

# SHASHANK PANDIT

Data Analyst | Data Scientist in Transition

+91-9137285727 | shashank.pandit774@gmail.com | [GitHub] | [LinkedIn] | [Portfolio]

## Profile Summary

Results-focused professional with over 7 years of experience in the Information Technology industry,

including significant expertise in data analysis, ETL processes, automation frameworks, and QA testing.

Skilled in Python, SQL, PySpark, and Selenium, with a strong focus on transitioning into data science roles.

Experience includes automating web application tests, building and enhancing automation frameworks,

and performing data analysis and visualization using Power BI.

## Key Skills

- Programming Languages: Python, Java, SQL
- Data Science Tools: Pandas, PySpark, SQLAlchemy, Power BI
- Automation Tools: Selenium Web Driver, TestNG, Rest Assured
- Data Visualization: Power BI, Matplotlib
- ETL Processes: Data Extraction, Transformation, and Loading (ETL)
- Frameworks & Concepts: BDD, Page Object Model, OOP, Data-Driven Framework
- Version Control & Collaboration: Git, Jenkins, Jira, Confluence
- Data Analysis: Python, SQL, PySpark, Power BI
- API Testing: REST API (Postman, Rest Assured)

- Cloud Platforms: Azure SQL

## Experience

QA Automation Engineer | Data Science Transition

Capgemini (CTE) | Client: Barclays | Pune, IN

03/2024 - 07/2024

- Automated daily tests for 8 eLearning courses using Selenium Web Driver, achieving a 20% increase in user engagement through issue resolution.
- Analyzed API logs to identify quiz attempts, extracting data and improving quiz completion metrics.
- Conducted ETL processes using Python and SQL for accurate data manipulation and loading.

Automation Test Engineer | Wipro | Client: Bandhan Bank | Kolkata, IN

01/2023 - 06/2023

- Developed automated test scripts using Selenium Web Driver and Java, reducing regression testing time by 80%.
- Conducted ETL testing on banking data, validating transformation rules and ensuring accurate data loads into target databases.
- Expanded test coverage by writing test cases for various banking modules and increasing automation for different functionalities.

Data Analyst (Internship) | NTT Data | Client: NSE | Remote

11/2020 - 12/2022

- Extracted and transformed large-scale datasets using SQL and Python, ensuring data accuracy for business reporting.
- Conducted manual ETL testing for financial data, validating data integrity at each stage of the pipeline.

- Automated test scripts using Selenium, contributing to a 30% reduction in testing time.

## **Data Science Projects**

### Airline Data ETL and Analysis Project

Skills: Python, Selenium, Pandas, PySpark, SQLAlchemy, Azure SQL

- Automated the extraction of airline data using Python and Selenium, cleaning and transforming the data for structured analysis.
- Loaded data into Azure SQL and conducted data analysis with PySpark to generate insights related to airline operations.
- Improved ETL processing time by 40%, providing key metrics for airline performance.

### Netflix Dataset Analysis

Skills: Python, Pandas, SQL, Power BI

- Conducted end-to-end ETL processes to clean and transform Netflix data, integrating results into Power BI dashboards.
- Created interactive visualizations and generated actionable insights on user preferences and content trends.

### Payroll Management System

Skills: Python, OOP, SQL

- Developed a payroll management system utilizing OOP concepts, optimizing inheritance, polymorphism, and modular architecture.
- Enhanced system efficiency, ensuring modular code readability and seamless data processing.

## **Education**

Master of Computer Applications (MCA)

Punjab Technical University

Bachelor of Computer Applications (BCA)

CCS University

### **Certifications**

- Data Science Certification - [Institute/Platform]
- Selenium Automation Testing Certification - [Institute/Platform]
- Power BI for Data Analytics - [Institute/Platform]