**Supply Chain Optimization with Python & Data Science: Reducing Waste in the Food Industry**

**Project Summary**

As part of an internal analytics initiative at FreshRoute Logistics Ltd., a national food distribution company operating six warehouses across the United Kingdom, I conducted an in-depth analysis to identify operational inefficiencies in the supply chain. By simulating real-world data and applying statistical and predictive techniques, I identified key drivers of spoilage and exposed regional disparities. I proposed data-driven recommendations to reduce waste and improve delivery performance.

**Problem Statement**

FreshRoute Logistics Ltd. was experiencing higher-than-expected spoilage rates across several locations. While average delays were within acceptable thresholds, losing perishable inventory suggested hidden inefficiencies. This project sought to answer:

* How strongly are delivery delays contributing to spoilage?
* Which regions are underperforming, and why?
* What data-driven actions can reduce operational waste and improve service reliability?

**Methodology**

I used Python to simulate 300 delivery records of supply chain data consisting of delivery delays (days) and spoilage rate(%) for six FreshRoute Logistics Warehouse locations: Edinburgh, Manchester, Birmingham, Bristol, Leeds, and London. The product categories included fruits, vegetables, Dairy, and Meat.

**Tools and Technologies used**

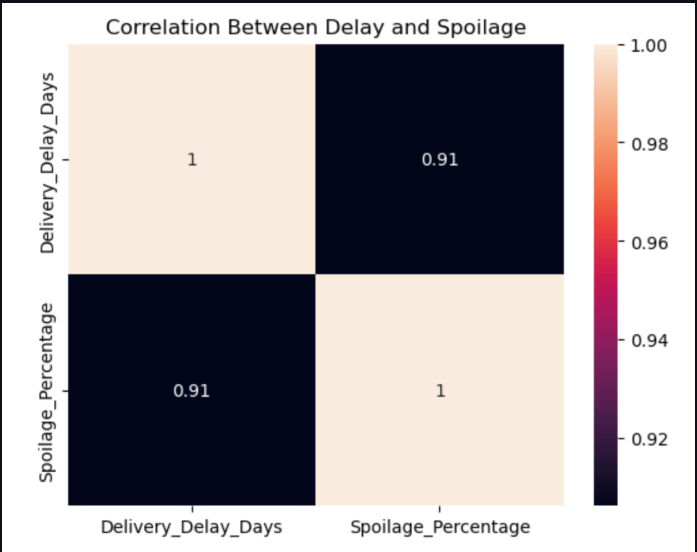
* Python (Pandas, NumPy, Seaborn, and Scikit-learn); Data Simulation & Analysis (Random, Matplotlib)
* Machine Learning (Linear Regression and Correlation Analysis)
* Visualisation (Plotly, matplotlib, and Seaborn)
* Notebook Environment: Jupyter

**Data analysis**

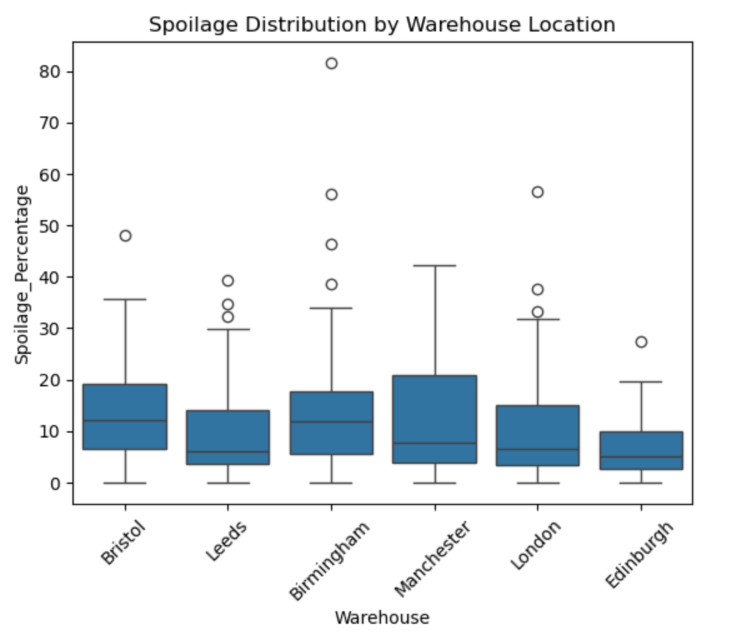
* Find the Relationship between delivery delay and spoilage percentage
* Find out the spoilage percentage per warehous

**Key Findings**

* There was a Positive correlation (~0.91) between delay and spoilage, which means that as the delivery delay increases, the spoilage also increases.



* Regional disparities existed:Edinburgh and London had significantly lower spoilage , while Birmingham and Bristol experienced higher spoilage despite similar delays.



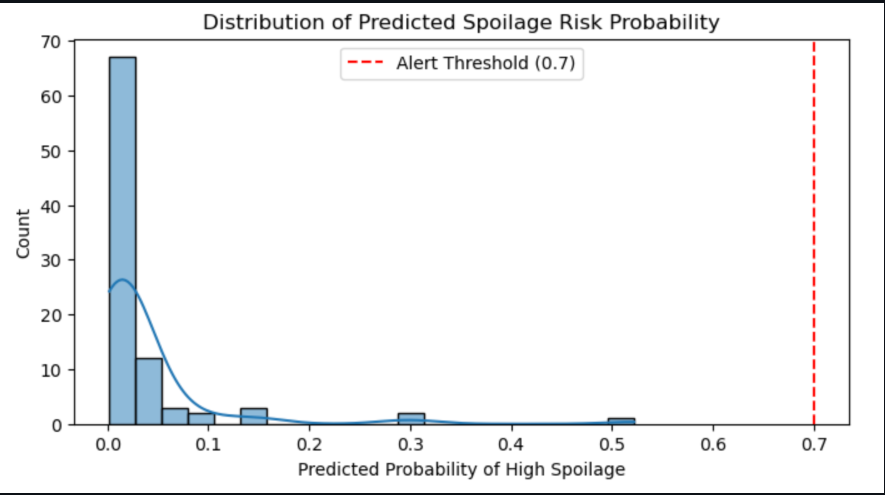
**Results and Recommendations**

**Short-term:**

* Edinburgh & London demonstrated best operational consistency; practices should be documented and scaled.
* Bristol & Birmingham require intervention: delivery scheduling improvements and cold chain enhancements.

**Mid-term:**

* Introduce a predictive model: Spoilage Risk Classifier.
* Goal: Predict whether a shipment is "High Risk" (likely to result in spoilage above a critical threshold, say 70%) before it arrives, using known pre-delivery metrics such as delay forecasts, region, and product type.



Importance: Allows operational teams to intervene proactively (reroute, repackage, prioritise delivery).

Example in use case:

The shipment of dairy to Birmingham was delayed by 3.2 days, and the expected arrival is on Friday.

The model predicts a 74% probability of high spoilage.

As a result:

* Alert sent to the cold storage team.
* The dispatch team reprioritises delivery or reroutes to the nearby warehouse.

**Next Steps**

(i) Incorporate external features: temperature, traffic, and weather forecasts to improve the model.

(ii) Calibrate the model per warehouse to allow location-specific tuning.

**Conclusion:**

This project gave me practical experience in using data analytics to solve real-world supply chain challenges. I learned how to simulate meaningful data, uncover relationships between delays and spoilage, and build a predictive model to flag high-risk shipments. One key challenge was balancing realistic data generation with model performance, but it taught me the importance of domain knowledge and clear communication. I’m excited to apply these skills to real business problems where data can drive efficiency and reduce waste.