

Modellierung der Kaltmiete

Vergleich Frankfurt am Main und Leipzig

Henrik Popp, Kai Herbst, Manuel Zeh

2024-01-25

Inhaltsverzeichnis

Aufgabenstellung	1
Einleitung	2
Datenerhebung	2
Explorative Datenanalyse	3
Modellierung	8
Zusammenfassung	9
Quellen und Hilfsmittel	9

Aufgabenstellung

Abschnitt	Aufgabe	Reiner Textumfang	Erledigt
Einleitung	Auf inhaltliche Aufgabenstellung eingehen	0,5 - 1 Seiten	[]
Datenerhebung	Wie wurden die Daten erhoben? (Suchfilter, Sortierung)	1 - 3 Sätze	[]
Explorative Datenanalyse	Analyse + eventl. Datenvorverarbeitung	1 - 2 Seiten	[]
Modellierung	Modellierung + Interpretation	1 - 2 Seiten.	[]
Zusammenfassung	Gemeinsam kurz zentrale Ergebnisse zusammenfassen + Auf Grenzen der Analyse eingehen	0,5 - 1 Seiten.	[]

- Hier auch noch Literatur recherchieren:
 - <https://de.statista.com/statistik/daten/studie/258635/umfrage/bruttokaltmiete-bewohnter-wohnungen-in-deutschland-nach-bundeslaendern/>
 - https://www.deutschlandatlas.bund.de/DE/Karten/Wie-wir-wohnen/040-Mieten.html#_6a54aw429
 - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8053893/>
 - <https://de.statista.com/statistik/daten/studie/262508/umfrage/mietpreise-in-frankfurt-am-main/>
 - <https://de.statista.com/statistik/daten/studie/1312743/umfrage/mieten-in-leipzig-nach-dem-baualter-der-wohnung/>
 - <https://de.statista.com/statistik/daten/studie/535299/umfrage/mietpreise-auf-dem-wohnungsmarkt-in-leipzig/>
 - <https://de.statista.com/statistik/daten/studie/1312730/umfrage/entwicklung-der-angebotsmieten-in-leipzig/>
 - https://link.springer.com/chapter/10.1007/978-3-658-11757-3_4
 - https://www.ifo.de/DocDL/ifoDD_14-06_03-10.pdf

Einleitung

In dieser Fallstudie sollen die Kaltmieten der beiden Städte Frankfurt am Main und Leipzig miteinander verglichen und modelliert werden. Ziel ist es, die verschiedenen möglichen Einflussfaktoren auf die Kaltmiete in den jeweiligen Städte zu bestimmen und eine Modellierung der Kaltmiete zu erstellen.

Zu Beginn wird auf die Datenerhebung eingegangen. Hier soll erklärt werden, woher die verarbeiteten Daten stammen und unter welchen Bedingungen die Daten erhoben wurden. Mit der explorativen Datenanalyse sollen dann die erhobenen Daten beschrieben und veranschaulicht werden. Hierbei wird die Vorverarbeitung der Daten beschrieben, im Anschluss wird mithilfe von Grafiken und dazugehörigen Interpretationen eine Datenanalyse erstellt. Dabei soll unter anderem herausgefunden werden, welche erhobenen Variablen den größten Einfluss auf die Kaltmiete einer Stadt haben oder wie hoch die eventuellen Unterschiede der Mieten in den beiden Städten sind. Den zentralen Teil des Dokuments stellt die Modellierung dar. Hier soll die Kaltmiete modelliert, also durch ein selbsterstelltes statistisches Modell dargestellt werden. Zudem wird das Modell interpretiert. Zum Abschluss werden die Ergebnisse in der Zusammenfassung aufgearbeitet und präsentiert.

Datenerhebung

Die Datenerhebung fand ausschließlich über den Online-Marktplatz für Wohnungen und Häuser *ImmobilienScout24* statt. Die untersuchten Objekte wurden dabei auf den Immobilientyp *Wohnung* beschränkt, was als Suchkriterium in der Suchleiste des Portals eingestellt werden kann. Weitere Suchkriterien haben sich auf den *Ort*, in diesem Fall Frankfurt am Main und Leipzig, und auf den Objekttyp, hier *Mieten*, beschränkt. Weitere

Kriterien wie *Anzahl der Zimmer*, *Fläche* oder einem *maximalen Preis* wurden auf den Standardeinstellungen belassen. Anschließend wurden je Ort der Reihe nach bis zu 45 Objekte in der von ImmobilienScout24 generierten Reihenfolge überprüft und in eine Excel-Datei aufgenommen, die im Folgendem als Basis für die Auswertung dient.

Aufgenommen in die Datenbasis wurden dabei die folgenden Variablen: der *Ort*, die *Kaltniete* in Euro, die *Wohnfläche* in Quadratmetern, das Angebot eines *Parkplatz*, die *Etage*, Anzahl der *Zimmer*, Vorhandensein eines *Balkon*, das *Baujahr* des Objektes, sowie der entsprechende Link zur Anzeige und dessen Abrufdatum.

Für die nachfolgenden Auswertungen und Analysen lesen wir zunächst die Excel-Datei ein:

```
# Pfad zur Excel-Datei erstellen
pfad_mieten <- here("Mieten.xlsx")
# Daten einlesen
mieten <- read_excel(pfad_mieten)
```

Über die Ausgabe der ersten sechs Einträge erhalten wir einen Einblick in die Daten:

```
# Obere 6 Beobachtungen
head(mieten)
```

```
# A tibble: 6 x 12
  Ort    Kaltmiete Wohnflaeche Parkplatz Etage Zimmer Balkon Einbaukueche Heizung
<chr>    <dbl>      <dbl> <chr>    <chr> <dbl> <chr>    <chr>      <chr>
1 Fran~    1800         70   ja      1      2 ja      ja      Fußbod~
2 Fran~    1500         60   ja      1      1 ja      ja      Zentra~
3 Fran~    2650        146.   ja      1      3 ja      ja      Fußbod~
4 Fran~    1700         94   ja      2      3 nein    ja      Fußbod~
5 Fran~    2000        113.   ja      3      4 ja      ja      Fußbod~
6 Fran~    1700        84.8   ja      3      3 ja      ja      Fußbod~
# i 3 more variables: Baujahr <chr>, Link <chr>, Abrufdatum <dtm>
```

Explorative Datenanalyse

Zu Beginn der explorativen Datenanalyse muss geprüft werden, ob die in der Datenquelle enthaltenen Daten auf eine bestimmte Art und Weise vorverarbeitet oder angepasst werden müssen. Hierzu kann zunächst mit `str(mieten)` die Struktur des Datensatzes angezeigt werden.

```
str(mieten)
```

```
tibble [100 x 12] (S3: tbl_df/tbl/data.frame)
 $ Ort      : chr [1:100] "Frankfurt" "Frankfurt" "Frankfurt" "Frankfurt" ...
```

```

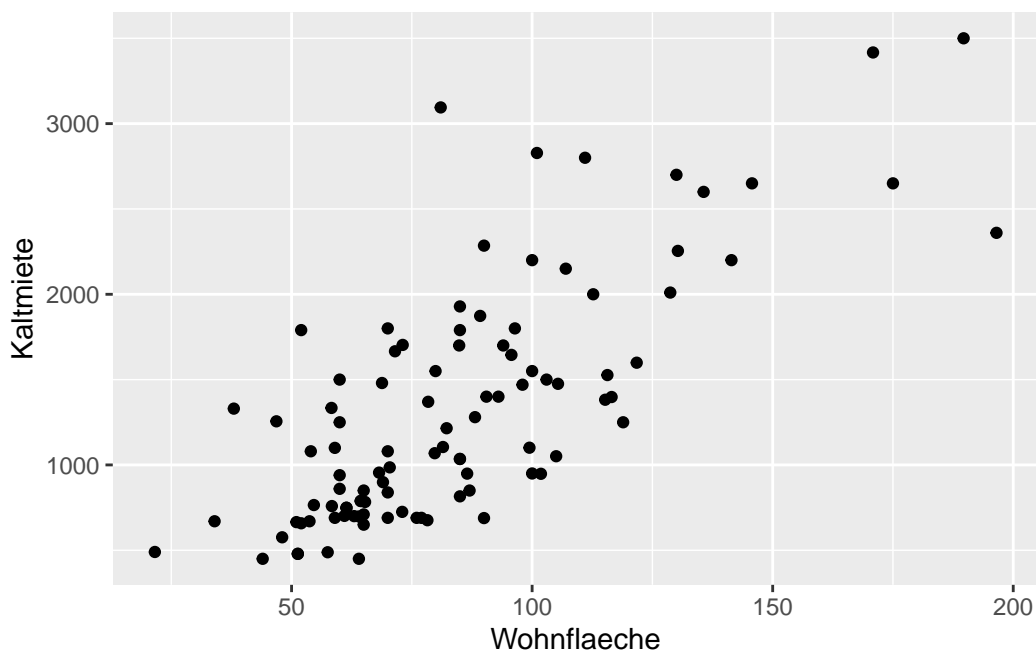
$ Kaltmiete      : num [1:100] 1800 1500 2650 1700 2000 1700 1480 2800 1080 2600 ...
$ Wohnflaeche    : num [1:100] 70 60 146 94 113 ...
$ Parkplatz      : chr [1:100] "ja" "ja" "ja" "ja" ...
$ Etage          : chr [1:100] "1" "1" "1" "2" ...
$ Zimmer         : num [1:100] 2 1 3 3 4 3 2 3 2 4 ...
$ Balkon         : chr [1:100] "ja" "ja" "ja" "nein" ...
$ Einbaukueche   : chr [1:100] "ja" "ja" "ja" "ja" ...
$ Heizung        : chr [1:100] "Fußbodenheizung" "Zentralheizung" "Fußbodenheizung" "Fußbodenheizung" ...
$ Baujahr        : chr [1:100] "2022" "1970" "2017" "-" ...
$ Link           : chr [1:100] "https://www.immobilienscout24.de/expose/136299839?referrer=R..."
$ Abrufdatum     : POSIXct[1:100], format: "2023-12-28" "2023-12-28" ...

```

Es kann festgestellt werden, dass im Datensatz sowohl kategoriale nominale Variablen wie **Heizung** oder **Zimmer**, als auch metrische verhältnisskalierte Variablen wie **Kaltmiete** oder **Wohnflaeche** auftreten. Zunächst werden keine Variablen angepasst bzw. Werte ersetzt, da für die späteren Diagramme die kategorial nominalen Variablen als Achsenbeschriftung gut verwendet werden können.

Zuerst soll auf den Zusammenhang von **Kaltmiete** und **Wohnflaeche** geschaut werden, bei zwei metrisch verhältnisskalierten Variablen bietet sich dafür ein Scatterplot an.

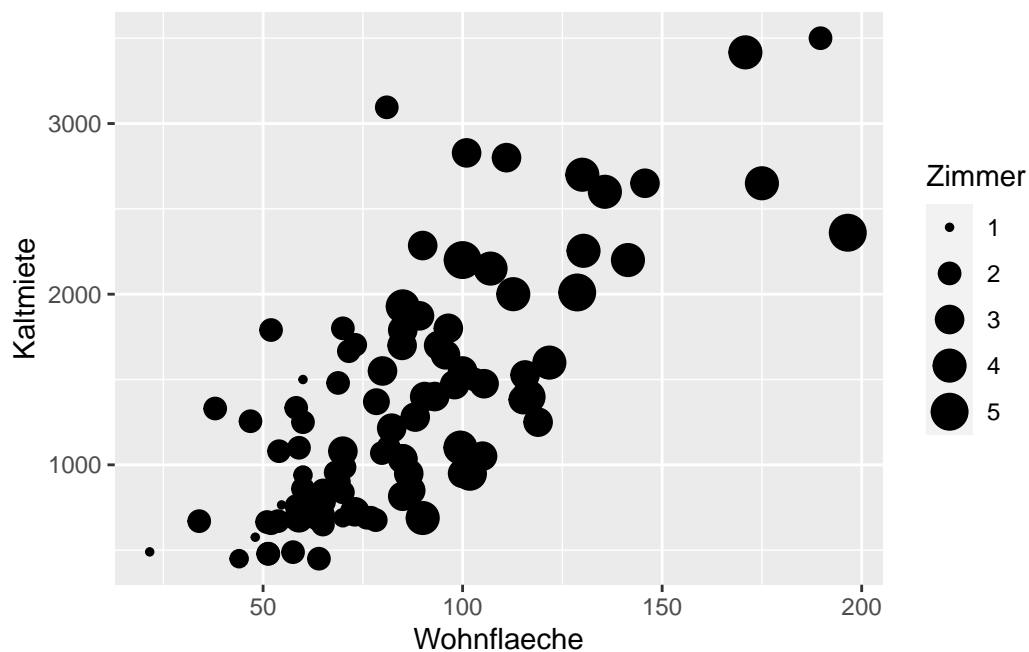
```
gf_point(Kaltmiete ~ Wohnflaeche, data = mieten)
```



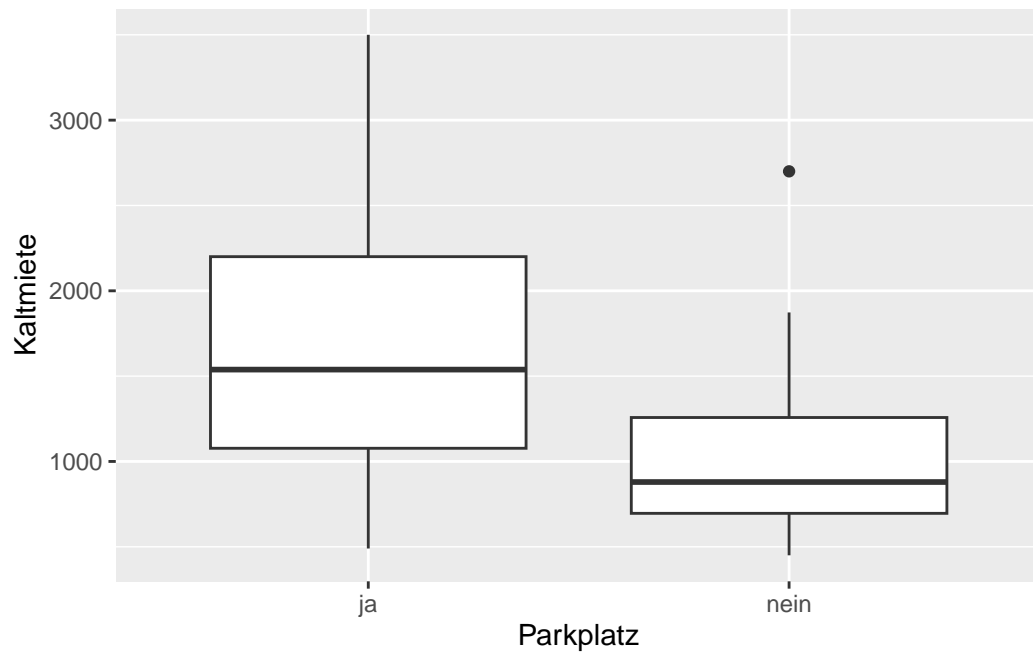
Grundsätzlich lässt sich ein positiver Zusammenhang zwischen **Kaltmiete** und **Wohnflaeche** erkennen, wobei die Streuung der Kaltmiete mit zunehmender Wohnfläche zunimmt. Als

nächstes soll das Scatterplot um die Anzahl an Zimmern erweitert werden, hierfür wird das obige Diagramm eingefärbt. Möchte man die Kaltmiete unter Betrachtung der Anzahl an Zimmern darstellen, bietet sich ein Streudiagramm an, da man eine stetige Variable (Kaltmiete) und eine diskrete Variable (Zimmer) hat.

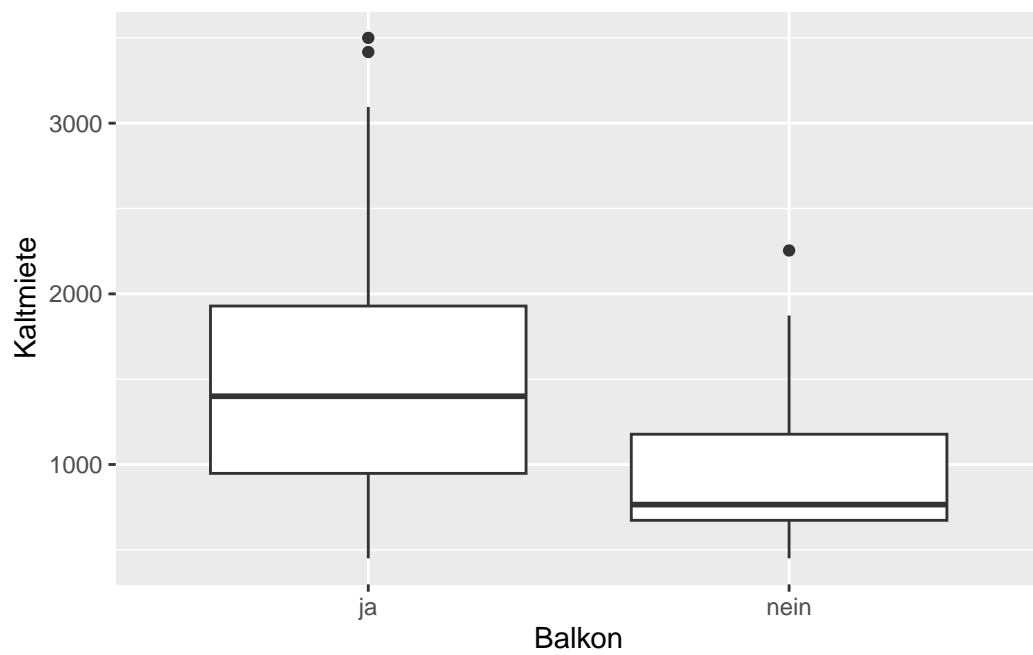
```
gf_point(Kaltmiete ~ Wohnflaeche, size = ~ Zimmer, data = mieten)
```



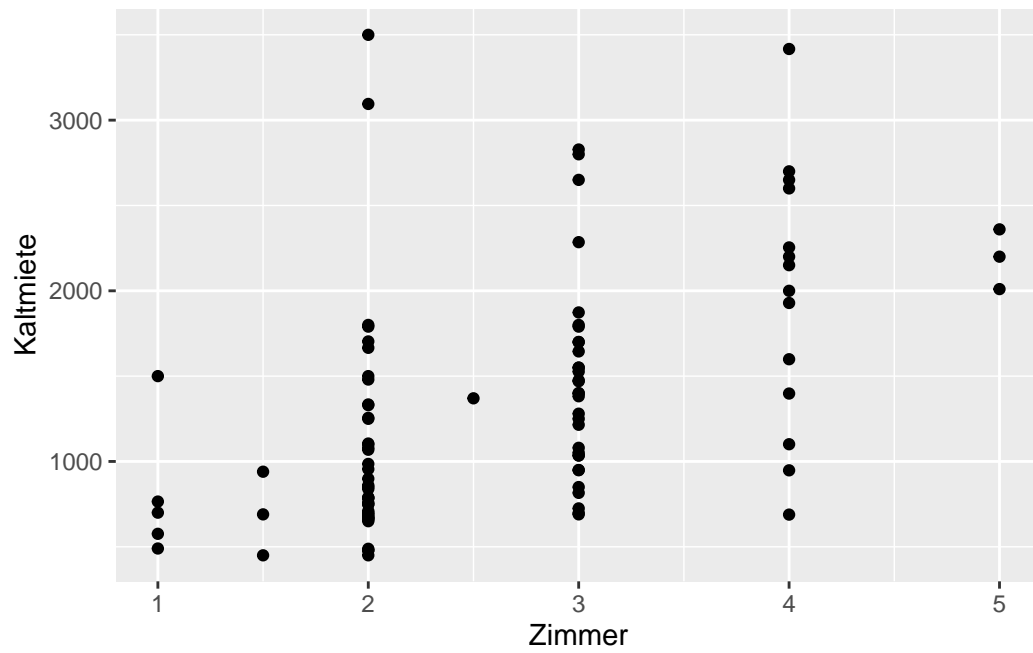
```
gf_boxplot(Kaltmiete ~ Parkplatz, data = mieten)
```



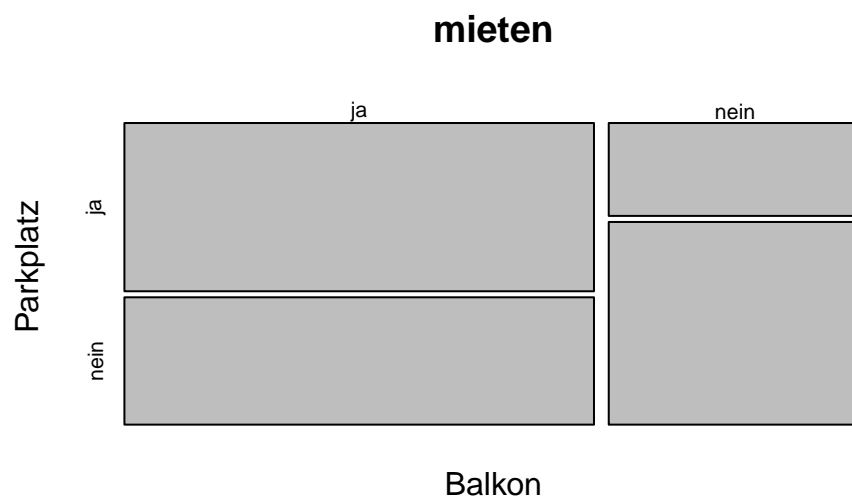
```
gf_boxplot(Kaltmiete ~ Balkon, data = mieten)
```



```
gf_point(Kaltmiete ~ Zimmer, data = mieten)
```



```
mosaicplot(Balkon ~ Parkplatz, data = mieten)
```



Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est

Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Modellierung

Modellieren Sie in diesem Abschnitt die Miete und interpretieren Sie Ihr Ergebnis.

Bei Einzelarbeiten sollte der reine Text (ohne Code, Abbildungen etc.) einen Umfang von ca. 0,5–1 Seiten haben, bei Gruppenarbeiten einen von ca. 1–2 Seiten.

Modell schätzen:

```
Modell_Mieten <- lm(Kaltemiete ~ 1, data = mieten)
summary(Modell_Mieten)
```

Call:

```
lm(formula = Kaltemiete ~ 1, data = mieten)
```

Residuals:

Min	1Q	Median	3Q	Max
-878.8	-572.3	-225.8	371.2	2171.2

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1328.75	70.11	18.95	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 701.1 on 99 degrees of freedom

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Zusammenfassung

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Quellen und Hilfsmittel

Führen Sie hier die verwendeten Hilfsmittel sowie die verwendete Literatur auf.