INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

The Whiskey Code

Behind the label do you believe in quantity or quality

HARD TO FIND® WHISKY
Selling Collectables Since 1991

About me

- I have recently completed my Bachelor of Computer Applications (BCA) from Parul University, Vadodara.
- I have prior experience in Python and data manipulation, so I developed a strong interest in the field of Data Science.
- Motivated by this passion, I have decided to pursue Data Science as my career path.
- Here are links to my GitHub and LinkedIn profiles.
  - https://www.linkedin.com/in/shubham-parmar-12164024a/
  - https://github.com/shubham132004

Problem Statement :

- Due to lack of structured data in the collectible whisky market leads to uncertain decisions.

- This project uses data analysis to uncover insights into pricing, quality, and trends — enabling smarter investments and sales.

Objective :

- The objective of this project is to perform an exploratory data analysis (EDA) on rare and collectible whisky data to identify patterns, trends, etc.

- By analyzing key attributes such as brand, type, alcohol content (ABV%), and pricing, the project aims to generate actionable insights that can support better decision-making for investors, and retailers.

Web Scrapping - Process

➤ Website : https://www.htfw.com/
   (Hard To Find Whiskey)

Steps Followed While Scrapping :

1. **Inspecting Structure :** Right click on your mouse and inspect the website to get various tags.

2. **Column Identification :** Selecting appropriate tags for extracting various categories like Brand , Price , Type , etc

3. **Libraries used while scrapping :**

   - **requests** – To send http request and get response from server
   - **BeautifulSoup** – For parsing the html content

4. **Other Libraries :**

   - **Pandas** – For data manipulation .
   - **Matplotlib and Seaborn** – For visualizing the data

INNOMATICS RESEARCH LABS

SUMMARY

Numerical Columns

1. **Price_Rupees** : Whiskeys Prices in INR
2. **ABV_Percent** : Shows how much percentage of alcohol whiskey contains
3. **Centi_liter** : Shows number of Centiliter each whiskey contains
4. **Mili_liter** : Shows number of Mililiter each whiskey contains

Categorical Columns

1. **Brand** : Contains all the whiskey brand names
2. **Varieties** : Contains all the varieties of whiskey
3. **Type** : Contain different types of whiskey names
4. **Year** : Shows how old the whiskeys are
5. **ABV_Category** : This shows how much alcoholic the whiskeys are and this are divided in Low , Medium and High .

Data Frame After Scrapping The Wesbite

	Brand	Varieties	Type	Centi_liter	Year	ABV_Percent	Price
0	Angel's Envy	Straight Port Cask Finish Bourbon	Bourbon	70cl		3% ABV	£59.95
1	Angel's Envy	Manhattan Bundle - Martini Rubino Vermouth & S...	Bourbon	70cl		3% ABV	£79.95
2	Buffalo Trace	Traveller Bourbon Whiskey (Chris	Bourbon	70cl		45% ABV	£31.95
3	Buffalo Trace	Kosher Wheat Recipe Bourbon	Bourbon	75cl		47% ABV	£43.95
4	Elijah Craig	Barrel Proof Kentucky Straight Bourbon	Barrel	70cl		8% ABV	£124.95
5	Maker's Mark	Kentucky Straight Bourbon	Kentucky	70cl		45% ABV	£35.95
6	Ben Holladay	Soft Red Wheat Bottled In Bond Straight Bourbon	Bourbon	70cl		50% ABV	£79.95
7	Booker's	2024-01 Batch Bourbon	Bourbon	70cl	7 Year Old	2% ABV	£89.95
8	Wild Turkey	Rare Breed Kentucky Straight Bourbon	Kentucky	70cl		4% ABV	£59.95
9	Wild Turkey	81 Kentucky Straight Bourbon	Kentucky	70cl		5% ABV	£29.95

Data Frame After Cleaning The Data

	Brand	Varieties	Type	Centi_liter	Year	ABV_Percent	Price_Rupees	ABV_Category	Mili_liter
0	Angel's Envy	Straight Port Cask Finish Bourbon	Bourbon	70	NAS	3	6810	Low	700
1	Angel's Envy	Manhattan Bundle - Martini Rubino Vermouth & S...	Bourbon	70	NAS	3	9082	Low	700
2	Buffalo Trace	Traveller Bourbon Whiskey (Chris	Bourbon	70	NAS	45	3630	Medium	700
3	Buffalo Trace	Kosher Wheat Recipe Bourbon	Bourbon	75	NAS	47	4993	High	750
4	Elijah Craig	Barrel Proof Kentucky Straight Bourbon	Barrel	70	NAS	8	14194	Low	700
5	Maker's Mark	Kentucky Straight Bourbon	Kentucky	70	NAS	45	4084	Medium	700
6	Ben Holladay	Soft Red Wheat Bottled In Bond Straight Bourbon	Bourbon	70	NAS	50	9082	High	700
7	Booker's	2024-01 Batch Bourbon	Bourbon	70	7 Year Old	2	10218	Low	700
8	Wild Turkey	Rare Breed Kentucky Straight Bourbon	Kentucky	70	NAS	4	6810	Low	700
9	Wild Turkey	81 Kentucky Straight Bourbon	Kentucky	70	NAS	5	3402	Low	700

Whiskey Types

- Bourbon
- Barrel
- Kentucky
- Tennessee
- Rye
- American

- Eagle
- Rum Finish
- Irish
- Very Rare
- Japanese
- Suntory
- Harmony

- Indian
- English
- Canadian
- Welsh
- Swedish
- Formula 1

INNOMATICS RESEARCH LABS
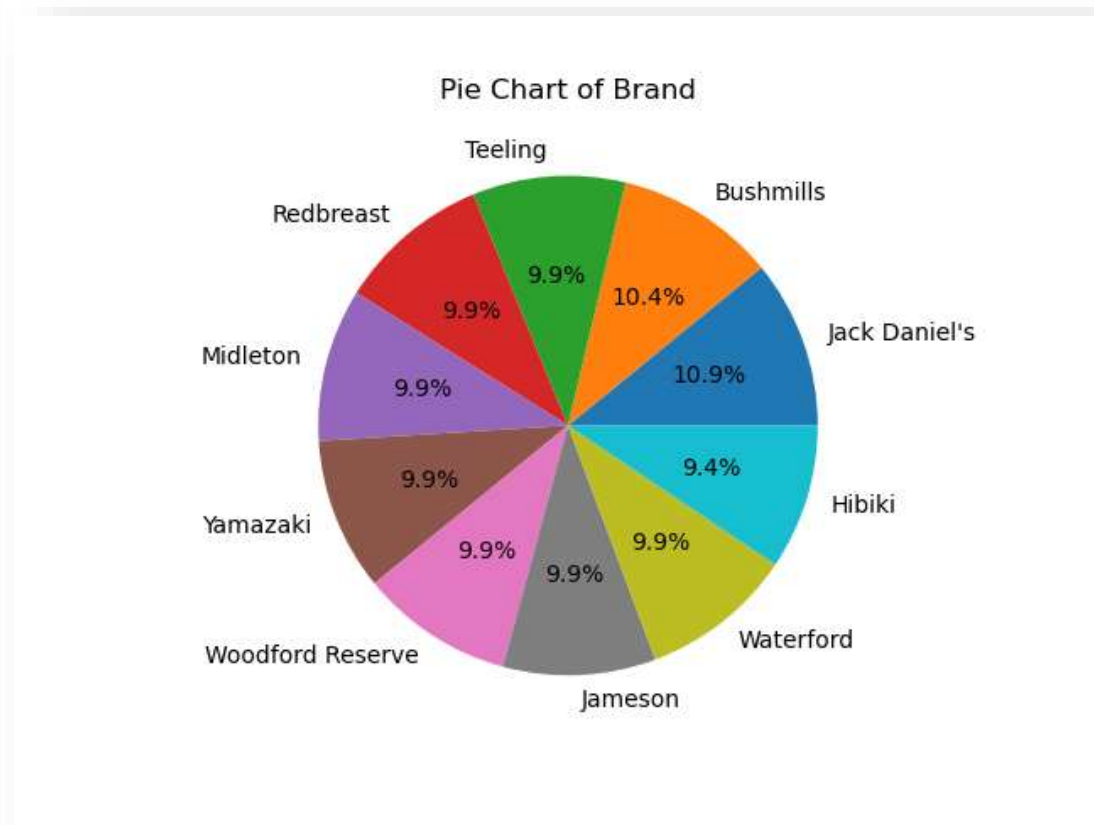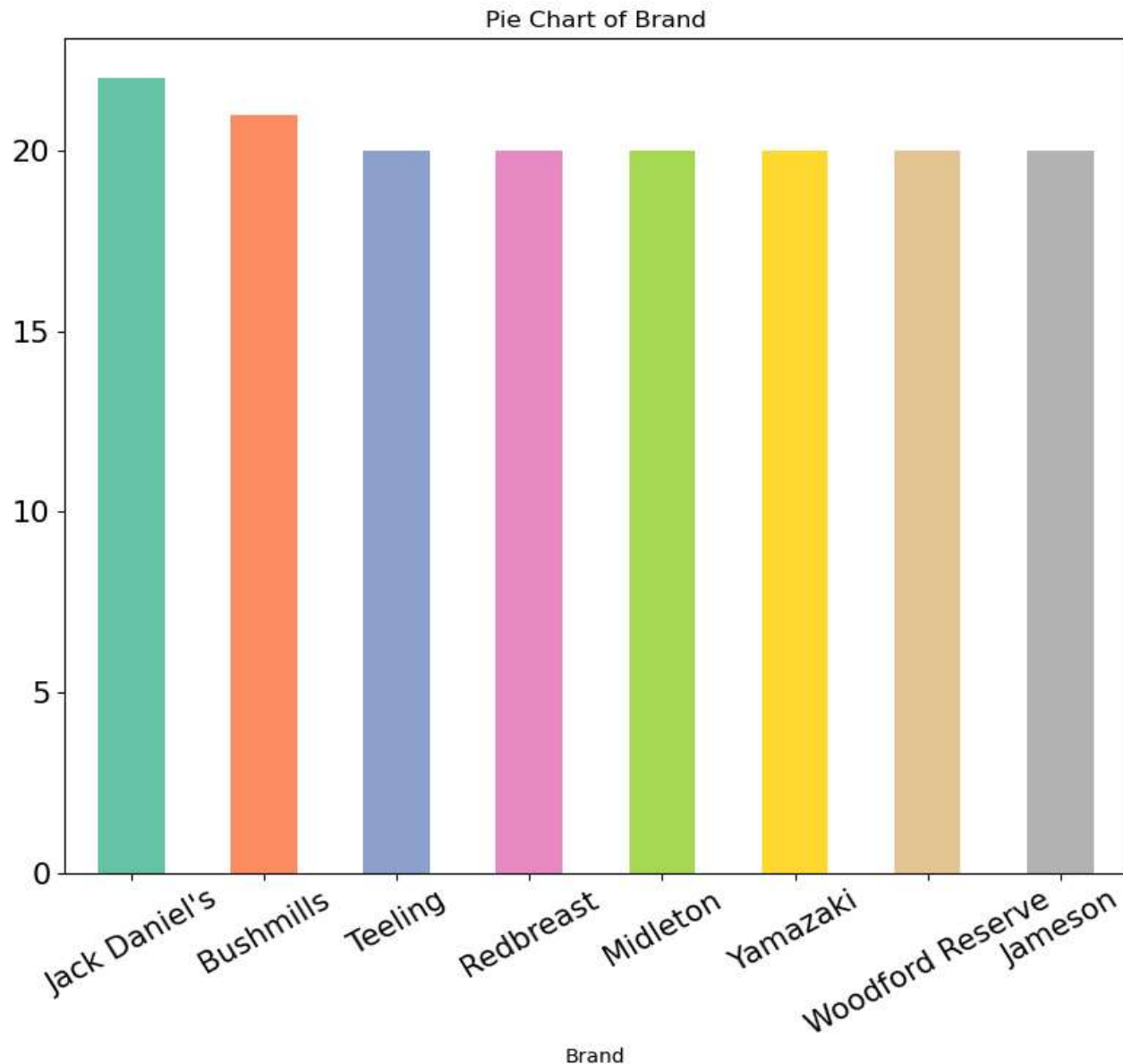
Univariate Graphs On Numerical Columns
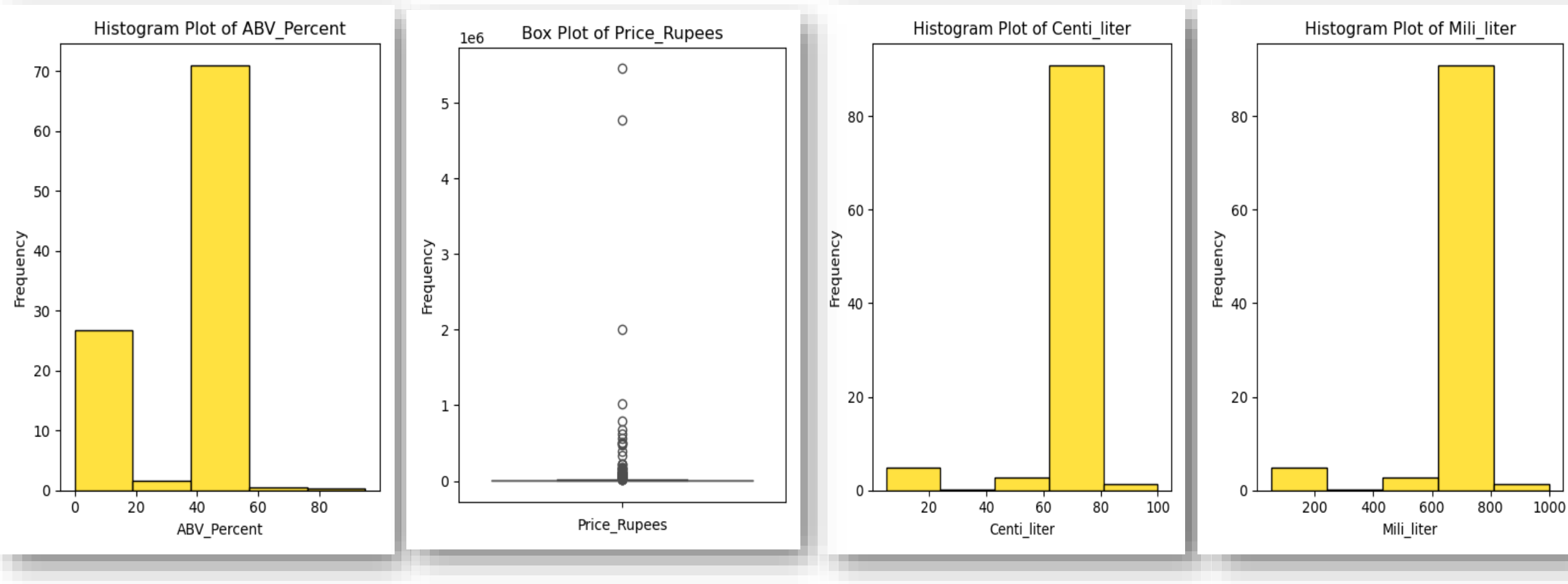
Observation : The above plot captures the distribution of data in numerical columns.
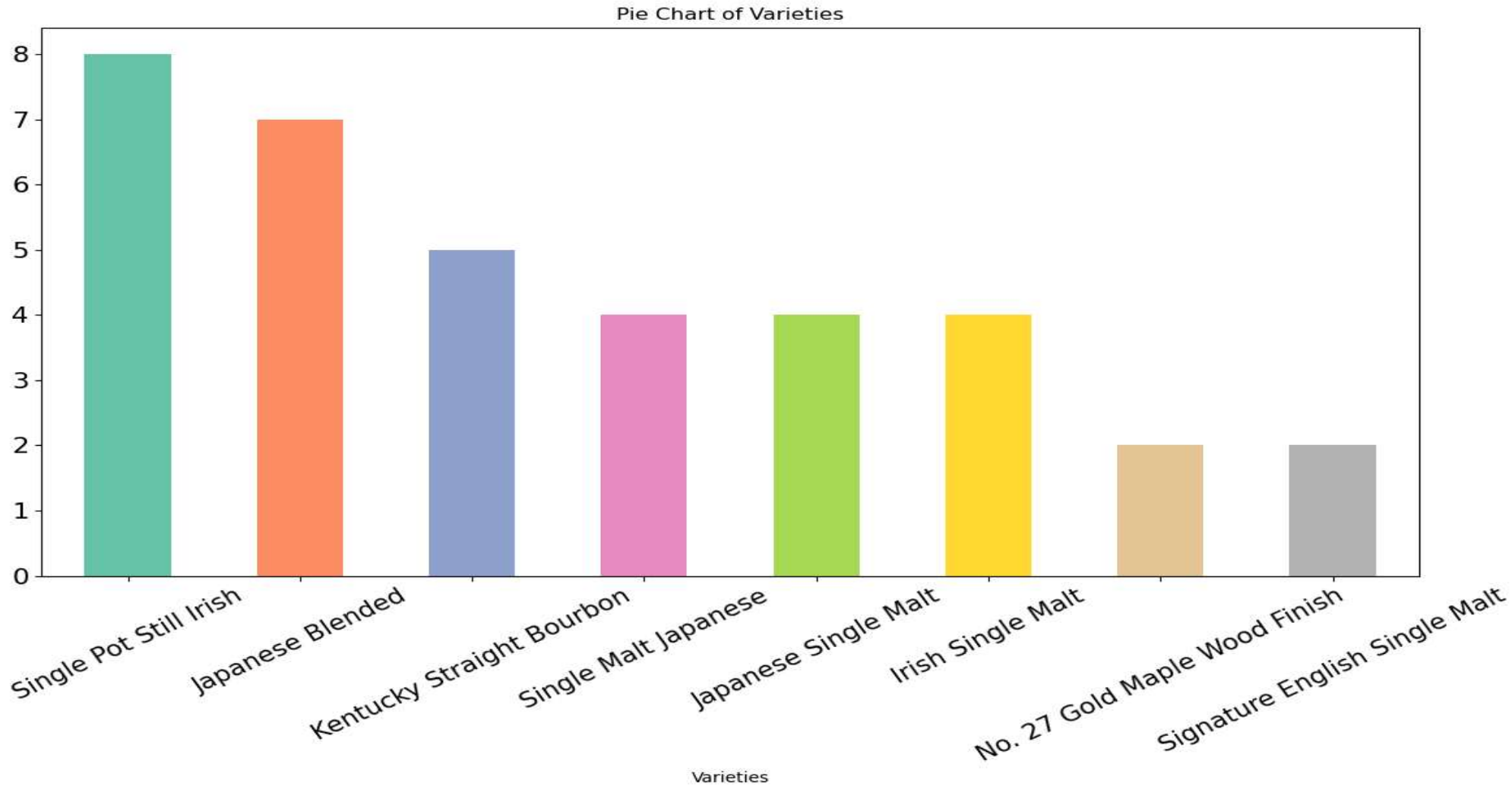
Here ABV refers to (Alcohol By Volume)
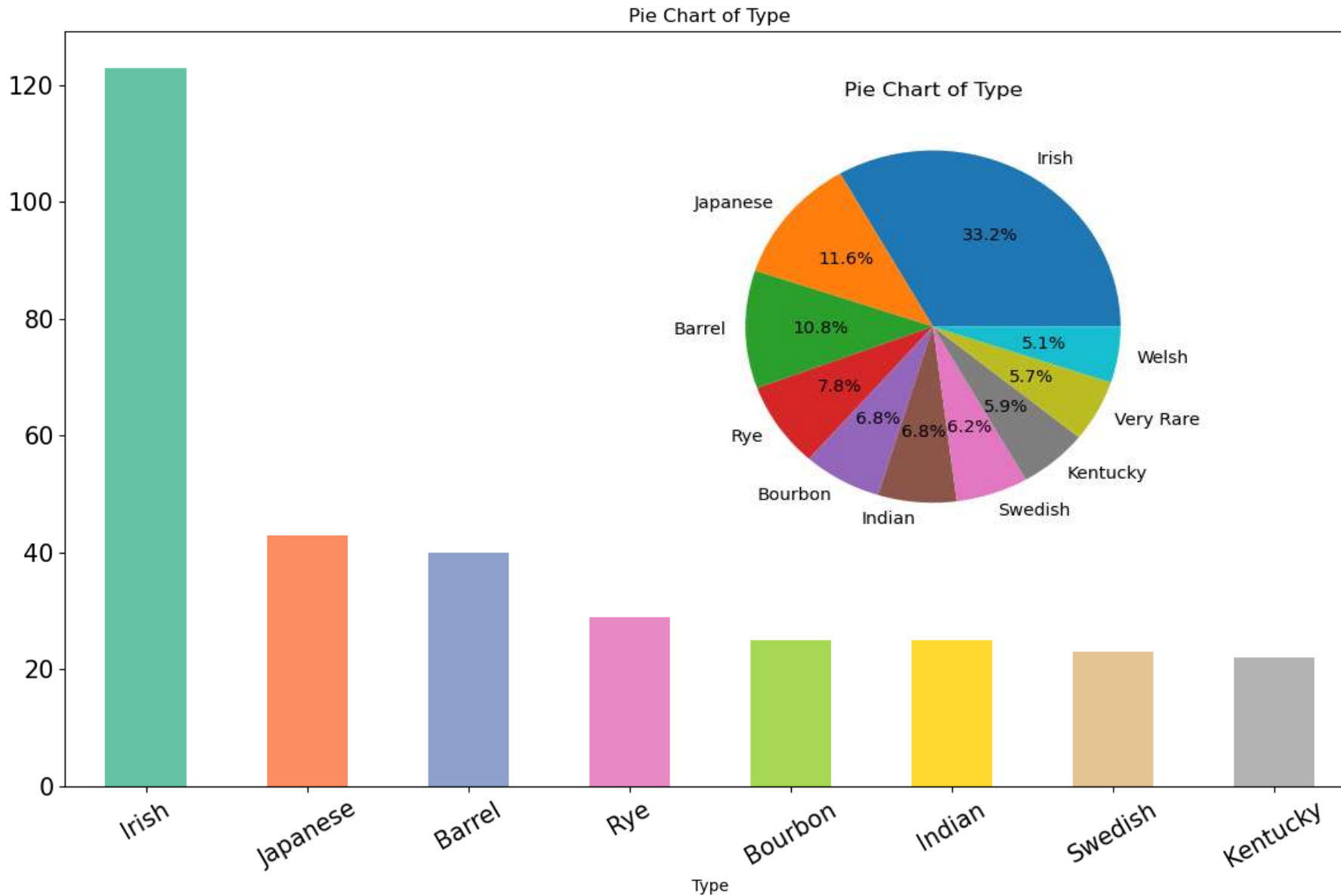
Univariate Graphs On Categorical Columns

Brands

Pie Chart of Brand

Pie Chart of Brand

Observation : Jack Daniel's holds the highest brand count, representing 10.9% of the total.
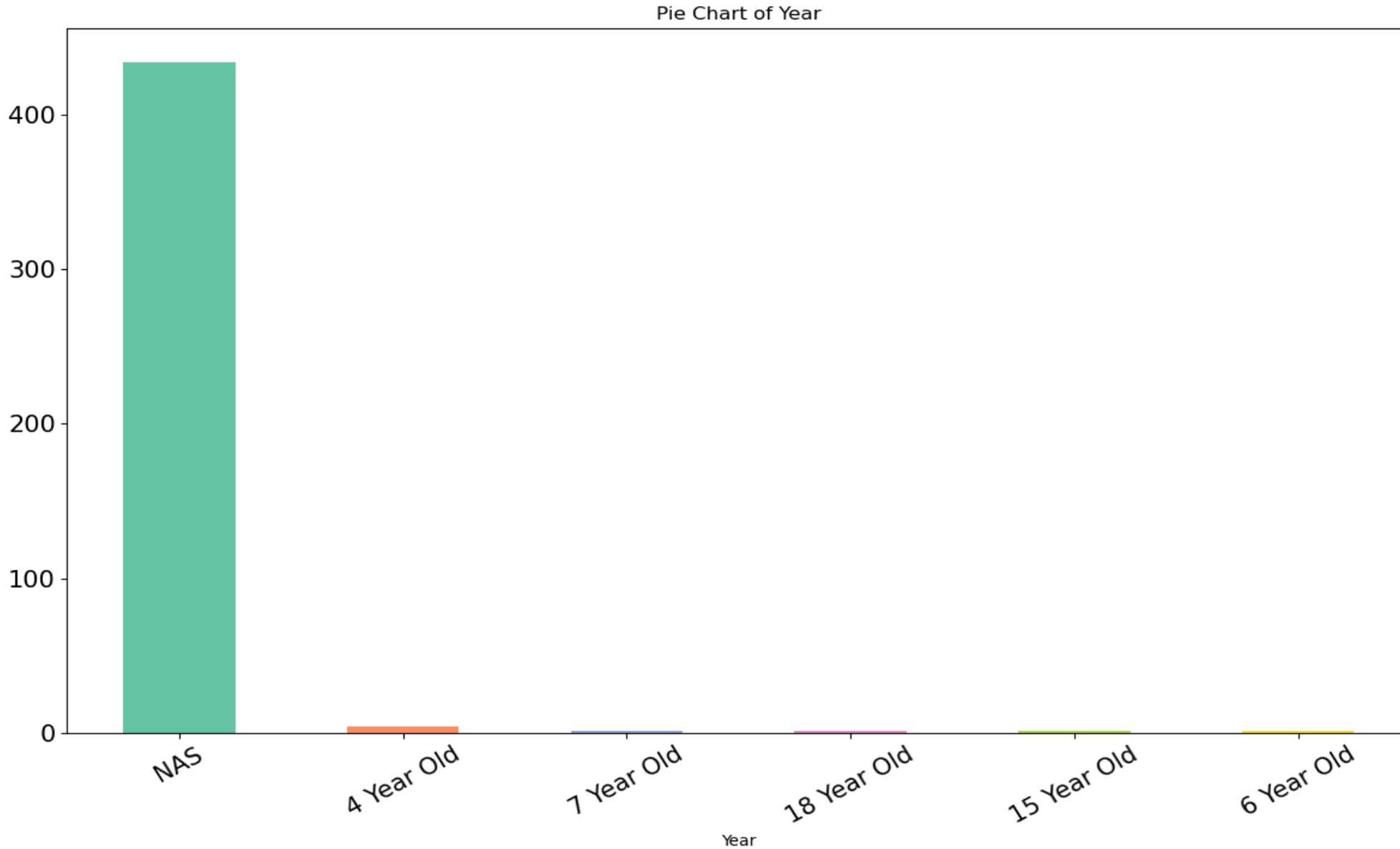
Varieties

Pie Chart of Varieties

Observation : Single Pot Still Irish Whiskey holds the highest share among all varieties at 20%

Type

Pie Chart of Type

Observation : Irish whiskey ranks as the most dominant type with 33.2 %

Year

Pie Chart of Year

Observation : In the 'Year' type , **NAS** holds the highest count with 98.2 %

NAS (NO AGE STATEMENT)

NAS :

- It means the whiskey does not display an age on the label (like "12 years" or "18 years"). But that doesn't mean it's low-quality — it's just a different approach

Why would a distillery release NAS whiskies? ▢

Flexibility in blending :
➢ They can use both younger and older barrels to achieve a desired flavor profile.

Supply and demand :
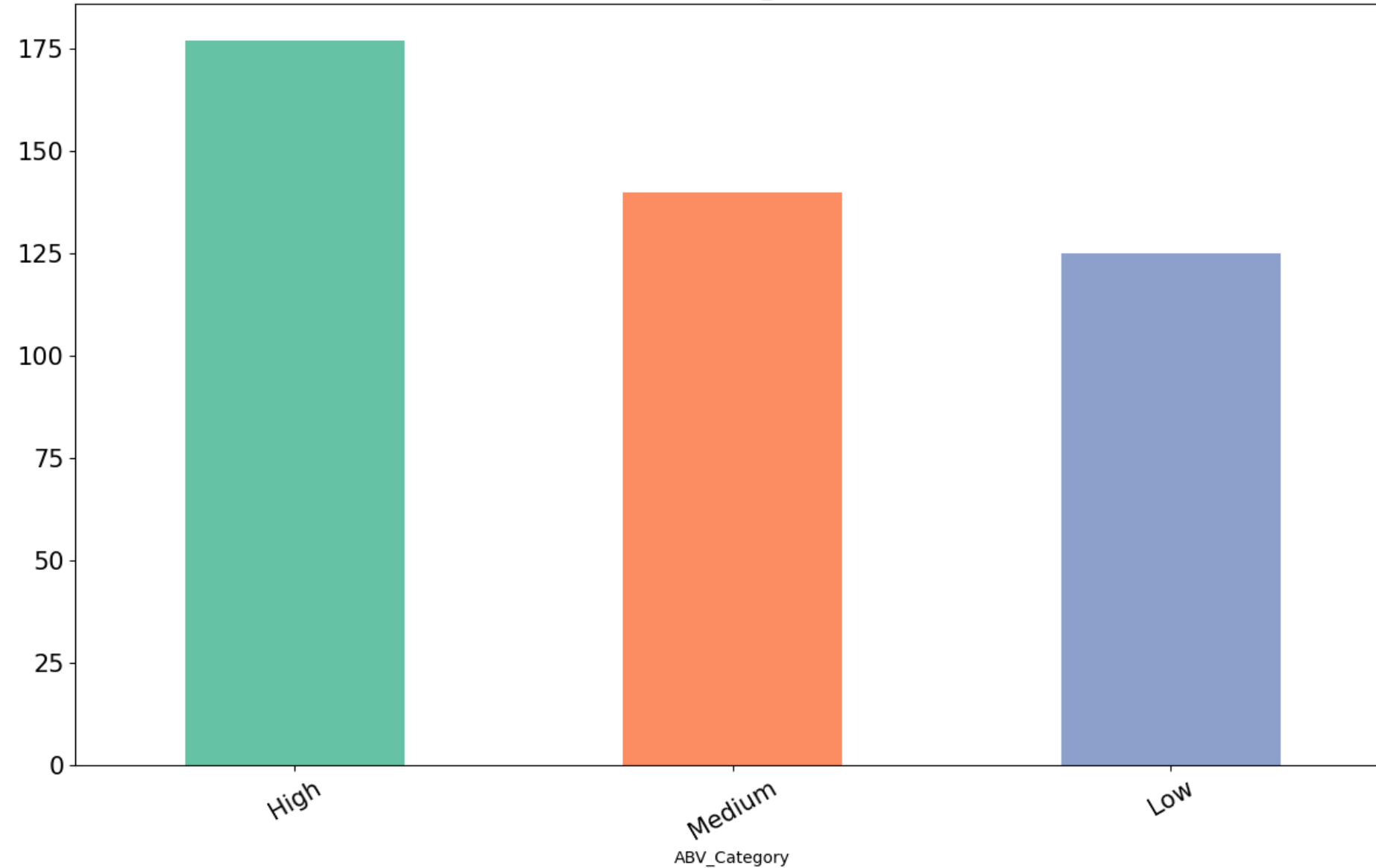➢ Older whiskey stocks are limited. NAS lets distilleries keep up with demand without waiting years for aging.

Focus on taste, not age :
➢ Some whiskies taste amazing even at a young age. NAS bottles invite drinkers to judge by flavor, not number.

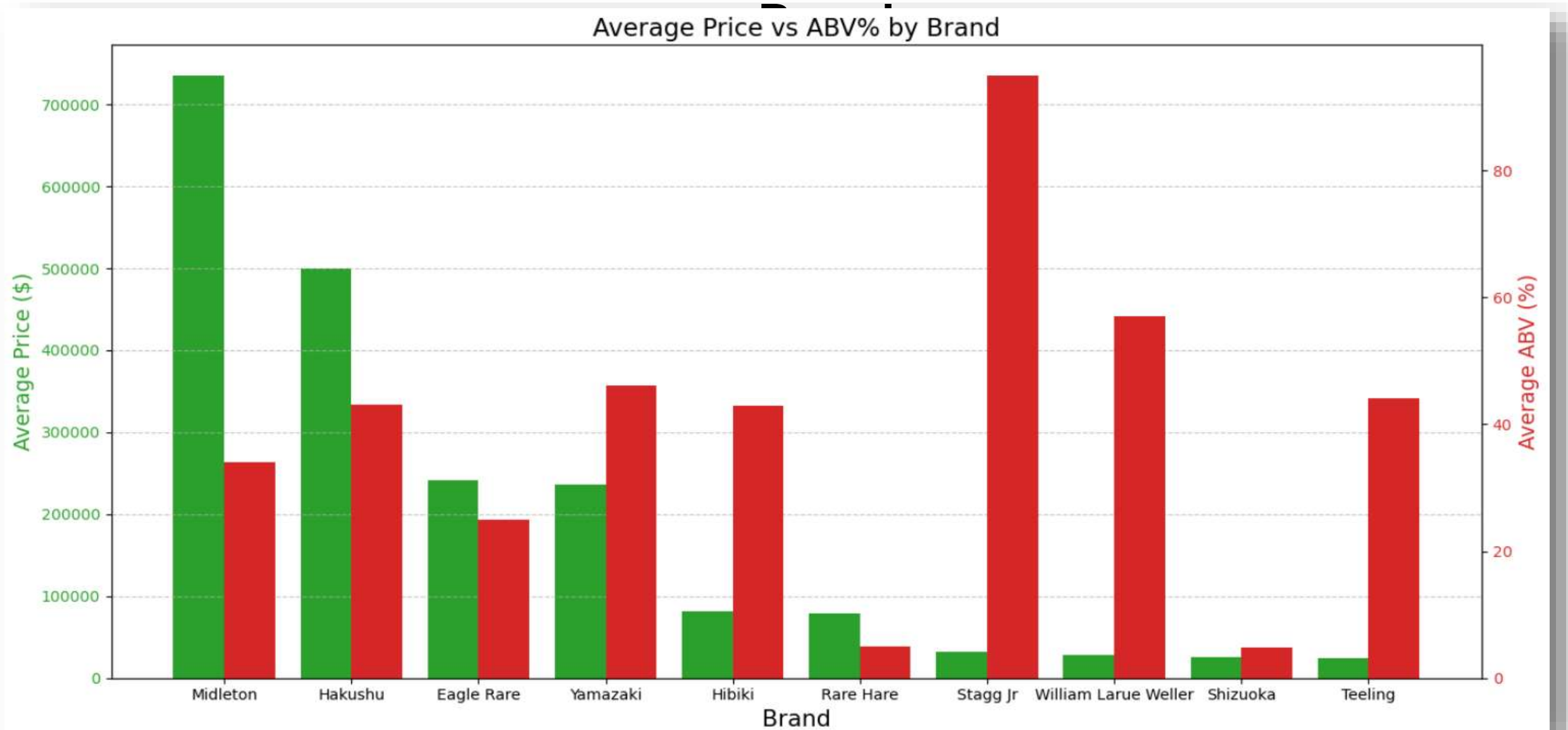# ABV Category

Pie Chart of ABV_Category

LOW : ABV% < 40%

MEDIUM : ABV% between 40% - 47%

HARD : ABV% > 47%

Observation : ABV Category with High% has the highest count

INNOMATICS
RESEARCH LABS

Is it true that higher-priced whiskies also have a higher ABV percentage?

Grouped Bar Chart of Avg Price vs ABV Percentage by

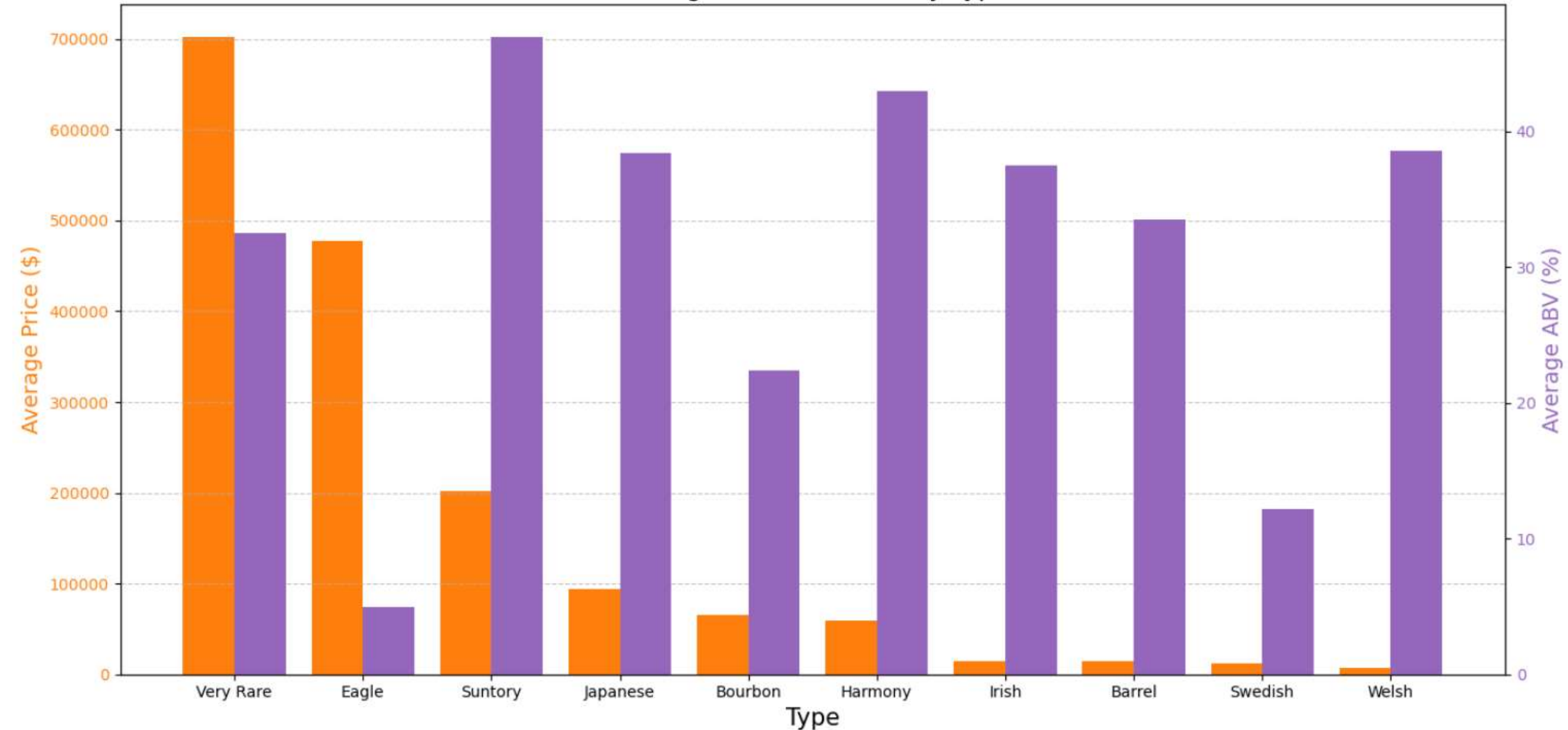**Observation :** Midelton has the highest price where Stagg jr has the highest alcoholic count.

Midleton Very Rare Silent Distillery Chapter 5

Cost : 54,52,794 rs

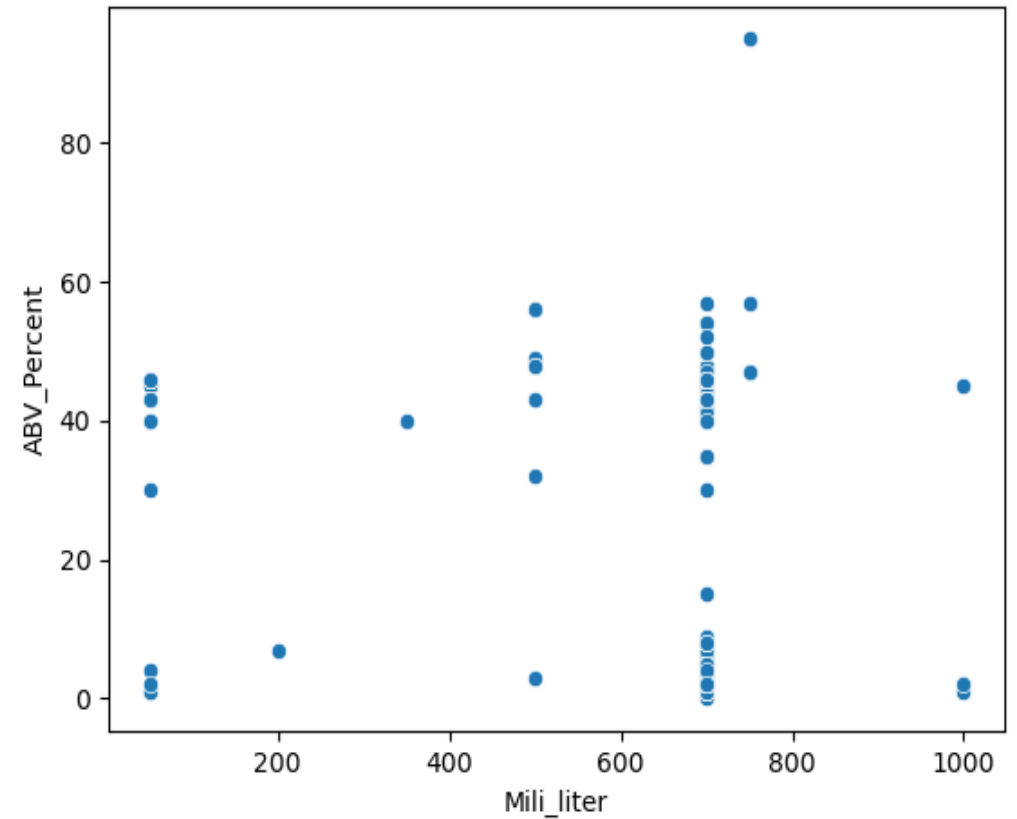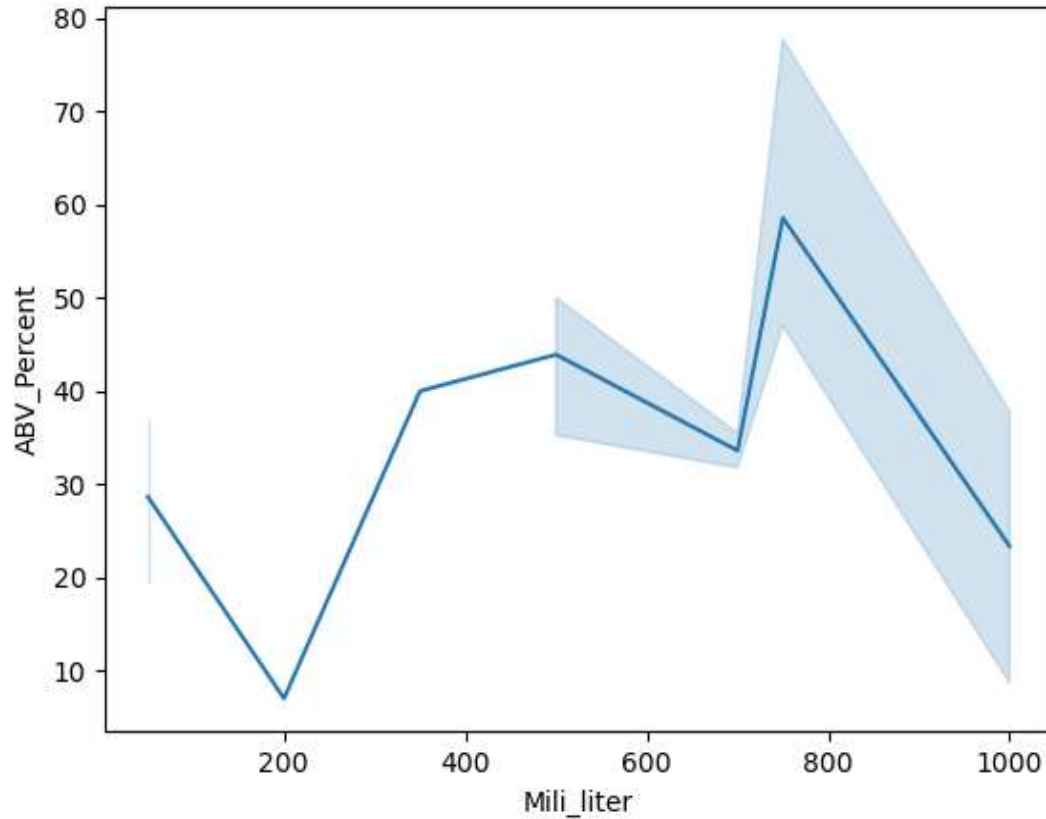# Grouped Bar Plot of Avg Price vs ABV Percentage by

Average Price vs ABV% by Type

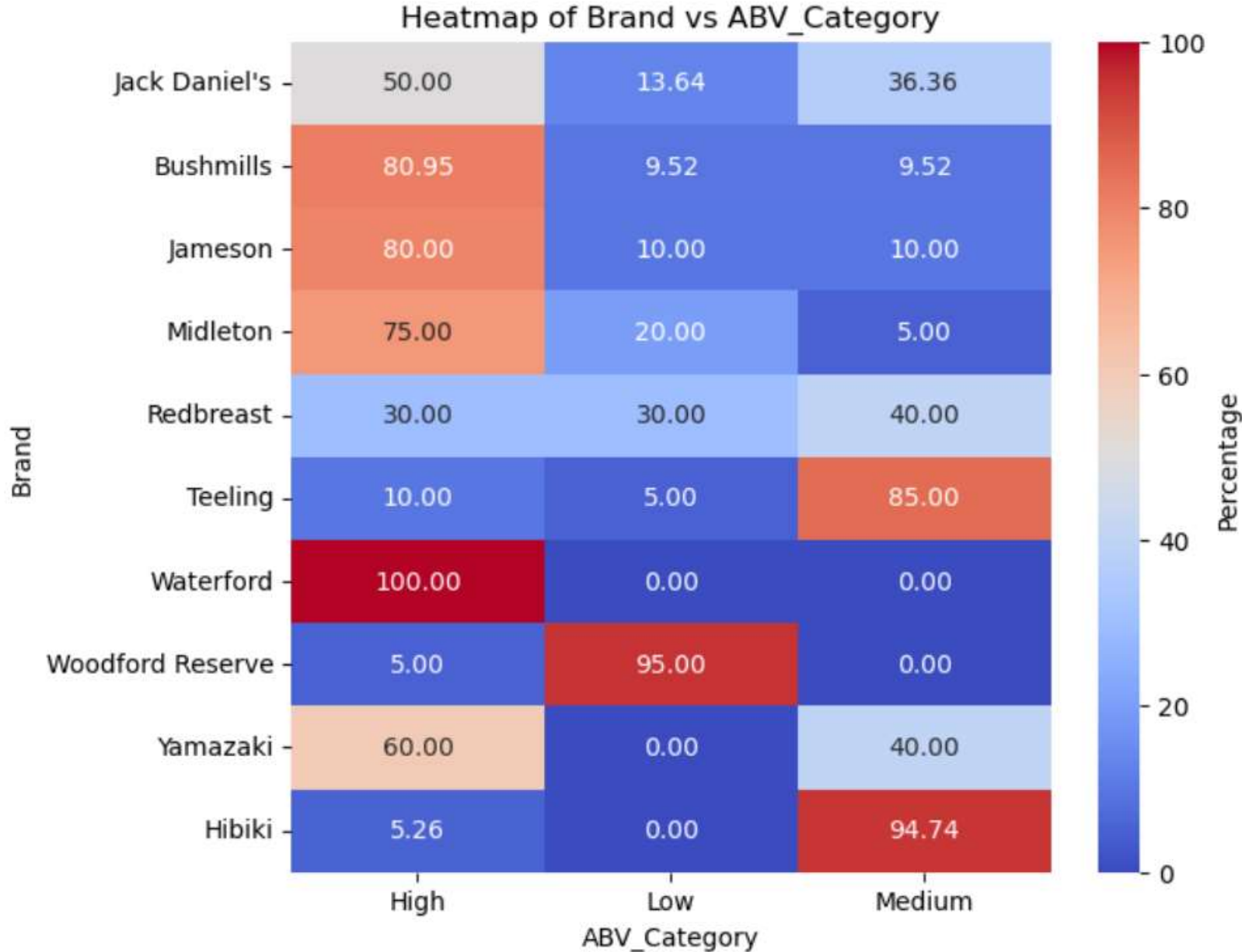**Observation :** Very Rare type has highest price whereas Suntory has highest ABV count

ABV Percent vs Mililiter (Num vs Num)

Observation : Mililiter between 600 to 800 have higher ABV_Percent

INNOMATICS
RESEARCH LABS

Heat Map on Brand vs ABV Category

Heatmap of Brand vs ABV_Category

	High	Low	Medium
Jack Daniel's	50.00	13.64	36.36
Bushmills	80.95	9.52	9.52
Jameson	80.00	10.00	10.00
Midleton	75.00	20.00	5.00
Redbreast	30.00	30.00	40.00
Teeling	10.00	5.00	85.00
Waterford	100.00	0.00	0.00
Woodford Reserve	5.00	95.00	0.00
Yamazaki	60.00	0.00	40.00
Hibiki	5.26	0.00	94.74

Obervation :

➢ Waterford → High
➢ WoodFord Reserve → Low
➢ Hibiki → Medium

INNOMATICS
RESEARCH LABS

Some Indian Brands

INNOMATICS RESEARCH LABS

Indian Brands : Brand vs Price
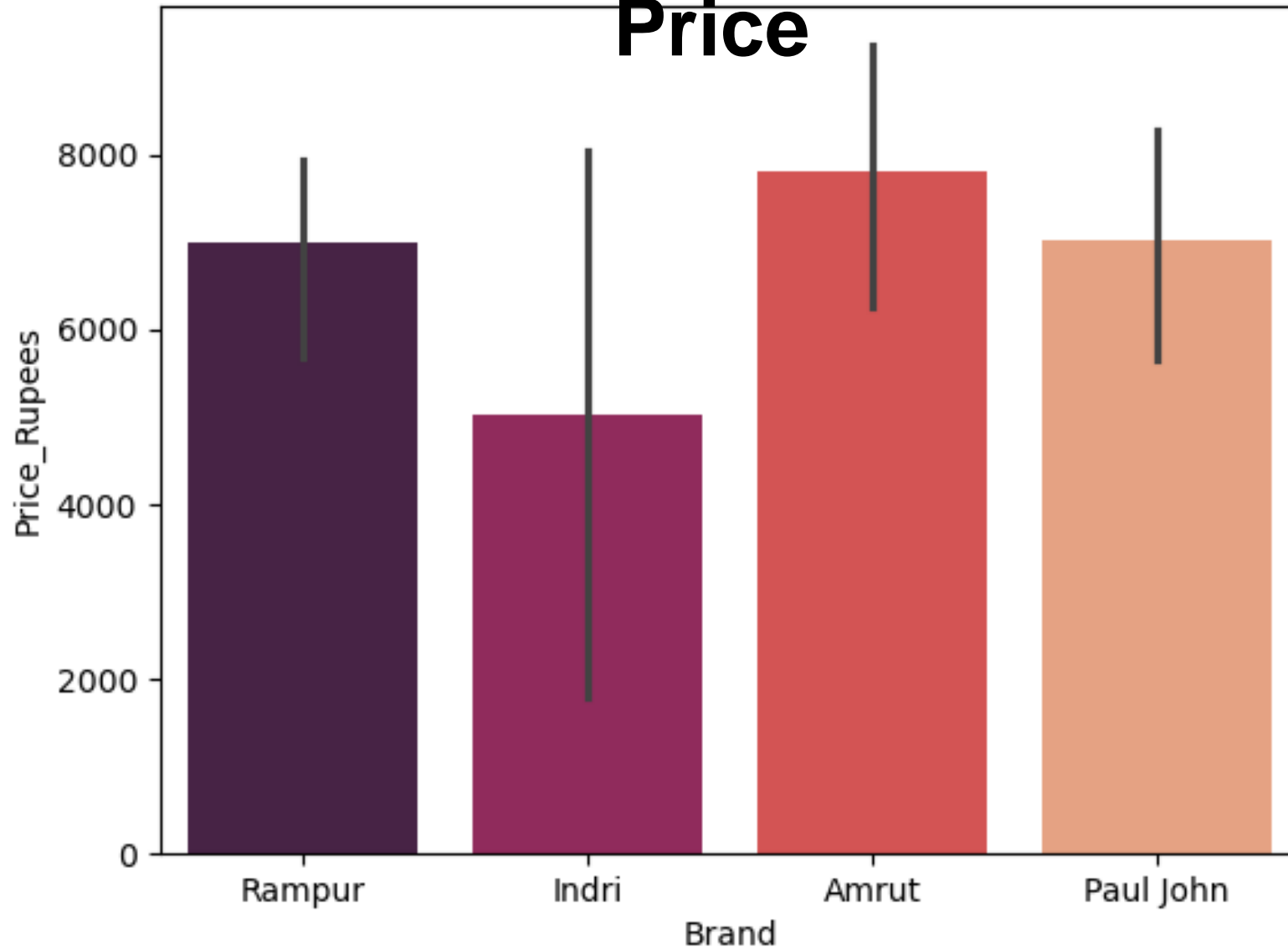
Observation : Amrut has the highest price among all Indian Brands

INNOMATICS RESEARCH LABS

Group Bar Plot of Price and ABV Percentage by Indian Brands

Comparison of Price and ABV Percentage by Brand
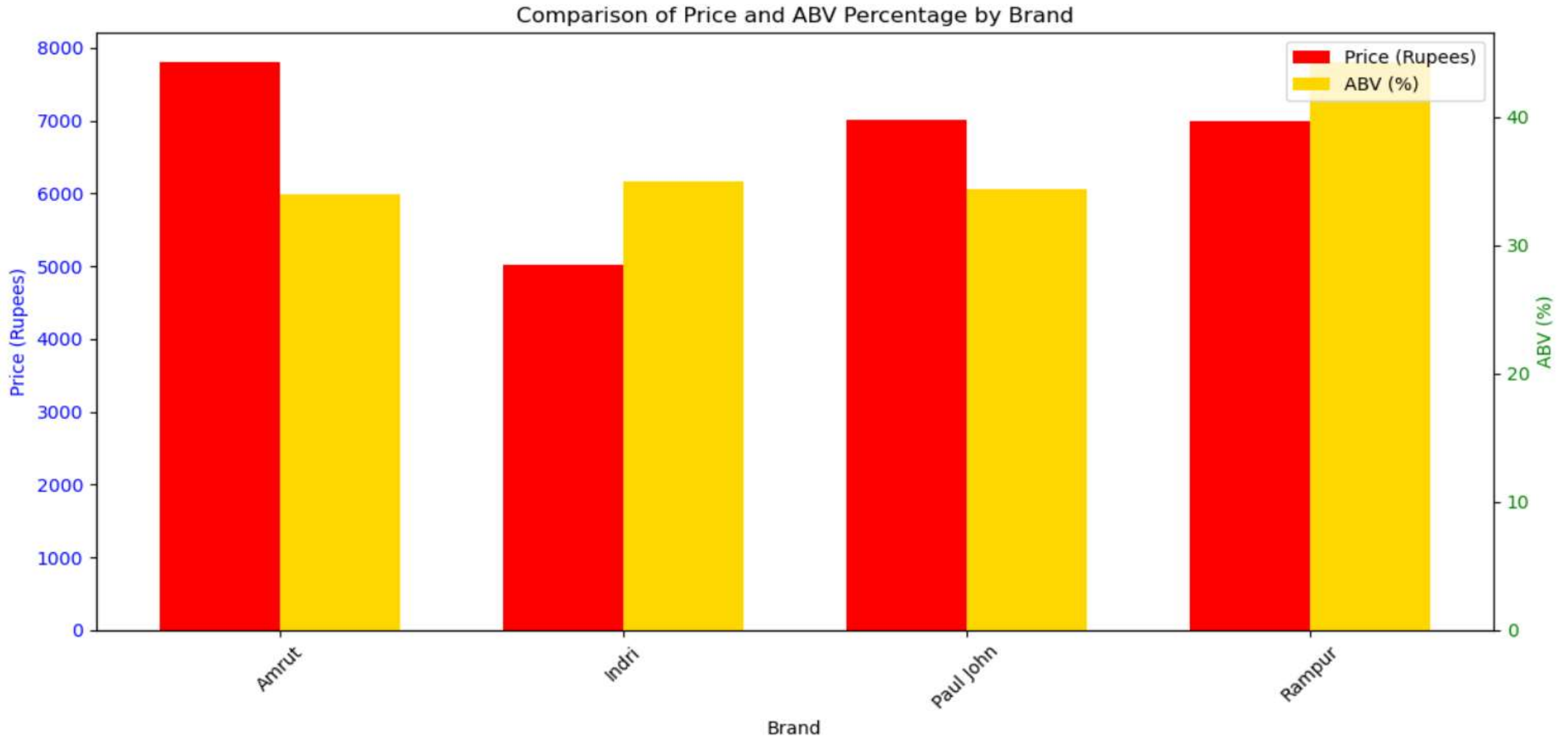
Observation : Amrut has the highest price where Rampur has the highest ABV Percentage

After Some Research I found

❖ **The Rampur Signature Reserve Single Malt Whisky.**

❖ **Cost : 5,00,000 rs .**

❖ **Currently there are 400 Bottles in the world.**

Conclusion

- In this messy whiskey market data Which reveals key trends in **Brands , Price , ABV Percentage , etc** .

- Which empowers smarter investment and sales decision through insights.

Future Enhancement For Growth Opportunities:

- Collect Data From Multiple Sources (Auctions , Retail Shops , etc)

- Analyse social media , blogs and reviews to track consumer sentiment towards towards different whiskey brands and varieties .

- We can build a dashboard using Power Bi or Tableau for users to explore whiskeys

- We can use selenium for better data extraction .

Consumer psychology plays a significant role in the whiskey market, influencing purchasing decisions through factors like brand loyalty, price perception, and emotional connections with brands and their stories.

QUESTIONS ?