# LE1

In diesem Kapitel setze ich mit den Grundlagen der Datenvisualisierungen auseinander. Dazu werde ich Visualisierungen erstellen und ein paar Grundsätze erläutern.

Als Datensatz verwende ich die Wetterdaten von der Wetterstation Mythenquai der Seepolizei zürich aus der Wettermonitor-Challenge, welche ich letztes Jahr absolviert habe. Die Daten sind hier zu finden: [https://data.stadt-zuerich.ch/dataset/sid\_wapo\_wetterstationen] (<https://data.stadt-zuerich.ch/dataset/sid_wapo_wetterstationen>). Der Datensatz erhält Messwerte über mehrere

Aufgabe in der Challenge war es, ein Dashboard für die aktuellen Daten zu erstellen. Ich möchte mich hier aber auf einen grösseren Zeitraum beziehen und werde Visualisierungen über mehrere Jahre erstellen.

## Barplots

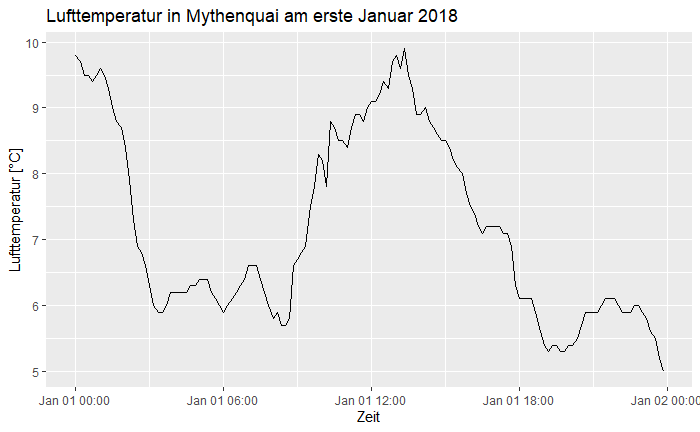
Chart, bar chart

Description automatically generated

In der Abbildung ist ersichtlich wieviel es jeden Monat geregnet hat im Jahr 2018.

Die Höhe der Balken zeigt den totalen Niederschlag des Monats in mm an. Da die Balken nebeneinander sind, kann der Betrachter schnell herausfinden in welchem Monat es am Meisten oder wenigsten geregnet hat. Bei einem Doughnut oder Kuchendiagramm wäre diese nicht mehr so einfach, da der Mensch Längen besser als Winkel vergleichen kann. Die zeitlich Sortierung der Balken ermöglicht dem Betrachter die Veränderung über den Verlauf des Jahres zu sehen.

## Zeitreihendiagramm



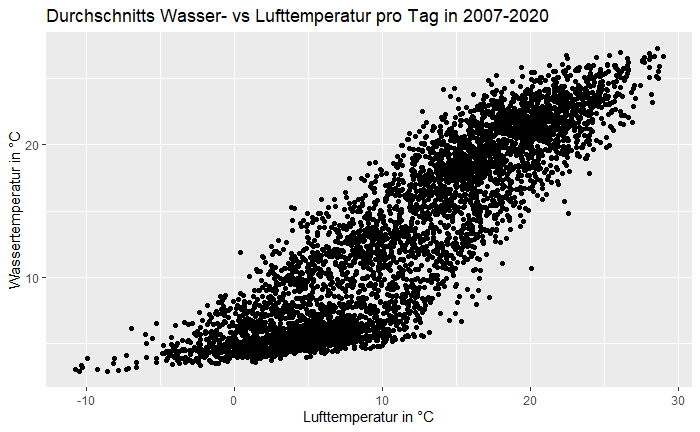
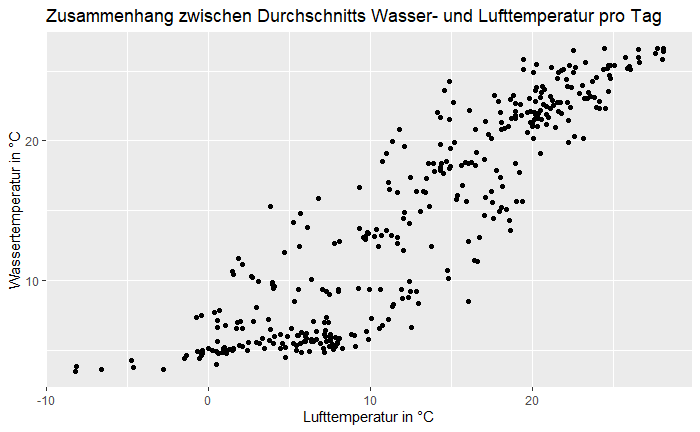
In der Abbildung 2 kann der Verlauf der Temperatur an einem Tag angesehen werden. Dazu eignet sich ein Zeitreihendiagramm oder allgemein auch «Liniendiagramm» genannt. Auf der y-Achse kann der Betrachter die Temperatur ablesen. Auf der x-Achse wird die Zeit sortiert dargestellt. Die verbundenen Linien machen dem Betrachter klar, dass die Daten einen temporalen Zusammenhang haben. Man könnte hier auch ein Balkendiagramm verwenden, aber erstens gäbe das hier sehr viele Balken (um genau zu sein 144) und zweitens ist der temporale Zusammenhang nicht mehr so präsent.

## Histogramm



In der Abbildung 3 ist die Verteilung des totalen Niederschlags pro Tag zu sehen. Die Breite der Balken fassen ein Intervall von «Niederschlag an einem Tag zusammen» zusammen, z.B. 0mm bis 5mm. Die Länge der Balken gibt die Anzahl Tage an in diesen es zwischen diesem Intervall geregnet hat. Der Betrachter sieht hier, dass es an vielen Tagen nicht oder wenig regnet und nur an ganz wenigen Tagen richtig viel regnet. Ein Histogramm ist optimal um die Verteilung von etwas darzustellen. TODO QUELLE

## Punktewolken

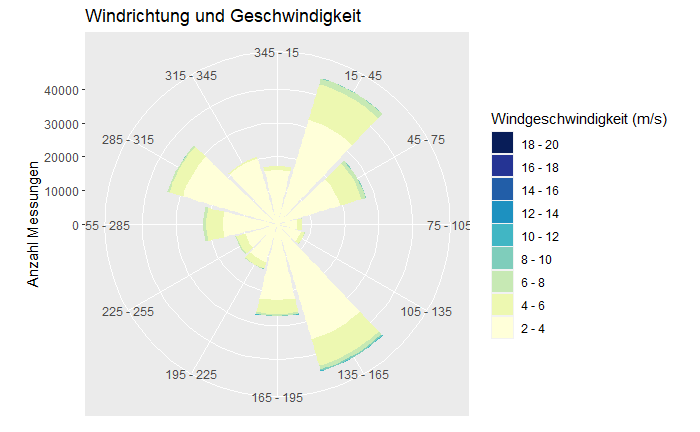
Chart, histogram

Description automatically generated

In der Abbildung 4 ist eine Punktewolke zusehen. Auf der x-Achse ist die Luftemperatur zusehen, auf der y-Achse die Wassertemperatur. Ein Punkt bezieht sich jeweil auf den Durchschnitt von einem Tag. Der Betrachter sieht sofort einen linearen Zusammenhang zwischen der Luftemperatur und der Wassertemperatur.

Wenn man die Daten von allen Jahren plottet, dann ist dies nicht mehr übersichtlich, wie in Abbildung 5 zu sehen ist. Bei so vielen Überschneidungen eignet sich eine Heatmap. Dort wird durch eine Farbe die Anzahl der Punkte in jedem Rechteck abgebildet (siehe Abbildung 6).

## Windrose



In der Abbildung 7 ist eine Windorse zu sehen. Der Betrachter kann hier mit Hilfe der Länge der Balken erkennen, von welcher Richtung der Wind am meisten weht. Mit Hilfe der Farben kann der Betrachter erkennen wie schnell der Wind geweht hat. Man könnte hier auch ein Balkendiagramm verwenden, es wäre für den Betrachter aber schwieriger die Himmelsrichtung zu erkennen und mit welchem Winkel die Balken zu einander stehen.