Interactive Visualisation Report

# LO1: Performance

Bei dieser Abgabe nutze ich die Programmiersprache R und lade mir die Library Plotly herunter. Plotly ist ideal für interaktive und dynamische Grafiken. Es gibt enorm viele Möglichkeiten, interaktive Grafiken zu gestalten. In den Modulen an der FHNW arbeiten wir meistens mit bereits gesäuberten und eher kleinen Datensets. Ich habe mir von Kaggle das Datenset Bikes heruntergeladen (Quelle: <https://www.kaggle.com/datasets/saurabhshahane/seoul-bike-sharing-demand-prediction>). Alle Grafiken, über die in diesem Report geschrieben wird, sind in meinem Github zu finden unter: [www.github.com/ChantalGisler/Interactive-Visualisation](http://www.github.com/ChantalGisler/Interactive-Visualisation). Ich persönlich empfehle, das HTML-File oder das R Notebook beim Lesen dieses Reports geöffnet zu haben, damit die Ausführungen hier auch Sinn ergeben. Ich werde sie hier auch dementsprechend erwähnen.

In diesem Datenset sind die Ausleihen von Fahrrädern in der südkoreanischen Hauptstadt Seoul enthalten, dazu diverse Wetterdaten zu jeder Stunde des Tages. Die Daten nutze ich für ein Projekt und eine Präsentation im Modul Programmieren in R. Die Datenmenge ist sehr gross, sie enthält 14 Spalten und 8760 Zeilen. Wenn ich hier also interaktive Grafiken haben möchte, geht es auch vor allem um die Performance. Denn in der heutigen Zeit muss alles sehr schnell gehen, das wird von allen erwartet – damit auch von dem Zielpublikum, das eine Visualisierung sehen möchte. Ich beobachte dieses Phänomen auch bei mir selbst: Wenn ein Youtube-Video länger als 5 Sekunden braucht, um zu laden und es abzuspielen, überlege ich mir noch einmal, ob ich es auch wirklich sehen möchte. In den meisten Fällen wechsle ich die Seite oder lege das Handy weg – das Interesse ist verschwunden. Daher ist es enorm wichtig, dass man die Performance bei den Grafiken möglichst hoch macht. Um die Performance zu erhöhen, gibt es mehrere Möglichkeiten, die ich jetzt genauer beschreiben werde.

Bevor man die Grafiken visualisiert, sollte man die Zahlen und Kenndaten zur Performance berücksichtigen. Konkret bedeutet das, dass ich unter anderem wissen muss, wie schnell der Prozessor und die Speicherkarte meines Computers oder Laptops ist. Umso kleiner, desto langsamer läuft das Gerät – vor allem, wenn man grosse Datenmengen verarbeiten möchte. Eine grosse Rolle spielt hier vor allem die CPU. Sie ist die zentrale Rechen- und Steuereinheit eines Computers, also eigentlich das «Gehirn» des Geräts, denn hier wird alles verarbeitet. Die Leistung meines Geräts kann ich unter Systeminformationen bei Prozessor abrufen. Die CPU meines Geräts liegt bei 1.60GHz und liegt damit beim normalen Standard. Umso höher die Anzahl GHz ist, desto schneller ist die CPU, sprich damit auch die Leistung des PCs. Das bedeutet auch, dass interaktive Visualisierungen je nach dem, wie viele Daten sie enthalten, auch schneller geladen werden. Die CPU in einem Laptop kann man nicht beeinflussen, ohne sich ein schnelleres oder leistungsfähigeres Gerät holen zu müssen.

Wenn schon von der Hardware, also den Geräten, gesprochen wird, muss auch der sogenannte Grafikprozessor, die sogenannte GPU, thematisiert werden. Sie ist hauptsächlich für die grafische Darstellung zuständig, beispielsweise von Videospielen, Filmen oder eben auch interaktiven Grafiken. Solche GPUs stellen die Leistung, sprich die Performance, für grafikrelevante Daten und Befehle bereit. Sie ist beispielsweise dafür zuständig, dass die Webseiten beim Surfen im Netz mit hoher Auflösung gezeigt werden, man beim Streamen alles genau erkennen kann und man bei 4k-Filmen jedes Detail genau erkennen kann. Ein Gerät mit einem guten Grafikprozessor kostet aber entsprechend auch mehr, wie man bei sogenannten Gamer-PCs sieht. Für Gamerinnen und Gamer ist eine hohe Auflösung wichtig, damit man den Gegner in jedem Spiel gut erkennen kann. Ein Beispiel: Bei einem gut programmierten und gut gestalteten Shootergame kann man das Aufblitzen der Waffen erkennen, bevor der Gegner die Waffe abfeuern und einen töten kann. Unter anderem deshalb brauchen Gamer ein Gerät mit hoher Auflösung, um diese kleinen Feinheiten erkennen zu können.

Es geht aber nicht nur um das Gerät, sondern auch darum, was genau man visualisieren möchte. Bei der Datenvisualisierung soll man sich vor dem Plotten überlegen: Muss ich wirklich jeden Datenpunkt einzeln zeigen, beispielsweise in einem Scatterplot? Oder wäre es nicht besser, wenn ich die Zahlen zusammenrechne und in einem Balkendiagramm zeige? Grundsätzlich gilt: Bei interaktiven Visualisierungen gibt es «schnellere» und «langsamere» Visualisierungsformen. Sprich: Umso aufwändiger eine Visualisierung ist und umso mehr sie können muss, desto mehr Leistung wird auch gebraucht, um sie schön darzustellen. Dasselbe gilt auch bei den Details: Umso detaillierter eine Visualisierung ist, desto mehr Leistung braucht sie auch, genau so wie der Gamer, der die aufblitzende Waffe erkennen muss. Um den Vergleich hier besonders gut darzustellen, habe ich dieselbe Grafik auf meinem normalen Alltagslaptop laufen lassen und die Zeit gestoppt. Anschliessend habe ich dieselbe Grafik auf dem Gamer-PC laufen lassen. Im vergangenen Jahr hatte dieser PC von allen auf dem Markt verfügbaren die beste Grafikkarte. Er eignet sich also nicht nur für Videospiele hervorragend, sondern auch für Grafiken und 3D-Modelle, die auf ihm wunderbar alles erkennbar machen und die meinen Laptop enorm schnell an seine Grenzen bringen.

In einem zweiten Versuch wollte ich aufzeigen, dass auch verschiedene Darstellungsformen interaktiver Grafiken von der Grafikkarte und dem PC abhängig sind. So wird das Balkendiagramm, in dem ich die Gesamtzahl aller ausgeliehenen Fahrräder nach Saison aufteile und darstelle, wesentlich schneller dargestellt. Die gemessene Zeit beträgt hier gerade mal 0.473356 Sekunden.

Anschliessend probiere ich es mit einer Grafik, in der mehr Daten vorkommen, nämlich einem Scatterplot, in dem jedes ausgeliehene Fahrrad als Punkt dargestellt wird und nach Saison in vier verschiedene Farben aufgeteilt wird. Und siehe da, diese aufwändigere Grafik braucht etwa dreimal so lange, wie das einfache Balkendiagramm: Es dauert ganze 1.329525 Sekunden, bis die Grafik dargestellt wird. Ähnlich sieht es mit dem interaktiven Linienplot aus, bei dem jede Linie einen Tag darstellt und auf der Y-Achse die Anzahl ausgeliehener Fahrräder liegt und auf der X-Achse die Uhrzeiten.

Für mich bedeutet das: Umso mehr ich in einer interaktiven Grafik darstellen möchte, desto langsamer wird sie. Aber auch das Gerät macht einen enormen Unterschied. Auf dem Gamer-PC werden die Visualisierungen deutlich schneller dargestellt. Ich bin mir bewusst, dass die Personen, die diesen Report lesen und das HTML-File oder das R Notebook dazu geöffnet haben und die Codes laufen lassen, bei der Messung der Zeit eventuell zu anderen Ergebnissen kommen. Deshalb habe ich die Ergebnisse der Zeitmessung hier auch bildlich dargestellt.

# LO2: Dashboard design principles

Damit eine interaktive Visualisierung für das Publikum möglichst gut verständlich ist, muss man einige Prinzipien verfolgen. Wie bei statischen Visualisierungen auch darf die interaktive Visualisierung nicht überfüllt sein und sie muss das Auge führen. Der Leser muss im Idealfall zuerst eine Übersicht erhalten, dann erst mit der Grafik herumspielen, also filtern und hineinzoomen. Aber wie kann man dem Leser diese Abfolge «aufzwingen» und wie kann es dem Leser im Idealfall sogar gefallen, sich mit der Grafik zu beschäftigen und ihm dabei auch gleich noch die Message hinter den Daten, sprich die Hauptaussage näherbringen und verständlich machen? Genau damit befasse ich mich in diesem Kapitel.

Wie bei jeder Visualisierung muss ich die Daten zunächst aufbereiten und mir überlegen, was genau ich zeigen möchte. Sprich: Was ist das Ziel der Visualisierung, was ist die Hauptaussage hinter den Daten? Wie bei statischen Visualisierungen überlege ich mir zuerst, was ich zeigen möchte. Danach kommen die Features. Eine interaktive Grafik ist ja schön und gut, aber man muss auch wissen, was genau sie aussagen möchte.

Ich muss mir auch bewusst sein, dass jede Grafik selbst ein sogenanntes Dashboard ist. Das bedeutet, ich als Urheberin dieser Grafik muss den Beobachter oder die Beobachterin durch die Grafik führen – ganz nach Ben Shneidermans Mantra: «Übersicht zuerst, dann zoomen und filtern und Details auf Verlangen». Shneiderman hat dieses Mantra in einem Bericht namens «The eyes have it: a taks by data type taxonomy for information visualizations». In der praktischen Anwendung bedeutet das Folgendes: Ich überlege mir zunächst, was für eine Grafik ich machen möchte und was sie anzeigen soll. Zu einer guten Grafik gehört auch der durchdachte Umgang mit Farben. Durch das Hinzufügen einer dritten Variable in Form von unterschiedlichen Farben, beispielsweise indem ich die Saisons farblich abtrenne, wird die Visualisierung für den Betrachtenden viel interessanter. Man kann so etwa lineare Abhängigkeiten erkennen, die ich durch eine Trendlinie zusätzlich verstärken kann. Auch das gehört zu einem guten Dashboard dazu, um den Betrachtenden für meine interaktive Grafik zu interessieren und ihm oder ihr einen guten Überblick zu geben. Das ist mir besonders wichtig: Bei einer interaktiven Grafik soll man unbedingt auch ohne zoomen, filtern und skalieren einen Überblick erhalten und die Aussage der Grafik verstehen können. Das Zoomen, Filtern und Skalieren ist aus meiner Sicht als Hilfsmittel gedacht, damit sich der Beobachter oder die Beobachterin noch intensiver mit den Daten befassen kann.

Wichtig zu wissen: Bei einer interaktiven Grafik wird die Computermaus auch immer als sogenannter Hover benutzt. Das heisst, wenn ich mit der Maus über eine Grafik schwebe, tauchen Zahlen auf. Bei einem Scatterplot sind es immer die Koordinaten des Datenpunktes, die dort auftauchen. Zunächst die X-Achse, dann die Y-Achse. Genauso ist es bei allen Grafiken, die eine X- und Y-Achse haben, beispielsweise auch Liniendiagrammen oder Boxplots. Bei einem Balkendiagramm wird immer zuerst die Höhe der Y-Achse angezeigt, die Zahl nach dem Komma ist die Nummer des Balkens. Diese sogenannte Hoverinfo kann ich ganz leicht verändern, aber auch hier muss ich mir vorher überlegen, was wichtig ist und was genau ich zeigen möchte. Dazu gehören nicht nur Werte, sondern auch Text. Das heisst, ich könnte hier auch Hinweise hineinlegen, um die Grafik noch interessanter zu machen. Das alles auszuführen, würde den Rahmen dieses Reports aber sprengen. In R gibt es enorm viele Möglichkeiten, interaktive Grafiken zu machen und die Daten interaktiv zu zeigen. Es gibt etwa die Funktion des Sliders, in den man gleichzeitig hineinzoomen kann. Grundsätzlich eine sehr coole Funktion, aber man muss auch wissen, wie man sie anwenden soll. Um meine Erklärungen zu zeigen, habe ich zwei Plots gemacht. Beide zeigen die Auswirkungen von Regen auf die Anzahl der Ausleihen von Fahrrädern in Seoul.

Zunächst habe ich einen Scatterplot gemacht, der auf der X-Achse zeigt, wie viele Milliliter Regen gefallen sind und auf der Y-Achse wie viele Fahrräder bei dem Regen ausgeliehen worden sind. Ich habe die Regenfälle auch nach Wochentag aufgeteilt und sie mit einer Trendlinie versehen, um zu zeigen, dass bei Regen immer weniger Fahrräder in Seoul ausgeliehen werden. Entstanden ist eine wirre Visualisierung, die zwar zeigt, dass umso mehr es regnet, desto weniger Fahrräder werden ausgeliehen. Aber die Aufteilung der Farben ist verwirrend und durch den Slider wird das gesamte Bild verzerrt und nicht richtig dargestellt. Da mehrere Trendlinien vorhanden sind, kann nicht genau gesehen werden, was hier überhaupt gezeigt werden soll. Ich persönlich erhalte davon Kopfschmerzen, umso länger ich auf die Grafik zeige und habe daher überhaupt keine Lust, mich näher damit zu befassen und ich gehe fest davon aus, dass man bei dieser Grafik auch gar nicht merkt, welche Möglichkeiten sie überhaupt bietet und was man alles sehen könnte. Ich lerne daraus, wie man es auf jeden Fall nicht machen soll. Die Grafik bietet mir keinen richtigen Überblick, da die Trendlinien übereinander liegen. Das heisst, ich müsste zunächst filtern, damit ich alle vier Trendlinien mit ihren Datenpunkten erkennen könnte – und das widerspricht Shneidermans Mantra. Denn eine Grafik muss von Anfang an übersichtlich sein, das ist für ihn der Heilige Gral, der bei jeder Visualisierung beachtet werden muss.

Deshalb gehe ich bei meinem positiven Beispiel anders vor: Ich beginne damit, die Visualisierung von Anfang an übersichtlich zu gestalten. Das bedeutet in meinem Beispiel: Weniger ist mehr. Wenn ich zeigen möchte, dass mehr Regen zu weniger Ausleihen führt, nutze ich besser einen Linienplot. Der Slider, den ich zum Schluss einsetzen möchte, eignet sich grundsätzlich sehr gut zum Vergleichen von zwei Werten, beispielsweise je ein Tag in einer anderen Saison oder wie oft die Fahrräder an den verschiedenen Wochentagen ausgeliehen werden. Hier benutze ich eine schönere Palette, damit man die Farben gut voneinander unterscheiden kann. Noch ist die Grafik normal und nicht interaktiv. Das ändert sich mit dem nächsten Schritt, bei dem ich den Slider hinzufüge. Der Code dafür ist sehr einfach:

Mit der Maus kann ich nun einen spezifischen Zeitpunkt auf der Grafik auswählen und erhalte alle Informationen zu den Ausleihen der Fahrräder an den verschiedenen Wochentagen. Mit dem Slider kann ich die Zeitspanne nun verändern und mir die Linienplots genauer ansehen. Ich kann durch einen Klick auf den Wochentag auch eine Wochentagslinie entfernen oder durch einen Doppelklick isolieren. Dass all das möglich ist, wusste ich anfangs nicht. Daher kann ich es auch nicht voraussetzen, dass der Leser oder die Leserin das weiss. Deshalb füge ich diese Informationen entweder in die Grafikbeschreibung oder den Untertitel ein. Ideal wäre hier ein Standarttext von R selbst, aber ich verstehe, dass hier der Aufwand auch zu gross ist. Da der Slider aber dort ist, wo sich normalerweise der Untertitel befindet, muss ich die entweder im Text um die Grafik zeigen, was sie kann oder im Grafiktext oder in einer Box nebenan.

# LO3: HCI basics

Nachdem ich jetzt weiss, wie ein gutes und übersichtliches Dashboard aussehen muss, komme ich jetzt zu dem, was eine interaktive Visualisierung richtig interessant macht: Die interaktiven und animierten Features, mit denen der oder die Betrachtende sich die interaktive Grafik genauer anschauen und sich darin vertiefen können. Konkret geht es um die Funktionen Zoom, Auswahl von einzelnen Daten und Filtern. Ich gehe dabei noch einen Schritt weiter und füge mit der Funktion «rangeslider» noch einen Slider hinzu, mit dem man in den Daten «surfen» kann. Hier kommt das Thema «Human Computer Interaction» (HCI) auf. Dazu habe ich mir das Video «Interaction Design for Data Visualization» angesehen, in dem der Facebookingenieur Miles McCrocklin über seine Erfahrungen mit Datenvisualiserungen und interaktiven Visualisierungen spricht (InfoQ, 2015).

Vorab muss ich aber wissen, für welche Art von Gerät ich die Grafik überhaupt mache. Wie bei Webseiten kommt es auch bei interaktiven Grafiken darauf an, worauf sie dargestellt werden. Haben meine Zuschauer eine Maus zur Verfügung, mit denen sie über die Punkte eines Scatterplots schweben und die Werte vergleichen können? Oder haben sie gar einen Smart Pen, mit dem sie in die Grafik hineinschreiben können. Aus diesem Grund muss ich mir bei besonderen Grafiken immer zuerst überlegen, für welches Publikum ich das eigentlich mache und vor allem wie ich es verfügbar mache. Wichtig: Ich persönlich bin nie der Endkonsument. Ich kenne die Daten ja schon, ich weiss, was die Visualisierung kann und was sie können muss. Ich muss sie also für Personen machen, die sich noch nie mit den Daten befasst haben. Als Journalistin habe ich gelernt, immer vom dümmsten möglichen Leser auszugehen, das heisst, Dinge ganz einfach zu erklären, damit jemand, der noch nie damit zu tun hatte, es auch versteht. Dieses Prinzip wende ich auch in der Datenvisualisierung an, wenn ich nicht sicher sein kann, dass es ausschliesslich für ein Publikum gedacht ist, das die Daten schon kennt. Grundsätzlich sind meine interaktiven Grafiken meist für eine Webseite bzw. einen Onlineblog gedacht, weshalb ich davon ausgehen muss, dass man sowohl mit Handy als auch mit Tablet und Laptop darauf zugreifen wird. Sprich: Nicht alle meine Leserinnen und Leser haben eine Maus zur Verfügung, mit denen sie sich die Grafik genau anschauen können. Aus diesem Grund habe ich auf Webseiten, auf denen ich interaktive Grafiken hochlade, ganz oben einen Disclaimer, in welchem ich erkläre, dass es mehr Sinn macht, die interaktiven Grafiken an einem Laptop mit einer Computermaus anzuschauen. Natürlich kann ich es niemandem vorschreiben, aber empfehlen. Und ich kann die Grafiken so machen, dass sie auch schön und interessant aussehen, ohne dass man in die Daten hineinzoomen oder sie skalieren kann.

Aus dem Modul «Storytelling mit Daten» und «Grundlagen der Datenvisualisierung» habe ich gelernt, dass man die Daten auf zwei Arten darstellen kann: Explorativ und explanativ. Explorativ bedeutet, dass ich den oder die Zuschauerin die Daten erforschen lasse. Das heisst, ich gebe den Nutzern einen Überblick und sie können mit verschiedenen Features selbst die Daten «erforschen» und ihre eigenen Schlüsse ziehen. Sie können mehr Daten hinzufügen oder extrahieren, worauf sie sich fokussieren, ist ihnen überlassen. Beim explanativen Storytelling führe ich die Nutzer durch meine Daten. Das heisst, sie können trotzdem hineinzoomen und mit den Daten interagieren, allerdings folgen sie dem roten Faden, den ich ausgelegt habe, um im Idealfall zum selben Schluss zu kommen. Natürlich ist auch bei der explorativen Erklärweise ein roter Faden dahinter, also ein Ziel, das ich als Urheberin der Visualisierungen ansteuere. Aber der Raum für Interpretationen ist offen und die Nutzer entscheiden selbst, wie weit sie sich mit den Daten befassen möchten und wie tief sie darin eintauchen. Das bedeutet für mich, dass ich mir immer bewusst sein muss, was ich mit meinen Visualisierungen aussagen möchte und wo welche Information «versteckt» sein soll. Ist die Hauptaussage nur einen Klick, ein Zoom oder eine Filterung von dem User entfernt? Und wie mache ich dem User das klar? Hier kommt auch wieder Ben Shneidermans Mantra ins Spiel: Der User muss einen Überblick haben, also verstehen, welche interaktiven Möglichkeiten er oder sie hat und dass er oder sie sie anwenden kann, um zur wesentlichen Aussage der Daten zu kommen. Im Projekt und der Challenge habe ich die Visualisierungen stets ein paar Freundinnen und Freunden gezeigt, die sich noch nie mit den Daten befasst haben und die nicht Data Science studieren, da die Präsentationen stets auf ein Laienpublikum ausgelegt waren. Ich habe sie gebeten, mir zu sagen, was sie sehen und erkennen. Das hat mir enorm dabei geholfen und das werde ich auch beibehalten.

Bei der Interaktion zwischen Computer und Mensch ist es von grosser Bedeutung, zu wissen, was das Publikum erwartet. Das bedeutet beispielsweise, dass ich Symbole bewusst auswählen muss. Wenn ich beispielsweise mit Daten über die Rega, die Schweiz oder das Rote Kreuz arbeite und ein Kreuz oder ein X auf einem roten Hintergrund für die Datenpunkte benutze, könnte man meinen, dass wenn man auf dieses Kreuz klickt, dass dann die Datenpunkte verschwinden oder man die interaktive Visualisierung beenden kann. Ich bin leider noch nicht so gut im Programmieren, dass ich eine Visualisierung darstellen könnte, bei der man auf etwas klickt und dann etwas Neues auftaucht. Aber ich bin mir der Darstellung bewusst und werde sie nicht vergessen.

Interaktive Visualisierungen müssen nicht unbedingt dafür eingesetzt werden, um die Daten zu analysieren und sie anzuschauen, sondern auch, um damit zu arbeiten. Momentan bin ich wissenschaftliche Assistentin beim Inselspital in Bern im Bereich Diabetologie. Aktuell setzen mein Team und ich uns mit Hyperglykämie auseinander, ein extrem tiefer Blutzucker, der bei den meisten Diabetikern nach einer Mahlzeit auftaucht. Wir haben unsere Testpersonen gebeten, in einem Tagebuch einzutragen, wann sie etwas gegessen haben und wann sie Symptome von Hyperglykämie haben. Wir wollten herausfinden, ob die Hyperglykämie auch nach kleinen Snacks kommt oder nur bei grossen Mahlzeiten. Leider haben die meisten Testpersonen nicht alle Mahlzeiten eingetragen. Im Nachhinein sagten sie gegenüber den Ärzten, dass sie beispielsweise bei ein paar Nüssen oder einem Apfel einfach nicht daran gedacht haben. Aus diesem Grund haben wir den Blutzucker als Liniendiagramm dargestellt und die Ärzte haben zusammen mit den Patienten von Hand Kreuze dort gemacht, wo die Patienten Snacks hatten. Die Grafik war ein grosser Erfolg, auch, weil die Ärzte es sich gewohnt sind, die Liniendiagramme ausgedruckt vor isch zu haben und Notizen darin zu machen. Ausdrucken war hier aber keine Option, da die Patienten ein Messgerät im Arm haben, das kontinuierlich den Stand des Blutzuckers an ein spezielles Messgerät schickt, das wiederum die Daten als Liniendiagram auf einer App anzeigt. Auch müssen die Daten in einer Datenbank festgehalten sein, da die Ergebnisse der Studie wiederholbar sein müssen. Aus Datenschutzgründen darf ich diese interaktive Grafik nicht zeigen, aber ich kann sie mit den Fahrrad-Daten von Seoul nachstellen. Dazu mache ich einen einfachen Linienplot, in dem ich die Schnee- und Regenmenge über ein Jahr in Seoul anzeige. Das Programm kann mir jetzt zwar nicht anzeigen, wann es gleichzeitig geregnet und geschneit hat, sprich es Schneeregen gab, aber ich kann das nun von Hand einzeichnen. Natürlich ist das nicht ganz so interessant wie die Daten aus dem Spital, aber es zeigt einmal mehr, was alles mit interaktiven Grafiken möglich ist und wie weit der Mensch am Computer hier eingreifen kann.

# LO4: Evaluation

Bevor ich eine Grafik mache, muss mir darüber klar werden, weshalb ich mich für eine interaktive Grafik entscheide und keine normale, statische mache. Ich muss mir darüber klar werden, was ich darstellen möchte und weshalb es von Vorteil ist, wenn sie interaktiv wäre. Die Darstellung einer statischen Grafik bleibt bestehen, was vor allem bei ausgedruckten Berichten wichtig ist, da man eine interaktive Grafik nur auf dem Bildschirm sehen kann. Bei statischen Grafiken wird das, was ich als Autorin des Berichts hervorhebe, auch von den Leserinnen und Lesern wahrgenommen. Das heisst: Ich entscheide, was wichtig ist und was hervorgehoben wird, die Person, die den Bericht liest, hat keine Möglichkeit, etwas anderes aus der Grafik hervorzuheben. Bei den interaktiven Grafiken hingegen hat der Leser oder die Leserin die Möglichkeit, die Grafik «upzudaten», also das hervorzuheben, was ihn oder sie interessiert. Mit der Hover-Funktion kann man sich spezifische Daten genauer anschauen und man kann Datenpunkte gruppieren. So entsteht ein völlig neuer Blickwinkel auf die Grafik. Das bedeutet umso mehr, dass ich den oder die Leserin durch die Grafik führen muss und ihm oder ihr auch zeigen muss, was in den Grafiken alles möglich ist. Bei interaktiven Grafiken muss ich umso mehr an den Endkonsumenten denken. Einerseits, dass er oder sie merkt, was man überhaupt alles in der Grafik hervorheben kann und nicht, ob man hineinzoomen oder sich einen Ausschnitt genauer ansehen kann. Das heisst, ich muss wissen, wer mein Zielpublikum ist und auch von welchem Gerät aus auf die Grafik zugegriffen wird. Passiert es mit einem Laptop, einem Smartphone oder einem Tablet? Hat der oder die Beobachtende eine Maus zur Verfügung oder nicht? Das sind allesamt Aspekte, an die ich bei statischen Grafiken nie gedacht habe, da ich davon ausgehen kann, dass die meisten statischen Grafiken sowohl auf einem grossen Bildschirm als auch auf einem kleinen sehr gut dargestellt werden. Grundsätzlich kannte ich dieses Problem schon von früher als ich als Journalistin arbeitete und eine eigene Webseite betrieb. Hier kannte ich diese Problematik vor allem mit Fotos und vor allem Videos, bei denen ich darauf achten musste, dass sie bereits bei der Bearbeitung auf die richtige Grösse achten musste, weil es sonst auf Smartphones verzogen oder falsch dargestellt werden würde.

Was mich bei diesem Modul aber besonders erstaunt und auch ein wenig geprägt hat, war, dass ich mich zum ersten Mal mit der Grafikkarte und der Speicherkarte meiner Geräte auseinandergesetzt habe. Bis jetzt hat mich Speicher nur interessiert, wenn ich ihn für etwas aktiv gebraucht habe, vor allem als externe Speicherkarte für meine vielen Fotos, bei meinem Smartphone, ebenfalls für Fotos. Bei den Laptops musste es immer einen CD-Abspieler haben und er musste möglichst leicht sein. Dass ich mich für solche «technischen» Sachen interessieren und es mich sogar faszinieren würde, hätte ich nicht erwartet. Ich habe tatsächlich etwas über Laptops und Geräte dazu gelernt – vor allem durch meine Freunde aus dem Informatikbereich. Das finde ich sehr spannend. Mein eigener Laptop pfeift aus dem letzten Loch, das heisst, ich muss mir bald einen neuen zulegen. Und dieses Mal werde ich mir einen mit guter Grafikkarte und viel Speicher holen, besonders da ich in den kommenden Modulen mit interaktiven Grafiken arbeiten werde und es sogar noch 3D-Modelle geben wird, die ich auf meinem Laptop gut dargestellt haben möchte, damit ich auch richtig damit arbeiten kann. Gleichzeitig habe ich festgestellt, dass es nicht nur für mich sondern auch für mein Publikum angenehmer ist, wenn ich ein schnelles und gutes Gerät habe. Ich habe die Dauer der Grafiken auf meinem alten Laptop und auf dem Gamer-PC vier Freundinnen, die nicht Data Science studieren, gezeigt und sie gefragt, welche Dauer für sie zu lang ist und mit welcher Visualisierung sie sich lieber befassen würden. Sie alle sagten alle unisono, dass sie eine schnellere Visualisierung bevorzugen. Eine Freundin erklärte mir das anhand eines Beispiels: Wenn bei Tiktok ein Video nicht schnell genug lädt, verliert sie das Interesse und scrollt weiter. Das würde sie auch bei Visualisierungen machen, wenn sie nicht schnell genug laden würden. Für mich heisst das: Wenn eine Grafik zu langsam ist, verliert mein Publikum das Interesse an der Grafik und das wäre schade.

Gleichzeitig konnte ich meine Programmierfähigkeiten durch dieses Modul ausbauen. Ich kann nun sehr coole Grafiken interaktiv gestalten und habe sogar angefangen, 3D-Modelle zu programmieren. Die stehen aber noch ganz am Anfang, dennoch bin ich sehr stolz. Ich selbst spiele gerne mit diesen Grafiken herum und muss mich manchmal zusammenreisen, um mich auf das Wesentliche zu fokussieren. R Studio bietet enorm viele Möglichkeiten und es gibt vor allem auch viele Arten von Code, der eine Visualisierung interaktiv macht. Ich habe viel mit den Kursen auf der Webseite DataCamp gearbeitet, wo auch über die Themen dieses Moduls gesprochen wurde, weshalb ich mich sogar noch intensiver damit befassen konnte. Besonders spannend fand ich das Thema rund um Ben Shneidermann. Ich habe eine ausführliche Zusammenfassung seines Berichts gelesen und dazu diverse Beispiele gesehen, wie man sie umsetzen kann. Deshalb schreibe ich mir im R Code immer auf, was genau ich zeigen möchte und wie die Visualisierung aussehen soll und vor allem was sie aussagen soll. Und ich habe begriffen, dass es enorm wichtig ist, daran zu denken, dass ich nicht die Endkonsumentin bin. Mir ist besonders ein Zitat vom Facebookprogrammiere Miles McCrocklin hängen geblieben, und zwar, dass interaktive Grafiken wie ein Tanz sind. Das heisst, nicht nur die Grafik muss den Zuschauenden dazu bringen, sich mit ihr zu befassen, sondern er muss sich auch von den Zuschauern ein Stück führen lassen und sie in die Richtung lenken können, in die ich als Urheberin das möchte.

Ich habe die Grafiken auch fünf Mitstudierenden gezeigt beziehungsweise habe sie einfach damit interagieren lassen, denn ich wollte herausfinden, wo der Fokus auf den Visualisierungen liegt, wenn sie interaktiv sind. Was mich hier ein wenig erstaunt hat: Alle finden die Möglichkeiten sehr cool und spielen gerne damit herum. Und genau hier beginnt das Problem: Indem sie «zu viel» damit herumspielen, realisieren sie teilweise gar nicht, worum es in den Grafiken überhaupt geht. Ich könnte eine sehr coole Grafik mit einem sehr wichtigen Inhalt machen, aber ich könnte mir nicht zu 100 Prozent sicher sein, dass die Information oder die Message dahinter wirklich bei ihnen ankommt oder sie sie überhaupt realisieren und sich dann in diesem Kontext mit der Grafik befassen. Deshalb ist es hier besonders wichtig, sein Publikum zu kennen und auch die Grafiken nicht ohne Kontext weiterzugeben. Für das Modul «Storytelling mit Daten» habe ich auch einige interaktive Grafiken gemacht. Und dadurch, dass sie vorher den Text gelesen haben, konnten sie aus meiner Sicht viel besser mit der Grafik interagieren. Natürlich kann man das nicht generalisieren, dazu bräuchte ich eine grössere Testgruppe von etwa 100 Personen, die ich bei der Interaktion mit mehreren verschiedenen interaktiven Grafiken beobachten könnte, aber mir hat es dennoch gezeigt, dass ich wissen muss, wer mein Publikum ist und was es von den interaktiven Grafiken erwartet. Das bedeutet für mich, dass Weniger teilweise doch Mehr ist, weil man sich so auf das Wesentliche, also die Message hinter den Daten und der interaktiven Grafik, fokussieren kann.

Für mich ist das eine wertvolle Erfahrung, da ich persönlich nach dem Programmieren den interaktiven Grafiken sehr viel Spass daran hatte, sie noch besser und noch interaktiver zu machen. Das bedeutet auch, dass ich mir bei Arbeiten, in denen ich Grafiken gestalte, genau vornehme, was ich zeigen möchte, was die Hauptaussage sein muss, und mir überlege, ob sie interaktiv sein muss, sprich ob es überhaupt einen Mehrwert bringen würde oder eben nicht. Mit R kann ich das sehr einfach herausfinden, da ich die Grafik ohnehin zuerst mit ggplot statisch gestalte und dann mit dem Befehl ggplotly erst interaktiv mache und programmiere, was sie überhaupt können muss. Natürlich kann man interaktive Grafiken in R auch von Anfang an interaktiv machen, aber ich habe für mich entdeckt, dass es so am besten funktioniert, da ich mich selbst dazu zwinge, zwischen einer statischen Grafik und einer interaktiven zu entscheiden und beide nebeneinander vor mir sehe und sie so auch vergleichen muss, gleichzeitig muss ich aber im Hinterkopf behalten, dass ich nicht das Zielpublikum bin. Ich habe in diesem Modul ebenfalls gelernt, dass ich mein Publikum eben sehr genau kennen muss. Und wenn ich es nicht kenne, muss ich die Grafiken möglichst einfach und ohne viel Firlefanz herum gestalten.

Ebenfalls wichtig zu wissen: Nur weil eine Grafik interaktiv ist und viel kann, heisst das nicht automatisch, dass es auch eine gute Grafik ist. Aus schlechtem Design wird auch eine schlechte Grafik, selbst wenn sie interaktiv ist und viele Möglichkeiten zur Interaktion bietet.