



Data Storm v6.0

— Preliminary Round —

Organized By

Rotaract Club of University of Moratuwa

May 7, 2025

Powered By

OCTAVE

Background to the Business

In our ABC insurance company, agents are the initial point of interaction with potential or prospective clients who show an interest in our services. Their performance in the early months can determine their long-term success.



Figure 1 : Insurance Company

Hackathon Challenge: Predict & Improve Insurance Agent Performance

We have two major goals:

- Predict which agents are at risk of not selling anything (i.e., going "NILL") in the following Month
- Have a clear understanding of the issue and try to improve the performance of all current agents by categorizing them and recommending personalized improvement plans.

Your task is to use real-world agent data to come up with data-driven answers to achieve both our goals.

EDA Part - Exploratory Data Analysis

1. What are the key metrics and distributions in the dataset? (Summary statistics)
2. How do sales patterns vary by month, and are there unexpected drops or spikes performance? (Use your time series knowledge from webinars)
3. How do all numerical features interact simultaneously? Multivariate Analysis)
4. How do individual agent trajectories evolve over time?
5. Innovative EDA (Explore more to uncover hidden insights).

Part 1 - Predict NILL Agents

Unfortunately, some agents fail to sell any policies in their following month — we call this One Month NILL. It is important to identify these agents as early as possible to provide them with guidance to act before it's too late.

What you'll be doing in part 1

- Forecast whether an agent will become One NILL or not (**Kaggle problem will cover this**).
- Identify the key factors that influence early agent performance.
- Suggest personalized SMART action plans (like training, mentoring, motivation) for at-risk agents.

Part 2 - Monitor and Improve Existing Agent Performance

Agents who are already working may be doing well, okay, or poorly — and that can change as well over time. Here we need to track their performance and help them grow to become champion agents.

What you'll be doing in part 2

- Analyze past and current performance data in the given dataset.
- Identify agents by categorizing them based on performance. (e.g., You can categorize High/Medium/Low performers).
- Recommend custom interventions for each group to help them improve.
- Track progress over time and identify if interventions are working.

Final Deliverables and Evaluation

Your submission should include the following:

For EDA: All code + charts in one notebook, with markdown explanations for insights.

For Part 1:

1. Trained prediction model that flags One Month NILL risks (**Kaggle problem will cover this**).
2. List of top factors affecting early performance.
3. Personalized action plan recommendation system for at-risk agents.

For Part 2:

1. Method to classify current agent performance (Low, Medium, High).
2. Intervention strategy based on performance category.
3. An optional progress tracker to measure changes over time.

Bonus: *A simple dashboard or visualization to show predictions, classifications, and recommendations.*

Category	Points
EDA	15 (3 each)
Accuracy of NILL prediction model	25 (Kaggle problem will cover this)
Insightfulness of influencing factors	10
Quality of personalized action plans	15
Effectiveness of classification logic	15
Thoughtfulness of interventions	10
Dashboard or visualization	10

Note: Please note that besides this marking criteria, your Kaggle rank will also be considered.

- You will be given a **36-hour time period** for submission, and no further extensions will be granted.
- According to the evaluation criteria mentioned in the above section, The top **10 teams** with the highest combined total will be selected as **finalists** of DataStorm v6.0

Tools Allowed: Use any tools you're comfortable with - Python, R, Excel, Power BI, Jupyter, Streamlit, etc.

Submission Checklist

1. Code files or notebooks
2. Clear explanation of your approach, Intervention logic and insights (README or PDF).
3. Trained models (or instructions to train)
4. Visualizations or dashboard (if built)