

Stratum Retail Group (retail operator)



Company archetype

Large retail and distribution company running both data-centre and cloud systems

Mission this quarter

Improve demand forecasting and pricing while protecting margin

CEO mandate

AI must improve profit. It must not create instability during peak trading periods.

CTO mandate

Roll out better forecasting models across regions without causing operational disruption.

AI ambition

Use existing infrastructure for steady workloads. Use cloud when extra capacity is needed, especially during seasonal peaks.

Constraints – the rules of the game

- Speed: moderate
 - Risk tolerance: low during peak sales periods
 - Budget pressure: high because margins are thin
 - Compliance and audit: medium
 - Organisational reality: strong production control culture; AI growing quickly
-

Red lines – must not happen

- Cloud costs spiking during peak trading
 - Forecast models changing too often and disrupting production planning
 - Margin improvements cancelled out by infrastructure growth
 - Data moving outside approved environments without oversight
-

Primary success definition – 90 days

Forecast accuracy improves. Infrastructure costs remain predictable and clearly linked to profit improvement.

FinOps governance posture – starting point

Likely strong already

- Change control in production
- Monitoring of core systems
- Operational discipline

Likely weak

- Clear lifecycle for AI models
 - Defined retraining windows
 - Visibility of cost per forecasting model
 - Rules for when to run workloads on-premise versus cloud
-

Suggested policy focus – pick three to five

- Clear stages for AI models before they affect production
 - Defined times when retraining is allowed
 - Cost tracking per forecasting model
 - Simple rule for workload placement between data centre and cloud
 - Controls for scaling during peak periods
-

Suggested rollout posture

- Start in one region or non-peak period
- Use notify-first before enforcing limits
- Allow exceptions, but make them visible and time-limited