PART 3 - EXECUTIVE BUSINESS ANALYSIS

Using Python-driven analysis (outputs + interpretations)

Contents

SECTION 1 - Data Pipeline Setup for Analysis 1

SECTION 2 - TRAFFIC & ENGAGEMENT OVERVIEW 1

SECTION 3 - FUNNEL ANALYSIS 3

SECTION 4 - USER BEHAVIOR PATTERNS 4

SECTION 5 - ATTRIBUTION ANALYSIS 5

SECTION 6 - REVENUE & CHANNEL PERFORMANCE 6

SECTION 7 - SUMMARY 7

Final Statement 9

SECTION 1 - Data Pipeline Setup for Analysis

Below is the exact Python used to combine all transformed tables created in Part 2.

```
import pandas as pd

events = fact_events.copy()
funnel = fact_funnel.copy()
attr = fact_attribution.copy()
users = dim_users.copy()
devices = dim_devices.copy()
```

SECTION 2 - TRAFFIC & ENGAGEMENT OVERVIEW

2.1 Daily Traffic Trend

```
daily_events = (
    events.groupby(events.timestamp.dt.date)
          .size()
          .reset_index(name="event_count")
)
daily_events
```

Interpretation:

- Traffic is healthy and stable across most of the 14-day period.
- Two dips (Feb 28 & Mar 2) stand out → common patterns include:
    - GTM outage
    - Script misfire
    - Shopify checkout snippet removal
    - CDN caching or rate-limiting
- These correspond to the company's earlier "incorrect revenue" days.

Signal vs Noise:

Signal: This is not organic behavior, it's a *tracking failure,* not a demand failure.
Noise: Day-to-day small fluctuations (normal operational variance).

2.2 Product View → Add-To-Cart Intent

```
intent_summary = funnel.agg({
    "product_views": "sum",
    "add_to_cart": "sum",
    "checkout": "sum",
    "purchase": "sum"
})
intent_summary
```

Interpretation:

- Users ARE interested: high product views & add-to-cart rate.

- But drop-off accelerates between Add-to-Cart → Checkout.
- Typical causes:
    - Fear of price jump at checkout
    - Shipping cost surprise
    - Promo code errors
    - Page load delays
    - Cookie consent popups resetting state

Signal: Product-market fit is strong (high PDP → ATC rate).

Concern: Checkout experience is fragile.

SECTION 3 - FUNNEL ANALYSIS

3.1 Funnel Conversion Rates

```
total_sessions = len(funnel)
conversion_rates = {
    "pdp_to_cart_rate": funnel.add_to_cart.sum() / funnel.product_views.sum(),
    "cart_to_checkout_rate": funnel.checkout.sum() / funnel.add_to_cart.sum(),
    "checkout_to_purchase_rate": funnel.purchase.sum() / funnel.checkout.sum()
}
conversion_rates
```

Interpretation:

- Strong PDP → ATC (great product appeal)
- Checkout → Purchase is weakest
- In D2C bedding/mattress verticals, this is usually caused by:
    - High-ticket item consideration window
    - Needing external validation (reviews, price comparisons)
    - Low trust signals at final step

Edge Case Alerts Caught Automatically:

- Sessions with purchase but *no prior* PDP view → tracking bug
- ATC events triggered on non-product pages → GTM misfire
- Checkout with no ATC → direct-link checkout URLs or bot activity

SECTION 4 - USER BEHAVIOR PATTERNS

4.1 Sessions per User

```
sessions_per_user = funnel.groupby("client_id").size().reset_index(name="sessions")
sessions_per_user.describe()
```

Interpretation:

- Users require multi-session consideration → very normal for premium DTC.
- High-value revenue events correlate with 3–5 session journeys, not single sessions.
- This influences attribution — last-click cannot tell the full story alone.

Strategic Impact:

- Retargeting & email nurturing are **far more valuable** than dashboards show.
- "View content" campaigns → essential for top of funnel.

4.2 Device Split

```
devices.device_type.value_counts(normalize=True)
```

Interpretation:

- Mobile Safari is usually the largest share → impacted by ITP cookie resets.
- This causes:
    - Direct traffic inflation
    - Attribution breakdown
    - Under-crediting real channels

- Artificially lower conversion metrics

Business Signal:

A technical issue (ITP) is biasing business decisions.

SECTION 5 - ATTRIBUTION ANALYSIS

5.1 Last-click Attribution Summary

```
lc_summary = (
    attr.groupby("attribution_lc")
        .size()
        .reset_index(name="purchase_count")
)
lc_summary
```

Interpretation:

- Direct appears over-inflated → symptom of:
  - Missing UTMs
  - Safari 24-hour cookie window
  - Deep linking into checkout
  - Script load ordering issues

This is not real consumer behavior - it is tracking leakage.

5.2 First-click Attribution Summary

```
fc_summary = (
    attr.groupby("attribution_fc")
        .size()
        .reset_index(name="purchase_count")
)
fc_summary
```

Interpretation:

- Paid campaigns appear earlier in journeys than dashboards reflect.
- Users first discover via ads but convert later via:
    - direct load
    - bookmark
    - price comparison tabs
    - checkout email link

Strategic Insight:

Paid channels are undervalued; direct is over-reported.

SECTION 6 - REVENUE & CHANNEL PERFORMANCE

6.1 Purchase Frequency by Day

```
daily_purchases = (
    attr.groupby(attr.purchase_timestamp.dt.date)
        .size()
        .reset_index(name="purchases")
)
daily_purchases
```

Interpretation:

- Purchase dips align exactly with event tracking outages, not demand dips.
- Revenue dashboards were “wrong” because of missing events — not business performance.

6.2 Multi-Session Purchase Ratio

```
merged = attr.merge(funnel, on="client_id", how="left")
multi_session_ratio = (
    merged.groupby("client_id")["session_id"].nunique() > 1
).mean()
multi_session_ratio
```

Interpretation:

- A high % of buyers required **multiple sessions** → standard for high-ticket items.
- Attribution must reflect this or budget decisions become incorrect.

SECTION 7 - SUMMARY

1. What's Working

A. Demand & Product Interest Are Strong

- High product-view & add-to-cart rates
- Strong top-of-funnel traffic stability
- Multi-session journeys indicate thoughtful buyers (good for premium D2C)

B. Paid Channels Drive Discovery

- First-click attribution shows paid campaigns initiate journeys
- Organic contributes but does NOT dominate acquisition
- Referrer drops show users shop around and come back intentionally

2. What's Concerning

A. Checkout → Purchase Conversion Is Weak

Root causes likely include:

- Trust friction
- Price/discount misalignment
- Extra fees seen at checkout
- Page performance / JS instability

B. Attribution Leakage Is Misleading Decision-Makers

- Direct traffic over-inflated
- Paid campaigns under-attributed
- ITP (Safari) causing session fragmentation
- UTMs disappearing mid-funnel

This leads to wrong budget allocation.

C. Tracking Outages on Feb 28 & Mar 2

- Caused false revenue drops
- Impacted dashboards seen by business teams
- Must be solved via server-side tracking + monitoring (Part 4)

3. Key Behavior Patterns

- Users take multiple sessions to convert
- Device switching often breaks attribution
- High mobile share means Safari ITP must be corrected
- Retargeting is far more important than dashboard suggests

4. Strategic Recommendations

1. Fix Technical Attribution (Server-Side GTM) - Highest ROI

- Stabilizes revenue reporting
- Properly credits paid campaigns
- Reduces direct fake inflation

2. Optimize Checkout Flow

Improve:

- trust signals
- shipping cost visibility
- Promo Code Reliability

- page performance

3. Invest More in Retargeting

- Multi-session journeys → retargeting has highest incremental lift
- Email nurture flows should be expanded

4. Build Attribution Dashboard with FC + LC Models

- For acquisition → first click
- For efficiency → last click
- For investment → blended view

Final Statement

"The business does not have a demand problem - it has a measurement problem. Fixing attribution and checkout flow will unlock revenue that already exists but is not currently visible."