



# DIS08 – Data Modeling

## 05 – Common Data Formats – Tutorial Session

---

Fabian Haak, Technische Hochschule Köln, Cologne, Germany

Version: 2021-11-19

# Topics for Today

1. Q&A

2. kein Quiz

3. Assignment/ Help

# Git Help

Bei Git(Hub) Problemen:

Funktioniert alles? Ist alles da, wo es sein sollte?

→ wahrscheinlich kann alles so bleiben wie es ist!

Wer es trotzdem genauer wissen möchte:

<https://www.atlassian.com/de/git/tutorials/undoing-changes>

How to clean up Git mess:

<http://justinhileman.info/article/git-pretty/git-pretty.pdf>

# Hinweise Zu Assignment 1

## Exercise 5.2: Cleaning the .csv file

- Sollten 2-4 Edits sein
- Nutzen Sie ein Mittel ihrer Wahl (Excel, Shell, Text Editor, Python,...)
- Tipp: Am ehesten in description Spalte notwendig, zur Not sind auch “sinnlose” Edits besser als keine Edits (solange Sie keine Inhalte drastisch ändern und die Ergebnisse von 5.3 beeinflussen)
- Wohl am einfachsten: Suchen und Ersetzen Funktionen  
→ Dokumentation z.B.:

```
## Edit 1
**Problem:** `Beschreibung des Problems oder Edits`
**Ersetzen:** `REGEX`
**Mit:** `REGEX2`
```

## **!Exercise 5.3 has to be solved in the Shell!**

- Manche der Aufgaben sollten in mehreren Commands gelöst werden.
- Sie dürfen die CSV in eine TSV Datei umwandeln falls dies hilft (z.B. mit Excel)
- Dokumentation ist wichtig, z.B.:

```
## Task 0
**Befehl:** `$ pwd`
**Ausgabe:** `/home/fabian`
**Erklärung:** Via pwd habe ich mir das aktuelle working directory ausgeben lassen.
```

# Hinweise Zu Assignment 1

## Exercise 5.3: Advanced Shell

- Aufgaben sind nicht nach Schwierigkeit sortiert, wenn Sie nicht weiter kommen, versuchen Sie die nächste Teilaufgabe!
- Praktischer Command, den Sie evtl noch nicht genutzt haben:
  - tr
    - tr REGEX1 REGEX2
      - Ersetzt in String Vorkommen von REGEX1 mit REGEX2 ZEICHENWEISE
      - z.B. echo "Hallo Welt" | tr H h
        - hallo Welt
      - tr -d REGEX
        - -d for delete characters
      - Matches characters
    - sed
      - sed 's/TERM1/TERM2/g' file
      - g at end for global (matches all occurrences)
      - z.B. echo "Hallo Welt" | sed 's/Hallo/Moin/g'
        - Moin Welt

# Allgemeine Hinweise

- Dokumentieren Sie! Was Sie nicht schreiben, können wir nicht bewerten. (Wir wollen aber auch nicht von jedem Klick lesen...)
- Wir bewerten die Dateien in ihrem Repository
  - Keine Sorge wegen Git-Chaos, solange alle geforderten Dateien vorhanden und richtig benannt sind
- Deadline ist der **26.11. um 20h**. Bitte einhalten!



# Open Questions



# Topics for Today

1. Q&A

2. kein Quiz

3. Assignment/ Help



# Quiz...

...zeit wird heute fürs Assignment genutzt!

# Topics for Today

1. Q&A

2. Quiz

2. Assignment/ Help