**Datenanalyse mit der Software TinkerPlots**

****

Inhalt

[Einführung in die Datenanalyse mit TinkerPlots 1](#_Toc26429508)

[Ihr erster Arbeitsauftrag 8](#_Toc26429509)

[Verteilung eines kategorialen Merkmals: Wie ist die Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW? 11](#_Toc26429510)

[Verteilung eines kategorialen Merkmals: Wie ist die Verteilung des Merkmals Wie\_zur\_Schule im Datensatz Grundschulen\_NRW? 17](#_Toc26429511)

[Untersuchung des Zusammenhangs zwischen zwei kategorialen Merkmalen: Inwiefern unterscheiden sich die Kinder die auf dem Dorf, in der Stadt oder in der Großstadt wohnen hinsichtlich der Art wie sie zur Schule kommen? 20](#_Toc26429512)

[Verteilung eines numerischen Merkmals : Wie ist die Verteilung des Merkmals „Anzahl\_Spiele\_auf\_Smartphone/Tablet“ im Datensatz Grundschulen\_NRW? 24](#_Toc26429513)

[Untersuchung des Zusammenhangs zwischen einem kategorialen und einem numerischen Merkmal: Inwiefern unterscheiden sich die Dritt- und die Viertklässler hinsichtlich der Anzahl ihrer Spiele auf dem Smartphone/Tablet? 28](#_Toc26429514)

[Nützliche Tipps zum Graph sowie Informationen zur Datentabelle 32](#_Toc26429515)

[Ihr zweiter Arbeitsauftrag 33](#_Toc26429516)

Dieses Kapitel soll eine Einführung in die Datenanalyse mit TinkerPlots beinhalten und mögliche Datenexplorationen mit TinkerPlots vorstellen. Es soll somit zum einen Einblick geben wie die Software TinkerPlots zur Datenanalyse eingesetzt und genutzt werden kann und zum anderen auch eine Hilfestellung und einen Leitfaden darstellen, sich in die Datenanalyse mit TinkerPlots einzuarbeiten. Arbeitsaufträge sind in diesem Kapitel in blauer Schrift gekennzeichnet, in grüner Schrift finden sich Verweise auf weitere Kapitel dieses Buches, in denen die hier in dieser Einführung thematisierten Inhalte noch einmal vertieft aufgriffen und mit unterrichtspraktischen Anregungen angereichert werden.

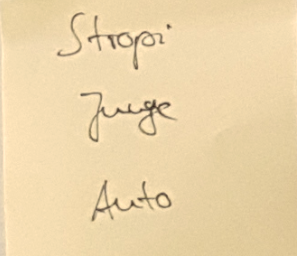
# Einführung in die Datenanalyse mit TinkerPlots

Eine wesentliche Grundphilosophie der Software TinkerPlots ist, dass sie die Daten in Form sogenannter Datenkarten speichert und an das Arbeiten und an die Operationen mit Datenkarten (Trennen, Ordnen, Stapeln) angeknüpft, so dass Graphiken im weiteren Sinne basierend auf den Grundoperationen „Stapeln“, „Ordnen“ und „Trennen“ erstellt werden können. Wir skizzieren diese Arbeitsweise an einem Beispiel, ohne zu sehr ins unterrichtspraktische Detail zu gehen (siehe aber Baustein 4 dieses Bandes).

Man stelle sich folgende Situation vor: Die Klasse 4b möchte sich am Beginn des Schuljahres besser kennenlernen und beschließt eine kurze klasseninterne Umfrage durchzuführen. Schnell haben Anna, Ben, Chris und Doro Ideen für mögliche Fragestellungen:

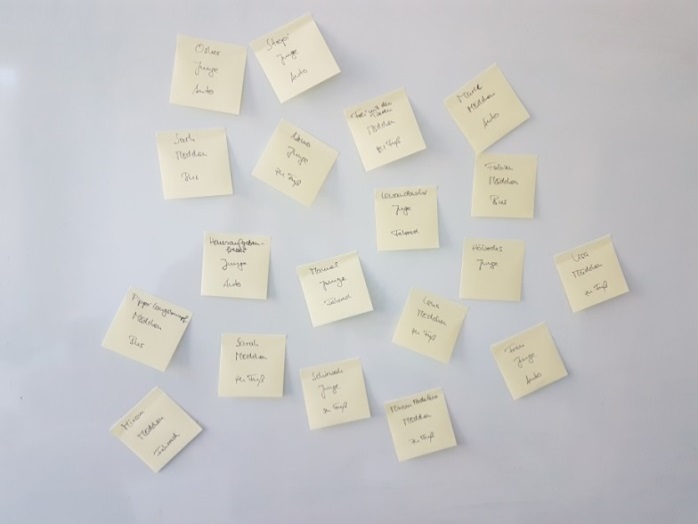
* Wie viele essen gerne Nudeln in unserer Klasse?
* Wie viele von uns haben ein Haustier?
* Nutzen die Jungen ihr Smartphone mehr als die Mädchen?
* Wie kommen die Kinder der Klasse 4b zur Schule?

Die Schülerinnen und Schüler der Klasse 4b wollen zunächst wissen, wie die Klasse zur Schule kommt. Man einigt sich im Vorfeld, dass es die folgenden Möglichkeiten gibt: Auto, Bus, Bahn, Fahrrad oder zu Fuß. Wenn mehr als ein Merkmal zutrifft, verabredet man, dass man das Verkehrsmittel angibt, in welchem man die meiste Zeit verbracht hat. Am besten man einigt sich auch darauf, dass man das Verkehrsmittel angibt, mit dem man an Morgen der Umfrage gekommen ist. Hierbei können die Schülerinnen und Schüler bereits Wichtiges zu statistischen Umfragen lernen. Jede Schülerin und jeder Schüler der Klasse 4b bekommt einen Klebe-Zettel (siehe z.B. Abbildung 1). Darauf schreibt sie/er einen Fantasienamen (um die Anonymität zu gewährleisten), sein Geschlecht, und wie er zur Schule kommt.



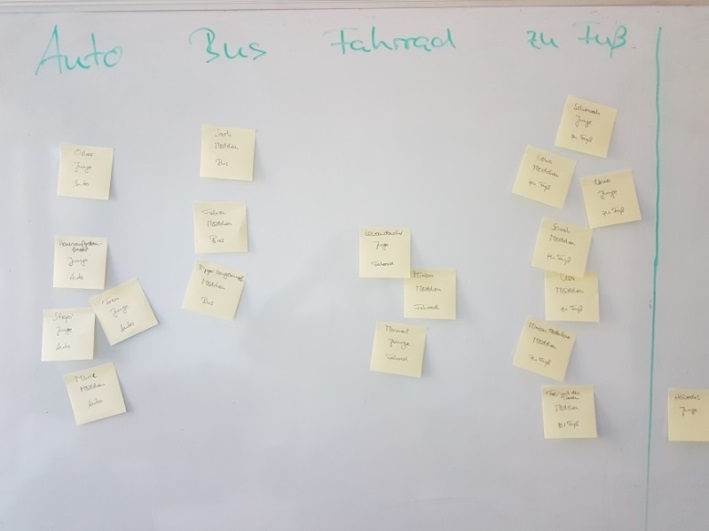
**Abbildung 1: Datenkarte von Stropi**

Nun können wir zunächst Fragestellungen („Wie viele Jungen und Mädchen gibt es in der Klasse 4b?“ oder „Wie kommen die Kinder der Klasse 4b zur Schule?“ oder „Mit welchem Verkehrsmittel kommen die meisten?“) anhand einfacher Operationen mit unseren Datenkarten beantworten. Wir versammeln die Schülerinnen und Schüler im Kinokreis und legen die Datenkarten aller Schülerinnen und Schüler (n=21) in die Mitte bzw. heften sie an die Tafel (Abbildung 2).



**Abbildung 2: Datenkarten der Klasse 4b ungeordnet an der Tafel**

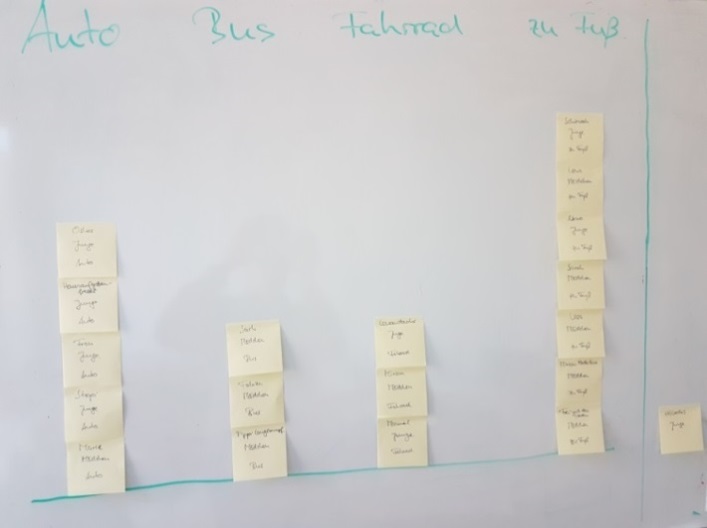
Um die Frage „Wie kommen die Kinder der Klasse 4b zur Schule?“ beantworten zu können, hilft ein Sortieren nach den verschiedenen Verkehrsmitteln (Abbildung 3).



**Abbildung 3: Datenkarten der Klasse 4b geordnet nach Verkehrsmittel**

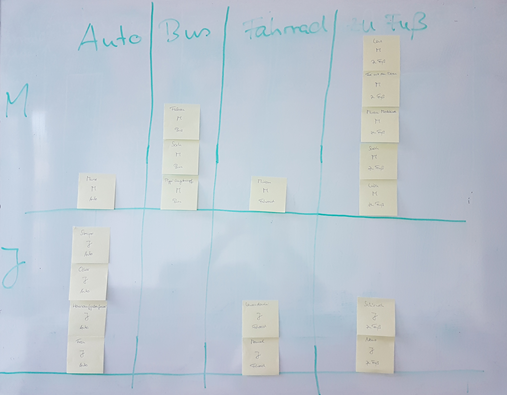
Wir haben dann ganz links die Datenkarten von Kindern, die mit dem Auto zur Schule gebracht werden und z.B. ganz rechts die Datenkarten von den Kindern, die zu Fuß zur Schule kommen.

Wir können uns durch Abzählen oder (nächster Ordnungsdurchgang) Stapeln der Datenkarten einen besseren Überblick über die verschiedenen Häufigkeiten verschaffen, wie die Kinder der Klasse 4b zur Schule kommen. Durch das Stapeln der Datenkarten haben wir dann mithilfe unserer Klebe-Zettel schon eine Art Säulendiagramm erstellt (Abbildung 4), in dem man besser die unterschiedlichen Häufigkeiten erkennen kann.



**Abbildung 4: Datenkarten-Säulendiagramm mit den Antworten zur Frage, wie man zur Schule kommt**

Nun hätten wir das unter Umständen auch noch mithilfe einer Strichliste hinbekommen. Die Klebezettel kann man nun aber umsortieren, indem man die Jungen von den Mädchen trennt (z.B. alle Mädchen oberhalb der Jungen anordnet und die entstandenen Lücken bei den Jungen durch Zusammenschieben schließt.



**Abbildung 5: Datenkarten-Säulendiagramm mit den Antworten zur Frage, wie man zur Schule kommt (getrennt nach dem Merkmal „Geschlecht“)**

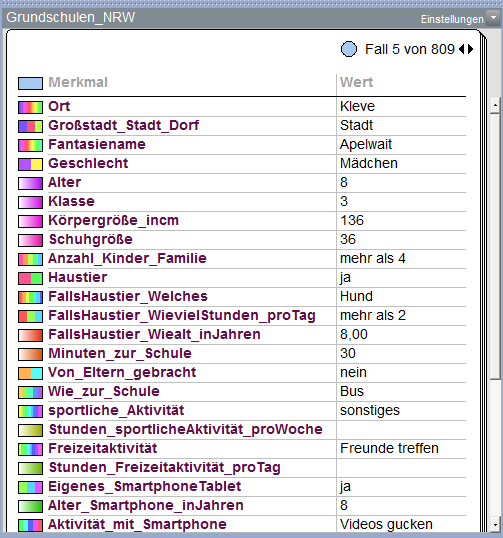
Dann kann man schauen, ob und welche Unterschiede es zwischen Mädchen und Jungen dabei gibt. Hätte man noch mehr Merkmale auf den Karten notiert, dann könnte man noch andere Fragestellungen durch Umsortieren beantworten. Das Ganze kann man sich natürlich auch mit anderen Merkmalen und zu anderen Fragestellungen vorstellen. Die einzigen Operationen, die man benötigt, sind im Wesentlichen das „Ordnen“, „Trennen“ und „Sortieren“ der Datenkarten.

Den Beispielen liegen statistische Fachbegriffe zugrunde: Geschlecht, Körpergröße und Verkehrsmittel sind *Merkmale*, die jeweils verschiedene Merkmalsausprägungen haben können. Aus den geordneten Datenkarten kann man z.B. die *Häufigkeiten* entnehmen, mit der bestimmte Ausprägungen vorkommen. Schaut man auf eine Darstellung als Ganzes, so kann man die auch als Darstellung der (Häufigkeits-) *Verteilung* eines Merkmals ansehen.

Was ist aber nun, wenn wir viel mehr als nur zwei oder drei Merkmale erhoben haben und was ist wenn wir mehrere Schülerinnen und Schüler befragen wollen, sagen wir unsere ganze Schule? Wie können wir die Daten dann effektiv auswerten? Natürlich kann man genauso vorgehen, wie es im obigen Abschnitt beschrieben wurde, jedoch wäre der Sortierungs- oder auch Trennprozess sehr langwierig. An dieser Stelle kann es nützlich sein, sich von einer Computersoftware unterstützen zu lassen. Die Software TinkerPlots dokumentiert Daten in Form von Datenkarten und bietet dieselben Operationen (Trennen, Stapeln, Ordnen) an, die wir auch bereits mit Datenkarten realisieren konnten.

*Der Einstieg in die Datenanalyse mit der Software TinkerPlots*

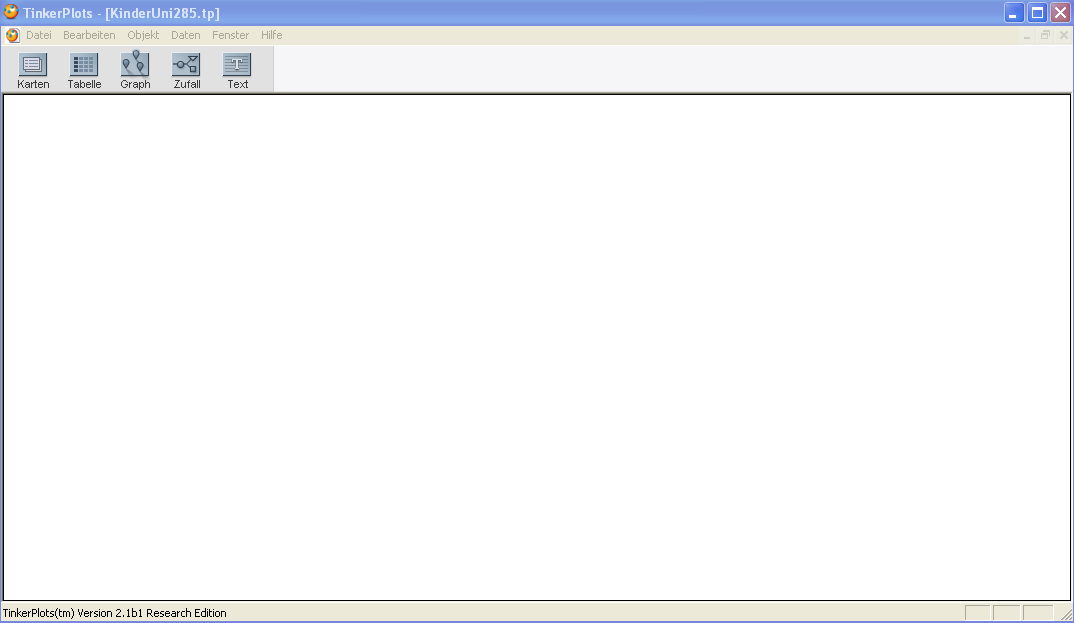
Wir betrachten als einführendes Beispiel in die Datenanalyse mit TinkerPlots die Analyse des Datensatzes „Grundschulen\_NRW“. Dieser wurde im Rahmen einer Bachelorarbeit im Jahr 2017 erhoben und enthält die Angaben zum Freizeit- und Medienverhalten von 809 Grundschülerinnen und Grundschülern in NRW. Die Stichprobe ist nicht repräsentativ für eine klar umrissene *Grundgesamtheit*. Diese Ergebnisse der Befragung sind in TinkerPlots, wie in unserem obigen Einführungsbeispiel, als *Datenkarten* gesammelt.



**Abbildung 6: Datenkarte von Apelwait im Datensatz Grundschulen\_NRW in TinkerPlots**

Schülerinnen und Schüler können eine Datenkarte erstmal so verstehen, dass links die jeweilige Frage in Kurzfassung steht, rechts die entsprechende Antwort der Schülerin, deren Datenkarte gerade gezeigt wird. Die Schülerin, die sich den Fantasienamen „Apelwait“ gegeben hat, ist der *Fall* Nr. 5 (siehe Abbildung 6). Die Fragen werden der Software auch als *Merkmal* bezeichnet, die Antwort als *Wert* (des Merkmals). Diese statistischen Fachbegriffe können hier im Kontext an Beispielen eingeführt und benutzt werden.

Zunächst einige einführende Sätze zur Bedienung der Software TinkerPlots. Wenn Sie das Programm TinkerPlots starten, so finden Sie folgenden Bildschirm vor (siehe Abbildung 7):



**Abbildung 7: Arbeitsfläche von TinkerPlots**

Ähnlich, wie man es in zum Beispiel in Windows gewohnt ist, gibt es eine Menüleiste, die die wichtigsten grundlegenden Funktionen, wie Datei öffnen, Datei speichern, Rückgängig-Funktion, Hilfemenü, usw. … beinhaltet. Ebenfalls gibt es das Arbeitsfeld und die Hauptleiste, in der man Datenkarten, Datentabellen, Graphen, Zufallsmaschinen und Textfelder in das Arbeitsfeld (weiße Fläche) ziehen kann.

 Dieses Icon steht für einen Datenkartenstapel. In ihm werden alle Fälle abgelegt. Jeder Fall bekommt seine eigene Datenkarte. Die verschiedenen Merkmale stehen untereinander und sind auf jeder Datenkarte gleich. Durch Betätigung des Knopfes erscheint ein Datenkartenstapel (siehe Abbildung 6)

 Alle Daten können auch in einer Tabelle angezeigt werden. Ein Fall befindet sich dann in einer Zeile, eine Zeile entspricht also einer Datenkarte des Datenkartenstapels. Die Merkmale der einzelnen Fälle stehen in den Spalten. Änderungen in der Tabelle führen automatisch zu Änderungen in den Datenkarten und umgekehrt.

Durch Betätigung des Knopfes und gleichzeitiges Ziehen in die Arbeitsfläche erscheint eine Datentabelle (siehe Abbildung 8)



**Abbildung 8: Datentabelle zum Datensatz Grundschulen\_NRW in TinkerPlots**

 Mit diesem Knopf wird eine Grafikfläche mit Daten erzeugt, aus denen man dann weitere Graphiken aus den Daten erzeugen. Dies ist das Hauptauswertungsobjekt in TinkerPlots.

 Hinter diesem Objekt verbirgt sich die Zufallsmaschine von TinkerPlots. Diese wird für alle Simulationen benötigt. Eine Tabelle mit den erzeugten Daten wird automatisch erstellt, sobald die Zufallsmaschine gestartet wird.

 Für Erklärungen, Ergebnisse, Hinweise oder sonstigen Text stehen Textfelder zur Verfügung.

In dieser Einführung wollen wir vor allem auf die Funktionen der

* Datenkarten
* Graphen
* Datentabelle

im Rahmen der Datenanalyse eingehen.

Zurück zu unserem Beispieldatensatz „Datensatz\_Grundschulen\_NRW“ und zu einem ersten Arbeitsauftrag (in blau).

# Ihr erster Arbeitsauftrag

Laden Sie die Software TinkerPlots von [www.tinkerplots.com/get](http://www.tinkerplots.com/get) herunter (Hilfestellungen zum Download und zur Installation der deutschen Version finden sich im Anhang).

Öffnen Sie den Beispieldatensatz „Datensatz\_Grundschulen\_NRW.tp“ in TinkerPlots. Nehmen Sie sich 15 Minuten Zeit, um erste Auswertungen mit TinkerPlots (anhand der Fragestellungen auf Seite 11-12) in diesem Datensatz eigenständig durchzuführen und machen Sie die Schritte, die auf den folgenden Seiten erklärt sind, nach. Die Anleitungen im Folgenden können Ihnen dabei helfen. Außerdem kann Ihnen die beiliegende Variablenliste helfen, einen Überblick über die verschiedenen Variablen im Datensatz zu erlangen.

In der Abbildung 6 sehen wir exemplarisch die Datenkarte der Schülerin Apelwait. Oben rechts kann man zwischen den verschiedenen Fällen (hier in diesem Falle sind das die verschiedenen 809 Kinder) hin- und herspringen. Dort können wir auch erkennen, dass wir 809 Datenkarten vorliegen haben.

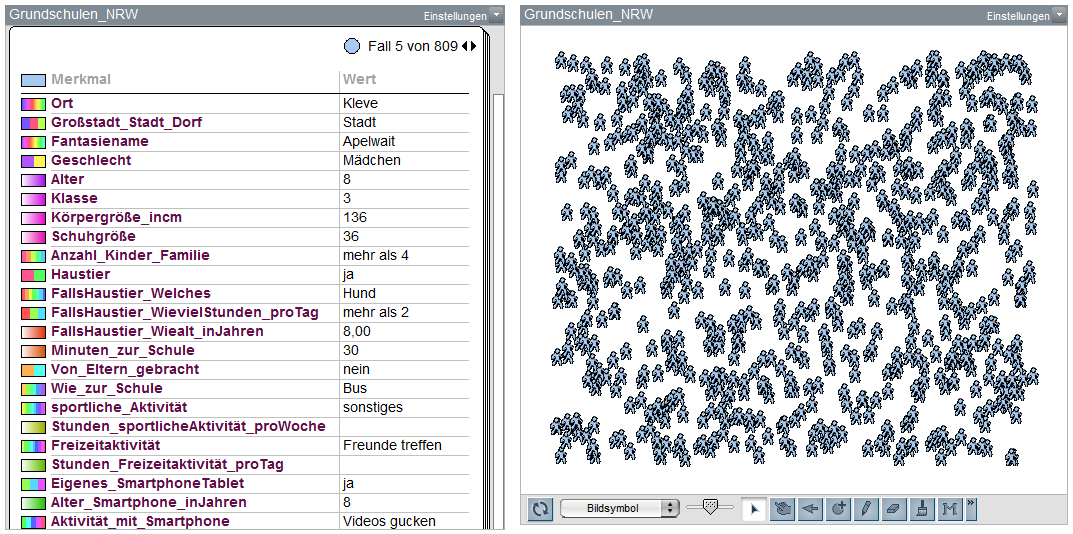
Wir können der Datenkarte (in Abbildung 6) einige Informationen entnehmen, u.a., dass die Schülerin Apelwait aus Kleve kommt, 8 Jahre alt ist, in die Klasse 3 geht und ca. 136 cm groß ist. Durch Klicken auf die Pfeiltasten rechts oben, lassen sich die Datenkarten durchblättern und Informationen über die anderen Schüler anzeigen.

Weiterhin erkennen wir, dass TinkerPlots in den Datenkarten in der Spalte Merkmal eine Farbkodierung angibt. Bei Geschlecht werden 2 Farben angezeigt, die für die beiden möglichen Ausprägungen „Mädchen“ und „Junge“ stehen. Bei Alter, Klasse und Körpergröße\_incm und allen anderen Merkmalen, die Zahlen als Werte haben können wird nur eine Farbe gewählt, die unterschiedliche Intensitäten haben kann, je höher die Zahl bei der Ausprägung, desto intensiver die Farbe. Der Typ der Farbkodierung ist also bei sogenannten *kategorialen Merkmalen* (die nur eine überschaubare Anzahl verschiedener Ausprägungen haben können) anders als bei den so genannten *numerischen* Merkmalen.

Die Daten auf den Datenkarten können verändert und ergänzt werden. Ebenfalls können Merkmale oder Fälle ergänzt werden.

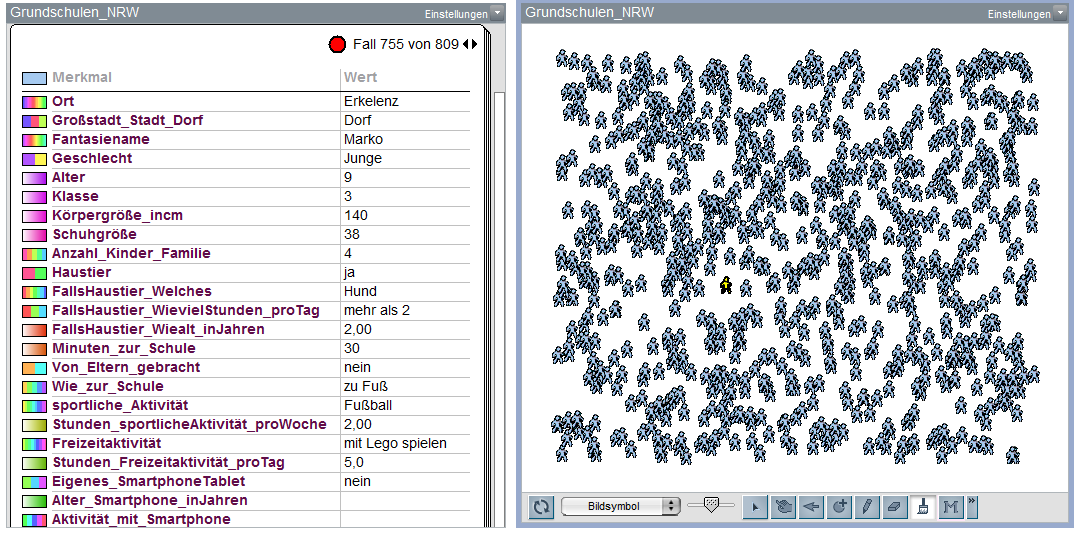
Ein wesentliches Element der Software zum Explorieren von Daten ist der „Graph“. Dieser ermöglicht eine Auswertung der Daten in Form von Sortieren und Anordnen (Ordnen, Trennen und Stapeln) der Datenkarten.

Um den Graph zu einem Datenkartensatz zu erstellen, wählt man in der Arbeitsleiste das Icon  und zieht dieses per „Drag & Drop“ zusätzlich zu den Datenkarten in den Arbeitsbereich. Es öffnet sich ein Fenster mit verschiedenen Personen-Symbol-Figuren (siehe Abbildung 9 rechts).



**Abbildung 9: Datenkartenstapel (links) und Graph (rechts) in TinkerPlots**

Jedes Symbol im Graph steht für eine Grundschülerin bzw. einen Grundschüler aus dem Datensatz. Wählen wir beispielsweise eine der „Personen“ im Graph aus, so können wir anhand der aktuellen Datenkarte die Daten zum Kind ablesen (Abbildung 10).



**Abbildung 10: Datenkartenstapel (links) und Graph (rechts) in TinkerPlots (Schüler Marko markiert)**

Wir können die Daten von „Marko“ in der Datenkarte ablesen: der Schüler Marko kommt aus Erkelenz, ist 9 Jahre alt, geht in die dritte Klasse und ist 140cm groß, etc.

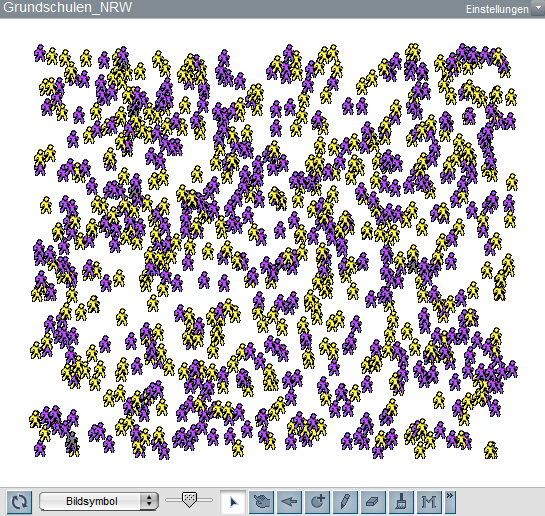
Nun wollen wir im Folgenden aber nicht einzelne Fälle betrachten, sondern die Verteilung von Merkmalen und Zusammenhänge zwischen Merkmalen untersuchen. Wir haben oben erfahren, dass man zwischen kategorialen Merkmalen und numerischen Merkmalen unterscheiden kann. Zunächst kann man dabei Verteilungen eines kategorialen oder eines numerischen Merkmals betrachten. Dann kann man in einem weiteren Schritt auch Zusammenhänge zwischen zwei kategorialen Merkmalen und zwischen einem kategorialen und einen numerischen Merkmal explorieren. Diese Idee soll unsere Exploration des Datensatzes *Grundschulen\_NRW.tp* und die Einführung in die Software TinkerPlots im Folgenden leiten – wir stellen die Datenexploration mit TinkerPlots anhand der folgenden Fragen vor. Die Bezeichnung in der Klammer drückt in abstrakter Weise aus, um welche(n) Typ(en) von Merkmalen es geht und wie viele Merkmale in einer Frage angesprochen werden. Diese Typisierung ist wichtig, da sich jeweils andere Auswertungsgraphiken anbieten.

1. Wie ist die Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW? (Verteilung eines kategorialen Merkmals)
2. Wie ist die Verteilung des Merkmals Wie\_zur\_Schule im Datensatz Grundschulen\_NRW? (Verteilung eines kategorialen Merkmals)
3. Inwiefern unterscheiden sich die Kinder, die auf dem Dorf, in der Stadt oder in der Großstadt wohnen, hinsichtlich der Art wie sie zur Schule kommen? (Untersuchung des Zusammenhangs zwischen zwei kategorialen Merkmalen)
4. Wie ist die Verteilung des Merkmals „Anzahl\_Spiele\_auf\_Smartphone/Tablet“ im Datensatz Grundschulen\_NRW? (Verteilung eines numerischen Merkmals)
5. Inwiefern unterscheiden sich die Dritt- und die Viertklässler hinsichtlich der Anzahl ihrer Spiele auf dem Smartphone/Tablet? (Untersuchung des Zusammenhangs zwischen einem kategorialen und einem numerischen Merkmal)

Beginnen wir mit der ersten Fragestellung…

## Verteilung eines kategorialen Merkmals: Wie ist die Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW?

Um diese erste Fragestellung zu explorieren, klicken wir zunächst das Merkmal Geschlecht im Datenkartenstapel an, so färben sich die Symbole der Jungen und die der Mädchen in unterschiedlichen Farben ein (Abbildung 11). Man weiß zunächst noch nicht, zu welchem Geschlecht welche Farbe gehört.



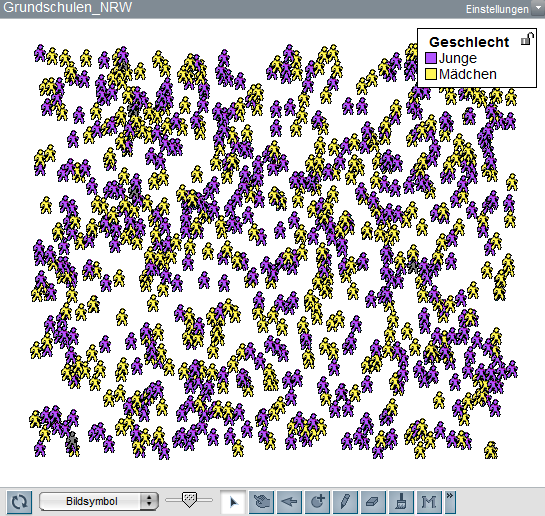
**Abbildung 11: TinkerPlots Graph zur Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW ungeordnet**

Ebenso kann man mit anderen Merkmalen verfahren. Will man nun zwischen Jungen und Mädchen unterscheiden können, oder besser gesagt –auf einfachem Wege- herausfinden, welches Geschlecht welcher Farbe zugeordnet ist, so kann sich der Legende-Funktion (Schalter ganz rechts) bedienen. Diese findet sich in der Arbeitsleiste zum Graphen (Abbildung 12).



**Abbildung 12: Arbeitsleiste zum Graphen in TinkerPlots**

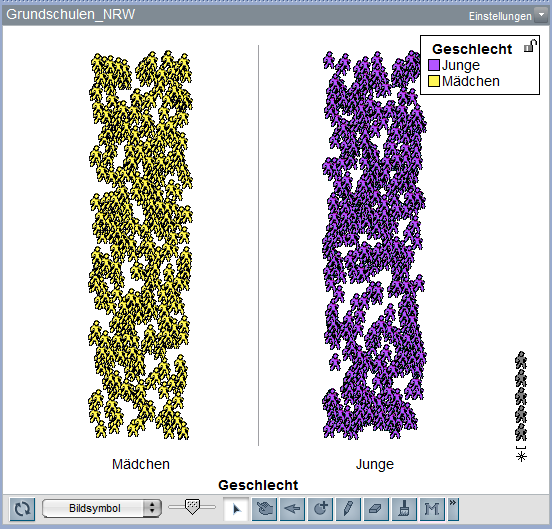
Durch Betätigung des Schalters „Legende“ (Abbildung 12, ganz rechts) erscheint im Graph rechts oben ein Fenster (siehe Abbildung 13), in welchem wir ablesen können, dass die gelben „Personen“ Mädchen und die „Personen“ in lila Jungen sind.



**Abbildung 13: TinkerPlots Graph zur Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW ungeordnet (mit Legende eingeblendet)**

Nun wollen wir Ordnung in die „Personen“ bringen. Dieses machen wir auf einer ersten Stufe, indem wir die Jungen und die Mädchen –ähnlich wie bei dem Arbeiten mit Datenkarten- „trennen“, um uns einen Überblick über die Verteilung des Merkmals „Geschlecht“ in diesem Datensatz zu verschaffen.

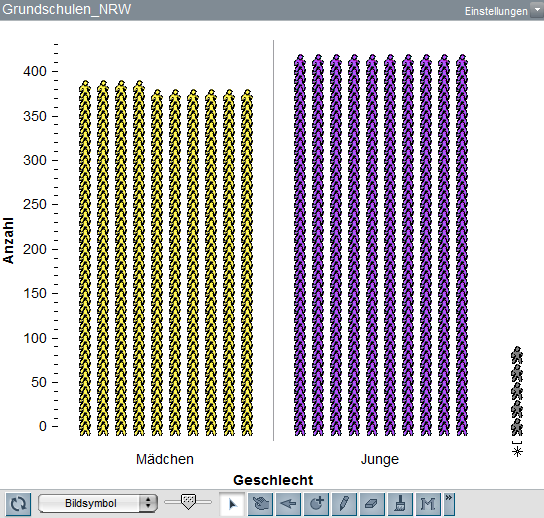
Dieses können wir mit der Funktion „Trennen“ realisieren: entweder durch Betätigung des Schalters  oder aber durch “Ziehen” eine der „Personen” im Graphen mit der Maus (siehe Abbildung 14).



**Abbildung 14: TinkerPlots Graph zur Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW.tp nach Mädchen und Jungen getrennt (mit Legende eingeblendet)**

Nun haben wir zwei Gruppen erstellt: In der linken Gruppe befinden sich die Mädchen, in der rechten die Jungen (die fünf grauen Personen-Symbole am rechten Rand repräsentieren Befragte, die keine Angabe zum Merkmal „Geschlecht“ gemacht haben. In der Statistik nennt man die „fehlende Werte“). Die „Personen“ sind in beiden Klassen leider noch sehr unsortiert und noch lässt sich nicht sagen, ob mehr Jungen oder mehr Mädchen in dem Datensatz sind.

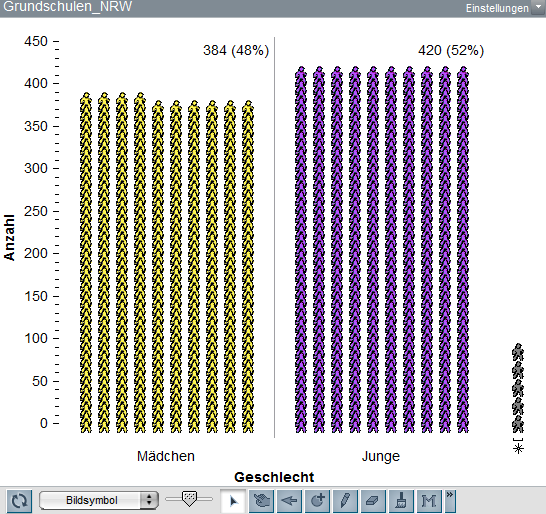
Durch die Funktion „Senkrecht Stapeln“  wollen wir die „Personen“ für eine bessere Übersicht „übereinander legen“ (Abbildung 15) – ähnlich wie wir es beim Aneinanderlegen der Datenkarten gemacht haben.



**Abbildung 15: TinkerPlots Graph zur Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW nach Mädchen und Jungen getrennt und gestapelt**

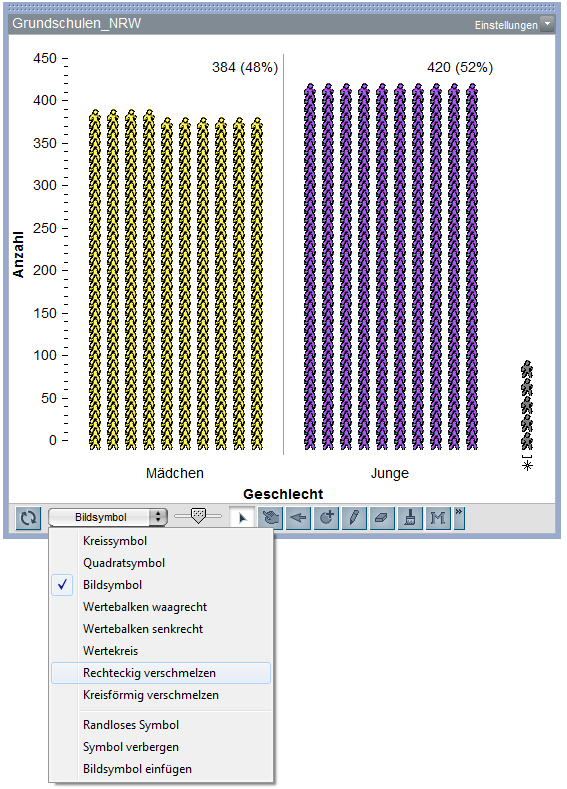
Wir haben nun ein Piktogramm erstellt, mit dem wir unsere Frage nach der Geschlechterverteilung in diesem Datensatz nun zumindest qualitativ („es gibt mehr Jungen in dem Datensatz als Mädchen“) beantworten können. An dieser Stelle ist anzumerken, dass aufgrund der Darstellung gegebenenfalls Irritationen entstehen können, da die Anzahl-Achse samt Anordnung der Personen-Symbole nicht den Konventionen entspricht. Während in der Klasse der Jungen alle zehn Personen-Symbole-Säulen gleichviele Elemente besitzen, sind die einzelnen Säulen in der Klasse der Mädchen nicht gleich hoch: die vier linken Säulen sind anzahlmäßig um 1 höher als die restlichen sechs Säulen. Das Ablesen der Anzahlen an der entsprechenden Skala ist somit erschwert und kann für Irritationen und Ungenauigkeiten sorgen. Man muss ferner erklären, dass sich die Anzahl nicht auf die Höhe einer einzelnen Säule bezieht, sondern auf die Gesamtzahl, so sind ja in den 10 Säulen für Jungen jeweils etwa nur 40 enthalten. Die Problematik kann ein wenig umgangen werden, indem man die Größe der Symbole verkleinert: dieses reduziert dann die Anzahl der Personen-Symbole-Säulen in beiden Klassen. Wollen wir die Anzahl-Frage quantitativ beantworten, so bedienen wir uns der „Anzahl“-Achse links bzw. der Funktion „Anzahl“  in der Arbeitsleiste des Graphen.

Diese liefert absolute und relative Häufigkeiten für einzelne Klassen (siehe Abbildung 16) und ermöglicht Aussagen wie „420 der Befragten sind Jungen und 384 sind Mädchen“ oder (mit Blick auf einen späteren Einsatz der Software, z.B. in der Sekundarstufe I) „52% der Befragten sind Jungen und 48% der Befragten sind Mädchen“. In der Grundschule würde man sich natürlich die Prozentwerte nicht anzeigen lassen.



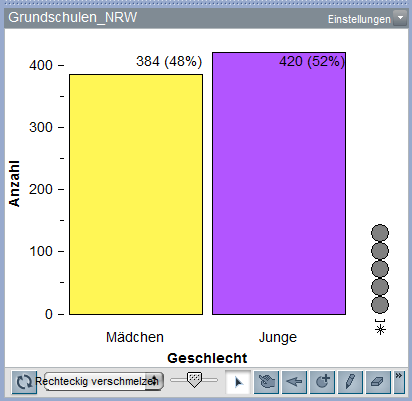
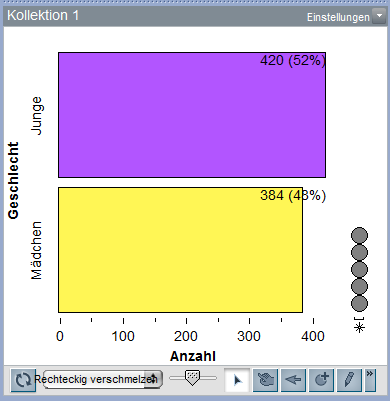
**Abbildung 16: TinkerPlots Graph zur Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW nach Mädchen und Jungen getrennt und gestapelt (mit absoluten und relativen Häufigkeiten)**

Will man nun vom Piktogramm zu eher konventionellen Graphiken, wie zum Säulendiagramm überschwenken, so können wir auch dieses realisieren, indem wir das Funktionsmenü des Graphen nutzen (Abbildung 17).



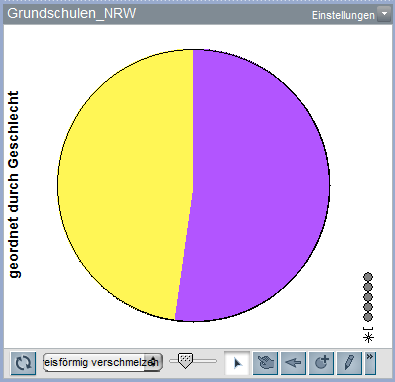
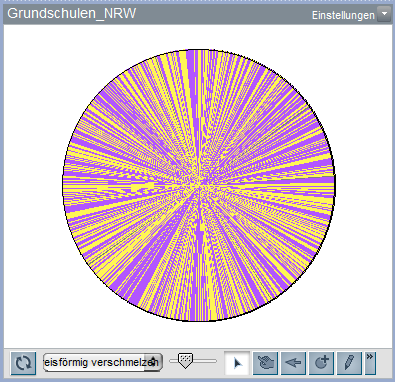
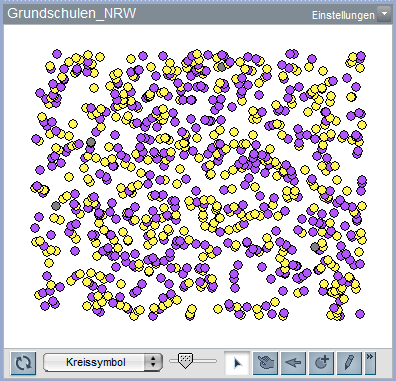
**Abbildung 17: TinkerPlots Graph zur Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW\_tp nach Mädchen und Jungen getrennt und gestapelt (mit absoluten und relativen Häufigkeiten) und Funktionsmenü des Graphen**

Durch Betätigen der Funktion „Rechteckig verschmelzen“, erhalten wir ein Säulendiagramm (Abbildung 18 links). Durch Drehung um 90° im Uhrzeigersinn (in TinkerPlots: Vertauschung der Achsen) erhalten wir ein Balkendiagramm zur Verteilung des Merkmals Geschlecht (Abbildung 18 rechts).

**Abbildung 18: Säulendiagramm zur Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW (links) und entsprechendes Balkendiagramm (rechts)**

Um ein weiteres für den Mathematikunterricht in der Primarstufe relevantes Diagramm, ein Kreisdiagramm, zu erstellen, mischen wir den Graph mit  einmal durch (Abb. 19 links), wählen dann „Kreisförmig verschmelzen“ (Abb. 19 Mitte). Jedem Fall ordnet die Software einen gleich großen Anteil zu, was für die Schülerinnen und Schüler verständlich ist, auch ohne dass sie über einen Bruchbegriff oder Winkelbegriff verfügen müssten. Das „Ordnen“ dann (Abb. 19 rechts) führt dann zum üblichen Kreisdiagramm, in dem sich aber die Schülerinnen und Schüler noch die einzelnen „Tortenstücke“ vorstellen können, auch wenn sie nicht explizit sichtbar sind.

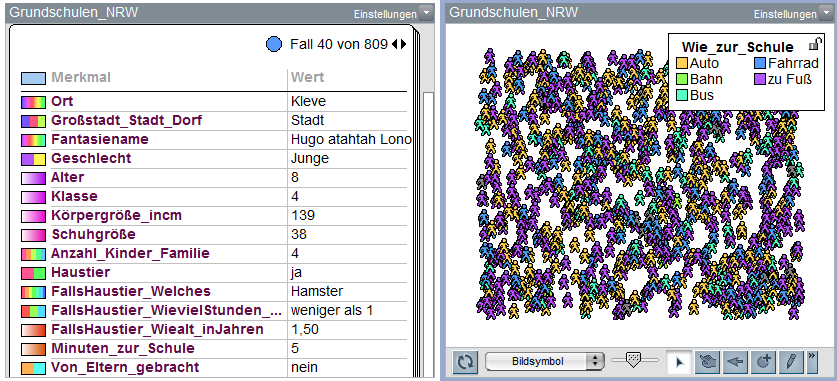


**Abbildung 19: Auf dem Weg zum Kreisdiagramm zur Verteilung des Merkmals Geschlecht im Datensatz Grundschulen\_NRW**

Anhand des Kreisdiagrammes (Abbildung 19, rechts) lässt sich sagen, dass etwas mehr als die Hälfte der Kinder im Datensatz Jungen sind. In TinkerPlots ist es nicht möglich, die Sektoren mit den entsprechenden relativen Häufigkeiten zu versehen (was in der Primarstufe ja auch irrelevant wäre). Vielmehr soll das Kreisdiagramm dazu einladen, qualitative Aussagen zu den Anteilen zu treffen, die Schülerinnen und Schüler intuitiv als Anteile an der Kreisfläche vergleichen können, ohne dass sie einen quantifizierten Anteilsbegriff benötigten. Nehmen wir nun ein kategoriales Merkmal, das mehr als nur zwei Ausprägungen hat, z.B. das Merkmal „Wie\_zur\_Schule“ mit den Ausprägungen „Auto“, „Bahn“, „Bus“, „Fahrrad“ und „zu Fuß“ und betrachten die folgende Fragestellung:

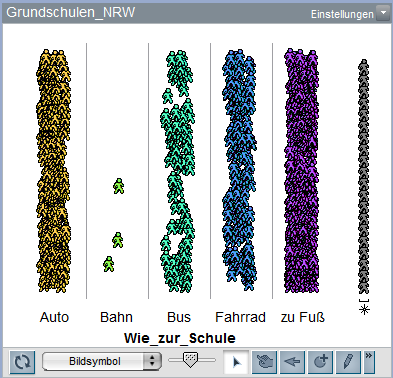
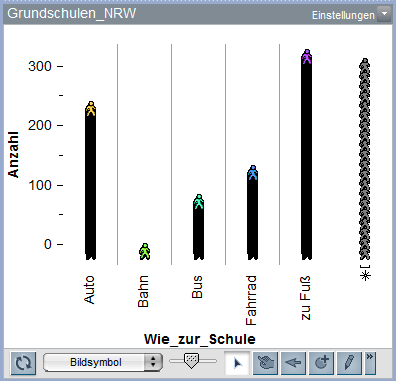
## Verteilung eines kategorialen Merkmals: Wie ist die Verteilung des Merkmals Wie\_zur\_Schule im Datensatz Grundschulen\_NRW?

Wir wählen das Merkmal „Wie\_zur\_Schule“ in den Datenkarten (Abbildung 20, links) aus und sehen, dass sich die „Personen“ gemäß der Ausprägungen einfärben (Abbildung 20, rechts).



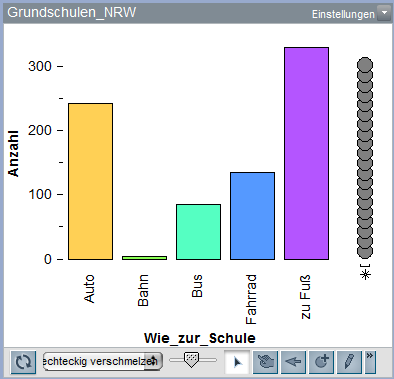
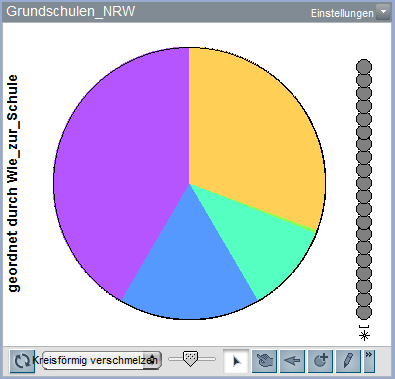
**Abbildung 20: Datenkarte in TinkerPlots (links) und Graph (eingefärbt nach Ausprägungen des Merkmals Wie\_zur\_Schule) rechts**

Wenn wir uns beispielsweise ein Bild darüber machen wollen, wie unsere befragten Schülerinnen und Schüler zur Schule kommen, so müssen wir wieder versuchen, Ordnung in die “Personen” zu bekommen – dieses geschieht wieder durch Anwendung von „Trennen“ (Abb. 21, links) und „Stapeln“ (Abb. 21, rechts).

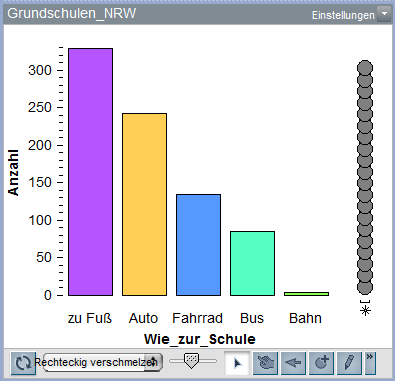
**Abbildung 21: TinkerPlots Graph zur Verteilung des Merkmals Wie\_zur\_Schule im Datensatz Grundschulen\_NRW nach Ausprägungen getrennt (links: nach Ausprägungen getrennt, rechts: nach Ausprägungen getrennt und gestapelt)**

Über die Funktion “Rechteckig verschmelzen” können wir dann wieder ein Säulendiagramm erstellen (Abbildung 22 links) oder aber auch (dann über die Funktion „Kreisförmig verschmelzen“) ein Kreisdiagramm (Abbildung 22 rechts):

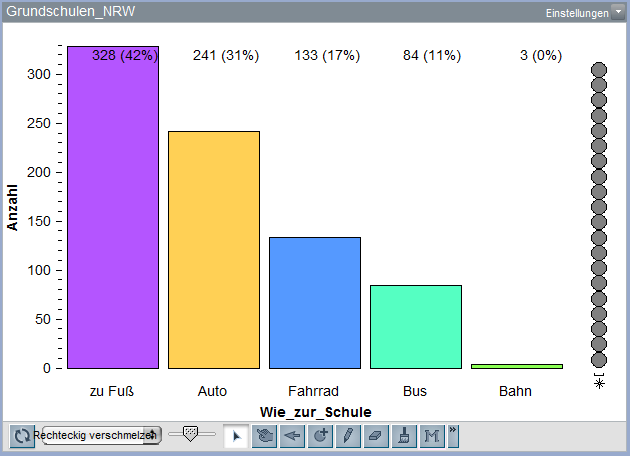
**Abbildung 22: Säulendiagramm zur Verteilung des Merkmals Wie\_zur\_Schule (links) und Kreisdiagramm zur Verteilung des Merkmals Wie\_zur\_Schule (rechts) im Datensatz Grundschulen\_NRW**

Eine weitere Modifikation beim Säulendiagramm kann vorgenommen werden, indem die Säulen verschoben werden. Will man beispielsweise etwas darüber aussagen, wie die Grundschülerinnen und Grundschüler in diesem Datensatz zur Schule kommen so kann es sich zunächst anbieten, die Säulen in Bezug auf ihre Häufigkeit entweder aufsteigend oder absteigend anzuordnen (Abbildung 23).



**Abbildung 23: Säulendiagramm (geordnet) zur Verteilung des Merkmals Wie\_zur\_Schule im Datensatz Grundschulen\_NRW**

Wir sehen, dass die meisten (mehr als 300) der Befragten zu Fuß zur Schule kommen, ebenfalls kommen viele der Befragten (fast 250) mit dem Auto, nur sehr wenige der Befragten nutzen die Bahn für den Schulweg. Diese qualitativen Aussagen lassen sich weiter verbessern, indem man absolute und relative Häufigkeiten einblendet (Abbildung 24).



**Abbildung 24: Säulendiagramm (geordnet) zur Verteilung des Merkmals Wie\_zur\_Schule im Datensatz Grundschulen\_NRW (mit absoluten und relativen Häufigkeiten eingeblendet)**

Nun können wir präziser sagen, dass 328 der Befragten (also etwas weniger als die Hälfte) zu Fuß zur Schule kommen, 241 der Befragten werden mit dem Auto zur Schule gebracht, etc., und nur drei Schüler nutzen die Bahn, um zur Schule zu kommen. Auch wenn in der Grundschule relative Häufikeiten tabu sind, kann z.B. die Nutzung des Kreisdiagramms in Abbildung 22 (rechts) eine erste Annäherung an besondere Anteile wie Hälfte oder Viertel beitragen. Beispielsweise könnte man bei Betrachtung der Kreissektoren sagen, dass etwas weniger als die Hälfte der Kinder zu Fuß zur Schule kommen und etwas mehr als ein Viertel der Kinder mit dem Auto zur Schule gebracht werden. Diese Angaben lassen sich unter Nutzung der Anzahl/Anteil-Funktion, die wir oben bereits vorgestellt haben, weiter präzisieren (siehe Abbildung 24). Hier lässt sich dann sagen, dass 42% der Kinder zu Fuß zur Schule kommen, 31% mit dem Auto, etc. An dieser Stelle bietet es sich an, weitere Nachforschungen vorzunehmen, weil sich weitere Fragen wie z.B. Was ist mit den drei Schülerinnen und Schülern, die mit der Bahn zur Schule kommen – haben diese drei Schüler vielleicht Gemeinsamkeiten? Gibt es Zusammenhänge zu anderen Merkmalen? Um dies zu beantworten, können hier unter anderem die Datenkarten der jeweiligen Schülerinnen und Schüler entsprechend auf mögliche Gemeinsamkeiten untersucht werden.

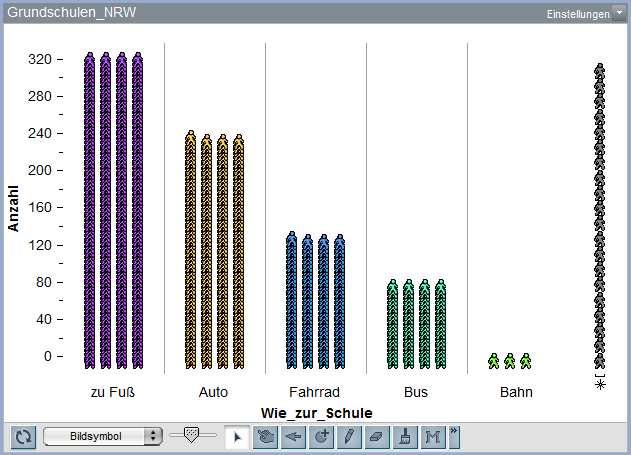
Ausblick: Weiterführung und Vertiefung sowie unterrichtspraktische Anregungen zur Einführung in die Datenanalyse mit TinkerPlots in Jahrgangsstufe 3/4 finden sich in Baustein 4: *Spielerisches Erlernen von Datenanalyse in der Grundschule - Über Lebendige Statistik und Datenkarten zur Datenanalyse mit TinkerPlots*

Bis jetzt haben wir uns bei unserer Analyse auf Fragestellungen beschränkt, die auf die Verteilung eines Merkmals abzielen. Wir gehen nun einen Schritt weiter und zeigen, wie man den Zusammenhang zwischen zwei kategorialen Merkmalen in TinkerPlots untersuchen kann. Die entsprechende Fragestellung lautet:

## Untersuchung des Zusammenhangs zwischen zwei kategorialen Merkmalen: Inwiefern unterscheiden sich die Kinder die auf dem Dorf, in der Stadt oder in der Großstadt wohnen hinsichtlich der Art wie sie zur Schule kommen?

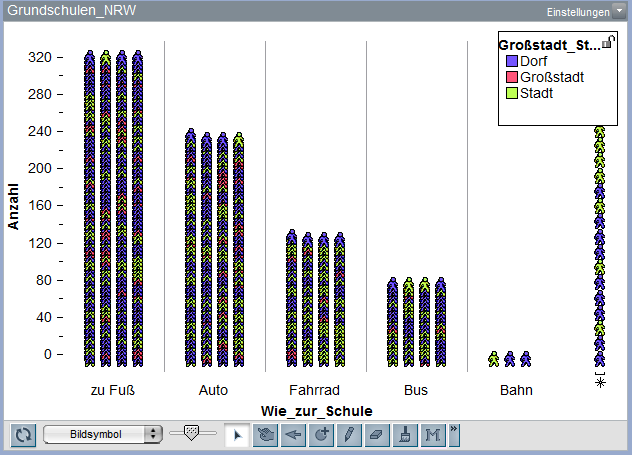
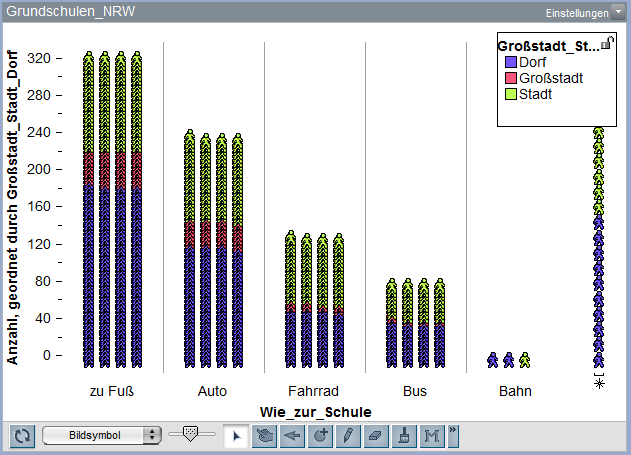
Im Unterricht könnten zu dieser Frage schon mal Vermutungen artikuliert werden, bevor man die Daten auswertet.

Nehmen wir hier noch einmal unser Piktogramm (siehe Abbildung 25), das uns die Verteilung des Merkmals „Wie\_zur\_Schule“ für den gesamten Datensatz zeigt:



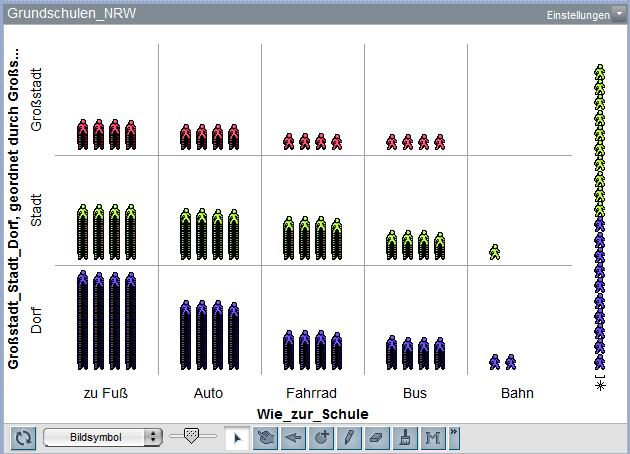
**Abbildung 25: TinkerPlots Graph zur Verteilung des Merkmals Wie\_zur\_Schule (getrennt und gestapelt) im Datensatz Grundschulen\_NRW**

Wir wählen in den Datenkarten das Merkmal „Großstadt\_Stadt\_Dorf“ aus: nun werden die Dorfkinder lila, die Stadtkinder grün und die Großstadtkinder rot eingefärbt (Abbildung 26, links). Auch hier kann “Ordnen” helfen (Abbildung 26, rechts). In diesem Fall ordnet die Software innerhalb der einzelnen Ausprägungen (zu Fuß, Auto, Fahrrad, Bus, Bahn) nach Dorf, Stadt und Großstadt.

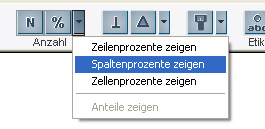
**Abbildung 26: TinkerPlots Graph zur Verteilung des Merkmals Wie\_zur\_Schule (getrennt und gestapelt) im Datensatz Grundschulen\_NRW – eingefärbt nach dem Merkmal Großstadt\_Stadt\_Dorf (links: ungeordnet, rechts: geordnet)**

Die Funktion „Trennen“ haben wir bis jetzt nur genutzt, um Klassen auf der „x-Achse“ zu generieren. Wir können aber auch in Richtung der „y-Achse“ trennen, d.h. in diesem Fall die Dorf-, Stadt- und Großstadtkinder trennen (siehe Abbildung 27).



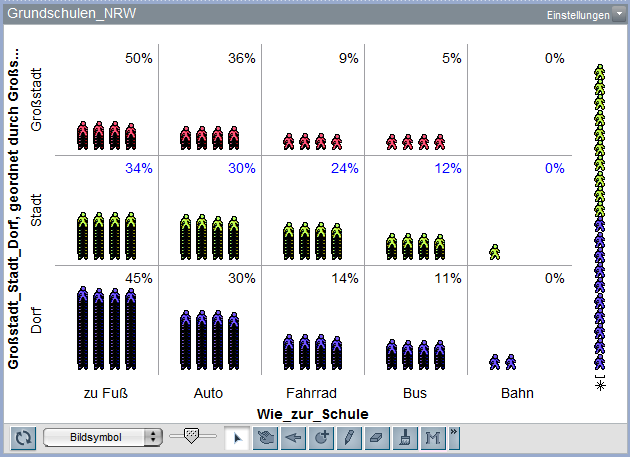
**Abbildung 27: TinkerPlots Graph zur Untersuchung des Zusammenhangs zwischen den Merkmalen Wie\_zur\_Schule und Großstadt\_Stadt\_Dorf im Datensatz Grundschulen\_NRW**

Nun haben wir drei Verteilungen des Merkmals “Wie\_zur\_Schule”: jeweils für die Dorfkinder, die Stadtkinder und für die Großstadtkinder. Durch Betrachtung der Form der jeweiligen Verteilung lassen sich jetzt schon erste Aussagen zu unserer Fragestellung treffen, noch genauer können wir sie mit unserer Anzahlfunktion  formulieren. Wir können nun mit absoluten Häufigkeiten aber auch mit relativen Häufigkeiten argumentieren. Die vergleichende Argumentation mit absoluten Häufigkeiten aufgrund der unterschiedlichen Gruppengrößen kann zu Fehlurteilen verleiten. Rein optisch kann man zeilenweise den Anteil der Fußgänger vergleichen, in der Stadt scheint der Anteil größer zu sein als im Dorf, die Anteile für Satdt und Großstadt sind optisch schwer zu vergleichen. Die Nutzung relativer Häufigkeiten (wenn auch nicht realisierbar im Mathematikunterricht der Primarstufe) kann weitere Einblicke bieten. Es lassen sich sogenannte Zeilen-, Spalten- und Zellenprozente mithilfe von TinkerPlots berechnen. Im Untermenü zu „Anzahl“ findet man hier die Möglichkeit sich verschiedene relative Häufigkeiten anzuzeigen (siehe Abbildung 28).



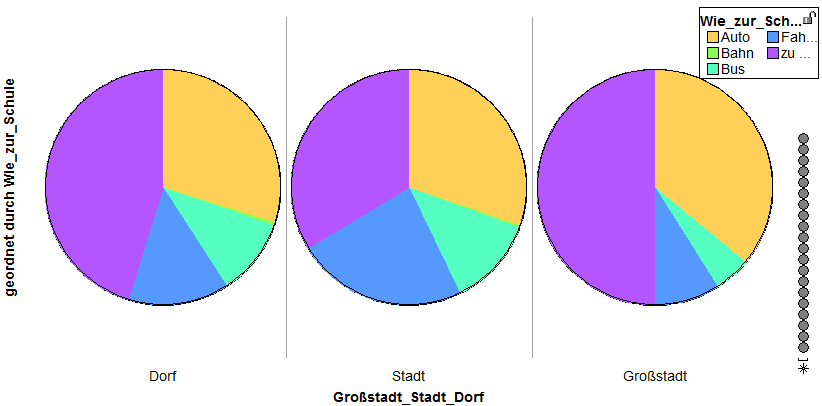
**Abbildung 28: Optionsmenü zur Auswahl der verschiedenen Prozente in TinkerPlots**

Für die Exploration unserer Fragestellung eignen sich vornehmlich in diesem Fall die Zeilenprozente (siehe Abbildung 29). Zeilenprozente bedeutet hier, dass man die Dorf-, Stadt- und Großstadtkinder als eigenständige Teilgruppe betrachtet und somit die Möglichkeit gegeben ist, z.B. in der Teilgruppe der Dorfkinder den Anteil der Kinder zu bestimmen, die zu Fuß zur Schule gehen (das sind in diesem Fall 45%, siehe Abbildung 29). Ähnlich könnte man mit Blick auf Abbildung 29 Aussagen treffen wie, dass z.B. 5% der Großstadtkinder mit dem Bus fahren.



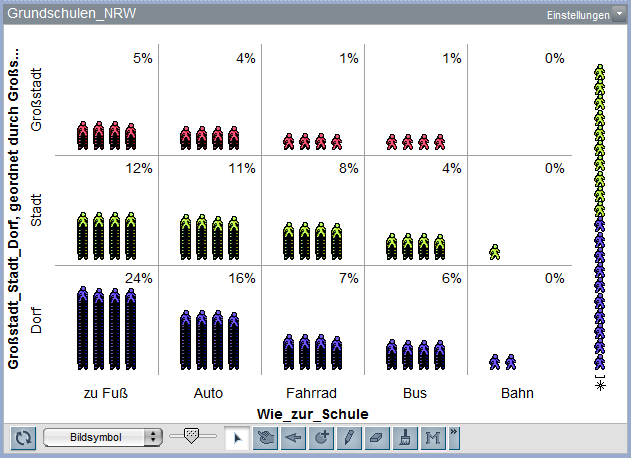
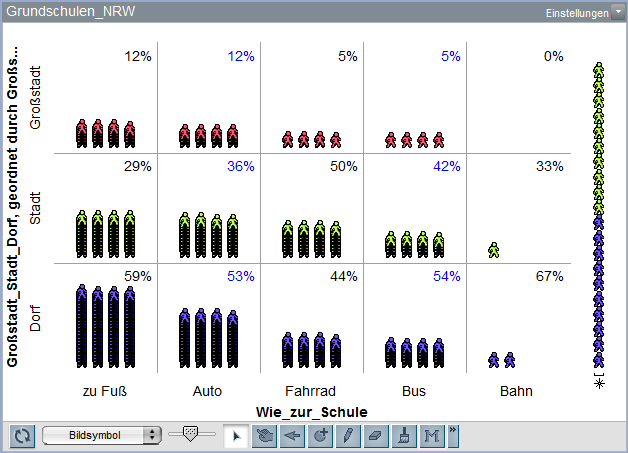
**Abbildung 29: TinkerPlots Graph zur Untersuchung des Zusammenhangs zwischen den Merkmalen Wie\_zur\_Schule und Großstadt\_Stadt\_Dorf im Datensatz Grundschulen\_NRW mit Zeilenprozente**

Insgesamt können wir mit Blick auf unsere Fragestellung sagen, dass es insbesondere Unterschiede bei der Fahrrad- und Busnutzung sowie bei den Fußgängern gibt. Wir sehen, dass 45% der Dorfkinder zu Fuß zur Schule gehen, in der Gruppe der Stadtkinder ist dieser Anteil mit 34% kleiner, in der Gruppe der Großstadtkinder ist er am größten: die Hälfte der befragten Großstadtkinder gehen zu Fuß zur Schule. Auch beim Fahrradfahren finden sich Unterschiede: während fast ein Viertel der Stadtkinder mit dem Fahrrad zur Schule fährt, beträgt dieser Anteil bei den Dorfkindern nur 14% und bei den Großstadtkindern sogar nur 9%. Da in der Grundschule kein Prozentbegriff vermittelt wird, sollte man einen qualitativen Anteilsvergleich hier mit Kreisdiagrammen machen (Abbildung 30).



**Abbildung 30: TinkerPlots Graph zur Untersuchung des Zusammenhangs zwischen den Merkmalen Wie\_zur\_Schule und Großstadt\_Stadt\_Dorf im Datensatz Grundschulen\_NRW mithilfe von Kreisdiagrammen**

Es lassen sich auch entsprechend die Spalten- (siehe Abbildung 31 links) und Zellenprozente (Anteil an der Gesamtgruppe) (siehe Abbildung 31 rechts) berechnen – diese sind bei der Exploration der obigen Fragestellung allerdings nicht zielführend.

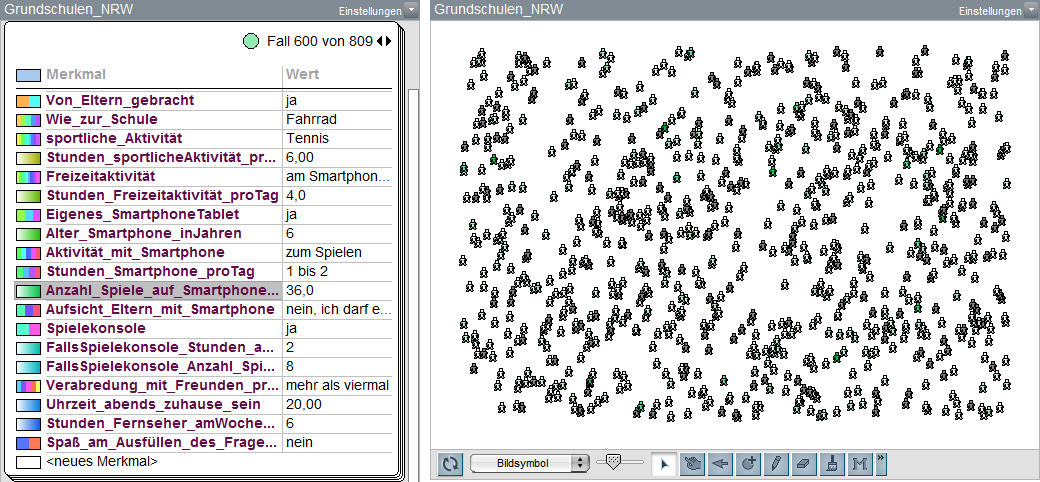


**Abbildung 31: TinkerPlots Graph zur Untersuchung des Zusammenhangs zwischen den Merkmalen Wie\_zur\_Schule und Großstadt\_Stadt\_Dorf im Datensatz Grundschulen\_NRW mit Spaltenprozenten (links) und Zellenprozenten (rechts)**

Auf weitere Details wollen wir bei dieser elementaren Einführung nicht weiter eingehen, sondern nun zur Analyse von Verteilungen numerischer Merkmale übergehen.

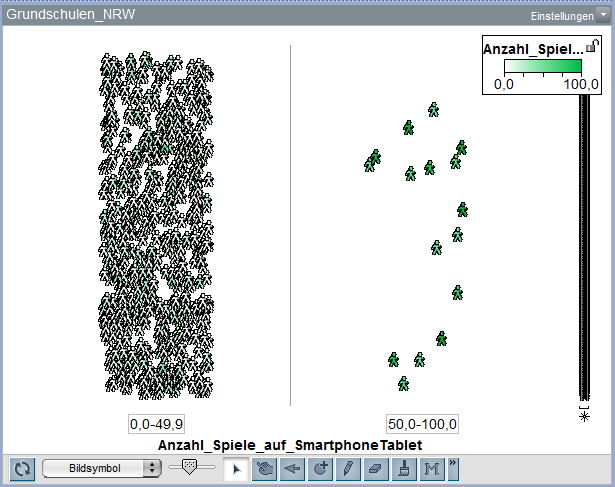
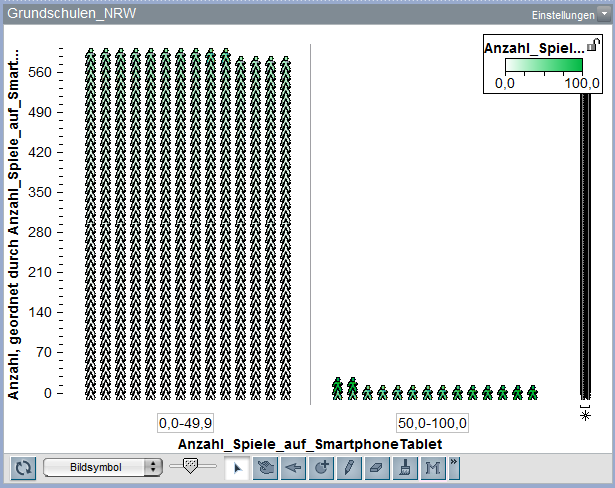
## Verteilung eines numerischen Merkmals : Wie ist die Verteilung des Merkmals „Anzahl\_Spiele\_auf\_Smartphone/Tablet“ im Datensatz Grundschulen\_NRW?

Wir haben bis hierhin nur die Verteilung bzw. Verteilungen von kategorialen Merkmalen (Geschlecht, Wie\_zur\_Schule) betrachtet: Wie sieht es aber mit numerischen Merkmalen aus? Wir wählen das Merkmal Anzahl\_Spiele\_auf\_Smartphone im Datenkartenstapel aus (siehe Abbildung 32).



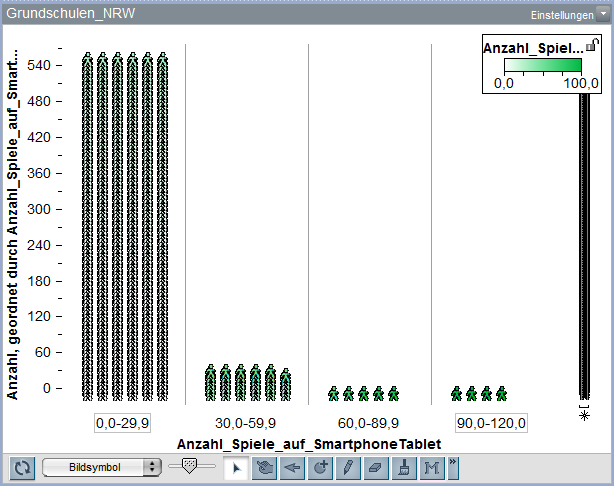
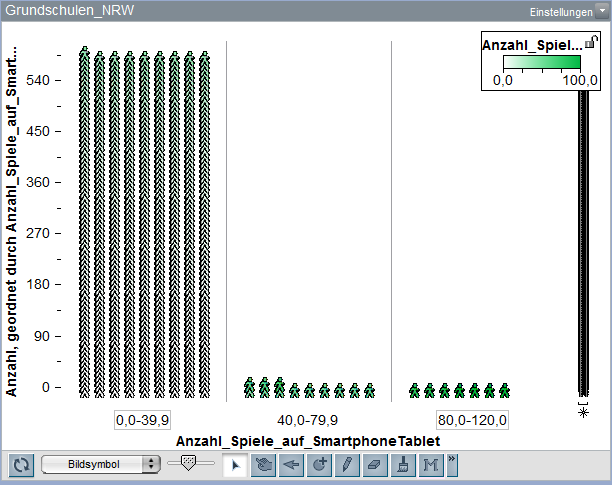
**Abbildung 32: Datenkarte und Graph in TinkerPlots bei der Exploration der Verteilung eines numerischen Merkmals im Datensatz Grundschulen\_NRW**

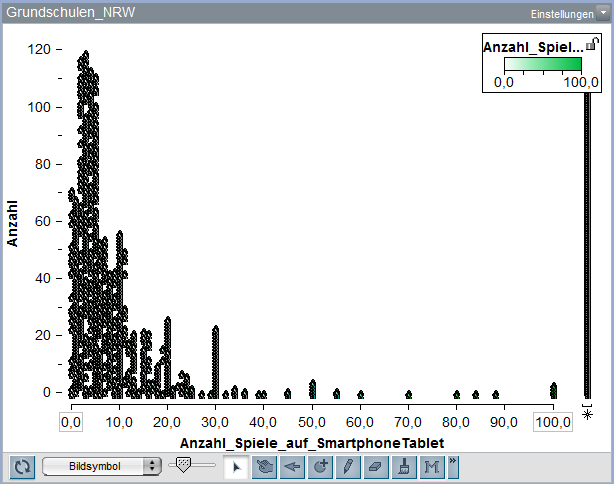
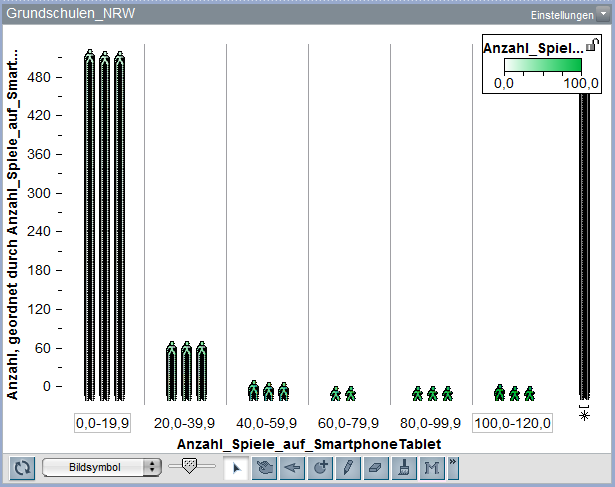
Die “Personen” sind nun grün eingefärbt, aber in unterschiedlichen Farbintensitäten – hier sehen wir nun wie in TinkerPlots kategoriale und numerische Merkmale anhand der Einfärbung unterschiedlich dargestellt werden. Bei den kategorialen Merkmalen hatten wir noch verschiedenen Ausprägungen verschiedene Farben zugeordnet. Nun -bei numerischen Merkmalen- sind die „Personen“ in einer Farbe eingefärbt, aber mit unterschiedlichen Intensitäten: desto stärker der Farbton ausgeprägt ist, desto mehr Spiele hat in diesem Fall das jeweilige Kind („Personen-Symbol“) auf dem Smartphone. Nun wollen wir eine Übersicht über die Verteilung des Merkmals „Anzahl\_Spiele\_auf\_Smartphone/Tablet“ bekommen – wie gehen wir vor? Wir nutzen wieder unsere Grundoperationen „Stapeln“, „Trennen“ und „Ordnen“, die wir analog auch mit unseren Datenkarten realisieren können. Mithilfe von „Trennen“ teilen wir die Kinder in zwei Klassen hinsichtlich der Anzahl ihrer Spiele ein (Abbildung 33 links). In der linken Klasse haben wir nun die Kinder, die zwischen 0 und 49 Spiele, in der rechten Klasse, die zwischen 50 und 100 Spiele haben. Auch hier kann „Stapeln“ und „Ordnen“ Übersicht verschaffen und uns zu einer übersichtlicheren Verteilung führen, wie wir (Abbildung 33 rechts) sehen können:

**Abbildung 33: TinkerPlots Graph zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmatphoneTablet (links: getrennt, rechts: getrennt und gestapelt)**

Durch wiederholte Anwendung der Operation Trennen können wir die Einteilung der Klassen immer weiter verfeinern, bis wir ein gestapeltes „Punktdiagramm“ haben (siehe Schrittfolge in Abbildung 34):



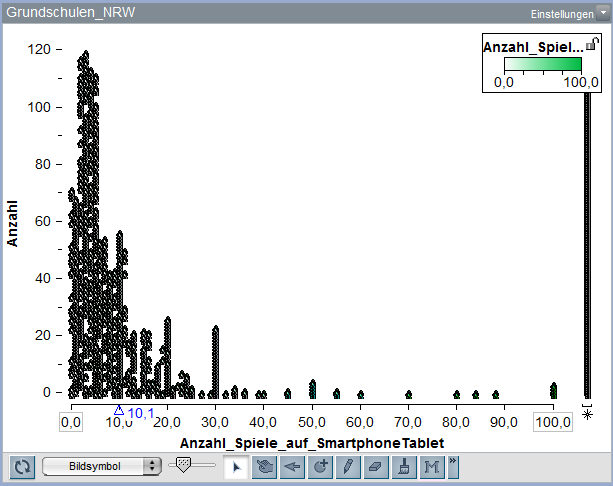
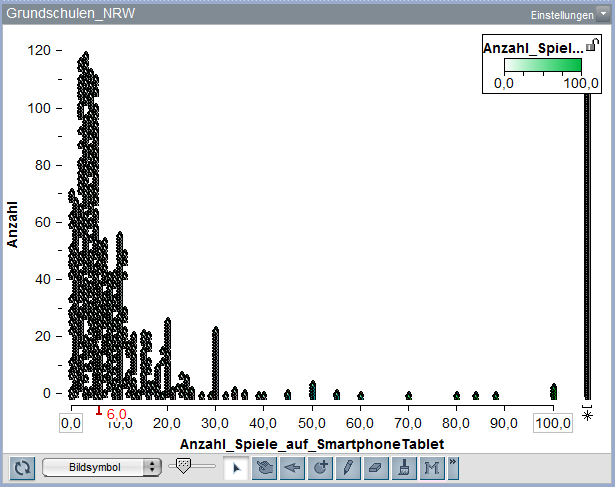


**Abbildung 34: Auf dem Weg zum Punktdiagramm zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet im Datensatz Grundschulen\_NRW**

Insbesondere anhand der TinkerPlots-Graphiken in Abbildung 34 kann man gut erkennen, dass die Verteilung des Merkmals „Anzahl\_Spiele\_auf\_SmartphoneTablet“ linkssteil/rechtsschief ist. Ein sehr großer Teil der befragten Schülerinnen und Schüler ballt sich auf der linken Seite des Graphen. Präziser ließe sich sagen, dass die meisten der Befragten 10 Spiele oder weniger auf ihrem Smartphone/Tablet haben. Es gibt weitere Auffälligkeiten in den Diagrammen: neben den „Ausreißern“, die 50 oder mehr Spiele auf ihrem Smartphone/Tablet haben, gibt es auch viele Kinder die genau 20 oder 30 Spiele auf dem Smartphone haben. Diesem Phänomen kann man näher auf den Grund gehen, indem man sich die Datenkarten der jeweiligen Kinder genauer anschaut.

Genauso könnte man versuchen, weitere Charakteristika der Verteilung „Anzahl\_Spiele\_auf\_SmartphoneTablet“ herauszufinden. Neben der Form ist das Zentrum einer Verteilung eines numerischen Merkmals sehr interessant, wenn man die Verteilung kurz und prägnant beschreiben möchte.

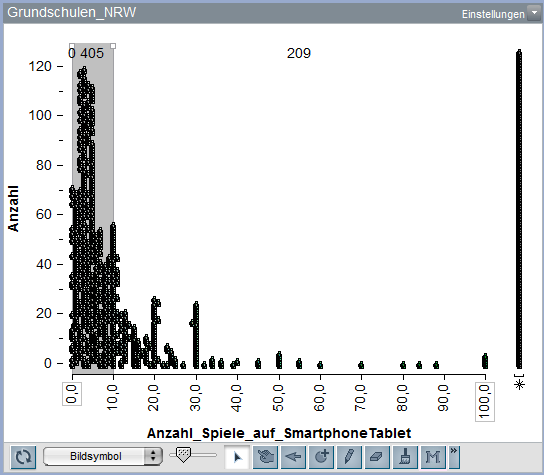
Mithilfe von TinkerPlots  können wir zunächst sowohl den Median als auch das arithmetische Mittel der Verteilung des Merkmals „Anzahl\_Spiele\_auf\_SmartphoneTablet“ berechnen (siehe Abbildung 35).



**Abbildung 35: Punktdiagramme zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet (links: mit Median, rechts: mit arithmetischen Mittel)**

Der Median als Zentrum der geordneten Datenreihe kann als Mittelwert bereits in der Grundschule thematisiert werden und bietet einen verlässlicheren Wert als beispielsweise das arithmetische Mittel, welches zum einen durch die Schiefe der hier vorliegenden Verteilung und zum anderen durch die hier vorliegenden Ausreißer deutlich größer als der Median ist, der das Zentrum der Verteilung besser kennzeichnet. Anhand von Abbildung 35 (links) können wir sagen, dass unsere befragten Schülerinnen und Schüler im Median 6 Spiele auf ihrem Smartphone/Tablet besitzen. Betrachtet man das arithmetische Mittel (siehe Abbildung 35 rechts), so lässt sich sagen, dass die befragten Schülerinnen und Schüler im Durchschnitt ungefähr 10 Spiele auf ihrem Smartphone haben.

Mithilfe von sogenannten Einteilern , die TinkerPlots zur Verfügung stellt, lassen sich weitere Fragen an die Verteilung stellen – und auch beantworten. Will man herausfinden wie viele der Befragten zwischen 0 und 10 Spielen auf dem Smartphone haben, so kann man diese Einteiler einfach über den gewünschten Bereich legen (dieser färbt sich dann grau ein) und in Kombination mit der Anzahl/Anteil-Funktion die entsprechende Anzahl bzw. den entsprechenden Anteil an Fällen in diesem Bereich durch die Software berechnen lassen (siehe Abbildung 36).



**Abbildung 36: Punktdiagramm zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet mit Einteiler**

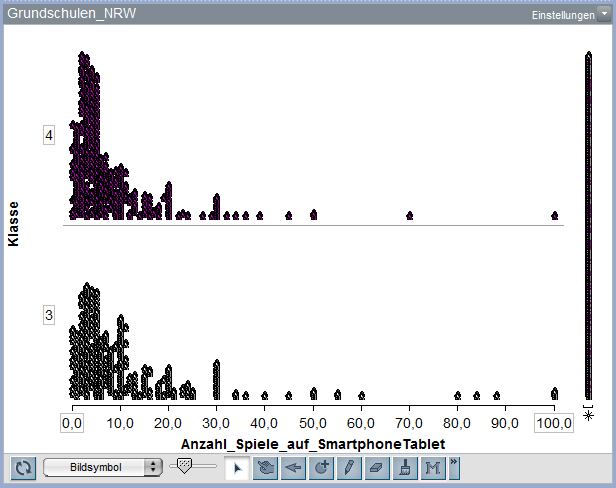
So können wir beispielsweise der Graphik in Abbildung 36 entnehmen, dass 405 (ca. 66%) der befragten Grundschulkinder unter 10 Spiele auf ihrem Smartphone/Tablet haben.

Ausblick: Weiterführung und Vertiefung sowie unterrichtspraktische Anregungen zum Erstellen und Interpretieren von gestapelten Punktdiagrammen finden sich in *Baustein 6: „Was bedeutet der Hügel da?“ - Eine Unterrichtssequenz zur Förderung des Lesens und Interpretierens von eindimensionalen Streudiagrammen in Klasse 4*

In einem weiteren Schritt könnte man wieder eine Frage nach Unterschieden aufwerfen, z.B. Unterschiede klassenspezifischer Art, d.h. inwiefern sich die Drittklässler und Viertklässler hinsichtlich der Anzahl ihrer Spiele auf dem Smartphone/Tablet unterscheiden, z.B.:

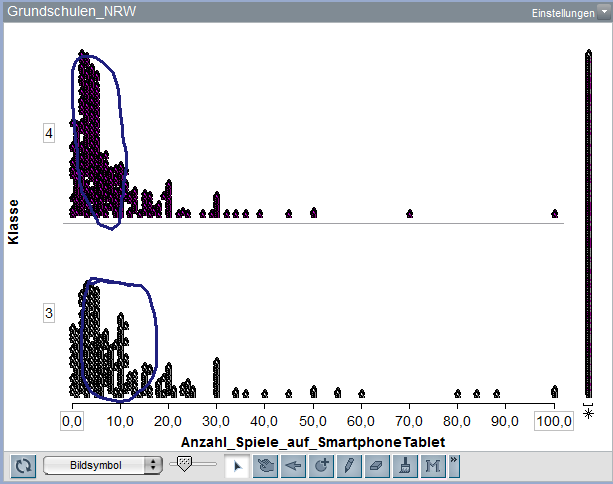
## Untersuchung des Zusammenhangs zwischen einem kategorialen und einem numerischen Merkmal: Inwiefern unterscheiden sich die Dritt- und die Viertklässler hinsichtlich der Anzahl ihrer Spiele auf dem Smartphone/Tablet?

Auch hier können wir genauso vorgehen, wie bei den kategorialen Variablen und jetzt in diesem Fall die Verteilung des Merkmals „Anzahl\_Spiele\_auf\_SmartphoneTablet“ zwischen den Drittklässlern und den Viertklässlern trennen (siehe Abbildung 37).



**Abbildung 37: Punktdiagramme zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet unterschieden nach dem Merkmal Klasse**

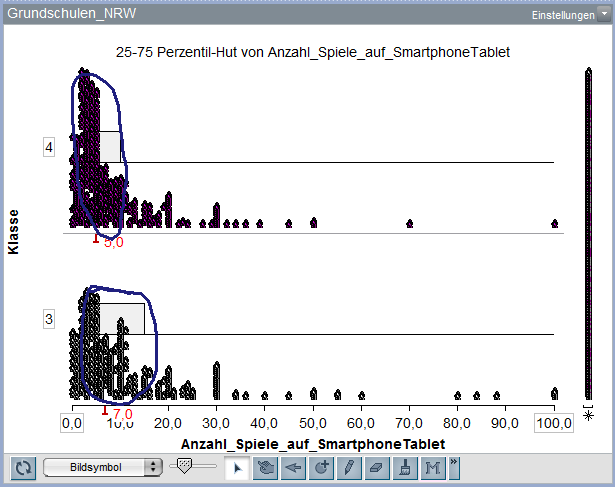
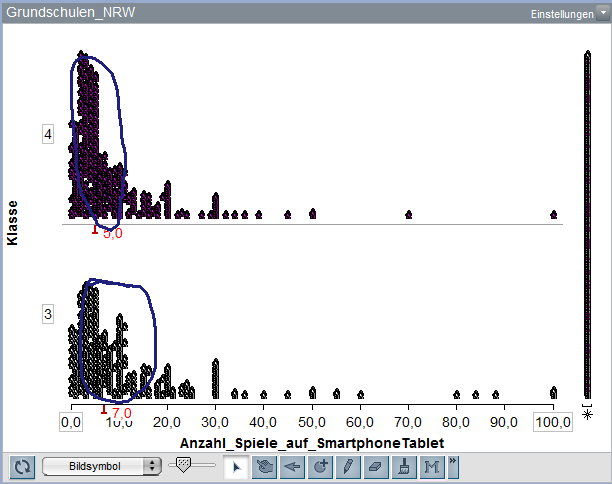
Um Unterschiede zwischen den beiden Verteilungen herauszuarbeiten, könnte man zunächst die Bereiche, in denen sich die meisten Datenpunkte ballen, markieren und die Lage dieser Ballen[[1]](#footnote-1) vergleichen. Das Stiftwerkzeug in TinkerPlots eignet sich gut dazu, diese so-genannten modalen Klumpen in den Verteilungen zu identifizieren (siehe Abbildung 38).



**Abbildung 38: Punktdiagramme zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet unterschieden nach dem Merkmal Klasse (mit eingezeichneten modalen Klumpen)**

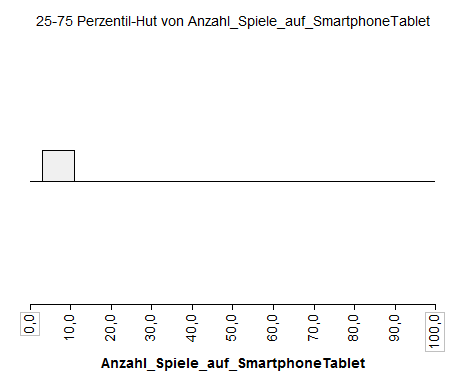
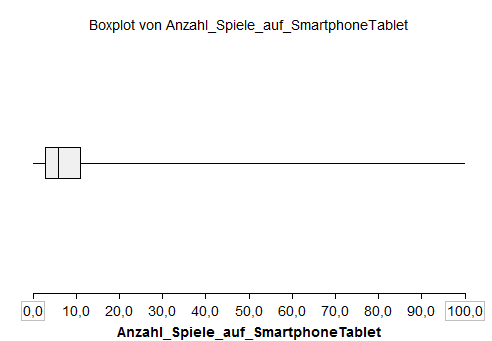
Ein erster qualitativer Blick zeigt, dass der so markierte modale Klumpen bei den Drittklässern breiter ist als der bei den Viertklässlern. Er liegt auch bei höheren Werten. Man kann das so ausdrücken, dass die Drittklässler tendenziell mehr Spiele auf dem Smartphone\_Tablet haben.

TinkerPlots bietet zum einen mit dem Median, aber auch mit der Nutzung von Hutplots Möglichkeiten an, diese Aussagen zu präzisieren, wie wir in Abbildung 39 nachvollziehen können.



**Abbildung 39: Punktdiagramme zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet unterschieden nach dem Merkmal Klasse (links mit modalen Klumpen & Median, rechts mit modalen Klumpen, Median und Hutplot)**

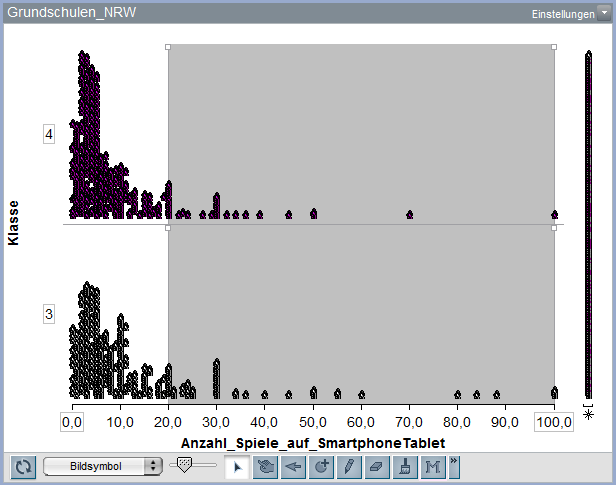
Der Median ist in der Gruppe der Drittklässler höher als in der Gruppe der Viertklässler (siehe Abbildung 39 links) was zu der Aussage veranlassen kann, dass die befragten Drittklässler im Median zwei Spiele mehr auf ihrem Smartphone/Tablet besitzen als die Viertklässler (7 statt 5). Es kann sich –gerade wenn es um den Vergleich von Verteilungen eines numerischen Merkmals geht - anbieten, sich auf spezielle Kennzahlen wie das erste Quartil (Das kleinste Viertel der Datenwerte ist kleiner als dieser oder gleich diesem Kennwert) den Median (die kleinste Hälfte der Datenwerte ist kleiner als dieser oder gleich diesem Kennwert), das dritte Quartil (Das kleinste Dreiviertel der Datenwerte ist kleiner als dieser oder gleich diesem Kennwert) sowie Minimum und Maximum der jeweiligen Verteilung zu beschränken. Eine Darstellung, die aus diesen fünf Kennzahlen besteht und somit eine Zusammenfassung robuster Lage- und Streumaße enthält, ist der Boxplot. Salopp gesprochen teilt der Boxplot eine aufsteigend geordnete Datenreihe in vier „Anteil-gleiche“ Abschnitte ein (Viertel:Viertel:Viertel:Viertel). Einen Boxplot zur Verteilung des Merkmals „Anzahl\_Spiele\_auf\_SmartphoneTablet“ sehen wir in Abbildung 40 (links).



**Abbildung 40: Boxplot (links) und Hutplot (rechts) zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet**

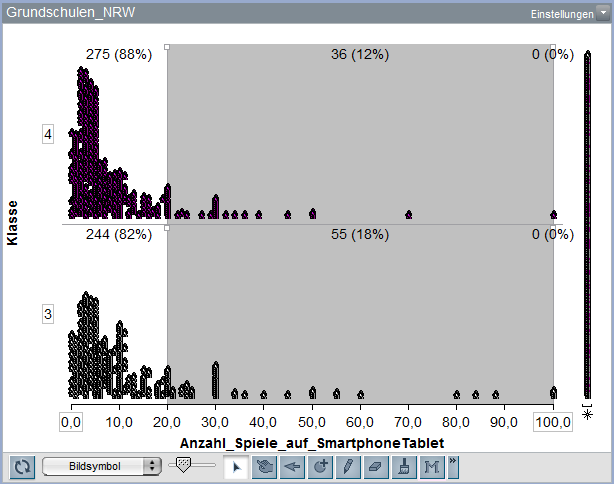
Als Vorstufe zum Boxplot ist der Hutplot (Abbildung 40 rechts) als Darstellung einer Verteilung eines numerischen Merkmals in TinkerPlots gegeben. Im Gegensatz zum Boxplot kann man den Median im Hutplot allerdings nicht ablesen. Außerdem teilt der Hutplot die Daten in das erste Viertel, die mittlere Hälfte („ die mittleren 50%“) und das letzte Viertel auf. Unter der Krone des Hutes liegt die mittlere Hälfte der Verteilung, an den Rändern bzw. unter der linken Krempe / rechten Krempe liegen jeweils die unteren / oberen Viertel der Daten. Die Hüte (Abbildung 39 rechts) festigen unseren Eindruck, den wir schon bei den modalen Klumpen gewonnen haben: der Hut (die mittlere Hälfte) der Verteilung der Drittklässler ist breiter, was auf eine größere Streuung der mittleren Hälfte darstellt. Außerdem ist die Krone des Huts bei den Drittklässlern mehr nach rechts verschoben als bei den Viertklässlern, was den Anlass zu der Aussage gibt, dass die Drittklässler tendenziell mehr Spiele auf dem Smartphone/Tablet haben als die Viertklässler.

Es schließen sich – wie oben bei der Betrachtung der Verteilung des Merkmals „Anzahl\_Spiele\_auf\_SmartphoneTablet“- weitere Fragen an die Verteilung an: wie viele Drittklässler haben mehr als 20 Spiele auf dem Smartphone/Tablet? Wie viele sind das bei den Viertklässlern? Und natürlich: Gibt es diesbezüglich Unterschiede zwischen beiden Gruppen? Um Fragen solcher Art beantworten zu können, eignen sich die „Einteiler“ , die wir schon bei der Exploration der Verteilung eines numerischen Merkmals kennengelernt haben. Mit diesem Werkzeug können wir nun z.B. die Anzahl und den Anteil der Kinder in Klasse 3 und 4 die mehr als 20 Spiele auf dem Smartphone/Tablet haben, ermitteln (Abbildung 41).



**Abbildung 41: Punktdiagramme zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet mit Einteiler**

Einblenden der Anzahlen und Anteile  liefert uns jetzt die Antwort auf unsere Fragen (Abbildung 42): 36 von 311 (12%) Viertklässlern haben 20 oder mehr Spiele auf dem Smartphone/Tablet, 55 von 299 (18%) Drittklässlern haben 20 oder mehr Spiele auf dem Smartphone/Tablet. [[2]](#footnote-2) An dieser Stelle sei noch einmal betont, dass im Mathematikunterricht der Grundschule der Anteilsbegriff als solcher natürlich noch nicht verwendet werden kann. Die Nutzung von bekannten prägnanten Anteilen (ein Viertel, die Hälfte, etc.) sowie die Nutzung von Kreisdiagrammen kann lediglich einer Propädeutik des Anteilsbegriffs (mit Blick auf die weiterführende Schule) bieten. Wir nutzen in unserer Einführung dennoch den Anteilsbegriff sowie die verschienenen Prozentarten, um das Datenanalysepotential der Software TinkerPlots aufzeigen zu können.



**Abbildung 42: Punktdiagramme zur Verteilung des Merkmals Anzahl\_Spiele\_auf\_SmartphoneTablet mit Einteiler (und absoluten und relativen Häufigkeiten)**

Ausblick: Weiterführung und Vertiefung sowie unterrichtspraktische Anregungen zum Verteilungsvergleich finden sich in Baustein 7: „*Springen schwere Papierfrösche weiter als leichte? Und warum haben eigentlich Drittklässler mehr Spiele auf dem Smartphone als Viertklässler?“ - Verteilungen vergleichen mit TinkerPlots (modale Klumpen, Hüte, Median)*

Unsere kleine elementare Einführung in die Datenanalyse mit TinkerPlots soll mit kurzen Ausführungen zum Graph-Menü sowie zur Datentabelle schließen. Außerdem bieten wir noch einen Arbeitsauftrag zur Festigung des hier Gelernten an.

# Nützliche Tipps zum Graph sowie Informationen zur Datentabelle

Es gibt noch weitere nützliche Eigenschaften, die die „Graph“-leiste bietet (Abbildung 43).



**Abbildung 43: „Graph“-Leiste in TinkerPlots**

Der Button „Neu mischen“ (roter Kreis) mischt den Graph durch (setzt ihn in sein ungeordnetes Anfangsstadium zurück), der Schieber (blauer Kreis) variiert die Symbolgröße. Mit dem Pfeil (grüner Kreis) lässt sich das Symbol verschieben und somit der Wert ändern, der Zeigefinger (gelber Kreis) setzt den zum Symbol zugehörigen Wert wieder zurück. Mit dem Kreuz (schwarzer Kreis) lässt sich ein Wert zum Graph hinzufügen. Mit dem Stift (orangener Kreis) können Markierungen im Graphen vorgenommen, mit dem Radiergummi (grauer Kreis) diese wieder entfernt werden. Mit Betätigung des Pfeils (weißer Kreis) kehrt man wieder zum „normalen“ Mauszeiger zurück.

**Die Datentabelle**

Um die Datentabelle eines Datensatzes zu erzeugen, können Sie diese durch Anwahl des Icons  in den Arbeitsbereich ziehen (Abbildung 44).



**Abbildung 44: Datentabelle zum Datensatz Grundschulen\_NRW in TinkerPlots**

Hier können Sie ebenfalls Merkmale und die entsprechenden Merkmalsausprägungen hinzufügen. Mit Anklicken von <neu> können Sie beispielsweise ein neues Merkmal erstellen, über den Rechtsklick erhält man die bekannten Operationen „ausschneiden“, „kopieren“ und „einfügen“.

Generell lassen sich in der Datentabelle ähnliche Einstellungen wie in den Datenkarten vornehmen. Insbesondere eignet sich die Datentabelle gut zur Dateneingabe.

# Ihr zweiter Arbeitsauftrag

Untersuchen Sie die folgenden drei Fragestellungen bzw. führen Sie die folgenden drei Aktivitäten im Datensatz\_Grundschulen\_NRW mit TinkerPlots durch:

1. Stellen Sie die Verteilung des Merkmals Eigenes SmartphoneTablet im Datensatz\_Grundschulen\_NRW dar und beschreiben Sie die Verteilung.
2. Besitzen eher die Jungen oder eher die Mädchen eine Spielekonsole? Untersuchen Sie, inwiefern sich die Jungen und Mädchen hinsichtlich des Besitzes einer Spielekonsole unterscheiden (Merkmale: Geschlecht, Spielekonsole).
3. Inwiefern unterscheiden sich die Jungen und die Mädchen im Datensatz hinsichtlich der Zeit, die sie am Wochenende am Fernseher verbringen? (Merkmale: Geschlecht, Stunden\_Fernseher\_amWochenende)

Gehen Sie bei Ihrer Datenanalyse mit TinkerPlots für die Exploration der Fragestellungen 1), 2) und 3) jeweils die folgenden Schritte i, ii und iii durch:

1. *Erzeugen Sie mit TinkerPlots aussagekräftige Graphiken, die Ihnen helfen, Ihre Fragestellung zu beantworten.*
2. *Beschreiben (Was ist dargestellt?) und interpretieren (Was bedeutet das für meine Fragestellung?) Sie die Graphiken, die Sie mit TinkerPlots erstellt haben. Nutzen Sie zur Dokumentation Ihrer Beschreibung und Ihrer Interpretation das Textfeld in TinkerPlots.*
3. *Beantworten Sie jeweils Ihre Fragestellung kurz und prägnant. Dokumentieren Sie Ihre Antworten in einem Textfeld in TinkerPlots.*

1. In der Fachdidaktik werden diese auch „modale Klumpen“ genannt. Dabei stellt der modale Klumpen eine intuitive Zusammenfassung der „Hauptgruppe“ der Daten da und berücksichtigt Lage und Streuung. Hierauf aufbauend kann dann später die Hauptgruppe standardisiert als die Lage der „mittleren Hälfte“ der Verteilung eingeführt werden. [↑](#footnote-ref-1)
2. Die Prozentangaben sind hier lediglich als Hintergrundinformation zu verstehen. In der Grundschule können diese als Gruppenvergleichsgrundlage nicht verwendet werden. Da die Gruppen (311 vs. 299) aber fast gleich groß sind, lässt sich auch anhand der absoluten Häufigkeiten (36 vs. 55) und dem Bezugnehmen mit der entsprechenden Gesamtheit „36 von 311“ und „55 von 299“ gut erahnen, dass bei den Drittklässlern tendenziell mehr Kinder sind, die 20 oder mehr Spiele auf ihrem Smartphone oder Tablet haben als bei den Viertklässlern. [↑](#footnote-ref-2)