 2023
Automation Anywhere Certified Advanced RPA Professional	Dec 2021

Projects

Automated Web Scraping with Python and Selenium <i>Python, Selenium, Pandas, NumPy</i>	January 2020
<ul style="list-style-type: none">Developed fully functional automatic bots using Python and Selenium to perform download operations from websites, streamlining data retrieval processes and enhancing efficiency.Developed a Python script to retrieve files from an S3 bucket, perform data manipulations, and upload the processed data back to the bucket, streamlining data processing workflows.Implemented error handling mechanisms and robust logging to ensure the stability and reliability of the bots, enabling continuous operation without human intervention.Utilized Python's flexibility and Selenium's powerful web automation capabilities to customize bot behaviour according to specific website structures and requirements, ensuring adaptability across various use cases.Designed and implemented an automated deployment pipeline using Python and AWS services, reducing manual intervention and improving deployment efficiency by 50%.	
Cloud Migration to AWS <i>Python, AWS Lambda, FastAPI, Gateway, Cloudwatch, CI/CD, GitHub, SFDC</i>	March 2021
<ul style="list-style-type: none">Implemented cloud-based solutions to optimize application performance, enhancing responsiveness and ensuring a robust user experience.Utilized AWS API Gateway to create, publish, and maintain secure APIs, enabling seamless communication between various components of the application architecture.Engineered and deployed Lambda functions to automate specific tasks, integrating CI/CD pipelines for code deployment across development, testing, and production environments. This streamlined critical processes within the application lifecycle.	
Data Pipeline Optimization with Azure Databricks <i>Python, Azure</i>	January 2021
<ul style="list-style-type: none">Developed and optimized data processing pipelines on Azure Databricks, ensuring adherence to quality standards for real-time reporting and data transformation processes.Documented and tracked all test results, issues, and resolutions using Jira to ensure smooth communication between business teams and developers.	
Sustainability Reporting for Healthcare <i>Python, Excel, Power BI</i>	May 2023
<ul style="list-style-type: none">Analyzed sustainability metrics within healthcare data and visualized the findings using Power BI. Automated data cleaning processes using Python.	
Statistical Analysis of Fertility Data <i>Python</i>	January 2022
<ul style="list-style-type: none">Applied statistical programming techniques using R and Python to analyze fertility treatment outcomes, enabling data-driven decisions in patient care.Performed comprehensive statistical analysis on fertility treatment datasets, focusing on data quality and standardization to support research-informed decisions.	
Embryo's Assesment <i>Python, TensorFlow, Neural Networks, PyTorch</i>	May 2022
<ul style="list-style-type: none">Developed an algorithm using Python and AI to Assess the time-lapse images of an embryo's life cycle.Implemented advanced neural network models such as VGG and Inception as part of the Convolutional Neural Network (CNN) architecture, contributing to accurate embryo classification.Provided coaching and training to team members on data quality standards and best practices in data science, supporting skill development aligned with program goals.	

Extracurricular(Experience During Studies)

Green Earth Initiative

Dec 2023

Volunteer Data Analyst

London, England

- Contributed to data analysis projects focused on climate change and environmental sustainability. Provided insights to help guide organizational strategies.

Showcase Group

Dec 2023 – Present

Event Assistant

London, England

- Lead and inspire temporary event staff, fostering a team environment that strives to provide exceptional experiences for both clients and customers.
- Support the execution of events primarily from our flagship store on Regent Street, as well as manage operations at various pop-up locations across the UK and internationally.
- Strong organisational skills and the ability to manage multiple tasks simultaneously in a dynamic, fast-paced environment.

Dishoom

July 2023 – Dec 2023

FrontOffice

Canary Wharf, London

- Collaborated seamlessly within cross-functional teams, ensuring smooth operations and optimal guest experiences.
- Maintained high standards of cleanliness and attention to detail in a fast-paced environment.
- Facilitated smooth transition during team handovers from AM shift to PM shift. Effectively communicate key information, ensuring a seamless transfer of responsibilities.

McDonald's

Apr 2022 – Aug 2023

Crew Member

London, England

- Provided exceptional customer service by ensuring accurate and efficient order taking, addressing customer inquiries, and resolving issues promptly.
- Collaborated with team members to ensure smooth operations during high-traffic periods, demonstrating adaptability and a strong work ethic.
- Demonstrated flexibility by readily taking on additional responsibilities as needed, showcasing adaptability and a proactive approach to teamwork.