**Exploring the Impact of Beer Characteristics and Styles on Overall Review Scores**

### 

### **Research Question:** To what extent do beer characteristics and beer styles influence the overall consumer review ratings of beers?

* ***Abstract***

The main aim of this study was to investigate how different characteristics of beer, such as Maltiness, Spiciness, Hoppiness, Saltiness, Sourness, Sweetness, Body, and Alcohol by Volume (ABV), impact the ratings given by consumers. The analysis provided insights into identifying which beer characteristics are most strongly associated with higher ratings (model 1) Additionally, we explored whether the style of beer using the top 10 reviewed beers had any influence on the overall score. Our findings revealed that certain characteristics such as Maltiness, Spiciness, Hoppiness, Body, and Alcohol by Volume (ABV) have a positive influence on overall beer ratings, while attributes like Sweetness and Sourness are linked to lower ratings.

In terms of beer styles (model 2), Stout, IPA, and Lambic are associated with higher ratings, while Lager tends to result in lower ratings. We believe that this analysis provides value by offering breweries insights into what their consumers prefer. By understanding which flavors and beer styles are most appealing, breweries can enhance their product offerings and reduce production of less desired lines. Focusing on the right balance of sensory attributes and developing beers in popular styles can improve customer satisfaction and lead to higher overall ratings. These insights can help breweries optimize their products for better market performance.

* ***Introduction***

The craft beer market has experienced tremendous growth in recent years, driven by consumers seeking unique, high-quality experiences that go beyond typical mass-market offerings. This project aims to explore the key factors that shape the enjoyment and perceived quality of craft beers, placing a particular emphasis on how different sensory characteristics affect consumer satisfaction. While beer consumption is common, not all beers provide the same level of enjoyment—something our team members have noticed firsthand. This prompted us to investigate what makes certain beers stand out and whether specific attributes can consistently influence consumer ratings.

Our main objective is to understand how various beer characteristics impact customer reviews, which we use as a measure of perceived quality. By doing so, we could provide valuable insights for brewers and marketers looking to improve their product offerings and better meet consumer preferences.

For this analysis, we considered several key attributes, such as Maltiness, Spiciness, Hoppiness, Saltiness, Sourness, Sweetness, Body, and Alcohol by Volume (ABV), to see which ones play a more dominant role in shaping the consumer experience. Additionally, we’re interested in whether changes in individual elements, like the alcohol percentage, have a predictable effect on the overall rating based on it having the highest coefficient and a high t-value.

**The hypothesis (H0) consists of a 1% increase in alcohol content would lead to a 0.1 unit increase in the overall rating of the beer,** assuming that alcohol content has a consistent and measurable impact on how consumers perceive beer quality. The **alternate hypothesis (H1)** suggests that the relationship between alcohol content and overall rating may not be related or linear. After conducting a thorough regression analysis, we found that the 0.1 coefficient was not within the 95% confidence interval for the relationship. This led us to reject the null hypothesis, indicating that **alcohol content alone may not** have a straightforward or predictable effect on consumer ratings.

Our results encourage further exploration into other factors that may play a bigger role in influencing beer ratings. For example, the balance of flavors, the interaction between taste and aroma, or even subtle attributes like mouthfeel and aftertaste might have a stronger impact on how consumers evaluate a craft beer. These insights could pave the way for brewers to experiment with different combinations of attributes to enhance the overall drinking experience. Additionally, marketers can use this information to highlight specific features in their promotions to appeal to target customer segments.

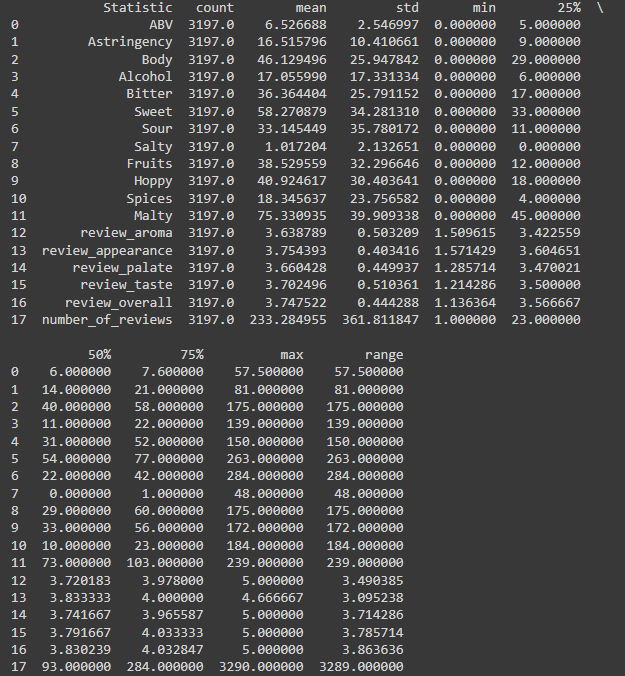
While we set out to find a straightforward link between alcohol content and beer quality, our results show that it’s not as simple as expected. This project sheds light on the need for a more nuanced approach when evaluating what makes a beer truly enjoyable. Understanding the interplay of various characteristics will be crucial for brewers looking to create products that stand out in an increasingly competitive craft beer market.

By pinpointing which attributes matter most, this analysis can **help** **brewers and marketers optimize their strategies and elevate the craft beer experience** for a diverse range of consumers.

* ***Analysis Discussion***

In the descriptive analysis, we explored the characteristics of different beers in a dataset by looking at different attributes such as its alcohol content, flavor profiles, and reviews. The goal was to understand how these factors vary and how they might affect overall beer ratings.

The summary table shown below provides basic statistics for each beer characteristic. For example, the Alcohol by Volume (ABV), which measures the alcohol strength of a beer, ranged from 0% to 57.5%, with an average of about 6.5%. This shows a wide range of alcohol content in beers in the data set.

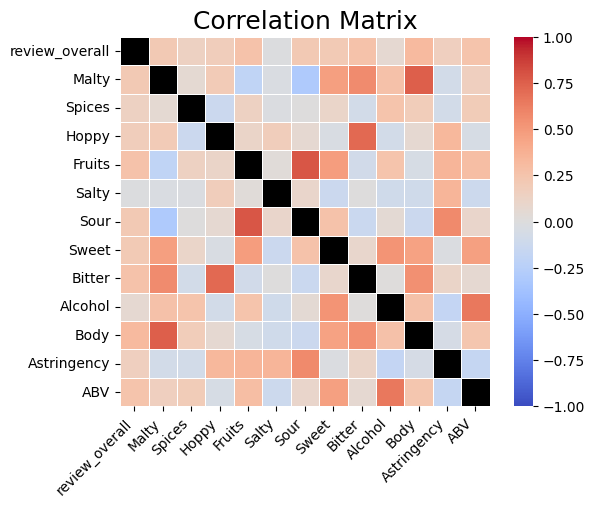


We also looked at flavors, such as sweet and bitter, which had a lot of variation, meaning the beers in the dataset have very different taste profiles. The overall review score had an average of around 3.75, showing that most beers received good, but not exceptional, ratings.

We used a histogram to display how the overall review scores were distributed. Most of the ratings clustered between 3.5 and 4.0, indicating that many beers were rated favorably. This means that while there were a few low or high ratings, most beers fell within this middle range, showing consistency in how consumers rated them.

**Correlation Matrix (Beer Characteristics)**

A **correlation matrix** was used to examine the relationships between different beer characteristics and its overall score of review. Some characteristics, like Malty, Hoppy, and Body, were positively related to higher overall ratings. This means that beers with these characteristics tended to score higher with consumers. On the other hand, characteristics like sweet and sour had a negative relationship with overall reviews, meaning that beers with sweeter or more sour flavors were generally rated lower.

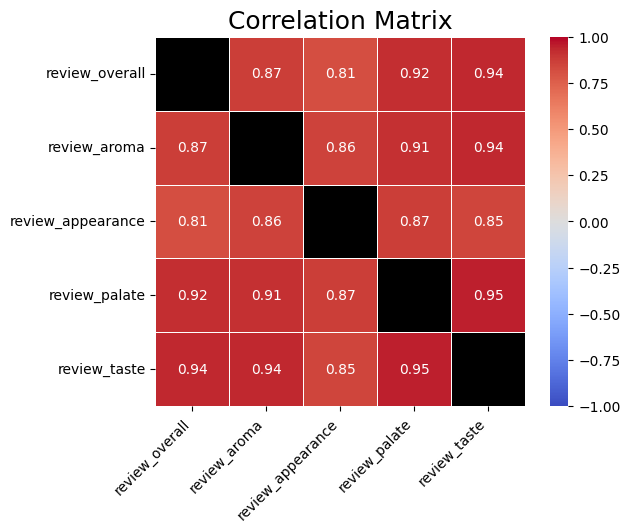


This heatmap shows correlations between various beer characteristics:

* The overall review score has moderate positive correlations with most attributes, particularly with "Malty" and "Alcohol".
* "Fruits" and "Salty" have a strong positive correlation, suggesting fruity beers tend to be saltier.
* "Bitter" and "Sweet" have a weak negative correlation, as expected for contrasting tastes.
* "Alcohol" and "ABV" (Alcohol By Volume) are positively correlated, which is logical.
* "Hoppy" and "Bitter" show a positive correlation, aligning with typical beer flavor profiles.

This heat map reveals complex relationships between beer attributes. Notable correlations include a strong positive link between fruitiness and saltiness, and between hoppiness and bitterness. The overall review score shows moderate positive correlations with most attributes, particularly maltiness and alcohol content. This suggests that these factors significantly influence overall ratings, while also highlighting the multifaceted nature of beer flavors and their interplay in consumer preferences.

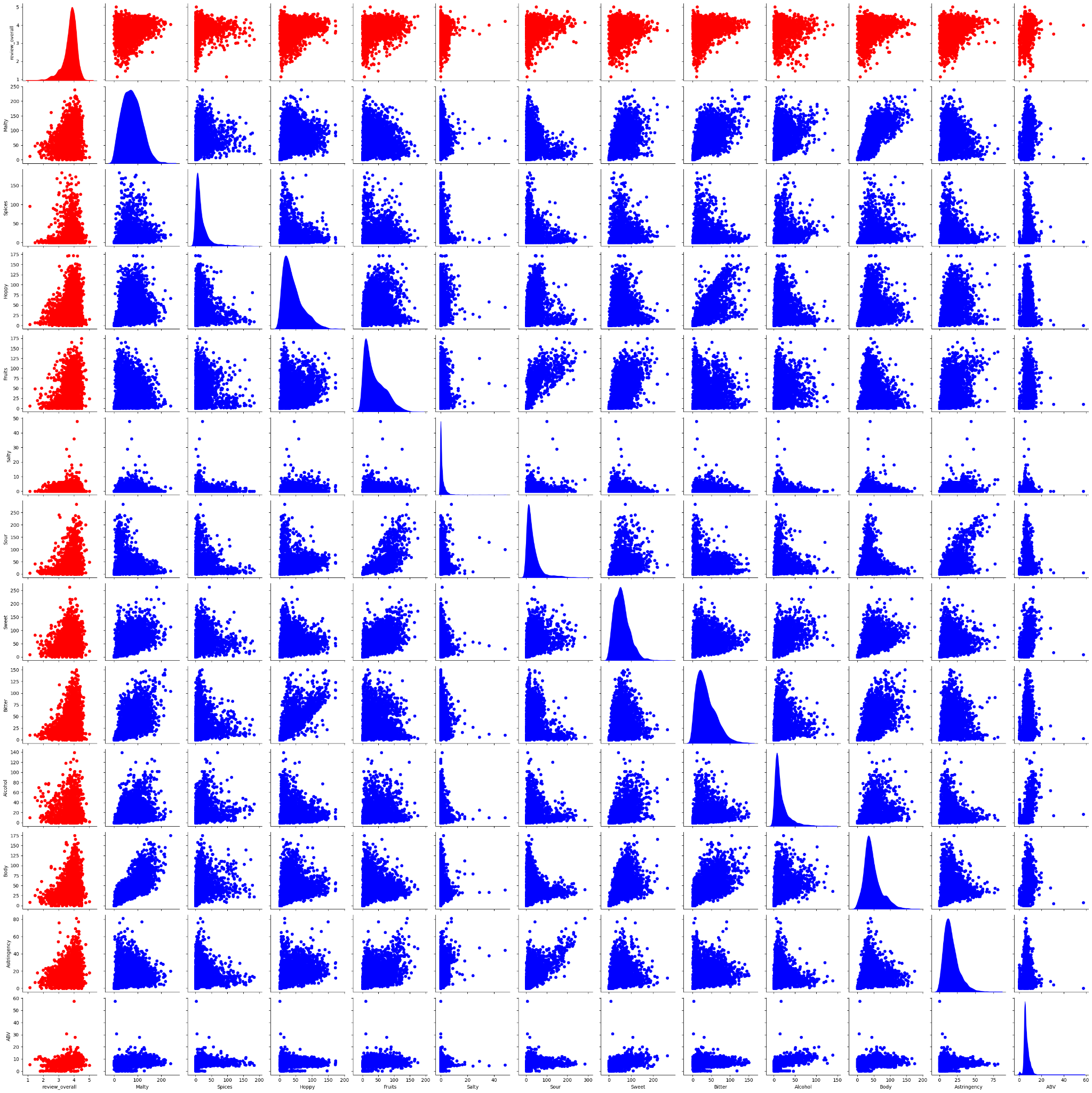
**Correlation Matrix (Review Components)**



The observed correlations in the review components are intriguing. The data indicates that high scores in specific flavor profiles are likely to lead to high overall review scores. All correlations exceed 0.8, demonstrating strong consistency among the various review elements. The most notable finding is the extremely high correlation of 0.95 between "review\_palate" and "review\_taste," emphasizing the importance of taste in shaping overall ratings.

Furthermore, "review\_overall" is most closely correlated with "review\_taste" at 0.94, highlighting the significant influence of taste on overall assessments. Despite slightly lower correlations, "review\_appearance" still remains strong at over 0.8. This high level of consistency among the review components suggests that reviewers tend to evaluate beers holistically, particularly emphasizing the strong relationship between taste perception and overall evaluation.

**Pairplot of Beer Characteristics**



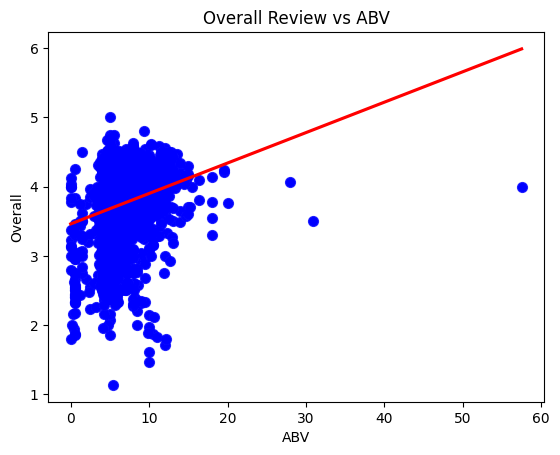
This pairplot shows relationships between multiple beer characteristics.

* Many variables show non-linear relationships, evident from the scatter plots.
* Some variables (e.g., "Sweet", "Sour") have discrete values, visible as distinct lines of points.
* "ABV" shows a roughly normal distribution, while others are more skewed.
* There are some outliers, particularly visible in plots involving "ABV".
* Some pairs (e.g., "Bitter" vs "Malty") show interesting patterns that might warrant further investigation.

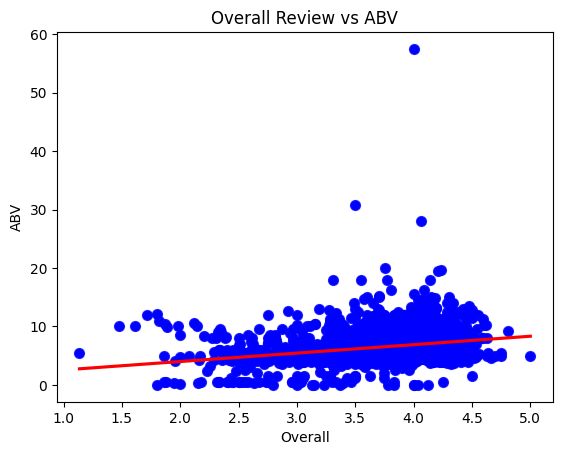
This visualization uncovers non-linear relationships between various beer attributes. Some variables, such as sweetness and sourness, exhibit distinct value patterns. The Alcohol By Volume (ABV) demonstrates a roughly normal distribution, while other attributes display more skewed distributions. The presence of outliers, particularly in ABV, is noticeable. These intricate relationships emphasize the importance of thorough consideration in modeling approaches and interpreting interactions among beer attributes.

**Overall Review vs ABV**

This scatter plot shows the relationship between ABV and overall review scores.



The scatter plot remains the same as the previous one, but the axis has been flipped to provide an alternative perspective while producing consistent results.



**Positive Correlation**: The slight upward slope of the regression line indicates that overall reviews improve as the ABV increases. This suggests that people tend to rate stronger beers slightly higher, possibly due to appreciating the bold flavors and experiences they offer.

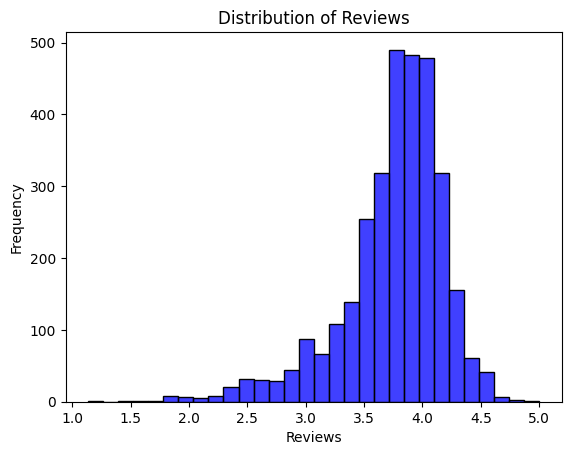
**Cluster at Lower ABV**: Most of the beers in the dataset have a lower ABV range (0-15%), which is in line with the fact that many mainstream beers have a moderate ABV.

**Outliers**: The few beers with very high ABV (30% and higher) have a mix of ratings. Some receive positive reviews while others receive less positive feedback. These high-ABV beers are often niche products, such as imperial stouts or specialty beers, which may appeal to only some reviewers.

Most beers cluster between 4-6% ABV with scores between 2.5-4.5. High-ABV outliers exist but don't necessarily receive higher scores. The plot reveals considerable variability in scores across all ABV levels, suggesting that while alcohol content influences ratings, other factors play significant roles in overall beer appreciation.

**Beer Review Distribution**

This histogram shows the distribution of overall review scores:



* The distribution is left-skewed, with most reviews clustered on the higher end of the scale.
* The peak of the distribution is around 3.75-4.0, suggesting generally positive reviews.
* There are very few reviews below 2.0, indicating that extremely negative reviews are rare.
* The distribution appears to be bimodal, with a smaller peak around 3.5 and a larger peak around 4.0.
* The range of reviews extends from about 1.0 to 5.0, covering the full scale, but with a clear preference for higher scores.

The distribution is roughly bell-shaped but with notable asymmetry. It peaks sharply around the 3.5-4.0 range, indicating that a large proportion of beers receive above-average ratings. There's a long left tail extending towards lower scores, suggesting that while poor ratings exist, they are relatively uncommon. The right tail drops off more steeply after 4.0, implying that perfect or near-perfect scores are also rare.

This distribution shape hints at a nuanced rating behavior where reviewers tend to differentiate more in the middle-to-high range, possibly reflecting the competitive nature of craft beer quality. The concentration of scores in the upper-middle range could also indicate a general satisfaction among beer drinkers or a potential bias towards positivity in the reviewing community.

We also looked at the number of reviews for each beer style. Some styles, like Lager, Stout, and WheatBeer, were much more common in the dataset, meaning these types of beers were more frequently reviewed. This gives an idea of which styles are most popular or widely available to consumers.

#### 

This analysis helps us understand the variety in beer flavors and styles, as well as their impact on how beers are rated. Consumers seem to prefer beers with malty or hoppy flavors, which often receive higher scores. In contrast, sweeter and sour beers are generally rated lower. Additionally, certain beer styles, such as Lager and Stout, are more common in the dataset, possibly reflecting their popularity or availability.

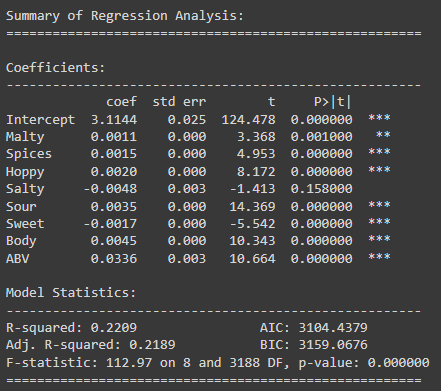
Overall, this descriptive analysis provides a good foundation for understanding what consumers like in their beer, which can be useful for breweries aiming to improve their products. By focusing on the most popular attributes and styles, breweries can create beers that are more likable.

* ***Modeling – Interpretation of Predictors***

In the analysis, two regression models were built to understand how different beer attributes and styles influence the overall review scores. The first model focused on sensory attributes (such as flavors and alcohol content), while the second model looked at beer styles (like Stout, IPA, and Lager) along with ABV (Alcohol by Volume). Below is an interpretation of the predictors and how they relate to the overall score in each model.

#### **Model 1: Characteristic Attributes**

This model assessed how various characteristics of the beer affect its overall rating.



**Malty:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: The positive coefficient for Malty indicates that beers with “malty” flavors tend to receive higher ratings. This suggests that consumers appreciate the richness that malty flavors provide in beer.

**Spices:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: The presence of spices positively influences overall scores, indicating that consumers enjoy the complexity spices add to the flavor profile of a beer.

**Hoppy:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: A positive relationship with overall ratings shows that beers with more hoppy flavors tend to be more appreciated, especially among consumers who favor bitterness or citrusy notes common in IPAs.

**Salty:**

* Statistical Significance: Not statistically significant (p-value > 0.05).
* Practical Importance: Despite being included in the model, salty flavors do not have a notable impact on overall ratings. This suggests that saltiness in beer is either uncommon or does not significantly affect consumer preferences.

**Sour:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: Sour beers tend to receive lower ratings, indicating that this flavor profile is less popular among general consumers. While sour beers may have a niche following, they are not favored by the broader audience.

**Sweet:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: Like sour, sweetness has a negative effect on the overall rating. Beers that are too sweet might be perceived as unbalanced, leading to lower consumer ratings.

**Body:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: A full-bodied beer is positively associated with higher ratings, suggesting that consumers appreciate beers with depth and texture, which contribute to a more satisfying drinking experience.

**ABV (Alcohol by Volume):**

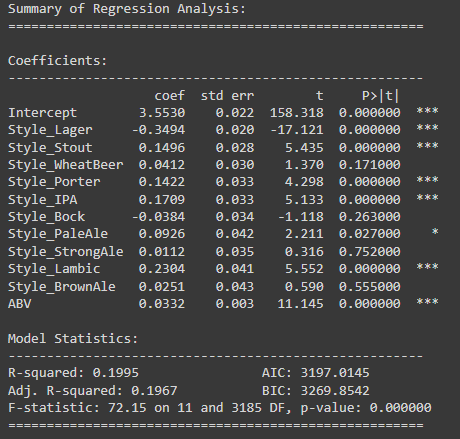
* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: Higher ABV positively affects the overall rating, showing that consumers tend to favor stronger beers. However, the magnitude of the effect is small, suggesting that while alcohol content matters, it is not the dominant factor in determining satisfaction.

#### **Model Fit for Characteristics:**

The model has an R-squared value of 0.2209, indicating that about 22% of the variance in the overall review score is explained by its characteristics. This suggests that while sensory characteristics play a significant role in determining beer ratings, other factors not included in this model also influence consumer preferences.

#### **Model 2: Beer Styles and ABV**

This model focused on how the ten most reviewed beer styles, along with ABV, affect overall reviews.



**Style Lager:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: Lager hurts the overall score. This suggests that while lager is a popular style, it receives lower ratings than styles like Stout or IPA.

**Style Stout:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: Stout positively influences the overall score, meaning consumers rate this style highly. Its robust flavors and body likely contribute to its popularity.

**Style WheatBeer:**

* Statistical Significance: Not statistically significant (p-value > 0.05).
* Practical Importance: WheatBeer does not significantly affect the overall rating. This suggests that while it may be popular, it doesn’t drastically impact consumer ratings, possibly due to its milder flavor profile.

**Style Porter:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: Porter positively affects ratings, indicating that this style is appreciated by consumers, likely due to its rich and malty characteristics.

**Style IPA:**

* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: IPAs are highly rated, likely due to their intense hoppy flavors. This shows that the bitterness and aromatic profiles of IPAs are well-received by consumers.

**Style Bock:**

* Statistical Significance: Not statistically significant (p-value > 0.05).
* Practical Importance: Like WheatBeer, Bock does not significantly affect ratings, indicating that it may not be as strongly preferred as other styles.

**Style PaleAle:**

* Statistical Significance: Marginally significant (p-value ~ 0.05).
* Practical Importance: Pale Ale slightly affects ratings, suggesting moderate consumer preference for this style.

**ABV (Alcohol by Volume):**

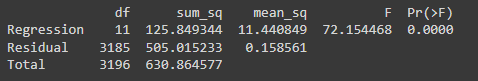
* Statistical Significance: Highly significant (p-value < 0.01).
* Practical Importance: As with the first model, ABV positively influences the score. Higher alcohol content tends to be associated with higher ratings, though the effect remains small.

This model has an R-squared value of 0.1995, meaning about 20% of the variation in overall review scores is explained by beer styles and ABV. While this is a modest fit, it shows that style plays a meaningful role in consumer preferences, although other factors likely contribute as well.

Both models reveal important insights about how different attributes and beer styles influence consumer ratings. Key sensory attributes such as Malty, Hoppy, and Body play a significant role in determining overall scores. Likewise, beer styles like Stout, IPA, and Porter are more likely to receive higher ratings, while Lager tends to have a negative effect. Although the models explain only about 20-22% of the variance in ratings, they highlight specific factors that breweries can focus on to improve their products and align more closely with consumer preferences.

* ***ANOVA Test Conclusion***

The ANOVA test aimed to assess if there are significant differences in overall review scores for beers, considering two factors, beer style and alcohol by volume (ABV). We analyzed these variables to understand their impact on consumer ratings.



#### **Results:**

The results of the ANOVA indicate a significant difference in beer review scores based on beer style and ABV levels. The F-value of 72.15 and a very low p-value of 0.0000 (significantly below the threshold of 0.05) provide strong evidence that the variations in overall review scores are statistically significant. This suggests that beer styles and alcohol content levels directly impact how consumers rate beers.

* Certain beer styles are more highly rated than others. For example, Stouts and IPAs generally receive higher scores, whereas Lagers receive lower ratings. This indicates that consumer preferences are strongly influenced by the characteristics associated with specific beer styles, such as flavor profiles, bitterness, and richness.
* Alcohol Content (ABV): The study shows that ABV strongly predicts overall review scores. Beers with higher ABV levels receive more positive ratings because of their rich flavor profiles. Consumers prefer beers with higher alcohol content, indicating that this factor is crucial to their overall satisfaction.

#### **Conclusion:**

Based on the ANOVA test results, it is confirmed that both the type of beer and its alcohol content significantly impact overall review scores. These findings are important for breweries, indicating that certain beer styles (such as Stouts and IPAs), and alcohol levels within a certain range are more likely to receive positive reviews.

Conversely, less popular styles like Lagers may need to be reconsidered regarding flavor or marketing to better suit consumer preferences. This highlights the significance of understanding and meeting consumer preferences in the brewing industry, ultimately making the work of breweries more meaningful.

Breweries can use this information to improve their offerings by producing beers that match consumer tastes and preferences. For example, they could prioritize beer styles with higher ratings or experiment with alcohol content to enhance consumer satisfaction and receive more positive reviews.

* ***Individual Significance Test Conclusion***

The individual significance test allowed us to assess the contribution of each predictor in the model to explain the overall review scores. By examining the p-values, we determined which predictors have a significant impact on the dependent variable, the overall score. Significant predictors have a p-value less than 0.05, which showed that their effect on the outcome was not because of random chance. Below is the interpretation of the key predictors and their relevance to the model.

**Malty:**

* Statistical Significance: With a p-value well below 0.01, Malty is a highly significant predictor in the model.
* Practical Importance: The positive coefficient for Malty indicates that beers with stronger malty flavors tend to receive higher ratings from consumers. This suggests that malty beers are generally well-liked and contribute positively to the overall score, making it a key factor for breweries aiming to improve product ratings.

**Spices:**

* Statistical Significance: Spices are also highly significant (p < 0.01).
* Practical Importance: The presence of spices is positively associated with overall review scores. Consumers appear to enjoy the complexity spices add to the beer, contributing to higher ratings. This highlights the importance of flavor complexity in determining consumer preferences.

**Hoppy:**

* Statistical Significance: Hoppy is a significant predictor (p < 0.01).
* Practical Importance: Beers with hoppy flavors, particularly common in IPAs, tend to receive higher scores. This reflects consumer preference for the bitterness and aromatic qualities typically found in hoppy beers. The positive effect of Hoppy suggests that it is a critical attribute in driving consumer satisfaction.

**Sweet:**

* Statistical Significance: Sweet is significant but has a negative effect (p < 0.01).
* Practical Importance: The negative coefficient for Sweet indicates that beers with stronger sweet flavors are less likely to receive high ratings. This suggests that consumers may find excessive sweetness unappealing, potentially impacting the balance of the beer. Breweries should be mindful of this when designing sweeter beer varieties.

**Sour:**

* Statistical Significance: Sour is a significant predictor with a negative effect (p < 0.01).
* Practical Importance: Sour beers generally receive lower ratings, indicating that this flavor profile may not appeal to the majority of consumers. While sour beers may have a niche market, the negative impact on overall scores suggests limited appeal among general consumers.

**Body:**

* Statistical Significance: Body is highly significant (p < 0.01).
* Practical Importance: The positive effect of Body suggests that beers with a fuller, richer texture tend to be rated higher. Consumers seem to value the depth and complexity that a fuller body brings, making it an important factor for breweries to consider when crafting their products.

**ABV (Alcohol by Volume):**

* Statistical Significance: ABV is a significant predictor (p < 0.01).
* Practical Importance: A higher ABV is associated with slightly higher overall ratings. While the effect size is relatively small, the positive relationship indicates that stronger beers are generally preferred by consumers, though it may not be the most critical factor compared to flavor attributes.

**Relevance to the Overall Model:**

**Statistically Significant Predictors:** Malty, Spices, Hoppy, Sour, Sweet, Body, and ABV substantially affect overall scores, supported by their low p-values (< 0.05).

**Not Statistically Significant:** Salty does not exhibit a statistically significant effect, suggesting it may not significantly contribute to determining overall scores.

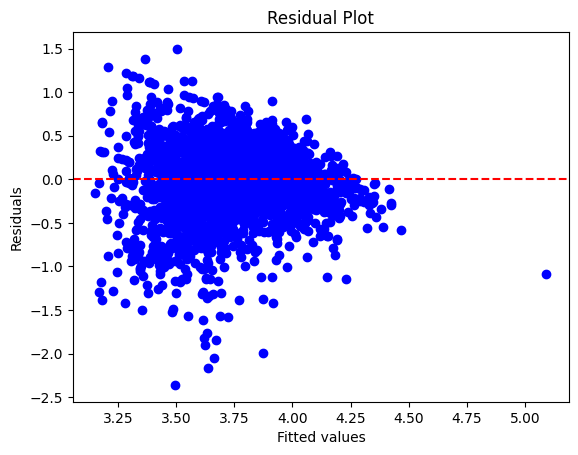
**Direction of Effects:** Positive coefficients indicate that increases in sensory attributes (Malty, Spices, Hoppy, Sour, Body, and ABV) correspond to higher overall scores, while Sweet and Salty yield adverse effects, signifying a reduction in the overall score. **However, we reject the Null Hypothesis (H0) due to the estimated effect of ABV (0.0336)** significantly differing from 0.1. This denotes compelling evidence that the impact of ABV on the overall score is not equivalent to 0.1; instead, it approximates 0.0336 per 1% increase in ABV.

* ***Diagnostics Discussion***

In regression analysis, it is critical to verify that certain assumptions about the model are met. These assumptions include the behavior of the residuals, the normality of errors, the absence of multicollinearity, and the linearity of relationships between the predictors and the dependent variable. In this assignment, we evaluated these assumptions through various diagnostic plots and statistical tests, all of which are essential for ensuring the reliability and validity of the regression results.

**Residual Plot (Model 1)**:

This plot shows the residuals (the difference between actual and predicted values) on the vertical axis and the fitted values (predicted by the model) on the horizontal axis.



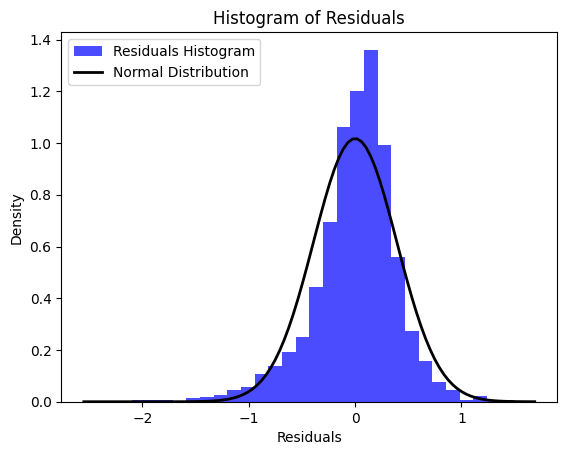
#### **Key Insights:**

The residuals mostly center around zero, indicating that the model has captured the main patterns in the data. However, there is a slight imbalance and wider spread as the predicted values increase, especially for higher beer scores. This suggests potential heteroscedasticity, where the variation of the residuals grows as the predicted values increase. This indicates that the model might underestimate or overestimate scores for beers with more extreme sensory profiles or alcohol content.

There is no clear non-random pattern, such as curvature, which suggests that the linear model is suitable for capturing the relationship between the predictors and the response variable. However, the increasing variance at higher fitted values may indicate the need for further model refinement, possibly through transformations or alternative modeling techniques.

#### **Histogram of Residuals (Model 1):**

The histogram of residuals displayed below assesses the normal distribution of residuals, an assumption of linear regression. The superimposed normal distribution curve serves as a visual reference for comparison.



#### 

#### **Key Insights:**

The histogram of residuals displayed below assesses the normal distribution of residuals, an assumption of linear regression. The superimposed normal distribution curve serves as a visual reference for comparison.

Minor Deviations: Although the deviations from normality are not significant, they should be noted. The slight skewness and the heavier left tail may suggest that certain predictors, especially those related to flavor profiles or ABV, may have non-linear relationships with overall ratings/scores. It is also possible that interactions between the predictors are not fully considered in the current model.

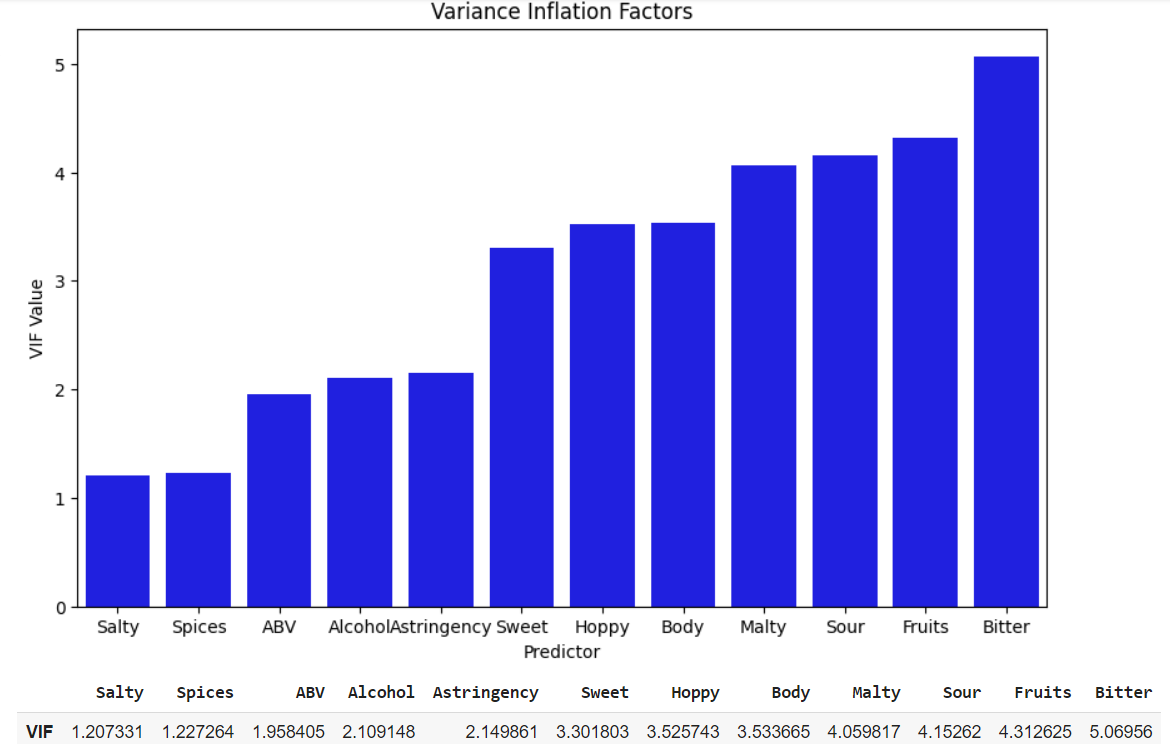
Overall, the diagnostic plots suggest that the regression model performs reasonably well in predicting overall beer ratings based on its characteristics. However, the presence of **heteroscedasticity** and slight **non-normality of residuals** indicates that further refinements could improve model accuracy.

**Variance Inflation Factors (VIF)**

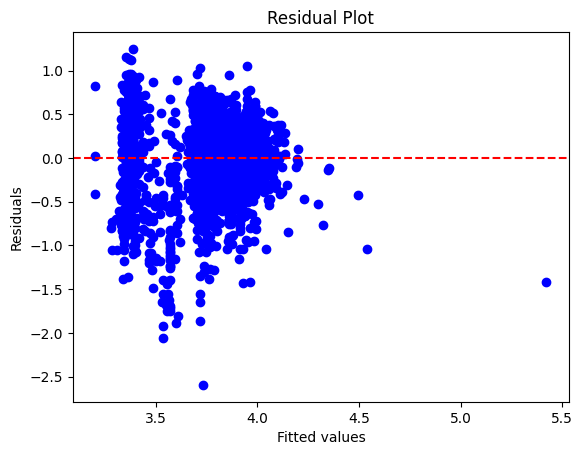
Our VIF bar chart indicates potential multicollinearity issues in the dataset. Most predictors have VIF values above 2, with "Bitter" showing the highest (>5). This suggests that some variables, especially bitterness, might be strongly correlated with others.

Variables like "Salty" and "Spices" show lower VIFs, indicating they're more independent. This analysis is crucial for understanding potential issues in predictive modeling and interpretation of individual variable effects.

* Most predictors have VIF values above 2, indicating some multicollinearity.
* "Bitter" has the highest VIF (>5), suggesting it might be strongly correlated with other predictors(Most likely with the hoppy variable as in the beer industry hops add bitterness to the sample).
* "Salty" and "Spices" have the lowest VIFs, close to 1, indicating they're relatively independent.
* Predictors with high VIFs (e.g., Bitter, Fruits, Sour) might need careful consideration in modeling to avoid overfitting.



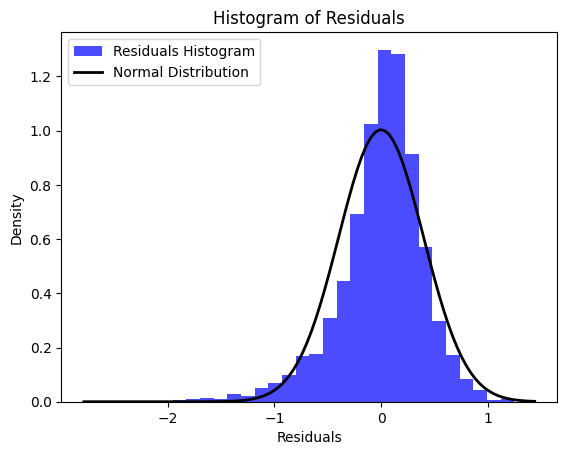
**Residual Plot (Model 2)**:



**Key Insights:**

These patterns could indicate issues such as non-linearity or heteroscedasticity. Ideally, the residuals should be randomly scattered around zero without any clear pattern. This implies that the model adequately captures the underlying relationships.

The residual plot also indicates a spread that is somewhat cone-shaped, which might suggest the presence of heteroscedasticity. This means that the variance of the residuals increases with the fitted values. It implies that the model might not fully account for the variability in the data for higher fitted values



The histogram of residuals reveals slight deviations from normality. While the residuals generally follow a normal distribution, as indicated by the superimposed normal curve, the slight left skewness and the presence of outliers on both tails suggest potential issues with model specification or unaccounted factors in the data.

In general, the diagnostic plots indicate that the regression model does a decent job of explaining the dependent variable, as most residuals are clustered around zero. However, the presence of heteroscedasticity and minor non-normality of residuals suggests areas for potential improvement in the model.

***Conclusion***

#### **Key Takeaways:**

This analysis aimed to understand how different characteristics of beer impact its review scores. Regression analysis, ANOVA tests, and diagnostic tests provided several important insights.

Remember the following information:

* Significant predictors: Key characteristics such as "Malty," "Spices," "Hoppy," "Body," and "ABV" were found to have a positive influence on overall review scores. Conversely, "Sweet" and "Sour" flavors negatively impacted scores, suggesting that consumers may prefer less sweet and sour beers.
* The impact of beer styles on ratings is significant. Stout, IPA, and Porter tend to receive higher ratings, while Lager negatively impacts scores. This indicates that consumers have distinct preferences for specific beer styles, and breweries should consider this when creating new products.
* Beers with higher alcohol content (ABV) tend to receive slightly higher scores, although the impact is relatively small. Consumers generally prefer beers with moderate to high alcohol content, but it's important to note that ABV alone is not the most significant factor influencing overall satisfaction.
* Although the alcohol by volume (ABV) showed a positive trend in predicting the overall score of the beer, we rejected the null hypothesis (Ho), which stated that for every 1% increase in alcohol, there is an additional 0.1 unit rating increase in the overall rating. Since 0.1 is not within the 95% confidence interval for B8, **we reject the null hypothesis and assume the alternative hypothesis (H1) that B8 does not equal 0.1.**

**Recommendation:**

Breweries can take specific strategic actions to improve their products and meet consumer preferences. Success in brewing involves finding the perfect balance between various sensory components. Here are some key actions breweries can take:

* **Master the Balance of Malty and Hoppy Flavors:** Beers with strong Malty and Hoppy profiles tend to receive higher ratings. However, balancing these flavors to create an enjoyable drinking experience is essential. Overemphasizing one flavor at the expense of the other can alienate certain consumer segments. By refining this balance, breweries can cater to a broader audience while maintaining a distinctive taste.
* Another important **strategic action for breweries is to balance sweetness and sourness** in their beers. Overpowering sweet and sour flavors can lower review scores. Breweries should aim to add complexity without overwhelming the palate. This approach can transform a potentially polarizing beer into one that appeals to a wider array of consumers.
* **Optimizing and evolving beer styles** is a crucial strategy for breweries. Stouts, IPAs, and Porters tend to receive higher ratings, suggesting that breweries should prioritize these styles. However, it's crucial to continually refine these styles to elevate them above the competition. Additionally, breweries might need to reimagine their approach to brewing Lagers by introducing subtle twists in flavor profiles to reinvigorate consumer interest.
* **Subtle Use of Alcohol by Volume (ABV)**: While ABV showed a positive correlation with overall beer ratings, it is important to note that the effect is nuanced and less significant than initially hypothesized. Instead of relying heavily on increasing alcohol content to boost ratings, breweries should focus on balancing ABV with other elements like flavor complexity and mouthfeel. Experimenting carefully with ABV is crucial to ensure that it complements, rather than overwhelms, the beer's other characteristics.

#### 

#### **Future Research Directions:**

Future research could focus on expanding the analysis in the following areas:

* **Consumer Demographics:** Demographic information such as age, gender, and location would be beneficial to include to gain more insights into specific consumer preferences. Different groups may have different tastes, which could help breweries target their products more effectively.
* **Seasonal Variations:** It would also be valuable to examine how preferences change seasonally. For example, certain beer styles like stouts may be more prevalent in the winter, while lighter beers such as lagers or wheat beers might be favored in the summer.
* **Longitudinal Studies:** A longer-term study could analyze changes in consumer preferences over time, especially as new beer trends emerge in the market.

**Limitations:**

While this analysis provides valuable insights, several limitations should be noted:

* **Heteroscedasticity:** The residuals plot indicated some heteroscedasticity, meaning that the variance of residuals was not constant across the fitted values. This could affect the efficiency of the regression estimates. Future models could address this by applying robust standard errors or transforming the dependent variable.
* **Limited Predictor Variables:** The analysis mainly focused on sensory attributes and beer styles. However, other factors, such as brand loyalty, packaging, and price, were not included. These factors could also play a significant role in consumer ratings.

#### **Final Thoughts:**

The analysis provides valuable insights to help breweries optimize their product offerings according to consumer preferences. By concentrating on popular flavor profiles and beer styles, breweries can **enhance** their likelihood of **receiving higher ratings** and **increasing customer satisfaction**. Future research, incorporating additional variables and refined models, can expand on these findings to offer even more detailed guidance for product development.