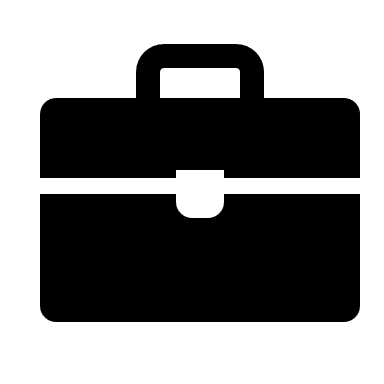
Puneeth Velivela

 (571)604-1146 |  Exp: 3 years | Fairfax, VA (Open to Relocate)

 [puneeth](mailto:puneeth)velivela07@gmail.com | A black background with a black square

AI-generated content may be incorrect. [www.linkedin.com/in/pvelivel](http://www.linkedin.com/in/pvelivel) |  <https://github.com/pvelivel> | A black background with a black square

AI-generated content may be incorrect. [Portfolio](https://pvelivel.github.io/)

# Professional summary

* Earned a Master’s in Data Analytics Engineering and mastered SQL, Python, and building Interactive Dashboards; presented insights to stakeholders, leading to improved data analysis.
* 3 years of hands-on experience transforming messy enterprise data into actionable insights at Infosys, building dashboards, streamlining data flows, and delivering reliable reporting across business units.
* Currently driving research in financial regime detection using unsupervised learning (VAE + GMM) under Prof. KC Chang; applying machine learning to real-world market data.

# PROFESSIONAL EXPERIENCE

**Graduate Research Assistant |** *Prof Kuo Chu Chang, George Mason University, Fairfax, VA* **Jun 2025 – Present**

* Enhanced financial market regime detection accuracy under Prof. KC Chang using unsupervised ML methods, incorporating VAE and GMM to refine the model's sensitivity to market volatility by 25%.
* Guided an asset allocation model using latent representations of S&P 500 signals, improving Sharpe ratio by 33.7%.

**Data Analyst Intern |** *Atlantic Flyway Council, Virginia* **Jan 2025 – May 2025**

* Built a scalable AWS Glue ETL pipeline that standardized 40+ years of bird colony data, enhancing data quality and reducing data processing time by 75% for stakeholders.
* Mapped population density of endangered bird species using geolocation data, revealing key habitats and informing conservation efforts across 15 states, resulting in higher quality data for stakeholders.
* Validated over 10,000 geolocation points and harmonized species names using AOU taxonomy.

**Data Analyst |** *Infosys Ltd, India* **Jan 2021 – Jun 2023**

* Designed and maintained SQL-based dashboards to track key business metrics and reporting KPIs.
* Prepared automated data flows using MuleSoft to transform XML and JSON data into structured formats.
* Developed Power BI reports for operations and finance teams, reducing manual reporting time by 50%.
* Uncovered a critical flaw in CAMT/XML message processing that caused duplicate transactions, and designed a validation system to flag similar errors, preventing financial loss.
* Collaborated with cross-functional teams to ensure accuracy and consistency across APIs and backend systems.

# EDUCATION

**Master of Science in Data Analytics Engineering |** George Mason University, Fairfax, VA | GPA: 4

**Bachelor of Science in Mechanical Engineering |** Koneru Lakshmaiah Educational Institution, India

# Academic Projects

**YouTube Trending Analytics |** [**Github Link**](https://github.com/pvelivel/aws-youtube-analytics-pipeline)

*PipelineAWS | Athena | QuickSight | Lambda | S3 | Parquet*

* Designed a serverless pipeline on AWS using Lambda and S3 to ingest YouTube trending data across 10 countries, enabling real-time analysis and identifying emerging trends.
* Processed raw JSON/CSV into Parquet using AWS Lambda, cataloged with Glue, and visualized via Athena and QuickSight**.**

**Diabetes Prediction Analysis |** [**Github Link**](https://github.com/pvelivel/diabetes-prediction-analysis)

*Python | Scikit-learn | Pandas | Logistic Regression | Clustering*

* Created ML models to predict diabetes risk based on clinical parameters.
* Applied logistic regression, decision trees, and K-means clustering to identify patient groups and key risk factors.

**Blinkit Sales Dashboard (Power BI) |** [**Github Link**](https://github.com/pvelivel/blinkit-sales-analysis-powerbi)

*Power BI | Excel | DAX | KPIs*

* Developed an interactive Power BI dashboard analyzing Blinkit's sales and customer satisfaction.
* Tracked KPIs like Total Sales, Avg. Rating, and Outlet Type to uncover insights and improve business strategy.

# SKILLS

* **Programming Languages:** Python, R, SQL, DataWeave
* **Data Visualization & BI Tools:** Tableau, Power BI, Python (Seaborn, Pandas, Matplotlib)
* **Databases:** MySQL, PostgreSQL, Relational Databases, Basic NoSQL (MongoDB)
* **Data Engineering & Integration:** MuleSoft Anypoint Studio, API Development, ETL Pipelines, Data Transformation, CI/CD Pipeline Automation, Batch Processing, Databricks (Notebooks, SQL Editor)
* **Machine Learning & Data Analysis:** Predictive Modelling, Logistic Regression, Decision Trees, Feature Engineering, Data Pre-processing, Data Warehousing Concepts, Report Optimization
* **Cloud & Monitoring Platforms:** MuleSoft CloudHub, Splunk, AWS (Basics: EC2, S3, Lambda), CI/CD Tools (Jenkins)
* **Version Control & Collaboration:** Git, GitHub, Jira, Postman.