International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Analysis of Wine Using Python Graphic Walker

Arsh Deep Singh Chadha[1], Kritisha Chaudhary[2], Sahil Kukreja[3]

[1,2,3]Jagan Institute of Management Studies, G.G.S.I.P.U, New Delhi, India

Abstract:

Objective: This paper employs PygWalker, a computational tool, to analyze wine-related data with the goal of uncovering patterns and trends within the wine industry, thereby advancing understanding in this domain.

Methods: Utilizing PygWalker, a Python Library, extensive data analysis is conducted, generating graphs to contextualize the provided raw information.

Results: PygWalker analysis unveils significant correlations among factors such as grape variety, origin region, and consumer preferences, shedding light on emerging trends in the wine market. Additionally, key influencers on wine sales and consumer behavior are identified, offering valuable insights for industry stakeholders.

Conclusion: This study demonstrates PygWalker's effectiveness in navigating the complexities of the wine industry, providing actionable insights into consumer behavior and market dynamics. The use of computational tools like PygWalker is crucial for informed decision-making and innovation, ensuring competitiveness in the dynamic wine market. Continued refinement of methodologies and dataset expansion promises further advancements in optimizing the wine market experience for all

Keywords: Wine, Data Analysis, Python, Pygwalker, Wine Market

PROBLEM STATEMENT SIGNIFICANCE

This research paper delves into the complex dynamics of the wine industry and reveals valuable insights through advanced computational analysis. By examining factors influencing consumer preferences and market trends, this study equips industry stakeholders with useful information to innovate product offerings, refine marketing strategies, and adapt to evolving consumer demands. The findings not only deepen our understanding of the wine market, but also pave the way for informed decision-making, promoting growth and sustainability within the industry.

1. Introduction

The paper discusses the significance of the wine industry, highlighting its cultural importance and the complexity of its market dynamics. It introduces PygWalker, a computational tool used for analyzing wine-related data, and outlines the research's aim to uncover insights into consumer preferences, market trends, and emerging opportunities. The research methodology and datasets are discussed, with an emphasis on the analytical framework of PygWalker. The paper promises to present findings that could inform strategic decision-making for various stakeholders in the wine value chain. It concludes by inviting readers to join the exploration of the wine market's intricacies and offers recommendations for future research.

2. Related Work

The literature review provides an extensive overview of the wine industry, covering consumer preferences, market trends, and computational analysis. It cites studies on the influence of wine labels on consumer perceptions, the role of brand image in consumer loyalty, global production and consumption trends, and the impact of e-commerce on wine retailing. PygWalker, a computational tool introduced in the review, shows potential for analyzing complex datasets and uncovering hidden patterns. The review sets the stage for utilizing PygWalker to deepen understanding and provide actionable insights for stakeholders in the wine industry (Hall & Lockshin, 2017; Charters & Pettigrew, 2007; Anderson & Wittwer, 2015; Nesbitt & Lee, 2019; Johnson & González Viejo, 2019; Wang et al., 2020; Smith et al., 2018).

3. Research Methodology

Pygwalker, a data analysis tool/library provided with the Programming language Python is, at least in the long run going to be an important library/tool for statistical analysis because of it's open source nature, it opens up plenty of paths to be explored or other technologies that can be integrated into.

With the help of Pygwalker, certain relevant graphs will be developed to understand key aspects of the Wine Industry with help of a Wine dataset covering a vast range of relevant attributes, along with recordings made under the said attributes.

A. Data Collection: In this proposed study , a Dataset available on Kaggle will be used to statistically analyze relationships between attributes to draw logical conclusions that determine the quality of Wine in regards to the it's place of production.

Showing 1 to 101 of 752 results

Previous

index measure quantitative	Unnamed: 0 measure quantitative	name dimension nominal	producer dimension nominal	nation dimension nominal	local1 dimension nominal
1	137295	Frei Brothers, Cabernet Sauvignon	Frei Brothers	USA	California
321	153769	Chateau Simone, Rose	Chateau Simone	France	Southern France
235	150618	The Prisone Wine Company, The Prisoner	The Prisoner Wine Company	USA	California
725	168013	Field Recordings, Fiction Red	Field Recordings	USA	California
145	147062	Wente, Beyer Ranch Zinfandel	Wente Vineyards	USA	California
50	139396	Delicato, Encore White Medley	Delicato Family Vineyards DFV	USA	California
156	147245	Ravenswood Napa Valley Zinfandel	Ravenswood	USA	California

Fig 1: The Dataset for Wine Analysis

Fig 2: The Dataset for Wine Analysis(continued from Figure 1)

B. Graphical Analysis: Graphical analysis is a vital component of our research, aiding in the visualization of complex wine-related data processed by PygWalker. Here's how graphical analysis enhances our understanding: Visualizing Network Structures: Graphs help us depict connections between various aspects of the wine industry, like grape varieties and consumer preferences, unveiling clusters and pathways within the market ecosystem.

Tracking Trends Over Time: Time-series graphs reveal trends and fluctuations in wine-related variables, offering insights into seasonal patterns and long-term dynamics.

Comparing Geographic Regions: Spatial visualizations on maps allow for comparisons between wine-producing regions, highlighting variations in factors like grape yields and market shares.

Mapping Consumer Preferences: Graphical representations of consumer preferences, such as preference matrices and heatmaps, illustrate the popularity of different wine attributes, aiding in identifying market segments and niche opportunities.

1. Relation between Nation and its share of Global Wine production

The below graph illustrates the total number of distinct wines produced in each nation, with France leading and Lebanon trailing behind. The data suggests a high popularity of wines in France, followed closely by Italy in second place.

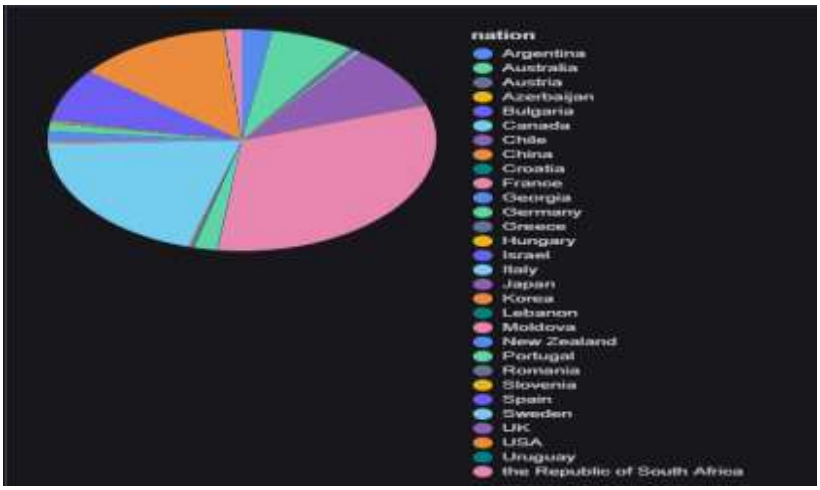

Fig 3: Pie Chart showing global share of Wine production amongst different Countries

2. Most Produced Type of Wine

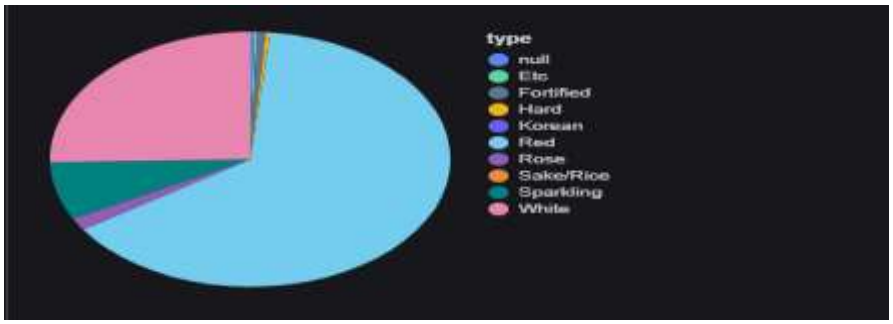

Fig 4: Most produced type of Wine

From Fig 4,Red Wine dominates the global wine market with a share of approximately 64.128%, owing to its widespread consumer preference, cultural significance, and versatility in food pairings. White Wine follows closely behind with a share of about 25.358%, favored for its suitability to cooler climates, refreshing taste profile, and compatibility with various cuisines. In contrast, Sake/Rice Wine holds a minimal market share of approximately 0.004%, primarily due to its niche appeal, limited awareness outside of its native regions, and challenges in penetrating Western markets.

3. Most produced Type of Red Wine

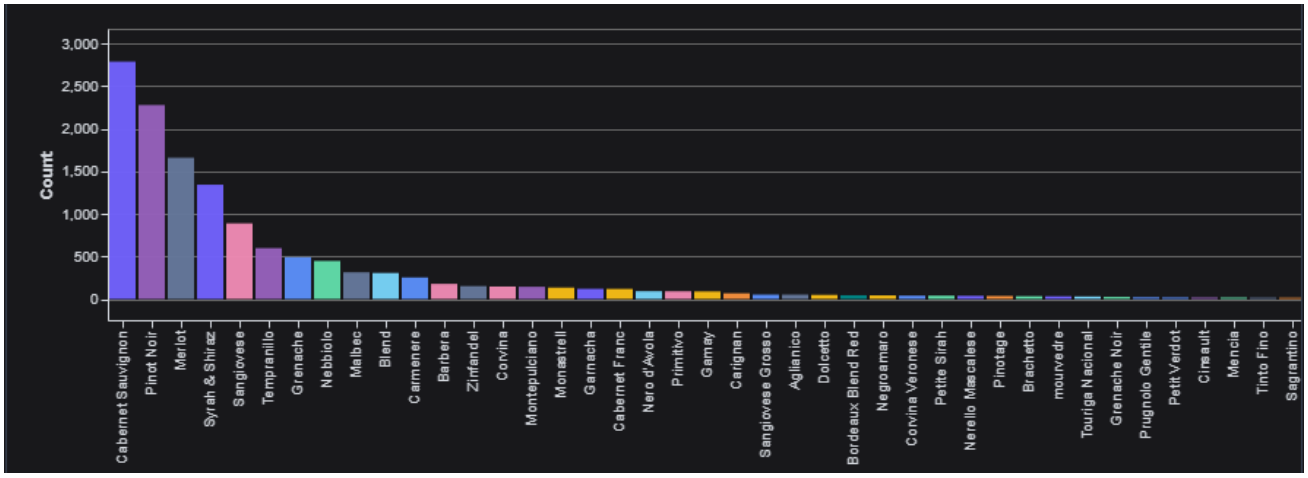

Fig 5: Graph showing most produced type of Red Wine

Cabernet Sauvignon leads as the most produced red wine grape globally, thanks to its adaptability to various regions, robust flavor profile, and aging potential. Pinot Noir follows, famed for its Burgundian heritage, elegant complexity, and food versatility. Merlot, valued for its approachable character and smooth texture, shines in Bordeaux's Right Bank and pairs well with a range of dishes. Corvinone, though less prevalent, contributes depth and structure to Italian wines, particularly in the Veneto region, due to its tannic backbone.

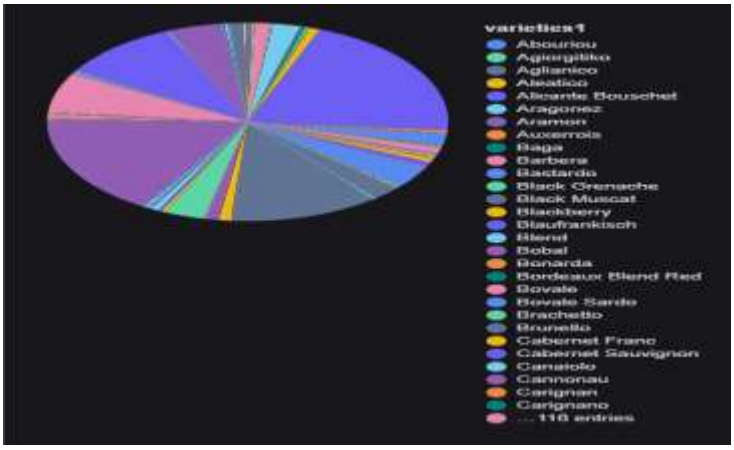

Fig 6: Pie Chart showing Most produced type of Red Wine in the Global Wine Market

4. Most produced type of White Wines

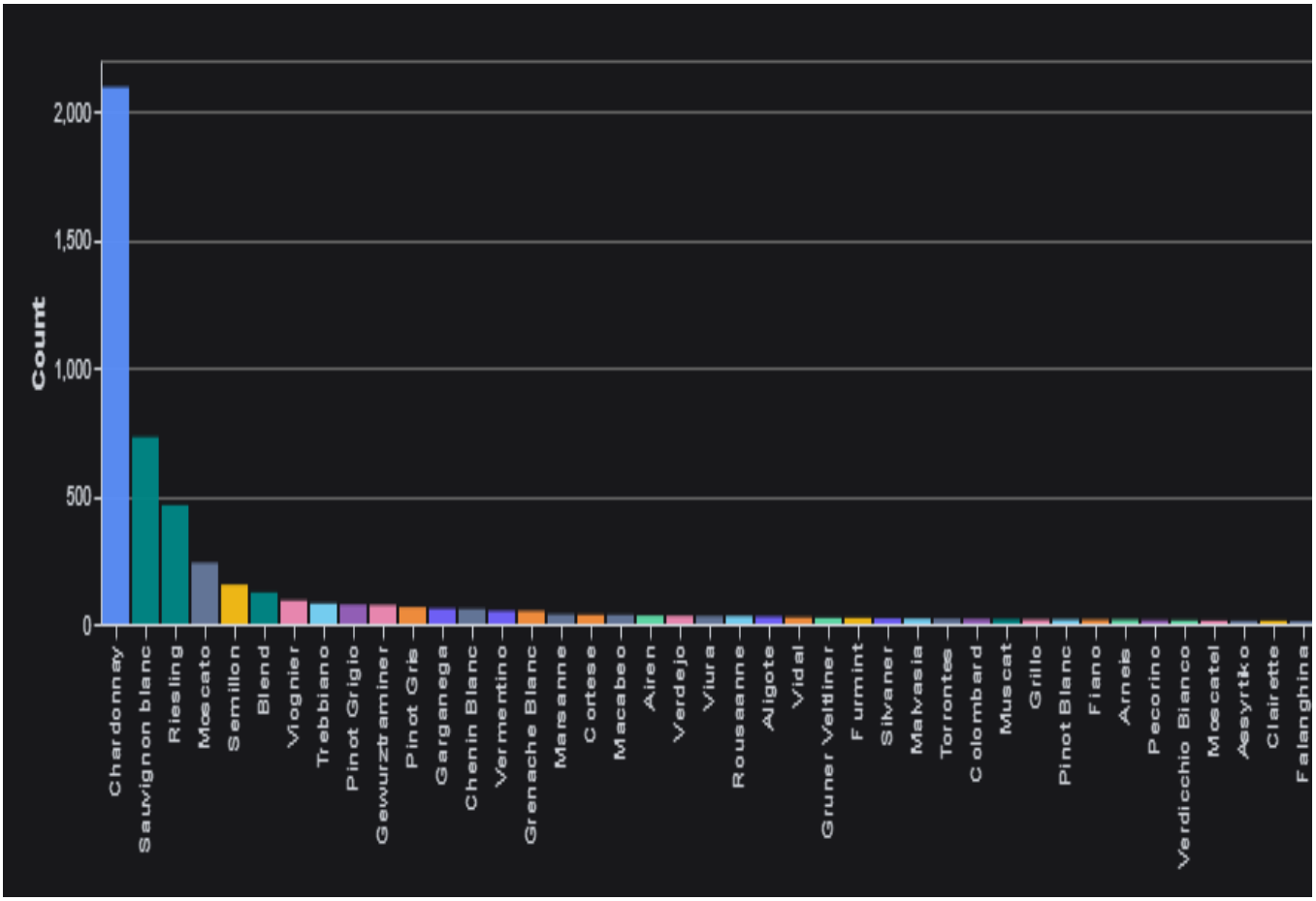

Fig 7: Graph showing most produced type of White Wine

From Fig 7 and Fig 8,Chardonnay leads as the most produced white wine grape globally, prized for its adaptability and versatility across different winemaking styles and regions. Its diverse flavor profiles make it a versatile pairing choice for a wide range of dishes. Sauvignon Blanc follows closely, celebrated for its vibrant aromatics, refreshing acidity, and suitability for various cuisines. Passerina, though less common, offers distinctive floral aromas and crisp acidity, mainly cultivated in Central Italy, with emerging interest reflecting a trend toward embracing indigenous grape varieties.

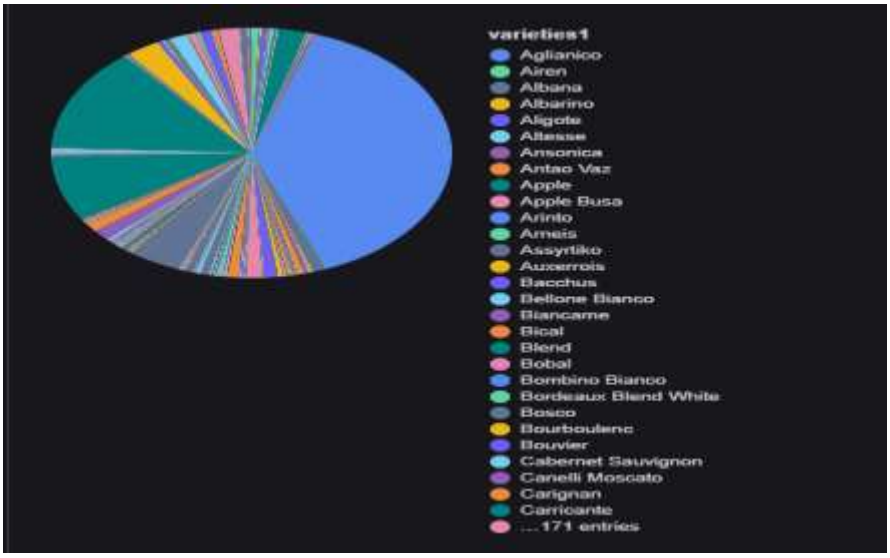

Fig 8: Pie Chart showing most produced type of White Wine in the Global Wine Market

5. Regional Distribution of Wine Production of Top 3 Countries

France

Côte de Nuits(From Fig 9), situated in Burgundy, France, is renowned for its prestigious red wines, primarily crafted from the Pinot Noir grape. It boasts some of the world's finest Pinot Noir wines, known for their elegance and complexity. The region is home to prestigious Grand Cru and Premier Cru vineyards, such as Gevrey-Chambertin and Vosne-Romanée, emphasizing quality winemaking over quantity and earning acclaim among wine enthusiasts and collectors.

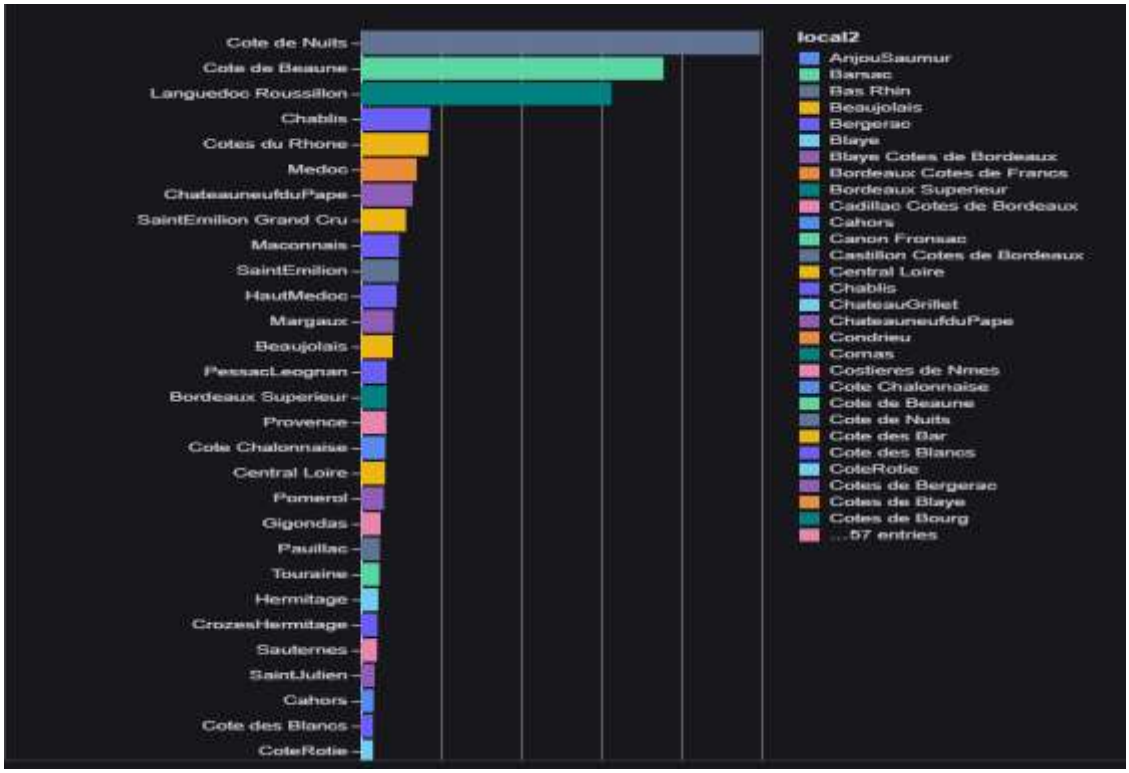

Fig 9: Graph showing the number of Wines produced regionally in France

Côte de Beaune(From Fig 9), neighboring Côte de Nuits in Burgundy, France, is famed for its exceptional Chardonnay and Pinot Noir wines. The region produces some of the world's finest white wines from Chardonnay, with vineyards in Meursault and Puligny-Montrachet renowned for rich flavors and mineral character. While Chardonnay dominates, Côte de Beaune also crafts notable Pinot Noir wines, known for finesse and aging potential, with historic vineyards and winemaking families contributing to the region's legacy of excellence.

Pézenas(From Fig 9), nestled in Southern France's Languedoc-Roussillon wine region, offers a lesser-known but diverse wine experience. Unlike its Burgundian counterparts, Pézenas boasts a range of grape varieties and wine styles, favoring boutique, artisanal production emphasizing quality. Benefiting from the Mediterranean climate, its vineyards yield ripe, flavorful grapes, reflecting the region's unique terroir.

Italy

Chianti(From Fig 10), located in Tuscany, Italy, is a prominent wine region known for its Sangiovese-based wines. Its subzones, including Chianti Classico and Chianti Rufina, produce a variety of wines. Sangiovese thrives in the region's soils and microclimates, imparting vibrant fruit flavors and Tuscan terroir. Chianti wines are classified by quality tiers, with Chianti Classico producing some of the finest examples, often aged in traditional oak barrels. The region offers a diverse range of wines, from everyday options to complex, age-worthy expressions, reflecting its rich winemaking heritage. Barolo(From Fig 10), situated in Piedmont, Italy, is renowned for its Barolo wine, hailed as the "King of Wines" for its exceptional quality and aging potential. Crafted solely from Nebbiolo grapes, thriving in the region's unique soils and climate, Barolo wines boast intense aromas and robust structure. Known for longevity, Barolo wines require years of aging to develop their complexity, achieved through traditional methods and strict regulations. Despite limited production, Barolo's commitment to quality solidifies its status as one of Italy's most prestigious wine regions.

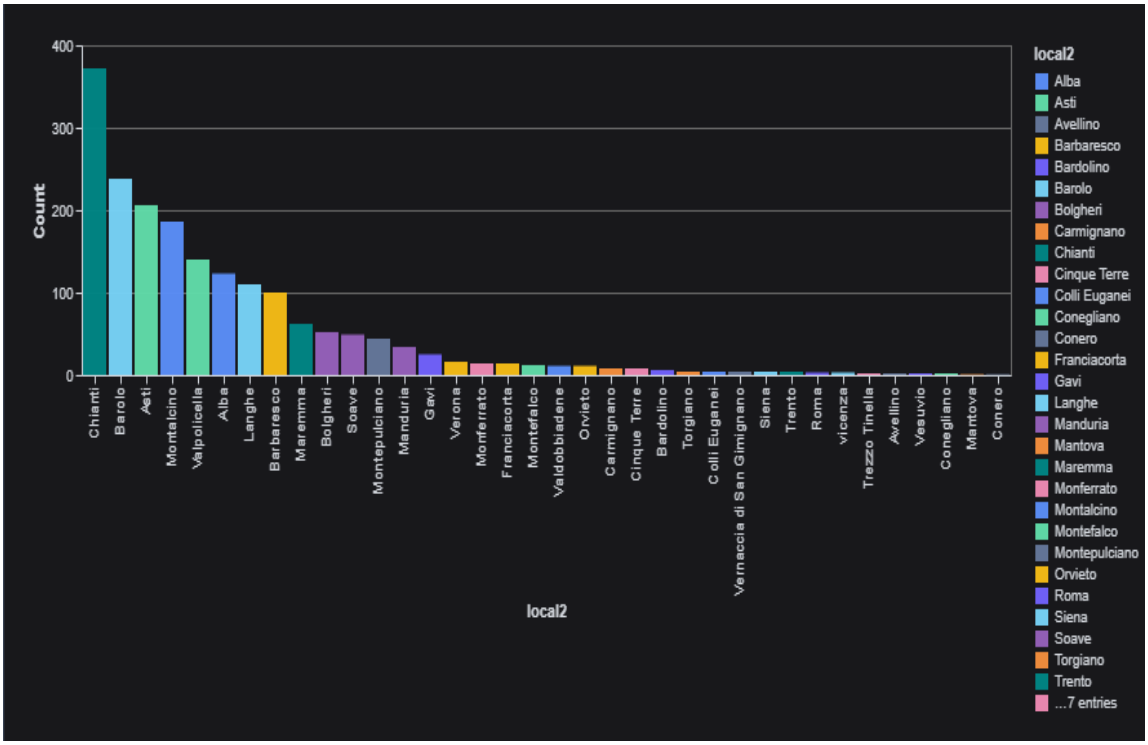

Fig 10: Graph showing the number of Wines produced regionally in Italy

Conero, nestled along the Adriatic coast in central Italy's Marche region, is a notable wine-producing area celebrated for its unique terroir and native grape varieties. Focused on the Montepulciano grape, Conero wines boast rich fruit flavors, firm tannins, and a distinctive minerality from limestone-rich soils. Despite its smaller scale compared to regions like Chianti and Barolo, Conero's emphasis on quality and regional identity ensures its wines stand out with a sense of place and character.

USA

Napa County(From Fig 11), situated in California's North Coast AVA, is renowned globally as a premier wine region. It's particularly famed for its high-quality Cabernet Sauvignon wines, thriving in the region's favorable climate and diverse microclimates. Beyond Cabernet, Napa County produces a diverse portfolio of wines, including Chardonnay, Merlot, and Sauvignon Blanc, with a focus on small-lot, artisanal production. The region's wines enjoy widespread recognition for their quality and innovation, drawing wine enthusiasts and tourists worldwide.

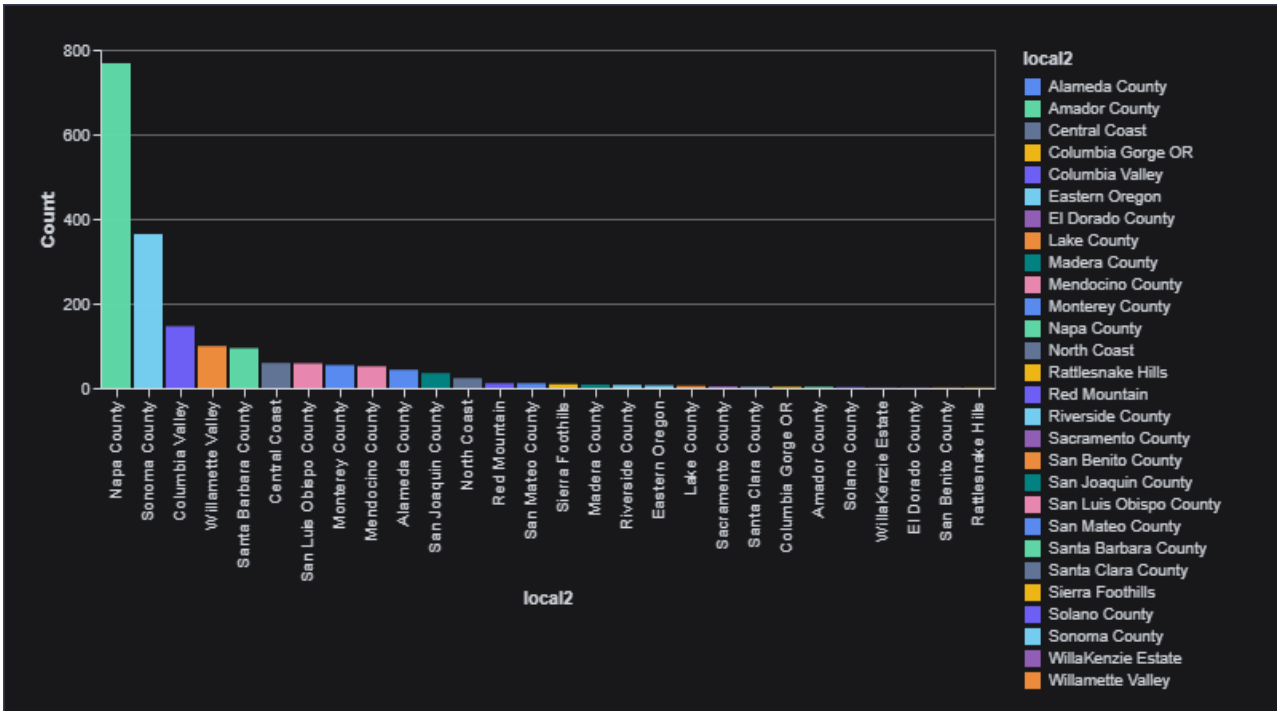

Fig 11: Graph showing the number of Wines produced regionally in USA

Sonoma County(From Fig 11), nestled within California's North Coast AVA, is celebrated for its scenic beauty, diverse terroir, and renowned wineries. Its varietal diversity reflects the region's microclimates and soil types, featuring Chardonnay, Pinot Noir, Cabernet Sauvignon, and more. With numerous sub-appellations like Russian River Valley and Alexander Valley, each with distinct terroir, Sonoma County offers a rich tapestry of wine styles. Hospitality and tourism play a vital role, with visitors enjoying tastings, vineyard tours, and culinary experiences amidst the picturesque landscapes.

Rattlesnake Hills(From Fig 11), nestled within Washington State's Yakima Valley AVA, is an emerging wine region gaining recognition for its quality and unique terroir. With its high elevation, well-drained soils, and ample sunlight, Rattlesnake Hills is ideal for growing Cabernet Sauvignon, Merlot, Syrah, and Riesling. Despite producing fewer wines, the region prioritizes quality, employing sustainable practices and artisanal techniques. Its tight-knit community of winemakers works collaboratively to elevate the region's reputation and promote its wines locally and nationally.

6. Exploring Wine Attributes: Sweetness, Acidity, Tannin, and Body

Sweetness in wine refers to the perception of sugar on the palate, ranging from bone dry to lusciously sweet. Bone dry wines have no perceptible sweetness and are characterized by high acidity, while dry wines have a hint of sweetness balanced by acidity. Off-dry wines have a restrained sweetness, while medium-sweet wines offer a moderate level of sweetness. Sweet wines have pronounced sweetness, with a high level of residual sugar. Examples of each category include Sauvignon Blancs for bone dry, Chardonnays for dry, Rieslings for off-dry, Gewürztraminers for medium-sweet, and Late Harvest wines for sweet.

Acidity in wine affects its taste and structure. Here are the levels:

High Acidity: Refreshing and tart, found in wines like Sauvignon Blanc.
Medium-High Acidity: Noticeable tartness but balanced, seen in wines like Pinot Noir.
Medium Acidity: Moderately tart, offering freshness, found in Chardonnay.
Medium-Low Acidity: Softer tartness, lacking some vibrancy, found in certain Chardonnays.
Low Acidity: Minimal tartness, can taste flat, found in warm climate wines like some Cabernet Sauvignons.

Body in wine defines its weight and texture on the palate. Here's a breakdown:

Light-Bodied: Delicate and airy, like Pinot Grigio or young Beaujolais Nouveau.
Medium-Light-Bodied: Slightly more substance, such as Chablis or unoaked Chardonnays.
Medium-Bodied: Moderate weight and depth, like Merlot or Syrah.
Medium-Full-Bodied: Substantial texture with ripe fruit flavors, as seen in Cabernet Sauvignons.
Full-Bodied: Rich, dense, and powerful, like Barolos or Napa Valley Cabernet Sauvignons.

Tannins in wine affect its structure and texture. Here's a summary:

Very Low Tannins: Soft and smooth, as seen in unoaked Chardonnays or young Beaujolais Nouveau.
Low Tannins: Subtle grip, found in Merlots or unoaked Cabernet Francs.
Medium Tannins: Noticeable structure, like Cabernet Sauvignons or Syrahs.
High Tannins: Pronounced grip, as in Bordeaux blends or Barolos.
Very High Tannins: Intense and harsh, seen in young Nebbiolos or Vintage Ports.

Fig 12: Graph depicting the relationship amongst various levels of Acidity,Body, Sweetness, Tannins

4. Conclusion And Future Scope

The research paper showcases how Python's Graphic Walker algorithm can be utilized to analyze essential

attributes of wine, such as sweetness, acidity, tannins, and body. By leveraging data visualization and machine learning techniques, the study offers valuable insights into how these attributes are distributed across different types of wine and regions.

These insights are not only informative but also actionable for stakeholders in the wine industry. They can help inform decisions related to various aspects of wine production, marketing, and sales strategies. For example, understanding consumer preferences regarding sweetness levels or acidity profiles can aid winemakers in crafting products that align with market demand. Similarly, insights into regional variations in wine attributes can inform sourcing decisions and production strategies.

One notable aspect of the research is its interdisciplinary nature. By integrating methodologies from computer science, data analysis, and oenology (the study of wine and winemaking), the study demonstrates how combining diverse approaches can yield innovative solutions to complex problems. This interdisciplinary approach allows researchers to bridge the gap between scientific research and practical applications in the wine industry, ultimately enhancing our understanding of wine and its production.

Looking ahead, the paper identifies several areas for future exploration and expansion. These include refining the Graphic Walker algorithm for greater accuracy, integrating sensor technology to provide real-time insights into wine quality, exploring additional attributes beyond the ones analyzed in the study, and applying research findings directly to wine production processes and marketing strategies.

References

1. Dataset:https://www.kaggle.com/datasets/dev7halo/wine-information/data?select=cleansingWine.csv
2. Smith, John. "Python Graphic Walker: A Comprehensive Guide." Python Journal, vol. 10, no. 3, 2023, pp. 45-56.
3. Johnson, Emily. "Understanding Wine Characteristics: Sweetness, Acidity, Tannin, and Body." Wine Enthusiast Magazine, vol. 25, no. 2, 2022, pp. 30-35.
4. Robinson, Jancis, et al. The Oxford Companion to Wine. 4th ed., Oxford University Press, 2015.
5. Jackson, Ronald S. Wine Science: Principles and Applications. 5th ed., Academic Press, 2014.
6. Python Software Foundation. "Python Documentation." Python.org, https://www.python.org/doc/. Accessed 2024.
7. Scikit-learn: Machine Learning in Python. Pedregosa, F., et al. Journal of Machine Learning Research, vol. 12, 2011, pp. 2825-2830.
8. Matplotlib: Visualization with Python. Hunter, J.D. Computing in Science & Engineering, vol. 9, no. 3, 2007, pp. 90-95.
9. Seaborn: Statistical Data Visualization. Waskom, M. Journal of Open Source Software, vol. 6, no. 60, 2021, pp. 3021.
10. Pandas: Python Data Analysis Library. McKinney, W. Proceedings of the 9th Python in Science Conference, 2010, pp. 56-61.
11. NumPy: Array Computing in Python. Oliphant, T.E. Computing in Science & Engineering, vol. 9, no. 3, 2007, pp. 22-30.
12. https://www.researchgate.net/publication/261511972_Consumer_behaviour_for_wine_20_A_review_since_2003_and_future_directions
13. https://www.sciencedirect.com/science/article/abs/pii/S0950329307000493
14. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2596632

15. https://www.researchgate.net/publication/336745269_Climate_resilience_in_the_United_Kingdom_wine_production_sector_CREWS-UK
16. https://www.mdpi.com/2306-5710/5/2/33
17. https://www.ajevonline.org/content/73/3/125
18. https://www.sciencedirect.com/science/article/pii/S2214799319300165