

# SCHOOL-SCOUT.DE

Unterrichtsmaterialien in digitaler und in gedruckter Form

**Auszug aus:**

*Daten-Spürnasen auf Spurensuche*

Das komplette Material finden Sie hier:

[School-Scout.de](https://www.school-scout.de)



Daniel Frischemeier & Rolf Biehler

# **Daten-Spürnasen auf Spurensuche**

Datenanalyse in der Grundschule  
mit digitalen Werkzeugen

**Klett | Kallmeyer**

<b>Einleitung</b> .....	<b>8</b>
<b>Teil 1: Fachliche Einführung in die Datenanalyse mit TinkerPlots und CODAP</b> .....	<b>13</b>
<b>1 Einführung in die Nutzung digitaler Werkzeuge zur Datenexploration</b> .....	<b>13</b>
1.1 Grundbegriffe: Standarddatentabelle, Merkmale, Merkmalsträger .....	13
1.2 Typen von Merkmalen .....	16
1.3 Datenkarten als Repräsentationsmittel .....	17
<b>2 Verteilungen kategorialer Merkmale</b> .....	<b>20</b>
2.1 Säulendiagramme und Balkendiagramme .....	20
2.1.1 Säulendiagramme und Balkendiagramme mit Datenkarten .....	20
2.1.2 Säulendiagramme mit TinkerPlots und CODAP: Überblick .....	24
2.1.3 Säulendiagramme mit TinkerPlots erstellen: Sortieren und Stapeln von elektronischen Datenkarten .....	25
2.1.4 Direkte Erzeugung von Säulendiagrammen in TinkerPlots und CODAP .....	31
2.1.5 Balkendiagramme mit TinkerPlots und CODAP .....	35
2.1.6 Lesen und Beschreiben von Säulen- und Balkendiagrammen, der Begriff der Häufigkeit .....	35
2.2 Kreisdiagramme und andere Anteilsdiagramme .....	39
<b>3 Zusammenhänge von zwei kategorialen Merkmalen</b> .....	<b>44</b>
3.1 Den Zusammenhang zweier kategorialer Merkmale darstellen .....	44
3.2 Softwaregestützte Vierfeldertafeln zum Zusammenhang zweier kategorialer Merkmale .....	50
3.2.1 Basisdarstellungen für absolute Häufigkeiten .....	50
3.2.2 Analysefragen zu absoluten Häufigkeiten bei numerischen und graphischen Vierfeldertafeln .....	52
3.2.3 Analysefragen bezogen auf unterschiedliche Anteile bei numerischen und graphischen Vierfeldertafeln .....	52
3.2.4 Darstellungen von Vierfeldertafeln mit eingeblendeten Anteilen .....	54
3.3 Komplexe Zusammenhänge zwischen zwei kategorialen Merkmalen explorieren .....	58

<b>4</b>	<b>Verteilung numerischer Merkmale</b>	<b>66</b>
4.1	Das Wertebalkendiagramm	66
4.2	Das gestapelte Punktdiagramm	70
4.2.1	Varianten des gestapelten Punktdiagramms	70
4.2.2	Erstellen eines gestapelten Punktdiagramms in TinkerPlots und in CODAP	72
4.2.3	Schrittweises Erstellen eines Punktdiagramms aus einem Wertebalkendiagramm in TinkerPlots	74
4.3	Klasseneinteilungen und Histogramme bei einem numerischen Merkmal	75
4.3.1	Histogramme in TinkerPlots	75
4.3.2	Histogramme in CODAP	78
4.3.3	Histogramme mit Wahl der Klasseneinteilung in TinkerPlots	80
4.3.4	Schrittweise Hinführung zum Histogramm in TinkerPlots	80
4.4	Mehrere Histogramme mit unterschiedlichen Klasseneinteilungen	85
4.5	Häufigkeiten in beliebigen Intervallen mit TinkerPlots und CODAP ermitteln	86
4.6	Kennzahlen für Verteilungen eines numerischen Merkmals	89
4.6.1	Überblick über die Kennzahlen für Verteilungen eines numerischen Merkmals	89
4.6.2	Der Median	90
4.6.3	Der Modalwert	95
4.6.4	Das arithmetische Mittel	96
4.6.5	Das arithmetische Mittel und der Median	105
4.7	Ein Hut für die Daten	107
4.7.1	Mittlere Hälfte eines Datensatzes	107
4.7.2	Quartile (Viertelwerte) und ihre Eigenschaften	110
4.7.3	Exkurs zur genauen Berechnung des unteren Quartils ( $Q_1$ ) und des oberen Quartils ( $Q_3$ )	112
<b>5</b>	<b>Vergleich von zwei numerischen Verteilungen</b>	<b>114</b>
5.1	Erste Vergleiche anhand von modalen Klumpen, der mittleren Hälfte und des Hutplots	114
5.2	Vergleiche unter Einbeziehung des Medians und des Boxplots	117
5.3	Vergleiche unter Einbeziehung des arithmetischen Mittelwerts	118
5.4	Abschließende Bemerkungen zum Vergleich von Verteilungen	118

<b>Teil 2: Unterrichtspraktische Anmerkungen und Ideen zur Förderung einer frühen Datenkompetenz</b>	<b>122</b>
<b>6 Statistische Fragen im Unterricht – Bemerkungen aus fachdidaktischer Perspektive</b>	<b>122</b>
6.1 Es stellen sich viele Fragen (Fragebogenfragen vs. Forscher:innenfragen)	123
6.2 Formulierung von Fragebogenfragen, Festlegung der Bedeutung von Merkmalen	125
6.3 Typen von Forscher:innenfragen	126
6.3.1 Ein kategoriales Merkmal	126
6.3.2 Ein numerisches Merkmal	127
6.3.3 Zwei kategoriale Merkmale	128
6.3.4 Ein kategoriales und ein numerisches Merkmal	128
<b>7 Verteilung kategorialer Merkmale im Unterricht</b>	<b>130</b>
7.1 Die Unterrichtssequenz im Überblick	131
7.2 Unterrichtsstunden 1 bis 2: Grundlagen und erste Datenexplorationen	132
7.2.1 Darum geht es in diesen Stunden	133
7.2.2 Möglicher Ablaufplan der Unterrichtsstunden 1 bis 2	134
7.2.3 Unterrichtsleitfaden	134
7.3 Unterrichtsstunden 3 bis 5: Datenanalyse mit Datenkarten	137
7.3.1 Darum geht es in diesen Stunden	137
7.3.2 Möglicher Ablaufplan der Unterrichtsstunden 3 bis 5	138
7.3.3 Unterrichtsleitfaden	138
7.4 Unterrichtsstunden 6 bis 8: Von den Datenkarten zur computergestützten Datenanalyse	146
7.4.1 Darum geht es in diesen Stunden	146
7.4.2 Möglicher Ablaufplan der Unterrichtsstunden 6 bis 8	148
7.4.3 Unterrichtsleitfaden	148
7.5 Unterrichtsstunden 9 bis 12: Statistische Projekte und umfangreiche Datensätze	160
7.5.1 Darum geht es in diesen Stunden	160
7.5.2 Möglicher Ablaufplan der Unterrichtsstunden 9 bis 12	162
7.5.3 Unterrichtsleitfaden	162

<b>8 Verteilung numerischer Variablen im Unterricht</b> .....	<b>178</b>
8.1 Die Unterrichtssequenz im Überblick .....	180
8.2 Unterrichtsstunden 1 bis 2: Lebendige Statistik mit Datenkarten .....	180
8.2.1 Darum geht es in diesen Stunden .....	180
8.2.2 Möglicher Ablaufplan der Unterrichtsstunden 1 bis 2 .....	180
8.2.3 Unterrichtsleitfaden .....	181
8.3 Unterrichtsstunden 3 bis 7: Exploration numerischer Daten in umfangreichen Datensätzen .....	189
8.3.1 Darum geht es in diesen Stunden .....	189
8.3.2 Möglicher Ablaufplan der Unterrichtsstunden 3 bis 7 .....	190
8.3.3 Unterrichtsleitfaden .....	190
8.4 Unterrichtsstunden 8 bis 9: Vergleich numerischer Verteilungen mit digitalen Werkzeugen .....	206
8.4.1 Darum geht es in diesen Stunden .....	206
8.4.2 Möglicher Ablaufplan der Unterrichtsstunden 8 bis 9 .....	206
8.4.3 Unterrichtsleitfaden .....	206
8.5 Unterrichtsstunden 10 bis 13: Numerische Daten in umfangreichen Datensätzen explorieren .....	210
8.5.1 Darum geht es in diesen Stunden .....	210
8.5.2 Möglicher Ablaufplan der Unterrichtsstunden 10 bis 13 .....	210
8.5.3 Unterrichtsleitfaden .....	211
<b>9 Abschließende Bemerkungen</b> .....	<b>216</b>
<b>10 Literaturverzeichnis</b> .....	<b>219</b>

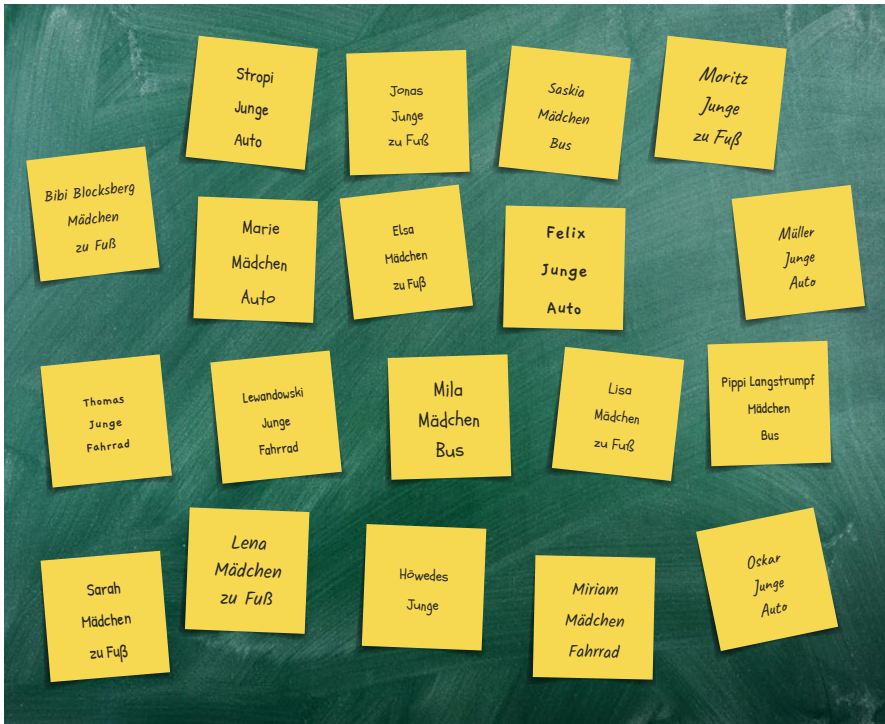


Abbildung 4: Datenkarten (Post-Its) der 19 Schüler:innen der Klasse 4b der Leonhard-Euler-Grundschule ungeordnet an der Tafel, Datensatz D#3

sie zur Schule kommt, rechts außen platziert. Dies führt zu einem *fehlenden Wert*, wie es in der Statistik heißt. Bei der Arbeit mit realen Daten muss immer mit fehlenden Werten gerechnet werden daher ist es sinnvoll, dies auch in einführenden Beispielen zu thematisieren.

Es ist sinnvoll, einem Merkmal, mit dem man arbeitet, auch einen Namen zu geben, z. B. *Fortbewegungsmittel*. Die vielleicht naheliegende Bezeichnung Verkehrsmittel würde hingegen nicht das „Zu Fuß gehen“ umfassen. Man könnte sich im Unterricht darauf einigen, unter Verkehrsmittel auch „zu Fuß“ als Ausprägung zuzulassen. Solche Umdeutungen umgangssprachlicher Begriffe lassen sich oft nicht vermeiden. Es ist typisch für die Statistik, dass die Bedeutungen von Merkmalen genau festgelegt werden müssen und die verwendeten Wörter dadurch oft eine etwas andere Bedeutung bekommen als im Alltag. Wir verwenden hier die schülernahe Merkmalsbezeichnung *Wie\_zur\_Schule*. Die Kurzform „Sortiere/Ordne nach *Wie\_zur\_Schule*“ (im weiteren Verlauf dieses Buches auch synonym mit dem Ausdruck „Trennen“ ver-

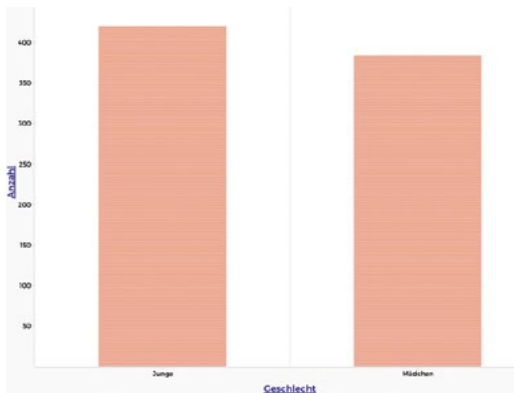


Abbildung 16: Säulendiagramm als Darstellung zur Verteilung des Merkmals *Geschlecht*, Datensatz D#5 in CODAP

möglich, über Trennen und Stapeln den Prozess zum Erstellen des Säulendiagramms schrittweise aufzubauen. Der Vorteil ist, dass man das gewünschte Diagramm in einem Schritt erhält.

Im nächsten Schritt kann dann durch das Kommando *Punkte rechteckig*, das dem *Verschmelzen* bei TinkerPlots entspricht, ein konventionelles Säulendiagramm mit Häufigkeitsachse erzeugt werden (siehe Abbildung 16).

Wie man die Verschmelzung realisiert und auch relative und absolute Häufigkeiten einblendet, zeigt das Menü, das man durch Klicken in den Graphen anzeigen lassen kann. Durch weiteres Klicken werden weitere Optionen sichtbar (Abbildung 17).

Die Reihenfolge der Ausprägungen kann geändert werden, indem die Beschriftung unter den Säulen angefasst und verschoben wird. Die zugehörigen Säulen verschieben sich mit.

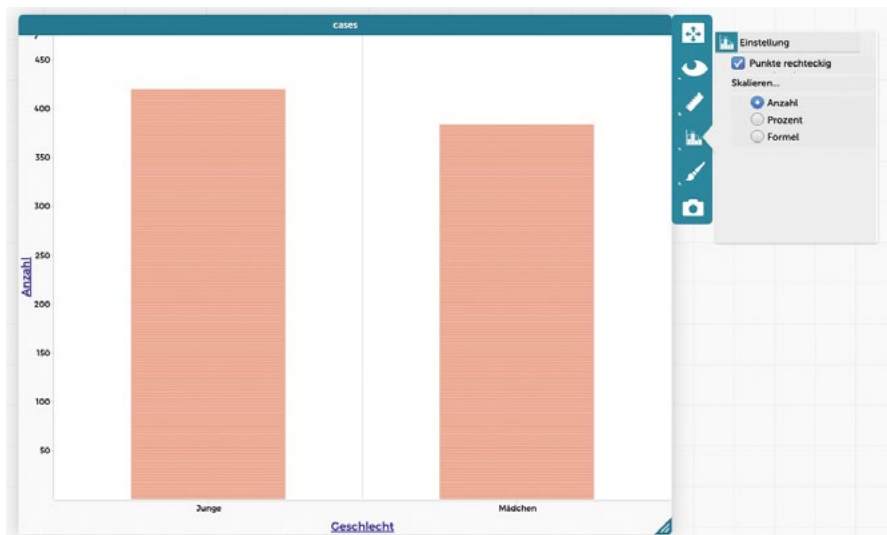


Abbildung 17: Säulendiagramm in CODAP mit Änderungsmenü. Datensatz D#5



## 4 Verteilung numerischer Merkmale

Numerische Merkmale sind beispielsweise *Körpergewicht*, *Körpergröße*, *Taschengeld in Euro*, *Anzahl von WhatsApp-Nachrichten am Tag* oder *Stunden Sport pro Woche*. Die wesentlichen Darstellungen der Verteilung numerischer Merkmale sind *Wertebalkendiagramme* und *gestapelte Punktdiagramme*. Wir erläutern gleich, was wir mit diesen Bezeichnungen meinen.

Ähnlich wie wir es bei den Daten der Klasse 4b (der Leonhard-Euler-Grundschule Klasse 4b – Stichprobe aus dem Datensatz Grundschüler\_innen\_NRW) in Kapitel 1 kennengelernt haben, lassen sich Ausprägungen numerischer

Merkmale auch auf Datenkarten notieren und in Standarddatentabellen repräsentieren. Die Körpergrößen der Kinder der Klasse 4b finden wir in Tabelle 16 notiert.

In TinkerPlots klicken wir für eine erste Darstellung das Merkmal *Körpergröße\_incm* an. Dann öffnen wir einen Graphen: Jeder:r Schüler:in wird wieder durch einen kleinen Kreis repräsentiert, der durch eine Grundfarbe eingefärbt wird je größer die Körpergröße, desto intensiver. Dies sieht man in Abbildung 42, wo wir zuvor noch als Beschriftung den Namen ausgewählt hatten. Wir können durch die im Vergleich zu *Nemo* schwächere Farbsättigung von *Hausaufgabenfreßer* erkennen, dass *Hausaufgabenfreßer* eine geringere *Körpergröße* hat als *Nemo*. Diese Information lässt sich natürlich auch der Datentabelle in TinkerPlots entnehmen.

Name	Körpergröße in cm
Nemo	149
Howedes	139
Lena	148
Frei mit den Tieren	143
Lewandowski	148
Lisa	136
Stropi	138
Sarh	143
Schirwah	142
Fabien	139
Pippi Langstumpf	132
Fran	136
Oskar	137
Miriam Madeleine	137
Marie	138
Manuel	156
Hausaufgabenfreßer	126
Sarah	142
Miriam	134

Tabelle 16: Übersicht über die Körpergrößen der Schüler:innen der Klasse 4b (*Leonhard\_Euler\_GS\_Klasse*) Datensatz D#10

### 4.1 Das Wertebalkendiagramm

Fragen wie „Wer ist das größte/kleinste Kind?“ können mithilfe von Tabelle 16 beantwortet werden. Die Farbabstufungen in Abbildung 42 lassen jedoch

Dieses Prinzip folgt aus dem Hebelgesetz, das auch für Abstände gilt, die nicht ganzzahlig sind. Zugleich entspricht diese Formulierung der ersten Formulierung der Ausgleichseigenschaft, wenn man Abweichungen durch Abstände zum Unterstützungspunkt ersetzt.

Die Deutung des arithmetischen Mittels als Gleichgewichts- oder Schwerpunkt einer Verteilung (auf einer Waage) ist eine inhaltlich über die Ausgleichseigenschaft hinausgehende weitere wichtige Deutung. Das Hebelgesetz wird in der Physik meist noch allgemeiner für beliebige Abstände und Gewichte angewendet. Aber in dem Spezialfall, dass nur ganzzahlige Abstände möglich sind und alle Kugeln das Gewicht 1 haben, ist unsere Formulierung dazu äquivalent. Den Gleichgewichtspunkt nennt man auch Schwerpunkt. Wir betrachten folgenden Datensatz (siehe Abbildung 78):

Man rechnet leicht nach, dass das arithmetische Mittel 6 ist, und in der Tat gilt für die Abstände  $6+4+2+2+1+1+1=3+3+3+4+4$ .

Es gibt einen Punkt, der den Wert 6 hat, also den Abstand 0. Diese Punkte werden bei der Berechnung nicht berücksichtigt.

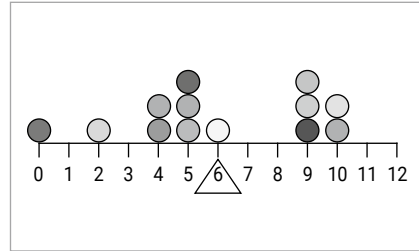


Abbildung 78: Veranschaulichung der Schwerpunkteigenschaft des arithmetischen Mittels (dritte Situation)

i

#### Schwerpunkteigenschaft des arithmetischen Mittelwerts

Wenn man Datenpunkte als gleichschwere Gewichte auf einem Brett mit einem darauf markierten Zahlenstrahl betrachtet, befindet sich das Brett im Gleichgewicht, wenn es im arithmetischen Mittel, dem Schwerpunkt, unterstützt wird.

Die Tatsache, dass für das arithmetische Mittel die Schwerpunkteigenschaft gilt, erklärt auch, warum TinkerPlots für das arithmetische Mittel das Dreieckssymbol (Unterstützungspunkt bei einer Waage) verwendet.

### 4.6.5 Das arithmetische Mittel und der Median

#### Arithmetisches Mittel und Median im Vergleich

Ein wesentlicher Punkt, der den Median vom arithmetischen Mittel unterscheidet, ist die Empfindlichkeit (des arithmetischen Mittels) bzw. die Robustheit (des Medians) gegenüber Ausreißern. Dies soll anhand des ersten

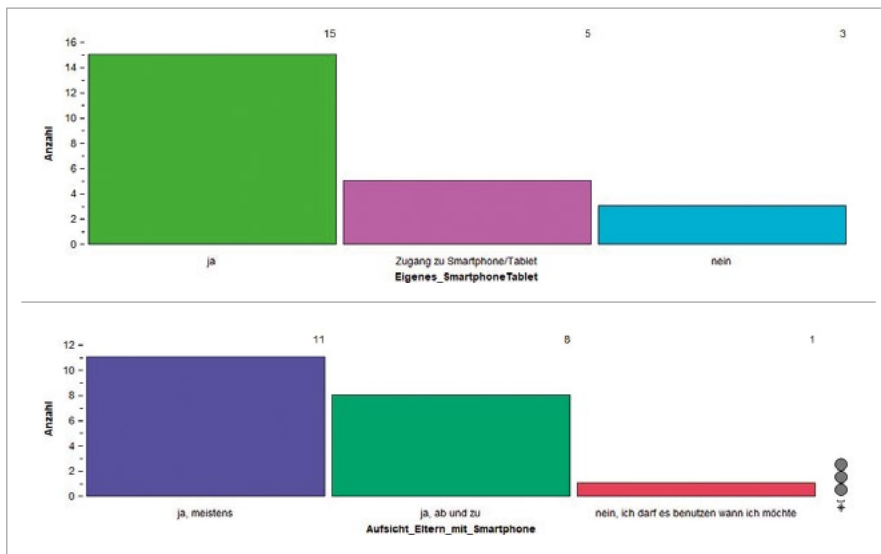


Abbildung 108: Säulendiagramm mit absoluten Häufigkeiten zur Verteilung des Merkmals *Eigenes SmartphoneTablet* – Datensatz *Emmy\_Noether\_GS\_Klasse4a* (oben) und Säulendiagramm mit absoluten Häufigkeiten zur Verteilung des Merkmals *Aufsicht\_Eltern\_mit\_Smartphone* (unten) – Datensatz *D#21* in TinkerPlots

direkt in die Arbeitsfläche von TinkerPlots oder CODAP einzutragen und so festzuhalten und anhand dieser zu präsentieren.

Generell lässt sich aus unseren Unterrichtsprojekten zur Einführung in die Datenanalyse mit TinkerPlots berichten, dass die Schüler:innen nur geringe Einstieghürden bei der Datenanalyse mit TinkerPlots zeigen und die Software zielführend nutzen können, um die Daten im Hinblick auf ihre Fragestellungen zu explorieren. Insbesondere die vorgeschalteten Aktivitäten zur Datenanalyse mit Lebendiger Statistik und die Arbeit mit Datenkarten un-

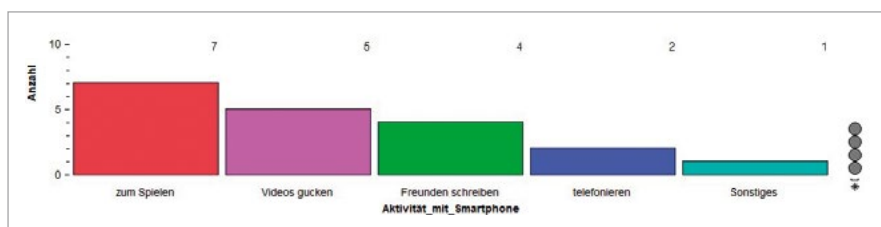


Abbildung 109: Säulendiagramm mit absoluten Häufigkeiten zur Verteilung des Merkmals *Aktivität\_mit\_SmartphoneTablet*- Datensatz *D#21* in TinkerPlots

- Wenn man den Ausreißer von 75 ignoriert (Zahl kann ja nicht stimmen), streuen die Jungen von 0 bis 15 und die Mädchen von 0 bis 12, kein sehr großer Unterschied.
- Es unterscheiden sich die Mädchen und die Jungen im Median um eine Stunde pro Tag (Jungen höher).
- Die mittlere Streubreite ist größer bei den Jungen, die Gruppe ist unterschiedlicher als die der Mädchen.
- Die Jungen haben eher mehr Freizeit am Tag als die Mädchen (Median ist um 1 größer, Box der Jungen endet bei 4, bei Mädchen bei 3).

	Jungen	Mädchen
Median	3	2
arithmetisches Mittel	–	–
Modalwert	1	1
Wertebereich (Minimum, Maximum)	0 bis 75	0 bis 12
Spannweite	75	12
Lage der Box/des Hutes, mittlere Hälfte	1 bis 4	1 bis 3
Lage des unteren Viertels	0 bis 1	0 bis 1
Lage des oberen Viertels	4 bis 75	3 bis 12
mittlere Streubreite	3	2
Stellen mit besonders großer Häufigkeit;	1, 2, 3, 4, 5	1, 2, 3, 4
sonstige Auffälligkeiten, Form der Verteilungen	--	--

Tabelle 38: Möglicher Verteilungsvergleichsplan der Spürnasen zur Vergleichssituation in Abbildung 143

Die Dokumentation der Ergebnisse kann dabei wie beim Spürnasenmodul zur Exploration der Verteilung kategorialer Variablen realisiert werden – dabei muss die Lehrkraft je nach Rahmenbedingungen ihre Entscheidungen und Abwägungen treffen:

- Die Schüler:innen können die Textfelder in TinkerPlots und CODAP nutzen, um ihre Ergebnisse und Erkenntnisse direkt in der Software festzuhalten.
- Es kann eine Textverarbeitungsdatei wie z. B. ein Word-Dokument genutzt werden. Dort müssen dann die TinkerPlots oder CODAP-Grafiken als Screenshots eingefügt werden.
- Es können mit Blick auf ein Datenposter einzelne Grafiken ausgedruckt werden, die dann zu einem Datenposter verarbeitet werden.

#### Phase 4: Vorstellen der Ergebnisse

In der letzten Phase des Spürnasenmoduls stellen die Schüler:innen ihr jeweiliges Projekt anhand einer der drei oben genannten Alternativen vor. Diese Phase kann nun ähnlich wie in Kapitel 7 gedacht werden – dabei kann die Präsentation der Ergebnisse entweder über den Projektor, das interaktive Whiteboard oder am ausgedruckten Datenposter erfolgen.

# SCHOOL-SCOUT.DE

Unterrichtsmaterialien in digitaler und in gedruckter Form

**Auszug aus:**

*Daten-Spürnasen auf Spurensuche*

Das komplette Material finden Sie hier:

[School-Scout.de](https://www.school-scout.de)

