

Grundlagen und Anwendung der Wahrscheinlichkeitstheorie

Wintersemester 2022/2023

L02 Projekt Datenanalyse

Gruppe 06

Björn Kämpfert
Felix Luis Fritsche
Benedict Casper Satter

Inhalt

Datensatz 3

3.1 Struktur und Inhalt

3.2 Skalenvariante

3.6 Verwendete Software und genutzte Funktionen

3.9 Modus, arithmetischer Mittelwert, Median

3.10 Spannweite

3.11 mittlere Abweichung vom Median

3.12 Stichprobenvarianz

3.13 Variationskoeffizient

3.14 Box-Whisker-Plot

3.15 Scatterplot

3.19 Zusammenfassung

3.20 / 3.21 Quartile und Dezile, Quartilabstand

3.22 Kovarianz

3.23 Korrelationskoeffizient

3.1 Struktur und Inhalt

Nach Absprache mit Dr.-Ing. Christian De Schryver haben wir für Datensatz 2 und 3 uns die Datensätze der Gruppe 05 besorgt, da unsere unrealistische Werte hatten.

Dieser Datensatz befasst sich mit Exportquoten der Jahre 2015 bis 2021. Die vorliegenden Dateien sind von 1 bis 7 durchnummeriert, wobei 2 und 6 nach der Bereinigung wegfallen. Jeder Zahl in der Datei ist entweder, wie in Datei a, eine Exportquote, in Prozent, zugeordnet, oder, wie in Datei b, eine Jahreszahl. Die Quelle und der Stand des Datensatzes sind in den letzten Zeilen angegeben, die Datei stammt vom Statistischen Bundesamt (Destatis), der angegebene Stand des Datensatzes ist der 12. September 2022.

3.4 Ergriffene Maßnahmen zur Bereinigung

Werte für die Jahre 2016 und 2020 fehlen jedoch, weshalb wir uns dazu entschieden haben, die Werte aus dem Datensatz zu entfernen.

3.6 Verwendete Software und genutzte Funktionen

Es liegen zwei Dateien, welche im Datensatz UTF 8-Kodiert sind, vor. Die beiden Tabellen sind in CSV-Dateien vorhanden, die Datei a, mit den Exportquoten, ist 149 Bytes groß und die Datei b, mit den Jahreszahlen, ist 136 Bytes groß . Die Trennung in den Urdateien verwendet als Dezimalzeichen ein Komma und als Trennzeichen ein Semikolon.

Zum Bearbeiten des Datensatzes wird Visual Studio Code benutzt, mit der Programmiersprache Python, dem Jupyter Notebook und den extra Bibliotheken pandas, numpy, matplotlib.

3.9 Modus, arithmetischer Mittelwert, Median

Modus: n.A., da es jeden Wert nur einmal gibt

Arithmetischer Mittelwert

Jahr	2018
Exportquote	47.02%

Median

Jahr	2018
Exportquote	47.00%

3.10 Spannweite

Jahr	6.0
Exportquote	0.6%

3.11 mittlere Abweichung vom Median

Jahr	1.6
Exportquote	0.184%

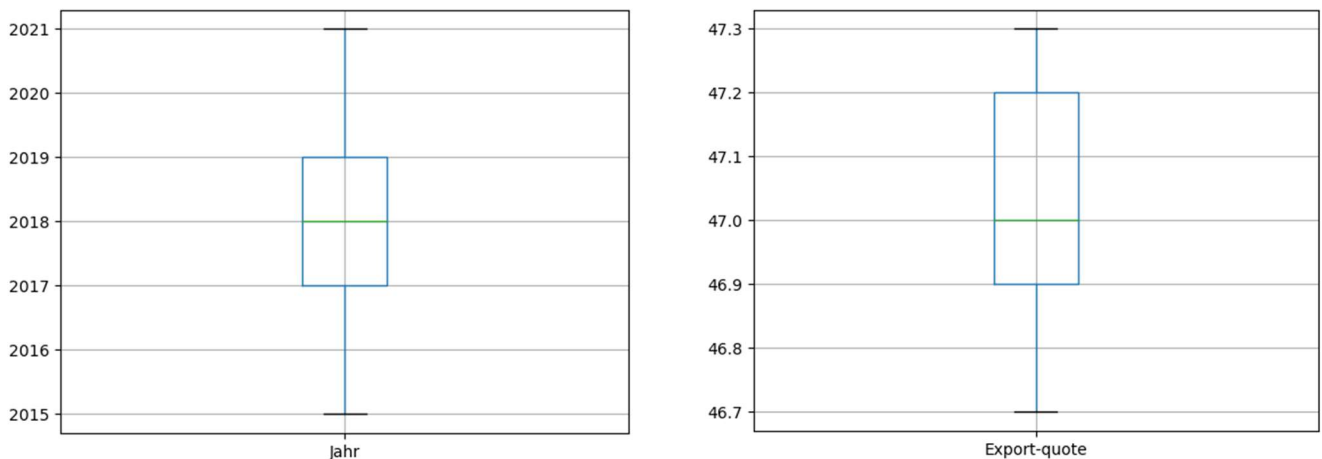
3.12 Stichprobenvarianz

Jahr	5000
Exportquote	0.057%

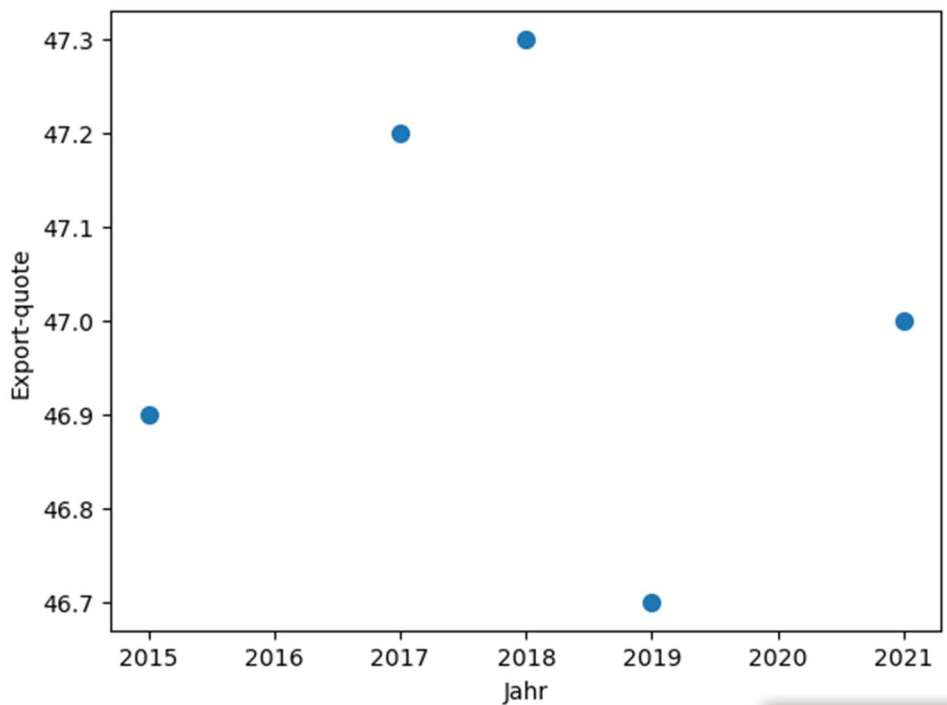
3.13 Variationskoeffizient

Jahr	0.111
Exportquote	0.507756%

3.14 Box-Whisker-Plot



3.15 Scatterplot



3.19 Zusammenfassung

Der Datensatz erfasst mit 7 Jahren einen eher kleinen Zeitraum, besonders wenn man beachtet, dass zwei der Jahre keine Angaben haben. Der arithmetische Mittelwert für die Exportquote liegt bei 47.02% und somit nur .02% vom Median entfernt, die mittlere Abweichung vom Median von nur 0.184% zeigt, dass sich die Exportquote über diesen Zeitraum nicht groß verändert hat, beziehungsweise dass die Werte alle nahe beieinander liegen. Die Spannweite der Exportquote beträgt auch nur 0.6 Prozent, wodurch diese Erkenntnis bestätigt wird. Der Variationskoeffizient fällt mit 0.507, aufgrund

der geringen Distanz der Werte, eher gering aus. Die Varianz von 0.057 zeigt einen geringen (quadrierten) Abstand der Werte vom arithmetischen Mittel.

3.20 / 3.21 Quartile und Dezile, Quartilabstand

Quartilsabstand: $Q3 - Q1 = 0.3\%$

	Jahreszahl	Exportquote
<i>min</i>	2015	46.7 %
<i>25%</i>	2017	46.9 %
<i>50%</i>	2018	47.0 %
<i>75%</i>	2019	47.2 %
<i>max</i>	2021	47.3 %

	Jahreszahl	Exportquote
<i>0.1</i>	2015.8	46.78 %
<i>0.2</i>	2016.6	46.86 %
<i>0.3</i>	2017.2	46.92 %
<i>0.4</i>	2017.6	46.96 %
<i>0.5</i>	2018.0	47.00 %
<i>0.6</i>	2018.4	47.08 %
<i>0.7</i>	2018.8	47.16 %
<i>0.8</i>	2019.4	47.22 %
<i>0.9</i>	2020.2	47.26 %
<i>1.0</i>	2021.0	47.30 %

3.22 Kovarianz

-0.05

3.23 Korrelationskoeffizient

-0.094

Quellenverzeichnis

Folien aus Vorlesung

<https://pandas.pydata.org/>

<https://studyflix.de/statistik/empirische-varianz-2016>

<https://www.geeksforgeeks.org/>

<https://matheguru.com/stochastik/kovarianz.html>