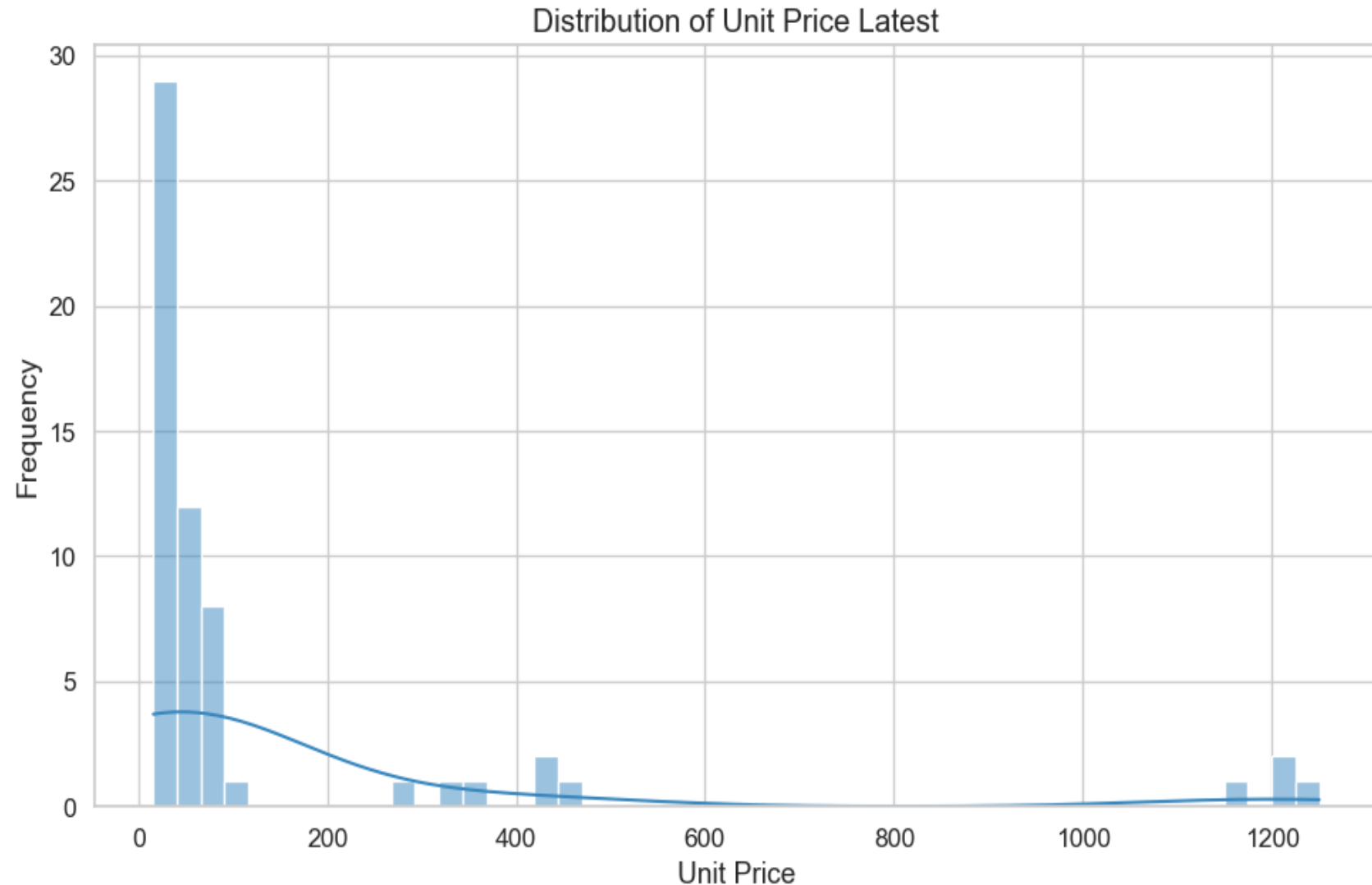


# "Exploratory Data Analysis of Chemical Vendor Pricing and Quality

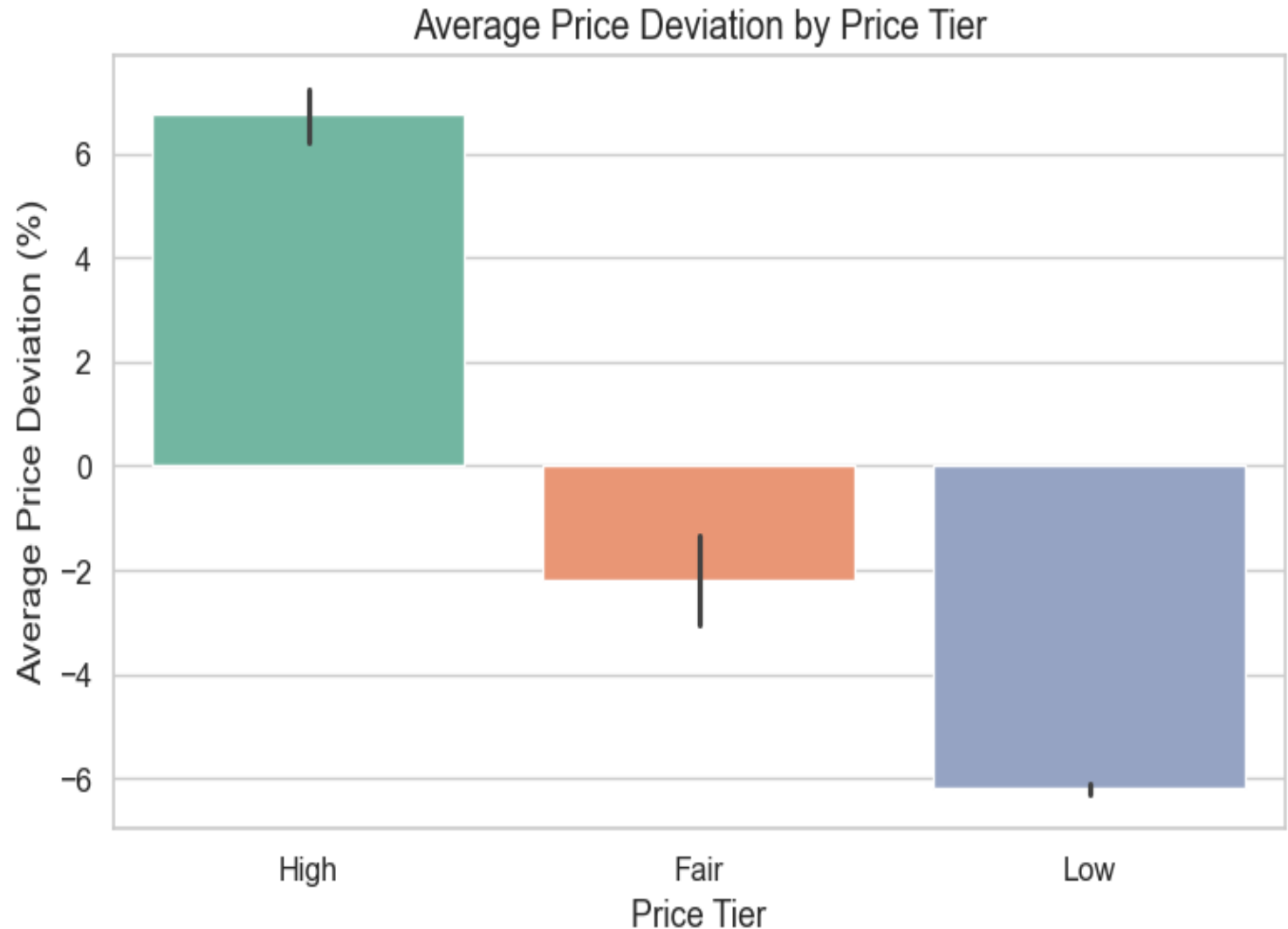
# Distribution of Latest Unit Price

- Most chemicals are priced in the lower range, with distribution heavily skewed to the right.
- A small number of specialty products command high prices, creating a long tail in the distribution.
- Indicates the presence of both commodity and high-value items in the dataset.



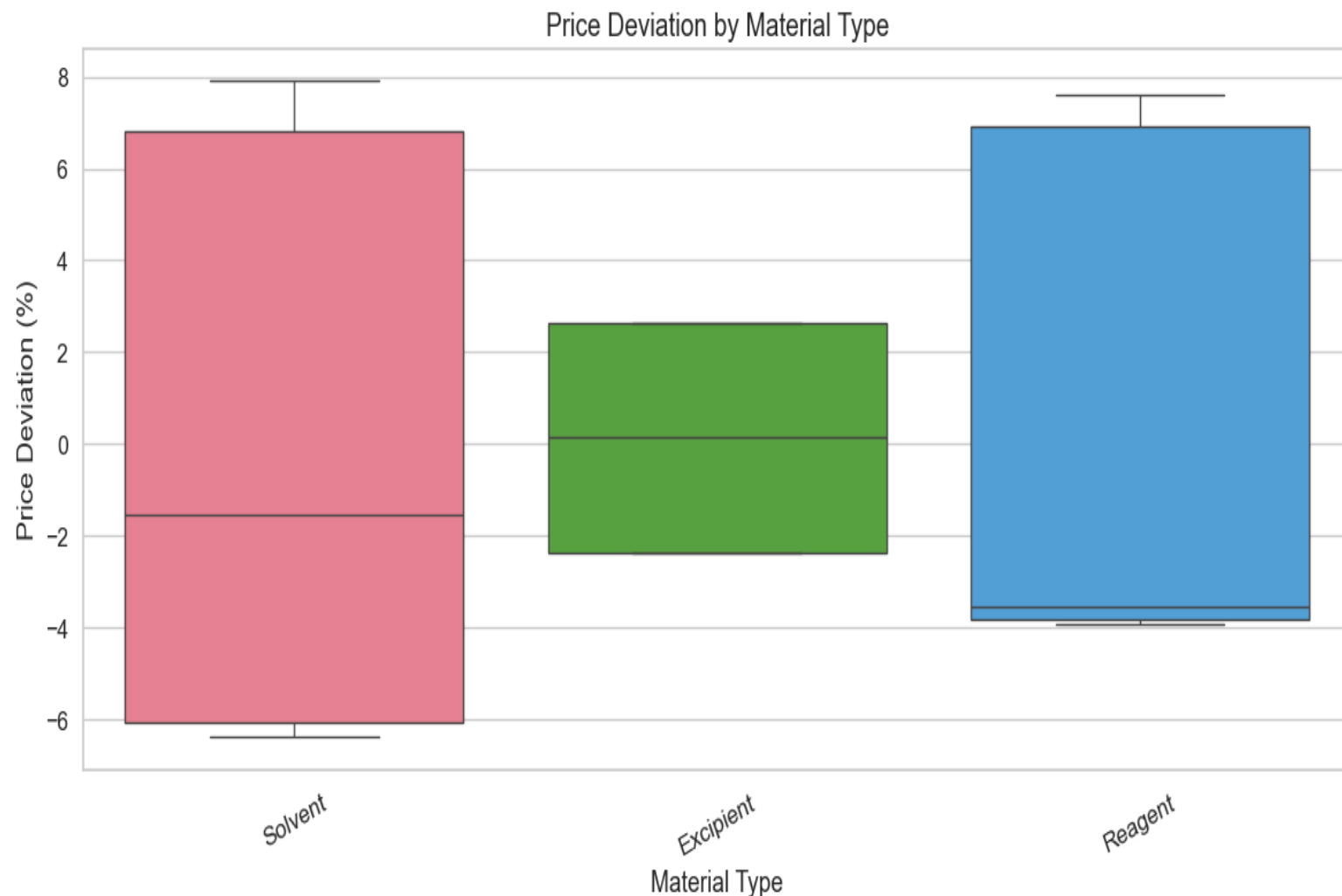
# Average Price Deviation by Price Tier

- High tier products have distinctly positive price deviations, reflecting premium pricing above market benchmarks.
- Fair tier items mostly align with benchmarks and show minor negative deviations.
- Low tier products consistently show negative deviations, indicating discounted vendor pricing.



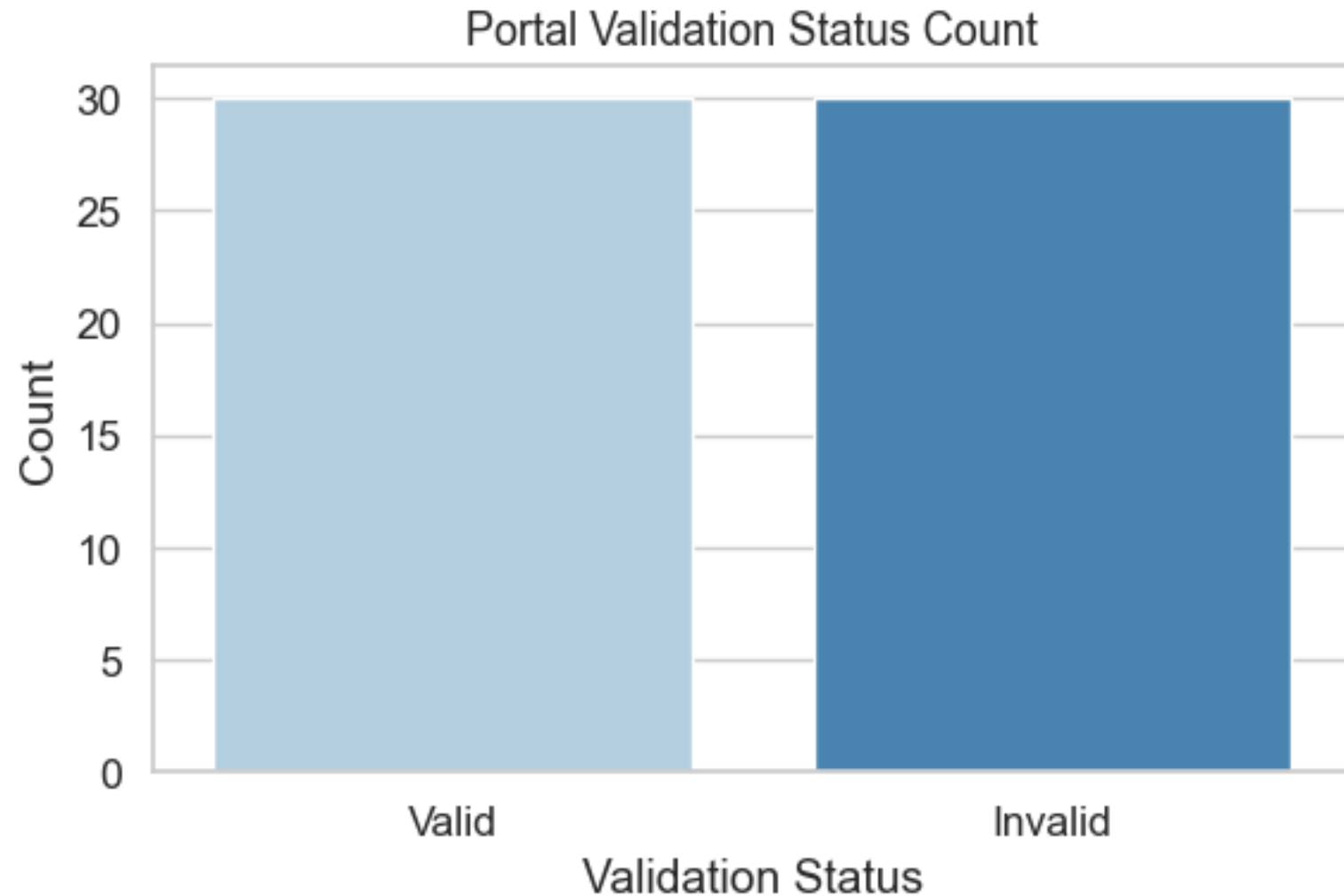
# Price Deviation by Material Type

- Solvents and reagents exhibit a wide spread in price deviation, indicating greater variability compared to benchmark prices.
- Excipients show lower variance, with deviations centered closer to zero, suggesting more stable pricing.
- Reagents tend to have consistently negative deviations, while solvents span both positive and negative.



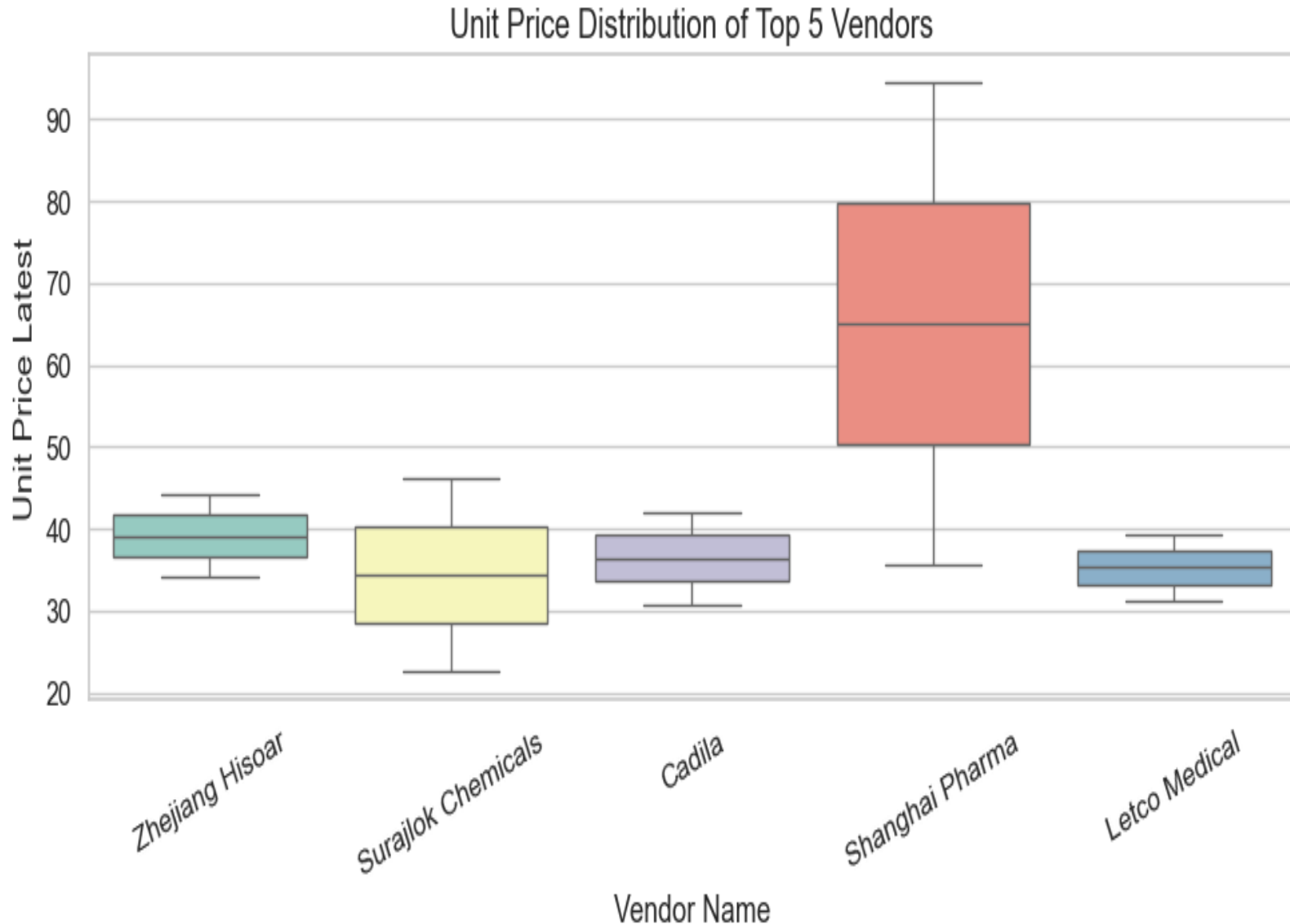
# Portal Validation Status Count

- The count of valid and invalid portal entries is nearly balanced, suggesting the data includes both highly reliable and potentially less trustworthy sources.
- Regular review of invalid entries may be needed for procurement accuracy.



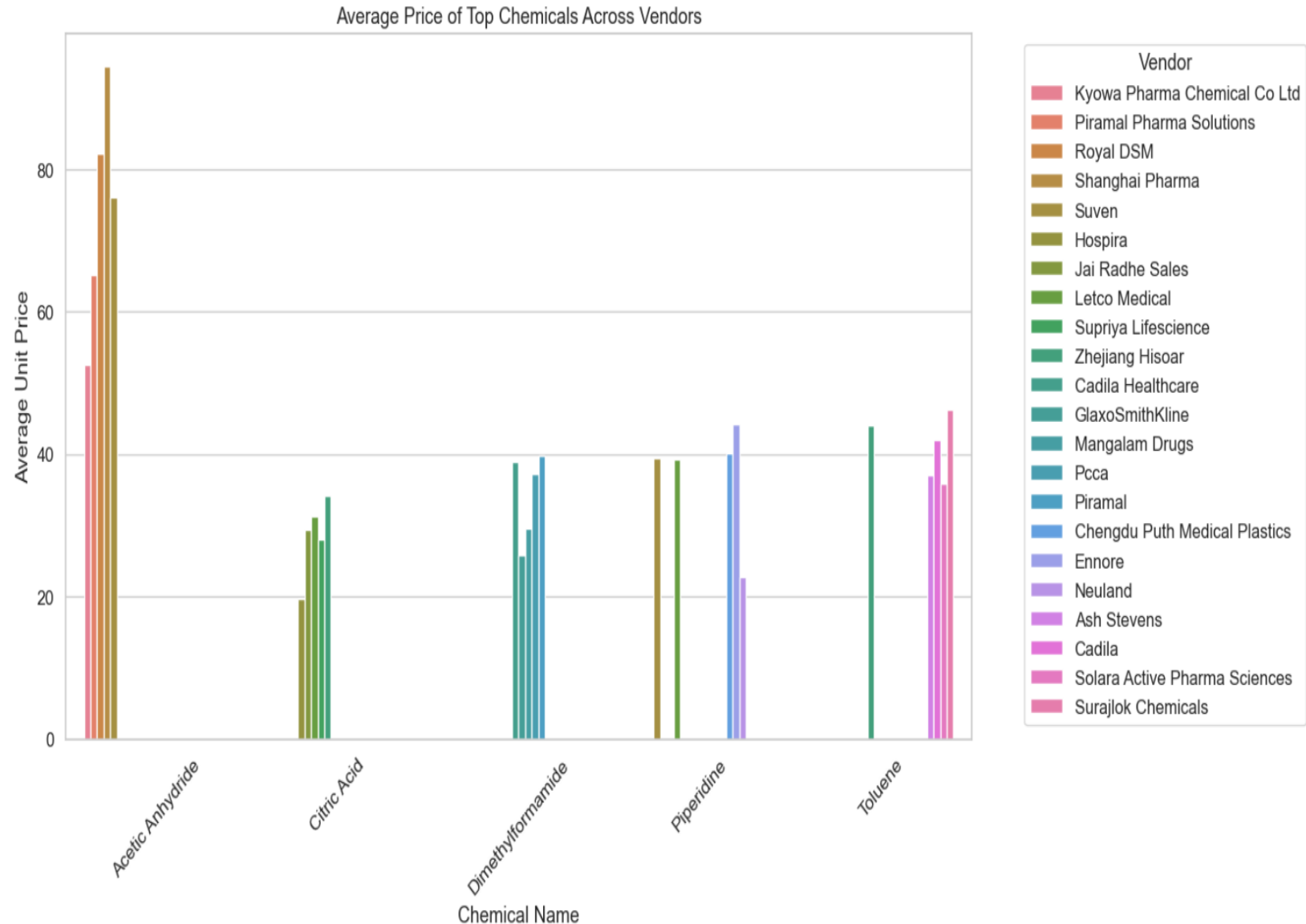
# Unit Price Distribution of Top 5 Vendors

- Most vendors have narrow, consistent price ranges.
- Shanghai Pharma shows wide price variation, suggesting diverse offerings or inconsistent pricing.



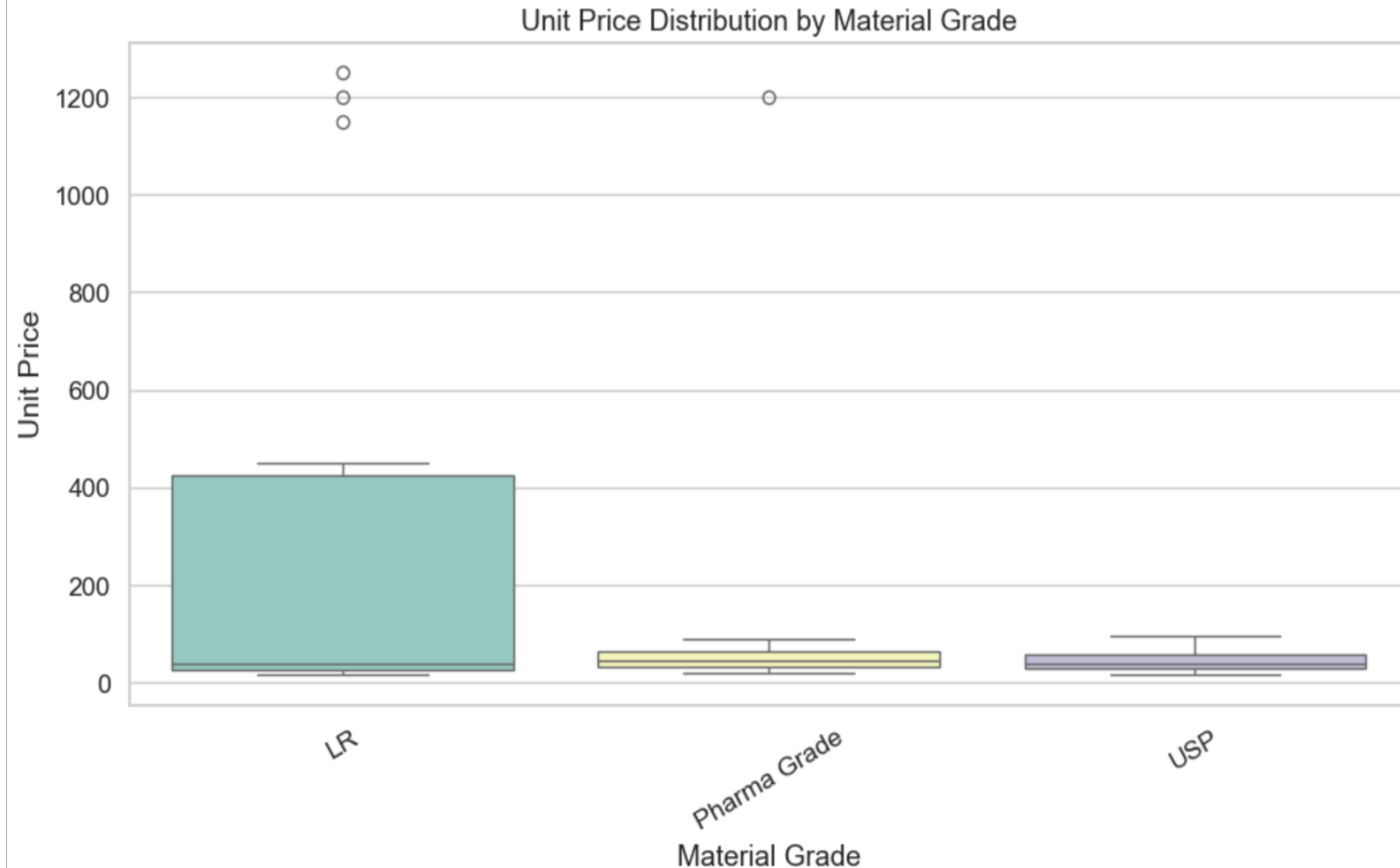
# Average Price of Top Chemicals Across Vendors

- Most vendors have narrow, consistent price ranges.
- Shanghai Pharma shows wide price variation, suggesting diverse offerings or inconsistent pricing.



# Unit Price Distribution by Material Grade

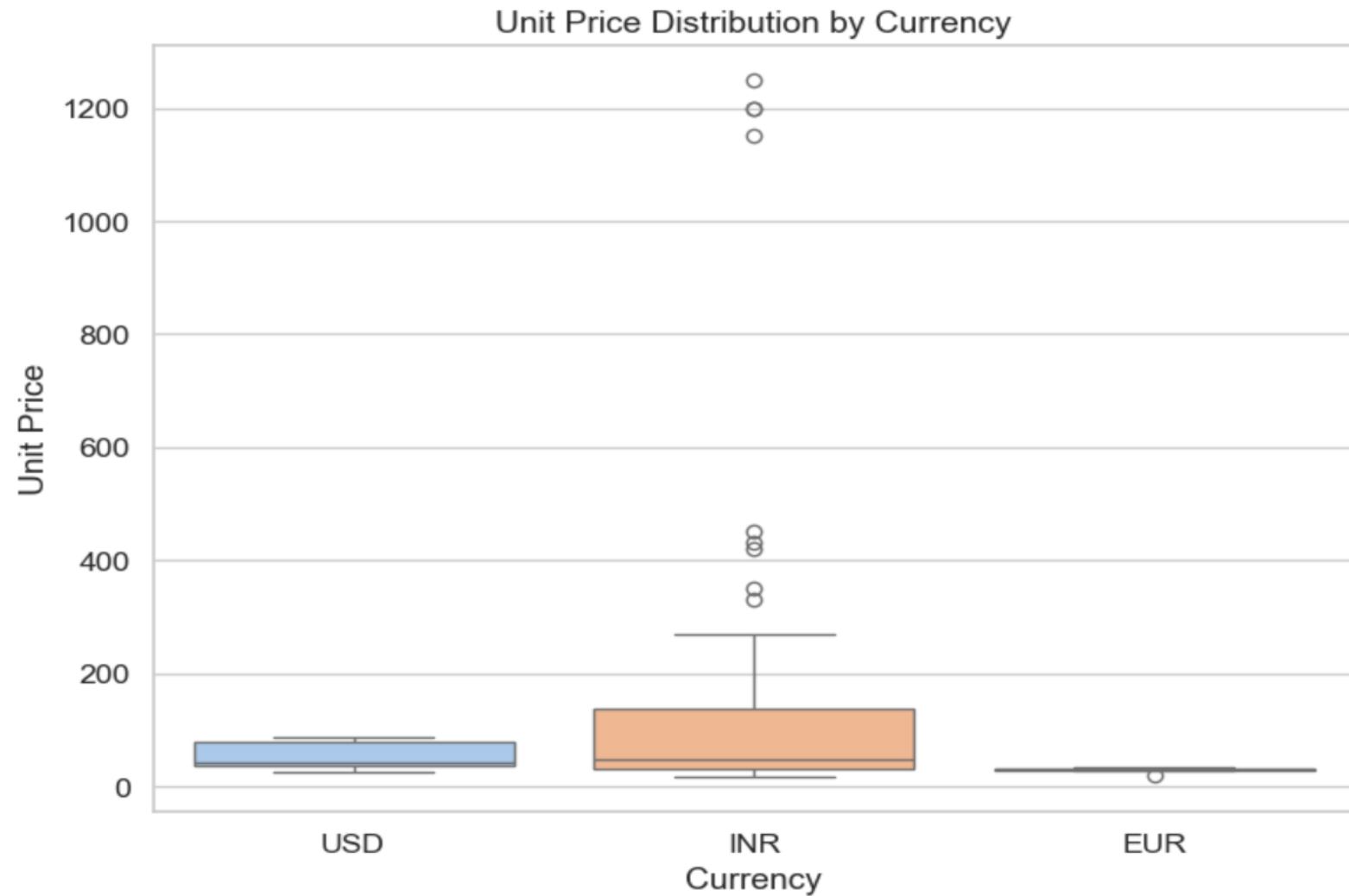
- LR (Laboratory Reagent) grade chemicals show higher prices on average and greater price dispersion, including several notable outliers at the very high end.
- Pharma Grade and USP chemicals have lower and more consistent pricing ranges compared to LR.
- The result suggests that higher-grade specifications, especially LR, can command a significant price premium and exhibit more market variability.





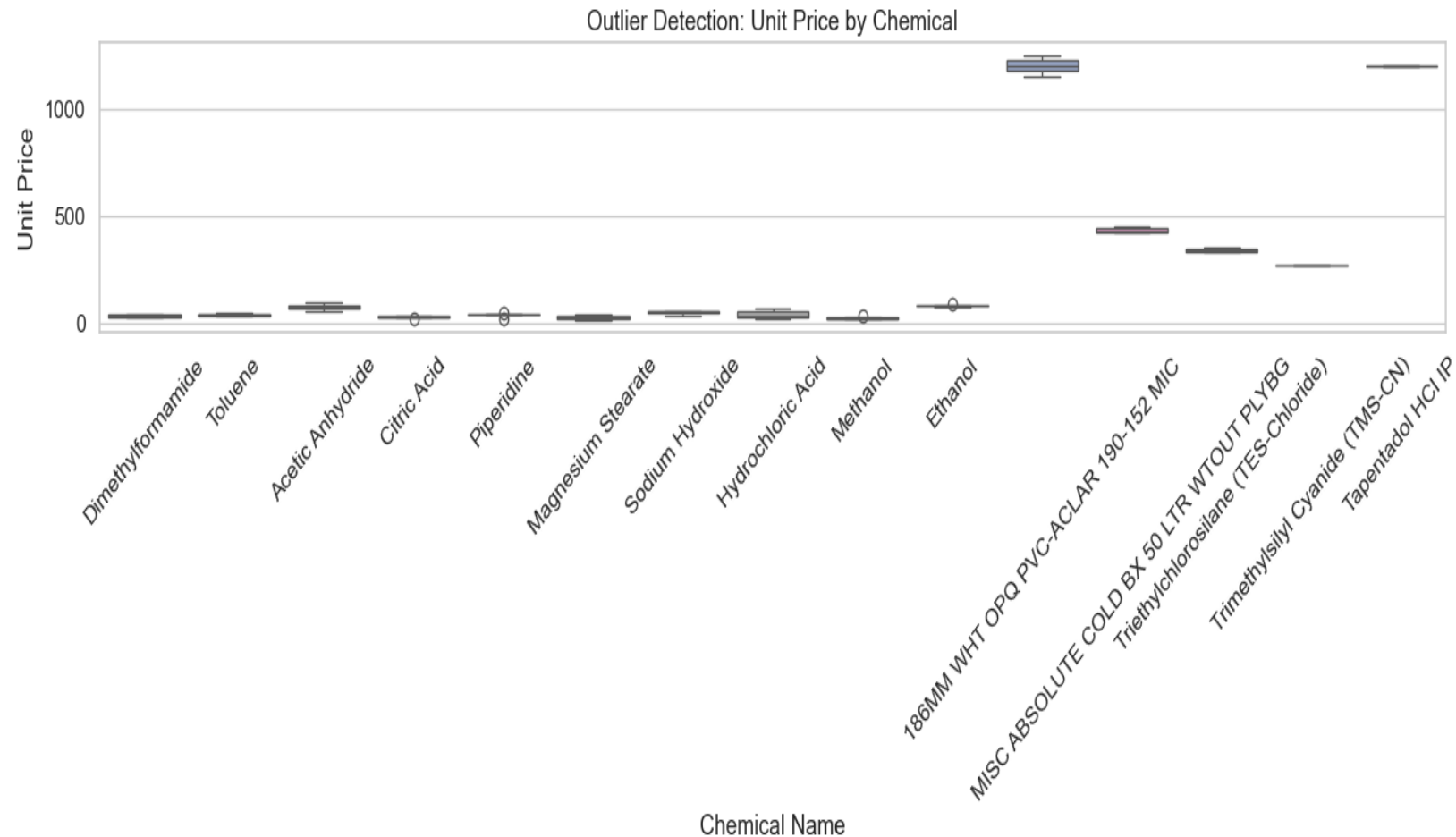
# Unit Price Distribution by Currency

- Prices quoted in INR display the widest price range and include the most extreme outliers, indicating occasional very high-value transactions.
- USD prices are generally moderate and more tightly clustered, suggesting more consistent pricing in international transactions.
- EUR transactions are comparatively few, with the tightest price band.
- Currency has a strong effect on apparent price distribution, likely reflecting both market/geographical and transactional factors.



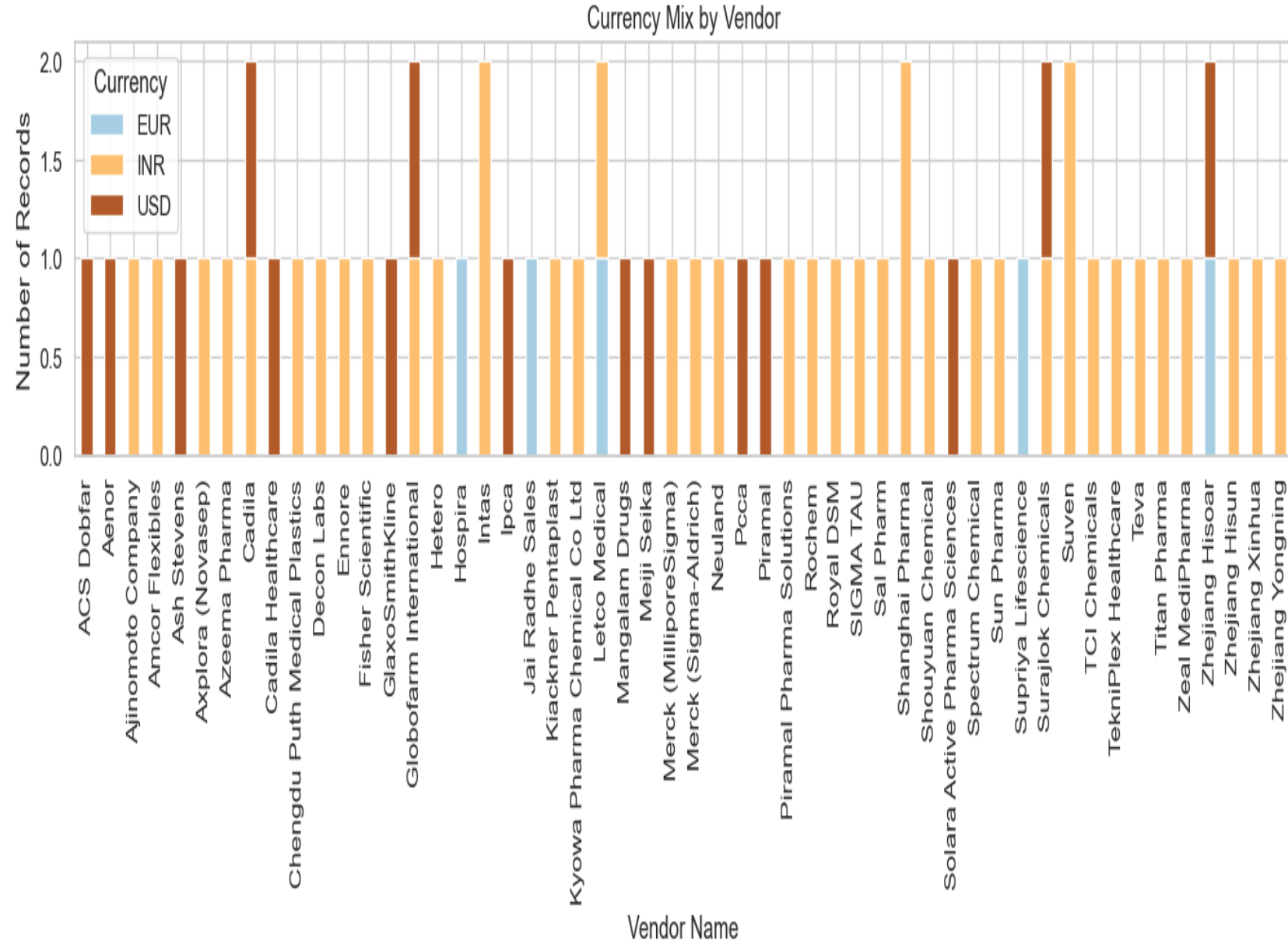
# Outlier Detection: Unit Price by Chemical

- There is a combination of raw material and packaging material into it and the cluster at relatively low and stable prices, but several specialty chemicals and packaging (such as “186MM WHT OPQ PVC-ACLAR 190×152 MIC” and “Tapentadol HCl IP”) stand out as extreme outliers with unit prices far above the median.
- These outliers significantly impact summary statistics and highlight where particularly expensive or premium products lie in the portfolio.
- Outlier detection helps flag chemicals whose pricing may warrant closer negotiation, quality review, or supplier diversification.



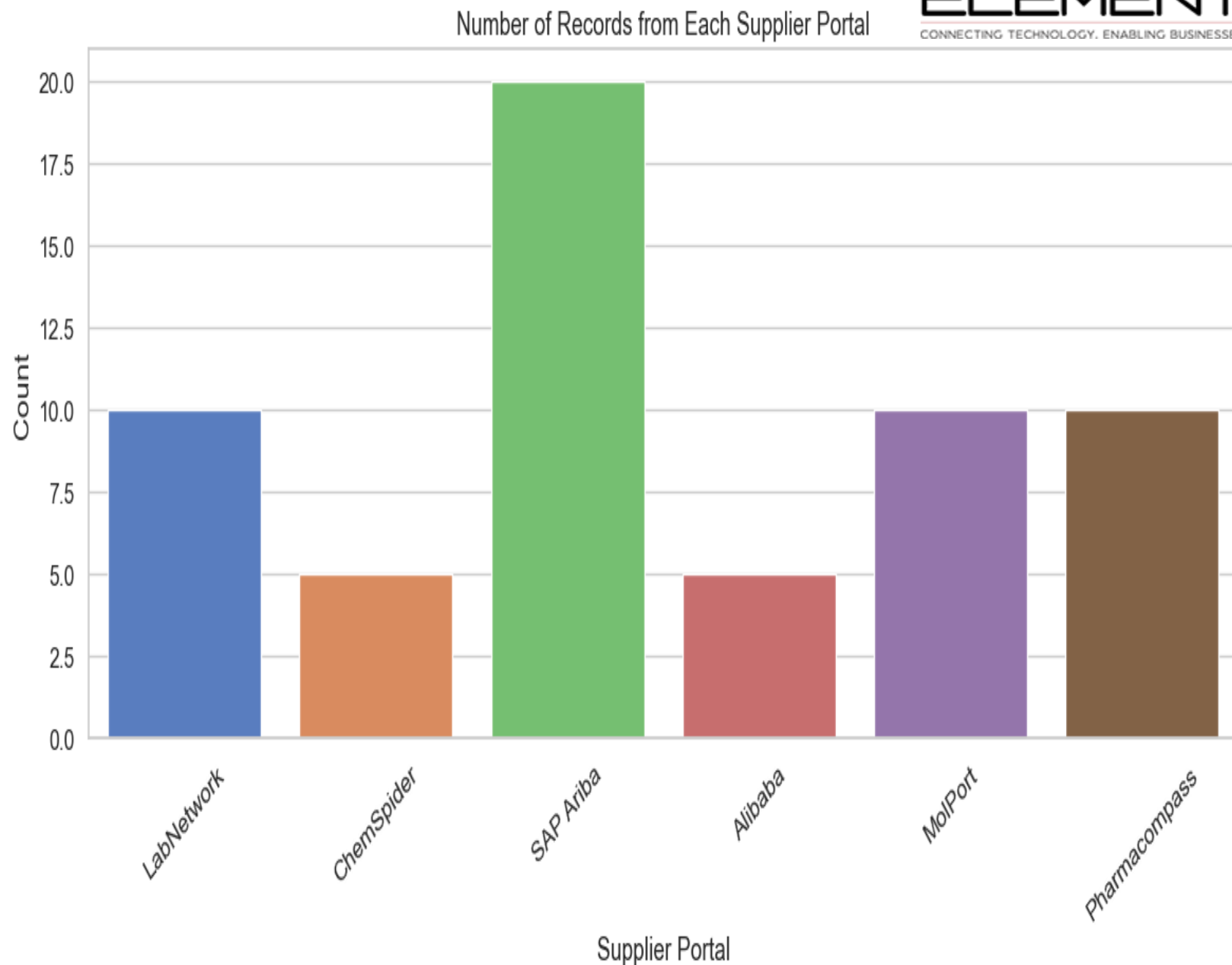
# Currency Mix by Vendor

- Most vendors quote prices in a single currency, but a few transact in multiple currencies (INR, USD, EUR), reflecting either a global customer base or flexible invoicing.
- INR and USD are the most commonly used currencies among vendors, with EUR less prevalent.
- Understanding each vendor's currency mix is valuable for procurement planning, as currency fluctuations and conversion costs may affect total spend.



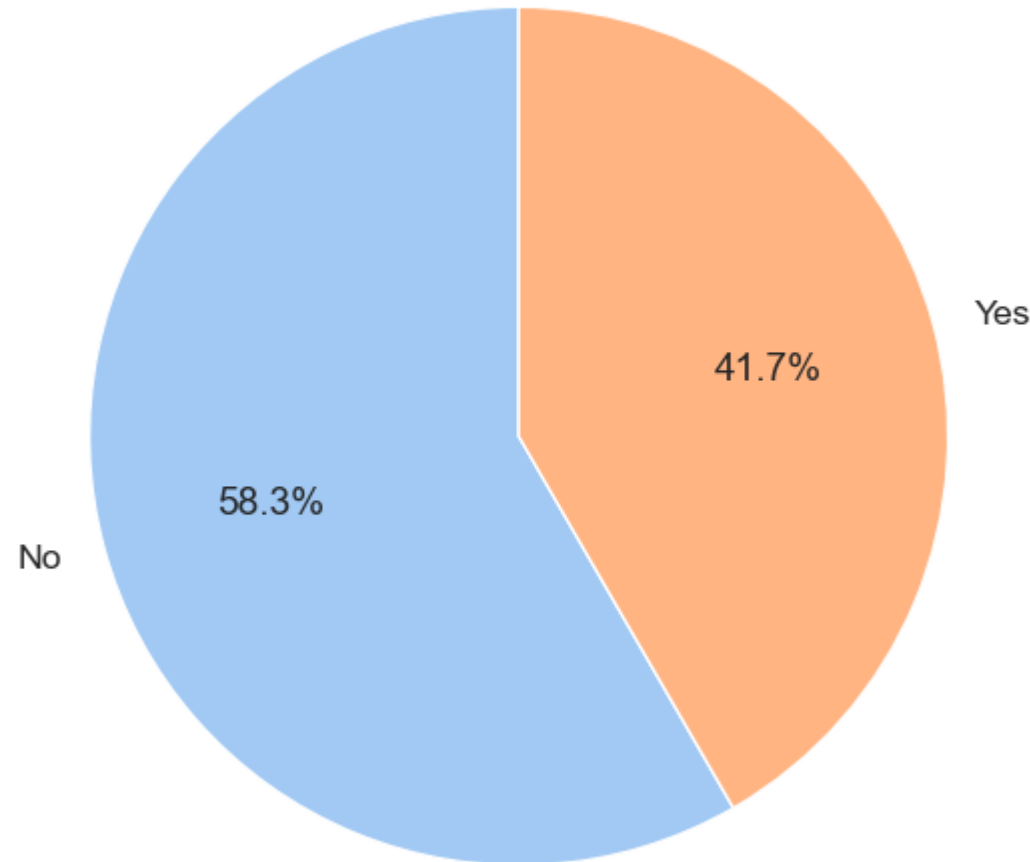
# Number of Records from Each Supplier Portal

- SAP Ariba is the leading data source, providing the most records, which highlights its importance in the procurement process.
- LabNetwork, MolPort, and Pharmacompass contribute a moderate number of records each, suggesting diversified sourcing.
- ChemSpider and Alibaba have the fewest entries, so data from these portals may have less overall influence in analyses or decision-making.



# GMP Compliance Proportion

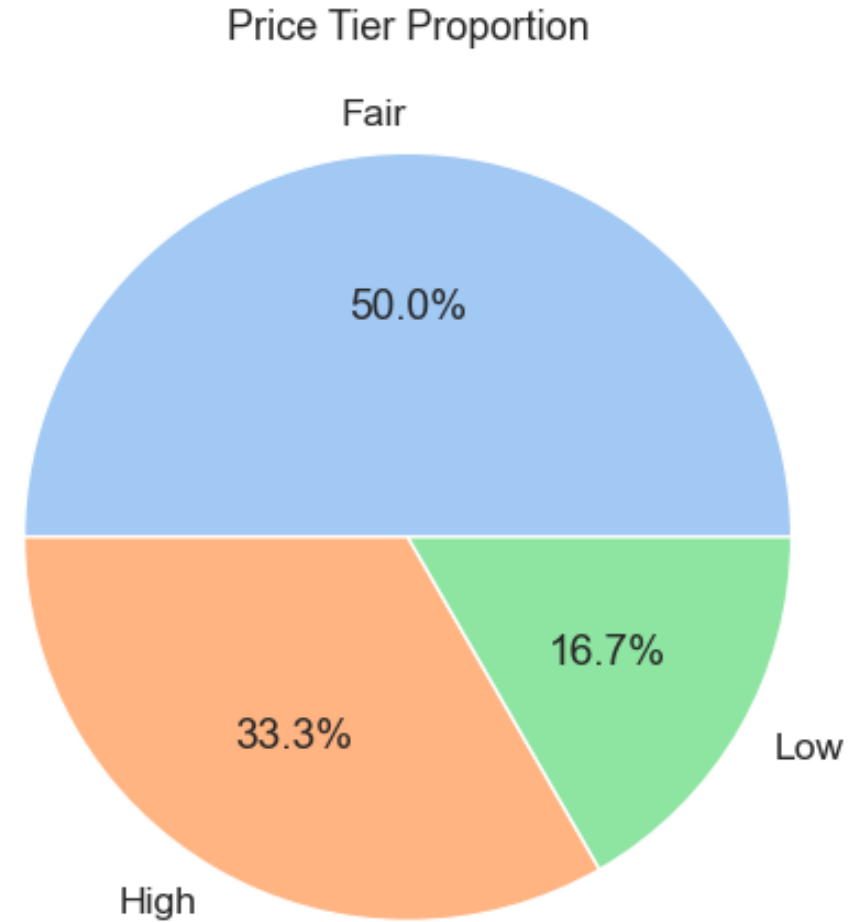
GMP Compliance Proportion



- The majority of materials in the dataset are not GMP compliant (58.3%), while a significant minority (41.7%) are GMP compliant.
- This indicates that although quality assurance is a priority for a significant portion, non-GMP materials still dominate, underscoring the need for stringent quality checks and risk management in sourcing.

# Price Tier Proportion

- Fair pricing tier dominates (50%), followed by High (33.3%) and Low (16.7%).
- Most products are priced near the benchmark; fewer are premium or discounted.



THANK YOU