SATYA CHIPPADA

Data Engineer | Python Developer | SWE | Project Coordinator
ETL • APIs • Analytics • Automation • Project Management
USA: +1-435-512-4598 | India: +91-6303553458
satych90@gmail.com | www.linkedin.com/in/satyachippada

PROFESSIONAL SUMMARY

Results-driven Python Data Engineer & IT-PM with about 5 years of experience designing, building, and scaling data-intensive and backend systems across fintech, energy trading, healthcare, and cloud platforms. Proven expertise in developing production-grade Python applications, end-to-end ETL pipelines, and analytics platforms using Python, PySpark, SQL, and REST APIs. Strong background in automation, performance optimisation, data quality frameworks, and CI/CD-driven delivery. Adept at collaborating with cross-functional stakeholders to deliver measurable business impact in Agile environments.

PROFESSIONAL EXPERIENCE

Vitol Inc, Houston, TX
Python Data Engineer & Project Coordinator, MIS Team | Dec 2019 – Jul 2021
- Designed, developed, and maintained production-grade Python applications supporting energy trading analytics and operational dashboards.
- Built scalable ETL pipelines using Python, PySpark and SQL to ingest, cleanse, validate, and transform large volumes of trade, pricing, and market data from the Bloomberg terminal.
- Implemented data quality checks, exception handling, logging, and audit trails to ensure accuracy, traceability, and compliance of financial data.
- Developed backend services and internal APIs in Python to expose analytics to downstream systems and BI tools.
- Automated end-to-end reporting workflows using Python scripts and schedulers like Cron and Apache Airflow, reducing manual effort and improving report turnaround time by ~40%.
- Led web scraping and data ingestion initiatives using Requests and BeautifulSoup, migrating legacy data sources into centralised SQL databases.
- Optimised SQL queries, indexes, and data models to improve performance for high-frequency and ad-hoc trading reports.
- Collaborated closely with traders, analysts, and product owners to translate business requirements into reliable technical solutions.
- Followed Agile development practices, participating in sprint planning, code reviews, and retrospectives.

Tech Stack: Python (Pandas, NumPy, REST APIs(Django, Flask & FastAPI)), PySpark, Apache Airflow, SQL, Git, Linux, Agile/Scrum

Google LLC (via eTouch Systems)
Full-Stack Engineer, Machine Intelligence Team | Aug 2018 – Jan 2019
- Queried and analysed large datasets using Google BigQuery, enabling insights into course adoption, user behaviour, and content effectiveness.
- Loaded data from an internally implemented data labelling initiative into BigQuery for analysis and usage in ML training.
- Wrote optimised SQL queries in BigQuery and integrated results into Python applications for downstream analytics and reporting.
- Collaborated with ML engineers to create Google Colaboratory, a free, hosted Jupyter notebook environment that requires no setup and has TensorFlow pre-installed, demo environments showcasing NLU and NLP models, and improved adoption across Google 1P engineering teams.
- Used Bootstrap to create responsive web applications, worked on building UI features for Google Conceptual Captions, and Google Natural Questions UIs for managing GARCON(Generation, Abstractization, Reading Comprehension On Neural architectures) research competitions.
- Followed Agile development practices, participating in sprint planning, code reviews, and iterative releases aligned with Google engineering standards.

Tech Stack: Google Cloud Platform (GCP), BigQuery, Python, SQL, JavaScript, TypeScript, ML/NLP, Tensorflow, Git, Internal Google Tools, Agile, SDLC

Goldman Sachs, Salt Lake City, UT
Full-Stack Engineer, Cloud Platforms & Virtualisation Team | Mar 2016 – Jul 2018
- Developed Python-based automation tools for cloud capacity management internally for cost optimisation, and infrastructure monitoring across 5,000+ servers.
- Built data processing and analytics pipelines using Python (pandas, NumPy) and ElasticSearch to aggregate and analyse large-scale server logs to detect EOL machines and decommission them.
- Designed and consumed REST APIs to integrate infrastructure data with internal dashboards and compliance platforms.
- Created monitoring and compliance dashboards using ElasticSearch, Kibana, and custom Python scripts for servers.
- Implemented CI/CD pipelines using Jenkins and Git, accelerating the deployment of Python services and automation jobs.
- Performed advanced data analysis and visualisation using matplotlib and seaborn to deliver executive-ready performance reports.
- Partnered with infrastructure, security, and compliance teams to meet audit and regulatory standards.
- Used the Pyramid framework for authentication and authorisation using ACLs (Access Control Lists)

Summer Intern, Cloud Computing & Virtualisation | Jun 2015 - Aug 2015
- Contributed to the implementation of Dynamic Compute support into host trackers, enabling system migration to the cloud while facilitating the use of Python scripts and SALT configuration management.

Tech Stack: Python (Pandas, NumPy, MatplotLib, Seaborn, REST APIs, Pyramid Framework), ELK (ElasticSearch, Kibana, Logstash), MongoDB, Jenkins, Git, Linux, SALT Configuration Management, Cloud Compute, Internal tools.

Catholic Health System, Buffalo, NY

IT Intern – Health Systems Analytics & IT Project Management | Jun 2024 – Sep 2024
- Designed and implemented data analysis workflows to extract, cleanse, and transform Epic EHR data, enabling actionable insights into MyChart patient portal adoption and engagement trends.
- Analysed large-scale patient encounter and appointment datasets using SQL and Python, supporting executive-level operational and revenue cycle reporting.
- Automated recurring analytics tasks using Python scripts, reducing manual data preparation effort and improving reporting turnaround time.
- Created interactive dashboards using Epic SlicerDicer outputs + Pulse graph outputs, translating raw EHR data into clinician- and leadership-friendly insights.
- Supported EHR optimisation initiatives by validating data consistency across Epic patient data entry modules and identifying process inefficiencies impacting patient experience.

Tech Stack: EPIC, Pulse, SlicerDicer, Python, SQL, Linux, Agile.

University at Buffalo, SUNY, Buffalo, NY

Student Project Coordinator & Operations Assistant (Part-time) | Sept 2024 – May 2025
- Coordinated academic operations at Goodyear Student Academic Centre (GSAC); oversaw staff and ensured front-desk efficiency.
- Assisted in course management for MGS 607: Technology & Innovation Management.

Volunteer, Kanya's, Andhra Pradesh, India

Jan 2022-May 2023
- Upcycled waste materials from local businesses and turned them into decorative art

<u>EDUCATION</u>

University at Buffalo, SUNY (NY, USA), **MBA (STEM)**, 2025 | GPA: 3.7/4
Utah State University (UT, USA), **MS**, **Management Information Systems**, 2015 | GPA: 3.87/4
NITW (Warangal, India), **B.Tech, Electrical & Electronics Engineering**, 2014

SKILLS

Programming: Python (Advanced), SQL, PySpark, Apache Spark, R
AI workflows and agents: Zapier, n8n, Make.com (Self-Learning)
AWS: S3, EC2, IAM roles & polices, Lambda functions, RDS Aurora, Redshift (Self-Learning)
GCP: Google Big Query, Google Cloud Build, Google Cloud Run (Self-Learning)
AI & Automation: LLM APIs (OpenAI, Gemini), Agentic workflow prototyping, NLP fundamentals, LangChain basics, Prompt Engineering
Backend: REST APIs, Flask, Django, Pyramid, Microservices Concepts, API design, Fast API
Data Engineering: ETL, Data Validation, Logging, Monitoring, Exception Handling
Databases: PostgreSQL, SQL Server, MongoDB
DevOps: Jenkins, Git, CI/CD, Linux, Bash
BI & Analytics: Tableau, Power BI, MS Excel, GCP Data Studio
Enterprise Tools, Project Management: Jira, Confluence

CERTIFICATIONS

Google Cloud: Intro to ML Language Processing
Foundations: Data, ML & AI
Analytics: Tableau (Credly)
Strategyn: JTBD + ODI Fundamentals