

The Holiday Game Plan

Prepared by: Linh Le

Externship: Breaking Games Analytics & Database Design Externship

Submitted: [Date]

How to use this document. Make a copy of this template at the start of Week 1. Each week, you'll add one section. By the end of the externship, this document — plus your HTML dashboard — is what you present to the Breaking Games team.

Don't worry about polish on day one. Some sections will be empty until later weeks. That's fine.

Note: This is one running document. You'll add to it a section at a time and paste the same link whenever a task asks for it.

Executive Summary

Fill this in last (during the capstone). Three to five sentences. The one-page version of "what did we find and what should we do."

[Your executive summary goes here.]

Week 1: Getting Into Breaking Games' Raw Data

1.1 Data inventory

For each of the six data sources, write what's in it, what questions it can answer, and what it can't answer alone.

File	What it contains
shopify_sales	What it has: One row per product (99 total) with units sold (net_items_sold) and the money columns: gross sales, discounts, returns, net sales, taxes, total sales. Also vendor and product type. Can answer: Best sellers by units or revenue, how big the catalog is, which products get discounted or returned a lot. Can't answer alone: Who bought, when, or where they came from. It's a lifetime total per product with no dates, so it can't show trends.

File	What it contains
shopify_orders	What it has: One row per order line (254 rows, 93 orders), July 1 to 29, 2025. Has order name, product, gross and net sales, discounts, shipping, taxes. 92 rows have no product title. Can answer: Daily sales, what an order looks like, shipping charges, and AOV for that month. Can't answer alone: Anything outside July 2025, or who the customer was. The blank product rows need cleaning before they're trusted, and it's only a month, so it can't back up full-year claims.
shopify_checkouts	What it has: One row per line item in a checkout (1,367 rows), not per session. Has customer, cart contents, subtotal, shipping, discount code, location, payment status, risk level, and a created_at timestamp. Can answer: How many checkouts started (1,062 once

File	What it contains
	you count distinct name), what's in abandoned carts, and when and where carts get started. Can't answer alone: Whether a checkout turned into an order, unless you match it to orders. COUNT(*) over-counts because of the line items, and some dates look off (the earliest is 2016), so check before trusting any time trend.
referrals	What it has: Traffic by referrer and landing page (8,712 rows): pageviews, add-to-cart rate, bounce rate, conversion rate, checkout sessions, plus the same numbers for the previous year. Can answer: Where visitors come from (Google, Facebook, Instagram, BoardGameGeek), which pages pull traffic, and how it compares to last year. Can't answer alone: Which ad drove a visit, or what a visitor bought. About a third

File	What it contains
	of rows have no referrer (direct or unknown), and it doesn't connect to products without a join on landing page.
meta_campaigns	What it has: One row per Meta (Facebook/Instagram) ad campaign (35 rows), Jan 1 to Jul 29, 2025: spend, reach, impressions, clicks, CTR, CPC, results, and cost per result. Total spend is \$37,200.94. Can answer: How much was spent, which campaigns were cheapest per result, and what's still running. Can't answer alone: Whether the ad money paid off. It has no revenue, and 29 of 35 campaigns have no results logged, so ROI needs sales data joined in.
meta_ad_sets	What it has: One row per ad set (8 rows), same columns as the campaigns table: spend, reach, clicks, results, cost per result. Covers targeting like Broad,

File	What it contains
	Interests, and Retargeting. Can answer: Which audience or targeting approach worked best (retargeting got the most spend, about \$5K). Can't answer alone: Which specific ad or product it sold. It's a different cut of the same ad spend, so don't add it to the campaigns table, or you'll count the money twice.

1.2 Five instant answers

The questions Mike asks every Monday morning. Paste the SQL query and the answer for each.

Q1: Top 10 products by units sold this period.

```
SELECT product_title, net_items_sold
FROM shopify_sales
ORDER BY net_items_sold DESC
LIMIT 10;
```

Answer: Top 10 below.

1. Dwellings of Eldervale - Oracle Hex Tile: 178 units
2. Dwellings of Eldervale - LEGENDARY Upgrade Kit: 175 units
3. Dwellings of Eldervale - Minotaur Mercenary Mini Expansion: 135 units
4. Dwellings of Eldervale - Shapeshifter Mercenary Mini Expansion: 134 units
5. Dwellings of Eldervale 2nd Edition: Standard: 129 units
6. Dwellings of Eldervale - Wooden Upgrade Treasure and Trove Tokens: 127 units

7. We're Doomed!: 104 units
8. Dwellings of Eldervale - DELUXE Upgrade Kit: 89 units
9. MIXTAPE: 83 units
10. The Big Box of Magic: 46 units

Reaction: Surprised me a little. Nearly the whole list is Dwellings of Eldervale add-ons, and the Oracle Hex Tile outsold the actual base game.

Q2: Top 10 products by total revenue.

```
SELECT product_title, total_sales
FROM shopify_sales
ORDER BY total_sales DESC
LIMIT 10;
```

Answer: Top 10 below.

1. Dwellings of Eldervale 2nd Edition: Standard: \$13,323.56
2. Dwellings of Eldervale - LEGENDARY Upgrade Kit: \$12,514.05
3. Dwellings of Eldervale - DELUXE Upgrade Kit: \$5,014.81
4. We're Doomed!: \$3,810.47
5. Dwellings of Eldervale - Wooden Upgrade Treasure and Trove Tokens: \$2,586.22
6. The Big Box of Magic: \$1,920.14
7. Dwellings of Eldervale - Exclusive Game Mat: \$1,874.13
8. MIXTAPE: \$1,733.39
9. Dwellings of Eldervale - Minotaur Mercenary Mini Expansion: \$1,370.38
10. Dwellings of Eldervale - Shapeshifter Mercenary Mini Expansion: \$1,361.17

Reaction: Different leader than the units list. The 2nd Edition base game wins on money at \$13,323.56 but sits at #5 on units, so cheap add-ons win on volume and the big boxes win on dollars.

Q3: How many products are in the catalog?

```
SELECT COUNT(*)
FROM shopify_sales;
```

Answer: 99 products

Reaction: No surprise. It matches the 99 rows I've seen since my first SELECT *.

Q4: What was the total Meta ad spend?

```
SELECT SUM(spend_usd)
FROM meta_campaigns;
```

Answer: \$37,200.94

Reaction: Matched what I expected: roughly \$37K, right where Mike's been quoting the ad budget.

Q5: How many unique checkout sessions started?

```
SELECT COUNT(DISTINCT name)
FROM shopify_checkouts;
```

Answer: 1,062 unique checkout sessions (COUNT(*) gives 1,367, but that's line items, not sessions)

Reaction: This one surprised me. The obvious COUNT(*) over-counts by about 300 because one checkout can hold several line items, so counting distinct names is what gives the real number.

Week 2 — Designing Your First Database

2.1 Schema diagram

Paste or sketch your schema after building dim_product. Show the dimension table and which fact tables now connect to it via product_id.

[Diagram or ASCII sketch here.]

2.2 The dim_product design decision

A short paragraph: what columns did you choose for dim_product, and why?

[Your reasoning]

[Screenshot of the final JOIN query from Step 5 and the top 5 results]

2.3 Connect the journey

Three queries that connect traffic to sales across the catalog. Paste each query and a one-line summary of what it returned.

Query 1 — Products with sales but mysterious origins. Products with sales rows but no referral traffic matching their slug.

[your query]

Result: [how many products, plus one or two examples]

Query 2 — Products with traffic but no sales. Traffic comes in, nobody buys.

[your query]

Result: [how many products, plus one or two examples]

Query 3 — The full story. Every product with `total_sales`, `total_pageviews`, and a pageviews-per-unit-sold ratio.

[your query]

Result: [paste your top 10 rows]

What you found — *a short paragraph covering:*

- Which products had high traffic but no sales?
- Which products had sales you couldn't trace to a source?
- What's the most interesting finding?

[Your paragraph.]

Week 3: What's Actually Selling

3.1 Product performance leaderboard

Query used:

[\[Add the query\]](#)

Top 25 products by revenue, with total sales, net items sold, and AOV for each.

Rank	Product	Total Sales	Net Items Sold	AOV
1				
...				
25				

3.2 Reorder recommendation

For Mike's November 1 reorder commit, your call:

[\[Add the summary\]](#)

Print heavy — reorder confidently (~10 products):

- [Product 1] — why: [your reason]
- [Product 2] — why: [your reason]
- ...

Print light — modest reorder (~5 products):

- [Product 1] — why: [your reason]
- ...

Don't print — let it go (~3 products):

- [Product 1] — why: [your reason]
 - ...
-

Week 4 — Following the Money

4.1 Channel breakdown

Where did the \$37,000 in Meta spend actually drive traffic? Where are customers coming from organically?

Channel	Pageviews	Checkout sessions	Conversion rate	Spend	Cost per pageview	Cost per checkout session
Paid Social (Meta)						
Organic						

Search						
Direct						
Community /Referral						
Email						
Other						

Key finding: [one sentence — e.g., “BoardGameGeek converts at about 2.9x Facebook’s rate on 54x less traffic, worth investing in community partnerships.”]

4.2 Meta campaign ROI

Top 5 best-performing campaigns:

Campaign	Spend	Results	Cost per result	CTR	CPC

Bottom 5 (the ones to question):

Campaign	Spend	Results	Why we should reconsider

4.3 Q4 ad allocation plan

Continue running (\$X total):

- [Campaign 1] — why: [your reason]
- ...

Cut or pause (\$X freed up):

- [Campaign 1] — why: [your reason]
- ...

New investment (where the freed dollars go):

- [Channel/strategy] — why: [your reason]

4.4 The uncomfortable finding

One paragraph: what's a result here that Mike won't want to hear, and how do you back it up with data?

[Your honest read.]

Week 5: Carts in the Wild

5.1 The real abandonment rate

Metric	Value
Unfinished checkout sessions in the export (2016-2025)	
Completed orders (July 2025, from <code>shopify_orders</code>)	
Naive abandonment rate (mismatched windows)	
Carts flagged as suspicious (your criteria)	

- **Time of day:** [pattern you noticed]
- **Day of week:** [pattern you noticed]
- **Cart size:** [pattern you noticed]
- **Discount usage:** [pattern you noticed]
- **Geography:** [pattern you noticed]

What looks off: The first thing I noticed that looks off is [X]. I'll investigate it as a bot/automation candidate in Step 3.

5.4 What Breaking Games can actually fix

Three concrete recommendations the team could act on before Q4:

1. [Recommendation 1] — backed by: [data point]
 2. [Recommendation 2] — backed by: [data point]
 3. [Recommendation 3] — backed by: [data point]
-

Capstone Project — The Holiday Game Plan

6.1 The dashboard

Paste a screenshot of your final HTML dashboard here. (Open [dashboard.html](#) in a browser → take a screenshot → drag in.)

[Dashboard screenshot]

6.2 The five numbers

#	KPI	Value	What it tells the team
1			
2			
3			
4			
5			

6.3 Key findings (three to five)

Each one: one sentence finding + one sentence "so what."

1. **Finding:** [your finding] **So what:** [the action it implies]
2. **Finding:** [your finding] **So what:** [the action it implies]
3. **Finding:** [your finding] **So what:** [the action it implies]
4. **Finding:** [your finding] **So what:** [the action it implies]
5. **Finding:** [your finding] **So what:** [the action it implies]

6.4 Recommendations to Breaking Games

The three things the team should do before Q4 — backed by your analysis.

1. [Recommendation]

- **Rationale:** [what the data shows]
- **Owner:** [who should drive it — Mike, Jess, ops, etc.]
- **Expected impact:** [your estimate]

2. [Recommendation]

- **Rationale:**

- **Owner:**
- **Expected impact:**

3. [Recommendation]

- **Rationale:**
- **Owner:**
- **Expected impact:**

6.5 What you'd need next

If you had another four weeks, what data or analysis would unlock the most additional value?

[Your wishlist.]