

---

# **WALRUS Data Analysis: An Exploration of Sales and Artist Demographics**

**Melissa Mickelson**  
**11/30/2023**

# Outline

---

- Introduction
- Exploratory Data Analysis
- Artist Demographics and Sales
- Key Findings and Implications
- Conclusion
- Appendix

# Introduction

---

WALRUS Art Gallery aims to understand the intricacies of its sales patterns and artist demographics over the past 3 years to optimize strategies for business success and support its diverse community of artists.

## Main Goal

Uncover actionable insights to enhance the gallery's operational efficiency, elevate artist visibility, and foster a more inclusive and successful art community.

## Hypotheses

- 1.The analysis will reveal specific days or seasons with heightened sales activity.
- 2.Diverse demographic insights will provide strategies for tailored promotions.
- 3.Statistical analysis will identify significant factors influencing sales, allowing for targeted improvements.

---

# Exploratory Data Analysis

---

# Methodology Overview

---

## Data Sources

- Two primary datasets utilized: 'WALRUS\_sales.csv' (sales data) and 'WALRUS\_artists.csv' (artist demographics).

## Data Cleaning

- Thorough cleaning process to address missing values and ensure data accuracy.
- Standardization of data formats for consistency.

## Data Merging

- Integration of sales and artist demographics datasets to enhance analytical capabilities.
- Merging performed based on the 'Artist' column, ensuring a seamless connection between sales and artist information.

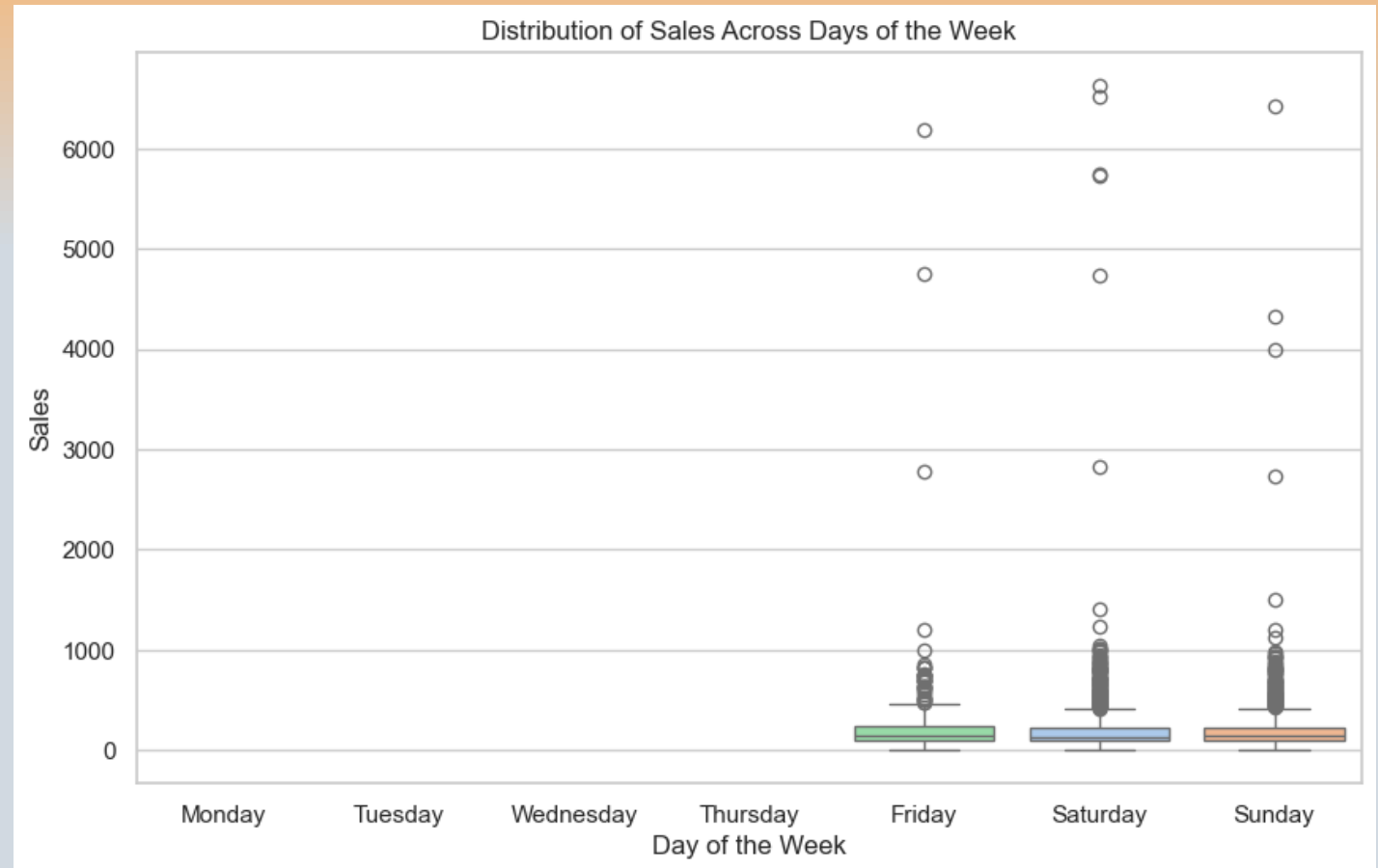
# Days of the Week Analysis

## Gallery Open Days

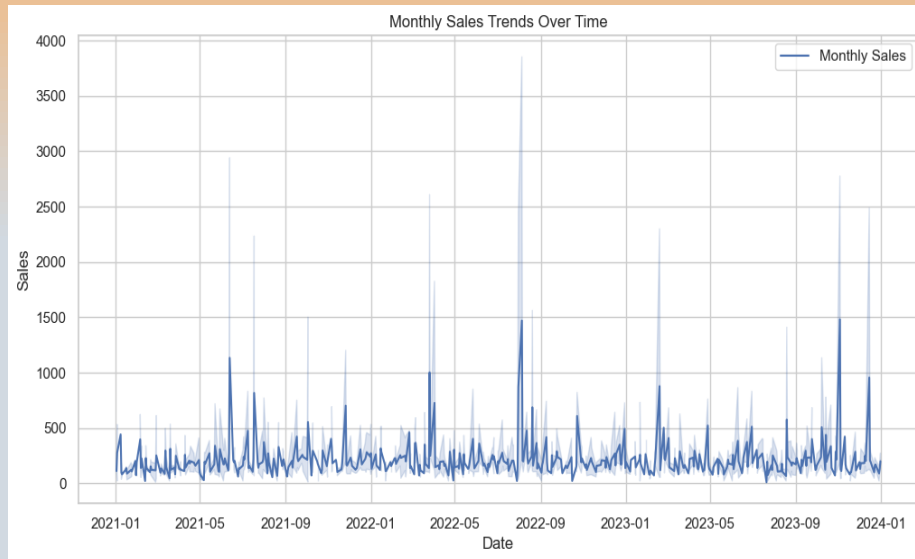
The gallery is open only on Fridays, Saturdays, and Sundays.

## Distribution of Sales Across Days of the Week

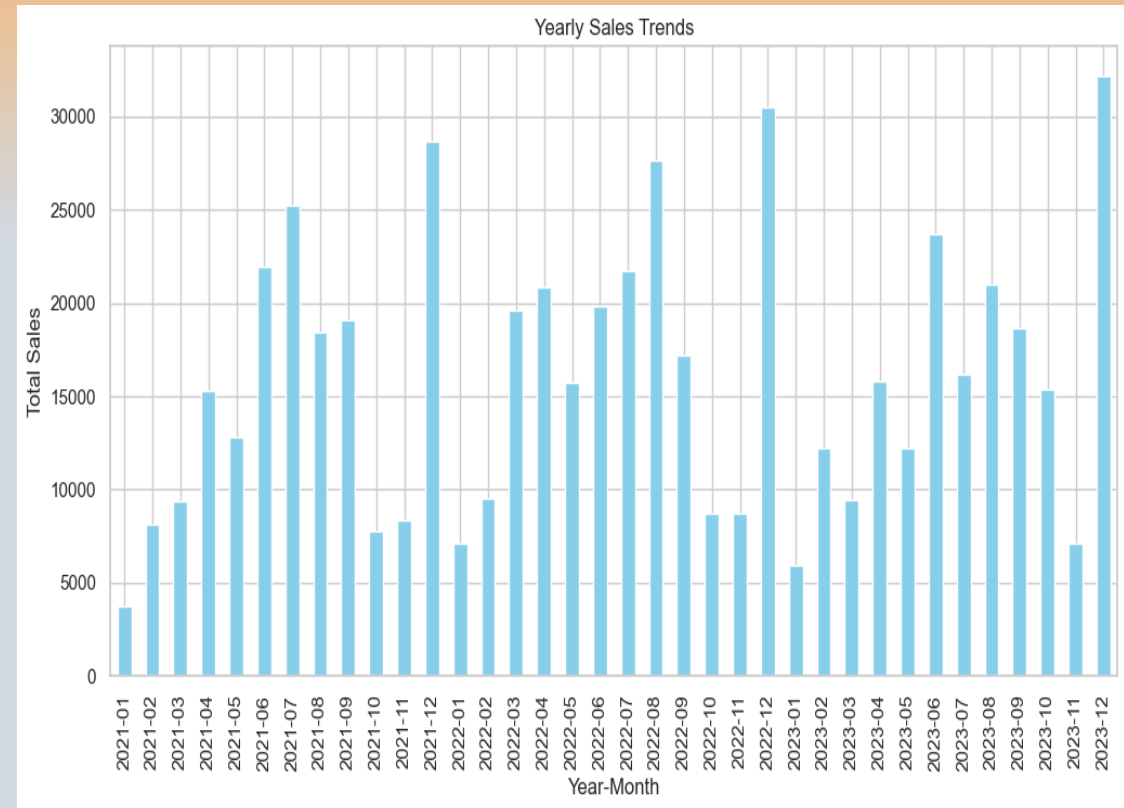
- The least sales are on Fridays and the most on Saturdays.
- Notably, Saturday has the highest value outliers with Sunday and Friday not far behind, indicating occasional exceptional sales on all days.



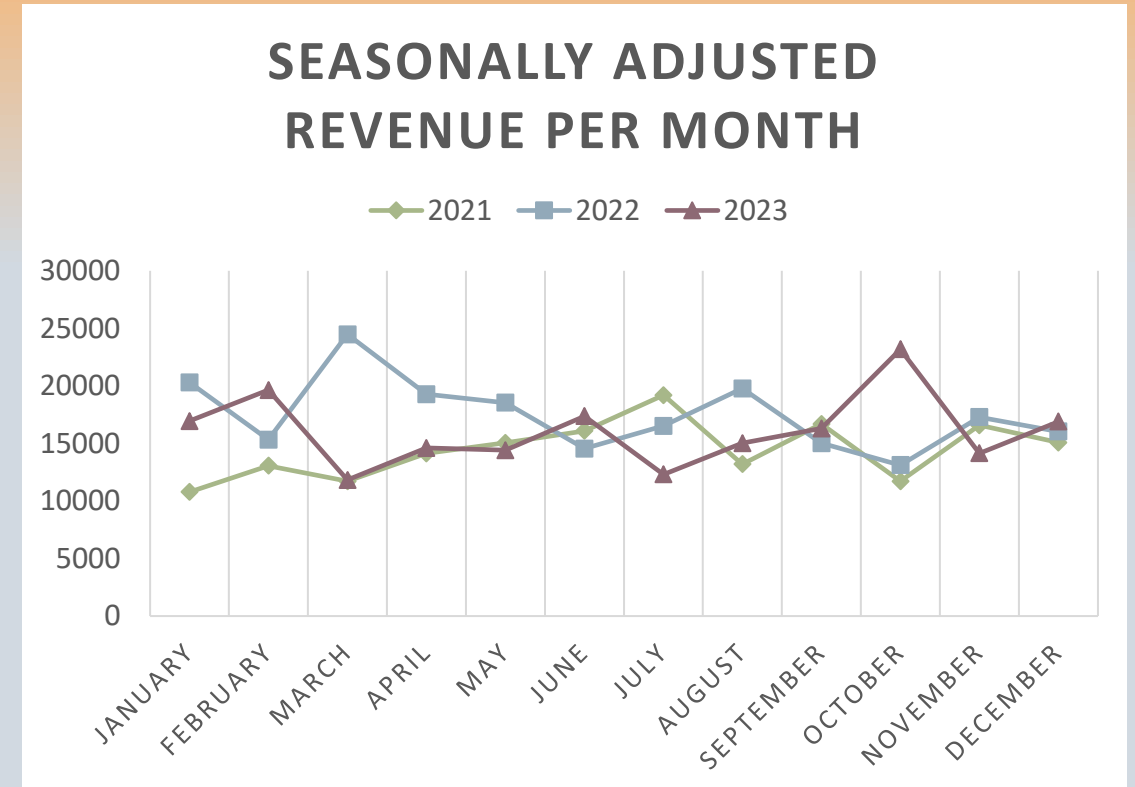
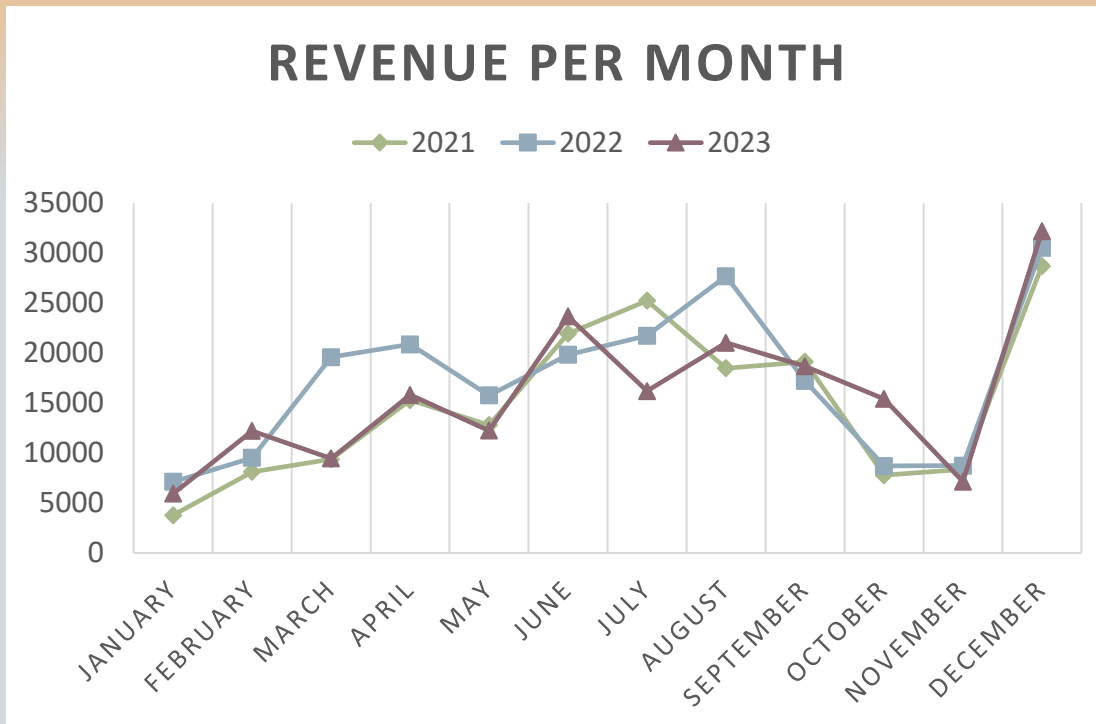
# Seasonal Trends



Sales peak around December and show a drop afterwards, with a moderate increase during the summer.



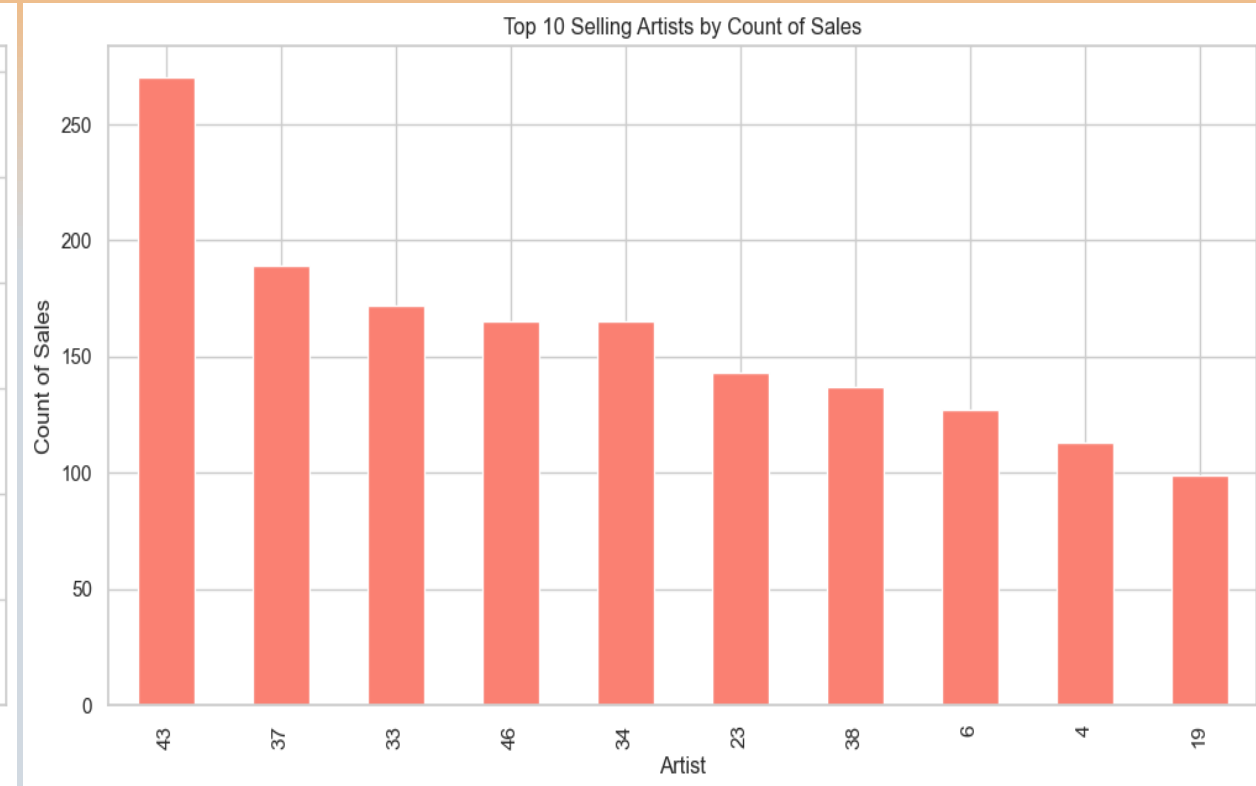
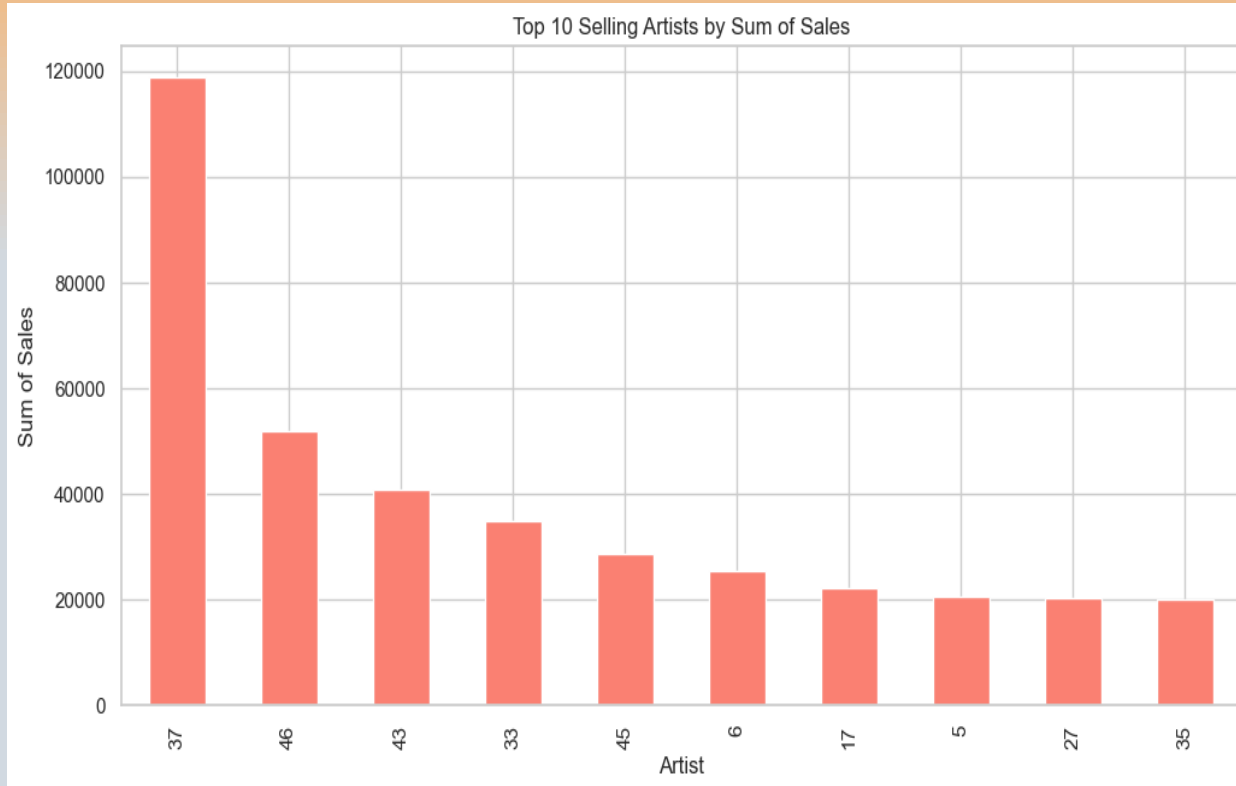
# Adjusted Sales



After adjusting for seasonality, revenue looks relatively stable over the past 3 years.



# Top Selling Artists



The highest-selling artist was not the most profitable.

---

# **Artist Demographics and Sales**

---

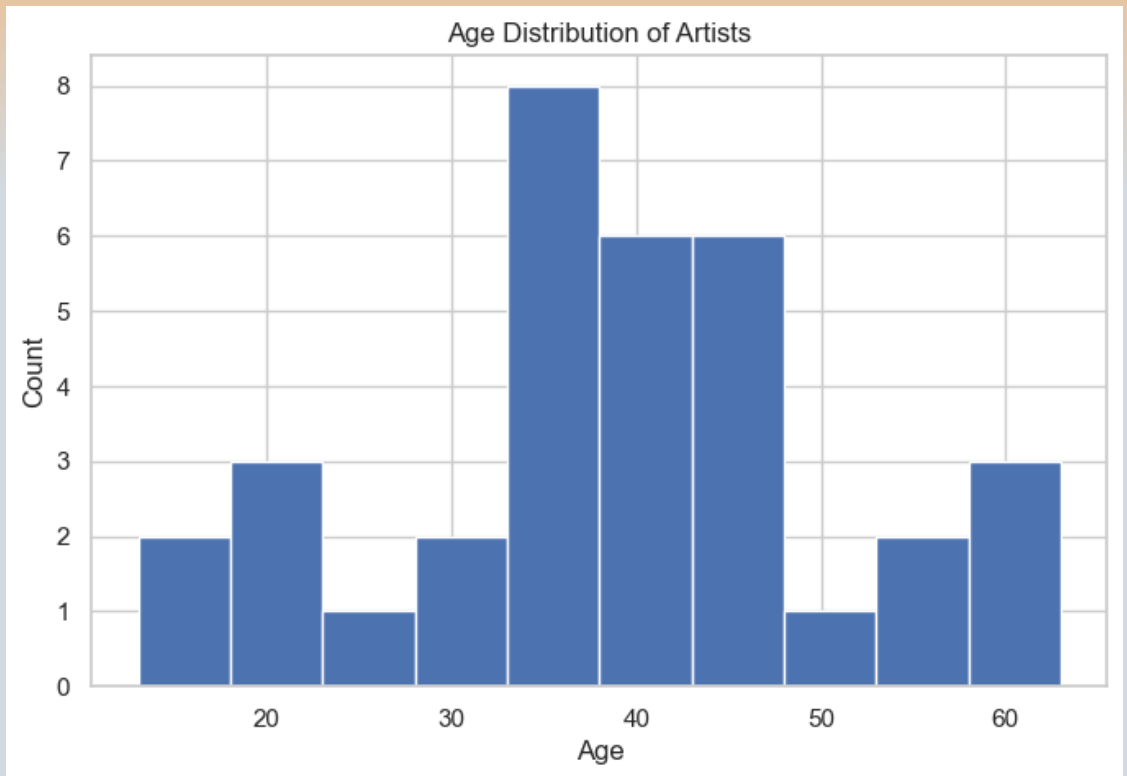
# The Artist Demographics Dataset

## Introduction to WALRUS Artists' Demographics

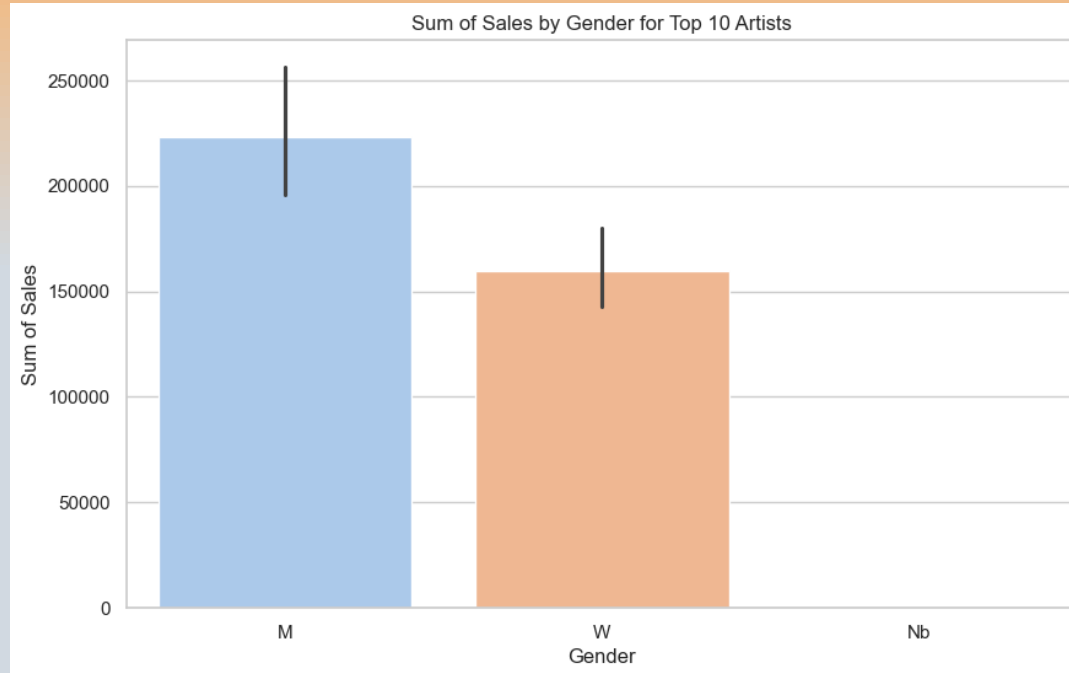
- Beyond sales figures, our second dataset captures crucial aspects of our artists' identities.
- Key columns include age, race, gender, LGBTQ identity, and promotion acceptance.

## Significance

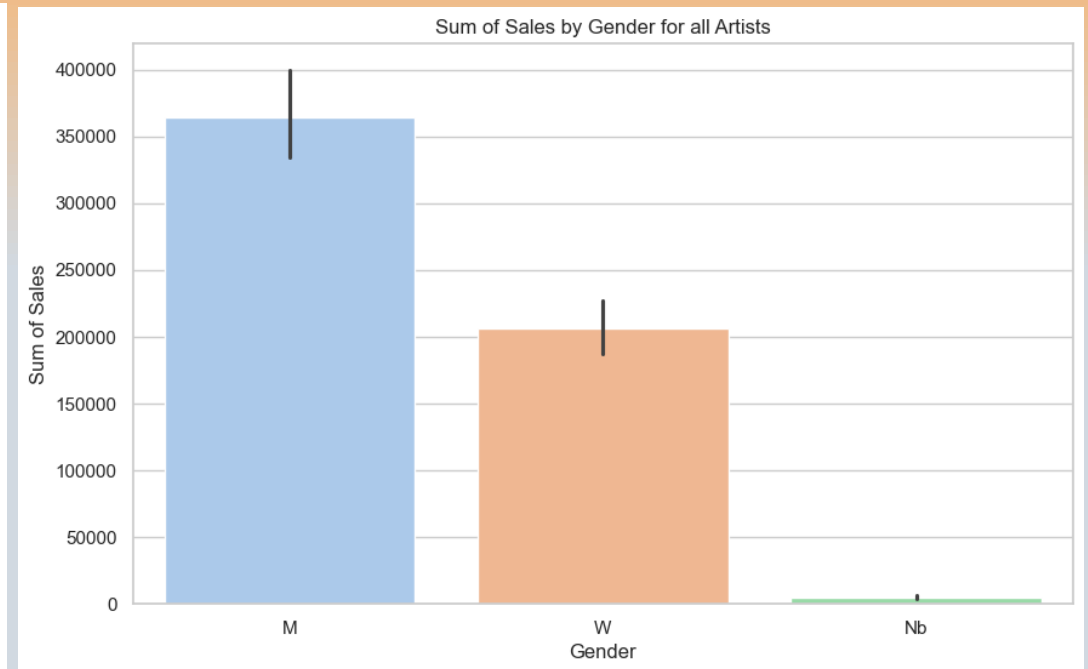
- Empowering artists by understanding their demographics fosters a more inclusive artistic community.
- Insights into artists' preferences regarding the use of demographic information for promotions.



# Sum of Sales by Gender

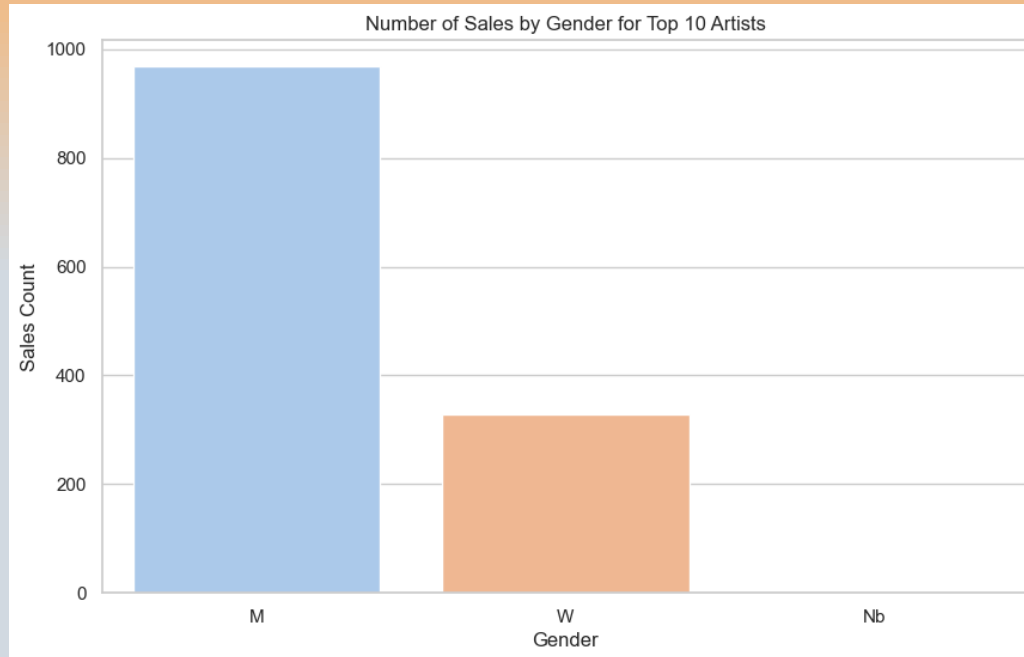


The top 10 showcases a gender-skewed distribution, suggesting potential biases in selection or varied pricing strategies.

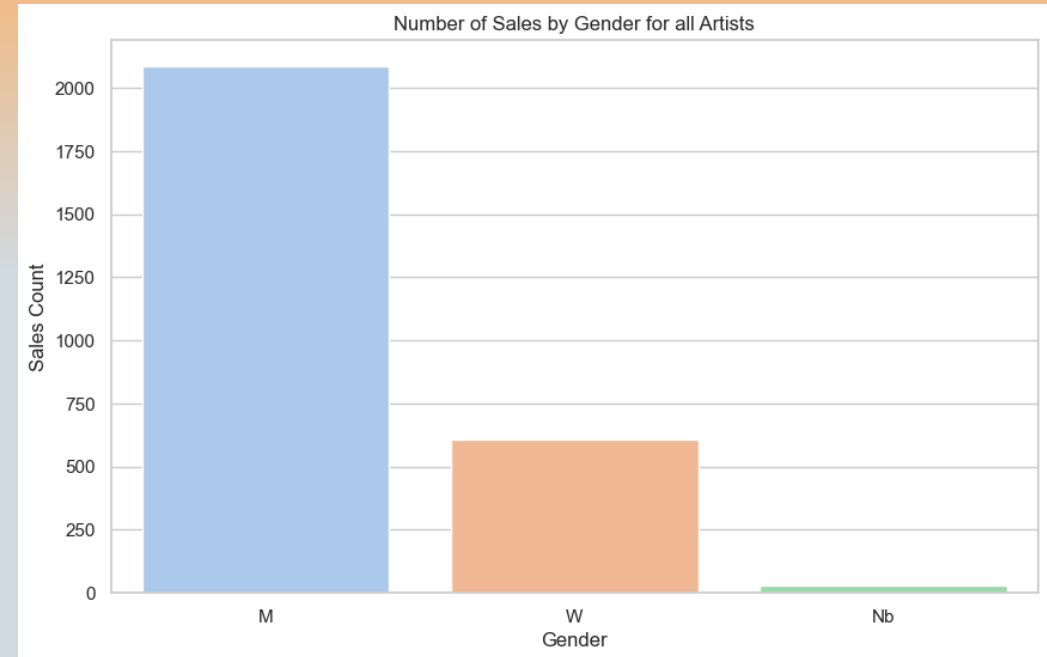


Aggregating sales reinforces gender disparities, hinting at systemic influences affecting artists across the board.

# Count of Sales by Gender



A skewed distribution in sales volume among the top 10 highlights potential biases impacting not only revenue but also the visibility of artists.



The collective sales volume reiterates gender disparities, urging a closer look at factors influencing the visibility and success of artists from different gender identities.

# Average Sales by Gender

## Statistical Analysis Highlights

### Men Sales

- Men exhibit significantly lower sale prices on average compared to women.

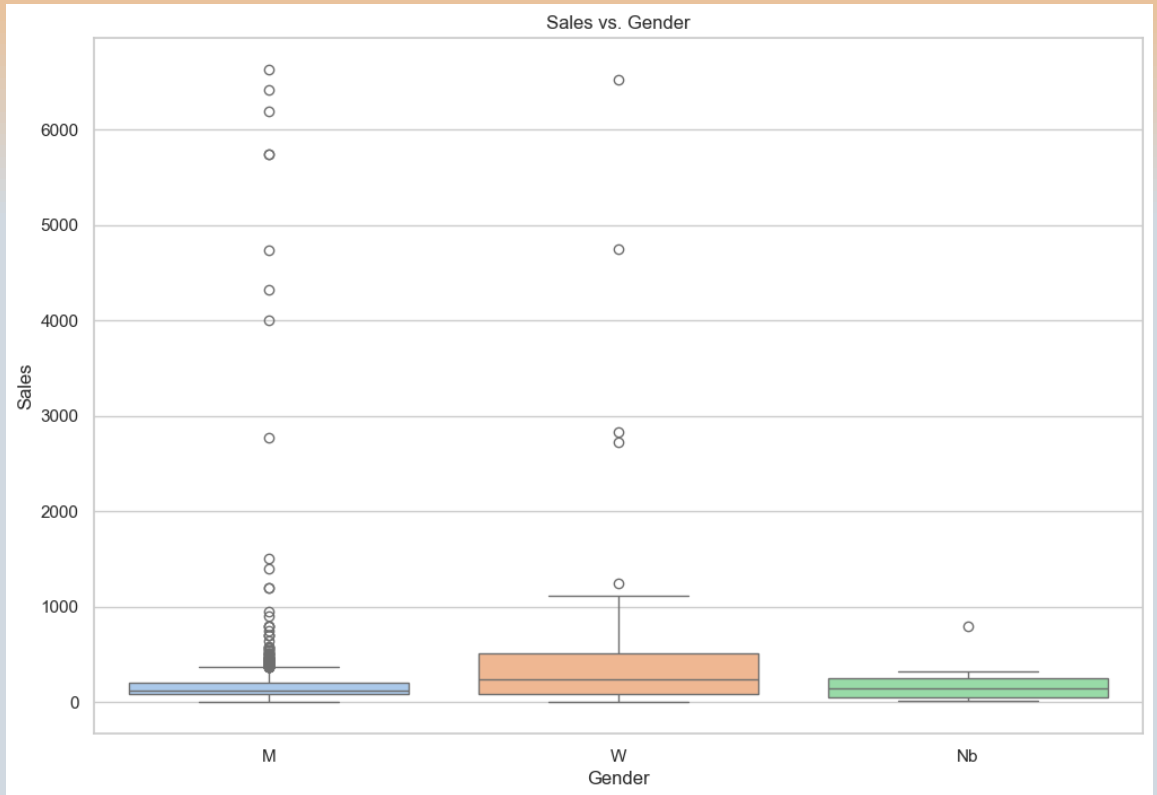
### Women Sales

- Women demonstrate significantly higher sale prices on average compared to men.

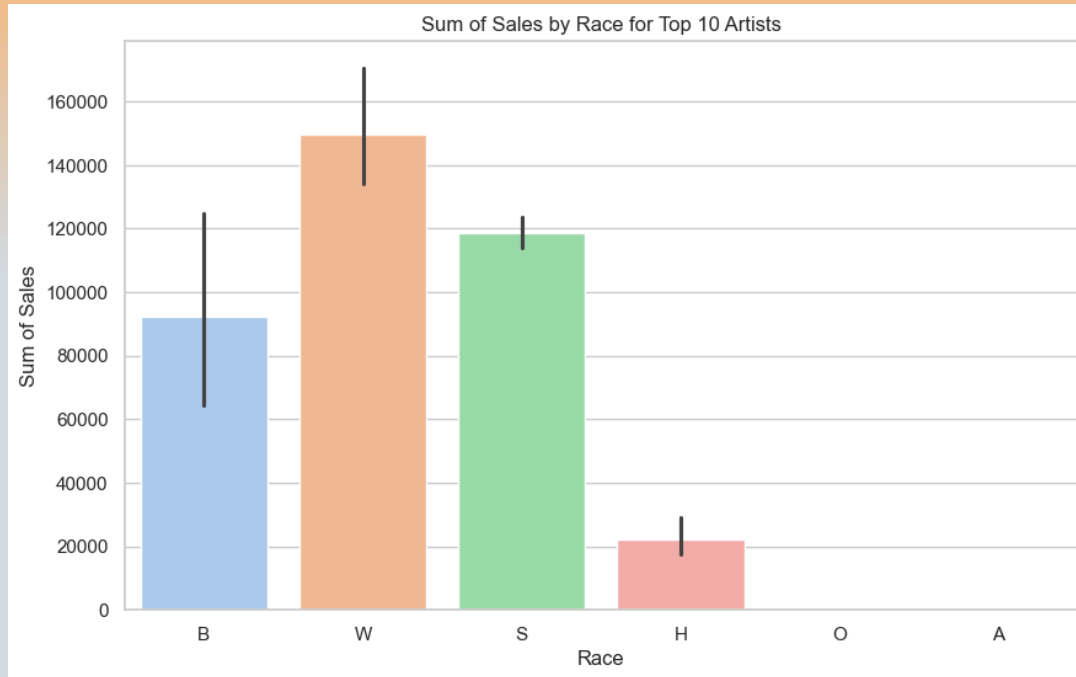
### Nonbinary Sales

- Nonbinary artists show no significant difference in sale prices compared to the overall dataset.

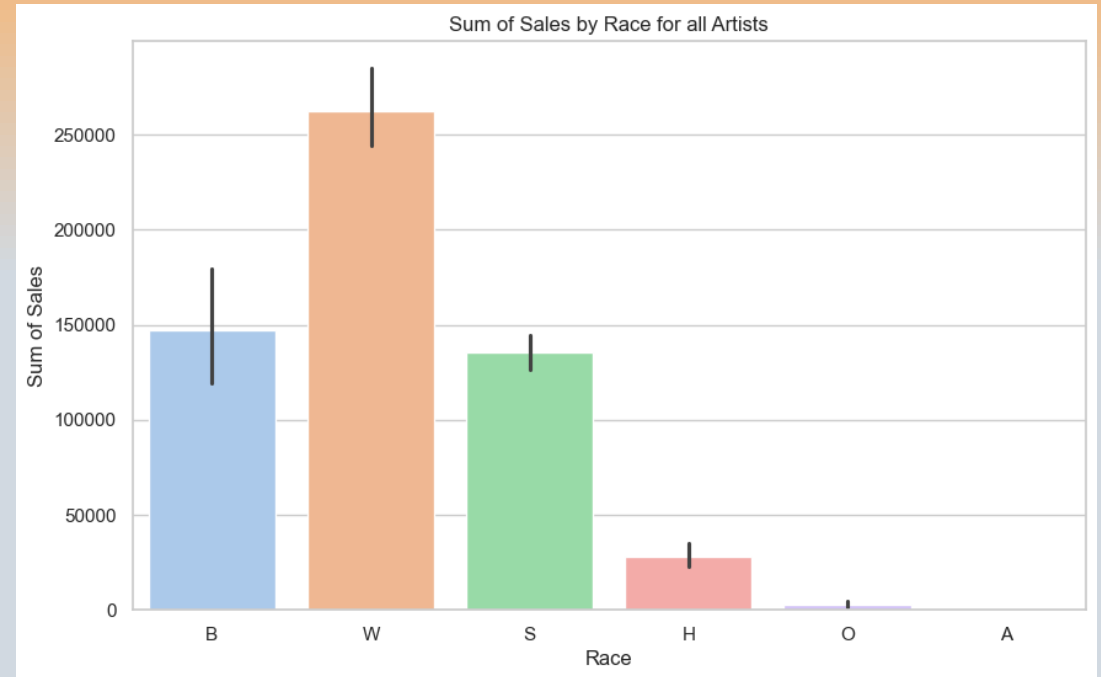
Though the range for sale prices is similar for men and women, the inner quartile for women is significantly higher.



# Sum of Sales by Race

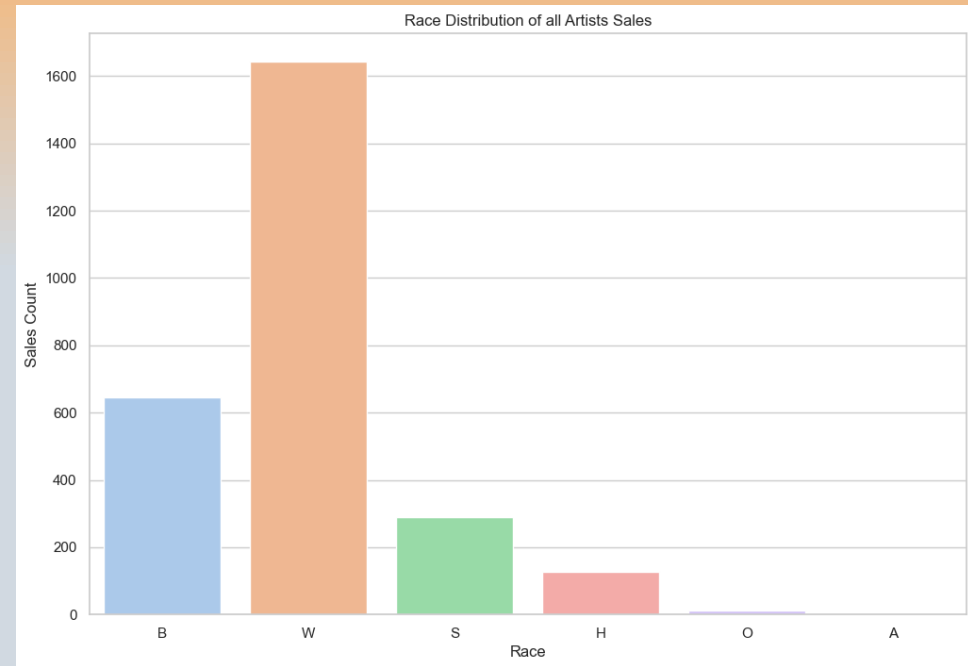
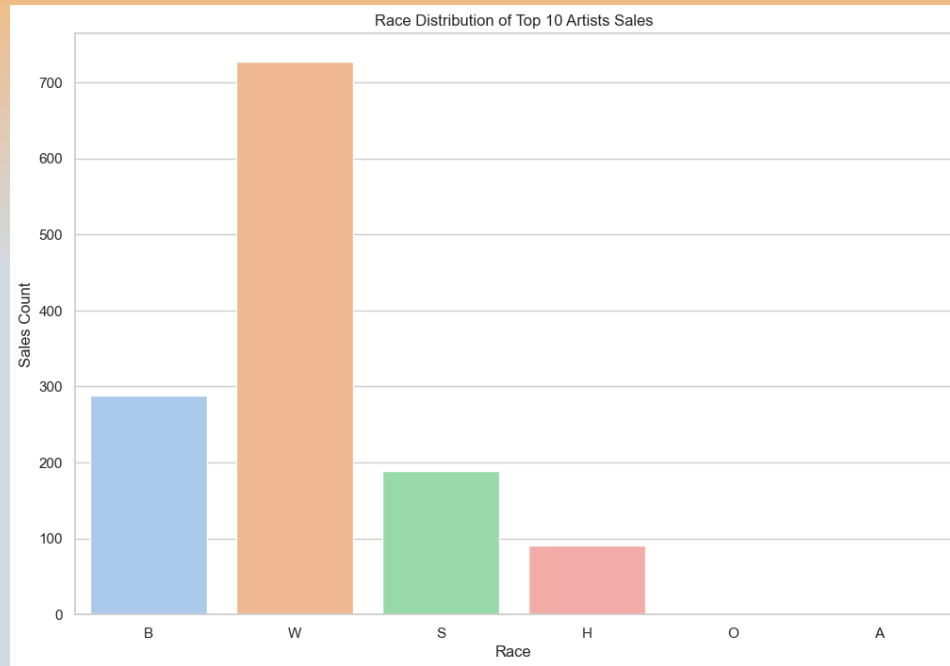


Asian (S) artists demonstrate substantial sales in the overall dataset and an even higher ratio in the top 10, showcasing a potential market demand.



White (W) artists lead in sales for both top 10 and all artists, indicating a consistent trend.

# Count of Sales by Race



The prominence of White (W) artists in sales highlights the need for equitable representation and promotional strategies for artists from other racial backgrounds.



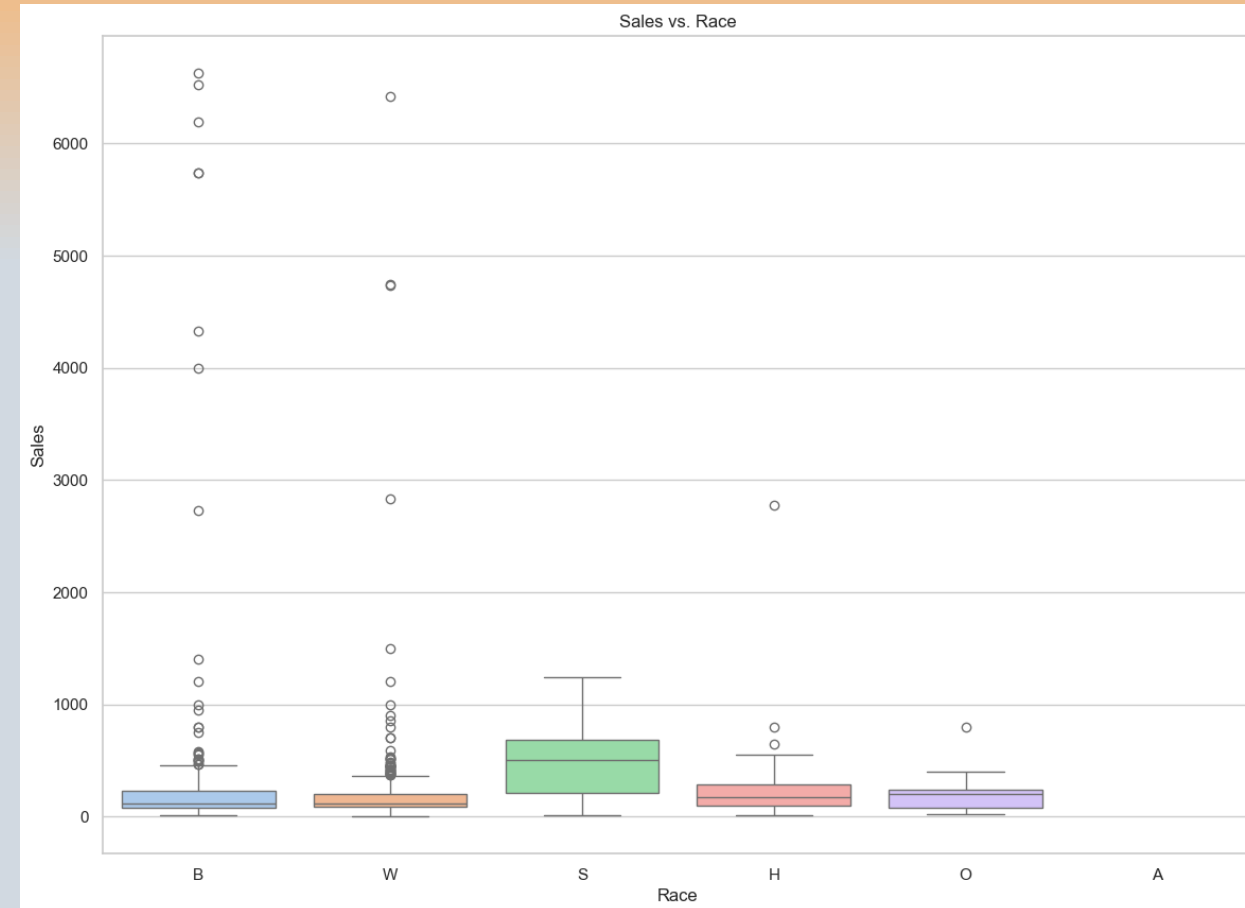
# Average Sales by Race

## Statistical Analysis Highlights

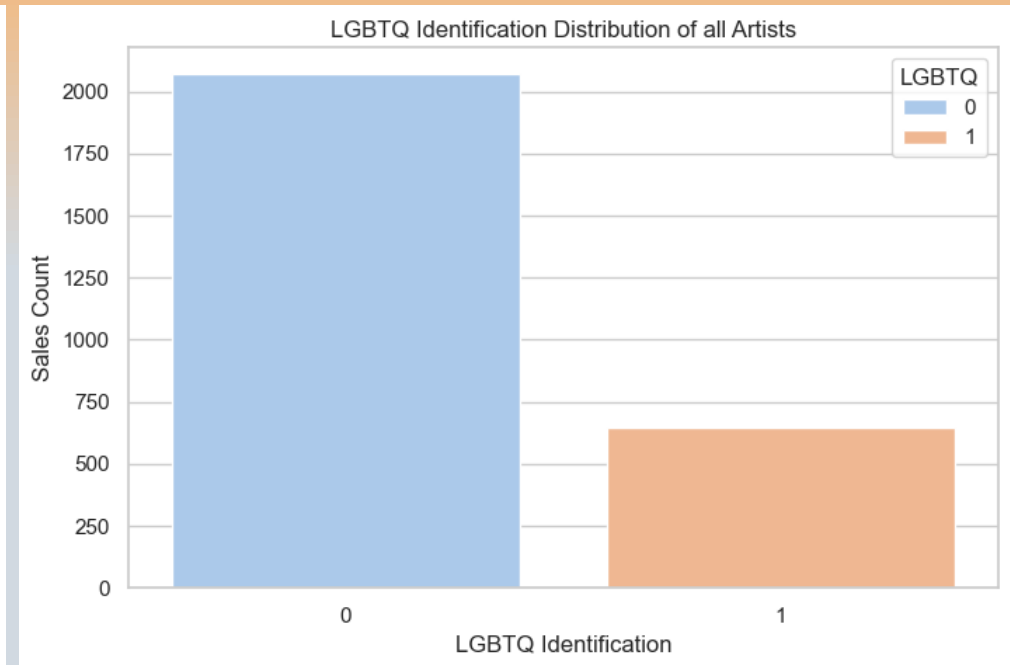
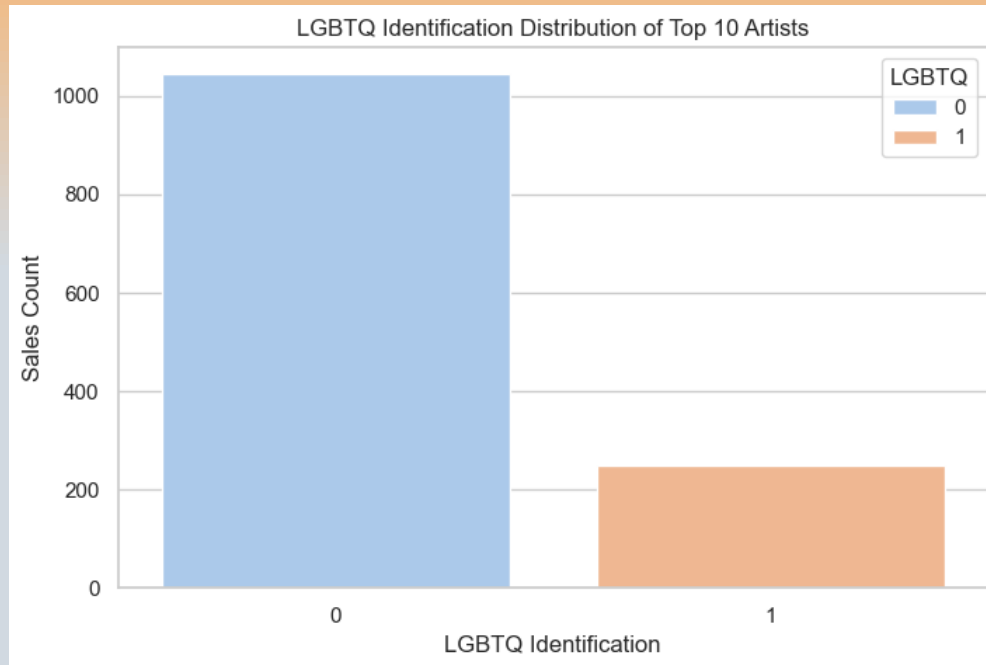
### Significant Disparities

- **White (W):** Statistically significant lower mean sales with a highly significant t-statistic, suggesting a need for targeted strategies.
- **Asian (S):** Demonstrates the highest mean sales and significant t-statistic, indicating potential market demand and profitability.

The analysis reveals notable disparities in mean sales among racial categories. Opportunities lie in further exploring and enhancing sales for Black (B), Hispanic (H), American Indian (A), and Other (O) artists.



# Sales by LGBTQ Identity



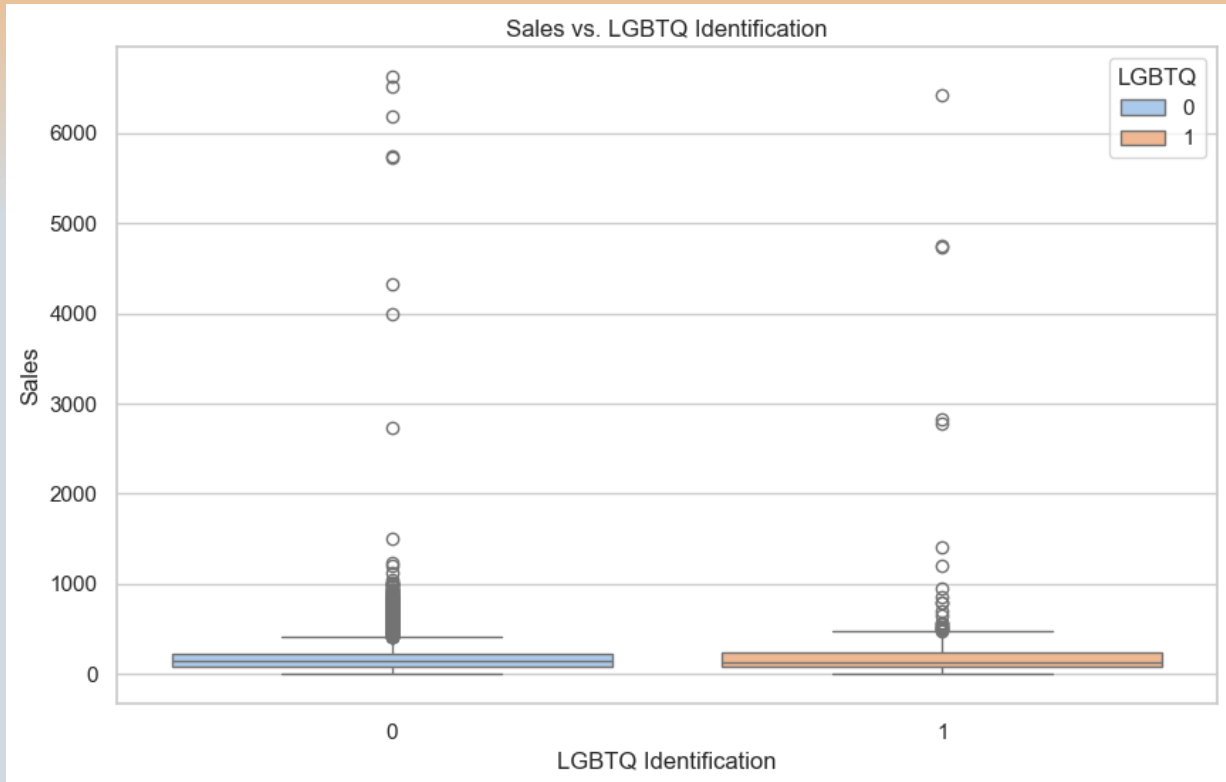
There is slightly less LGBTQ representation in the sales of the top 10 artists when compared to the sales of all artists.

# Average Sales by LGBTQ Identity

## Statistical Analysis Highlights

The analysis indicates that LGBTQ+ identity does not significantly influence art sales.

Both LGBTQ+ and non-LGBTQ artists demonstrate similar mean sales, range, and inner quartile, suggesting a level playing field in terms of market acceptance.



---

# **Key Findings, Implications, and Hypotheses Review**

---

# Key Findings

---

## Top Artists Strategy

- Focus on collaboration and support for high-performing artists, especially '37,' '46,' '43,' and '33.'
- Investigate and replicate success factors of top-selling artists, particularly engaging with '37' for insights.

## Sales Optimization Strategies

- Invest in themed seasonal promotions, especially during peak months like December and summer.
- Investigate the factors contributing to exceptional Saturday sales and implement targeted promotions.

## Demographic Insights

### Gender Differences

- Significant differences in sales based on gender.
- Though men have more sales, their sales are less profitable than women. The factors behind this need to be further explored.

### Race-Specific Strategies

- Significant differences in sales across racial categories.
- Tailor strategies for each racial group, emphasizing diversity, collaboration, and thematic exhibitions. Further investigate the profitability of Asian artists.

# Implications

---

## **Inclusive Promotion**

- Leverage statistical differences to emphasize diversity in promotional strategies.
- Tailor promotions for different demographic groups to foster an inclusive art community.

## **Ongoing Exploration**

- Encourage continuous exploration and data collection for nuanced insights over time.
- Implement adaptive strategies based on evolving market dynamics and trends.

## **Limitations and Future Directions**

- Acknowledge the small sample size limitation, especially with certain demographics.
- Investigate the strategies contributing to the higher sales volume for Men and White artists and look for ways to apply it to other demographics.
- Investigate the higher average sales price for Women and Asian artists. Look for strategies that can be applied to other demographics.

# Conclusion

---

## **Hypothesis 1:** Specific Days or Seasons with Heightened Sales Activity

Supported by findings on exceptional Saturday sales and the recommendation to explore themed seasonal promotions.

## **Hypothesis 2:** Diverse Demographic Insights for Tailored Promotions

Validated by significant differences in sales across gender and racial categories, leading to tailored promotional strategies.

## **Hypothesis 3:** Statistical Analysis Identifying Significant Factors

Confirmed through the identification of influential factors, guiding targeted improvements in sales strategies, such as some artists having larger quantities of sales compared to others having larger sale prices.

## **Summary**

Focus on lacking demographics when bringing in new artists, prioritize diversity, inclusion, and continuous exploration for overall gallery success.

# Appendix

---

Jupyter Notebook – EDA and Visualizations

[https://github.com/mnmickelson/portfolio...WALRUS\\_SalesAnalysis.ipynb](https://github.com/mnmickelson/portfolio...WALRUS_SalesAnalysis.ipynb)

| Race | Count | Mean | Std | Min | Max  |
|------|-------|------|-----|-----|------|
| W    | 1644  | 160  | 254 | 20  | 6420 |
| B    | 646   | 227  | 597 | 8   | 6625 |
| S    | 290   | 468  | 283 | 10  | 1240 |
| H    | 128   | 218  | 270 | 12  | 2775 |
| O    | 13    | 217  | 210 | 20  | 800  |

| LGBTQ | Count | Mean | Std | Min | Max  |
|-------|-------|------|-----|-----|------|
| 0     | 2074  | 214  | 371 | 20  | 6625 |
| 1     | 647   | 203  | 408 | 8   | 6420 |

| Gender | Count | Mean | Std | Min | Max  |
|--------|-------|------|-----|-----|------|
| M      | 2058  | 175  | 357 | 20  | 6625 |
| W      | 606   | 341  | 157 | 8   | 6520 |
| Nb     | 30    | 172  | 432 | 20  | 800  |



# Appendix(continued)

Excel Sheet –  
[https://github.com/mnmickelson/portfolio...WALRUS\\_DeSeasonalization.xlsx](https://github.com/mnmickelson/portfolio...WALRUS_DeSeasonalization.xlsx)

| Monthly Revenue |       |       |       |
|-----------------|-------|-------|-------|
| Month           | 2021  | 2022  | 2023  |
| January         | 3780  | 7120  | 5940  |
| February        | 8113  | 9520  | 12200 |
| March           | 9370  | 19575 | 9450  |
| April           | 15283 | 20850 | 15795 |
| May             | 12785 | 15750 | 12240 |
| June            | 21930 | 19820 | 23660 |
| July            | 25225 | 21715 | 16183 |
| August          | 18470 | 27670 | 21012 |
| September       | 19105 | 17195 | 18665 |
| October         | 7780  | 8705  | 15395 |
| November        | 8370  | 8730  | 7140  |
| December        | 28693 | 30515 | 32155 |

| Seasonally Adjusted Monthly Revenue |          |          |          |
|-------------------------------------|----------|----------|----------|
| Month                               | 2021     | 2022     | 2023     |
| January                             | 10772.55 | 20291.15 | 16928.29 |
| February                            | 13051.29 | 15314.71 | 19626.00 |
| March                               | 11712.07 | 24467.86 | 11812.07 |
| April                               | 14124.59 | 19269.63 | 14597.78 |
| May                                 | 15047.89 | 18537.68 | 14406.43 |
| June                                | 9380.49  | 14542.14 | 17359.59 |
| July                                | 16673.23 | 16509.77 | 12303.83 |
| August                              | 18027.73 | 19775.12 | 15016.80 |
| September                           | 16126.85 | 15013.60 | 16297.11 |
| October                             | 28760.58 | 13104.47 | 23175.56 |
| November                            | 15403.37 | 17284.25 | 14136.26 |
| December                            | 4396.67  | 16029.20 | 16890.68 |

# Appendix(continued)

---

## Race:

Statistical Analysis for B vs. Sales:  
t-statistic: 0.8401495680367369, p-value:  
0.40088421936944707

Statistical Analysis for W vs. Sales:  
t-statistic: -4.917895216011172, p-value:  
9.06968165672657e-07

Statistical Analysis for S vs. Sales:  
t-statistic: 11.15660810569325, p-value:  
2.3742625963404517e-28

Statistical Analysis for H vs. Sales:  
t-statistic: 0.1783829577521755, p-value:  
0.8584349241298375

Statistical Analysis for O vs. Sales:  
t-statistic: 0.05000274595452199, p-value:  
0.9601238553371794

## LGBTQ:

Statistical Analysis for 1 vs. Sales:  
t-statistic: -0.48785960964587965,  
p-value: 0.6256810600824065

Statistical Analysis for 0 vs. Sales:  
t-statistic: 0.2341699430975336, p-  
value: 0.8148630498725435

## Promotion Acceptance:

Statistical Analysis for 1 vs. Sales:  
t-statistic: -0.011246336624944507,  
p-value: 0.9910273545062369

Statistical Analysis for 0 vs. Sales:  
t-statistic: 0.034209683694690844,  
p-value: 0.9727121448993024

## Gender:

Statistical Analysis for M vs. Sales:  
t-statistic: -3.4362397263117463,  
p-value: 0.000594848689030012

Statistical Analysis for W vs. Sales:  
t-statistic: 7.386556901707277, p-  
value: 1.8954739208764997e-13

Statistical Analysis for Nb vs. Sales:  
t-statistic: -0.5784280389384268,  
p-value: 0.5630225509300639