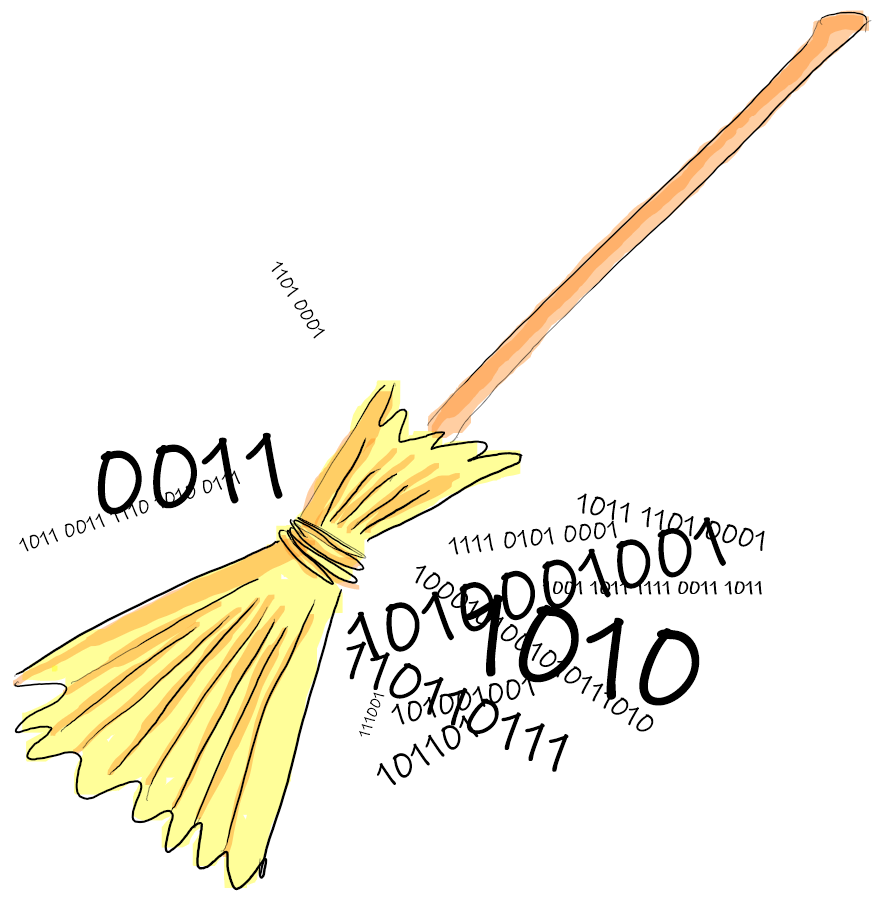
# Lern- und Arbeitsauftrag *LA\_162\_6903\_DatenAufbereiten.docx*

|  |  |
| --- | --- |
| Titel | Daten aufbereiten |
| Modul | 162 Informatiker/in EFZ |
| Autor / Version | Hanspeter Vogt, Jürg Haller, Alexander Flick,  überarbeitet von Lars Meyer / V2.1 |
| Hilfsmittel: | PR\_162\_DatenAufbereiten.pptx  LA\_162\_6903\_DatenAufbereiten-1\_Abgabe.txt  LA\_162\_6903\_DatenAufbereiten-2\_Abgabe.txt  (LA\_162\_6903\_DatenAufbereiten-3\_Abgabe.zip) |
| Nachweis | Leistungsbeurteilung zu Ende des Moduls |
| Sozialform | Einzelarbeit / Partnerarbeit |
| Leistungsziele | LZ2.1 / LZ2.2 |

## Ausgangslage

Hat Ihre Analyse der Datenqualität ergeben, dass behebbare Defizite bestehen, geht es nun ans Bereinigen.



## Aufgabenstellung

Sie eigenen Sich Werkzeuge und Methoden an, Daten so aufzubereiten, dass sie für die weitere Auswertung bereit sind.

### Vorbereitung: Theorie

Studieren Sie die Präsentation «PR\_162\_DatenAufbereiten».

### Teilaufgabe 1: Textfile aufbereiten

Oftmals liegen Daten nur in Textform (Papier oder Textdokument) vor und müssen für eine Auswertung mindestens in eine tabellarische Form überführt werden. Dabei müssen die Datentypen bestimmt und gesetzt werden.

Im vorliegenden Fall haben Sie die Datei *LA\_162\_6903\_DatenAufbereiten-1\_Abgabe.txt* erhalten mit dem Auftrag, diese aufzubereiten.

1. Analysieren Sie die Datenqualität der erhaltenen Daten und halten diese schriftlich fest.

Vollständigkeit: Mit dem Autofilter kann man nachschauen, ob alle Kantone enthalten sind. -> Es sind alle enthalten. Beim Durchschauen der Tabelle, tauchten auch keine Leeren Zeilen auf.

Aktualität: Die Daten stammen aus dem Jahr 2004 und sind daher nicht mehr aktuell.

Herkunft: Die Daten stammen vom CH Bundesamt für Statistik-> Vertrauenswürdige Quelle

Redundanz: Alle Spalten mit der Bezeichnung Total sind überflüssig, da sie eigentlich mit den anderen Spalten ausgerechnet werden können.

Konsistenz: Die ist eigentlich vorgegeben aber sie muss in der Kantonsspalte noch vervollständigt werden.

1. Wählen Sie eine sinnvolle Form für die Daten und begründen Sie ihre Wahl schriftlich.

Für diese Aufgabe wähle ich Excel, weil man die Kantone falls nötig einfach mit der Autovervollständigung fertigstellen kann.

1. Bereiten Sie die Daten für eine weitere Auswertung auf und halten Sie Ihre Vorgehensweise schriftlich fest.
2. Vergleichen Sie ihr Resultat mit dem des\*r Lernpartners\*in.

### Teilaufgabe 2: Textfile aufbereiten und ergänzen

Im vorliegenden Fall erhalten Sie von der Verwaltung der BBB die Stundenpläne des Schuljahres 06/07 im Textdokument *LA\_162\_6903\_DatenAufbereiten-2\_Abgabe.txt*.

1. Analysieren Sie die Datenqualität der erhaltenen Daten und halten diese schriftlich fest.
2. Importieren Sie die Datei in Excel.
3. Bereiten Sie die Daten auf.
4. Erstellen Sie eine weitere Spalte, die die Klassennamen auf Redundanzen überprüft (Formel nutzen). Dabei soll der Klassenname mit dem Beruf und dem Jahrgang verglichen werden.
5. Vergleichen Sie ihr Resultat mit dem des\*r Lernpartners\*in.

## Gütekriterien

Der Lern- und Arbeitsauftrag ist erfüllt, wenn …

…Sie ein Textfile in tabellarische Form gebracht und für die weitere Auswertung aufbereitet haben.

…Sie gegebene Daten mit zusätzlichen Hilfsspalten ergänzen können.

## Zusätzliche Angaben zum Auftrag

Für das Aufbereiten einer Excel-tabelle ist die Funktion Suchen & Ersetzen äusserst sinnvoll. Mittels Platzhaltern oder regulären Ausdrücken lassen sich auch komplexere Zelleninhalte verändern.

<https://support.office.com/de-ch/article/Suchen-und-Ersetzen-von-Text-mit-regul%C3%A4ren-Ausdr%C3%BCcken-fortgeschrittene-Anwender-eeaa03b0-e9f3-4921-b1e8-85b0ad1c427f?ui=de-DE&rs=de-CH&ad=CH>

*Quelle Stationenverzeichnis:*

<http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&_pageLabel=_dwdwww_spezielle_nutzer_energiewirtschaft_historisch&T26607173141161345039102gsbDocumentPath=Navigation%2FOeffentlichkeit%2FKlima__Umwelt%2FKlimadaten%2Fkldaten__kostenfrei%2Fstationsverzeichnis__weltweit__node.html%3F__nnn%3Dtrue>

*Quelle monatliche Klimadaten:*

<http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&_pageLabel=_dwdwww_spezielle_nutzer_energiewirtschaft_historisch&T26607173141161345039102gsbDocumentPath=Navigation%2FOeffentlichkeit%2FKlima__Umwelt%2FKlimadaten%2Fkldaten__kostenfrei%2Fkldat__welt__monat__node.html%3F__nnn%3Dtrue>

## Mögliche Erweiterungsaufträge

### Teilaufgabe 3 (optional): HTML-Files importieren, aufbereiten und als CSV exportieren

Der Deutsche Wetterdienst stellt für ausgewählte Messstationen aus der ganzen Welt die Klimadaten der wichtigsten Wetterelemente online und frei zur Verfügung.

Leider lassen sich spezielle Daten nur Online in Form von HTML-Tabellen anzeigen. Eine gezielte Auswertung soll mittels Excel realisiert werden. Dazu müssen die Daten überführt werden. Informationen zu den Daten sind in den beiden HTML-Dateien enthalten (*LA\_162\_6903\_DatenAufbereiten-3\_Abgabe.zip*).

1. Analysieren Sie die Datenqualität der erhaltenen Daten und halten diese schriftlich fest.
2. Überführen Sie die Klimadaten des Monats Dezember 2014 nach Excel und bereiten Sie die Tabelle auf. Bitte nutzen Sie nur deutsche Begriffe in der neuen Datei.
3. Exportieren Sie die aufbereitete Excel-Datei im csv-Format.
4. Überführen Sie das Stationenverzeichnis nach Excel und bereiten Sie die Tabelle auf. Bitte nutzen Sie nur deutsche Begriffe in der neuen Datei.
5. Exportieren Sie die aufbereitete Excel-Datei in eine csv-Datei.