

Navya A

Navyamanvi99@gmail.com | +91 8951362595 | Bengaluru, IN

CAREER OBJECTIVE

Detail-oriented and results-driven Data Engineer and Python Developer with 6+ years of experience in designing, building, and optimizing scalable data pipelines and cloud-based data solutions. Seeking to contribute my expertise in Apache Kafka, Spark, Airflow, AWS, Snowflake, and Python to a forward thinking organization focused on data-driven innovation and operational efficiency.

PROFESSIONAL EXPERIENCE

- Experienced Data Engineer and Python Developer with expertise in building real-time data pipelines and scalable data processing systems.
- Proficient in streaming technologies like Apache Kafka and Apache Spark for handling large-scale, real-time data ingestion and processing.
- Skilled in designing and implementing ETL workflows using Airflow to automate complex data pipelines and ensure data quality.
- Strong knowledge of cloud data platforms including Snowflake, AWS S3, Redshift, and AWS Lambda for storage, analytics, and serverless computing.
- Developed RESTful APIs using Django and Flask to expose data services and integrate with internal applications.
- Expertise in data manipulation, validation, and transformation using Python libraries such as Pandas and NumPy.
- Hands-on experience with containerization using Docker and deployment automation through CI/CD pipelines with Git, Bitbucket, and Jenkins.
- Proficient in SQL and PostgreSQL for both transactional and analytical database management.
- Built and maintained interactive dashboards and visualizations using BI tools including Tableau and Power BI to support data-driven decisions.
- Familiar with Agile methodologies and tools such as Jira and Notion for project management, collaboration, and sprint planning.
- Adept at writing unit tests and maintaining code quality to ensure robust and maintainable software solutions.

WORK EXPERIENCE

- Worked as **Assistant Manager** at **Britannia**, Bengaluru | December 2024 - Present
- Worked as **Senior Data Analyst** at **Kinesso India Pvt Ltd**, Bengaluru | November 2021 - December 2024
- Worked as **Data Analyst** at **PayPal**, Bengaluru | October 2020 - October 2021
- Worked as **Data Consultant** at **L& T**, Bengaluru | October 2019 - June 2020
- Worked as **Associate Data Analyst** at **Photon**, Bengaluru | April 2017 - September 2019

TECHNICAL SKILLS

- Languages & Tools : **Python(Pandas, Numpy, Scikit-learn), SQL, Bash, DAX, Git, Jenkins, Bitbucket, Docker, Jira, Notion, Power Apps**
- Databases : **PostgreSQL, MySQL, MongoDB, Redshift, Snowflake**
- Frameworks : **Django, Flask, Apache Spark, Apache Kafka, Airflow**
- Cloud Platforms : **AWS (EC2, S3, Lambda, IAM, Redshift, Snowflake), Azure**
- Machine Learning : **Regression, p-value, Hypothesis Testing, Time Series Analysis**
- IDEs : **Visual Studio Code, PyCharm, Jupyter Notebook**
- Operating Systems : **Windows, Linux**

EDUCATION

- **Bachelor of Engineering (BE)** | Visvesvaraya Technological University, Belgaum (SIT)

KEY PROJECTS

Project 4: Sales and Demand Forecasting for Food Outlets

Technologies: Python, Pandas, NumPy, Apache Spark, Kafka, PostgreSQL, AWS S3, Airflow, Django, Git, Jira, Bitbucket, CI/CD, Power BI, Docker, Linux.

Description:

Developed a sales and demand forecasting system to optimize inventory and reduce wastage for multiple food outlets. Collected and processed historical sales and customer data using Kafka and Apache Spark. Cleaned and transformed data using Pandas and NumPy, storing the results in PostgreSQL for fast querying. Automated ETL pipelines with Airflow and built Django APIs to provide forecasting results to internal dashboards. Created interactive Power BI reports to visualize sales trends and demand predictions. Managed deployments using Docker and ensured code quality with CI/CD pipelines. Collaborated with cross-functional teams in Agile sprints using Jira.

Role and Responsibilities:

- Ingested and processed large volumes of sales and customer data via Kafka and Spark.
- Performed data cleaning, transformation, and feature engineering with Pandas and NumPy.
- Stored and managed transactional and forecast data in PostgreSQL.
- Automated data workflows and pipeline scheduling using Airflow.
- Developed REST APIs with Django to expose forecasting data to dashboards.
- Created Power BI dashboards to visualize demand trends and support decision-making.
- Containerized applications using Docker for scalable deployments.
- Maintained source code and CI/CD pipelines using Git, Bitbucket, and Jira.
- Participated in Agile ceremonies for planning and progress tracking.

Project 3: Churn Prediction Engine for Telecom Customers

Technologies: Python, Pandas, NumPy, Apache Spark, Kafka, Snowflake, AWS S3, Airflow, SQL, Django, Git, Jira, Bitbucket, CI/CD, Power BI, Notion, Power Apps, Linux.

Description:

Developed a churn prediction engine to identify telecom customers likely to leave the service. Collected customer usage data via Kafka and processed it using Spark for feature extraction. Cleaned and stored processed data in Snowflake, with raw logs archived in AWS S3. Created Django APIs to serve churn scores to CRM systems. Automated data pipelines using Airflow and managed documentation and collaboration via Notion and Power Apps. Built interactive dashboards using Power BI to visualize churn trends and support business decisions. Managed code and deployments with Git, Bitbucket, and CI/CD pipelines. Worked in Agile sprints coordinated through Jira.

Role and Responsibilities:

- Ingested and processed large volumes of customer activity data using Kafka and Spark.
- Engineered predictive features like usage frequency, complaints, and recharge patterns with Python.
- Loaded clean data into Snowflake for analysis and archived raw data in AWS S3.
- Developed REST APIs using Django to expose churn prediction results to internal teams.
- Automated ETL workflows and data validation with Airflow.
- Created Power BI dashboards for visualization and reporting of churn insights.
- Used Notion and Power Apps for project documentation and cross-team collaboration.
- Maintained version control and continuous integration through Git, Bitbucket, and CI/CD.
- Participated in Agile ceremonies with Jira for sprint planning and issue tracking.

Project 2: Secure Payment Gateway Integration

Technologies: Python, Pandas, NumPy, Apache Spark, Kafka, Snowflake, AWS S3, Airflow, SQL, Django, Git, Jira, Bitbucket, CI/CD, Power BI, Notion, Power Apps, Linux

Description:

Developed a secure payment gateway integration to process online transactions with high reliability and security. Built RESTful APIs using Django and utilized AWS Lambda for serverless transaction processing. Used Pandas and NumPy for data validation and transformation tasks. Stored transactional data in PostgreSQL for fast retrieval and Redshift for analytics and reporting. Managed raw data and backups in AWS S3. Scheduled cron jobs for automated reconciliation and settlement. Ensured robust code quality with unit testing and deployed applications via Docker containers. Secured API endpoints using AWS API Gateway. Maintained source control and automated deployments with Git, Bitbucket, and CI/CD in an Agile setup.

Role and Responsibilities:

- Developed Django APIs for secure payment processing and transaction management.
- Implemented AWS Lambda functions for efficient serverless workflows.
- Used Pandas and NumPy for data cleaning, validation, and transformation.

- Managed transactional data storage in PostgreSQL and analytical data in Redshift.
- Scheduled and maintained cron jobs for automated batch reconciliation.
- Created and maintained unit tests to ensure high code quality.
- Containerized applications using Docker for consistent deployments.
- Handled code versioning and CI/CD pipeline using Git, Bitbucket, and Jira.
- Collaborated with Agile teams for sprint planning and issue tracking.

Project 1: Real-Time Credit Card Transaction Monitoring System

Technologies: Python, Apache Kafka, Apache Spark (Streaming), Snowflake, AWS S3, Airflow, SQL, Flask, Git, Jira, Bitbucket, CI/CD, Tableau, Linux.

Description:

Designed and implemented a real-time transaction monitoring system for detecting suspicious credit card activity. Integrated Kafka for live data streaming and Spark for rule-based processing. Stored processed data in Snowflake and archived raw data in AWS S3. Developed Flask APIs to expose flagged transactions. Automated ETL workflows using Airflow. Used Tableau to create fraud analytics dashboards. Maintained deployment pipelines using Git, Bitbucket, and CI/CD tools. Worked in Agile sprints tracked via Jira.

Role and Responsibilities:

- Developed real-time data pipelines using Kafka and Spark Streaming.
- Ingested, processed, and transformed transaction data based on fraud rules.
- Stored and queried structured data in Snowflake, archived raw data in S3.
- Built Flask APIs to provide transaction insights to internal tools.
- Automated workflows with Airflow for scheduling and monitoring.
- Collaborated with analytics team to create Tableau dashboards.
- Managed code with Git, Bitbucket, and integrated CI/CD processes.
- Worked in Agile with Jira for sprint planning, tracking, and delivery.

DECLARATION

I hereby declare that the information provided above is true to the best of my knowledge and belief.

Navya A
