Fleet Maintenance Triage Agent

Results deck

- Objective: prioritize vehicles for service using predicted near-term breakdown risk and operational cost.
- Reproducible pipeline: validate → EDA/stats → train → triage snapshot → latest aliases → deck/PDF.
- Slides are built from exported artifacts in outputs/ (figures + tables).

How to interpret

- This deck is fully reproducible: rerun the pipeline to refresh numbers and visuals.
- The ‘snapshot’ in the footer tells you which triage run the deck is based on.

Agenda

- Data quality & dataset overview
- EDA + reliability/statistics (KM + Cox)
- Model performance (PR + calibration) + metrics
- Fleet triage view + recommended queue
- Guardrails + similar-case evidence
- Appendix: additional artifacts

How to interpret

- Main slides tell the story end-to-end.
- Appendix contains extra plots/tables auto-included from outputs/.

Pipeline

One-command reproducible workflow

- Run full pipeline: .\scripts\run.ps1 report
- Or rebuild deck only: .\scripts\run.ps1 deck

How to interpret

- Validate catches data issues early (schema, ranges, missingness).
- EDA/Stats gives context and risk drivers (KM + Cox).
- Train exports PR/calibration + metrics; Triage exports queue + guardrails + evidence.

Executive summary

0.053

Best AUPRC

0.049

Best Brier

hgb / 30d

Model / Horizon

42 / 8

Actionable / Abstain

How to interpret

- AUPRC is best for rare events (higher is better).
- Brier summarizes probability quality (lower is better).
- Actionable vs Abstain shows how often guardrails trigger manual review.

Data quality

Validation + dataset overview

Data quality
Validation summary

Validation summary

item	value
vehicle_day rows	9000
vehicle_day cols	30
dtc_event rows	3155
work_order rows	11
checks passed	31
checks failed	0

Great Expectations checks

How to interpret

- Each row is a data quality rule (schema/range/missingness).
- All checks should pass in normal operation; failures mean unreliable downstream results.
- Use this slide to justify that the pipeline is safe to run automatically.

Key takeaways
Data quality

- Goal: ensure inputs are consistent (types, ranges, missingness) before modeling.
- Checks passed: 31/31 (pipeline gate).
- If checks fail, rerun data generation / fix upstream ingestion before trusting any metrics.
- This is the foundation for a production-quality ML/agent workflow.

EDA & statistics

Operational patterns and risk factors

EDA overview

Dataset + operational signals

EDA overview

metric	value
vehicle_day rows	9000
vehicle_day cols	31
n_vehicles	50
date_min	2024-01-01
date_max	2024-06-28
n_days	180
dtc_event rows	3155
dtc events per vehicle-day	0.35056
work_order rows	11
label prevalence breakdown_7d	0.00967
label prevalence breakdown_30d	0.03522

Dataset summary

Work orders by subsystem

subsystem	n_work_orders	mean_downtime_days	mean_parts_lead_time_days
brakes_wheel_end	5	6.000	1.000
cooling	3	3.000	3.800
aftertreatment	1	4.000	2.000
drivetrain	1	3.000	0.000
electrical	1	2.000	3.000

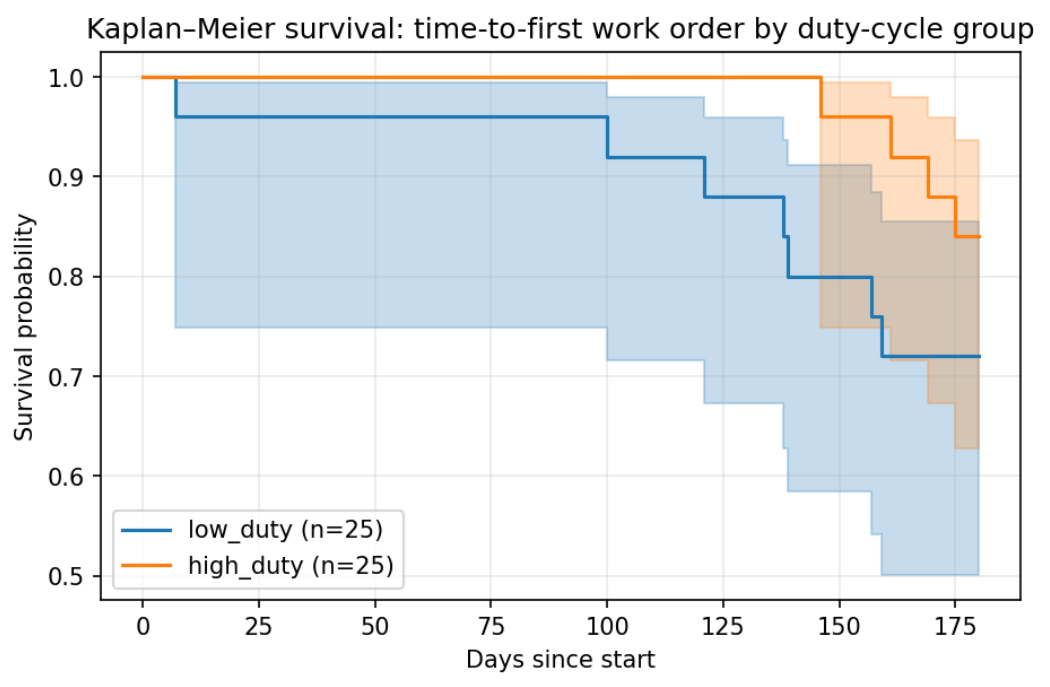
Maintenance demand

How to interpret

- EDA overview: sanity-check volume, coverage, and key variables (are values plausible?).
- Work orders: higher counts indicate where maintenance load is concentrated.
- Use this to connect the model to real operations (what breaks, where, and how often).

Reliability & risk factors

Survival + hazard ratios



Kaplan–Meier by duty-cycle group

Cox hazard ratios

covariate	coef	se_coef	p	hazard_ratio	hr_ci95_low	hr_ci95_high
duty_cycle_mean	-8.565	19.75	0.665	0.000	0.000	12,379,210,443,134
ambient_temp_mean	0.834	0.980	0.395	2.302	0.337	15.71
dtc_rate_mean	0.989	4.338	0.820	2.689	0.001	13,241
spc_alert_share	-7.025	4.658	0.131	0.001	0.000	8.196

Risk multipliers (HR)

How to interpret

- Kaplan–Meier: lower curve means faster failures; separation suggests duty cycle impacts reliability.
- Cox hazard ratios: HR>1 increases failure hazard, HR<1 decreases; bigger deviation from 1 = stronger effect.
- These results help explain *why* certain vehicles get prioritized (not only that they do).

Key takeaways

EDA & statistics

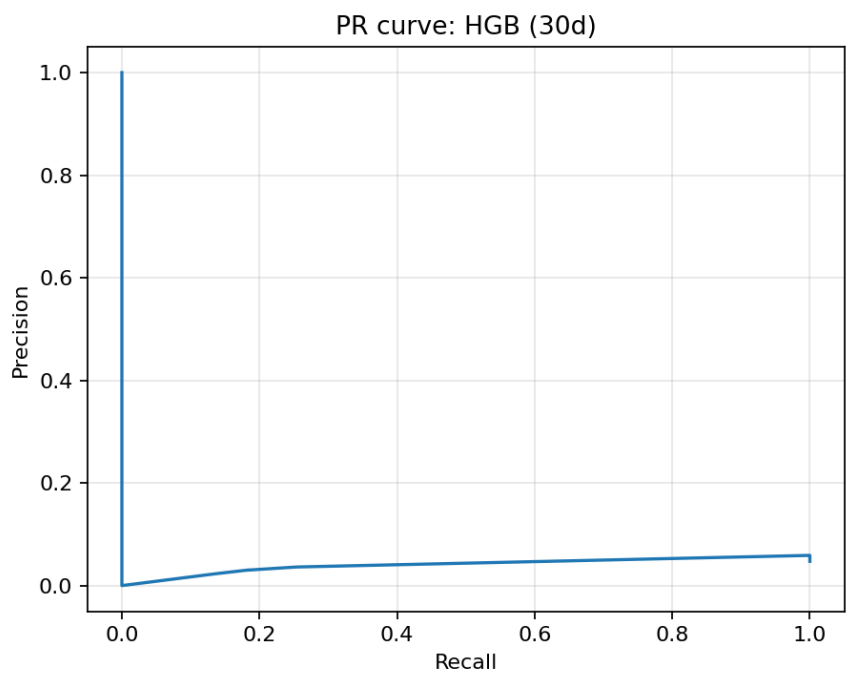
- EDA connects model outputs to fleet reality: coverage, usage patterns, and failure/service signals.
- Highest maintenance demand: brakes_wheel_end (5 work orders).
- Highest maintenance demand: cooling (3 work orders).
- Top risk-increasing factor: dtc_rate_mean (HR≈2.69; >1 increases hazard).
- Top risk-increasing factor: ambient_temp_mean (HR≈2.30; >1 increases hazard).
- Use these insights to propose targeted interventions (routes, duty cycles, subsystems).

Model performance

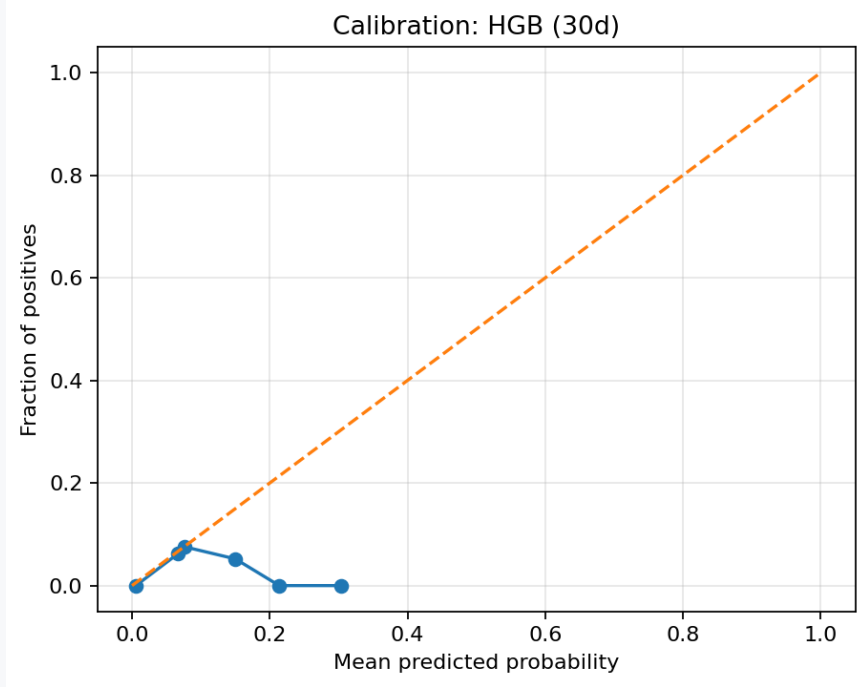
Calibration + ranking quality

Model performance

hgb • 30d



Precision–Recall



Calibration

How to interpret

- Precision–Recall: higher is better; focus on precision at the recall region you need operationally.
- Calibration: points close to the diagonal mean predicted probabilities match real-world frequencies.
- Good ranking (PR) helps choose *which* vehicles; good calibration helps decide *how urgent* they are.

Metrics summary

Model ranking + operational precision

Risk metrics (sorted)

task	model	calibrated	auprc	brier	precision_at_k_mondays_
risk_30d	logreg	1	0.053	0.049	0.025
risk_30d	hgb	1	0.051	0.055	0.000
risk_7d	logreg	1	0.028	0.025	0.075
risk_7d	hgb	1	0.027	0.027	0.025

Overall ranking

Precision@K (Mondays)

monday_date	precision_at_k	k	horizon_days	model
2024-06-03	0.100	10	7	logreg
2024-06-10	0.100	10	7	logreg
2024-06-17	0.100	10	7	logreg
2024-06-24	0.000	10	7	logreg
2024-06-03	0.000	10	7	hgb
2024-06-10	0.000	10	7	hgb
2024-06-17	0.100	10	7	hgb
2024-06-24	0.000	10	7	hgb
2024-06-03	0.000	10	30	logreg
2024-06-10	0.000	10	30	logreg
2024-06-17	0.100	10	30	logreg
2024-06-24	0.000	10	30	logreg
2024-06-03	0.000	10	30	hgb
2024-06-10	0.000	10	30	hgb

Showing first 14 rows (truncated).

Top-K quality

How to interpret

- Risk metrics: compare models; higher AUPRC is better for rare events.
- Precision@K: if K is your service capacity, this approximates how many queued vehicles are truly high-risk.
- Use these tables to justify the selected model and why it’s operationally useful.

Key takeaways

Model performance

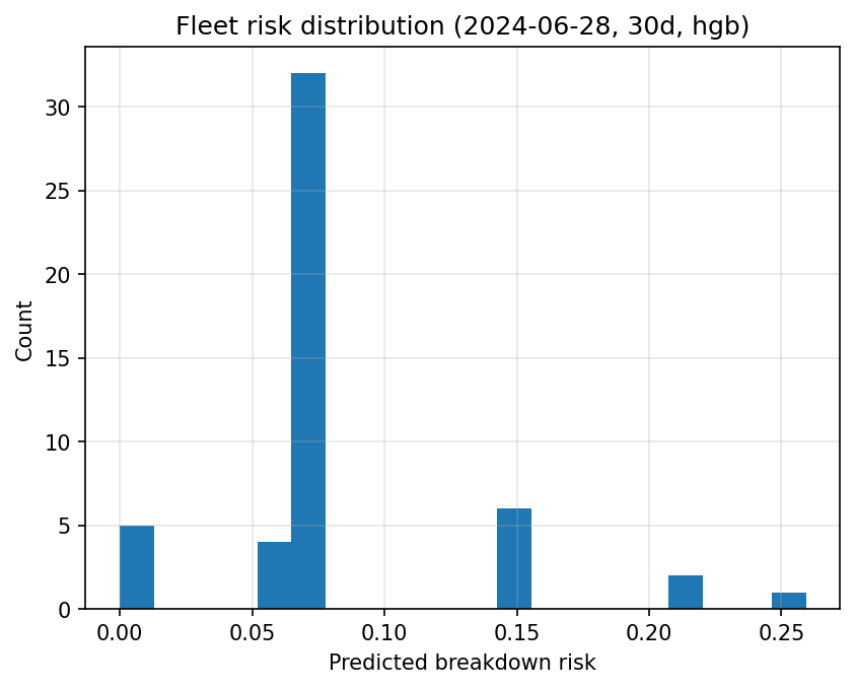
- We care about both: (1) ranking the right vehicles, and (2) reliable probabilities for urgency.
- Best ranking quality (AUPRC): 0.053 for risk_30d / logreg.
- Lower Brier is better (probabilities). Best Brier: 0.049.
- PR supports ‘who to service first’; calibration supports ‘how risky is this vehicle’.
- Precision@K maps model performance to a real service queue size.

Fleet triage

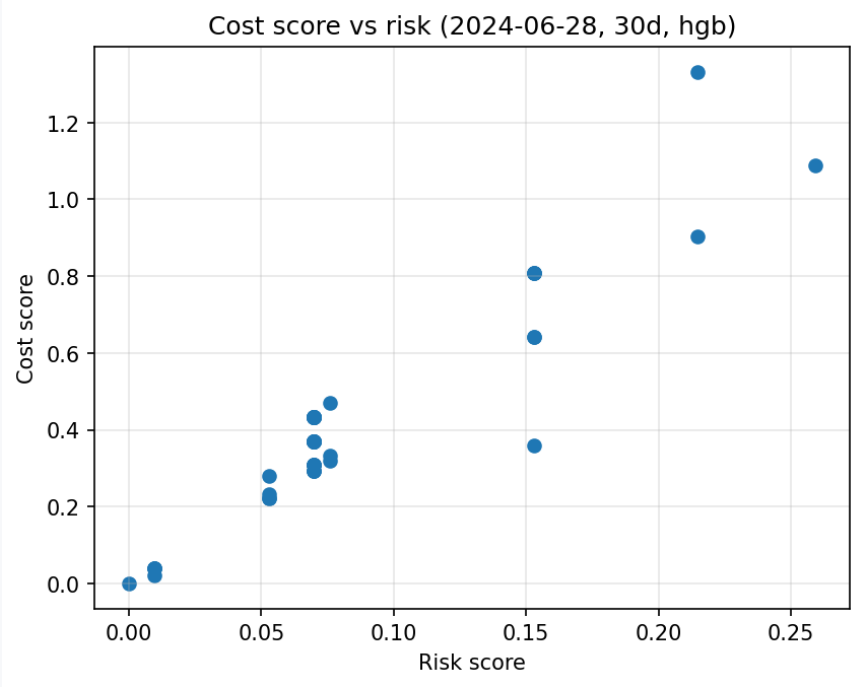
Decision view + recommended actions

Fleet decision view

Risk distribution + decision trade-off



Fleet risk distribution (latest)



Cost vs risk (latest)

How to interpret

- Histogram: look for a small high-risk tail (few vehicles drive most of the risk).
- Cost vs risk: upper-right suggests high priority (high risk and high expected cost impact).
- This view motivates why prioritization beats ‘first come, first served’.

Triage summary

How many vehicles are actionable today?

50	42	8	10
Fleet size	Actionable	Abstain	Queue (K)

How to interpret

- Actionable = passes guardrails; candidates for automatic ranking and scheduling.
- Abstain = OOD/low confidence; should be reviewed manually or require more data.
- Queue (K) maps to daily/weekly service capacity.

Mean risk by route

Where risk concentrates operationally

Mean risk by route (latest)

route_type	mean_risk
mixed	0.109
highway	0.070
shop	0.070
urban	0.060

Targeted planning by route

How to interpret

- Higher mean risk routes may indicate harsher operating conditions or specific usage profiles.
- This supports operational mitigation: route-specific inspections, spares, or preventive maintenance.
- Treat as a planning layer (aggregate view), not a substitute for per-vehicle ranking.

Recommended service queue

Top-K prioritized vehicles

Service queue (latest)

vehicle_id	route_type	duty_cycle	dtc_count_to	spc_alert_toc	risk_pct	risk_band	ood_flag	subsystem_g	cost_score	recommended
V0016	mixed	0.458	0	0	21.50	Low	False	aftertreatment	1.332	No action (routi
V0013	mixed	0.546	2	1	21.50	Low	False	drivetrain	0.902	No action (routi
V0048	mixed	0.744	0	1	15.30	Low	False	cooling	0.807	No action (routi
V0014	mixed	0.330	3	1	15.30	Low	False	cooling	0.807	No action (routi
V0042	highway	0.602	0	1	15.30	Low	False	cooling	0.807	No action (routi
V0020	mixed	0.854	1	1	15.30	Low	False	drivetrain	0.642	No action (routi
V0025	highway	0.728	0	1	7.600	Low	False	aftertreatment	0.470	No action (routi
V0003	highway	0.642	2	1	7.000	Low	False	aftertreatment	0.433	No action (routi
V0001	mixed	0.349	0	1	7.000	Low	False	aftertreatment	0.433	No action (routi
V0002	highway	0.755	0	1	7.000	Low	False	aftertreatment	0.433	No action (routi

Ranked by cost impact (latest snapshot)

How to interpret

- Rows are ordered by the chosen ranking objective (e.g., cost impact).
- Top of the list = highest expected value of servicing now (given risk + cost).
- Use as an input to scheduling; final decision can consider constraints (parts, location, etc.).

Guardrails

Abstentions (OOD / low confidence)

Abstained vehicles (latest)

vehicle_id	route_type	risk_pct	risk_band	ood_flag	abstain_reasons
V0004	mixed	25.90	Low	True	ood
V0012	urban	15.30	Low	True	ood
V0026	mixed	7.000	Low	True	ood
V0035	shop	7.000	Low	True	ood
V0045	mixed	7.000	Low	True	ood
V0036	urban	7.000	Low	True	ood
V0046	mixed	7.000	Low	True	ood
V0028	urban	0.900	Low	True	ood

Flagged for manual review

How to interpret

- Abstentions are not ‘low risk’; they are ‘low trust’.
- Common causes: out-of-distribution features, missing signals, or probability uncertainty.
- Operationally: route these cases to technicians/engineers for targeted diagnosis.

Similar-case evidence

Nearest-neighbor retrieval for transparency

Similar-cases index

rank	vehicle_id	alias_tex	alias_csv
1	V0042	similar_cases_latest_01.tex	similar_cases_latest_01.csv
2	V0014	similar_cases_latest_02.tex	similar_cases_latest_02.csv
3	V0048	similar_cases_latest_03.tex	similar_cases_latest_03.csv
4	V0013	similar_cases_latest_04.tex	similar_cases_latest_04.csv
5	V0016	similar_cases_latest_05.tex	similar_cases_latest_05.csv

Index

Example evidence (similar_cases_latest_01.csv)

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0046	2024-06-02	highway	0.618	0	1	0	0	0.308	2.246
V0003	2024-06-14	highway	0.653	0	1	0	0	0.306	2.273
V0047	2024-05-12	highway	0.459	0	1	0	0	0.303	2.298
V0042	2024-06-25	highway	0.583	0	1	0	0	0.303	2.305
V0047	2024-06-10	highway	0.511	0	1	0	0	0.292	2.426

Example

How to interpret

- Index lists nearest historical cases for a queried vehicle (similar signals/usage profile).
- Use evidence to explain the ranking: ‘this looks like past failures of type X’.
- This improves trust and helps engineers validate whether the recommendation makes sense.

Key takeaways

Fleet triage

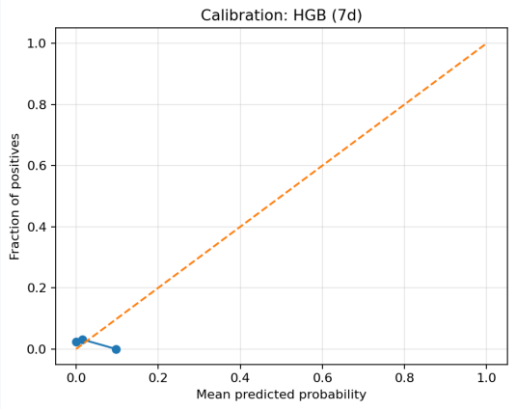
- Triage converts model outputs into operational actions: queue, abstentions, and explanations.
- Actionable vs Abstain: 42 vs 8.
- Interpretation: abstentions represent uncertainty/OOD and should be manually reviewed.
- Risk distribution often shows a high-risk tail → prioritization yields high ROI.
- Service queue is a ranked shortlist; similar-case evidence supports transparency and debugging.

Appendix

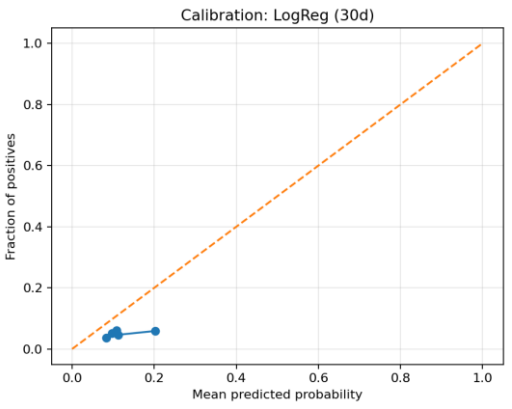
Additional figures and tables (auto)

Appendix: Additional figures

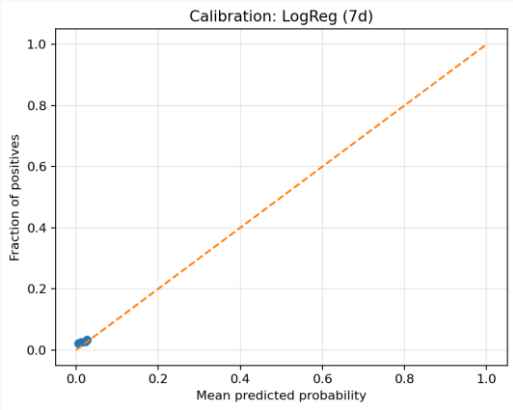
1/2



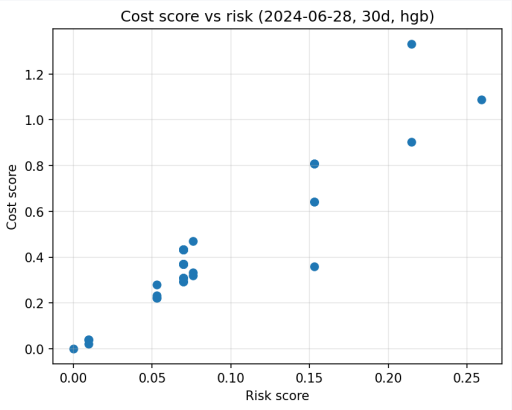
calibration_hgb_7d.png



calibration_logreg_30d.png

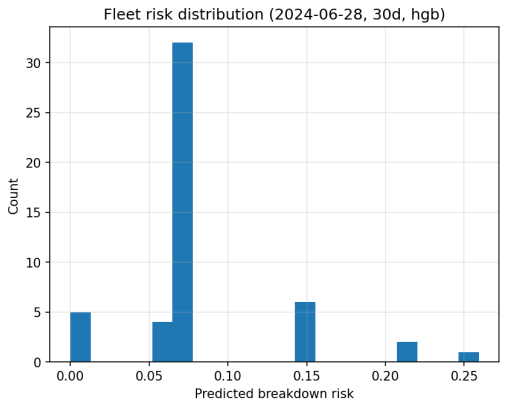


calibration_logreg_7d.png

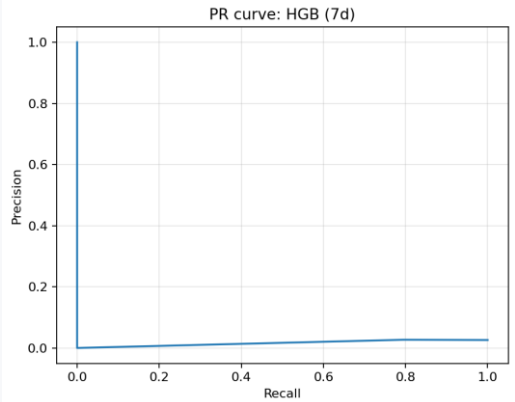


fleet_cost_vs_risk_2024-06-28_30d_hgb_cost.png

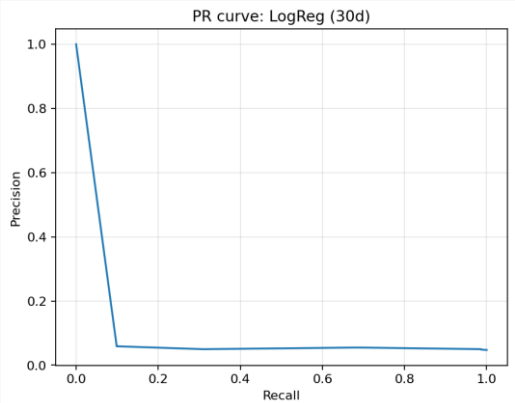
Appendix: Additional figures
2/2



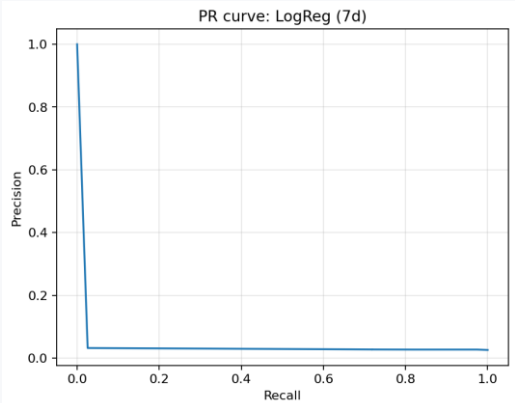
fleet_risk_hist_2024-06-28_30d_hgb_cost.png



pr_hgb_7d.png



pr_logreg_30d.png



pr_logreg_7d.png

Appendix: Additional table

abstained_vehicles_2024-06-28_30d_hgb_cost.csv

abstained_vehicles_2024-06-28_30d_hgb_cost.csv

vehicle_id	route_type	risk_pct	risk_band	ood_flag	abstain_reasons
V0004	mixed	25.90	Low	True	ood
V0012	urban	15.30	Low	True	ood
V0026	mixed	7.000	Low	True	ood
V0035	shop	7.000	Low	True	ood
V0045	mixed	7.000	Low	True	ood
V0036	urban	7.000	Low	True	ood
V0046	mixed	7.000	Low	True	ood
V0028	urban	0.900	Low	True	ood

abstained_vehicles_2024-06-28_30d_hgb_cost.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

mean_risk_by_route_2024-06-28_30d_hgb_cost.csv

mean_risk_by_route_2024-06-28_30d_hgb_cost.csv

route_type	mean_risk
mixed	0.109
highway	0.070
shop	0.070
urban	0.060

mean_risk_by_route_2024-06-28_30d_hgb_cost.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

service_queue_2024-06-28_30d_hgb_cost.csv

service_queue_2024-06-28_30d_hgb_cost.csv

vehicle_id	route_type	duty_cycle	dtc_count_to	spc_alert_toc	risk_pct	risk_band	ood_flag	subsystem_g	cost_score	recommended
V0016	mixed	0.458	0	0	21.50	Low	False	aftertreatmen	1.332	No action (routi
V0013	mixed	0.546	2	1	21.50	Low	False	drivetrain	0.902	No action (routi
V0048	mixed	0.744	0	1	15.30	Low	False	cooling	0.807	No action (routi
V0014	mixed	0.330	3	1	15.30	Low	False	cooling	0.807	No action (routi
V0042	highway	0.602	0	1	15.30	Low	False	cooling	0.807	No action (routi
V0020	mixed	0.854	1	1	15.30	Low	False	drivetrain	0.642	No action (routi
V0025	highway	0.728	0	1	7.600	Low	False	aftertreatmen	0.470	No action (routi
V0003	highway	0.642	2	1	7.000	Low	False	aftertreatmen	0.433	No action (routi
V0001	mixed	0.349	0	1	7.000	Low	False	aftertreatmen	0.433	No action (routi
V0002	highway	0.755	0	1	7.000	Low	False	aftertreatmen	0.433	No action (routi

service_queue_2024-06-28_30d_hgb_cost.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_2024-06-28_30d_hgb_cost_V0013.csv

similar_cases_2024-06-28_30d_hgb_cost_V0013.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0009	2024-06-01	mixed	0.703	1	1	1	1	0.278	2.596
V0021	2024-04-11	highway	0.649	2	1	0	0	0.250	3.008
V0030	2024-05-19	mixed	0.543	3	1	0	0	0.249	3.012
V0017	2024-06-18	highway	0.714	2	1	0	0	0.239	3.179
V0045	2024-05-23	mixed	0.496	2	1	0	0	0.236	3.243

similar_cases_2024-06-28_30d_hgb_cost_V0013.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_2024-06-28_30d_hgb_cost_V0014.csv

similar_cases_2024-06-28_30d_hgb_cost_V0014.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0046	2024-05-30	mixed	0.462	3	1	0	0	0.244	3.105
V0017	2024-05-06	mixed	0.404	3	1	0	1	0.241	3.151
V0004	2024-05-21	mixed	0.419	3	1	0	0	0.228	3.378
V0011	2024-04-22	mixed	0.277	2	1	0	0	0.225	3.450
V0021	2024-06-07	mixed	0.392	3	1	0	0	0.216	3.625

similar_cases_2024-06-28_30d_hgb_cost_V0014.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_2024-06-28_30d_hgb_cost_V0016.csv

similar_cases_2024-06-28_30d_hgb_cost_V0016.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0044	2024-04-07	urban	0.353	0	0	0	0	0.252	2.968
V0011	2024-03-11	mixed	0.247	0	0	0	0	0.239	3.184
V0031	2024-03-30	urban	0.362	0	0	0	0	0.237	3.225
V0004	2024-02-26	mixed	0.413	0	0	0	0	0.227	3.413
V0020	2024-03-07	mixed	0.366	0	0	0	0	0.225	3.435

similar_cases_2024-06-28_30d_hgb_cost_V0016.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_2024-06-28_30d_hgb_cost_V0042.csv

similar_cases_2024-06-28_30d_hgb_cost_V0042.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0046	2024-06-02	highway	0.618	0	1	0	0	0.308	2.246
V0003	2024-06-14	highway	0.653	0	1	0	0	0.306	2.273
V0047	2024-05-12	highway	0.459	0	1	0	0	0.303	2.298
V0042	2024-06-25	highway	0.583	0	1	0	0	0.303	2.305
V0047	2024-06-10	highway	0.511	0	1	0	0	0.292	2.426

similar_cases_2024-06-28_30d_hgb_cost_V0042.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_2024-06-28_30d_hgb_cost_V0048.csv

similar_cases_2024-06-28_30d_hgb_cost_V0048.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0048	2024-05-13	highway	0.885	0	1	0	0	0.272	2.671
V0024	2024-05-15	mixed	0.751	0	1	0	0	0.244	3.101
V0042	2024-05-14	highway	0.790	0	1	0	0	0.235	3.250
V0030	2024-06-05	mixed	0.703	0	1	0	0	0.234	3.268
V0041	2024-04-30	mixed	0.776	0	1	0	0	0.229	3.372

similar_cases_2024-06-28_30d_hgb_cost_V0048.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_latest_02.csv

similar_cases_latest_02.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0046	2024-05-30	mixed	0.462	3	1	0	0	0.244	3.105
V0017	2024-05-06	mixed	0.404	3	1	0	1	0.241	3.151
V0004	2024-05-21	mixed	0.419	3	1	0	0	0.228	3.378
V0011	2024-04-22	mixed	0.277	2	1	0	0	0.225	3.450
V0021	2024-06-07	mixed	0.392	3	1	0	0	0.216	3.625

similar_cases_latest_02.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_latest_03.csv

similar_cases_latest_03.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0048	2024-05-13	highway	0.885	0	1	0	0	0.272	2.671
V0024	2024-05-15	mixed	0.751	0	1	0	0	0.244	3.101
V0042	2024-05-14	highway	0.790	0	1	0	0	0.235	3.250
V0030	2024-06-05	mixed	0.703	0	1	0	0	0.234	3.268
V0041	2024-04-30	mixed	0.776	0	1	0	0	0.229	3.372

similar_cases_latest_03.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_latest_04.csv

similar_cases_latest_04.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0009	2024-06-01	mixed	0.703	1	1	1	1	0.278	2.596
V0021	2024-04-11	highway	0.649	2	1	0	0	0.250	3.008
V0030	2024-05-19	mixed	0.543	3	1	0	0	0.249	3.012
V0017	2024-06-18	highway	0.714	2	1	0	0	0.239	3.179
V0045	2024-05-23	mixed	0.496	2	1	0	0	0.236	3.243

similar_cases_latest_04.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.

Appendix: Additional table

similar_cases_latest_05.csv

similar_cases_latest_05.csv

vehicle_id	date	route_type	duty_cycle	dtc_count_tod	spc_alert_toda	breakdown_7c	breakdown_3(	similarity	distance
V0044	2024-04-07	urban	0.353	0	0	0	0	0.252	2.968
V0011	2024-03-11	mixed	0.247	0	0	0	0	0.239	3.184
V0031	2024-03-30	urban	0.362	0	0	0	0	0.237	3.225
V0004	2024-02-26	mixed	0.413	0	0	0	0	0.227	3.413
V0020	2024-03-07	mixed	0.366	0	0	0	0	0.225	3.435

similar_cases_latest_05.csv

How to interpret

- This table is auto-included from outputs/tables.
- Use it as supporting detail; keep the main story in the earlier slides.