

DSML Project Paper

Iris Lüthi

Maja Velkova

Yannik Zimmermann

Contents

1	Einleitung	2
2	Data Source und Datasets	2
2.1	US Pollution	2
2.2	US Weather Events	2
2.3	Global Air Quality Index (AQI)	2
2.4	Temperature of Major Cities	2
3	Analyse	2
4	Resultate und Diskussion	2
5	Fazit	2
5.1	Including Plots	3

1 Einleitung

Im Rahmen des Moduls Data Science and Machine Learning DSML werden die Studierenden der Berner Fachhochschule dazu angehalten eine empirische Arbeit auf dem Gebiet der Datenanalyse oder des maschinellen Lernens unter Verwendung von R oder Python zu erarbeiten.

Die Autoren haben sich für zwei verschiedene Datensets entschieden. Zum einen repräsentiert das erste Datenset die US-Wetterereignisse zwischen den Jahren 2016 bis 2021 und das zweite Datenset die US-Umweltverschmutzung zwischen den Jahren 2000 bis 2021. Die Autoren möchten in der folgenden Arbeit herausfinden, ob es einen Zusammenhang zwischen der Wirkung des Wetters und der Umweltverschmutzung gibt. Um die Analyse einzugrenzen, werden nur die Staaten Texas, California, und New York analysiert. Es wurden Hypothesen/Fragen am Anfang aufgestellt, welche anschliessend durch die importierten Datensätze auf deren Richtigkeit überprüft wurden.

2 Data Source und Datasets

2.1 US Pollution

2.2 US Weather Events

2.3 Global Air Quality Index (AQI)

2.4 Temperature of Major Cities

3 Analyse

4 Resultate und Diskussion

5 Fazit

Literaturverzeichnis

Selbstständigkeitserklärung

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

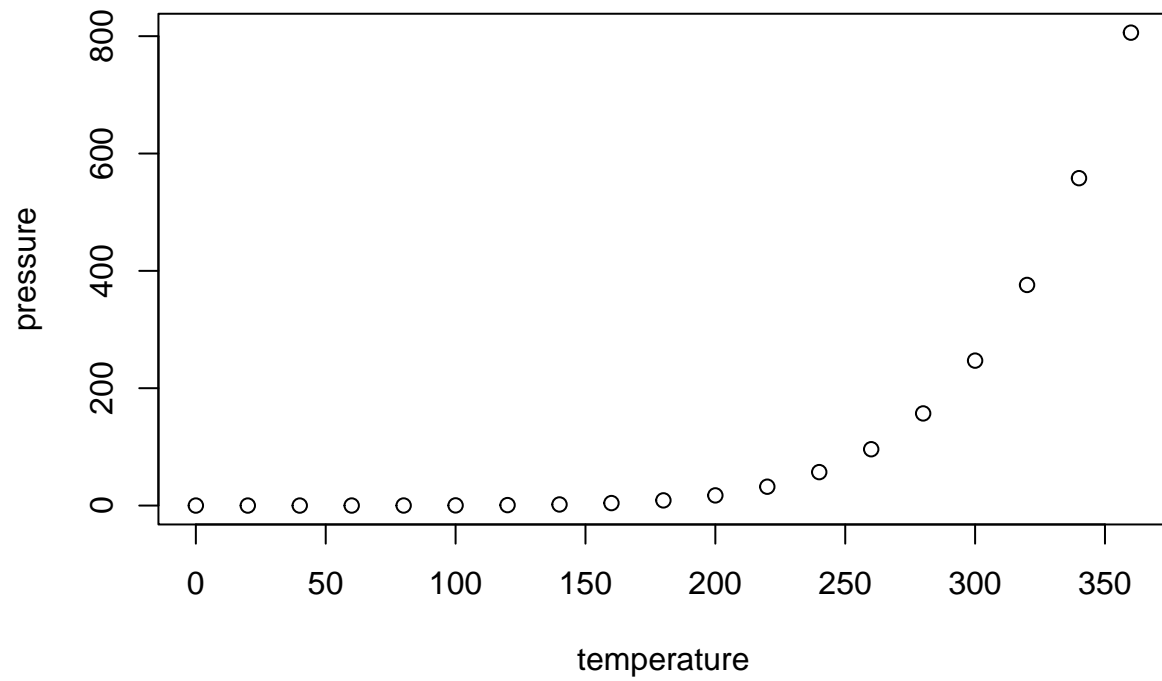
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

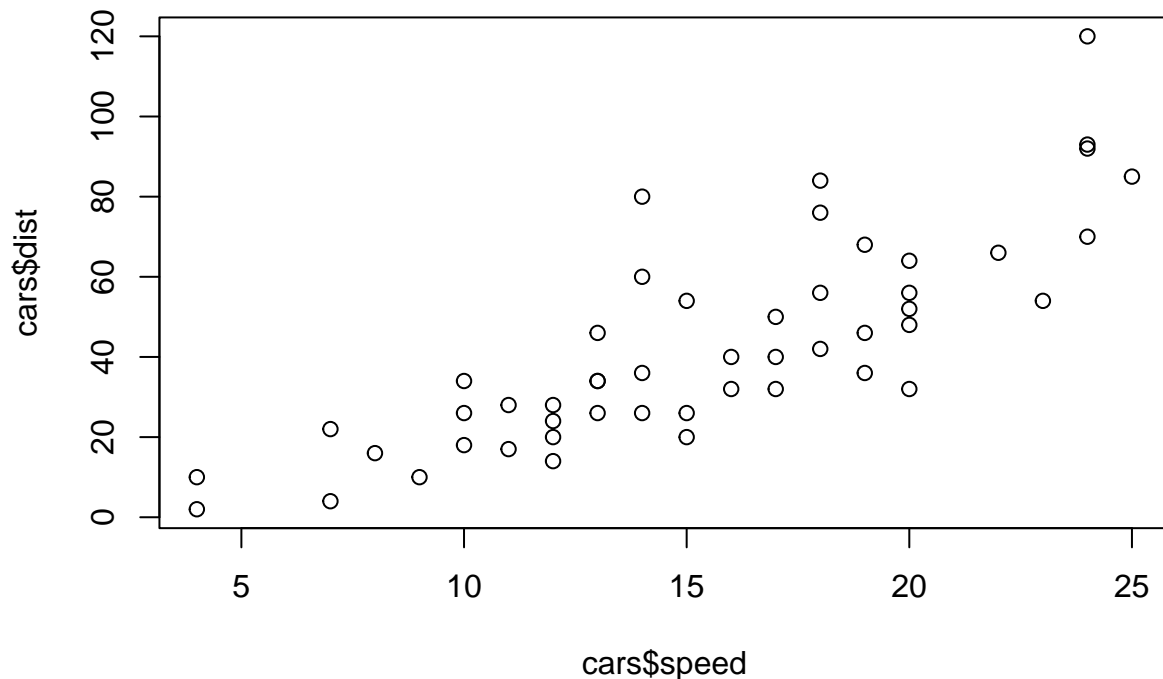
```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   : 2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

5.1 Including Plots

You can also embed plots, for example:





Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

Fragen in Zusammenhang mit Unwetter: - Hat sich die Anzahl Unwetter (definieren was mit Unwetter gemeint ist) über die fünf Jahre erhöht? - Wenn ja, welche Art von Unwetter hat zugenommen? Welche Art von Unwetter hat abgenommen/ist weniger vorgekommen? - Besteht eine Korrelation zwischen der erhöhten Anzahl Unwetter (definieren welches Unwetter) und der erhöhten/tieferen Durchschnittstemperaturen? - Hat die Frequenz der Unwetter zu- oder abgenommen? - Frage kann noch nach Unwetterart und Ort aufgeteilt werden. - Hat die Stärke der Unwetter zu- oder abgenommen? - Frage kann noch nach Unwetterart und Ort aufgeteilt werden. - Hat die Wassermenge an Regentagen allgemein zu- oder abgenommen? - Haben sich die Wetterereignissen örtlich verschoben? (Zuerst Standard aufzeigen, danach mögliche Änderung) - Hat sich die Anzahl Regentagen verändert? - Hat sich die Anzahl Sonnentage verändert? - usw

Fragen in Zusammenhang mit Temperaturänderungen: - Wurde eine generelle stetige Temperaturänderung über die fünf Jahre gemessen? - Wenn ja, wurde es allgemein kälter oder wärmer? - Wurde eine allgemeine Temperaturerhöhung über die Jahre 2016-2021 gemessen? - Wenn ja, wie hoch und wie rasch hat sich die Temperatur geändert? (Diagramm) - In welchem Staat, wurde die stärkste Temperaturänderung gemessen? - Wie stark weicht die Temperaturänderung von der natürlichen Erderwärmung ab?

Fragen in Zusammenhang mit den Gasen: - Welches Gas hat allgemein am stärksten zugenommen? - Gleiche Frage in Bezug zu den einzelnen Staaten und Jahren - Kann aufgrund der Erhöhung eines bestimmten Gases eine Erhöhung/Verstärkung einer bestimmten Unwetterart erkannt werden?