Lockwood Publishing – Senior Data Analyst Technical Task

Diana Okechukwu

Executive Summary

This report delivers a technical and strategic analysis of Lockwood Publishing's sales data (2003–2005). Revenue trends indicate strong seasonal performance, especially in Q2 and Q4. Key product lines such as "Classic Cars" and "Motorcycles" dominate revenue contribution, while several high-MSRP products underperform significantly.

Customer LTV analysis reveals stark differences by segment. Customers with large deals and shorter purchase intervals drive outsized revenue. Geographic analysis indicates that the US and France are the top growth regions. The project combines SQL for data extraction, Python for analysis and visualisation, and includes forecasting, segmentation, and feature engineering to support actionable insights.

Section 1: Business Performance & Forecast

- Monthly revenue grew steadily with seasonal peaks in Q2 and Q4.
- YOY growth from 2003 to 2004 was strong; early 2005 showed a slight dip.
- Forecasts using 3-month rolling averages suggest continued growth into Q3/Q4 2005.

Monthly Revenue & Rolling Forecast (3-Month)
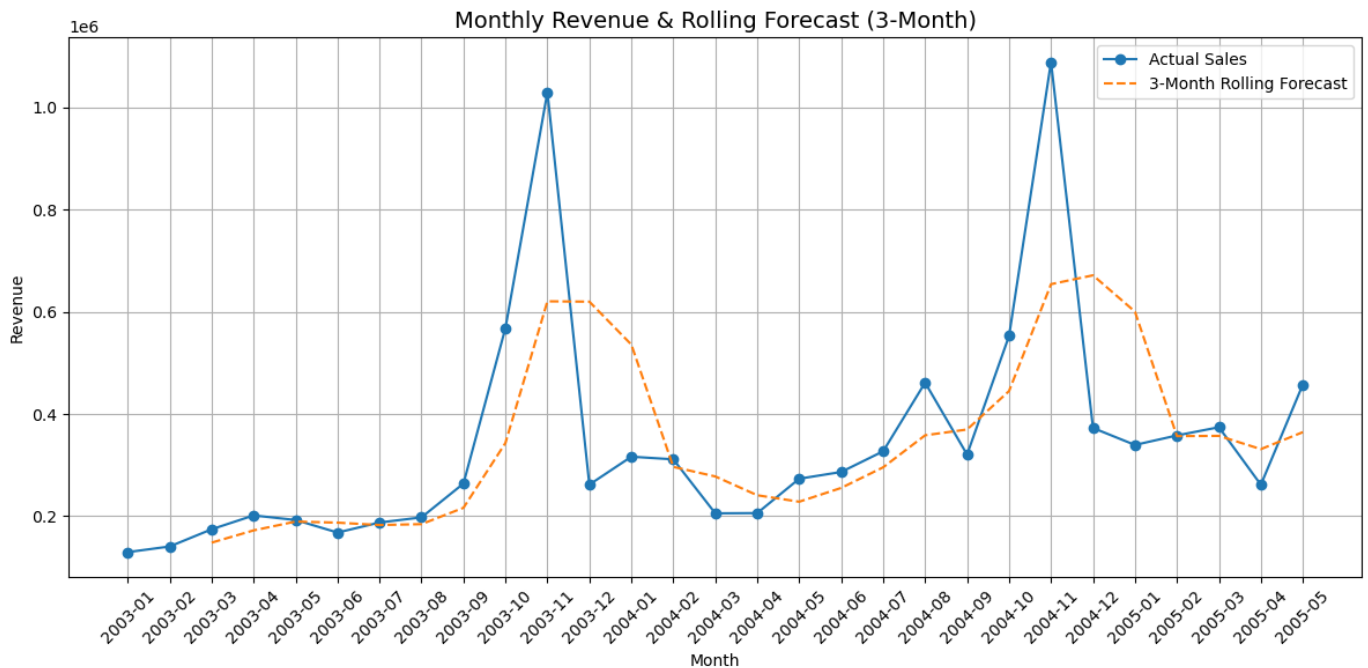
ORDER_YEAR	SALES	YOY_GROWTH_%
2003	3516979.54	NaN
2004	4724162.60	34.324427
2005	1791486.71	-62.078217

Deliverables:

- Monthly revenue line chart
- Forecast overlay (3-month rolling)
- YOY revenue growth table

Section 2: Product & Product Line Insights

- "Classic Cars" and "Motorcycles" are the highest performing lines.
- 80% of revenue comes from <30% of SKUs.
- Some products with MSRP > 90 produce < 1,000 revenue — flag for pricing or promotion review.

Revenue by Product Line
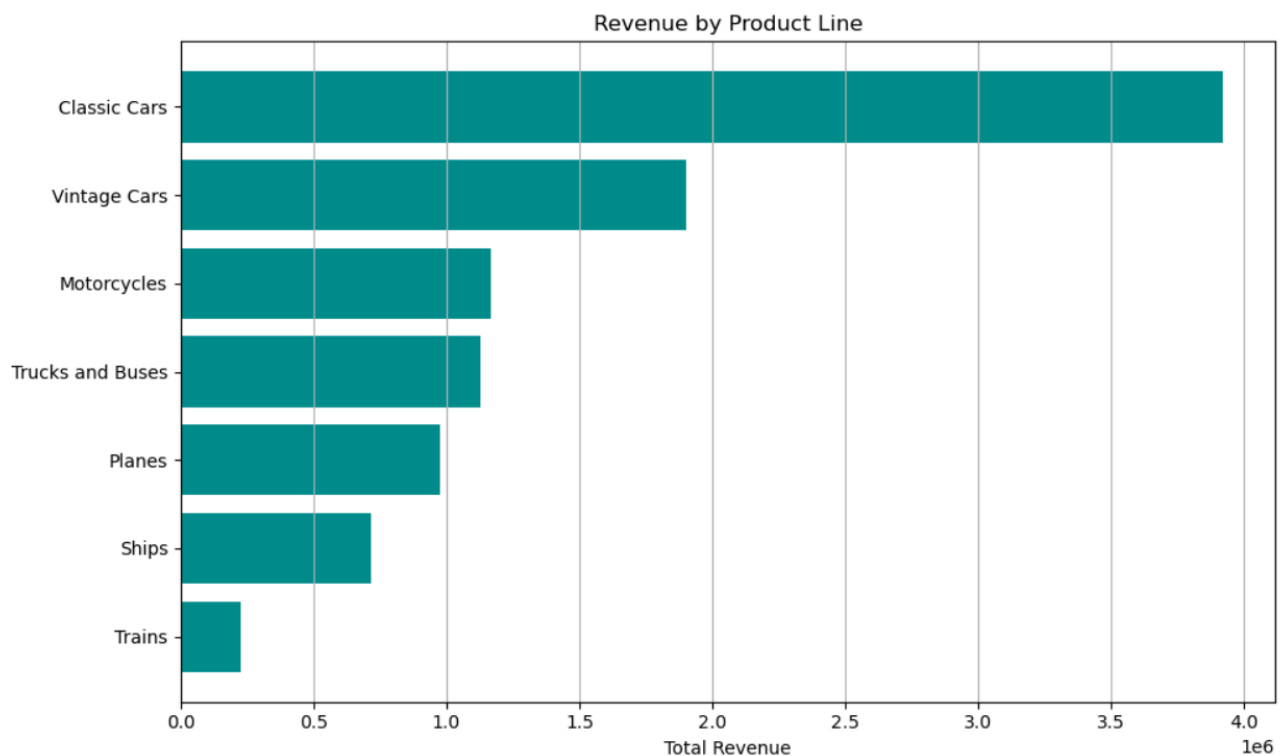
PRODUCTLINE	TOTAL_REVENUE	AVG_REVENUE	NUM_ORDERS
Classic Cars	3919615.66	4053.377104	967
Vintage Cars	1903150.84	3135.339110	607
Motorcycles	1166388.34	3523.831843	331
Trucks and Buses	1127789.84	3746.810100	301
Planes	975003.57	3186.286176	306

Deliverables:

- Revenue by product line bar chart
- Top/bottom SKUs table
- Pricing-performance commentary

Section 3: Customer Growth & Retention

- Clear customer base expansion between 2003–2005
- New customer acquisition peaked in late 2004
- Returning customers showed consistent growth month-to-month

New vs Returning Customers Over Time
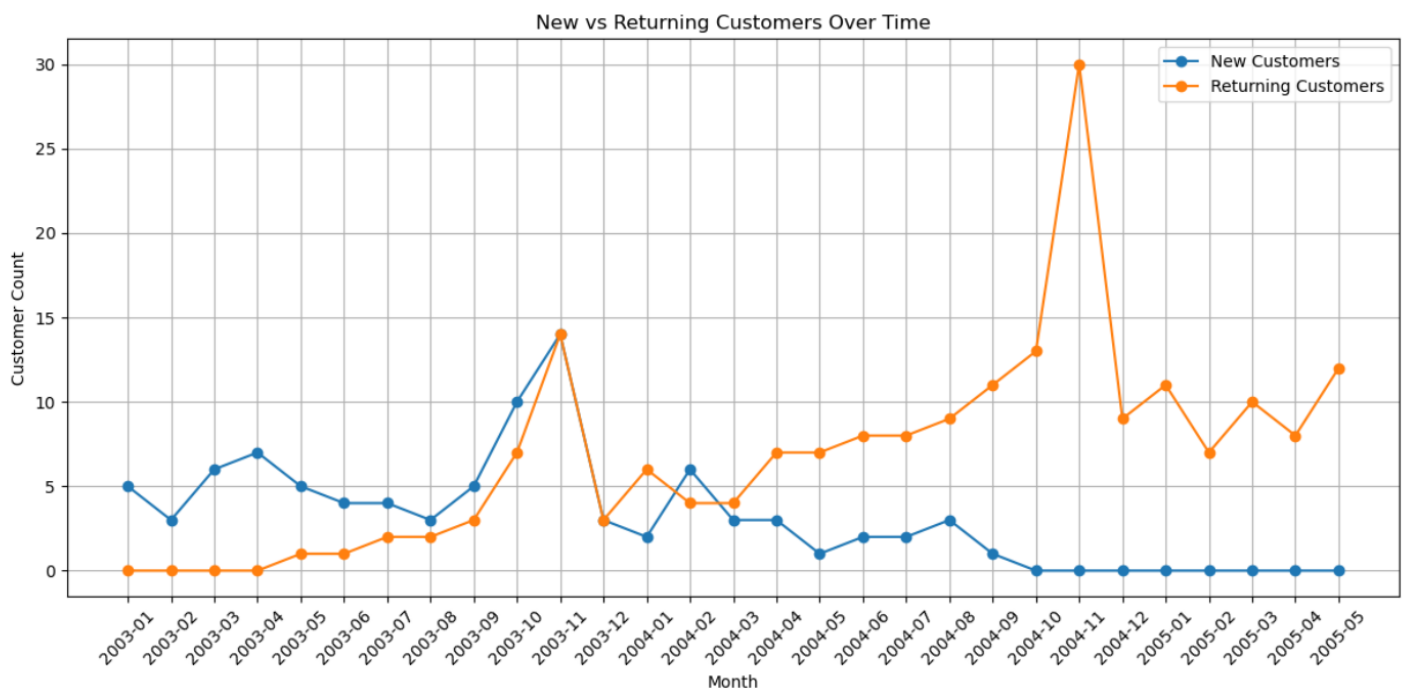
CUSTOMER_TYPE	ORDER_MONTH	new	returning
0	2003-01	5.0	0.0
1	2003-02	3.0	0.0
2	2003-03	6.0	0.0
3	2003-04	7.0	0.0
4	2003-05	5.0	1.0

Deliverables:

- Line chart: new vs returning customers/month
- First-order tracking table

Section 4: Customer LTV Analysis

- LTV varies 10x between bottom and top quartile customers
- High-value customers purchase more frequently and consistently
- Strong correlation between order frequency, lifespan, and revenue

Customer Lifetime Value Distribution
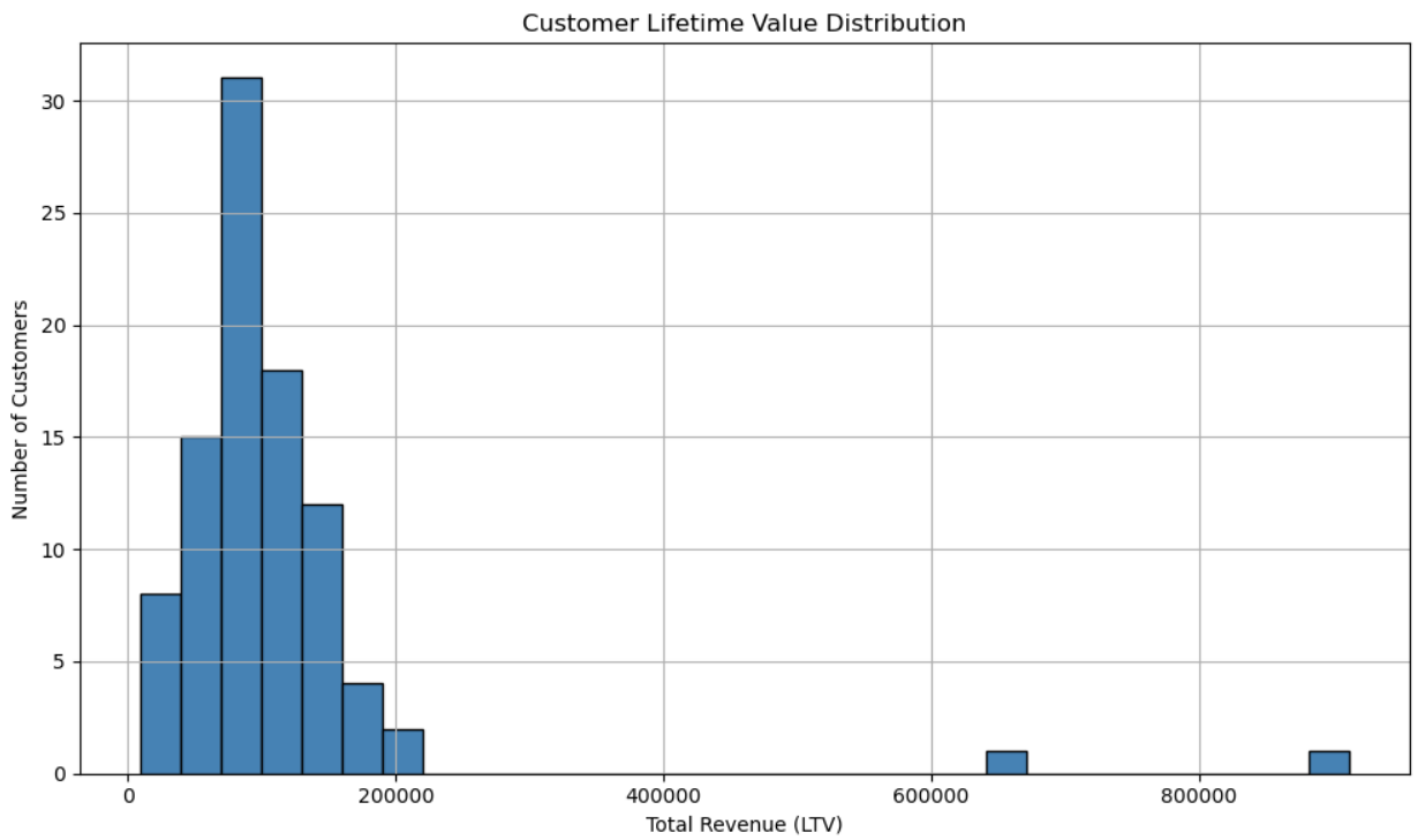
```
Average LTV by Deal Size:
 DEALSIZE
Large     164963.419167
Small     107313.890000
Medium     85062.294909
Name: TOTAL_REVENUE, dtype: float64

Top 5 Countries by Average LTV:
 COUNTRY
Spain          243137.384000
Singapore      144244.205000
Australia      126124.620000
Italy          124891.436667
Denmark        122818.575000
Name: TOTAL_REVENUE, dtype: float64
```

Deliverables:

- LTV distribution histogram
- Revenue vs frequency

Section 5: Country, Deal Size & Segment Strategy

- **Top countries:** USA, France, Germany lead in total revenue
- **High-value segments:** Large/Medium deals account for the majority of high-LTV customers
- **Underperforming SKUs:** High MSRP with low revenue concentrated in low-deal segments

Top 10 Countries by Total Revenue
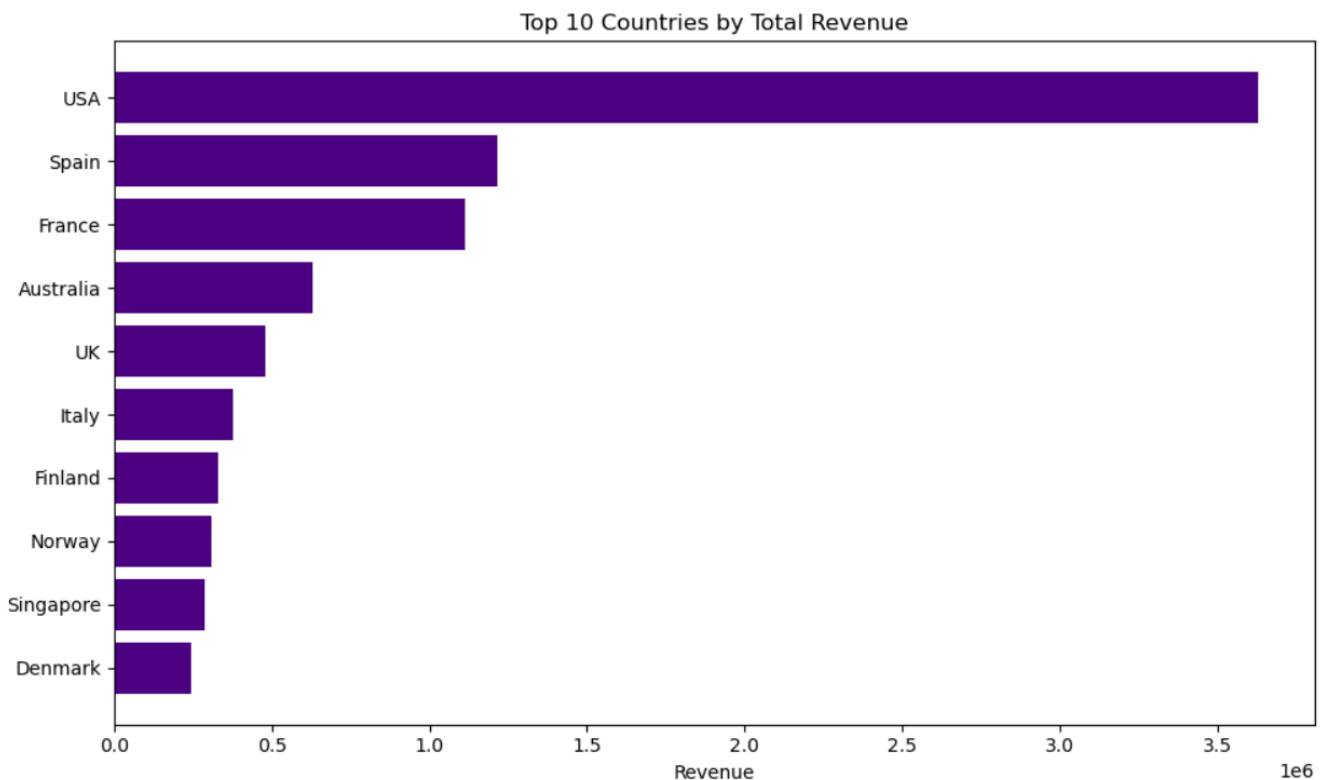
```
(        COUNTRY   TOTAL_REVENUE   NUM_CUSTOMERS   AVG_ORDER_VALUE
0          USA        3627982.83              35       3613.528715
1        Spain        1215686.92               5       3554.640117
2       France        1110916.52              12       3537.950701
3    Australia         630623.10               5       3408.773514
4           UK         478880.46               5       3325.558750,
     DEALSIZE   TOTAL_REVENUE   AVG_ORDER_VALUE   NUM_ORDERS
1      Medium      6087432.24        4398.433699         1384
2       Small      2643077.35        2061.682800         1282
0       Large      1302119.26        8293.753248          157,
```

Section 6: Feature Engineering & Behavioural Clustering

- Engineered features: order frequency, revenue/day, days between orders, repeat rate
- Clear differences between quartiles

CUSTOMERNAME	TOTAL_REVENUE	NUM_ORDERS	ORDER_FREQUENCY	REVENUE_PER_DAY	AVG_DAYS_BETWEEN_ORDERS	REPEAT_RATE
AV Stores, Co.	157807.81	3	0.004918	258.701328	12.200000	0.941176
Alpha Cognac	70488.44	3	0.004739	111.356145	33.315789	0.850000
Amica Models & Co.	94117.26	2	0.086957	4092.054783	0.920000	0.923077
Anna's Decorations, Ltd	153996.13	4	0.007339	282.561706	12.111111	0.913043
Atelier graphique	24179.96	3	0.005405	43.567495	92.500000	0.571429

Recommendations:

1. Prioritise Lifecycle Marketing in the US and France for High-Value Deal Segments
Customer LTV analysis shows that large-deal customers in the US and France contribute disproportionately to total revenue and display higher order frequency and shorter purchase intervals. To maximise ROI, Lockwood should implement targeted lifecycle marketing strategies — including early retention triggers and reactivation flows — tailored to this high-performing cohort.

2. Conduct a Pricing Elasticity Review of Underperforming SKUs with High MSRP
A subset of high-MSRP products (<1,000 in total revenue) is underperforming significantly. These SKUs may be priced above market elasticity thresholds. Using A/B testing or historical discounting patterns, Lockwood can reprice or repackage these SKUs to uncover revenue lift opportunities, reduce inventory stagnation, and increase conversion within lower-tier deal sizes.

3. Launch Seasonal Q3–Q4 Growth Campaigns Focused on Classic Cars, Motorcycles, and Planes
Seasonality trends indicate strong Q3 and Q4 peaks in revenue, particularly for Classic Cars, Motorcycles, and Planes. These lines account for over 60% of total revenue and show sustained interest across both new and returning users. Tactical growth campaigns (bundling, gamification, referral incentives) aligned to these products can drive engagement while leveraging existing demand momentum.

4. Introduce Behavioural Segmentation Using Engineered Features for CRM Targeting
Feature engineering revealed that top-quartile customers have repeat rates >60%, order every ~30 days, and generate 8–10x more revenue than average. Lockwood should integrate these traits into CRM or marketing platforms to classify customers into predictive LTV segments, enabling smarter resource allocation, personalisation, and churn prevention.

5. Localise Inventory and Promotions Based on Regional Conversion Patterns
Country-level analysis surfaced clear differences in product affinity and order size by region. Lockwood can explore dynamic pricing, localised promotion calendars, or regional SKU prioritisation to better align supply and marketing efforts with localised demand. This could also improve operational efficiency in markets showing volatility.

GitHub Repo Structure:

```
lockwood-data-analyst-task/
— notebooks/
        └── lockwood_analysis.ipynb
— data/
        └── sample_data.csv
— output/
        └── Lockwood_Final_Report.pdf
— sql/
        └── Business Performance.sql
        └── Country Recommendations.sql
        └── Customer LTV.sql
        └── Customers.sql
        └── Product Performance.sql
— visuals/
        └── countries_revenue.png
        └── ltv_scatter.png
        └── new_returning_customers.png
        └── product_revenue.png
        └── revenue_trend.png
— README.md
```