Dashboard Walkthrough and Technical Summary

Purpose of this document: give you a clear script for your video walkthrough, while also explaining th technical work behind the dashboard.

Project objective: reduce late deliveries by diagnosing where delay happens (seller processing vs carr transit), identifying high-risk routes, and simulating a better ETA/EDD policy.

Data used: Olist raw datasets in data/raw; primary analysis grain is order-seller leg. Delivered sample size is 97,811 records.

What was implemented:

- Reusable analytics pipeline in src/utils.py for canonical-table creation, feature engineering, route ranking, and EDD simulation.

- Four notebooks for staged analysis: data quality, delivery definition, root-cause analysis, and modeling/solution.

- Streamlit dashboard (app.py) for stakeholder storytelling and scenario testing.

Video Script Flow

Video walkthrough script (suggested sequence):

1) Open with business context: "A late delivery is when actual delivery days exceed promised delivery days."

2) KPI snapshot: delivered orders = 97,811; late rate = 8.0%; average promised = 23.8 days; average actual = 12.5 days.

3) Explain trend panel: monthly miss rate (%) shows whether reliability is improving or deteriorating over time.

4) Explain root-cause panel: compare on-time vs late cohorts to isolate whether seller handling time or carrier transit dominates misses.

5) Explain hotspot-routes panel: top state-to-state lanes by failure rate (with minimum volume thresh so actions target high-impact lanes, not noisy low-volume edges.

6) Explain solution simulator: adjust quantile level and risk weighting to test trade-off between fewer misses and longer promised windows.

7) Close with actions for operations, carrier management, and CX messaging.

Results and Interpretation

Key analytical findings from current run:

- Overall late rate: 8.0%.

- On-time cohort: seller processing 3.0 days, carrier transit 7.9 days.

- Late cohort: seller processing 5.8 days, carrier transit 25.7 days.

- Interpretation: both components increase for late orders, but carrier transit shows the larger jump in this dataset.

Top 5 failure routes (min volume = 100):

- SP->AL: failure 26.1%, volume 257, mean delay -7.3 days

- MA->SP: failure 25.0%, volume 124, mean delay -9.2 days

- SP->MA: failure 21.1%, volume 497, mean delay -8.6 days

- SP->PI: failure 18.1%, volume 331, mean delay -9.9 days

- PR->BA: failure 16.7%, volume 144, mean delay -9.6 days

How the Solution Works + Trade-offs

How the dashboard solution simulator works (technical but plain language):

Step A: Build a baseline ETA from historical quantiles of seller processing time and carrier transit time.

Step B: Compute risk score = weighted blend of route historical late rate and seller historical late rate.

Step C: Add a risk buffer (0 to 3 days) based on risk-score bucket.

Step D: Compare current policy vs proposed policy on late-rate and promised-day impact.

Important demo insight:

- At quantile 0.80, proposed late rate is 16.1% vs current 7.5% (worse by 8.6 pp).

- At quantile 0.95, proposed late rate is 3.4% vs current 7.5% (improves by 4.1 pp), with +3.8 promise days.

Use this to explain trade-offs: stronger reliability usually requires longer promise windows.

ETA confidence bands at q=0.95 baseline: P50=25.4 days, P80=37.4 days, P90=41.3 days.

Stakeholder Closing and Key

Suggested closing statement for stakeholders:

"We now have a transparent reliability framework: identify where delay happens, prioritize the worst lanes, and tune ETA policy based on explicit service-vs-speed trade-offs."

Abbreviation key:

- ETA: Estimated Time of Arrival (promised delivery date).
- EDD: Estimated Delivery Date logic used to compute promises.
- RCA: Root Cause Analysis.
- CX: Customer Experience.
- SLA: Service Level Agreement.
- pp: Percentage points (difference between two percentages).

Files to reference during your video:

- app.py (dashboard)
- src/utils.py (pipeline logic)
- notebooks/00_data_quality.ipynb
- notebooks/01_define_late_delivery.ipynb
- notebooks/02_root_cause_analysis.ipynb
- notebooks/03_nps_impact_and_solutions.ipynb