Blinkit Grocery Sales Analysis

Comprehensive Analysis of Sales Performance,
Customer Satisfaction & Inventory Distribution

Tools: Python | PostgreSQL | Power BI

EXECUTIVE SUMMARY

This data analytics project delivers actionable insights into Blinkit's grocery delivery operations through comprehensive analysis of 8,523 transactions across 10 outlets. Using Python, SQL, and Power BI, the analysis reveals critical patterns in sales performance, customer satisfaction, and inventory distribution.

KEY METRICS AT A GLANCE

Total Revenue	**$1.20 Million**
Average Transaction	$141
Customer Rating	3.9/5.0
Product Range	8,523 items across 16 categories

Critical Business Findings

- Tier 3 locations generate 42.3% of revenue, outperforming metropolitan areas
- Supermarket Type1 format contributes 65.5% of total sales
- Low-fat products dominate with 64.6% market share
- Top 2 categories (Fruits & Vegetables, Snack Foods) drive 30% of revenue
- Customer satisfaction consistent at 3.97/5.0 but below 4.0 excellence threshold

1. PROJECT OVERVIEW

1.1 Business Problem

Blinkit operates in India's competitive last-minute grocery delivery market. To maintain market leadership and sustainable growth, the organization needs data-driven insights to:

- Identify high-performing product categories and outlet configurations
- Understand geographical performance patterns across city tiers
- Optimize inventory distribution and product visibility
- Balance customer satisfaction with operational efficiency

1.2 Project Objectives

- Analyze sales performance across products, outlets, and locations
- Evaluate customer satisfaction metrics and correlations
- Assess inventory distribution efficiency
- Develop strategic recommendations for business optimization

1.3 Dataset Overview

Attribute	Details
Total Records	8,523 transactions
Unique Products	1,559 items
Outlets	10 locations
Product Categories	16 types
Time Period	Outlet establishments 2012-2022
Key Metrics	Sales, ratings, visibility, location tiers

2. METHODOLOGY & TOOLS

2.1 Analytical Approach

This project employs a three-layer analytical framework combining statistical analysis, database querying, and interactive visualization:

Layer	Tool	Purpose
Data Processing & EDA	Python (Pandas, NumPy)	Cleaning, exploration, statistical analysis
Database Analysis	PostgreSQL + SQL	Advanced queries, aggregations, insights
Visualization & BI	Power BI	Interactive dashboards, KPIs, reporting

2.2 Analysis Workflow

1. **Data Acquisition:** Imported Excel dataset into Python environment
2. **Data Cleaning:** Standardized columns, handled 1,463 missing values in item_weight
3. **Exploratory Analysis:** Generated descriptive statistics, distributions, correlations
4. **Database Setup:** Loaded data into PostgreSQL using SQLAlchemy
5. **SQL Analysis:** Executed 13 analytical queries for business intelligence
6. **Dashboard Creation:** Built Power BI dashboard with 12+ visualizations
7. **Insight Synthesis:** Derived actionable recommendations from multi-layer analysis

3. PYTHON EXPLORATORY DATA ANALYSIS

3.1 Data Quality Assessment

Initial data examination revealed high quality with minimal issues:

- No duplicate records detected
- 1,463 missing values in item_weight (17.2%) - acceptable for analysis
- All categorical fields properly encoded
- No outliers requiring removal in sales or rating distributions

3.2 Key Statistical Findings

Metric	Value	Insight
Total Sales	$1,201,681.49	Strong revenue base
Avg Sales/Transaction	$140.99	Consistent basket size
Avg Customer Rating	3.97/5.0	Good but improvable
Sales Std Deviation	Low variance	Stable transaction values
Rating Distribution	Concentrated 3-5	Generally satisfied customers

3.3 Inventory Distribution Analysis

Outlet inventory diversity:

- OUT027: 935 items (highest diversity)
- OUT013, OUT046, OUT049: 930+ items (strong range)
- OUT019: 528 items (specialized/smaller format)
- Average: 852 items per outlet

```
blinkit_df.describe()
```

	outlet_establishment_year	item_visibility	item_weight	sales	rating
count	8523.000000	8523.000000	8523.000000	8523.000000	8523.000000
mean	2016.450546	0.066132	12.813390	140.992783	3.965857
std	3.189396	0.051598	4.241384	62.275067	0.605651
min	2011.000000	0.000000	4.555000	31.290000	1.000000
25%	2014.000000	0.026989	9.310000	93.826500	4.000000
50%	2016.000000	0.053931	12.850000	143.012800	4.000000
75%	2018.000000	0.094585	16.000000	185.643700	4.200000
max	2022.000000	0.328391	21.350000	266.888400	5.000000

4. SQL DATABASE ANALYSIS

4.1 Sales Performance by Category

Category	Total Sales	% of Total
Fruits & Vegetables	$178,123	14.8%
Snack Foods	$175,433	14.6%
Household	$135,965	11.3%
Frozen Foods	$119,203	9.9%
Dairy	$101,868	8.5%

Key Insight: Top 5 categories contribute 59% of revenue, indicating healthy diversification with clear leaders.

4.2 Geographical Performance Analysis

Location Tier	Total Sales	Market Share	Performance
Tier 3	$472,131	42.27%	★★★ Exceptional
Tier 2	$393,151	35.21%	★★ Strong
Tier 1	$336,400	30.13%	★ Good

Critical Finding: Counter-intuitive result - smaller cities outperform metros, suggesting lower competition, better unit economics, and successful market positioning in emerging markets.

4.3 Fat Content Preference

Product Type	Sales	Market Share
Low Fat	$776,317	64.6%
Regular	$425,365	35.4%

	item_fat_content text	total_sales numeric	sales_percentage numeric
1	Regular	425361.80	35.40
2	Low Fat	776319.69	64.60

5. OUTLET PERFORMANCE ANALYSIS

5.1 Outlet Type Comparison

Outlet Type	Total Sales	% Share	Avg Sales	Avg Rating	Items
Supermarket Type1	$787,550	65.5%	$141	3.96	5,577
Grocery Store	$151,939	12.6%	$140	3.99	1,083
Supermarket Type2	$131,478	10.9%	$142	3.97	928
Supermarket Type3	$130,715	10.9%	$140	3.95	935

5.2 Key Findings

- Supermarket Type1 dominates with 65.5% revenue despite representing 40% of outlets
- Average transaction value consistent ($140-142) across all formats
- Grocery Stores achieve highest satisfaction (3.99) but lowest total sales
- Type1 outlets benefit from scale advantages and broader product range
- Customer satisfaction shows no correlation with outlet sales volume

6. POWER BI DASHBOARD INSIGHTS

6.1 Dashboard Architecture

The Power BI dashboard provides real-time business intelligence through:

- Four primary KPI cards: Total Sales, Average Sales, Item Count, Average Rating
- Time-series analysis: Outlet establishment trend showing $205K peak in 2018
- Geographical breakdown: Tier performance with interactive filtering
- Category performance: Horizontal bar chart ranking all 16 categories
- Outlet size distribution: Donut chart showing Medium (42.3%), Small (37%), High (20.7%)
- Fat content analysis: Stacked bars comparing Low Fat vs Regular across tiers
- Detailed metrics table: Comprehensive outlet type comparison

6.2 Interactive Features

- Dynamic filters: Location tier, outlet size, item type slicers
- Cross-filtering: Clicking any element filters entire dashboard
- Drill-down capability: Navigate from outlet type to individual locations

- Trend analysis: Historical performance visualization
- Comparative views: Side-by-side outlet and category comparisons

7. INVENTORY & VISIBILITY ANALYSIS

7.1 Item Visibility Patterns

Item Type	Avg Visibility	Avg Sales	Correlation
Breakfast	0.0857	$173	High vis, strong sales
Starchy Foods	0.0676	$178	Moderate vis, highest sales
Frozen Foods	0.0656	$164	Lower vis, good sales
Health & Hygiene	0.0552	$148	Lowest vis, lower sales

Counterintuitive Finding: No strong positive correlation between visibility and sales. Starchy Foods lead in sales with moderate visibility, while Breakfast items have highest visibility but lower sales. This suggests purchasing decisions driven by necessity rather than display prominence.

7.2 Inventory Efficiency Metrics

- Average inventory breadth: 852 items per outlet
- High-performing outlets maintain 930+ SKUs
- Smaller outlets (528 items) show focused product selection
- Visibility range: 0.055-0.086 across categories
- No significant correlation between inventory size and sales performance

7.3 Optimization Opportunities

Analysis reveals three key optimization areas:

Space Reallocation: Reallocate premium shelf space from high-visibility breakfast items to high-sales starchy foods and frozen items

Category Focus: Increase depth in top-performing categories (Fruits & Veg, Snacks) rather than breadth across all categories

Visibility Testing: Conduct A/B testing on Health & Hygiene products to determine if increased visibility drives sales growth

8. CUSTOMER SATISFACTION ANALYSIS

8.1 Overall Satisfaction Metrics

Metric	Value	Assessment
Overall Average Rating	3.97/5.0	Good, but below 4.0 excellence
Rating Range	3.95 - 3.99	Very consistent across outlets
Standard Deviation	Low	Predictable customer experience
Distribution	Concentrated 3-5	Few low ratings

8.2 Satisfaction by Outlet Type

- Grocery Store: 3.99 (highest) - personalized service, smaller format
- Supermarket Type2: 3.97 - balanced size and service
- Supermarket Type1: 3.96 - scale advantage with consistent quality
- Supermarket Type3: 3.95 - standard performance

8.3 Critical Insights

- Minimal variation (0.04 points) suggests standardized service quality
- All formats performing in narrow band indicates strong operational consistency
- Grocery Stores lead despite lowest sales, suggesting customer preference for experience over selection
- No strong correlation between rating and sales volume
- Scores below 4.0 threshold leave room for competitive differentiation

RECOMMENDATION: Implement targeted initiatives to push ratings from 3.97 to 4.2+. Focus on: (1) Delivery time reliability, (2) Product quality consistency, (3) Customer service training, (4) Post-delivery follow-up. Even a 0.2-point improvement can significantly impact brand perception.

9. KEY INSIGHTS & DISCOVERIES

9.1 Sales Performance Insights

INSIGHT 1: Tier 3 Market Dominance

Finding: Tier 3 locations generate $472K (42.3% of revenue), outperforming Tier 1 metros by 40%

Implication: Lower competition, better unit economics, and strong product-market fit in emerging markets

Significance: Contradicts traditional retail wisdom favoring metropolitan concentration

INSIGHT 2: Format Efficiency Gap

Finding: Supermarket Type1 generates 65.5% of revenue with consistent $141 avg transaction

Implication: Scale economies and broader product range drive superior performance

Significance: Format standardization opportunity for other outlet types

INSIGHT 3: Category Concentration

Finding: Fruits & Vegetables ($178K) and Snack Foods ($175K) drive 30% of total revenue

Implication: Fresh produce indicates quality trust; snacks show impulse purchase success

Significance: These categories are critical to customer frequency and basket building

9.2 Consumer Behavior Insights

INSIGHT 4: Health-Conscious Consumer Base

Finding: Low-fat products command 64.6% market share (nearly 2:1 ratio)

Implication: Strong preference for healthier options among urban, health-conscious demographic

Significance: Opportunity to expand organic and wellness product lines

10. STRATEGIC RECOMMENDATIONS

Based on comprehensive analysis across Python, SQL, and Power BI, the following prioritized recommendations are proposed:

HIGH PRIORITY RECOMMENDATION 1

Accelerate Tier 3 Market Expansion

Action: Target 5-7 new Supermarket Type1 outlets in high-growth Tier 3 cities within 18 months

Rationale: Tier 3 locations outperform by 40% with lower competition and better unit economics

Expected Impact: 25-30% increase in total sales; improved market penetration in underserved regions

Implementation: Conduct market research → Site selection → Pilot 2-3 locations → Scale based on performance

Timeline: Q1-Q2: Research; Q3-Q4: First pilots; Year 2: Scale rollout

HIGH PRIORITY RECOMMENDATION 2

Supermarket Type1 Format Optimization & Replication

Action: Document and standardize Supermarket Type1 best practices; convert underperforming formats where feasible

Rationale: 65.5% sales contribution demonstrates clear format superiority

Expected Impact: 15-20% improvement in operational efficiency across network

Implementation: Best practice audit → Training programs → Gradual format optimization → Performance tracking

Quick Win: Apply Type1 inventory management practices to Type2 and Type3 outlets immediately

MEDIUM PRIORITY RECOMMENDATION 3

Customer Experience Enhancement Program

Action: Launch comprehensive program to elevate ratings from 3.97 to 4.2+

Focus Areas: Delivery reliability, quality assurance, service training, post-delivery engagement

Expected Impact: Improved retention, enhanced brand perception, premium pricing opportunity

Implementation: Mystery shopper program → Staff training → Feedback loops → Incentive alignment

Success Metric: Achieve 4.2+ average rating within 12 months

MEDIUM PRIORITY RECOMMENDATION 4

Category Management & Product Line Expansion

Action: Strengthen inventory in top categories; expand low-fat/organic offerings; optimize shelf space

Rationale: Top 2 categories drive 30% of revenue; 64.6% low-fat preference shows health trend

Expected Impact: 10-15% increase in basket size; 8-12% growth in health category

Implementation: SKU rationalization → Supplier partnerships → Planogram redesign → Marketing campaigns

Quick Win: Reallocate shelf space from low-performing to high-velocity categories immediately

11. IMPLEMENTATION ROADMAP

11.1 Phased Execution Plan

PHASE 1: IMMEDIATE ACTIONS (0-3 Months)

- □ Conduct Supermarket Type1 best practice documentation
- □ Begin Tier 3 market research for expansion targets
- □ Launch customer experience audit and training program
- □ Implement shelf space reallocation based on visibility analysis
- □ Expand low-fat and organic product SKUs by 15%

PHASE 2: SHORT-TERM INITIATIVES (3-6 Months)

- □ Open first 2-3 pilot outlets in selected Tier 3 cities
- □ Roll out Type1 operational practices to other formats
- □ Achieve 4.0+ customer rating through experience improvements
- □ Complete category management optimization
- □ Establish performance monitoring dashboard

PHASE 3: LONG-TERM STRATEGY (6-18 Months)

- □ Scale Tier 3 expansion to 5-7 total new outlets
- □ Achieve 4.2+ customer rating consistently
- □ Complete format optimization across outlet network
- □ Launch advanced analytics for predictive inventory
- □ Expand health-focused product lines by 30%

11.2 Success Metrics & KPIs

KPI	Current	Target (12M)	Target (18M)
Total Sales	$1.20M	$1.50M	$1.65M
Customer Rating	3.97	4.0+	4.2+

Tier 3 Revenue Share	42.3%	48%	52%
Type1 Contribution	65.5%	70%	72%
Low-Fat Market Share	64.6%	67%	70%

12. RISKS & MITIGATION STRATEGIES

TIER 3 EXPANSION RISK

Description: New markets may not replicate current Tier 3 success

Probability: Medium | Impact: High

Mitigation: Pilot approach with 2-3 locations; thorough market research; exit criteria defined upfront

CUSTOMER RATING PLATEAU

Description: Rating improvements may stall due to operational constraints

Probability: Medium | Impact: Medium

Mitigation: Multi-pronged approach targeting delivery, quality, service; continuous feedback loops

FORMAT CONVERSION RESISTANCE

Description: Existing outlets may resist Type1 standardization

Probability: Low | Impact: Medium

Mitigation: Change management program; pilot conversions; demonstrate ROI before scaling

COMPETITIVE RESPONSE

Description: Competitors may increase Tier 3 presence

Probability: High | Impact: Medium

Mitigation: First-mover advantage through rapid expansion; build brand loyalty; operational excellence

12.1 Contingency Plans

- If Tier 3 pilots underperform: Reassess location selection; focus on strengthening existing markets
- If rating improvement stalls: Increase investment in technology solutions; consider external consultants
- If format conversion faces resistance: Maintain multi-format strategy; focus on operational improvements within each type
- If competitive pressure intensifies: Accelerate expansion timeline; enhance differentiation through service quality

13. CONCLUSION

13.1 Project Summary

This comprehensive data analytics project analyzed 8,523 transactions across 10 Blinkit outlets using Python, SQL, and Power BI. The multi-layered analytical approach uncovered critical insights that challenge conventional retail wisdom and reveal significant optimization opportunities.

13.2 Key Takeaways

1. Tier 3 dominance represents Blinkit's competitive advantage and primary growth opportunity

2. Supermarket Type1 format provides the operational blueprint for future expansion

3. Health-conscious consumer base (64.6% low-fat preference) drives product strategy

4. Customer satisfaction consistency (3.97) indicates operational maturity but improvement potential

5. Category concentration (top 2 = 30% revenue) offers clear focus areas for inventory optimization

13.3 Strategic Formula for Success

Supermarket Type1 Format
+
Tier 3 Locations
+
Low-Fat/Health Products
=
Optimal Growth Strategy

13.4 Expected Outcomes

With successful implementation of recommendations:

- 25-30% revenue growth through Tier 3 expansion
- 15-20% operational efficiency improvement via format optimization

- Customer rating elevation to 4.2+ for competitive differentiation
- 10-15% basket size increase through category management
- Market leadership position in emerging city grocery delivery

13.5 Next Steps

- Present findings to executive leadership for strategic alignment
- Secure budget allocation for Tier 3 expansion and technology initiatives
- Establish cross-functional implementation team
- Develop detailed project plans for high-priority recommendations
- Set up real-time monitoring dashboard for ongoing performance tracking