Walmart’s Hurricane Data: Pop-Tarts,

DS2002 Jason Williamson

Florida, 2004

It was the evening before Hurricane Frances made landfall in Florida. The aisles of a Walmart Supercenter in Orlando buzzed with anxious shoppers. Parents pushed carts stacked with bottled water, flashlights, and batteries. But what surprised the store manager most were the empty shelves in the breakfast aisle. By 9 p.m., the entire stock of Strawberry Pop-Tarts had vanished. “We sold out faster than flashlights,” he muttered, shaking his head. He didn’t know it then, but his store was part of a pattern that Walmart’s data scientists in Bentonville were about to uncover—and elevate into company legend.

Introduction: The Storm Approaches

On a humid September morning in Bentonville, Arkansas, the data science team at Walmart was already buzzing with activity. Hurricane Frances—a Category 3 storm—was forecasted to make landfall in Florida in less than a week. Executives knew the drill: panic buying would sweep the state, store shelves would empty, and supply chains would be stretched to their limits. But this time, Walmart had a weapon it hadn’t possessed a decade earlier: the ability to mine billions of rows of transaction data to anticipate customer needs.

Inside the company’s Operations Center, rows of screens displayed weather models, sales dashboards, and supply chain alerts. Managers asked the same question over and over: *What exactly will people want this time?* The difference between guessing and knowing could mean millions in revenue—and, more importantly, whether Walmart delivered for its customers when they needed it most.

The Data Landscape

By 2004, Walmart had already become legendary for its mastery of information systems. Its **Retail Link** platform gave suppliers direct access to sales data, while the company’s data warehouse—then among the largest in the world—captured terabytes of point-of-sale (POS)

records. Every item scanned at checkout across thousands of stores was logged, along with store IDs, timestamps, and payment information.

Yet the sheer scale of data wasn't Walmart's biggest challenge. The problem was **messiness**:

- **Fragmented SKUs:** Strawberry Pop-Tarts alone appeared under multiple codes depending on packaging, supplier, or store region.
- **Incomplete Data:** Registers sometimes dropped timestamps or mis-logged locations.
- **External Feeds:** Weather data, critical for storm planning, arrived in inconsistent formats and time zones.

An analyst described the problem bluntly: *"We're not drowning in water yet—we're drowning in data."*

(See Exhibit 1 for a sample of raw transaction data.)

ETL: Wrestling Order from Chaos

To act quickly, Walmart leaned on its **ETL pipelines**—Extract, Transform, and Load. The sequence was simple in theory but grueling in execution:

1. **Extract** – Data poured in from store registers, distribution centers, and weather services. This raw input formed the foundation.
2. **Transform** – Here, analysts fought the real battle:
    - Duplicates were removed.
    - SKU codes were standardized across regions.
    - Missing timestamps were imputed or flagged.
    - Weather alerts were aligned with store-level sales data.

    One junior analyst admitted, *"Half of my job feels like detective work—figuring out if PT-12 and SKU#459812 are the same Pop-Tart or two different ones."*

3. **Load** – Cleaned data was loaded into Walmart's enterprise warehouse, making it usable for querying and analysis.

Each decision carried consequences. Should all Pop-Tarts SKUs be aggregated as one category, or treated separately by packaging? Would an error in timestamp alignment skew the correlation between weather alerts and spikes in sales? The integrity of every later insight depended on these choices.

(See Exhibit 5 for a schematic of Walmart's ETL pipeline.)

The Surprising Insight

When the team finally ran its queries, the expected patterns emerged. Bottled water, batteries, flashlights—all surged as the storm drew near. But one chart made the analysts stop and laugh: **Strawberry Pop-Tarts sales increased sevenfold before hurricanes.**

"Wait—Pop-Tarts?" one senior analyst said, pointing at the screen. "That's… not what I expected."

Another replied, *"Well, they're cheap, non-perishable, and you don't even need a toaster. Makes sense when you think about it."*

Beer sales also spiked, but the Pop-Tarts discovery was the outlier that proved invaluable. Regardless of the why, the data was clear: people wanted Pop-Tarts.

Within hours, Walmart's supply chain managers began routing trucks to Florida stores, ensuring shelves would be stocked with both essentials and the now-famous toaster pastries.

(See Exhibit 2 for pre-hurricane sales trends.)

From Data to Business Impact

The decision paid off:

- **Customer Satisfaction:** Shoppers found the products they craved, reducing frustration.
- **Supply Chain Efficiency:** Strategic stocking reduced last-minute emergency shipments.
- **Financial Gains:** Sales increased without significant waste, protecting margins.
- **Cultural Legacy:** Inside Walmart, the "Pop-Tarts story" became folklore—a symbol of data science's power to reveal non-obvious truths.

The case demonstrated not only the value of advanced analytics but also the critical importance of **trustworthy data pipelines.** Without careful wrangling, Pop-Tarts could have remained hidden in the noise—or worse, a flawed ETL process might have falsely inflated their demand.

Scaling the Model

As storms became more frequent and severe in the 2010s, Walmart executives faced new questions. Could the Pop-Tarts insight be replicated across other regions? Were there equally hidden gems in the data—products whose demand spiked under unusual conditions? Could Walmart use machine learning models to predict pre-storm demand patterns nationwide, or should it remain cautious, responding to each storm as a unique case?

During one strategy session, voices clashed:

- **Operations Chief:** *"We can't risk empty shelves. If the data says Pop-Tarts, we ship Pop-Tarts. Customers remember who delivered."*
- **Finance VP:** *"But if we overstock and the storm shifts, we're stuck with pallets of unsold product. That's millions in waste."*
- **Chief Data Officer:** *"The risk isn't just the model—it's the pipeline. If our ETL misclassifies products, we'll make the wrong call. We need cleaner wrangling before fancier algorithms."*
- **CEO:** *"Our reputation is on the line. When families come to us in a crisis, are we stocking what they truly need—or gambling on anomalies?"*

The debate underscored a new tension: how to balance quirky insights with enterprise-wide strategy.

What now???

The room grew quiet as the latest dashboard came up. Another storm was forming in the Gulf. Alongside the familiar spikes in water, batteries, and Pop-Tarts, a new anomaly appeared: **board game sales were rising in storm-affected regions.**

One analyst smiled: *"Makes sense. Families stuck at home need something to do besides watch the rain."*
The Finance VP shook his head: *"Or it's a data artifact. Maybe Monopoly and 'card game bundles' got lumped together. We'd be shipping cardboard for nothing."*

The CEO looked around the table. "We've got 48 hours to decide. Do we bet on board games—or not?"

Walmart’s Hurricane Data: Pop-Tarts, Preparation, and the Power of Data

---

Exhibit 1: Walmart Transaction Data Sample (Simplified)

Transaction ID	Store ID	Timestamp	SKU Code	Product Name	Quantity	Notes
TXN001	FL-239	2023-09-12 08:15:00	PT-12	Pop-Tarts Strawberry	3	SKU variant by region
TXN002	FL-239	2023-09-12 08:16:00	SKU#459812	POPTART-STRAWBERRY	1	Different SKU
TXN003	FL-105	2023-09-12 08:17:00	BATT-100	AA Batteries (4-pack)	5	Core emergency item
TXN004	FL-239	2023-09-12 08:19:00	BEER-24	Light Beer (24-pack)	2	Sales spike observed
TXN005	FL-330	2023-09-12 08:25:00	POPTART-STRAWBERRY	Pop-Tarts Strawberry	2	Inconsistent naming
TXN006	FL-239	NULL	PT-12	Pop-Tarts Strawberry	4	Missing timestamp

---

Exhibit 2: Pre-Hurricane Sales Trend (Selected Categories)

Average daily unit sales across Florida stores (baseline = normal week; hurricane = week before landfall)

Product Category	Baseline Sales (Units/Day)	Hurricane Sales (Units/Day)	% Increase
Bottled Water	12,000	48,000	+300%
Batteries	8,500	34,000	+300%
Flashlights	1,200	4,800	+300%
Beer	6,000	18,000	+200%

Product Category	Baseline Sales (Units/Day)	Hurricane Sales (Units/Day)	% Increase
Pop-Tarts (Strawberry)	2,000	14,000	**+600%**

Exhibit 3: Weather Feed (Simplified Extract)

Date	Time	Location	Alert Level	Notes
2023-09-12	08:00	Florida	Advisory	Tropical storm watch issued
2023-09-13	14:00	Florida	Warning	Hurricane upgraded to Category 2
2023-09-14	06:00	Florida	Emergency	Hurricane upgraded to Category 3

Integration Issue: Weather feed timestamps logged in GMT; sales data logged in EST. Misalignment required correction.

Exhibit 4: Supply Chain Constraints (Simplified)

Distribution Center	Avg. Daily Shipments	Surge Capacity (Max)	Notes
Jacksonville, FL	500	1,000	Nearest to storm region
Atlanta, GA	400	700	Secondary supply option
Dallas, TX	300	600	Long-haul, higher cost

Exhibit 5: Data Pipeline Workflow (ETL)

```
graph TD
A[Raw POS Data] --> B[Extract]
B --> C[Transform - Deduplicate]
C --> D[Transform - Standardize SKUs]
D --> E[Transform - Align Timestamps]
E --> F[Transform - Join with Weather Data]
F --> G[Load into Data Warehouse]
G --> H[Analytics Dashboard]
H --> I[Business Decision: Stock Pop-Tarts]
```

Exhibit 6: Emerging Pattern – Board Games Sales

Product Category	Baseline Sales (Units/Day)	Hurricane Sales (Units/Day)	% Increase
Board Games	1,000	3,500	+250%

Product Category	Baseline Sales (Units/Day)	Hurricane Sales (Units/Day)	% Increase
Playing Cards	800	2,200	+175%
Streaming Devices	500	1,400	+180%

Question: Are these signals genuine, or artifacts of mis-coded SKUs and data lag?
