

Case Study Discussion

Sales Forecasting — Nigerian Retail

What this project does: Predicts how much money a retail chain will make each day.

Data used: 800,000 sales receipts from shops across 15 cities (January – October 2024)

1. Business Context

Running a retail chain in Nigeria is hard work. Every day, hundreds of thousands of customers walk into stores and make purchases — but the amount of money coming in each day is never exactly the same. Some days are busier, some are slower.

The problem is: **if you don't know in advance how much money is coming in, you make bad decisions.**

For example:

- You might order too many products, and some go to waste
- You might have too many or too few cashiers working on a given day
- You might run out of cash at the register when customers need change
- You might overspend on a marketing campaign without knowing if it actually helped

This project was built to solve that. The goal was to look at past sales history and use it to accurately predict what sales will look like tomorrow, next week, or next month — so the business can plan ahead with confidence.

2. Problem Statement

The core question this project answers is simple:

“Based on how much we sold in the past, how much will we sell tomorrow?”

The challenge is that you can't just guess. The prediction needs to be accurate enough to act on — whether that means ordering stock, scheduling staff, or setting a cash budget.

This project used 10 months of real sales data from a retail chain to build a system that can answer that question reliably.

The target: Get predictions that are wrong by less than 10% on average — good enough to make real decisions.

3. Data Understanding

The dataset contains **800,000 individual sales receipts** collected from shops across Nigeria.

What	Details
Number of receipts	800,000
Time period covered	January 1 – October 27, 2024
Cities covered	15 cities across Nigeria
Amount per receipt	₦1,000 – ₦200,000
Total sales per day	₦12.6 million – ₦25.5 million (average: ₦19.1 million)

Each receipt tells us: which store it came from, which city, what date, how many items were bought, how much was paid, and whether the customer paid with cash, card, or mobile money.

Key observations from the data:

- Daily sales are fairly steady — no wild spikes or crashes
- Half of all payments (50%) are made in cash
- 40% of customers pay by card, and 10% use mobile money

4. How the Prediction Was Built

4.1 Step 1: Simplify the Data

Instead of looking at 800,000 individual receipts, the project grouped them by day — adding up all the money made on each day to get a single daily total. This gave a clean picture of how revenue moved day by day.

4.2 Step 2: Find Useful Patterns

To predict tomorrow's sales, the system was taught to look at useful clues from the past:

- **What day of the week is it?** (Monday, Tuesday, etc.)
- **What month is it?**
- **What week of the year is it?**
- **Is it a weekend?**
- **How much did we make exactly 7 days ago?** (same weekday last week)
- **How much did we make 14 days ago?**
- **What was the average sales over the last 7 days?**

These clues help the system understand patterns — for example, if sales tend to be higher in certain weeks, or if last week's numbers are a good guide for this week.

4.3 Step 3: Train a Prediction Engine

A prediction tool called **XGBoost** was used. Think of it like a very experienced analyst who has studied thousands of past scenarios and learned the patterns well enough to make

good guesses about the future. It was trained on the first 8 months of data, then tested on the last 2 months to see how accurate it was.

4.4 Step 4: Test It on Real Data It Had Never Seen

The system was tested against actual sales figures it had never seen before. This is important — like giving a student an exam they haven't seen the answers to. Only then do you know how good the predictions really are.

5. How Accurate Is It?

Measure	Result	Plain English
Average daily error	N1,359,537	On a typical day, the prediction is off by about N1.4 million
Worst-case error	N1,717,428	Occasionally, predictions miss by up to N1.7 million
Accuracy rate	92.82%	The model is correct to within 7.18% on average

What does 7.18% error mean in practice?

On a day when the store makes N19 million, the prediction would typically land somewhere between N17.6M and N20.4M. That's tight enough to:

- Decide how much stock to order
- Plan how many staff to schedule
- Set how much cash should be available at each branch

The target was to stay under a 10% error rate. This system comfortably achieves that at **7.18%**, meaning the predictions are reliable enough to base real business decisions on.

6. What We Learned

6.1 About the Business

1. **Sales are steady throughout the week.** There is no “busy weekend” or “slow Monday” pattern. Revenue is consistent every day of the week — which means you don’t need to staff up just because it’s Friday or Saturday.
2. **Daily revenue is predictable.** Sales stay within a range of N12.6M to N25.5M with no sudden drops or spikes. This is actually good news — it means forecasting works well here.
3. **Most customers still pay with cash.** 50% of all sales are cash transactions. This means the business needs to always have enough physical cash available — especially during busy periods.
4. **Digital payments are growing.** 40% card + 10% mobile money shows that customers are slowly shifting away from cash. This trend is worth watching.

6.2 About the Predictions

1. **The best clue about tomorrow is what happened last week.** Looking back 7 and 14 days gives the system strong signals about what to expect. History repeats itself in retail.
2. **Weekly position matters a lot.** Which week of the year it is turns out to be the strongest single signal — suggesting there are slow and busy periods spread across the calendar.
3. **Weekends make no meaningful difference.** Whether it's Saturday or Sunday has zero effect on how much the store makes. Sales don't spike on weekends in this dataset.

7. How This Helps the Business

Ordering Stock

If you know Tuesday will likely bring in N20M in sales, you can estimate how many products will be needed and order accordingly — reducing waste and avoiding empty shelves.

Scheduling Staff

If predictions show a slow week coming up, you can roster fewer staff. If a high-revenue period is forecast, you bring in extra hands. This saves money and improves customer service.

Managing Cash

Half the money coming in is physical cash. Knowing in advance when high-cash days are coming allows branches to request cash top-ups, avoiding a situation where tills run dry.

Measuring Campaigns

If a promotion runs during a week where the model predicted N19M but you actually made N24M, you can confidently say the campaign drove that N5M difference — rather than guessing.

8. Where This Falls Short

1. **It doesn't know about special events.** Public holidays, Eid, Christmas, new term shopping — none of these were included. So the model might be off during those periods.
2. **It only looks at the whole chain, not individual stores.** The prediction is for total national revenue, not for Lagos or Abuja separately. Store managers can't use this directly for their own branches yet.
3. **Ten months of data isn't enough to capture yearly patterns.** To truly understand annual cycles (e.g., pre-Christmas spikes), you'd need at least 2–3 years of data.
4. **It doesn't give a range, only a single number.** A manager might want to know “we expect N19M, but it could be anywhere from N17M to N21M.” Right now, the system only gives one number, not a confidence range.

9. What Should Happen Next

9.1 Make the Predictions Better

- Feed in public holidays so the model knows when unusual sales patterns are expected
- Include data about promotions and discount campaigns
- Add city-level economic signals (e.g., fuel prices, local events)

9.2 Make It More Useful

- Build separate forecasts for each city or store — so regional managers can actually act on the numbers
- Create a simple dashboard where non-technical staff can view tomorrow's revenue forecast without needing to run any code

9.3 Make It an Ongoing Tool

- Set it up to automatically update every day with new sales data
- Check the accuracy monthly and retrain if it starts drifting
- Send alerts to management if predictions are consistently far off

10. Conclusion

This project started with 800,000 raw sales receipts and turned them into a working prediction system that can tell a retail chain what to expect in daily revenue — with a 92.82% accuracy rate.

The real value isn't the technology. The value is that **managers no longer have to guess**. They can plan stock orders, staff rosters, cash reserves, and marketing campaigns based on actual predicted numbers — not instinct.

With some refinements (especially adding holiday data and store-level breakdowns), this system could become a core part of how the business plans and operates every single day.

Document prepared for portfolio presentation and business case discussions.