Head of Analytics Assignment

Question 1: E-commerce Funnel Analysis and Insights

1) What We're Solving and Why

- **Goal:** Identify where customers drop off, quantify the value of those leaks, and design experiments to lift conversion and revenue.
- **Why it matters:** Optimizing without understanding traffic quality vs. UX issues risks fixing the wrong thing. Analysis must be measurement-sound: session-level attribution, no duplicate revenue, and tagging sanity.

What We Measure and How:

- **Canonical Funnel (GA4 standard events, strict, time-ordered):** `visit {first_visit | session_start} → view_item → add_to_cart → begin_checkout → purchase`
- **Session grain:** `(user_pseudo_id, ga_session_number)` → avoids cross-device ambiguity.
- **Revenue canon:** `event_value_in_usd` (complete on purchases).
- **Dedup rule:** unique `transaction_id`, keep last by event_time (no duplicates observed).
- **Attribution:** raw `traffic_source_medium/source/name` (cleaned only in presentation).
- **Validation:** `pagetype` sparse; retained for sanity checks only.

📎[Github Repository] 📎[Python pipeline link] 📎[SQL reference]

The Data We Used:

- **Event sequencing:** `event_time`, `event_name`, `event_bundle_sequence_id`
- **Identity/session:** `user_pseudo_id`, `ga_session_number`
- **Engagement:** `param.session_engaged (session-level)`, `param.engaged_session_event (event-level)`
- **Attribution:** `traffic_source_medium`, `traffic_source_source`, `traffic_source_name`

- **E-com specifics:** `transaction_id`, `event_value_in_usd`

Data hygiene checks:

- IDs stable, timestamps ordered.
- Revenue integrity: no inflation, AOV consistent.
- ~29% of purchases missing `begin_checkout` → *tagging gap* or *non-linear paths* (e.g., direct-to-checkout).

2) Key Findings and Insights

- **Overall Funnel (Session-Level) -** *Click to see plot*
  - Sessions: **5,427**
  - Step CRs: visit→view **35.9%**, view→add **16.1%** (*weakest link*), add→checkout **69.0%**, checkout→purchase **29.6%**
  - Visit→Purchase (closed): **1.18%**
  - Revenue: **$124,414** • AOV: **$1,413.80**
- **Daily trend:** Stable conversions across May 31–June 10. Variations match traffic mix (organic/affiliates weaker on weekends). *Click to see plot*
- **Channel snapshot:** ***Strong*:** cpc/google (1.76%), direct (1.57%), ***Weak*:** organic/google (0.54%), affiliates (0.45%) → likely LP mismatch or lower intent. *Click to see plot*
- **Engagement lens:** Engaged sessions: **3.31% CVR vs 1.18% overall (~2.8× higher)** → quality of traffic and on-page engagement matter more than checkout UX. *Click to see plot*
- **Notable anomalies:** June 8/10: lower product-view→cart rates (0.40) driven by organic/affiliate spikes.
- **High-value carts (> $3k):** flow more smoothly to checkout, suggesting premium buyers are less price-sensitive.
- **Measurement sanity:** 29% of purchases missing `begin_checkout`. Likely under-tagging or direct-to-checkout flows. Reported both strict funnel and purchase-after-view to surface risk. *Click to see plot*

3) Executive Summary

- Funnel converts **1.18%**. Major leak at PDP commitment (16%).
- Checkout not primary bottleneck (30% CR, below benchmark but secondary).
- Engagement quality > UX polish. Engaged traffic converts 3× better.
- Channels uneven. Direct/CPC strong, Organic/Affiliates drag down.
- Fix tagging (29% missing begin_checkout).

So what? Focus first on **pre-cart reassurance (fit, delivery, cost transparency)**. Then optimize **checkout trust & ease**, and finally tackle **channel alignment & tagging hygiene**.

4) A/B Test Hypotheses

Assumptions: ~15k sessions/month • AOV = $1,413.80 • Baseline CVR = 1.18% • Baseline revenue ≈ $250k/month
Impact model: $ impact = baseline revenue × relative lift × coverage

Top 3 to Run First (fastest ROI on biggest leaks)

1. **PDP Risk Reversal**

   Why: Biggest leak is view→add (16%); high-AOV buyers need reassurance.
   Test: Add delivery guarantees, “no hidden fees,” return/warranty badges, inline reviews above CTA.
   Impact*: +15–25% **view→add** → +$38k–$63k/mo*

2. **Price Transparency**

   Why: Checkout drop-offs linked to hidden costs.
   Test: Show all-in price (tax/shipping/discounts) on PDP/cart.
   Impact*: +5–10% **checkout→purchase** → +$12k–$25k/mo*

3. **CTA Placement + Social Proof**

   Why: Users miss/add fatigue; reviews not near CTA.
   Test: Move “Add to Cart” above fold, near reviews.
   Impact*: +8–15% **view→add** → +$20k–$38k/mo*

Medium Priority (high ROI, medium effort): Size & Fit Assistant, BNPL badges, Express Wallets, Adaptive Checkout, One-Click Checkout.

Longer-Term Bets (strategic, infra-heavy) - Intent-responsive PDPs, channel-specific LPs.

Bottom Line

Treat **PDP clarity and reassurance as the growth engine**, optimize **acquisition quality**, and **tighten tagging** so wins are measurable and durable. Done right, these changes can lift monthly revenue by **$100k+ without adding new traffic spend**.

Question 2: Analysis of PDP Heatmap

The mobile PDP heatmap reveals top-heavy engagement, with high taps on promos and trust elements indicating deal-seeking and reassurance needs, but sparse activity below the fold suggests scroll fatigue and potential friction. Key patterns show users prioritizing visuals and social proof, yet distractions like competing offers may delay add-to-cart actions. The following table analyzes 7 main areas, linking observations to user behaviors, impacts, and validation methods.

Heatmap Observation	**Behavioral Hypothesis (why users do this)**	**Potential Impact on Add-to-Cart**	**Data to Validate / Measure**
Promo overload: global banner ("$750 off mattress") + hero "15% Off" dominate taps.	Dual offers trigger coupon FOMO and trust confusion; users chase codes instead of the PDP flow.	Negative. Attention hijack → slower time-to-ATC(common 15-20% mobile drop-off per Baymard 2025), higher exits.	**Quant**: Track clicks on promo banners and hero by traffic source; run A/B test with one simple offer on the page; measure view-to-add rate, scroll to CTA percent, and rage clicks. **Qual**: Use pop-up polls asking "Was the offer clear?"; watch session videos for users looping on promos.
Hidden CTA below the fold; variants get taps but no visible next step.	Greedy fold pushes commitment cue out of sight; users customize without seeing a clear action.	High negative. Depresses ATC, increases latency.	**Quant**: Measure CTA show rate (e.g., above fold), sticky add-to-cart views/clicks, time to add-to-cart, and add-to-checkout rate. **Qual**: Watch videos for scroll stops near options; do quick tests with 5 mobile users ("Find the Add to Cart button").
Very few hero taps; those that exist cluster mid-image (face/center), not on arrows.	Users expect to swipe/zoom; center taps = failed gesture (arrows too small/out of thumb zone).	Negative today; upside if fixed. Interactive hero deepens product conviction.	**Quant**: Track swipes, zooms, and carousel moves on hero; compare view-to-add rate and time-to-add before/after changes. **Qual**: Watch videos for failed swipes; do fast user tests on if swipe/zoom feels obvious.
Reviews, FAQs, badges ("Trial, Warranty, Free Shipping") are hot.	Risk reversal + social proof drive pre-cart commitment on high-AOV items.	Positive if adjacent to CTA. Under-leveraged if buried/slow. Lifting add-to-cart by 20-25% if positive reviews (Contentsquare	**Quant**: Track review clicks leading to add-to-cart lift; A/B test moving stars/badges next to CTA; measure page bounce rate. **Qual**: Use short surveys asking "What convinced you?"; watch videos of users seeking trust info.

		benchmark)	
Phone/chat icons see frequent taps.	Shoppers escalate to human reassurance (fit, delivery, returns) when PDP doesn't answer quickly.	Mixed: Saves some conversions via chat but signals PDP gaps, increasing abandonment (10-15% mobile friction per SoftTeco).	**Quant**: Track chat/call clicks to purchase rate; A/B test adding quick answers above CTA; measure add-to-cart change and fewer chats. **Qual**: Review chat logs for common questions; poll after chat "Was your question answered?"
Tabs ("Features/12 Reasons") get clicks; long text blocks are cold.	Mobile users skim, not read; dense copy wastes intent.	Neutral→negative today; Negative: Unread value props hurt consideration-to-add; shorten to bullets for +12% engagement (Nielsen Norman).	**Quant**: Track tab views to add-to-cart rate and scroll depth; A/B test bullets/icons vs. long paragraphs. **Qual**: Do 5-second tests on what users remember from benefits; interview users on how clear it was.
Heavy promo focus may be campaign-driven.	Deal-primed visitors (coupon/affiliate) arrive looking for offers; weak/hidden CTA makes promo the only action.	Negative if they never progress; positive if promo is unified with CTA.	**Quant**: Group users by traffic source; track banner-clickers' add-to-cart rate and revenue per session; A/B test custom landing pages with one promo + CTA. **Qual**: Check if ad messages match the page; poll users "What brought you here today?"

Bottom Line

Simplify promos, surface CTA, bring trust cues forward, enable intuitive hero interaction.

Question 3: Analysis of Exclusive Discount Popup A/B Test Results

1) Funnel Context

- **Typical funnel**: Awareness → Interest → Consideration → Intent → Purchase.
- **Puffy**: site visit → PDP views → add-to-cart → checkout → purchase.
- **Popup placement:** Puffy's popup fires **immediately on arrival** (before product exploration). After an email entry, it **requests a phone number**. This creates a **micro-funnel at the very top (Awareness → Data Capture)** where the very first interaction is about user data, not product discovery.
- **Why it matters:** Early interruption can either prime or derail. Asking both email + SMS pushes users to "just email," reducing SMS capture. SMS subs are **3–5× more valuable**.
- **How to judge success:** Incremental RPV & RPS (email vs. SMS) over 14–30 days, not raw signups.

2) Key Metrics from the Data

- **Signups:**
  - **Email-only:** Control 1.56% → Variant 1.86% (**+18.8%**). Not significant (p≈0.24). 17k more samples needed.
  - **Email + Phone:** Control 0.86% → Variant 0.58% (**−31.8%**). Not significant (p≈0.40), but the drop is large.
  - **So what:** More emails, but **fewer SMS signups** (higher value). Net: funnel tilts toward lower-value leads.
- **Conversions:**
  - Unique CR flat (1.55% → 1.54%).
  - Total CR flat/slight up (1.71% → 1.75%).
  - **So what:** Popup does not create more buyers.
- **Monetization:**
  - **RPV:** $23.29 → $21.24 (−8.8%).
  - **AOV:** $1,421.60 → $1,345.83 (−5.3%).
  - **So what:** Visitors with the popup are **less valuable per order and per visit**—consistent with coupon-driven browsing.
- **Behavioral hints:**
  - Add-to-Cart: −5.7% (down).
  - Buy-Now: +2.4% (slight up).
  - Checkout: +1.3% (flat).

- **So what:** Popup may distract some shoppers at PDP, without offsetting lift later.

Net Insight: ✅ More emails, ❌ fewer SMS, ↔ conversions, ❌ RPV & AOV down. More leads ≠ more money.

3) Supporting Data Needs

1. **Post-Signup Monetization:** Track 14–90 day revenue per signup, open/click rates, unsubscribes. Clarifies if extra emails convert long term.

2. **Margin & Discount Impact:** % of orders with popup discount codes, average discount depth, margin-adjusted RPV/AOV.

3. **Segmentation:** Break down results by device, traffic source, and visitor type (new vs returning).

Critical gap: Without lifecycle monetization, we can't tell if these new signups are incremental or just "coupon hunters."

4) Final Recommendation

Do not roll out the mystery discount popup as-is.
*With wide CIs and non-significant flags, we shouldn't claim certainty—but the **direction** of the money metrics is unfavorable.*

Rationale:
- **Funnel fit:** Grows the top (emails) but weakens SMS capture (higher value).
- **Conversion impact:** No lift at checkout or purchase.
- **Money signals:** RPV and AOV both fall → visitors are less valuable.
- **Risk:** Trading quality (SMS) for quantity (emails).

Immediate Next Actions:
1. **Refine popup timing/design:** Test fixed discount vs. mystery. Trigger on exit-intent or scroll, not instantly.
2. **Protect SMS capture:** Use progressive capture (email first, then SMS later).
3. **Reduce interruption:** Frequency cap (1/7 days) and suppress popup for existing subscribers. Avoid "infinite discount hunting" loops that erode brand trust.

Bottom Line

The mystery popup is a risky volume play. It delivers more signups but at the cost of fewer high-value SMS, lower RPV, and lower AOV. Unless long-term revenue proves otherwise, Puffy risks diluting customer value instead of growing it.

Question 4: Process & AI Usage

How I worked.
I approached each question as a business problem, not just a data exercise. I clarified funnel logic (session vs. user, revenue de-dupe), ran the numbers, and overlaid them with behavioral and psychological insights. Every output was shaped as if going to a board deck—clear, prioritized, and tied to revenue impact.

Where AI helped.

- **Idea generation:** Brainstormed leaks, heatmap behaviors, A/B test angles.
- **Code & analysis:** Drafted Python/SQL; I validated manually.
- **Refinement:** Compressed long analysis into <500 words exec summaries.
- **Reporting:** Automated charts, markdown, PDFs.

Where I took the lead.

- Decided measurement approach (eg. session vs. user, RPV vs. signup count).
- Interpreted tricky signals (e.g., higher signups but lower RPV, slides on face for heatmap, etc).
- Prioritized hypotheses by expected revenue lift, not novelty.
- Flagged missing pieces (margin-adjusted RPV, long-term monetization) that AI couldn't infer.
- **Did manual QA**: tested the site myself and found the discount popup bug—still showing after coming from a discount email. This risks confusing users, creating a "**greed loop,**" and eroding brand trust.

Why it matters.
I don't use AI as an "answer machine." I use it as a copilot—fast for code, structure, and drafts—while I bring strategy, judgment, and context. The result: analysis that's not just technically sound, but business-ready and actionable.

Tools - Openai, Grok, Claude, Julius.ai, gamma.