Executive Summary

The following project is built for an end-to-end analytics solution to assess portfolio health, collection risk and collection performance for Boreas Auto Finance.

Business Problem

The organization lacked unified visibility of contracts and repayments for its business team, as a contract's responsibility of repayment (collection) rests with the business team for the first 18 months post disbursement After 18 months a contract it is transferred to the collection team.

Need to monitor collection efficiency, understand overdue, DPD days trend and understand if there is any correlation between tenure and DPD buckets.

Difficulty in identifying high-risk business managers and the concerned branches.

Objective

To enable portfolio level and managerial level risk monitoring for the business team to analyse risk at the very source.

Dataset Overview

9,436 loan contracts | 113k+ repayment records | Period: Jan-2023 to Nov-2024

Note: ER diagram for the schema has been uploaded to the repository.

Table	Description
Contracts	Loan details
Repayment	EMI- transactions
Branches	Branch and Business Managers mapping
Dealers	Dealer master
Brokers	Broker master
Executives	Business Executives mapping

Assumptions & Limitations:

The dataset is synthetically generated for learning and demonstration purposes.

The dataset represents the collection phase/pattern of the business teams of the Banks/NBFCs within the auto loans segment.
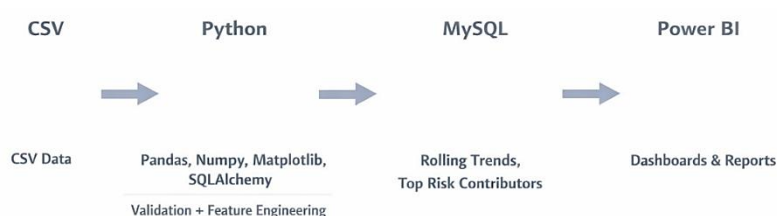
The dataset assumes responsibility of collection rests with the business team for the initial 18 months after the Delivery Order is processed.

The repayment history is limited to 680 days (January 2023 to November 2024).

Each contract has a fixed EMI due date 5th/10th/15th of a month.

Actual payment timestamps are unavailable; therefore, all repayments (inclusive of the delayed payments) are assumed to occur on the scheduled due date of each month.

Tech Stack Used:

CSV	Python	MySQL	Power BI
CSV Data	Pandas, Numpy, Matplotlib, SQLAlchemy	Rolling Trends, Top Risk Contributors	Dashboards & Reports
	Validation + Feature Engineering		

Key Feature Engineering:

Channel Type: Self/Dealer/Broker.

DPD Buckets: Regular/S1/S2/NPA.

## Treatment of Outliers and Null Values:

Note: Outliers were not treated as data issues, since extreme values represent real high-exposure loans relevant to the portfolio risk.

A total of 284 contracts, which is approximately 3%, have missing customer names; these records were kept to maintain comprehensive repayment histories and to prevent the loss of risk indicators.

Regulatory fields such as Payment Mode, PSL Tag, and RC Verification were filled in as 'Unknown' rather than employing statistical techniques, since these fields denote categorical compliance characteristics of the business.

The absence of Dealer and Broker identifiers was deliberately maintained and utilized to create a Channel Type feature (Self / Dealer / Broker), facilitating subsequent channel-level downstream analysis.

Analytics Finding & Business Insights:
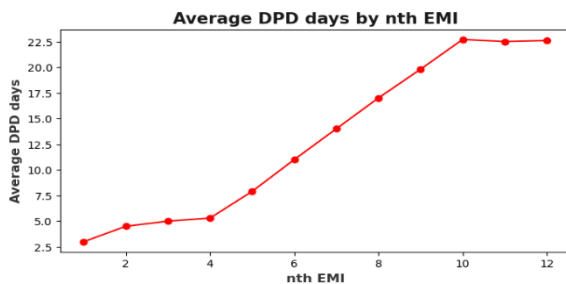
Throughout the data, there is an increase of repayment fatigue in the later loan lifecycle indicated by increase in average DPD days sharply post 4th EMI as the later EMIs show higher DPD days.

```
-------------------------------------------
Correlation between Tenure and DPD buckets:
-------------------------------------------
                 tenure_months   dpd_bucket_num
tenure_months         1.000000         0.005343
dpd_bucket_num        0.005343         1.000000


-------------------------------------------
Spearman ρ = 0.005
-------------------------------------------
```

There is no significant relationship between tenure of the loan and DPD days buckets.
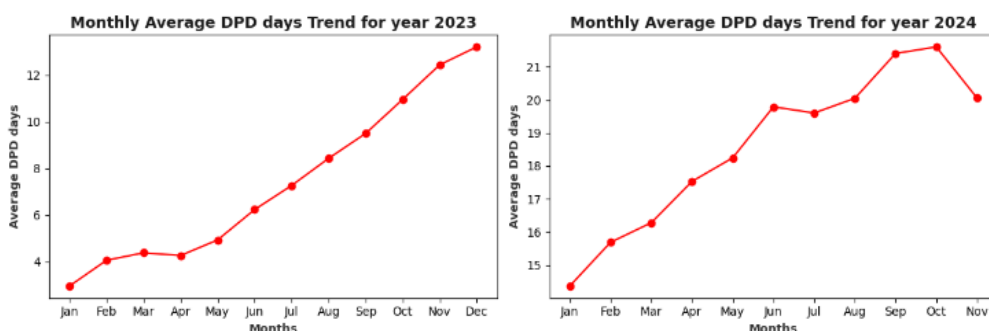
There is a steady rise in average DPD days through 2023 which continues to rise with a higher base in 2024 driven by persistent delinquent contracts where delays are accumulating over time.

month	total_overdue	rolling_3m_avg_overdue
2023-01	1444641.5	1444641.5
2023-02	3515512.5	2480077
2023-03	5692527.5	3550893.83
2023-04	6771382	5326474
2023-05	11807800.5	8090570
2023-06	18926824.5	12502002.33
2023-07	27063277	19265967.33
2023-08	38507190	28165763.83
2023-09	51661943.5	39077470.17
2023-10	67579660	52582931.17
2023-11	87332461	68858021.5
2023-12	108388157.5	87766759.5
2024-01	107223580.5	100981399.67
2024-02	109638186.5	108416641.5
2024-03	103894478.5	106918748.5
2024-04	100148040	104560235
2024-05	93830557.5	99291025.33
2024-06	89381436	94453344.5
2024-07	73554371	85588788.17
2024-08	62019205.5	74985004.17
2024-09	51514433	62362669.83
2024-10	36543232	50025623.5
2024-11	19347736.5	35801800.5

Overdue peaked in early 2024 and later declined.

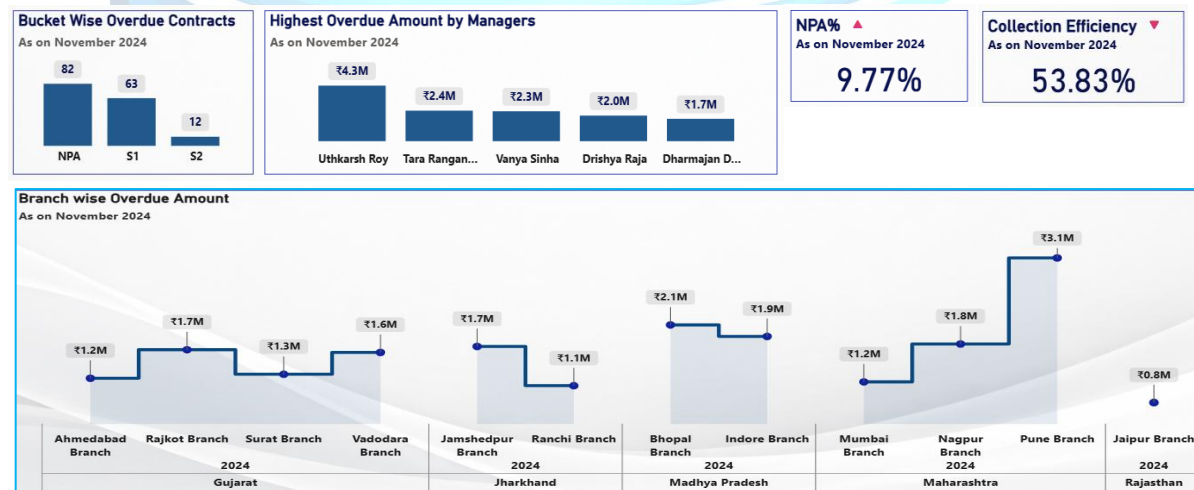| month	total_amount_recovered	rolling_3m_avg_amount_recovered
2023-01	22123237.5	22123237.5
2023-02	42152900	32138068.75
2023-03	65854119	43376752.17
2023-04	90121133.5	66042717.5
2023-05	110592128.5	88855793.67
2023-06	130872434	110528565.33
2023-07	153053878.5	131506147
2023-08	173027818	152318043.5
2023-09	194807137.5	173629611.33
2023-10	214104049.5	193979668.33
2023-11	231732965	213548050.67
2023-12	254857619.5	233564878
2024-01	233253118.5	239947901
2024-02	211129624.5	233080120.83
2024-03	190111350	211498031
2024-04	166462885.5	189234620
2024-05	145673572	167415935.83
2024-06	123971181	145369212.83
2024-07	104391947.5	124678900.17
2024-08	83371590	103911572.83
2024-09	61530966	83098167.83
2024-10	41279430.5	62060662.17
2024-11	22555478	41788624.83

Similar patterns were followed with the recovery amount. It improved throughout early 2024, then declined in later months indicating weakened collection momentum, which aligns with consistent rise in DPD days throughout the portfolio. It primarily indicates resolution of the partial set of overdue contracts and rising accumulation of *hard-to-collect* risk.

Portfolio Health Snapshot:

Total Finance Amount
Contracts: 9,436
₹9.8bn

Total Actual Overdue Amount
Contracts: 157
₹19.3M

PSL% ▼
Contracts: 1,852
19.80%

Average LTV%
80.00%

- Portfolio size stands at INR 9.8 billion with a total of 9,436 contracts.
- PSL (Priority Sector Lending) credit exposure is very low (20%), targeted is 45%.
- Average LTV (Loan to Value) is 80% signifying a conservative lending relative to asset value.
- Despite large portfolio size overdue exposure is concentrated in a small subset of contracts.

Risk Oversight Snapshot:

Bucket Wise Overdue Contracts
As on November 2024
NPA 82 | S1 63 | S2 12

Highest Overdue Amount by Managers
As on November 2024
Uthkarsh Roy ₹4.3M | Tara Rangan… ₹2.4M | Vanya Sinha ₹2.3M | Drishya Raja ₹2.0M | Dharmajan D… ₹1.7M

NPA% ▲
As on November 2024
9.77%

Collection Efficiency ▼
As on November 2024
53.83%

Branch wise Overdue Amount
As on November 2024

Gujarat				Jharkhand		Madhya Pradesh		Maharashtra			Rajasthan
Ahmedabad Branch ₹1.2M	Rajkot Branch ₹1.7M	Surat Branch ₹1.3M	Vadodara Branch ₹1.6M	Jamshedpur Branch ₹1.7M	Ranchi Branch ₹1.1M	Bhopal Branch ₹2.1M	Indore Branch ₹1.9M	Mumbai Branch ₹1.2M	Nagpur Branch ₹1.8M	Pune Branch ₹3.1M	Jaipur Branch ₹0.8M

- As on **November 2024, NPAs are severely high (9.77%)** as compared to the industry range.
- The NPA/S1/S2 buckets point towards a rolling stress on collection by the business team.
- **Collection Efficiency at 53%** indicates a crisis for the month of November 2024 which requires focused operational intervention.
- **Highest overdue** in branches is **held by Pune (INR 3.1 Million)** and in Managers it is held my Mr.Uthkarsh Roy. (INR 4.3 Million) followed by Tara Ranganathan (INR 2.4 Million) and Vanya Sinha (INR 2.3 Million) highlighting accountability clusters.

Manager Wise Oversight Snapshot:
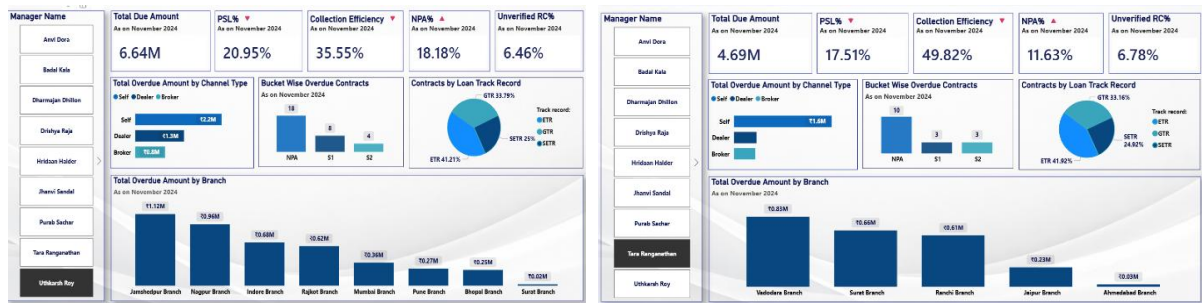
- *Mr.Uthkarsh holds the highest overdue* in the portfolio with *35% Collection Efficiency & 18% NPA* (with a total number of 18 NPAs).
- *Mr.Tara holds the 2nd highest overdue* in the portfolio with *49% Collection Efficiency & 11% NPA* (with a total number of 10 NPAs).

Business Recommendations:

Immediate Collection Stabilization: Initiate focused recovery drives for overdue contracts, prioritize NPA and corresponding overdue buckets.

Manager and Branch risk governance: Introduce daily/weekly scorecards (comprising Collection efficiency, and signalling towards upcoming overdue cases inclusive of DPD days buckets).

Channel risk control: Introduce strict underwriting and risk control norms under Dealer and Broker sourced leads and scrutiny for repeated customer as per loan track records.

Lifecycle based suggestions: As DPD days shoot up post 4th EMI introduced early reminders and follow ups starting 2nd EMI. Introduce risk-based early transfer to collections (before 18 months) for persistently overdue contracts. Resort to repossession in severe/critical cases.

Links:

Power BI Dashboard link:

https://app.powerbi.com/links/ONmEj4I1oK?ctid=25d120d6-b16a-44af-bb4b-ac24bb3751c9&pbi_source=linkShare&bookmarkGuid=2efb46df-a44a-4cea-a7d1-e0b9a360dc9e

LinkedIn: https://www.linkedin.com/in/vedant28bhardwaj

GitHub: vedantbhardwaj10 (Vedant Bhardwaj)