

Project Fashion Insight

The background features a large, semi-transparent circular graphic with a textured, light orange and yellow gradient. Overlaid on this graphic is the word "alo" in a bold, black, sans-serif font.

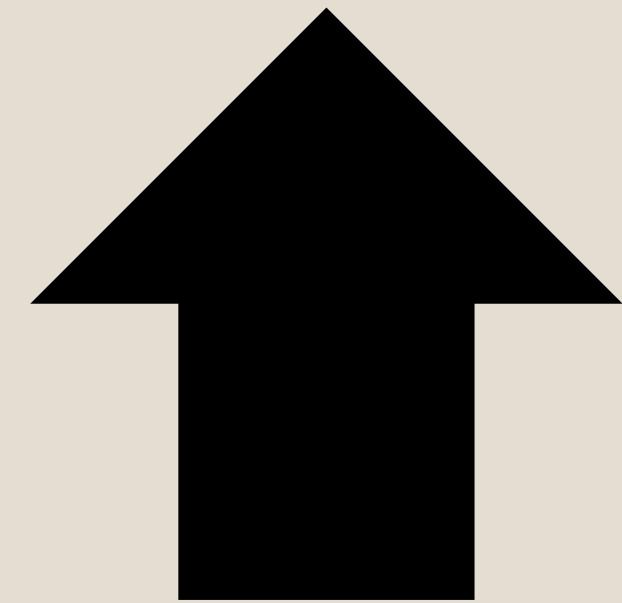
alo

Kali, Taleen, Kendall, Lilly, Karen

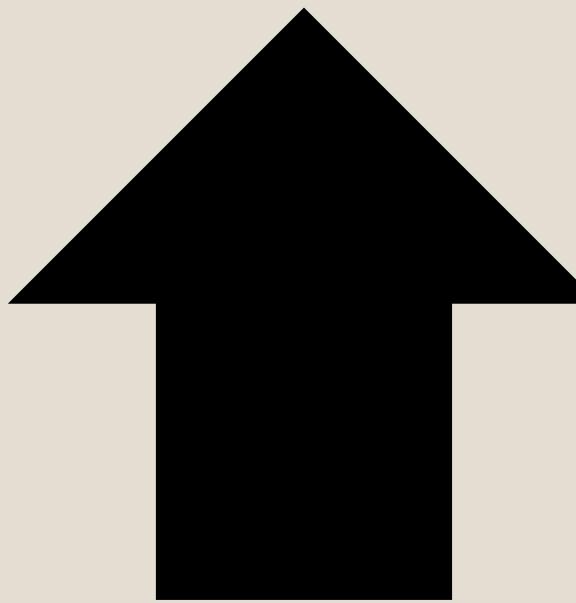
WHERE WE LEFT
OFF...



AVERAGE CUSTOMER REVIEW RATINGS



3.33-10.00%



WITHIN 6 MONTHS

REFLECTS



Alo Yoga

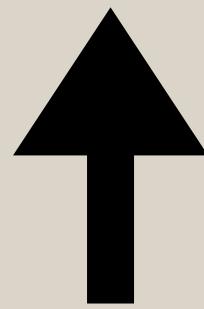
CUSTOMER SATISFACTION

IMPACTS



Alo Yoga

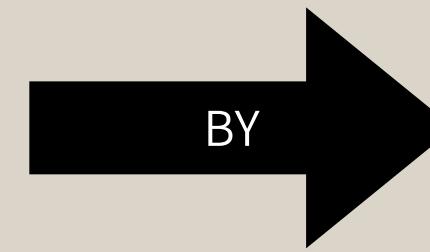
REVENUE



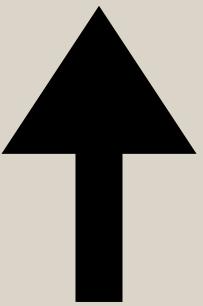
IMPROVED
CUSTOMER
SATISFACTION



INCREASE IN
REVENUE



INCREASING AVG
CUSTOMER REVIEW
RATING



FIND THE MOST INFLUENTIAL ASPECT
OF CUSTOMER REVIEW RATINGS



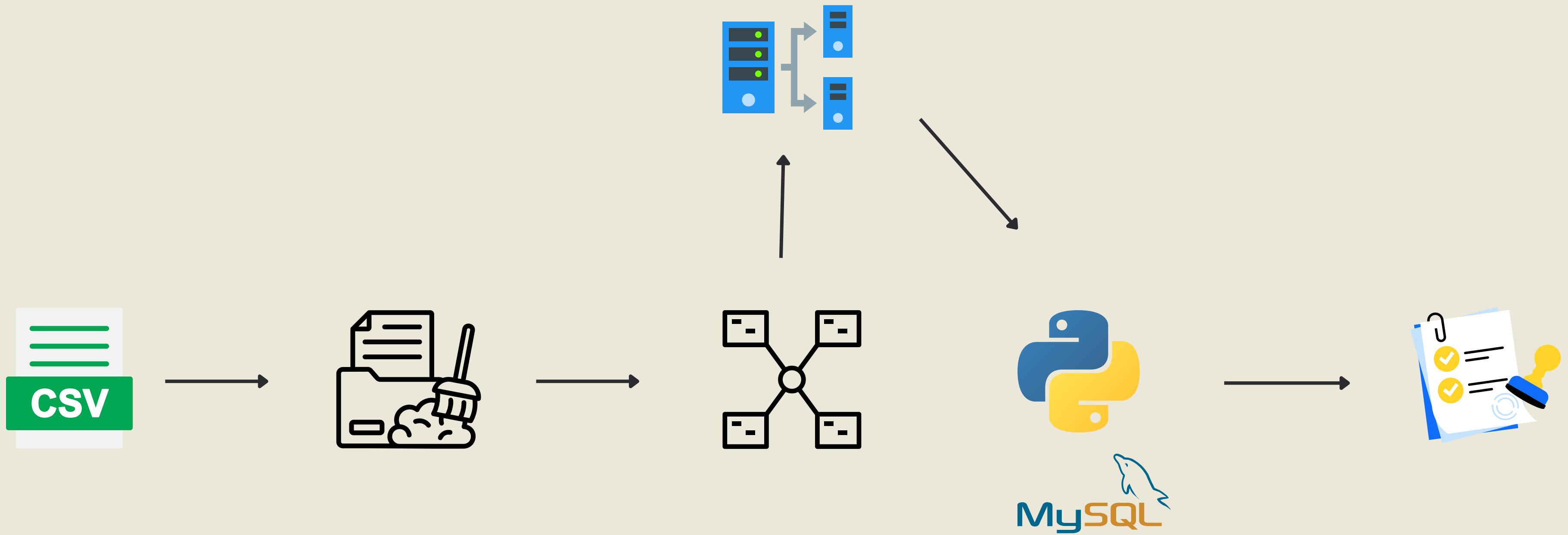
CREATE PERSONALIZATION AND
MARKETING STRATEGIES



Etl Pipeline - diagram



Etl Pipeline - diagram



Automated Etl Pipeline -



INDUSTRIALIZE

The screenshot shows the Windows Task Scheduler interface. On the left, a list of scheduled tasks is displayed, including Adobe Acrobat Update, ArcGIS Pro In..., ETL_Automat..., MicrosoftEd..., OneDrive Re..., OneDrive Sta..., and OneDrive Sta... The 'ALO_ETL_automated_pipeline' task is selected. On the right, a detailed view of this task is shown, including tabs for General, Triggers, Actions, Conditions, and Settings. The General tab shows the task name as 'ALO_ETL_automated_pipeline', location as '\', and author as 'DESKTOP-CJB7C70\Taleen Barakeh'. The Triggers tab is active, showing a single trigger named 'Create Task' set to run daily at 10:24 PM. The Actions tab indicates the task runs a PowerShell script. The Conditions tab and Settings tab are also visible.

This screenshot shows the 'New Trigger' dialog box. The 'Begin the task' dropdown is set to 'On a schedule'. Under 'Settings', the 'Daily' radio button is selected, with 'Start' set to 4/23/2024 at 10:24:13 PM and 'Recur every' set to 1 day. The 'Enabled' checkbox is checked. Advanced settings include 'Repeat task every' set to 1 hour for indefinitely, and other options like random delay and duration. The 'OK' and 'Cancel' buttons are at the bottom.

The screenshot shows the Windows Event Viewer with the 'Event 106, TaskScheduler' entry highlighted. The event details show an info level event from 4/23/2024 at 10:24:13 PM. The message states: 'User "DESKTOP-CJB7C70\Taleen Barakeh" registered Task Scheduler task "ALO_ETL_automated_pipeline".' The event is categorized as 'Task registered' and has an operational code. The log name is 'Microsoft-Windows-TaskScheduler/Operational'.

Automated Etl Pipeline - Running

INDUSTRIALIZE

The screenshot shows the Task Scheduler interface with two tasks listed:

Name	Status	Triggers	Next Run Time	Last Run Time	Last Run Result
Adobe Acrobat	Ready	Multiple triggers defined	4/24/2024 9:00:00 AM	4/23/2024 6:27:16 PM	The operation completed successfully.
ALO_ETL_auto	Run...	At 10:24 PM every day - After triggered, repeat every 1 hour indefinitely.	4/24/2024 1:24:13 AM	4/24/2024 1:03:50 AM	The task is currently running.

Below the table, there are tabs for General, Triggers, Actions, Conditions, Settings, and History. A note says: "When you create a task, you can specify the conditions that will trigger the task. To change these triggers, open the task property pages using the Properties command." A detailed view of the Triggers tab for the ALO_ETL_auto task is shown, listing a single daily trigger: "At 10:24 PM every day - After triggered, repeat every 1 hour indefinitely." with the status "Enabled".

The screenshot shows the Event Log interface with 13 events listed:

Level	Date and Time	Event...	Task Category	Operational Code	Correlation Id
Info	4/24/2024 1:03:54 AM	102	Task completed	(2)	c4dca452-73...
Info	4/24/2024 1:03:54 AM	201	Action complet...	(2)	c4dca452-73...
Info	4/24/2024 1:03:50 AM	110	Task triggered ...	Info	c4dca452-73...
Info	4/24/2024 1:03:50 AM	200	Action started	(1)	c4dca452-73...
Info	4/24/2024 1:03:50 AM	100	Task Started	(1)	c4dca452-73...
Info	4/24/2024 1:03:50 AM	129	Created Task Pr...	Info	
Info	4/23/2024 10:38:34 PM	102	Task completed	(2)	fac13a34-90...
Info	4/23/2024 10:38:34 PM	201	Action complet...	(2)	fac13a34-90...
Info	4/23/2024 10:38:29 PM	110	Task triggered ...	Info	fac13a34-90...
Info	4/23/2024 10:38:29 PM	200	Action started	(1)	fac13a34-90...
Info	4/23/2024 10:38:29 PM	100	Task Started	(1)	fac13a34-90...
Info	4/23/2024 10:38:29 PM	129	Created Task Pr...	Info	
Info	4/22/2024 10:27:25 PM	106	Task registered	Info	

Below the event list, a specific event is selected: "Event 102, TaskScheduler". It has tabs for General and Details.

Automated Pipeline & Industrial Phase



Standardization

Scalability

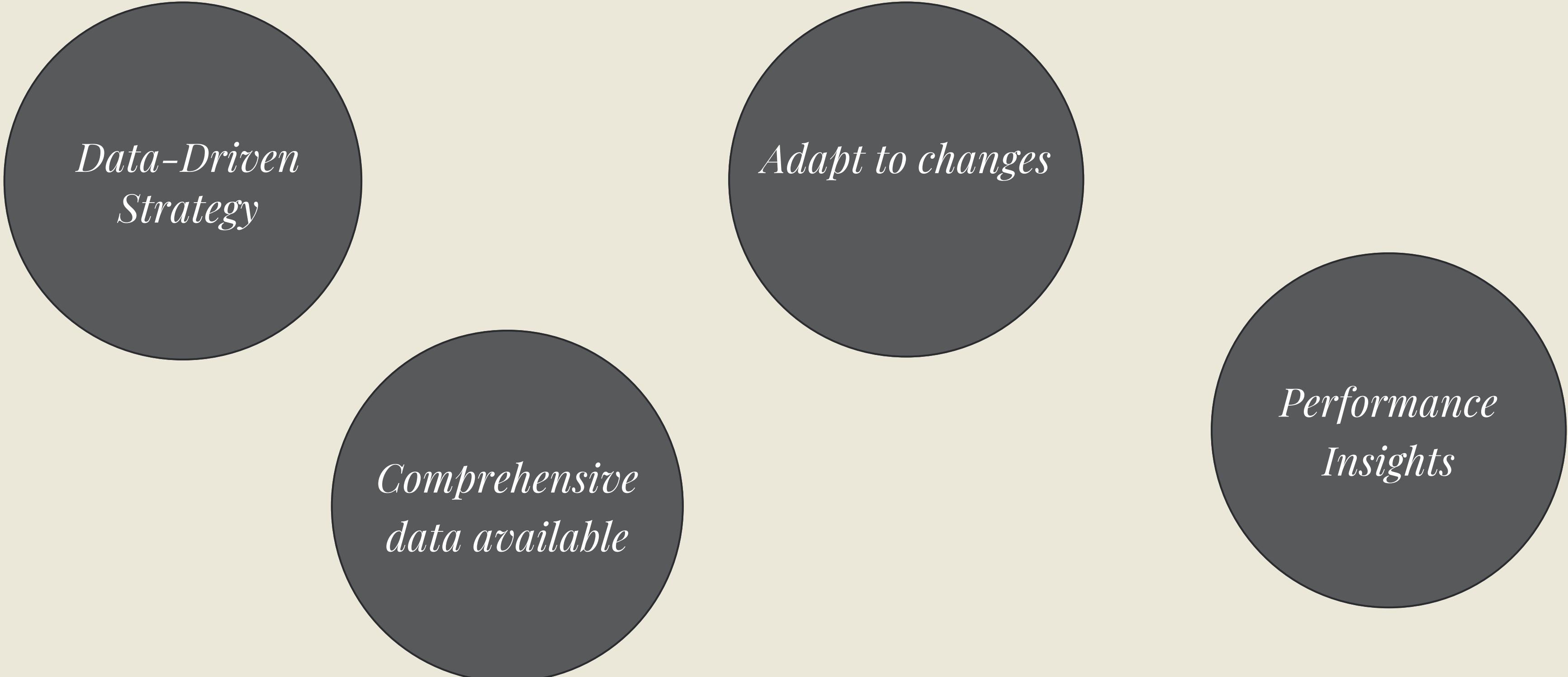
Reliability

Efficiency

Monitoring



How Automated ETL Pipeline Helps with Alo Yoga:



*Data-Driven
Strategy*

Adapt to changes

*Comprehensive
data available*

*Performance
Insights*



Predictive Model Training



Typical Predictive Models

Linear Regression

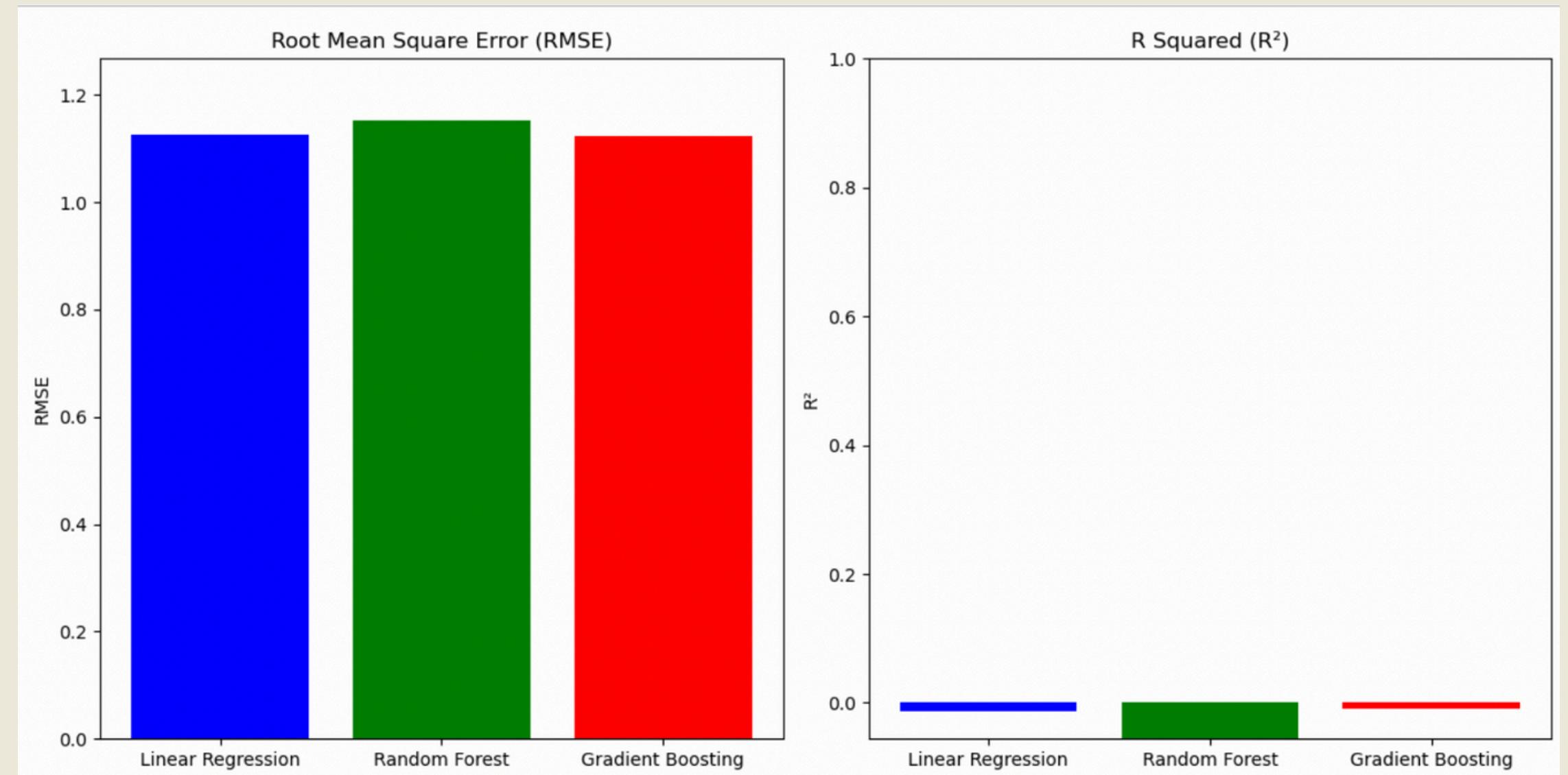
- RSME = 1.125
- R^2 = -0.012

Random Forest Regressor

- RSME = 1.152
- R^2 = -0.0126

Gradient Booster Regressor

- RSME = 1.122
- R^2 = -0.008



All underperform a simple model

None are great predictors of review rating

Methodology

XGBoost

XGBoost Regressor

- Objective - minimizing squared error

Grid Search to Find Best Hyperparameters

- **max_depth** - control depth of trees and prevent overfitting
- **min_child_weight** - regularization parameter
- **subsample** and **colsample_bytree** - sampling the dataset
- **n_estimators** and **learning_rate** - control the number of boosting rounds and the learning step size

The XGBoost logo is displayed in a large, bold, blue sans-serif font. The word "XGBoost" is written in a single line, with the "X" being smaller than the other letters.

729 unique combinations of hyperparameters

3 values for each parameter

3-fold cross validation

Results

Best Parameters and Hyperparameters



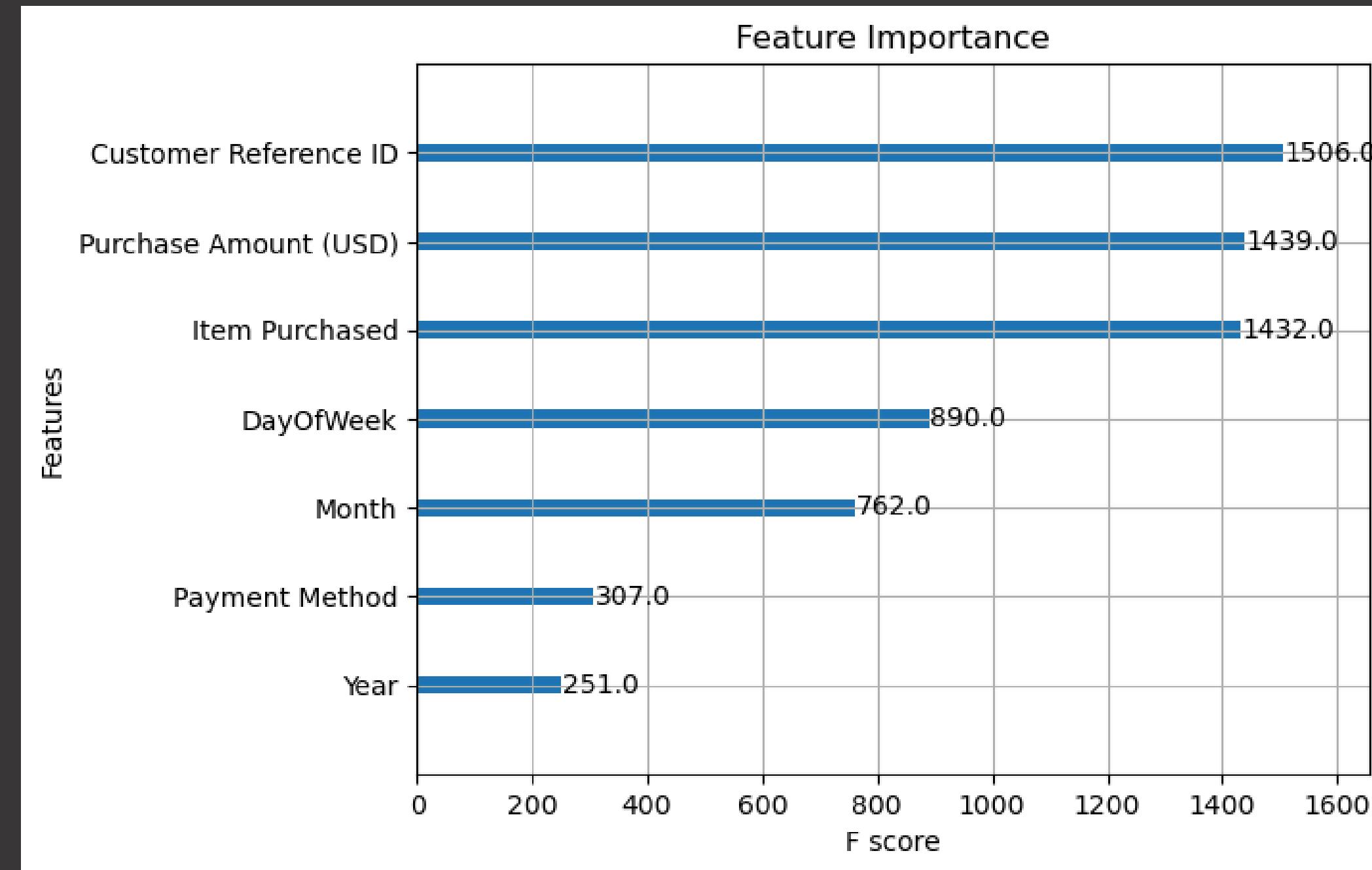
Optimal Hyperparameters

- **max_depth** = 7
- **min_child_weight** = 1
- **colsample_bytree** and **subsample** = 0.5
- **n_estimators** = 100
- **learning_rate** = 0.01

Validation Set MAE = 0.929

Test Set MAE = 0.933

Feature Importance



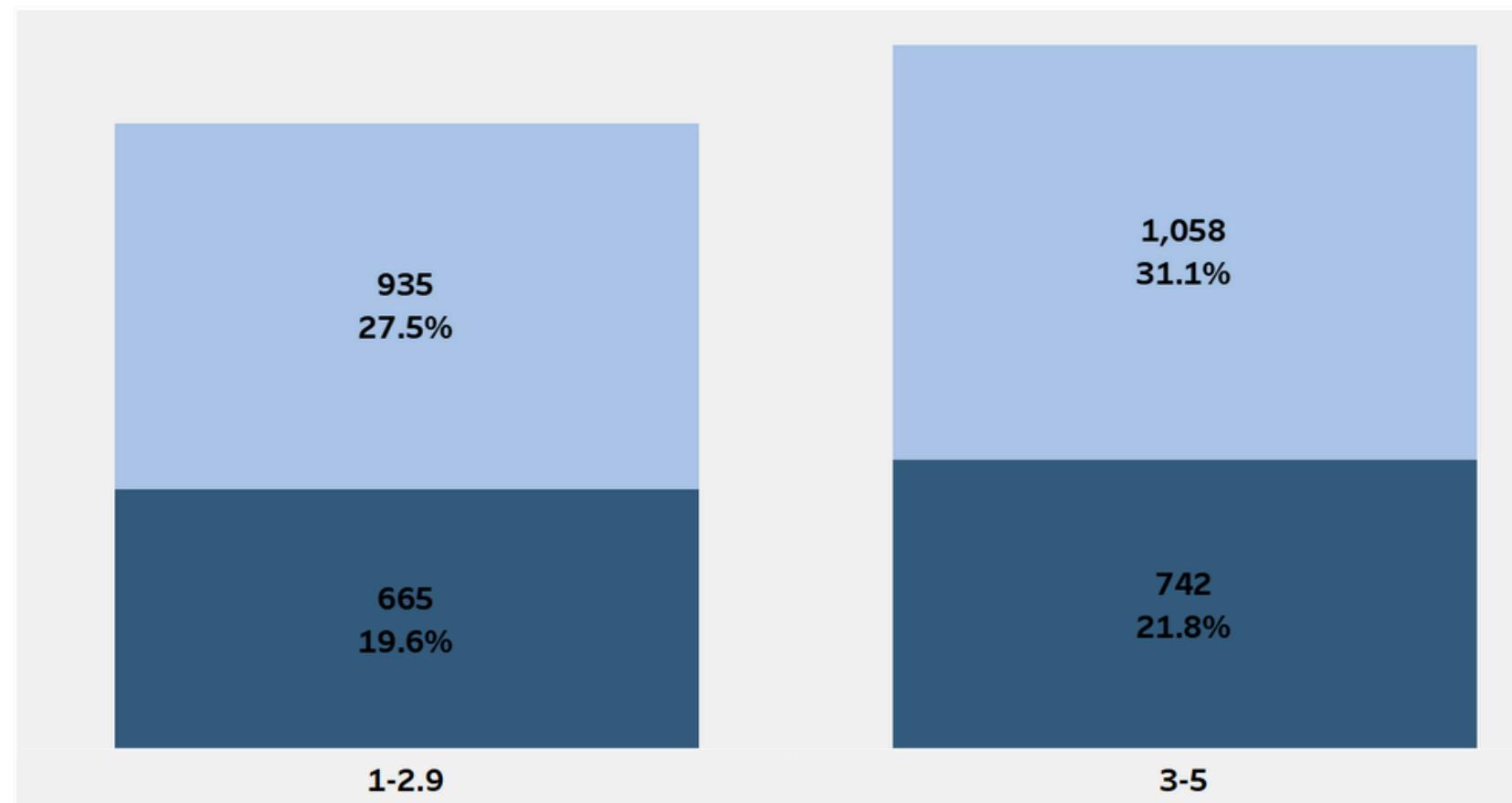


Diagnostic Analytics

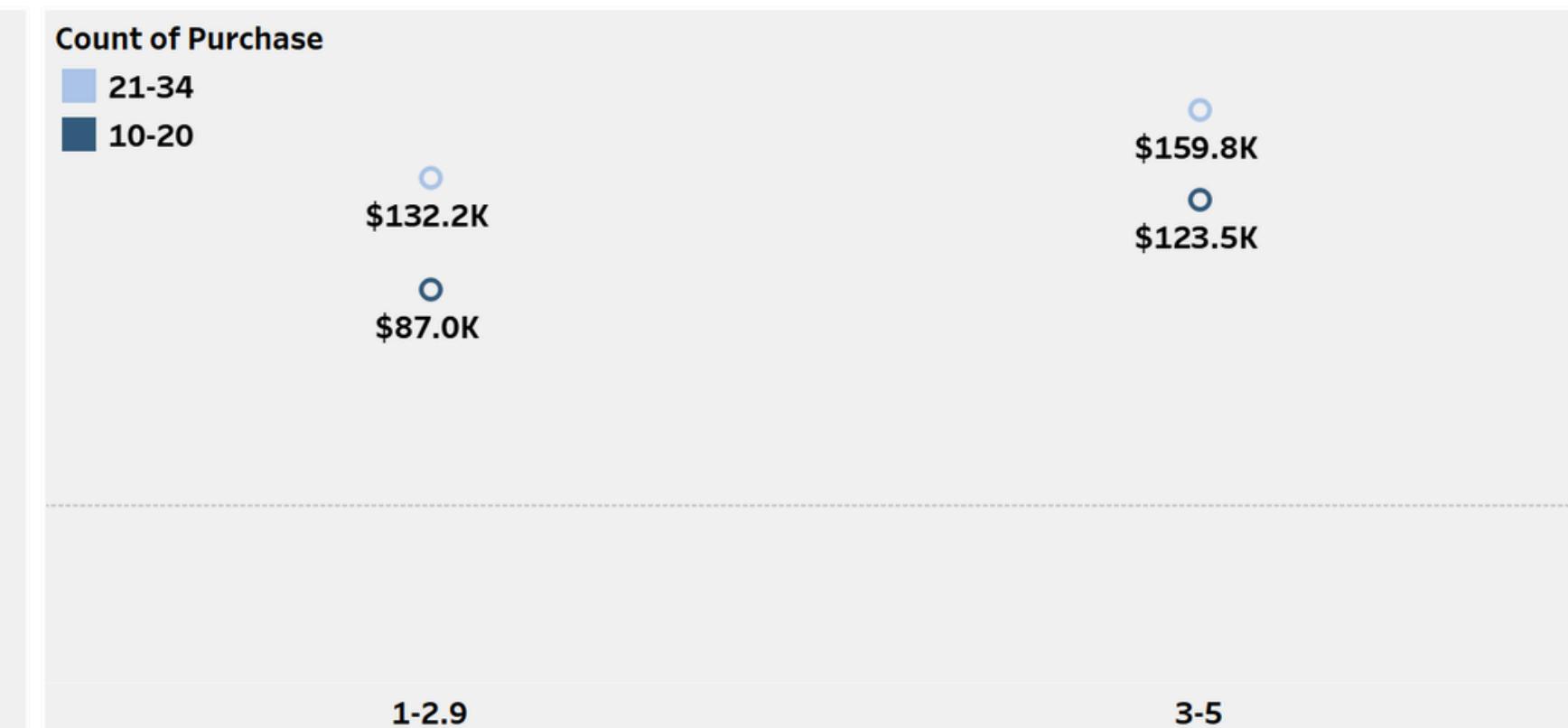


Fashion Company Customer Satisfaction Diagnostic Dashboard

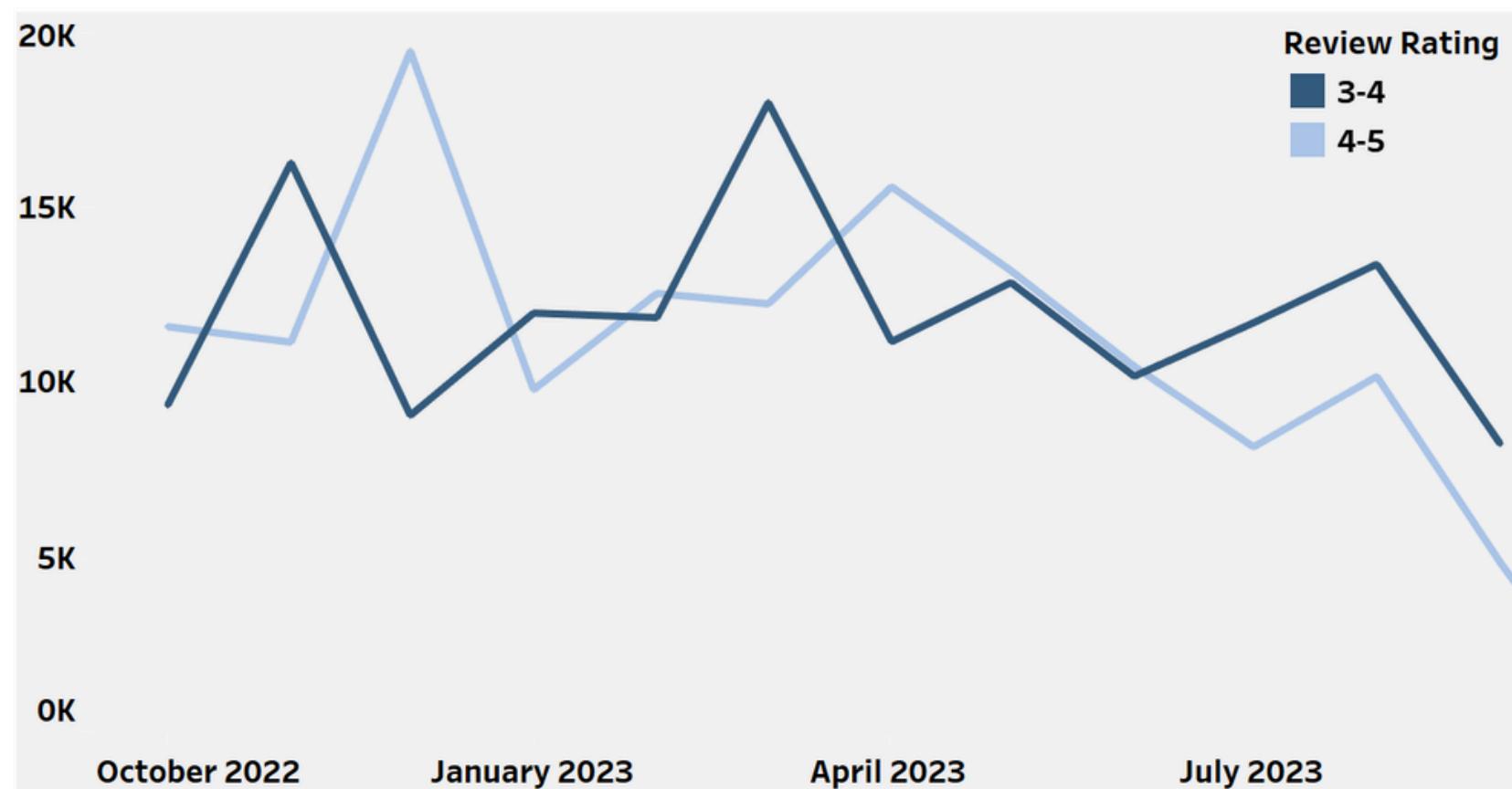
Count of Repeat Purchases by Ratings



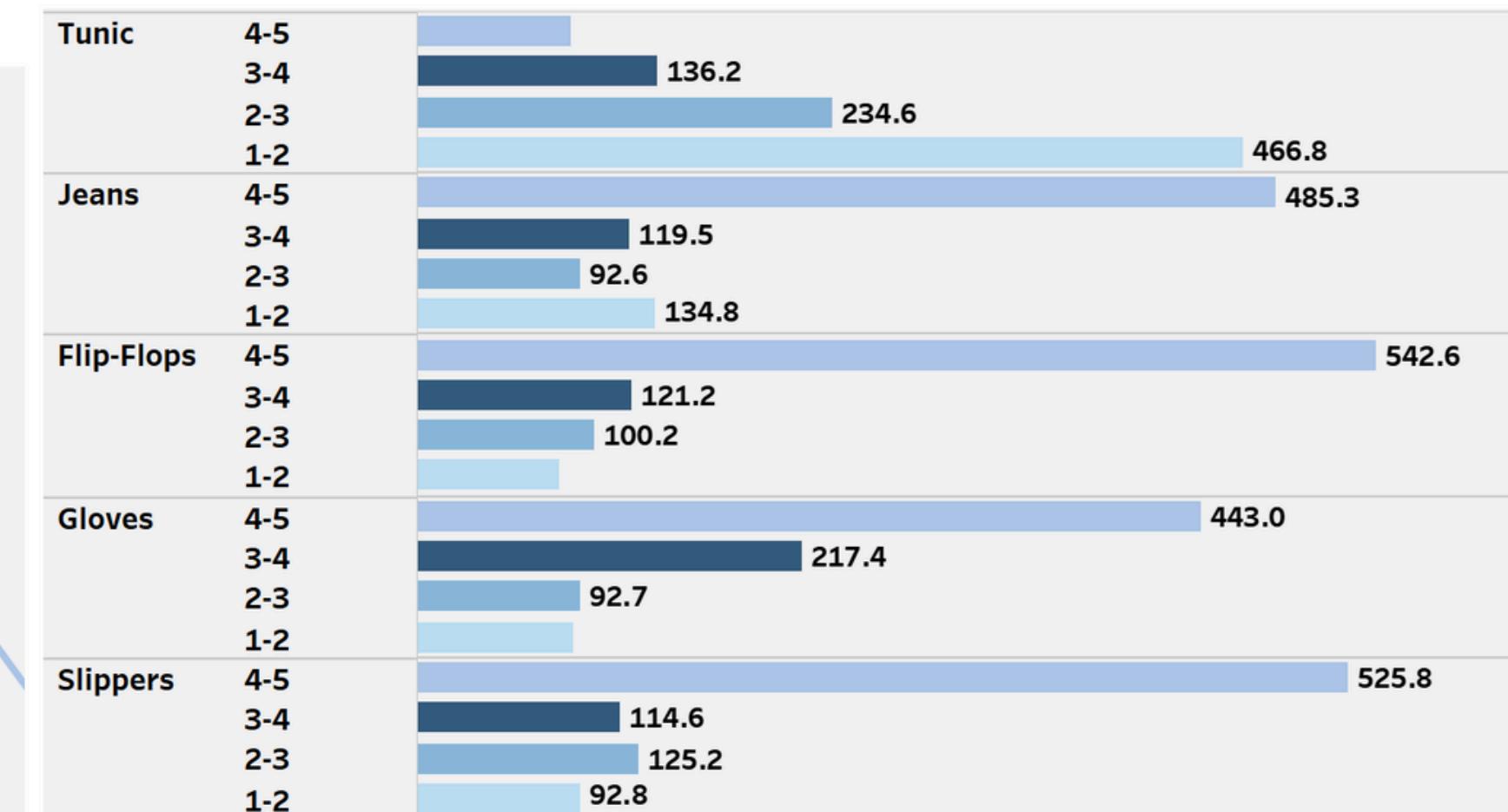
Total Revenue by Rating



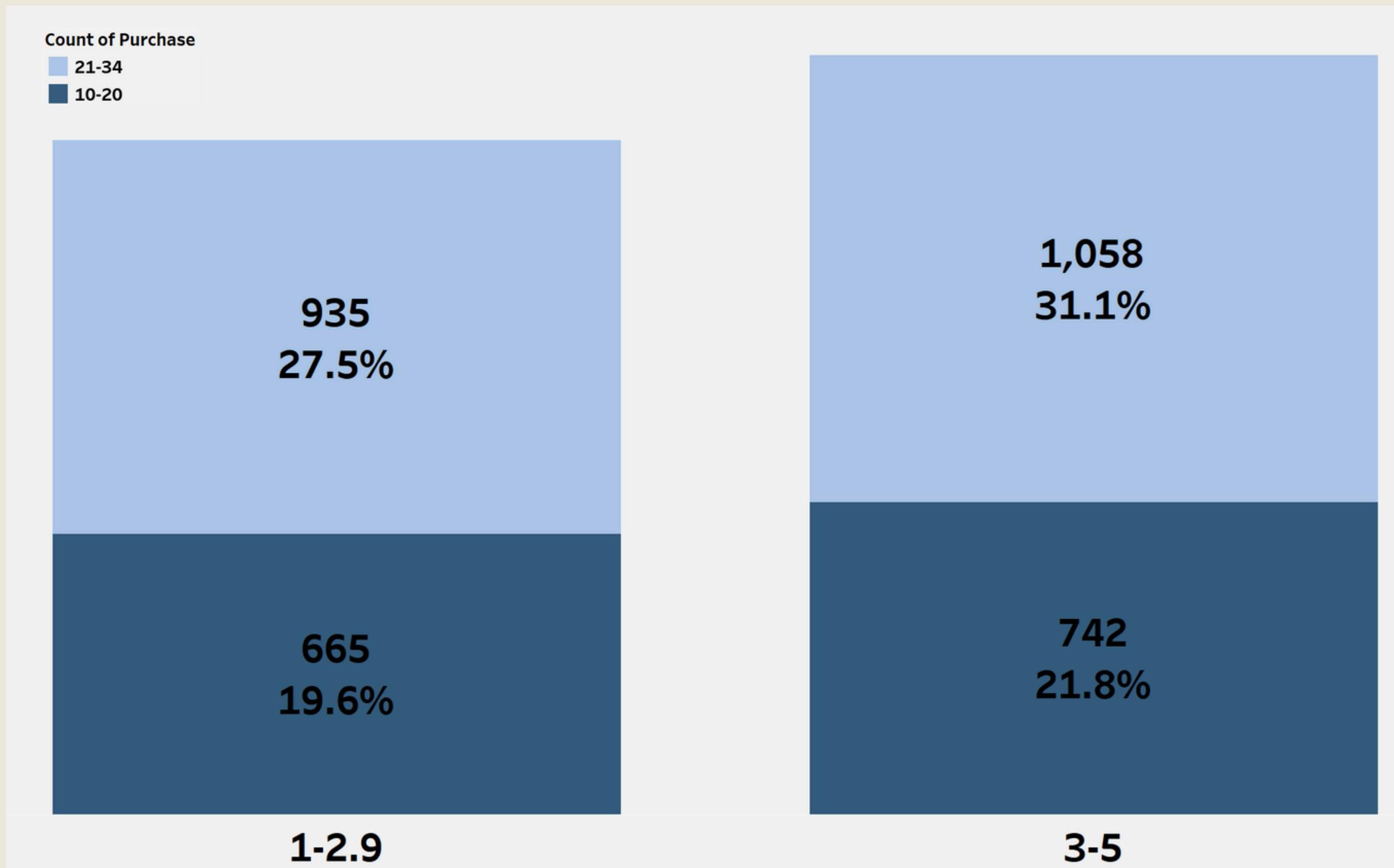
Trend of Total Revenue by Ratings (3-4 and 4-5)



Average Revenue of Top 5 Items



Importance of Loyalty: Repeat Purchases are More Frequent in Higher Rating Categories



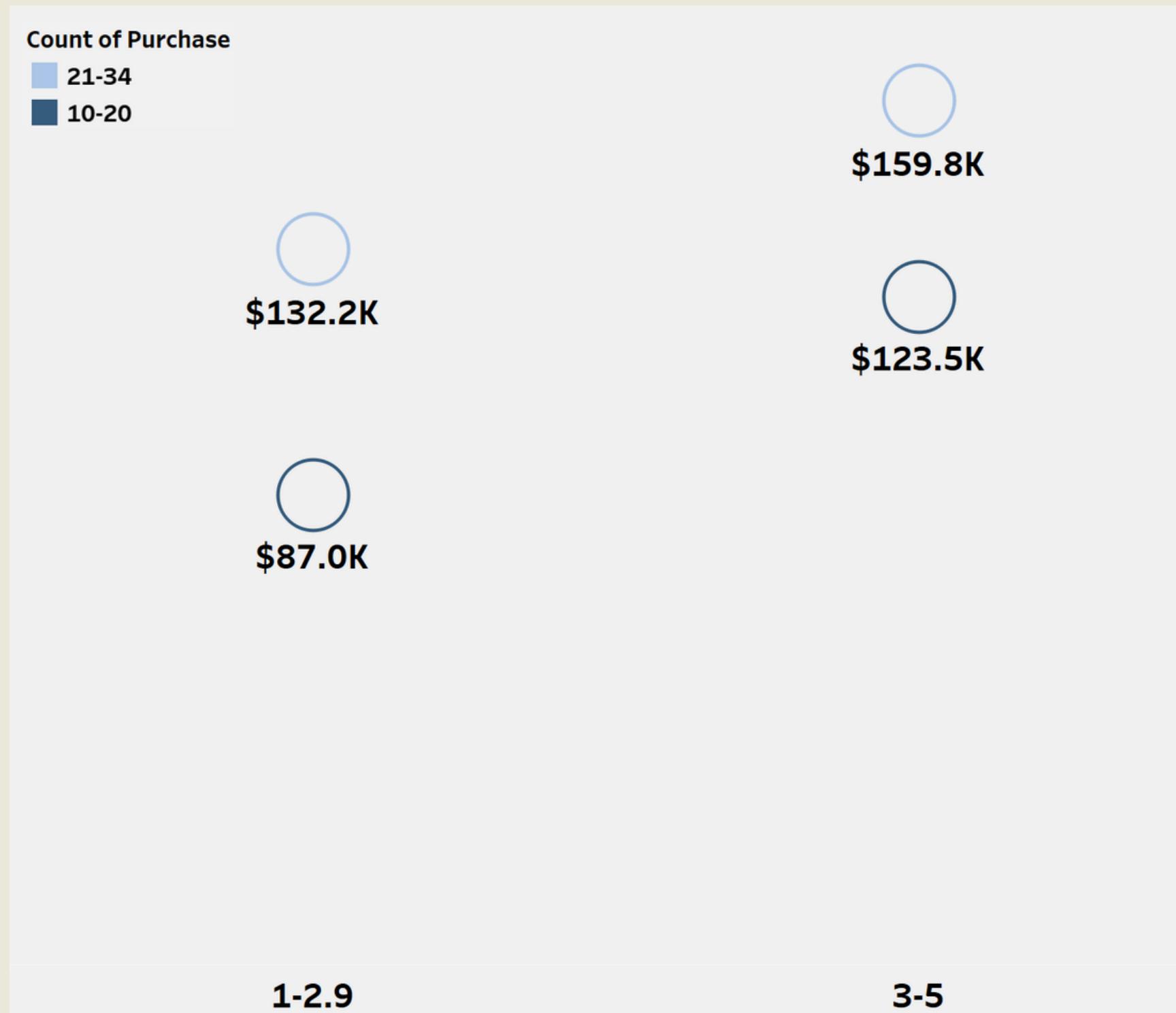
Recommendation

- Increase product quality
- Incentivize with loyalty programs

Prediction

- Increasing purchase rates will increase revenue from high rated items

Repeat-Customer Purchases Also Drive More Revenue



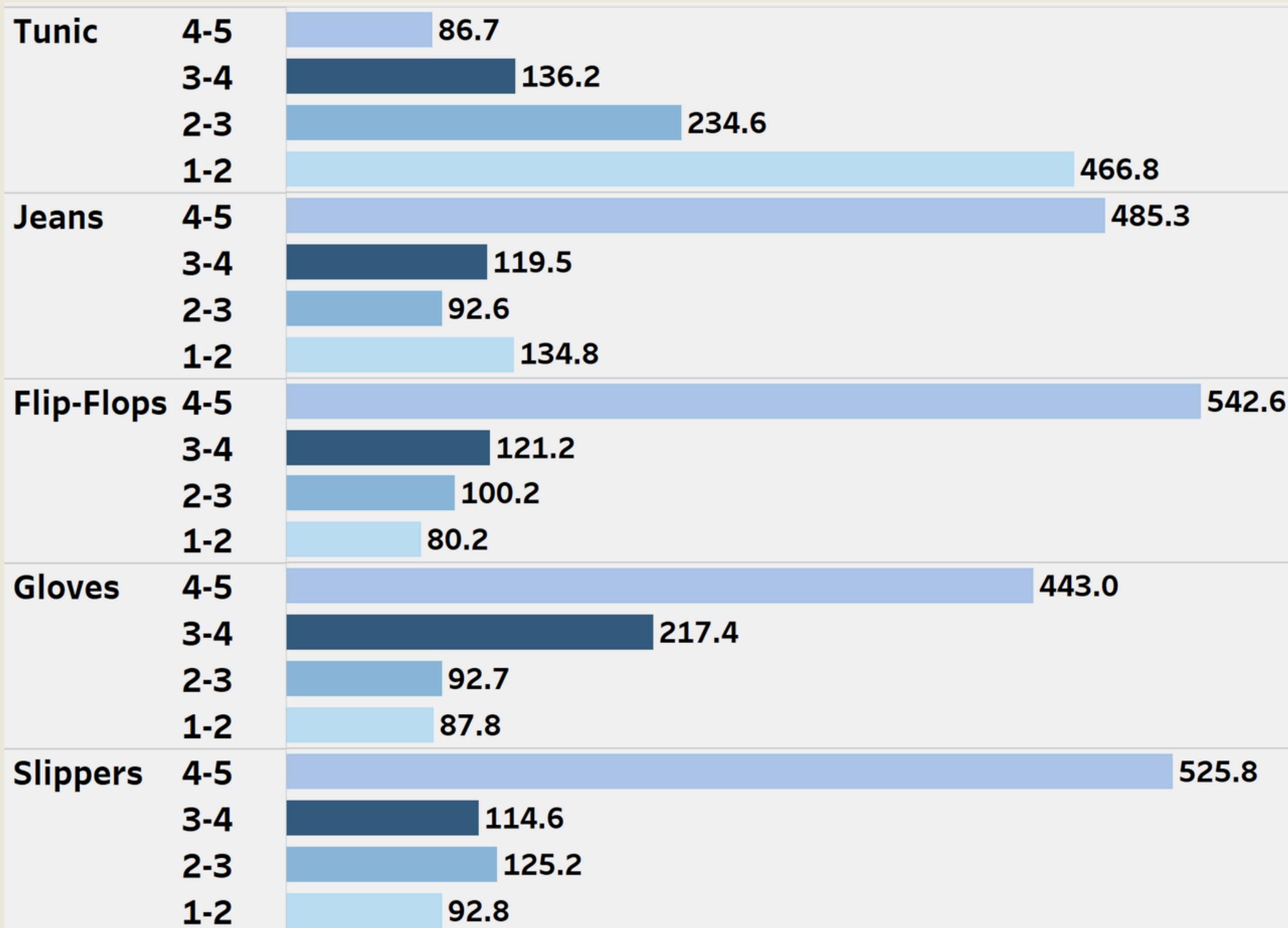
Recommendation

- Focus on quality improvements
- Increase customer service efforts for lower-rated items

Prediction

- Revenue from lower-rated item will align with revenue of higher-rated items

Average Revenue of Top 5 Items Correlates Mostly with Ratings



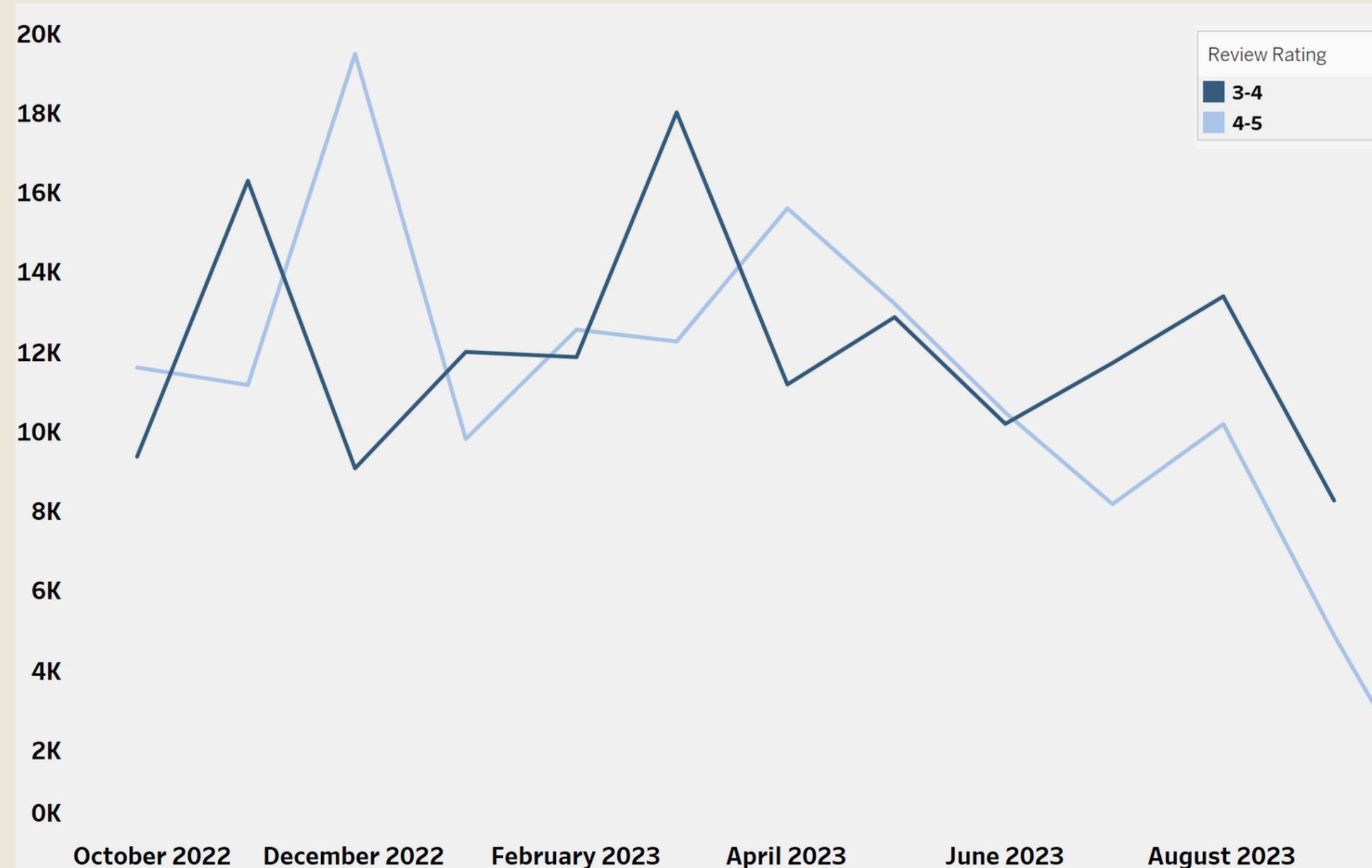
Recommendation

- Focus marketing on high revenue products

Prediction

- Focus on top-performing items will likely increase revenue concentration

Revenue from 4-5 Rating Group is more Nuanced



Recommendation

- Personalized marketing campaigns during low points

Prediction

- Increasing sales volume from previously low sales times will increase revenue and customer satisfaction



A/B testing & more



Key Metrics

Key Metric

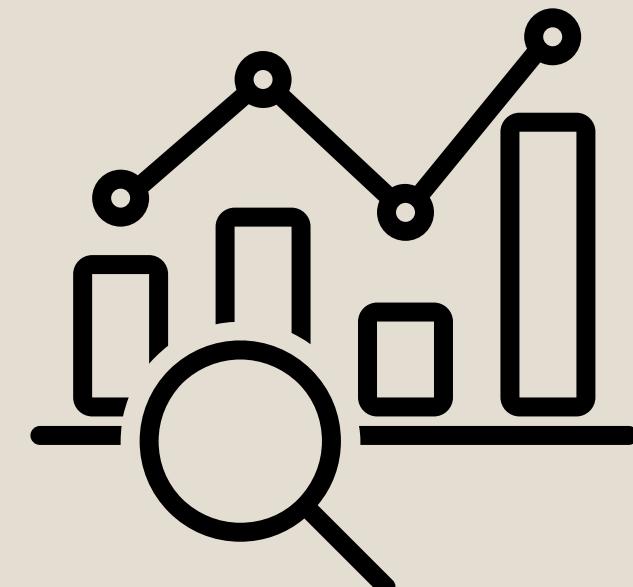
- *Conversion Rate*

Alignment

- *More effective in driving sales. Increasing revenue and customer engagement.*

Why?

- *Personalized recommendations*



Hypotheses for Success

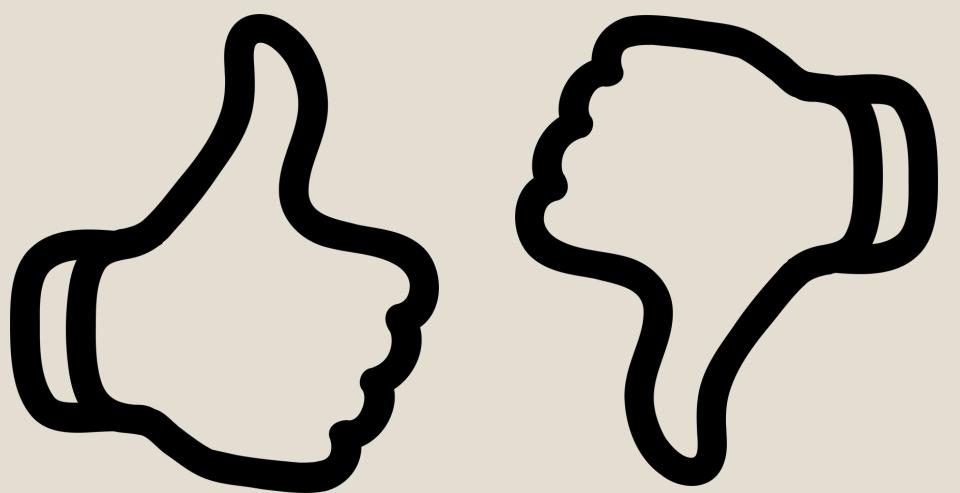


H_0 = Compared to generic recommendations, personalized recommendations has no affect on user engagement

H_1 = Personalized recommendations results in higher user engagement

**Utilize AI
for...**

- Machine Learning
- Mention Time Learning
- Monitoring and ptimization



Seamless User Randomization



JavaScript random number generator

Balanced distribution 50/50 split

Persistent grouping

Users retain their assigned group for study

Sample Size for Reliable Insights

Establishing our foundation:

A baseline average product rating of 3.0

Minimum Detectible Effect:

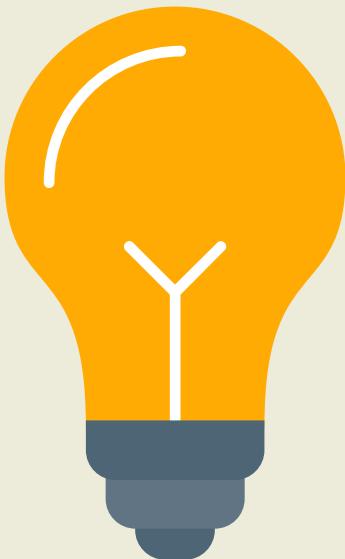
Detecting a 0.1-0.2 increase in ratings.

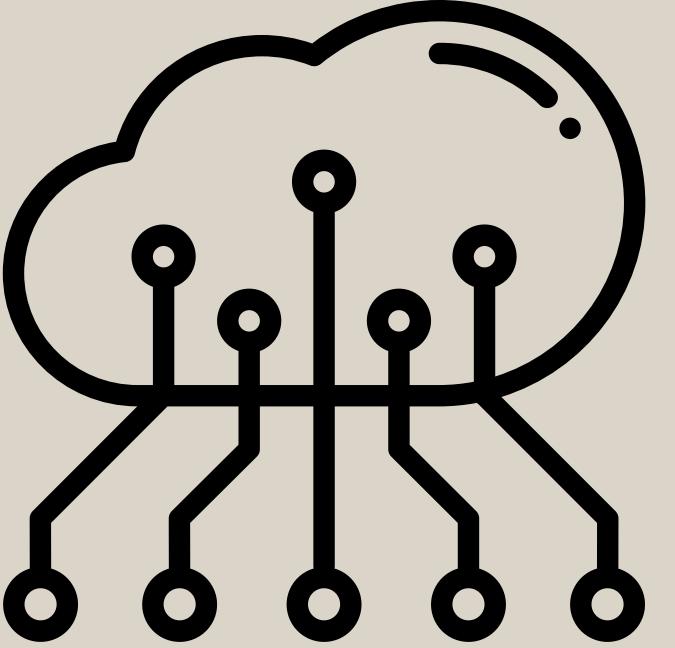
Calculated precision:

80% power to catch true effects, 5% α for result confidence

Strategic allocation:

1,000 visitors daily a 4-week duration for data collection





Technical Requirements & Timeline

Timeline

Phase	Weeks	Description
Preparation	1-2	Configure Test Platform
Execution	4	Run A/B testing with users
Analysis	1-2	Challenge null hypothesis
Reporting	1	Recommendations to stakeholders

Technical Requirements

- Testing Platform
- Integration
- Data Tracking
- Local Storage

Commitment to Ethical Testing and Bias Management



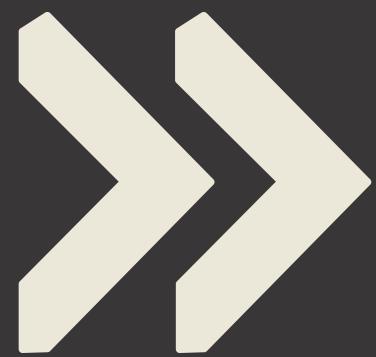
Transparency & Informed Consent



Bias Identification & Correction

Responsive Adjustment

From Data To Decisions



Collection, Analysis, & Significance

Data Collection & Segmentation

- Google Analytics
- Cookie-stored IDs



Analytical Depth Significance

- Python and Excel
- Statistical significance α of 0.05

DATA MANIFESTO

Differentiate

Cultivating a Data-Informed Culture

Data Literacy as a Foundation

- Employees
- Training

Advancing Analytics Maturity

- Invest
- Enhance capabilities

Data-Informed Decision Making

- Cross-functional teams
- Integrate insights

Sustain Competitive Advantage

- Anticipate market trends
- Feedback loops



Commitment to Data Excellence

improve customer experience
and drive growth!

Vision for the Future

Every decision strengthens
our market position!

