HALLOWEEN CHALLENGE

Top 3 Candies

#mavenhalloweenchallenge

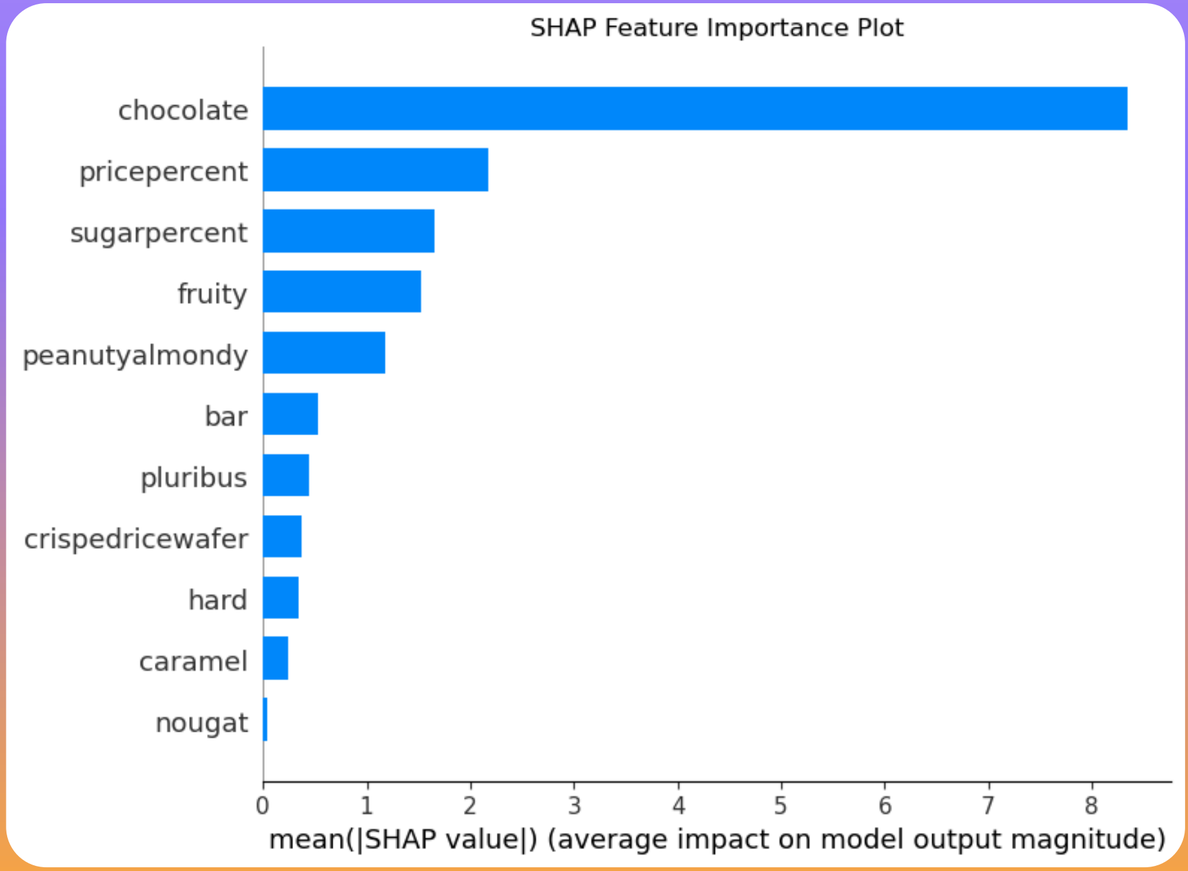
HOW ARE TOP-3 SELECTED ?

- The **Top 3 Candies** are selected by combining feature-based predictions (***SHAP Values***) and actual performance (***winpercent***).

- **Key Metrics:**
  - **Composite Score** (based on feature importance)
  - **winpercent** (actual popularity score)

FEATURE IMPORTANCE USING SHAP VALUES

- **SHAP *(SHapley Additive exPlanations)*** values are used to determine the importance of each feature. SHAP values show how much each feature (e.g., c*hocolate, fruity, sugarpercent*) contributes to predicting ***winpercent***.



- ***chocolate, pricepercent, and sugarpercent*** are the most important factors contributing to predicting *winpercent*.

- A **Random Forest Regressor** is used for determining the SHAP Importance values

COMPOSITE SCORE CALCULATION

- The **Composite Score** is calculated by the following formula :

$$\text{Composite Score} = \sum(\text{Feature Value} \times \text{Feature Weight})$$

- Each feature value is multiplied by its SHAP importance weight.

- For each candy, the **feature values** (such as chocolate or sugarpercent) were multiplied by their corresponding **feature weights**. This step ensured that more influential features had a proportionally larger effect on the overall score.

COMBINING COMPOSITE SCORE AND WINPERCENT

- Combining the Composite Score and Win-Percent into one **Final Score.**

- The win-percent is an observed or actual performance metric for each candy. It's an essential component because it directly shows how popular or well-liked the candy is. However, instead of relying on this alone, we combine it with the composite score to account for both the observed popularity and the theoretical score based on features.

- Weights are assigned to the composite score and the win-percent based on their perceived importance. In this case a weight of **0.6** for the *winpercent* and **0.4** for *Composite Score* is considered because ***winpercent*** is given more importance compared to ***Composite Score.***

- Both ***winpercent*** and ***Composite Score*** are normalized to bring different scales onto a comparable range of 0 to 1.

CALCULATING FINAL SCORE TO DECIDE WINNERS

- The **Final Score** is calculated as follows :

$$\text{Final Score} = (\text{Normalized Winpercent} \times \text{Winpercent Weight}) + (\text{Normalized Composite Score} \times \text{Composite Score Weight})$$

- For example, a product with a normalized winpercent of 0.8 and a normalized composite score of 0.7, the final score would be calculated as:

Final Score = (0.8×0.6)+(0.7×0.4)
= 0.48+0.28
= 0.76

RANKING THE WINNERS 👑 BY FINAL SCORE

- The products are ranked based on the **Final Score**, with the top-ranked products having the highest final score. The ranking is done in descending order, meaning the product with the highest final score gets the top rank (Rank 1), and so on.

- The **Winners** based on the **Final_Score** are :

competitorname	Final_Score	Final_Rank
Reese's Peanut Butter cup	0.970389	1.0
Twix	0.952858	2.0
Snickers	0.914072	3.0

SWEET TREATS FOR THE SWEETEST TOOTH !

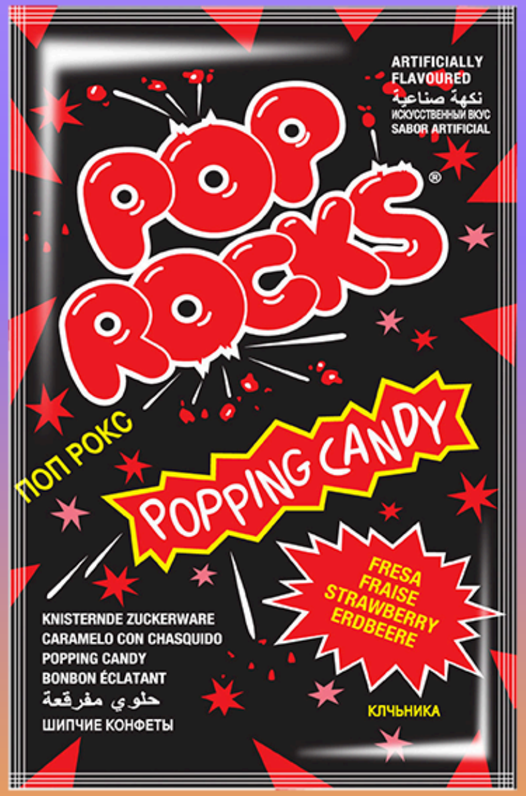
- In a similar way if we consider a combination of ***sugarpercent*** *and* ***Composite Score*** *to get a similar score as Final_Score, which we can call it* ***Sweet_Tooth_Score***, The Top 3 Candies for people with a sweet tooth are :

BUDGET-FRIENDLY $ AND DELICIOUS !

- In this case we consider a combination of ***pricepercent*** *and* ***Composite Score*** *to get a final score, which we can call it* ***Price_Sensitive_Score****, The Top 3 Best Value Candies for Price-Sensitive Shoppers* are :

TOP 3 FRUITY TREATS 🍊 FLAVOR-PACKED AND FUN !

- In this case we consider ***Composite Score*** *for all candies with fruit flavour, and then ranked the Top 3 Best Candies* :



KEY INSIGHTS & CONCLUSION

- Identified the Top 3 candies for different consumer preferences: Overall Winners, those with a sweet tooth, price-sensitive consumers, and fruity candy using a combination of data analysis techniques.

- Analyzed binary features (like chocolate, fruity) and continuous features (like *sugar percent, price percent*) to understand their impact on candy popularity (*winpercent*).

- Used **SHAP** values to determine the most influential features in predicting candy popularity. Key drivers included *chocolate ,sugar percent & pricepercent*.

- Developed a Composite Score by multiplying each feature's value with its SHAP-derived weight, providing a theoretical ranking of candies based on their attributes.

- Balanced the Composite Score with the real-world winpercent by using a weighted formula (*Final Score*), ensuring the rankings considered both theoretical predictions and consumer preferences.

Top 3 Overall Winners ⭐

- Reese's Peanut Buttercups
- Twix
- Snickers

Top 3 Budget-Friendly Candies $

- Tootsie Rolls
- Sixlets
- Hershey's Kisses

Top 3 Candies for a Sweet Tooth 🍬

- Reese's stuffed with pieces
- Milky Way Simply Caramel
- Whoppers

Top 3 Fruity Candies 🍊

- Tootsie Pop
- Ring Pop
- Pop Rocks

THANK YOU !