

## Übung zur Korrelation und Regression

### Migranten in Großstädten in NRW

	Einwohner (in T)	Anteil mit Migrationshintergrund (in %)
Düsseldorf	586	32,5
Duisburg	488	30,1
Essen	566	24,6
Wuppertal	343	32,1
Bonn	306	30,6
Köln	1006	31,6
Aachen	236	30
Oberhausen	210	23,2
Münster	290	19,5
Dortmund	571	28,2
Bochum	362	22,9

Quelle: Zensus 2011

Gibt es einen Zusammenhang zwischen der Einwohnerzahl und dem Anteil der Menschen mit Migrationshintergrund? Berechne zur Untersuchung der Frage den Pearsonschen Korrelationskoeffizient und stelle ein Regressionsmodell auf. Was ist ein geeigneter Wert für die Aussagekraft des Modells und wie aussagekräftig ist das vorliegende?

### Religion in Bonn und Umgebung

	Einwohner (in T)	Anteil der römisch-katholischen Menschen (in %)
Bonn	306	40,52
Euskirchen	55	60
Brühl	44	50
Wachtberg	19	52,63
Bornheim	46	56,52
Siegburg	39	41,03
Sankt Augustin	54	44,44
Troisdorf	72	44,44
Hürth	55	50,91

Quelle: Zensus 2011

Eine häufig in sozialen Studien auftauchende Annahme geht davon aus, dass in ländlicheren Regionen der Anteil der einer Religion Zugehörigen höher liegt als in urbanen Gegenden. Können die vorliegenden Daten zum Anteil der römisch-katholischen Personen in ausgewählten (Klein-) Städten in der Bonner Umgebung diese Hypothese bestätigen? Untersuche anhand des Korrelationskoeffizienten sowie eines Regressionsmodells. Worin könnten die Ursachen liegen?

## Arbeit und Einkommen

	Arbeitslosenquote	Durchschnittseinkommen 2009 (in T)
Bonn	6,2	24,9
Düsseldorf	8,8	28,7
Duisburg	12,7	17,3
Essen	12,1	21,5
Gelsenkirchen	12,4	16,4
Köln	9,3	24,9
Paderborn	5,8	21,4
Oberhausen	10,1	18,2
Dortmund	13,2	19,1
Bochum	11,2	20,4

Quelle: Agentur für Arbeit/ Land NRW

Zwischen Arbeitslosenquote und Einkommen ist grundsätzlich ein starker Zusammenhang zu vermuten. Diskutiere: Welches ist die abhängige, welches die unabhängige Variable? Ist ein positiver oder negativer Zusammenhang zu erwarten? Überprüfe die Zusammenhangsannahme anhand des Korrelationskoeffizienten.

Aus der Berechnung der obigen Daten ergibt sich eine Steigung  $b$  von  $-0,89$  und ein Achsenabschnitt  $a$  von  $30,36$ . Stelle die zugehörige Regressionsgleichung auf.

Aus Krefeld ist eine Arbeitslosenquote von  $8,9$  bekannt, wie hoch müsste nach dem Modell das Durchschnittseinkommen liegen? Wie hoch müsste die Arbeitslosigkeit in Siegen sein, wenn das Regressionsmodell ein durchschnittliches Einkommen von  $25.630$  Euro prognostiziert?

## Managervergütungen

	Jahresüberschuss 2010 (in Mrd)	Vorstandsgehälter 2010 (in Mio)
Allianz	7,17	3,78
BASF	7,37	3,21
Beiersdorf	0,55	0,86
Daimler	6,63	4,39
Deutsche Bank	3,98	6,53
Fresenius	1,24	1,75
Metro	1,63	3,42
RWE	4,98	3,38
Siemens	5,81	4,15

Quelle: DSW-Studie 2011/ Studie Ernst & Young

Untersuche Überschuss und durchschnittliche Vorstandsgehälter (inklusive Boni) der neun aufgeführten DAX-Unternehmen. Gibt es einen Zusammenhang? Was ist die abhängige, was die unabhängige Variable.

Stelle ein Regressionsmodell auf. Wie lässt sich das Modell interpretieren? Welches Vorstandsgehalt müsste VW, mit einem Jahresumsatz von  $8,99$  Mrd. Euro, bezahlen?

## Übung zu Chi-Quadrat und Cramer's V

### Bundestagswahl 2009 – Wähler in T

	<b>CDU/CSU</b>	<b>SPD</b>	<b>FDP</b>	<b>Linke</b>	<b>Grüne</b>	<b>Summen</b>
<b>Männer</b>	6512	4832	3388	2794	1970	<b>19496</b>
<b>Frauen</b>	6686	5158	2928	2362	2674	<b>19808</b>
<b>Summen</b>	<b>13198</b>	<b>9990</b>	<b>6316</b>	<b>5156</b>	<b>4644</b>	<b>39304</b>

Quelle: Bundeswahlleitung

Haben Frauen und Männer unterschiedliche Parteipräferenzen? Offensichtlich ja. Aber gibt es auch signifikant messbare Unterschiede? Berechne zur Beantwortung dieser Frage den Chi-Quadrat-Wert der gegebenen Kreuztabelle und untersuche, ob ein Effekt vorliegt. Wie lässt sich das Ergebnis interpretieren?

### Cannabiskonsum im Jahr 2009

	<b>18-20 Jahre</b>	<b>21-24 Jahre</b>	<b>25-29 Jahre</b>	<b>30-39 Jahre</b>	<b>40-49 Jahre</b>	<b>Summen</b>
<b>Habe konsumiert</b>	16	14	11	4	4	<b>49</b>
<b>Habe nicht konsumiert</b>	84	86	89	96	96	<b>451</b>
<b>Summen</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>500</b>

In einer Umfrage zur Nutzung legaler und illegaler Drogen wurden aus verschiedenen Altersgruppen jeweils 100 Personen befragt. Was lässt sich beim Konsum von Cannabis bezüglich der Altersgruppen sagen. Welcher Zusammenhang kann vermutet werden und bestätigt ein Test diese Vermutung?

### Theoretische Chi-Quadrat-Werte

$\phi$	<b>0,050</b>	<b>0,010</b>	<b>0,005</b>	<b>0,001</b>
<b>1</b>	3,841	6,635	7,879	10,828
<b>2</b>	5,991	9,210	10,597	13,816
<b>3</b>	7,815	11,345	12,838	16,266
<b>4</b>	9,488	13,277	14,860	18,467
<b>5</b>	11,070	15,086	16,750	20,515
<b>6</b>	12,592	16,812	18,548	22,458
<b>7</b>	14,067	18,475	20,278	24,322
<b>8</b>	15,507	20,090	21,955	26,124
<b>9</b>	16,919	21,666	23,589	27,877
<b>10</b>	18,307	23,209	25,188	29,588
<b>11</b>	19,675	24,725	26,757	31,264
<b>12</b>	21,026	26,217	28,300	32,909
<b>13</b>	22,362	27,688	29,819	34,528
<b>14</b>	23,685	29,141	31,319	36,123
<b>15</b>	24,996	30,578	32,801	37,697

	Angst Stellenverlust				Gesamtsumme
	nein	ja, ich befürchte, arbeitslos zu werden	ja, ich befürchte, meine Stelle wechseln zu muessen	ja, ich befürchte beides	
öffentlicher Dienst	409	20	44	23	496
bundes- oder landeseigenes Unternehmen	63	12	11	9	95
gemeinnütziges Unternehmen	191	32	20	28	271
privatwirtschaftliches Unternehmen	1279	161	185	125	1750
	1942	225	260	185	2612

Quelle: Wahlkampfpanel 2013

2612 Personen in unterschiedlichen Beschäftigungsfeldern wurden im Jahr 2013 nach ihren Ängsten zu einem möglichen Stellenverlust befragt. Welcher Zusammenhang ist zu erwarten? Erlaubt die Erhebung eine signifikante Aussage? Und falls ja, kann ein relevanter Effekt festgestellt werden?

		Demokratiezufriedenheit					Gesamtsumme
		sehr zufrieden	zufrieden	teils/teils	unzufrieden	sehr unzufrieden	
Arbeitslosigkeit	ja	57	364	844	380	165	1810
vergangene 10 Jahre	nein	159	1050	1290	442	122	3063
Gesamtsumme		216	1414	2134	822	287	4873

Quelle: Wahlkampfpanel 2013

Besteht ein Zusammenhang zwischen Arbeitslosigkeitserfahrung und Zufriedenheit mit der Demokratie? Berechne den Chi-Quadrat-Wert der Tabelle sowie die Effektstärke.