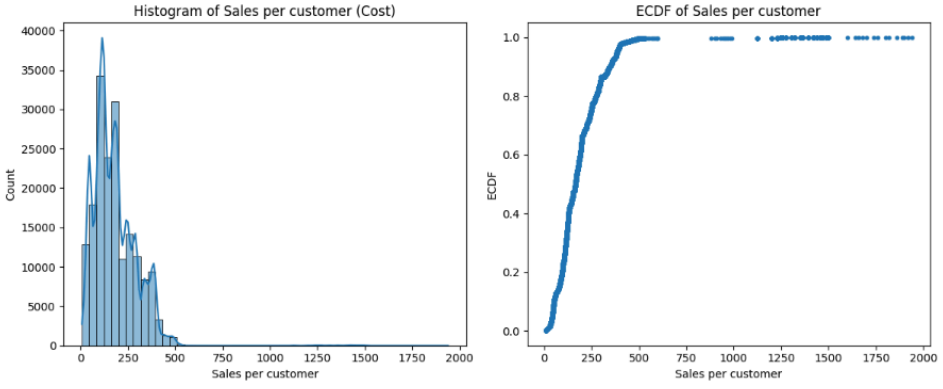
- **Project Title: AI-Driven Risk Management and Decision Support for Sustainable Supply Chains**
- **Project Start Date:** 1st June 2025
- **Names of project associates/ students (if student, mentions degree and year):**
    - Puneet, BS in Data Science, Final Year
    - Manish Kumawat, Intern
- **Problem Statement/ Description (3-4 lines):** Supply chains face unpredictable disruptions that can severely impact service levels and costs. Stochastic risk assessment enables FedEx SMART Centre to quantify these uncertainties using probabilistic models. This approach supports robust decision-making under uncertainty, aligning with FedEx's goal of building resilient, sustainable, and agile logistics solutions.

Progress against project objectives:

S. No	Objective	Progress Status
1.	Risk Assessment Framework	● Working on designing a stochastic risk assessment model integrating probability distributions of supply chain disruptions. ● Found a dataset and worked on the code to implement Monte Carlo simulations to evaluate potential outcomes of supply chain decisions under uncertainty.
2.	AI-Based Decision Support System	Yet to do
3.	Predictive Analytics for Disruptions	Yet to do
4.	Sustainable and Resilient Logistics Strategies	Yet to do

Work completed in the project thus far:

1. Initial Research and Planning
    a. Identified and documented publicly available datasets relevant to supply chain disruption modeling.
    b. Researched stochastic risk assessment models (Stochastic Programming, Markov Chains, Bayesian Networks, Newsvendor) with pros, cons, and application domains.
2. Paper Review and Concept Extraction
    a. Reviewed a key paper on "Stochastic Integrated Supplier Selection and Disruption Risk Assessment under Ripple Effect."
    b. Summarized its objective function, decision variables, constraints, and uncertainty modeling. Extracted and documented the Bayesian Network structure for risk propagation and ripple-effect modeling.
3. Literature Comparison
    a. Created a matrix comparing four major stochastic modeling approaches, highlighting application domains, pros, cons, and references.
4. Seminar Participation and Summaries
    a. Attended the FedEx/IIT Kharagpur seminar on Variance-Adjusted Cosine Similarity (VACS) and summarized key takeaways, including mathematical formulation and potential application to similarity/clustering tasks.
    b. Attended the SMART FedEx Seminar Series, summarizing insights on logistics challenges in India, global best practices, and the need for localized solutions.
5. Data Preparation (Phase 2)
    a. Reviewed the selected dataset (SupplyGraph or equivalent) provided by the team. Created a clear data dictionary mapping original fields to risk variables (e.g., frequency, severity, cost).
    b. Identified data quality issues (missing values, outliers) and flagged them in a notebook.
    c. Conducted exploratory data analysis (EDA), including histograms and ECDFs to understand variable distributions and tail behavior. For demonstration, I have attached a histogram and ECDF graph for 'Sales per customer.' Drafted a short note on applying VACS to disruption similarity tasks.
    d.



6. Collaboration and Repository Management
    a. Coordinated work with team members to avoid duplication.
    b. Uploaded artifacts (data dictionary, EDA notebooks) to a shared Google Drive and GitHub repository.
    c. Maintained logical folder structures and basic documentation in the repository.

Specific work done this month:

Same as above, since the project began this month.

Project Outcome for the month:

- Completion of Phase 1 (Literature & Concept Review):
  - Achieved a comprehensive understanding of stochastic risk assessment methods applicable to supply chain disruptions.
  - Extracted and documented the mathematical framework, Bayesian Network structure, and ripple-effect modeling from the selected paper.
  - Developed a comparison matrix of major stochastic modeling approaches to guide model selection.
- Data Source Identification and Initial Preparation:
  - Identified and documented publicly available datasets suitable for modeling supply chain disruptions.
  - Selected and reviewed a target dataset (e.g., SupplyGraph) for use in modeling tasks.
  - Mapped dataset fields to disruption risk variables with a clear data dictionary.
  - Flagged data quality issues for further cleaning and transformation.
- Preliminary Data Analysis and Exploration:
  - Conducted exploratory data analysis (EDA), producing summary statistics, histograms, and empirical distribution plots.
  - Identified heavy-tailed and multimodal behaviors in variables, informing candidate distribution fitting.
- Seminar Insights and Knowledge Transfer:
  - Summarized key learnings from expert seminars on Variance-Adjusted Cosine Similarity (VACS) and the SMART FedEx Logistics series.
  - Drafted a short note on the potential use of VACS for disruption-similarity analysis in later project phases.
- Collaboration and Documentation:
  - Established a shared Google Drive folder and GitHub repository for all code, documents, and analyses.
  - Ensured that notebooks, data dictionaries, and seminar summaries were uploaded, verified, and organized for team use.

Next Steps (Work Plan for next month):

Finalize and document the data-cleaning pipeline

- Complete prepare_data.py with all imputation and cleaning rules.
- Version and archive the cleaned dataset in /data/clean.
- Add clear docstrings and usage instructions.

Complete and validate the distribution fitting

- Fit selected distributions (Exponential, Weibull, Lognormal, Pareto) to inter-arrival times, severity proxy, and lead-time variance.
- Produce parameter tables with KS, AIC, and BIC metrics.
- Review and select best-fit distributions for Monte Carlo inputs.

Develop and refine the Monte Carlo simulation module (v0.2)

- Use fitted distributions to simulate 10,000–50,000 scenarios.
- Add code for scenario sampling and cost aggregation.
- Plot risk exposure metrics (e.g., mean, percentiles, tail risk).

Design and prototype risk index calculation

- Define the initial formula or method for the composite risk index.
- Implement in modular code (src/calculate_risk_index.py).
- Test on simulated output to validate behavior.

Initial AI-based Decision Support System planning

- Start literature review on RL methods for decision support.
- Outline RL model structure (state, action, reward definitions).
- Document potential data needs and simulation integration.

Polish and expand EDA deliverables

- Add markdown interpretations on distribution fits and residual analysis.
- Include plots comparing empirical vs. fitted distributions.
- Push final plots to /reports/figures.

Maintain and improve collaboration/documentation

- Ensure updated notebooks and scripts are pushed to GitHub with clear README files.
- Tag all deliverables with version numbers.
- Document all assumptions, methods, and intermediate results in a monthly report.

Weekly sync meetings

- Schedule 1–2 brief team check-ins to review distribution fitting, simulation design, and RL planning.
- Capture and share concise meeting notes in the repo.

Project Progress:

Project Timeline and current work activity:

Phase/Work Tasks	Sub-Tasks	June	July	August	Month	Month
3.1 Risk Assessment Framework	Literature review and model design	Completed				
	Paper extraction & modeling approach	Completed				
	Data sourcing and initial mapping	Completed				
	Data cleaning and quality assessment	In Progress	Completed			
	Exploratory data analysis and distribution fitting	In Progress	Completed			
	Monte Carlo simulation prototype	Start	In Progress			
	Multi-factor risk index design			Start		
	Risk index testing and			Start		

	refinement					
AI-Based Decision Support System	RL model planning and literature review		Start	In Progress		
	Data preparation for RL models			Start		
	Initial RL model development			Start		
	Testing and refinement of RL models			Start		

Work Contributions by individual team members in the project:

1. **Puneet**
    a. Conducted a detailed literature review on stochastic risk assessment models disrupting supply chains.
    b. Extracted and documented the mathematical framework, decision variables, constraints, and uncertainty modeling from a key research paper.
    c. Created a comparison matrix of four stochastic modeling approaches (Stochastic Programming, Markov Chains, Bayesian Networks, and Newsvendor models) with application domains, pros, and cons.
    d. Mapped dataset fields to disrupt risk variables by developing a clear data dictionary.
    e. Flagged data quality issues (missing values, outliers) in a dedicated notebook.
    f. Performed exploratory data analysis (EDA), generating histograms and ECDFs for risk variables.
    g. Drafted a one-page brief on the potential use of Variance-Adjusted Cosine Similarity (VACS) for disruption similarity and clustering tasks.
    h. Summarized insights from two relevant seminars (on VACS and SMART FedEx Logistics Series).
    i. Organized and uploaded artifacts (notebooks, documents) to the shared Google Drive and GitHub repository.

2. **Manish**
    a. Researched and shortlisted available benchmark datasets suitable for supply chain disruption modeling (e.g., SupplyGraph, TAC-SCM, EM-DAT) publicly.

b. Collected and documented detailed metadata about the selected dataset in a structured format.
c. Provided the initial dataset for the team's review.
d. Began developing a reproducible data-cleaning script in Python (pandas) to handle missing values, normalize units, and produce a tidy output.
e. Set up and maintained the shared GitHub repository structure (including /data, /notebooks, /src directories).
f. Planned distribution fitting for key risk variables (including model selection, parameter estimation, and goodness-of-fit metrics) as a next-step deliverable.
g. Participated in mid-week coordination meetings to align dataset field mapping and cleaning strategies.