

# Management Summary

## Ziel

- Bestimmung der Eigenschaften der neuen Süßigkeit.
- Maximierung des voraussichtlichen Gewinns der neuen Süßigkeit.

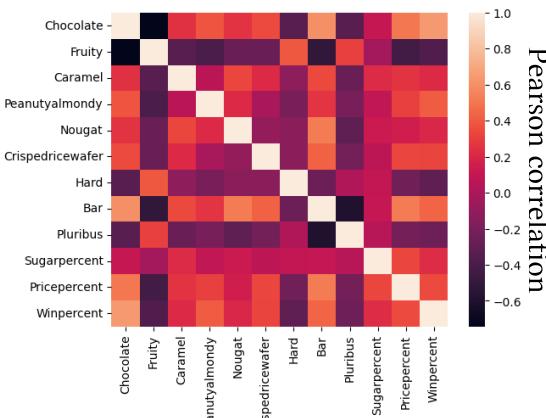
## Vorgehensweise explorative Datenanalyse

- Korrelationsanalyse zwischen den Eigenschaften. Die Grafik a. gibt erste Erkenntnisse, dass der Gewinn (Winpercent) die höchste Korrelation mit Schokolade hat.
- Betrachte die verschiedenen Eigenschaften und schaue ob Süßigkeiten mit oder ohne dieser Eigenschaft einen Einfluss auf den Gewinn hat (Grafik b. und c.).
- Die Verteilung der binären Variablen ist in Grafik b. zusammen mit der Signifikanz wiedergespiegelt. Die numerischen Variablen zusammen mit der Pearson Korrelation sind in Grafik c. zu sehen. Hier sehen wir eine schwache lineare Korrelation.

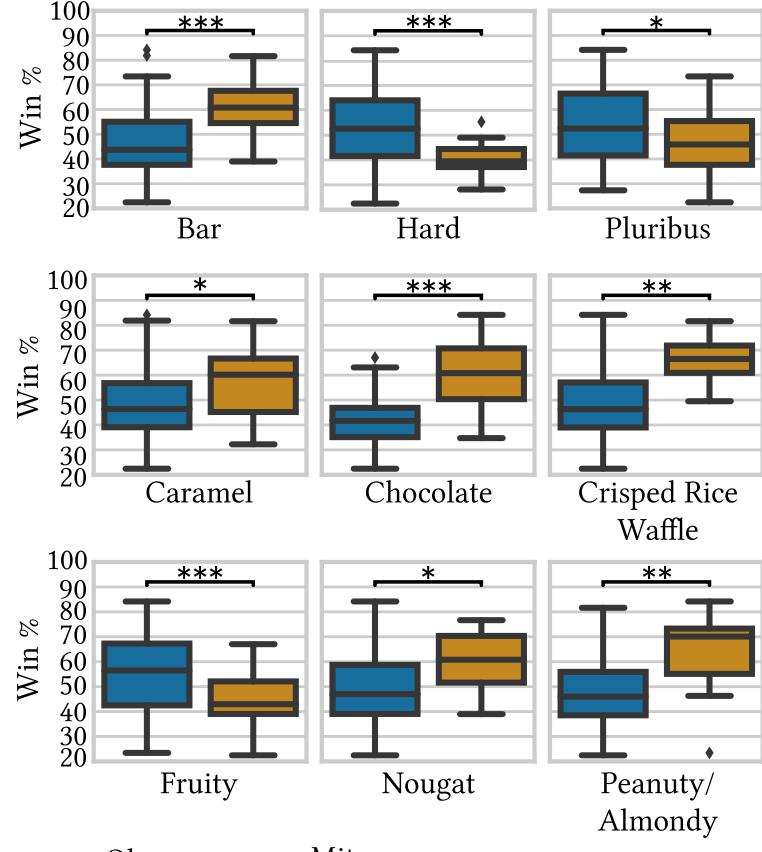
## Schlussfolgerung mittels explorativer Datenanalyse

- Basierend auf erster explorativer Analyse sollte unsere neue Süßigkeit folgende Eigenschaften haben: *Bar, nicht hart, Schokolade, Knusperige Reiswaffel, nicht fruchtig und Erdnuss/ Mandel.*

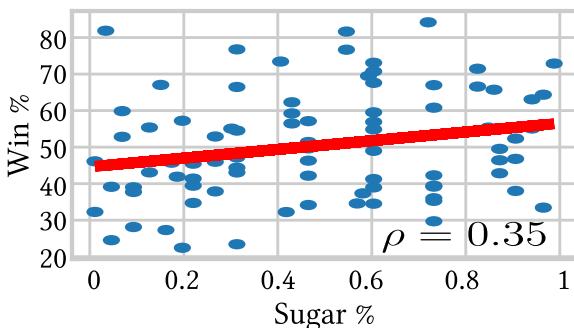
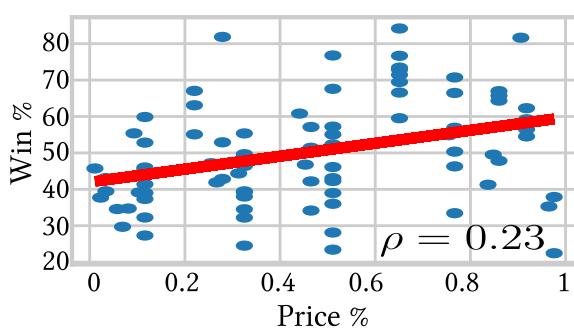
a.



b.



c.



■ Ohne Eigenschaft ■ Mit Eigenschaft

Welchs *t*-test significance: not significant (n.s.)

\*\*\*  $P < 10^{-3}$

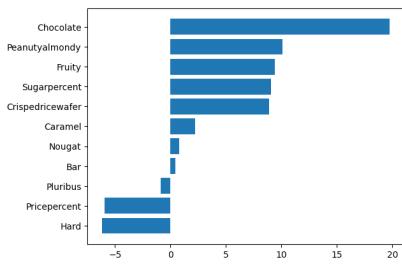
\*\*  $P < 10^{-2}$

\*  $P < 10^{-1}$

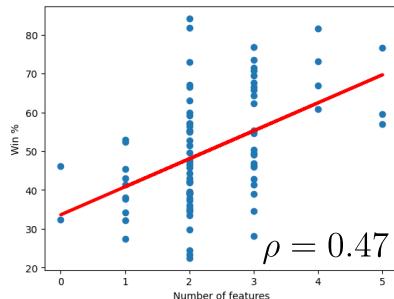
## Feature Analyse mittels Linearer Regression - Vorgehensweise

- Beschreibe den Gewinn durch alle anderen Eigenschaften im Datensatz mittels Linearer Regression (Bestimmtheitsmaß:  $R^2 = 0.57$ ).
- Identifizierte Eigenschaften mit größten Koeffizienten: Hoher Koeffizient beschreibt größere Auswirkung auf Gewinn, falls die Eigenschaft vorhanden ist (Grafik d.). Negativer Koeffizient besagt, dass die vorhandene Eigenschaft einen negativen Einfluss auf den Gewinn hat.
- Berechne P-Werte der Features und selektiere Features mit  $P < 0.05$ : Schokolade, Erdnuss/ Mandel, fruchtig
- Verifizierte ob Anzahl an Eigenschaften einer Süßigkeit einen Einfluss auf Gewinn haben (Grafik e.). Positive Pearson-Korrelation.
- Überprüfe mittels Key Driver Analyse welche Feature den stärksten Einfluss auf Gewinn haben (Grafik f.).

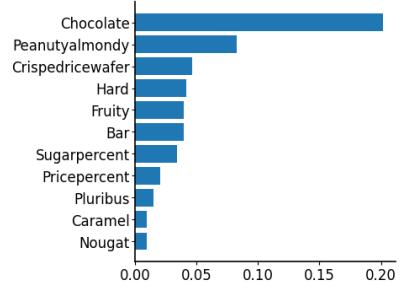
d.



e.



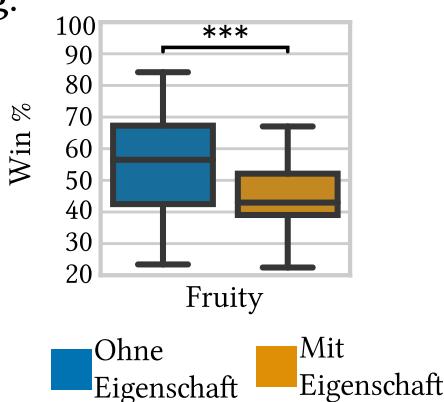
f.



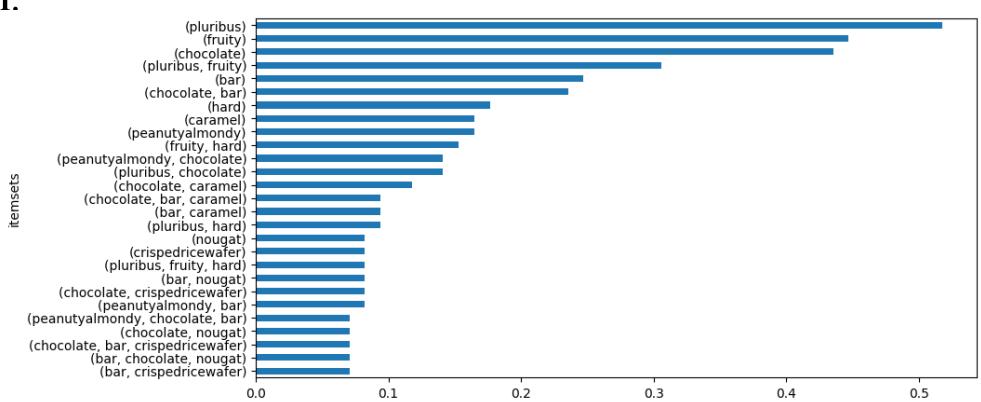
## Erläuterung der Notwendigkeit der Key-Driver-Analyse

- Lineare Regression empfiehlt die fruchtige Eigenschaft obwohl Plot g. vorschlägt, dass zusätzliche fruchtige Eigenschaft negativen Einfluss auf Gewinn hat.
- Fruchtig oft in Kombination mit pluribus (mehrere Süßigkeiten pro Packung) und hard (Grafik h.).
- Beide Eigenschaften wirken sich jedoch negativ auf Gewinn aus (Grafik b. und d.).
- Fruchtig bekommt höheren Koeffizienten zugewiesen → Modell versucht negativen Einfluss der anderen Features/ Kombinationen auszugleichen
- Daher Empfehlung der ersten vier Features mittels Key Driver Analyse

g.



h.



## Finale Schlussfolgerung

Unsere neue Süßigkeit sollte basierend auf der Marktanalyse folgende Eigenschaften haben:

- Schokolade
- Erdnuss/ Mandel
- Knusperige Reiswaffel
- Nicht hart