

Auswertung - canadian_beers.csv

1 & 2)

Aus welchen Beobachtungen setzen sich die Daten zusammen?
 Beschreiben Sie diese kurz und überlegen Sie sich, welches Skalenniveau die einzelnen Beobachtungen haben. & Wie setzen sich die einzelnen Beobachtungen zusammen? Verwenden Sie Methoden für univariate Datensätze sinnvoll.

Variable	Skalenniveau	Methode	Beobachtung
rank	ordinal	minimum, maximum, median	rank, from 1 to 100, as rated by BeerAdvocate.com users
name	nominal	count	name of the beer
brewery	nominal	count	the brewery responsible for this delicious creation
style	nominal	count	the style of the beer
abv	Verhältnis	average, mean, minimum, maximum, sd	alcohol by volume (%)
score	Intervall	average, mean, minimum, maximum, sd	overall score determined by BeerAdvocate.com users
ratings	Verhältnis	average, mean, minimum, maximum, sd	number of ratings

Skalenniveau: nominal (keine Rangordnung), ordinal (Rangfolge), Intervall (Gleichheit der Differenzen), Verhältnis (Wohldefinierter Nullpunkt)

3)

Ermitteln Sie Kennzahlen für die einzelnen Beobachtungen, die Ihnen sinnvoll erscheinen. (Lokation, Modus, Varianzen, Quantile...)

```
Frage 3):
```

```
---
   rank    abv   score  ratings
count 100.00 100.00 100.00 100.00
mean  50.50  8.27  4.14  513.60
std   29.01  1.93  0.11 1312.09
min   1.00  4.40  3.91  50.00
25%  25.75  6.57  4.07  91.50
50%  50.50  8.55  4.14 135.50
75%  75.25 10.00  4.20 271.00
max 100.00 12.20  4.48 10021.00
```

Varianz von rank: 841.67

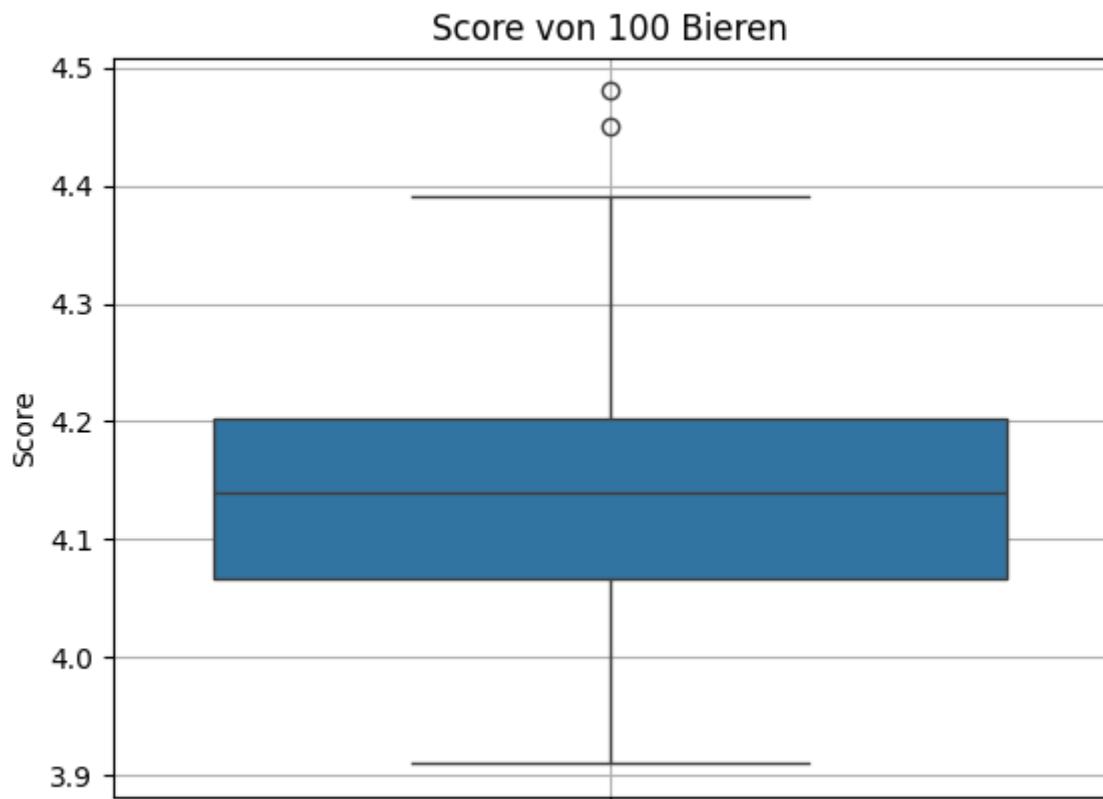
Varianz von abv: 3.73

Varianz von score: 0.01

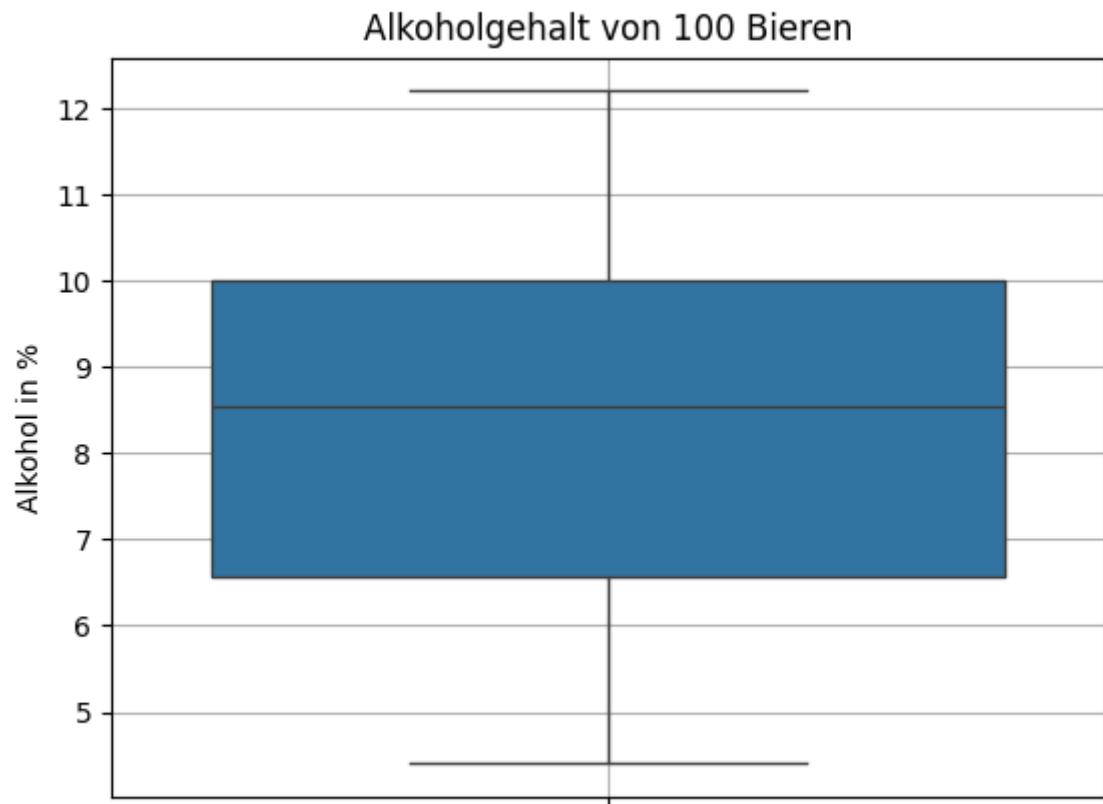
Varianz von ratings: 1721577.37

```
---
```

Auswertung Frage 3 - Anzahl, Median, Standardabweichung, min-Wert, Quartile (25%, 50% 75%) und max-Wert.
Darunter die jeweiligen Varianzen.



Boxplot - Score von 100 Bieren



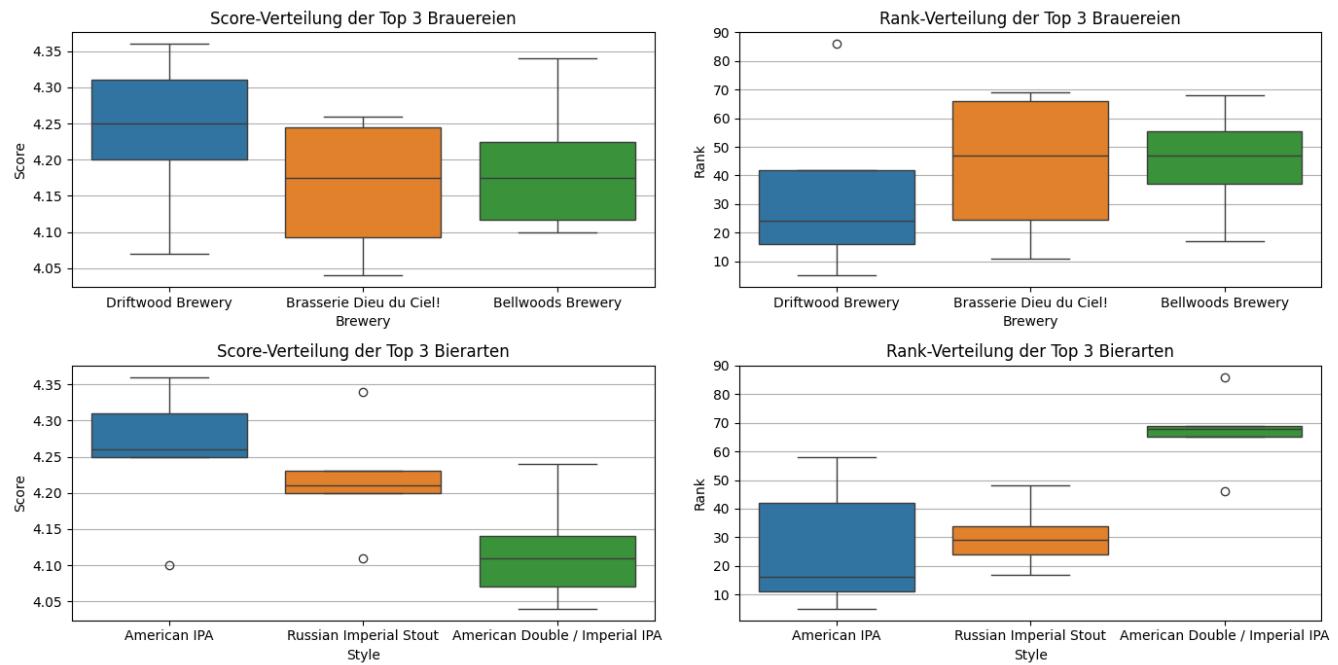
Boxplot - Alkoholgehalt von 100 Bieren

4)

Ermitteln Sie die 3 häufigsten Brauereien und Bier Arten, die in diesem Datensatz vorkommen. Wie verteilen sich Rang und Score auf diese Brauereien und Arten von Bier? Das Bier welcher Brauerei und welcher Art würden Sie jemandem empfehlen, wenn Sie sich nach dem Score richten? (Betrachten Sie hier nur die 3 häufigsten Arten)

brewery	style	score	rank
Bellwoods Brewery	American Double / Imperial IPA	4.12	57.0
	American IPA	4.10	58.0
	Russian Imperial Stout	4.26	33.0
Brasserie Dieu du Ciel!	American Double / Imperial IPA	4.14	67.0
	American IPA	4.26	11.0
	Russian Imperial Stout	4.11	29.0
Driftwood Brewery	American Double / Imperial IPA	4.07	86.0
	American IPA	4.31	21.0
	Russian Imperial Stout	4.20	24.0

Auswertung Frage 4 - zeigt brewery und style gruppiert, mit Median für score und rank



Multiplot - Top 3 Rank und Score Auswertung nach Brauereien und Bierarten

Analyse

Brauereien

Betrachtet man die erste Reihe der Plots, sieht man, dass die Brauerei "Driftwood Brewery" (*blau*) sowohl beim Score als auch bei der Rank-Verteilung (*niedriger Rank ist besser*) besser abschneidet als die beiden anderen.

Bierarten

Betrachtet man die zweite Reihe der Plots, sieht man, dass die Bierart "American IPA" (*blau*) sowohl beim Score als auch der Rank-Verteilung deutlich besser abschneidet als die beiden anderen. Die Bierart "American Double / Imperial IPA" (*grün*) schneidet mit Abstand in beiden Kategorien am schlechtesten ab.

Empfehlung

Wenn man die beiden "Score"-Plots betrachtet, ergibt sich eine eindeutige Empfehlung für die Bierart "American IPA" der Brauerei "Driftwood Brewery", da beide mit Abstand am besten abschneiden.

5)

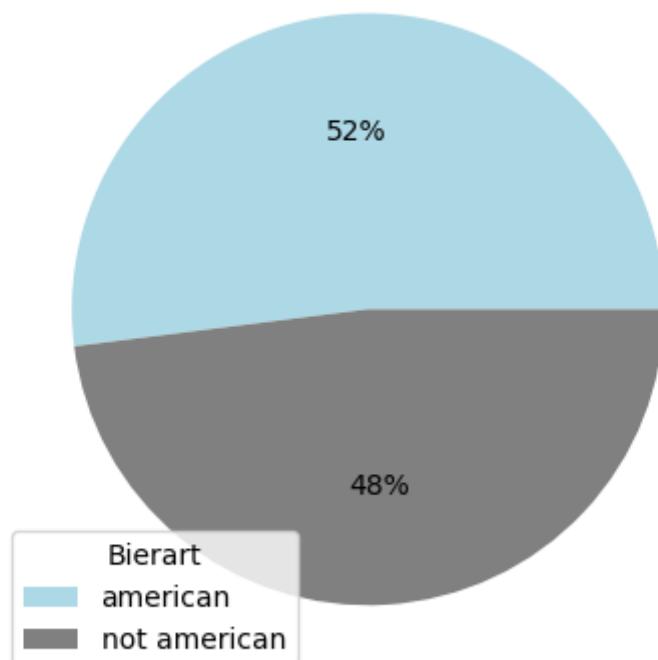
Fügen Sie eine Variable ein, die anzeigt, ob es sich um ein american style beer handelt (Tipp: Variable style). Wie sehen score und rank in diesem Fall aus?

Frage 5):

	score	rank
isAmericanStyle		
0	4.13	50.04
1	4.16	51.00

Auswertung Frage 5 - zeigt gruppiert, ob ein Bier "american style" ist, mit den Spalten score und rank

Verteilung amerikanischer Biere mit nicht amerikanischen Bieren



Piechart - Verteilung American/Non-American (zeigt nur Verteilung - score und rank nicht einbezogen)

Analyse

Die Auswertung zeigt das "american style"-Biere nicht unbedingt besser sind als andere. (Bezug auf Quelle: Auswertung Frage 5)

6)

Untersuchen Sie die Unterschiede zwischen 'Ale' und 'Stout' Bieren.

Zur Untersuchung der Unterschiede zwischen 'Ale' 'Stoudt' erstellen wir aus der Variablen "style" eine Variable beeType. Wir berücksichtigen somit nur Ale- und Stoudt-Biere. Für diese Teilmenge werden die Mittelwerte von Score, Alkoholgehalt (abv), Rank und Anzahl Ratings nach beerType berechnet und zusätzlich Boxplots für Score und Alkoholgehalt erstellt.

Man sieht, dass Stoudt-Biere im Durchschnitt einen höheren Alkoholgehalt als Ales haben. Die Scores von Ales und Stouts liegen ähnlich hoch, Stouts tendenziell leicht höher. Beim Rank schneiden Stouts im Mittel etwas besser ab (niedrigerer Rang). Ales haben im Durchschnitt mehr Bewertungen als Stouts.

Insgesamt gibt es leichte Unterschiede zugunsten der Stouts (stärker und etwas besser bewertet), aber keine extremen Abweichungen.

Auswertungen Frage 6

Frage 6):

Anzahl Ale- und Stout-Biere:

beerType

Anzahl Ale- und Stout-Biere:

beerType

Stout 21

Ale 19

Stout 21

Ale 19

Ale 19

Name: count, dtype: int64

Name: count, dtype: int64

Mittelwerte von Score, Alkoholgehalt, Rank und Ratings nach Biertyp:

Mittelwerte von Score, Alkoholgehalt, Rank und Ratings nach Biertyp:

	score	abv	rank	ratings
--	-------	-----	------	---------

	score	abv	rank	ratings
--	-------	-----	------	---------

beerType

beerType

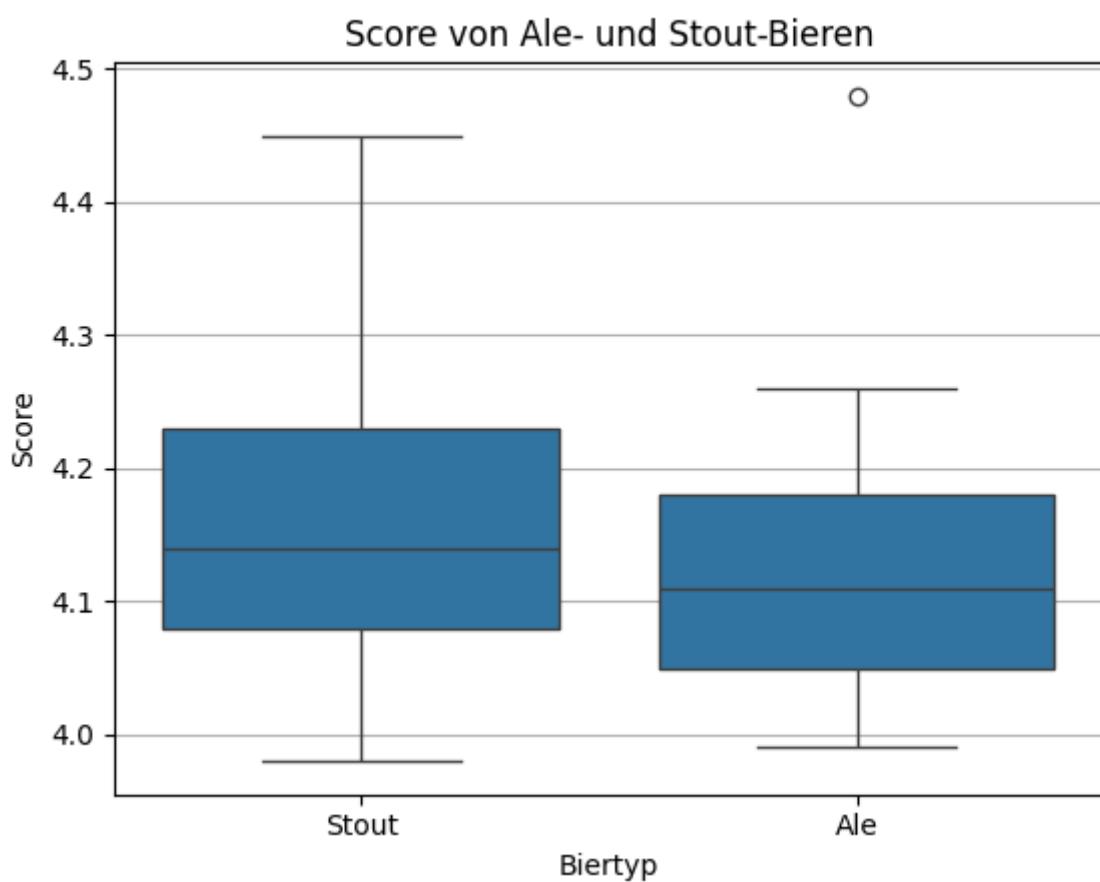
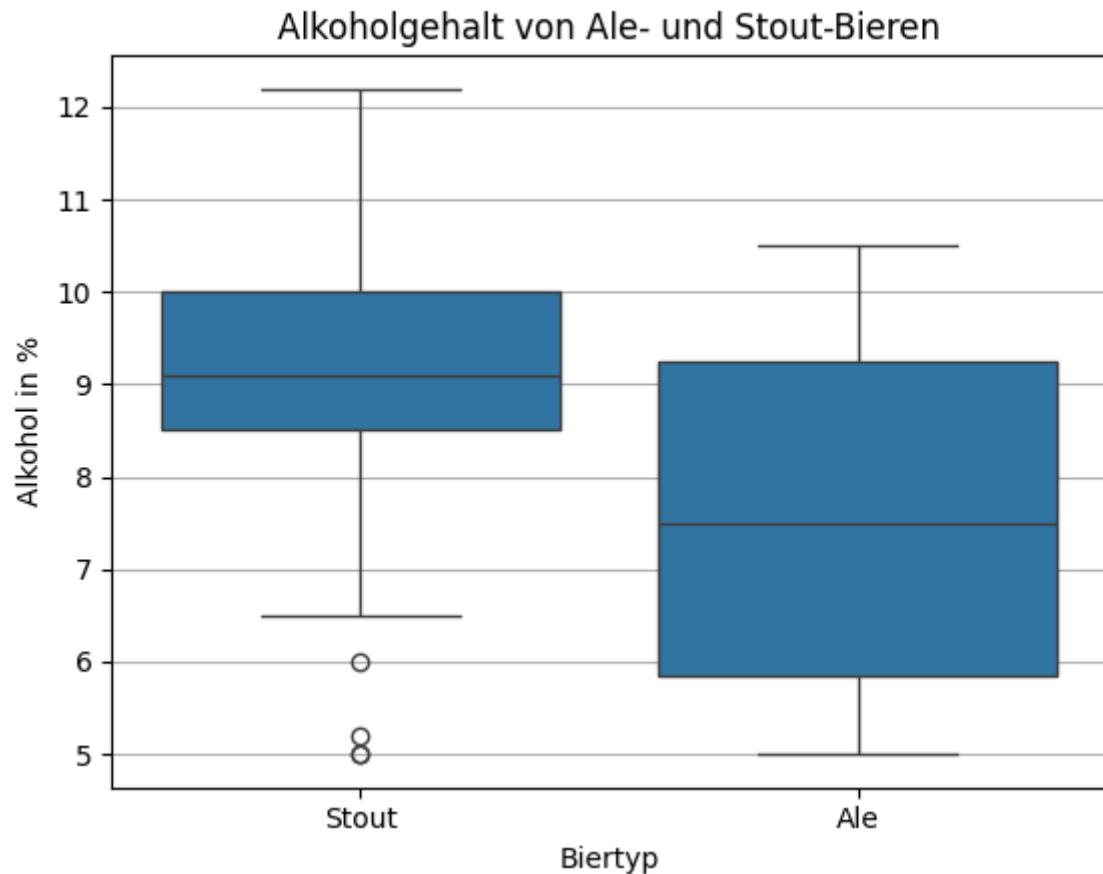
Ale	4.13	7.62	50.16	934.11
-----	------	------	-------	--------

Ale	4.13	7.62	50.16	934.11
-----	------	------	-------	--------

Stout	4.17	8.87	46.71	531.00
-------	------	------	-------	--------

Stout	4.17	8.87	46.71	531.00
-------	------	------	-------	--------

Plots Auswertung Frage 6

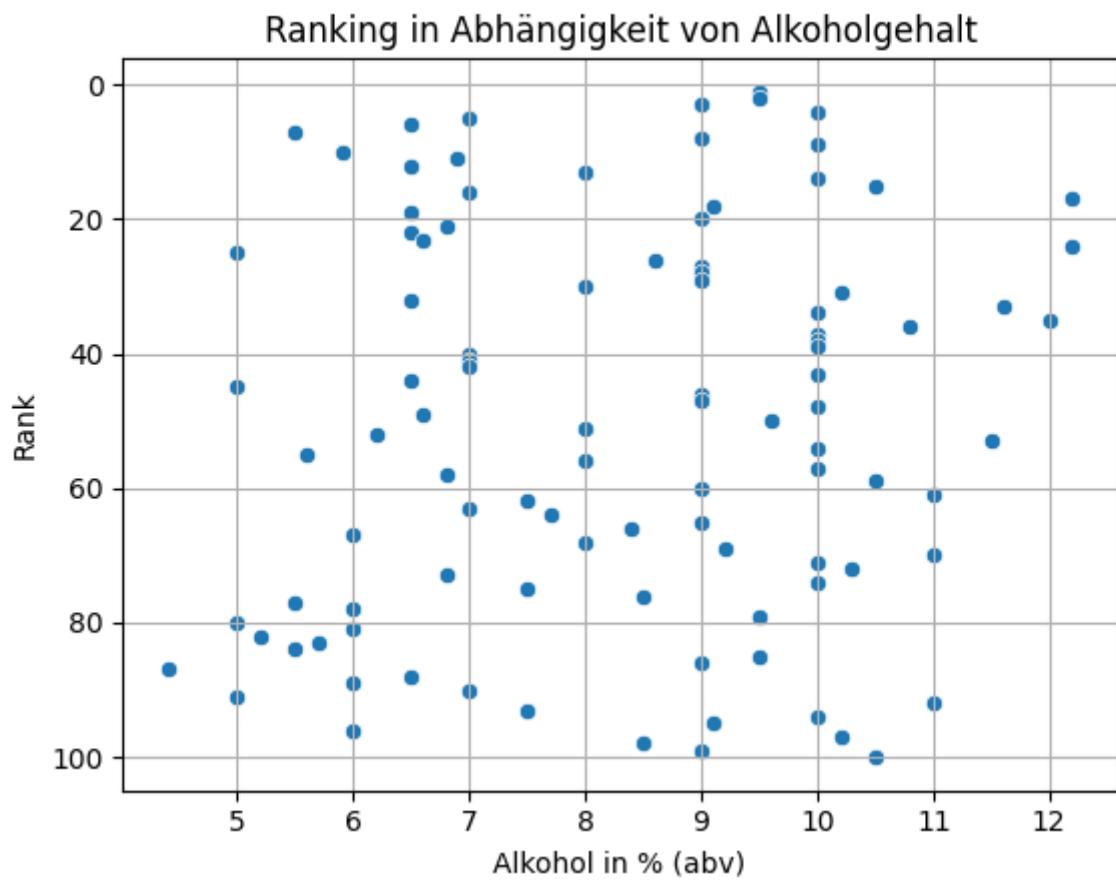


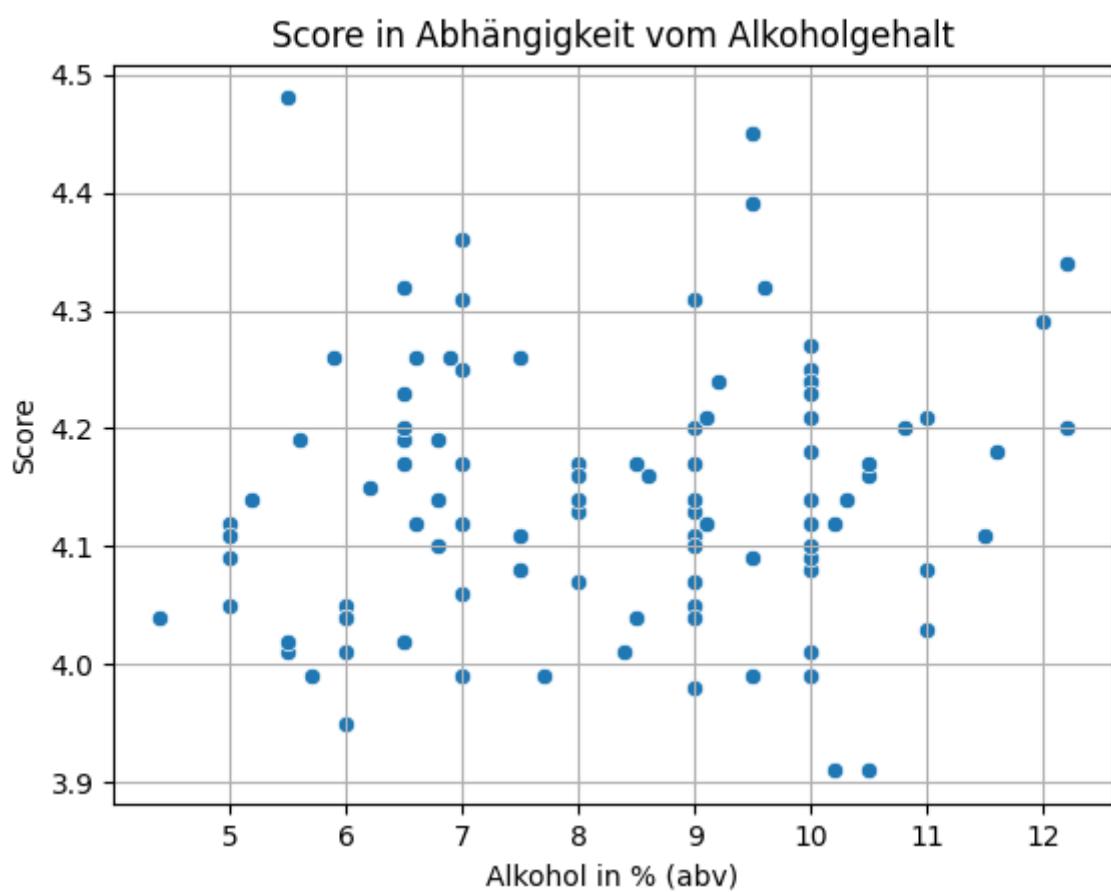
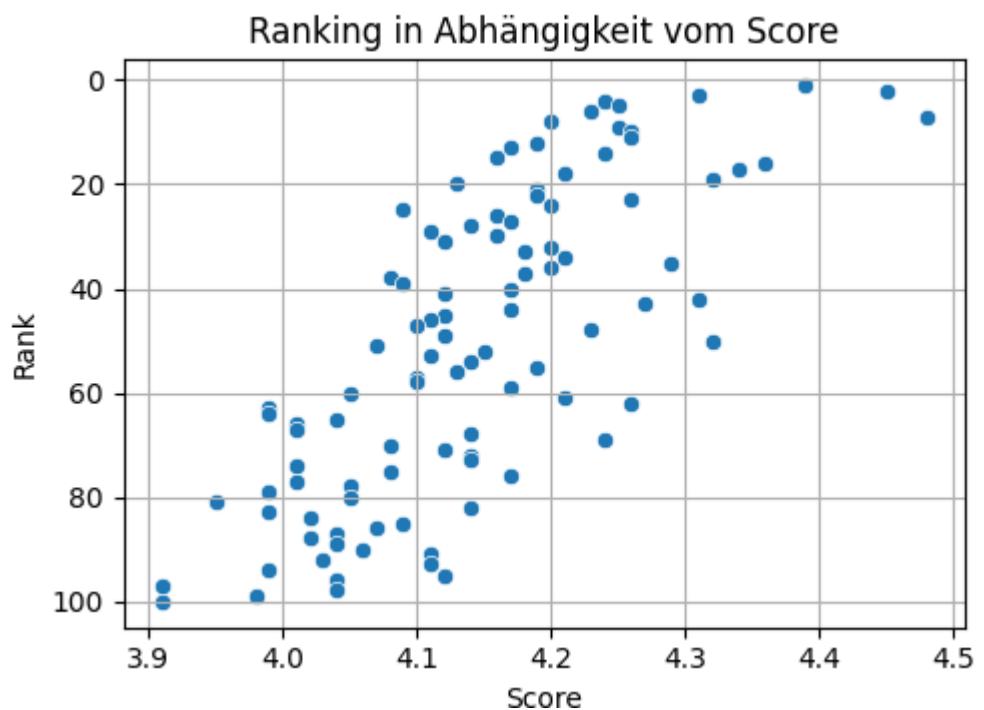
7)

Wie verhält sich das Ranking in Abhängigkeit vom Score und Alkoholgehalt? Ermitteln Sie, falls passend, den Korrelationskoeffizienten.

Zwischen Rank und Score gibt es einen schlechten Zusammenhang (je höher der Score, desto besser bzw. niedriger der Rank), während der Alkoholgehalt fast keinen Einfluss auf das Ranking hat, was man in den Scatterplots gut sieht. Beim Score ist ein deutlicher Trend erkennbar, beim Alkoholgehalt sehen die Punkte eher zufällig verteilt aus.

Plots Auswertung Frage 7





8)

Untersuchen Sie den Score in Abhangigkeit der anderen Variablen.
Welche Variablen haben Ihrer Meinung nach einen Einfluss?

Der Score wird hauptsächlich durch die Bewertung selbst und somit indirekt durch das Ranking bestimmt. Einen leichten Einfluss haben auch die Anzahl der Ratings sowie der Bierstil. Der reine Alkoholgehalt spielt dagegen nur eine kleine Rolle.

Plots Auswertung Frage 8

