

# Yaswanth Pati

+1-609-613-1289 | [yaswanth.pati01@gmail.com](mailto:yaswanth.pati01@gmail.com) | <https://linkedin.com/in/yaswanthpati/>  
<https://github.com/Yaswanth-pati>

## SUMMARY

Data Science professional with over 3 years of experience and M.S. in Data Science in statistical analysis, predictive modeling, and data visualization. Demonstrated success in enhancing operational efficiency and reducing processing times by implementing data-driven solutions using Python, SQL, and R. Skilled in translating business challenges into actionable insights while leveraging dashboard tools like Power BI and Tableau.

## EDUCATION

### Saint Peter's university

Feb 2023 - Feb 2025

M.S, Data Science

- GPA: 3.9/4.0

## PROFESSIONAL EXPERIENCE

### Elegantek

Jun 2024 - Dec 2024

Data science intern

MI, USA

- Developed predictive models using machine learning algorithms, achieving a 20% improvement in operational efficiency while formulating testable hypotheses for further analysis.
- Conducted comprehensive data preprocessing and feature engineering on large datasets, enhancing model accuracy by 15%.
- Applied NLP powered sentiment analysis on unstructured customer comments to extract actionable insights, resulting in a 25% increase in engagement.
- Created interactive dashboards using Power BI and Tableau to translate complex data patterns into clear visual insights for stakeholders.
- Deployed ML models with Flask and Azure to enable real-time analysis, effectively querying and manipulating structured data.

### Saint Peters University

Feb 2023 – Nov 2023

Graduate Research Assistant - Data Analyst

Jersey, USA

- Developed Conducted statistical analysis and data visualization for faculty-led research projects, utilizing Python, R, and SQL to uncover patterns and trends in over 250,000 rows of raw data.
- Designed and maintained 4 interactive dashboards using Tableau and Power BI to support university decision-making and research tracking.
- Developed and deployed machine learning models for projects focused on public health and behavioral analytics, improving prediction accuracy by over 20%.

### Infosys Limited

Jan 2021 - Jul 2022

Data Analyst

Hyderabad, India

- Cleaned and analyzed raw datasets using Python and SQL to identify key trends that informed strategic business initiatives and boosted operational efficiency by 30%.
- Streamlined data pipelines to optimize processing, reducing monthly processing time by 20 hours.
- Designed interactive, web-based dashboards integrating SQL databases with Tableau and Power BI to effectively visualize statistical findings and evaluate business hypotheses.
- Collaborated with cross-functional teams in an Agile environment to develop data-driven web solutions, accelerating project delivery by 30%.

## NOTABLE PROJECTS

### Multi-Model Text Summarization Using NLP Models Feb 2025

- Created an NLP-based text summarization system using BART,T5 and Pegasus models.
- Built a Flask API to process user input and generate summaries dynamically.
- Utilized similarity scores using TF-IDF, BERT Cosine and Word2Vec to assess the similarity of the summaries generated by the 3 models.
- Built an interactive UI with real-time text summarization and model comparison.

### Route Planning for sensitive populations Nov 2024

- Innovated a comprehensive mapping tool integrating live data on pollutants from roadways via Azure Maps API; directly contributed to reducing daily commute-related respiratory issues by informing users about cleaner travel options available.
- Developed a comprehensive route visualization system integrating air quality indicators and climatic factors; enabled for users to access critical environmental information directly from interactive tooltips.
- Engineered interactive features facilitating intuitive vocal inputs for geographic routes; streamlined data retrieval processes reduced average query handling time from two minutes to just under fifty seconds per request.

### **Wildfire Prediction Using ML Nov 2024**

- Devised an advanced predictive model utilizing brightness values from the MODIS and VIIRS satellite dataset, achieving a classification accuracy of over 95% through rigorous cross-validation against real-world wildfire conditions.
- Achieved high accuracy in intensity classification using a Random Forest Classifier, with results validated through cross validation. Established a Confidence Level Classification Model, achieving 97% accuracy, leveraging features like brightness, FRP, and track dimensions.

### **TECHNICAL SKILLS**

---

- **Programming Languages:** Python, R, SAS
- **Machine Learning:** Supervised and Unsupervised Learning, Regression analysis, Classification, Clustering, Feature selection, Ensemble methods
- **Natural Language Processing:** Tokenization, Word Embeddings, NER, Sentiment Analysis, Transformer Models, BERT
- **Data Analysis Tools:** Tableau, Power Bi, Excel, Azure, Microsoft Word, Microsoft Access, Microsoft PowerPoint
- **Big Data Tools:** Hadoop, Hive, Map reduce, Apache Spark
- **Tools:** Git, Angular framework, Node.js, Jupyter notebook, GitHub, Databricks
- **Work Methodologies:** Agile Environment
- **UI:** HTML, CSS, JavaScript, AngularJS, Unit testing
- **Soft Skills:** Analytical, Problem-solving, Critical Thinking, Decision Making, Communication Skills
- **Database Management Systems:** SQL, MySQL, PostgreSQL
- **Analytics & Industry Knowledge:** Statistical Methods, Data Analysis, Statistical Testing, Evaluation Designs, Data Science

### **CERTIFICATIONS**

---

**Generative AI :** LinkedIn (2025)

**Databricks Fundamentals :** Databricks (2025)

**Databricks Generative AI Fundamentals :** Databricks (2025)

**SQL certification:** Frontlines Edu Tech Private Limited (Apr 2024)

**Google Data Analytics:** Coursera (Jun 2022)