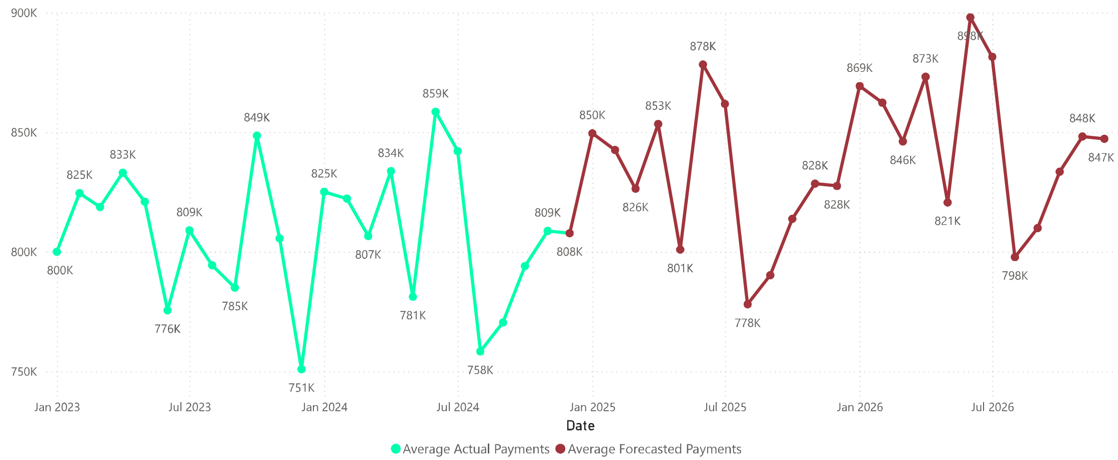
Retail Banking Forecast Suite

Aziz Akhtar

Average Payment Volume	Average Loan Apps	Average Deposit Balance	Average Churn Rate	Average App Logins	Average Loss Rate
$807K	300	5M	17.43%	1.20K	9.57%

Average Monthly Payment per Account – 24-Month Actual & 24-Month SARIMA Forecast



SARIMA FORECASTINC CHART

- **The first trio** (1, 1, 1) are the **non-seasonal** parts of the model, often written as ARIMA(p,d,q):
  - **p = 1** → "AR(1)" means we include last month's value in predicting this month.
  - **d = 1** → "I(1)" means we difference the series once (we look at month-to-month changes rather than raw levels).
  - **q = 1** → "MA(1)" means we include last month's forecast error (the amount we missed by) to correct this month's prediction.
- **The second quartet** (0, 1, 1, 12) are the **seasonal** parts, written as (P,D,Q,s):
  - **P = 0** → no seasonal autoregressive term (we're not leaning on the value from 12 months ago).
  - **D = 1** → one seasonal difference (we compare this month to the same month last year).
  - **Q = 1** → one seasonal moving-average term (we correct based on the forecast error from the same month last year).
  - **s = 12** → the season length is 12 months (i.e. yearly seasonality).

So altogether **SARIMAX(1,1,1)x(0,1,1,12)** means:
"Forecast using last month's level (AR1), last month's error (MA1), after removing month-to-month change (I1), and also remove yearly change (seasonal D1) plus correct by last year's error (seasonal MA1), with a 12-month seasonal cycle."

Dep. Variable: Forecasted_Payment_Volume
No. Observations: 24
Model: SARIMAX(1, 1, 1)x(0, 1, 1, 12)
Log Likelihood: 0.000
AIC: 8.000
BIC: nan

ar.L1 = 0.13
0.1273
"We lean about 13% on last month's actual average."
ma.L1 = –0.60
–0.6008
"We subtract 60% of last month's error to correct us."
seasonal MA = 0
0
"Skipping the 'same month last year' check."
sigma² ≈ 3.65 × 10⁹
3.65 billion
"size of the noise, how often data spikes"

Log Likelihood = 0: Just a technical score of fit; here with tiny data it doesn't mean much.
AIC = 8: A cheat-score that says "low number means simple yet OK fit."
BIC = NaN: We didn't have enough data for this test, so ignore it.

Fit Metrics simplified:

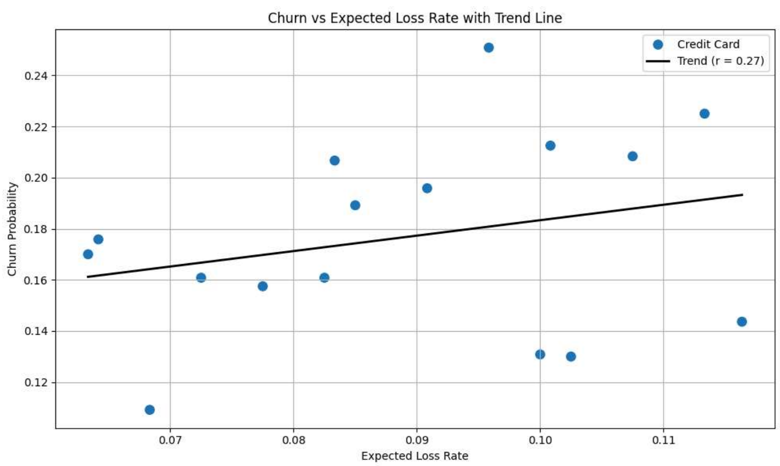
1. **I told the model**: "Use last month's average, plus fix 60% of last month's mistake. Don't bother with last year's same month."
2. **The model found** that leaning a bit on last month and correcting for most of last month's error gave the best simple forecast.
3. **Due to small sample data set (2 years)** some of the usual check-ups (like BIC or confidence tests) aren't reliable—so you'll see weird "NaN" or infinite values in those spots.

Bottom line: Our forecast line (the red one) is just you + 13% of last month – 60% of last error, month after month. It's simple but works OK for short runs when you don't have tons of data.

Retail Banking Forecast Suite

Aziz Akhtar

Average Payment Volume	Average Loan Apps	Average Deposit Balance	Average Churn Rate	Average App Logins	Average Loss Rate
$807K	300	5M	17.43%	1.20K	9.57%



Product Filter - Graph Only

Multiple selections

```
# from your linregress call
slope     ≈ 0.25    # how steep the line is
intercept ≈ 0.135   # where the line crosses the Y-axis
r_value   ≈ 0.21    # correlation coefficient
p_value   ≈ 0.08    # probability this slope could be zero
stderr    ≈ 0.10    # uncertainty on the slope
n_points  = len(grouped)
r_squared = r_value**2  # ≈ 0.044
```

slope
0.25
For every extra 1% in loss rate, churn goes up by 0.25%.

intercept
0.135 (13.5%)
If loss rate were 0%, you'd still lose 13.5% of customers.

r_value
0.21
Loss rate & churn have a small, positive link

r^2
0.044 (4.4%)
Only 4.4% of churn swings are explained by loss rate changes.

p_value
0.08 (8%)
There's an 8% chance this slope happened by random luck.

stderr
0.10
slope (0.25) could plausibly be off by ±0.10 in either dir.

n_points
~64
You ran this on all quarter×product dots (e.g. 16 quarters × 4 products).

TLDR
- **Slope 0.25**: If you bump expected loss from, say, 10 % to 11 %, expect churn up by 0.25 %.
- **Intercept 13.5 %**: Even at zero loss rate, you'd still bleed ~13.5 % of customers.
- **r = 0.21**: There's a tiny positive vibe between loss and churn—but it's weak.
- **r^2 = 4.4 %**: Loss rate barely explains churn (95.6 % of churn is about other stuff).
- **p = 0.08**: It's not "statistically significant" at the usual 5 % cutoff—so maybe the link is just noise.
- **stderr 0.10**: Our slope estimate isn't super-precise; it could be as low as 0.15 or as high as 0.35.

riskier loans see slightly more churn, but only a tiny bit—and most churn has nothing to do with loss rate.