

## Teil III: Statistik

- ⇒ Alle Fragen sind zu beantworten.
- ⇒ Es können keine oder mehrere Antworten richtig sein.
- ⇒ Eine Frage ist NUR dann richtig beantwortet, wenn ALLE richtigen Antworten angekreuzt wurden.
- ⇒ Wird eine oder mehrere Antworten zu viel oder zu wenig angekreuzt, gilt die Frage als **falsch** beantwortet!
- ⇒ Bei den Fragen 1. bis 15. sind 3 Punkte, bei Frage 16. fünf Punkte, also insgesamt 50 Punkte zu erreichen.

**1. Welche Möglichkeiten haben Sie, um den Zusammenhang zwischen zwei metrischen Variablen darzustellen. [3 Punkte]**

- Scatterplot
- Pearson-Korrelation
- Gar keine, weil die Variablen mindestens nominal skaliert sein müssen
- Balkendiagramm

**2. Für eine statistische Analyse wird eine Reihe von Merkmalen erfasst. Notieren Sie jeweils mit N, O, I bzw. R, ob es sich um eine nominale, ordinale, intervall- oder ratio-Skala handelt. [3 Punkte]**

- Gemeinde, in der der/die Befragte wohnt
- Einschätzung der Arbeit von Kanzlerin Merkel (ausgezeichnet, gut, nicht gut, sehr schlecht)
- Nationalität bei Geburt
- Intelligenzquotient
- Haushaltseinkommen
- Bildung (ohne 10. Klasse – 10. Klasse – Abitur)

**3. Kreuzen Sie alle möglichen Antworten an. Mit einem Histogramm wird Folgendes dargestellt:**

[3 Punkte]

- Häufigkeiten von stetigen metrischen Merkmalen
- Häufigkeiten von stetigen ordinalen Merkmalen
- Die Signifikanz einer bivariaten Verteilung
- Die Schiefe von Verteilungen

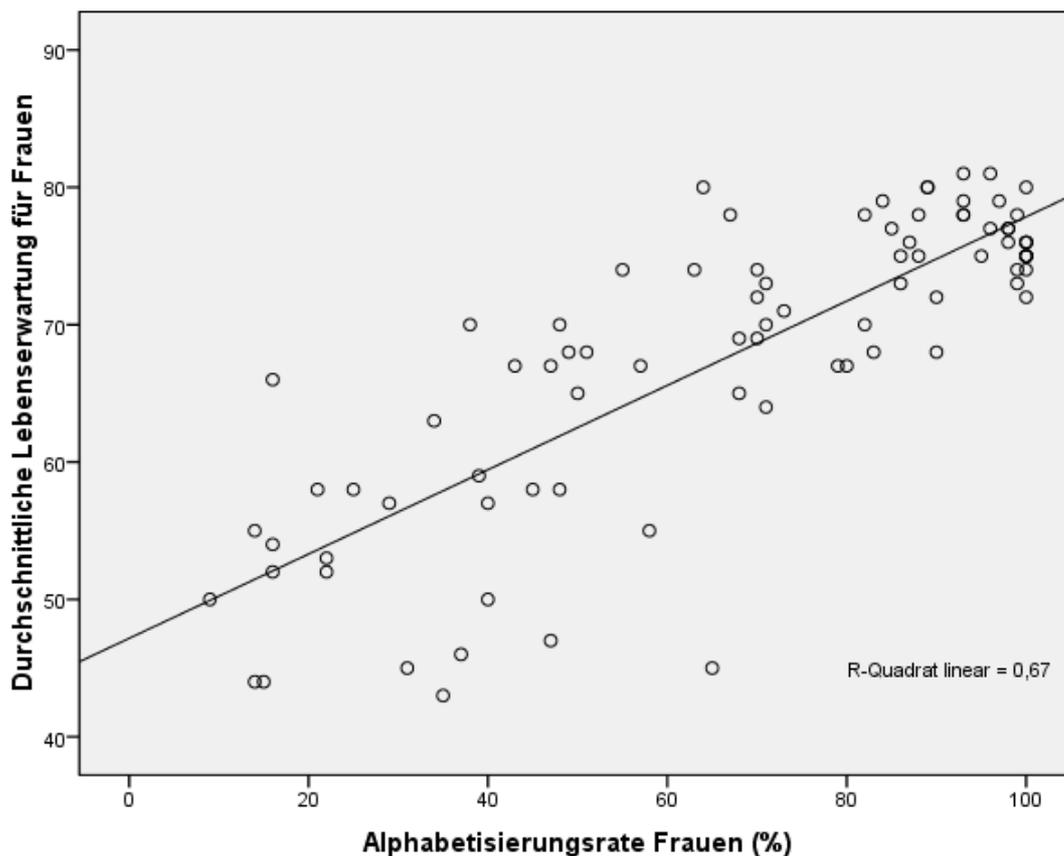
**4. Eine repräsentative Befragung in Greifswald erfasste 600 Studenten und 800 Angestellte. Davon gaben 520 Studenten sowie 400 Angestellte an mit ihrer Tätigkeit zufrieden zu sein. Die anderen waren unzufrieden. Kreuzen Sie alle möglichen Antworten an. [3 Punkte]**

- Der Prozentsatz der Zufriedenen ist bei den Studenten höher als bei den Angestellten.
- Der Anteil der Unzufriedenen ist bei Studenten und Angestellten gleich.
- Die Mehrheit der Unzufriedenen sind Angestellte.
- Die absolute Mehrheit der Unzufriedenen ist höher als die der Zufriedenen.

**5. Was ist bei Ausreißern zu beachten? Kreuzen Sie alle möglichen Antworten an. [3 Punkte]**

- Ausreißer sind Extremwerte und steigern die Signifikanz.
- Ausreißer können durch Boxplots identifiziert werden.
- Ausreißer können Ergebnisse verzerren.
- Ausreißer verbessern die Ergebnisse.

6. Erläutern Sie folgende Grafik [3 Punkte]:

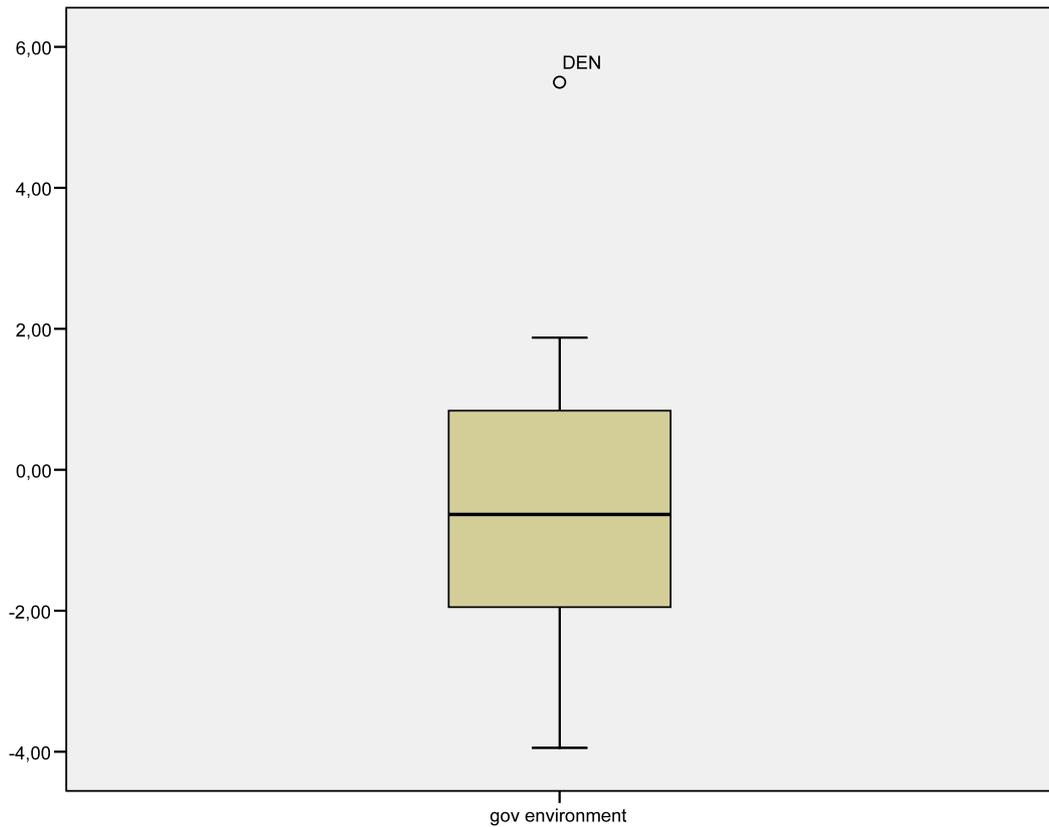


- Es handelt sich um ein Streudiagramm.
- Es besteht ein positiver Zusammenhang zwischen beiden Variablen.
- Es besteht ein negativer Zusammenhang zwischen beiden Variablen.
- Es besteht gar kein Zusammenhang zwischen beiden Variablen.
- Zwei Drittel der Varianz der einen Variable kann durch die Varianz der anderen Variable erklärt werden.
- Ausreißer verzerren den Zusammenhang zu stark.

7. Was ist bei der Korrelation zu beachten? [3 Punkte]

- Eine Korrelation beweist nicht zwingend einen kausalen Zusammenhang.
- Wenn Korrelation vorliegt, liegt auch immer Kausalität vor.
- Die Pearson-Korrelation setzt mindestens intervallskalierte Variablen voraus.
- Wenn der Korrelationswert unter 0,05 liegt, kann von einer Mittelwertgleichheit ausgegangen werden.
- Die Spearman-Korrelation basiert auf dem Rangvergleich.

**8. Interpretieren Sie folgende Grafik [3 Punkte]:**



- Es handelt sich um ein Histogramm.
- Aufgrund von Ausreißern hat die Grafik keine Erklärungskraft.
- Der Median liegt bei ungefähr -1.
- Der arithmetische Mittelwert liegt bei ungefähr -1.

**9. Ordnen Sie folgenden Beispielen die jeweils korrekten grafischen Darstellungsformen zu (jeweils mit Linien verbinden) [3 Punkte]:**

- |  |   |   |                            |
|--|---|---|----------------------------|
| Wahlabsicht in Abhängigkeit vom Geschlecht | - | - | Einfaches Balkendiagramm   |
| Einkommen in Abhängigkeit vom Bundesland   | - | - | Histogramm                 |
| Zustimmung zur Bundesregierung             | - | - | Gruppiertes Balkendiagramm |
| Bruttoinlandsprodukt aller Länder          | - | - | Gruppiertes Boxplot        |

**10. Was gilt für den Median?** [3 Punkte]

- Um ihn zu berechnen, bedarf es mindestens nominalskalierten Daten.
- Ist anfälliger für Ausreißer als das arithmetische Mittel.
- Liegt bei einer symmetrischen Verteilung rechts vom Modus.
- Der Median bezeichnet den Wert einer Verteilung, der genau in der Mitte liegt, über und unter dem also jeweils 50 Prozent der Werte liegen.

**11. Was ist bei der linearen Regression zu beachten?** [3 Punkte]

- Die abhängige Variable muss mindestens intervallskaliert sein.
- Die Beta-Werte geben die Wichtigkeit der unabhängigen Variablen für die Erklärung der abhängigen Variable an.
- Es darf keine Beziehung zwischen der abhängigen und der unabhängigen Variable vorliegen.
- Es soll keine Heteroskedastizität vorliegen.

**12. Was kennzeichnet den Determinationskoeffizienten  $R^2$  ?** [3 Punkte]

- Entspricht dem Quotient aus erklärter Streuung und Gesamtstreuung.
- Gibt mit 100 multipliziert den Anteil der erklärten Varianz der abhängigen Variable durch die unabhängigen Variablen wieder.
- Er ist unempfindlich gegenüber der Anzahl der unabhängigen Variablen und der Anzahl der Fälle.
- $R^2$  gibt die Güte des Modells wieder.
- Korrigiertes  $R^2$  gibt die Signifikanz des Modells wieder.

13. In einer Umweltstudie soll eine Korrelation Zusammenhänge zwischen Wirtschaftswachstum, EU-Mitgliedschaft (Dummy für Mitglieder=1; Nichtmitglieder=0), Umweltperformanz und Regierungsanteil linker Parteien darstellen. Kreuzen Sie die zutreffenden Aussagen an. [3 Punkte]

Correlations

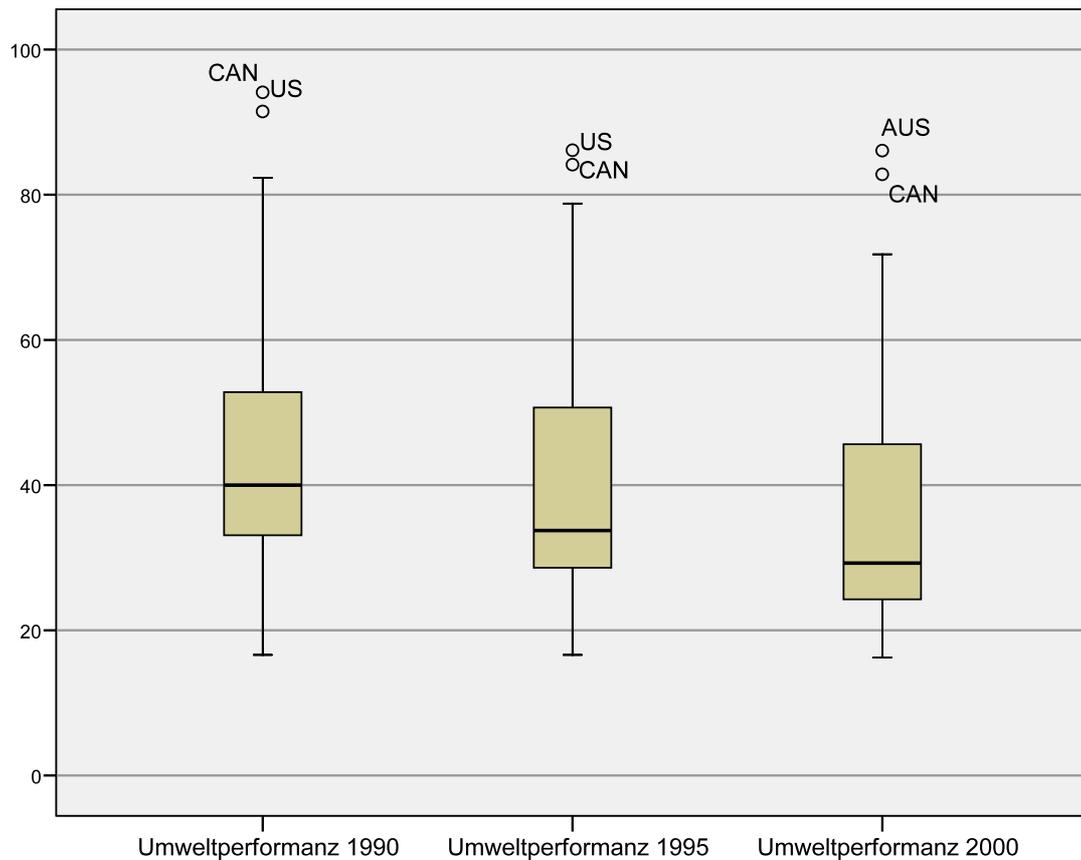
		Wirtschaftswachstum	Mitgliedschaft EU (Dummy)	Regierungsanteil linker Parteien	Umweltperformanz
Wirtschaftswachstum	Pearson Correlation	1	,138	,545*	,661**
	Sig. (2-tailed)		,549	,011	,001
	N	21	21	21	21
Mitgliedschaft EU (Dummy)	Pearson Correlation	,138	1	-,151	-,038
	Sig. (2-tailed)	,549		,514	,869
	N	21	21	21	21
Regierungsanteil linker Parteien	Pearson Correlation	,545*	-,151	1	,779**
	Sig. (2-tailed)	,011	,514		,000
	N	21	21	21	21
Umweltperformanz	Pearson Correlation	,661**	-,038	,779**	1
	Sig. (2-tailed)	,001	,869	,000	
	N	21	21	21	21

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

- Mit steigendem Wirtschaftswachstum nimmt die Umweltperformanz zu. Dies ist auf dem 1%-Niveau signifikant.
- Zwischen Umweltperformanz und EU-Mitgliedschaft besteht der größte signifikante Zusammenhang.
- Bei dieser Studie werden 21 Länder untersucht.
- Es besteht ein stärkerer Zusammenhang zwischen Umweltperformanz und Regierungsanteil linker Parteien, als zwischen Umweltperformanz und Wirtschaftswachstum.
- Eine Korrelation nach Pearson zwischen der Dummyvariable für EU-Mitgliedschaft und dem Wirtschaftswachstum ist aufgrund des Skalenniveaus unzulässig.

14. Für eine Umweltstudie wurde die Umweltperformanz von 21 Ländern zu drei Zeitpunkten verglichen. Kreuzen Sie die zutreffenden Aussagen an. [3 Punkte]



- Über die Zeit nimmt die Umweltperformanz tendenziell ab.
- Zu allen drei Zeitpunkten bestehen Ausreißer.
- Der Median bleibt stabil unter einem Wert von 20.
- Es hätte ein gruppiertes Balkendiagramm verwendet werden müssen, um die Zusammenhänge darstellen zu können.
- Zu allen drei Zeitpunkten befinden sich mindestens 75% der Fälle unter einem Wert von 60.

15. Welches sind Unterschiede und Gemeinsamkeiten von Spearman- und Pearson-Korrelation? [3 Punkte]

- Messen den Zusammenhang zwischen 2 Variablen.
- Spearman stellt die Stärke und Pearson die Richtung des Zusammenhangs dar.
- Pearson ist das Quadrat aus Spearman.
- Spearman darf nur bei metrisch skalierten Variablen angewendet werden.
- Bei beiden kann der Korrelationskoeffizient zwischen -1 und 1 liegen.

**16. Eine lineare Regression soll den Zusammenhang zwischen Wirtschaftswachstum, EU-Mitgliedschaft und Umweltperformanz aufzeigen. Kreuzen Sie die zutreffenden Aussagen an. [5 Punkte]**

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,674 <sup>a</sup>	,454	,393	,77903332

a. Predictors: (Constant), Wirtschaftswachstum, Mitgliedschaft EU (Dummy)

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9,076	2	4,538	7,477	,004 <sup>a</sup>
	Residual	10,924	18	,607		
	Total	20,000	20			

a. Predictors: (Constant), Wirtschaftswachstum, Mitgliedschaft EU (Dummy)

b. Dependent Variable: Umweltperformanz

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,336	,270		1,241	,231
	Mitgliedschaft EU (Dummy)	-,261	,347	-,132	-,753	,461
	Wirtschaftswachstum	,315	,082	,679	3,861	,001

a. Dependent Variable: Umweltperformanz

- Es wurde ein Modell gebildet, in dem das Wirtschaftswachstum durch die Umweltperformanz und die Mitgliedschaft in der EU erklärt wird.
- Eine EU-Mitgliedschaft führt zu einer geringeren Umweltperformanz. Allerdings ist dieses Ergebnis nicht auf dem 5%-Niveau signifikant.
- Das Gesamtmodell erklärt über ein Drittel der Varianz der abhängigen Variable.
- Das Modell hätte nicht angewandt werden dürfen, da die Anzahl der Freiheitsgrade nicht ausreichend ist.