WINE ANALYSIS

BIDV PROJECT

SUBMITTED TO: MITALI SAIGAL
SUBMITTED BY: SHREYA 22CSU515
ANKIT DAS 22CSU216

Index

Acknowledgment

We want to express our sincere gratitude to The Northcap University, Gurugram for providing the necessary resources, facilities, and support for the successful completion of this project.

A special thanks to my project guide, Mitali Sehgal, for her invaluable guidance and insightful feedback throughout the research process. Her mentorship has been a key factor in the completion of this work.

I am deeply thankful to Francesco Fontanella and his team for providing the DARWIN dataset, which is pivotal in developing predictive models for Alzheimer's disease detection. The dataset offered rich insights into handwriting analysis, greatly aiding the research.

Abstract

Wine is a universally cherished beverage, celebrated for its complexity, cultural heritage, and economic significance. Understanding the intricate relationships between wine attributes—ranging from chemical composition to sensory perception and market dynamics—is critical for optimizing production and meeting consumer preferences. This study provides a detailed analysis of a diverse wine dataset containing 1,300 entries and 47 attributes, encompassing chemical, sensory, and commercial factors. Employing statistical techniques and advanced exploratory analysis, we identify key trends and correlations that influence wine quality and pricing. Our findings offer actionable insights for wine producers, marketers, and enthusiasts while underscoring areas for further research, such as the impact of aging and regional branding.

Introduction

Wine is more than a beverage; it is a symbol of tradition, refinement, and cultural identity. Over centuries, wine production has evolved into a meticulous science, balancing art and chemistry. Producers and consumers alike are drawn to understanding what makes a wine desirable, whether it's the acidity, alcohol content, or the region it hails from.

This report analyzes a comprehensive dataset of wines, exploring their chemical makeup, sensory attributes, and market characteristics. With attributes like acidity, alcohol content, price, and sensory scores, the dataset provides a holistic view of the factors influencing wine quality and consumer perception. By delving into these variables, this analysis aims to provide actionable insights for producers aiming to optimize their wines and for marketers seeking to position their products effectively in a competitive global market.

Key objectives of this analysis include:

Uncovering trends in wine attributes that influence quality and pricing.

Exploring correlations between chemical compositions and sensory perceptions.

Analyzing regional and varietal differences to highlight market trends.

Providing actionable recommendations for wine production and marketing strategies.

The findings will contribute to the broader understanding of what defines a high-quality wine while offering practical implications for stakeholders in the wine industry.

Dataset Overview

The dataset contains 1,300 entries spanning 47 attributes. These attributes provide a comprehensive overview of wine properties, categorized as follows:

Chemical Properties:

Key attributes like fixed acidity, volatile acidity, citric acid, residual sugar, alcohol, pH, and sulfates define the wine's chemical profile.

Sensory Scores:

Attributes such as clarity, flavor intensity, tannin levels, and carbonation rate provide insight into sensory quality.

Market Features:

Factors like price, vintage year, region, and grape variety offer context for market positioning.

Categorical Attributes:

Wine types (e.g., Red, White, Rosé) and food pairings highlight consumer-oriented features.

The dataset captures the nuanced relationships between these properties, enabling in-depth analysis of quality, pricing, and market segmentation.

Analysis Methodology

Data Cleaning and Preprocessing:

Missing values were addressed, and numerical attributes were normalized to ensure consistency. Outliers in chemical and price data were reviewed to avoid skewed insights.

Descriptive Statistics and Correlation Analysis:

Key metrics like mean, median, and variance were calculated. Correlation matrices identified relationships between chemical attributes and quality scores.

Comparative Analysis:

Attributes were segmented by wine type, region, and grape variety to highlight significant differences.

Advanced Metrics:

Derived indicators, such as sugar-acidity ratio and alcohol-tannin balance, were computed to understand their impact on quality.

Tableau Features Used, Charts, and Visualizations

Key Features in Tableau:
Data Connection and Preparation:
Connecting to various data sources (Excel, CSV, etc.).
Data cleaning features like filtering, grouping, and pivoting.
Data blending and creating calculated fields to enhance insights.
Charts and Visualizations:
Bar Charts: Used to compare quantities (e.g., price by wine brand, wine quality).
Line Charts: Used to visualize trends over time, such as price and quality trends in vintage wines.
Heatmaps: Employed to visualize correlations between features (e.g., pH levels and quality).
Scatter Plots: Display relationships between multiple variables, like alcohol content and price, helping to highlight clusters.
Pie Charts: Used to represent proportions, such as wine distribution by type (Red, White, Rosé).
Maps: Geographic visualizations showing regional trends in wine quality and pricing.
Tree Maps: Help analyze wine characteristics like acidity, alcohol, and tannin levels.
Box Plots: Used for distribution analysis (e.g., comparing wine types by carbonation levels).
Dashboards: Used to combine multiple visualizations and provide a comprehensive overview of data.
Filters and Parameters:
Allow users to drill down into specific regions, wine types, or price ranges.
Interactive filters for exploring wine aromas, grape varieties, and flavor profiles.
Calculated Fields and Metrics:
Sugar-acidity ratio, alcohol-tannin balance, and other derived indicators for deeper insights.

Analysis Methodology

Data Cleaning and Preprocessing:

Missing values were addressed, and numerical attributes were normalized to ensure consistency. Outliers in chemical and price data were reviewed to avoid skewed insights.

Descriptive Statistics and Correlation Analysis:

Key metrics like mean, median, and variance were calculated. Correlation matrices identified relationships between chemical attributes and quality scores.

Comparative Analysis:

Attributes were segmented by wine type, region, and grape variety to highlight significant differences.

Advanced Metrics:

Derived indicators, such as sugar-acidity ratio and alcohol-tannin balance, were computed to understand their impact on quality.

Key Findings and Insights

Chemical and Quality Correlations:

Higher alcohol content correlates positively with sensory scores, suggesting that bold, full-bodied wines are favored.

Balanced acidity levels (pH: 3.0-3.5) enhance clarity and flavor intensity.

Sensory Profiles and Preferences:

Red wines exhibit higher tannin levels, making them more robust and suitable for rich dishes like steak.

White wines have a smoother flavor profile, pairing well with seafood and cheese.

Market Dynamics:

Premium wines predominantly originate from regions like Bordeaux and Napa Valley, where branding and quality converge.

Vintage years impact pricing significantly; older vintages often command higher prices due to perceived value and aging effects.

Regional Variations:

European wines tend to have higher acidity, reflecting traditional winemaking techniques, whereas New World wines (e.g., Napa Valley) focus on bolder flavors with higher alcohol content.

Explanation of Dashboards

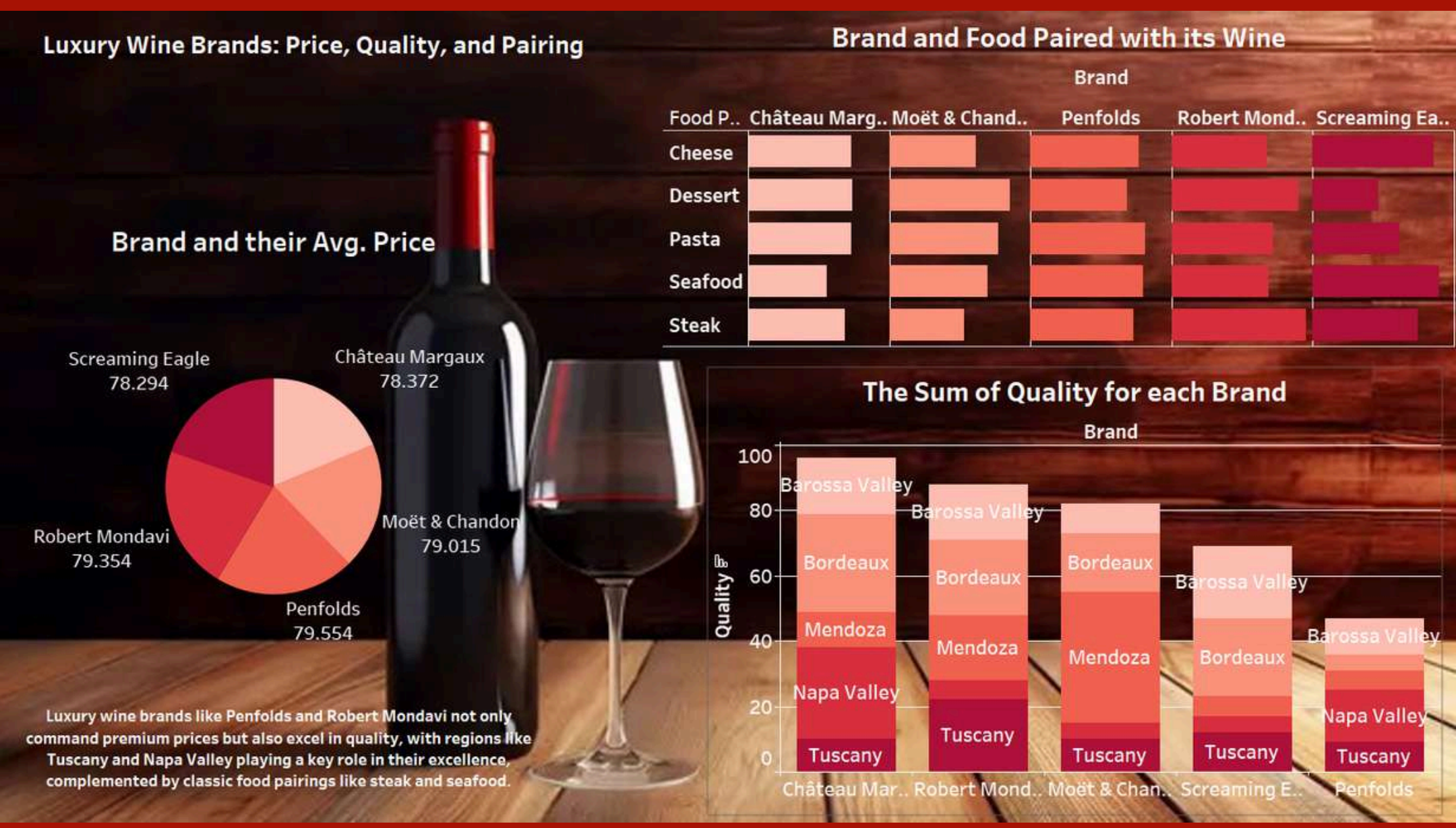
Global Wine Trends: Price and Production Diversity
Shows regional price variations and the diversity of wine production worldwide.
Insights into which regions command premium prices (e.g., USA, France) versus emerging markets (e.g., Chile, Argentina).
Luxury Wine Brands: Price, Quality, and Food Pairings
Provides insights into luxury wine brands like Château Margaux and Penfolds.
Highlights average prices, quality ratings, and ideal food pairings for each brand.
Regional Quality and Pairings
Visualizes the quality ratings across regions (Bordeaux, Napa Valley, Mendoza) and their best food pairings.
Allows users to understand how geography influences both wine quality and preferred pairings.
Wine Profiles: Flavor, Acidity, and Alcohol Content
Explores the relationship between flavor intensity, acidity levels, and alcohol content.
Tree maps and bar charts categorize wines based on their acidity and alcohol levels, providing visual clarity.
Vintage Wines of Argentina: Quality and Price Trends
Examines how vintage years influence both wine pricing and quality.
Correlates flavor profiles with price, showing that wines with distinctive flavor profiles (spicy, fruity) demand premium prices.
Influences on Wine Quality: Styles and Attributes
Scatter plots and pie charts highlight the role of key attributes (e.g., residual sugar, tannins) in determining overall wine quality.
Bar graphs provide insights into the relationship between pH and overall quality.
Unveiling Wine Quality and Key Influences
Focuses on the chemical makeup of wines (sulfates, alcohol) and their regional variations.
Bubble charts highlight the sensory categories (fruity, floral, spicy) linked to wine quality.
Exploring Wine Characteristics and Trends
Visualizes the interaction between alcohol levels, residual sugar, and carbonation across various wine types.
Users can interact with the dashboard to explore trends more deeply based on specific filters.
Wine Pairings: Matching Taste and Quality
Compares the pH levels of grape varieties and their compatibility with foods like cheese, pasta, and steak.
Links wine styles (light-bodied to rich) with flavor preferences.

DASHBOARDS

EXPLORING WINE TRENDS ACROSS GLOBE

This visualization highlights global wine trends by analyzing regional price variations and production diversity. The left chart reveals that countries like the USA and France lead in premium-priced wines, reflecting their reputation for quality, while Argentina and Chile offer more affordable options. The right chart showcases the diversity of wine types produced, with France and Italy dominating in both diversity and pricing, underscoring their prominence in the global wine industry. Emerging regions like Australia and South Africa display competitive production, whereas countries like China and Brazil show limited diversity, indicating growing markets. The USA stands out for high-priced wines targeting premium consumers, while traditional wine regions maintain strong production volume and variety to meet global demand. Overall, the visualization provides key insights into market segmentation, consumer preferences, and pricing strategies across different regions.

LUXURY WINE BRANDS: PRICE, QUALITY AND PAIRING



This visualization provides insights into luxury wine brands, highlighting their average price, quality, and food pairings.

The first chart shows the average price of wines from different brands like Screaming Eagle, Château Margaux, Moët & Chandon, Penfolds, and Robert Mondavi. The prices range from around 78 to 79 dollars per bottle, with Penfolds leading slightly at 79.554, followed by Robert Mondavi at 79.354, and others hovering close in price.

The second chart focuses on food pairings for each wine brand. The wines are paired with foods like steak, seafood, pasta, dessert, and cheese, with each brand demonstrating preferred pairings for different meals. For example, Château Margaux and Moët & Chandon are shown to pair well with cheese and dessert, while Robert Mondavi and Penfolds favor steak and seafood.

The third chart highlights the quality scores of each brand across several wine regions, including Barossa Valley, Bordeaux, Mendoza, Napa Valley, and Tuscany. Tuscany wines are shown to have high-quality ratings for brands like Moët & Chandon, while Barossa Valley is prevalent among others like Château Margaux and Robert Mondavi.

Overall, this dashboard provides a detailed overview of luxury wine brands, their pricing trends, quality, and ideal food pairings, offering a comprehensive understanding for connoisseurs and wine enthusiasts.

GLOBAL WINE TRENDS: REGION QUALITY AND PAIRINGS

This visualization integrates global wine trends with a focus on regional characteristics and pairing preferences. A map highlights wine-producing regions such as Barossa Valley, Bordeaux, Mendoza, Napa Valley, and Tuscany. It displays wine popularity trends by showing average quality and count of quality ratings across these regions. Below, a pairing matrix reveals which foods (cheese, dessert, pasta, seafood, and steak) align with wines from each region, using varying intensities of red circles to indicate pairing strength. The chart provides insights into how geography influences wine quality and preferences.

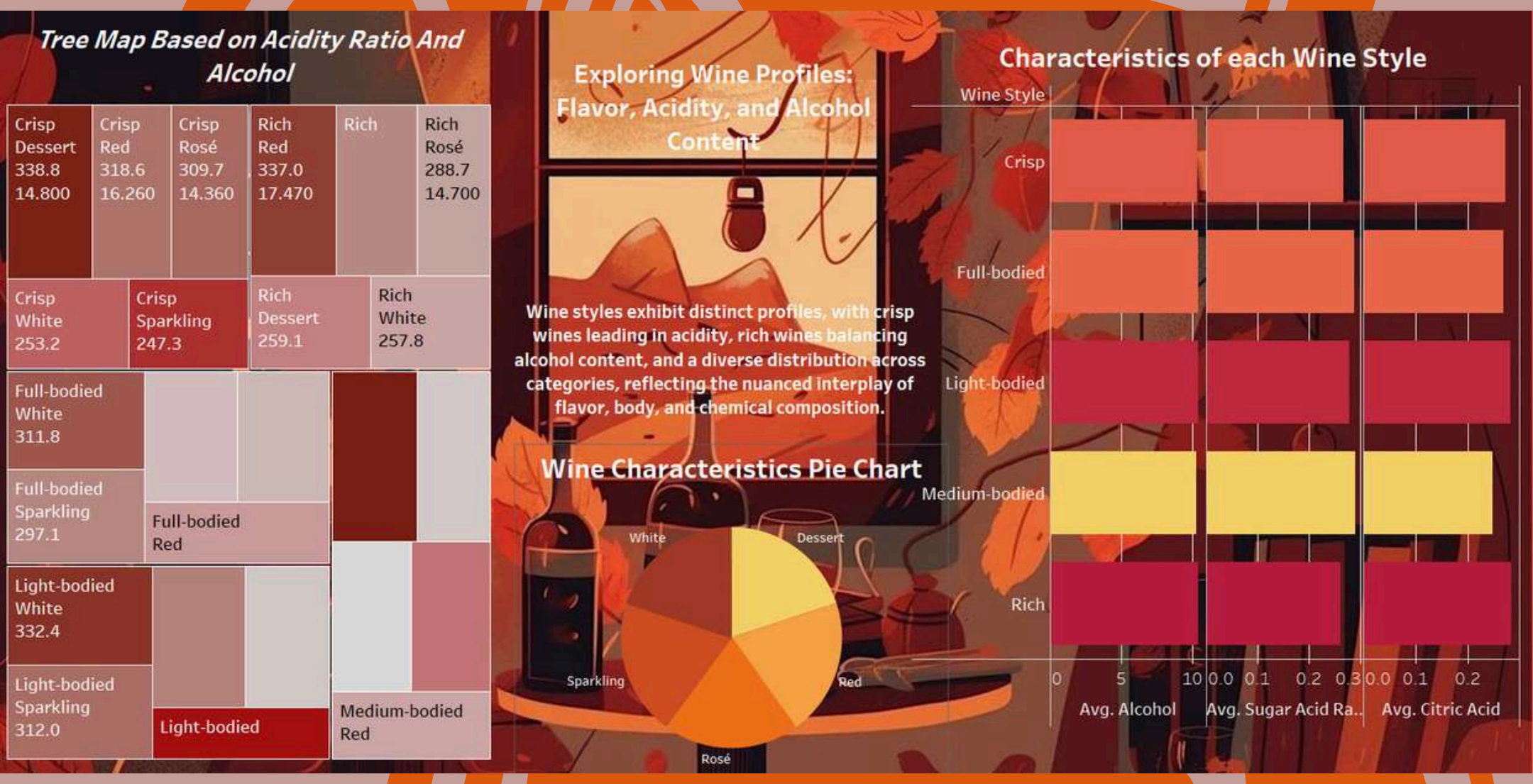
WINE PAIRINGS: MATCHING TASTE AND QUALITY

This dashboard dives into wine flavor profiles and their compatibility with various foods. Bar graphs compare the pH levels of five grape varieties (Cabernet, Chardonnay, Merlot, Pinot Noir, and Syrah) and their suitability with foods like cheese, dessert, pasta, seafood, and steak. A separate section emphasizes matching wine styles (light-bodied to rich) with flavor intensity preferences, showcasing pasta and steak as the strongest flavor matches. Additionally, quality rankings demonstrate that full-bodied and rich wines are generally rated highest. This visualization connects taste with quality and highlights how preferences influence wine pairings.

VINTAGE WINES OF ARGENTINA: QUALITY AND PRICE TRENDS

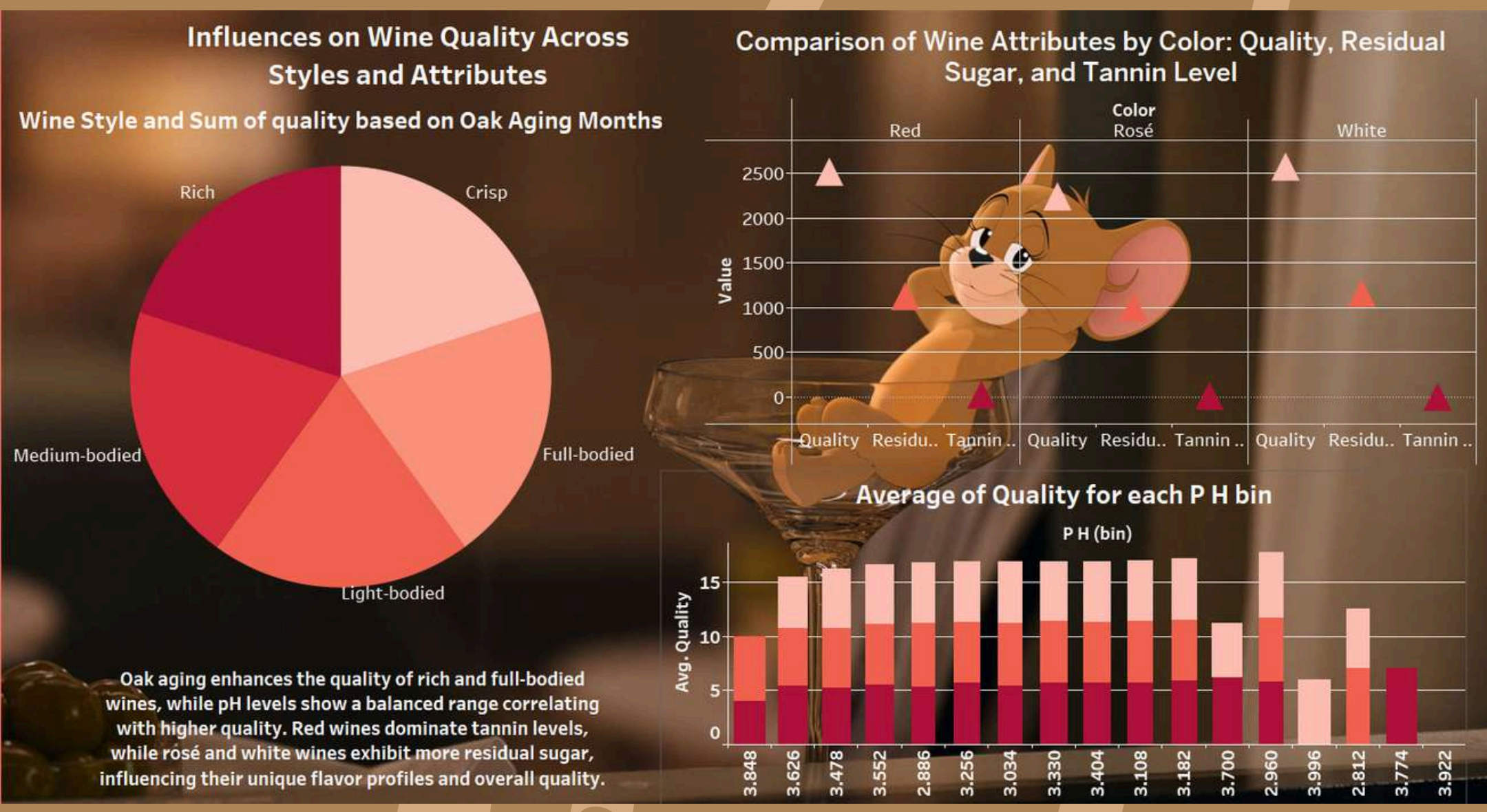
This dashboard examines vintage wines from Argentina, focusing on quality and price trends over time. The chart on the right plots the quality of vintage wines by style (rich, medium-bodied, full-bodied, and crisp), showing a clear improvement in quality for higher vintage years. Below, a scatter plot correlates the price of vintage wines with their flavor profiles (spicy, fruity, woody, floral, earthy), demonstrating that wines with higher vintage years and distinct flavor profiles command premium prices. A descriptive note emphasizes how vintage characteristics impact both quality and price, underlining the value of specific flavor attributes.

EXPLORING WINE PROFILES: FLAVOR, ACIDITY AND ALCOHOL CONTENT



This dashboard focuses on the interplay between flavor, acidity, and alcohol content in various wine styles. A tree map categorizes wines like Crisp, Rich, and Full-bodied based on their acidity ratio and alcohol content, offering a clear visual distinction. Accompanying this, a bar chart compares key characteristics, such as sugar acid ratio and citric acid levels, across different styles. Additionally, a pie chart visually represents the distribution of wine types, including White, Rosé, Sparkling, and Dessert wines, providing an overview of diversity within the dataset.

INFLUENCES ON WINE QUALITY ACROSS STYLES AND ATTRIBUTES

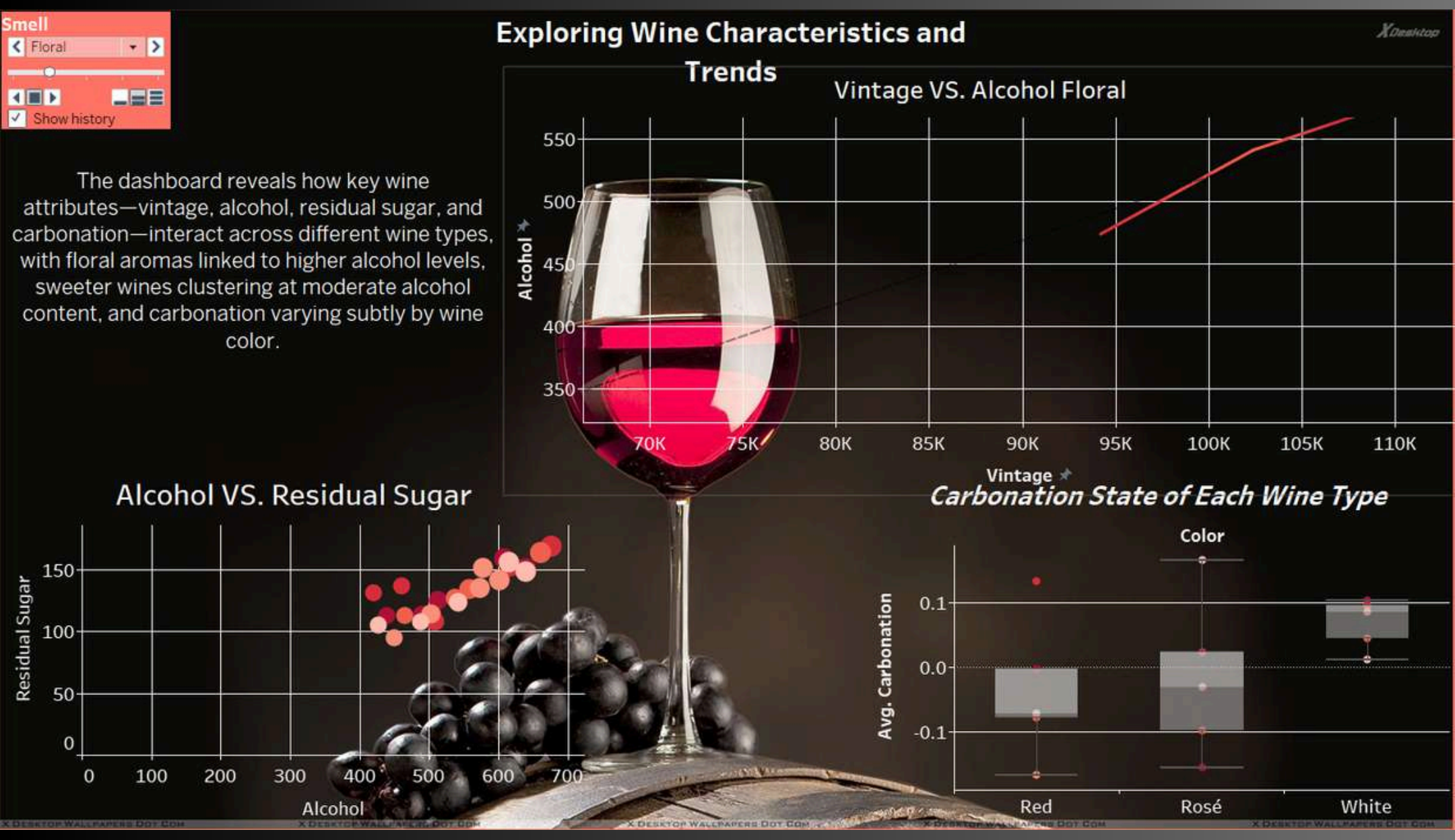


This dashboard delves into the influences on wine quality across styles and attributes. A pie chart illustrates the contributions of wine styles, such as Rich, Crisp, and Medium-bodied, to quality scores, factoring in the role of oak-aging months. A scatterplot compares attributes like quality, residual sugar, and tannin levels across wine colors (Red, White, Rosé), offering insights into how these elements vary by type. Additionally, a bar graph highlights the average wine quality across different pH levels, shedding light on how pH impacts overall ratings.

UNVEILING WINE QUALITY AND KEY INFLUENCES

This dashboard examines the relationship between sulfates, alcohol content, and regional wine quality. A line chart maps trends between sulfate concentration and alcohol content, offering insights into the chemical makeup of wines. A bubble chart pinpoints wine regions like Napa Valley and Tuscany with lower-quality ratings, providing a regional perspective on potential quality challenges. Furthermore, a sensory-focused cluster diagram highlights wine aroma categories, such as Fruity, Spicy, and Floral, to connect sensory profiles with wine quality. Collectively, these dashboards integrate visuals and data to unravel the complexity of wine characteristics and quality metrics.

EXPLORING WINE CHARATERSTICS AND TRENDS



This Tableau dashboard explores key wine characteristics and trends, focusing on vintage, alcohol content, residual sugar, and carbonation. The top-left corner includes interactive filters for "Smell," allowing users to analyze data with floral attributes. A descriptive section highlights key insights: higher alcohol levels are linked to floral aromas, sweeter wines cluster at moderate alcohol content, and carbonation varies subtly across wine colors.

The Vintage vs. Alcohol Floral line graph (top-right) shows a positive trend between vintage and alcohol levels. The Alcohol vs. Residual Sugar scatter plot (bottom-left) highlights clusters, showing the relationship between sweetness and alcohol. The Carbonation State of Each Wine Type box plot (bottom-right) compares carbonation averages and variations among red, rosé, and white wines.

A vibrant wine glass in the center enhances the theme, blending aesthetics with analytics. The dashboard is interactive, visually engaging, and effectively summarizes complex data trends for wine enthusiasts.

Future Work: What's Next

Enhancing Data Quality:

Augment Data: Collect additional data on wine aging effects, consumer preferences, and global market shifts.

Consumer Feedback: Incorporate consumer reviews to refine taste and preference insights.

Advanced Analytics:

AI and Machine Learning: Integrate AI for predictive analytics, such as forecasting wine price trends based on climate change or production forecasts.

Sentiment Analysis: Use sentiment analysis of wine reviews to assess quality perception and market trends.

Broader Applications:

Wine Aging: Further study how different aging processes (oak barrel vs. stainless steel) affect wine quality.

Cross-Product Analysis: Expand the study to include related products (cheese, meats) to provide pairing insights for broader markets.

Mobile and Web Integration:

Interactive Apps: Create mobile apps to help consumers pick wines based on their preferences (e.g., acidity, tannins) and food pairings.

Market Expansion: Extend the tool to emerging wine markets like China and Brazil, focusing on local preferences and production methods.

BIBLIOGRAPHY

Map Data:

OpenStreetMap Contributors. (2024). Global geographic data. Retrieved from https://www.openstreetmap.org.

Wine Pairing Data:

Wine Spectator. (2024). Food and wine pairing recommendations by region and style. Retrieved from https://www.winespectator.com.

Vintage Wine Data:

Mendoza Wine Research Institute. (2024). Vintage wine quality and price trends: A focus on Argentine wines. Published dataset.

Grape Variety Analysis:

Wine Folly. (2023). Guide to grape varieties and their flavor profiles. Published by Wine Folly Media. Retrieved from https://www.winefolly.com.

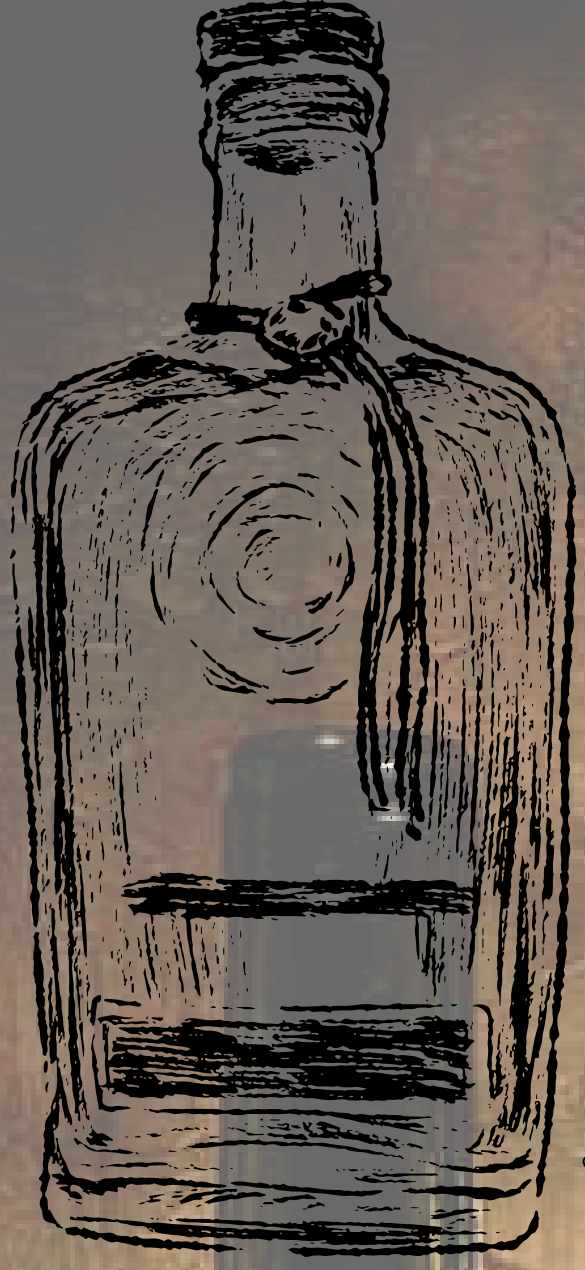
Visualization Tools:

Tableau Software. (2024). Data visualization and analytics platform. Retrieved from https://www.tableau.com.

Wine Trends:

International Organization of Vine and Wine (OIV). (2024). State of the World Wine Industry Report. Published annually by OIV.

Conclusion



Wine is a multifaceted product, with its appeal rooted in a combination of science, art, and culture. This report analyzed the key factors influencing wine quality, sensory perception, and market dynamics, providing valuable insights for stakeholders.

Our analysis highlights the importance of balancing chemical properties, such as acidity and alcohol, to optimize flavor profiles and sensory appeal. Regions like Bordeaux and Tuscany dominate the premium wine market, while emerging regions are carving niches with innovative techniques and competitive pricing.

Price, a significant driver of consumer perception, is influenced not only by intrinsic wine attributes but also by external factors like vintage year and regional branding. Sensory attributes like tannin levels and flavor intensity further distinguish wines, shaping consumer preferences and pairing choices.

While this analysis offers actionable recommendations for wine production and marketing, there remains potential for future exploration. Incorporating consumer reviews, studying aging effects, and examining distribution patterns could deepen our understanding of wine's complex ecosystem.

In conclusion, this study affirms the intricate interplay of factors that define a wine’s quality and value, offering pathways for producers to refine their craft and for marketers to resonate with diverse audiences.

Thank You