

- Betriebliche Software hat nur begrenzte Auswertungsmöglichkeiten.
- Datenbankfunktionen werden gefürchtet statt genutzt.
- Tabellenkalkulation ermöglicht intuitives Arbeiten.
- Tabellenkalkulation bietet brauchbare Grafiken.
- Tabellenkalkulation passt gut zu Präsentationssoftware.

automatischer Abruf bei Bedarf

- Datenbankverbindung (z.B. mit ODBC)
- Webabfrage (SOAP...)

Austausch von Dateien

- Text mit fester Breite
- CSV
- Edifact
- XML

feste Breite Text steht in gleich breiten Spalten

CSV Felder sind unterschiedlich breit, mit Trennzeichen dazwischen

EDIFACT Genormtes CSV-Format für wirtschaftliche Anwendungen

XML genormter Format-»Baukasten«

- breit verfügbar
- nur lose definiert
- maschinelle Datenprüfung schwierig

Beispiel 1

<i>Kunde;Datum;Betrag</i>
Maier;13.11.2017;99,00 €

Beispiel 2

<i>Customer;Date;Amount</i>
Maier,11/13/17,EUR 99.00

Beispiel 3

<i>Kunde;Tag;Wert</i>
"Maier";2017-11-13;99

- Ursprünglich *Comma-separated Values*: Felder werden mit Kommas getrennt.
- Zeichensätze unterscheiden sich nach Betriebssystem, Region und Sprache.
- Auch Dezimaltrennzeichen, Tausendertrenner, Datumsformate unterscheiden sich nach Region und Sprache.
- Was macht man, wenn der Feldtrenner in einem Feld vorkommt?

CSV schreiben ist einfach, CSV lesen ist kompliziert.

- basiert auf CSV
- weltweite Norm für Datenaustausch im Handel
- genau festgelegte Datenbeschreibung
- umfangreicher Standard (über 500 Seiten)
- aufwendig umzusetzen
- nicht jeder erdenkliche Geschäftsvorfall kann abgebildet werden

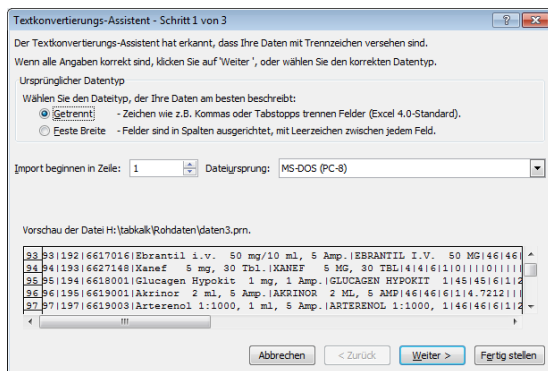
Sie finden ein Beispiel im Wikipedia-Artikel zu EDIFACT.

- XML (*extensible markup language*) ist ein *Baukasten* für Dateiformate.
- Es gibt Werkzeuge für automatische Datenprüfung und Konvertierung.
- Das erstmalige Einrichten eines Dateiformates ist aufwendig.
- Wiederholte Nutzung ist elegant und schnell.
- XML ist die Zukunft – seit 20 Jahren.

```
<Buchung>  
  <Kunde>Maier</Kunde>  
  <Datum>  
    <Tag>13</Tag>  
    <Monat>11</Monat>  
    <Jahr>2017</Jahr>  
  </Datum>  
  <Betrag Waehrung="Euro">99.00</Betrag>  
</Buchung>
```


CSV in Excel importieren I

- Gehen Sie auf *Datei öffnen*. Stellen Sie den Datentyp auf *Textdatei*.
- Im ersten Schritt des Assistenten prüfen Sie,
 - ob die Daten feste Breite oder Trennzeichen haben,
 - ob die Umlaute stimmen,
 - ob es Kopfzeilen gibt, die Sie nicht wollen.



- Unerwünschte Kopfzeilen können Sie auch nach dem Import löschen.
- Wenn es verschiedene Feldtrenner gibt oder Sie den falschen gewählt haben, arbeiten Sie nach dem Import mit *Daten/Text in Spalten* nach.
- Der Zeichensatz *muss* stimmen, das lässt sich im Nachhinein nicht reparieren.

Es gibt historisch bedingt unterschiedliche Zeichenkodierungen. Wird beim Import die falsche Kodierung eingestellt, sind Umlaute (äöü) und Sonderzeichen nicht lesbar.

- Suchen Sie in der Vorschau eine Zeile mit Umlauten (falls in den Daten welche vorkommen können) oder Sonderzeichen.
- Probieren Sie geeignete Zeichenkodierungen aus, bis die Darstellung stimmt.
- Notieren Sie sich die Einstellung für das nächste Mal.

In Deutschland sind die folgenden Zeichensätze verbreitet:

DOS850 bei älterer Software

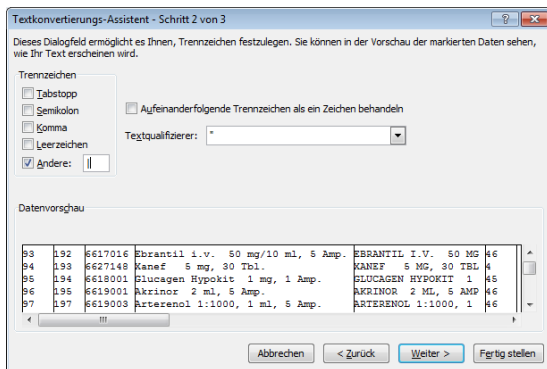
Windows-1252 bei Programmen aus der Microsoft-Welt

UTF-8 bei Daten aus Java-Programmen, Linux-Systemen und dem WWW.

Wenn in Ihrem Importmaterial Umlaute mit zwei Zeichen dargestellt werden, wählen Sie UTF-8.

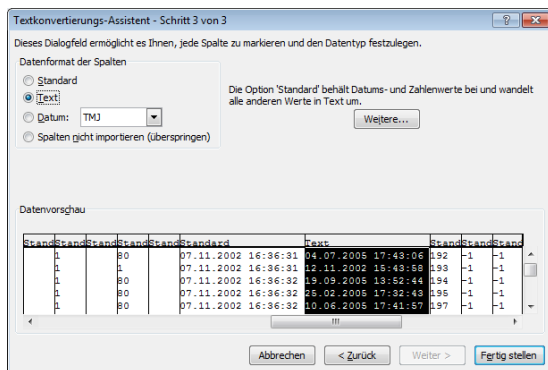
CSV in Excel importieren III

- Im zweiten Schritt definieren Sie die Trennzeichen oder setzen die Spaltentrenner.



CSV in Excel importieren IV

- Im dritten Schritt können Sie das Format einzelner Spalten gezielt auswählen.
 - Falls nichts geeignetes dabei ist, importieren Sie die betroffene Spalte als *Text*.
 - Falls Sie eine Spalte nicht brauchen, überspringen Sie sie.



Manche Spaltenformate werden von Excel falsch interpretiert:

- Eine Zahl mit Punkt (wie 1.2) wird als Datum gedeutet.
- Ein Gleichheitszeichen am Feldanfang wird als Formel interpretiert.
- Bei Datumswerten können Sie das richtige Datumsformat aussuchen.

Importieren Sie falsch interpretierbare Spalten als *Text*.

Sie können ein Datum nicht rückwirkend in eine Gliederungsangabe oder Bestellnummer zurückwandeln, Formeln können Sicherheitslücken ausnutzen.

- Wenn Überschriften fehlen: Geben Sie jeder Spalte eine Überschrift.
- Geben Sie Ihrem Datenbereich einen Namen.
- Kontrollieren Sie die Daten.
- Falls Sie eine Spalte nicht korrekt bearbeiten konnten:
 - Nach rechts bringen und mit *Text in Spalten* zerlegen.
 - Mit Hilfe von Excel-Funktionen bearbeiten.
- Bei groben Import-Fehlern: Import wiederholen.
- Falls Sie einen Import öfters machen müssen: Dokumentieren Sie Ihre Schritte.

- Neuere Excel-Versionen überspringen den Assistenten, wenn der Dateiname mit .csv endet.
Abhilfe: Ändern Sie die Endung vor dem Import auf .prn oder .txt, falle es nicht passt.
- In schweren Fällen bearbeiten Sie die Datei vorab mit einem Texteditor (nicht mit einer Textverarbeitung wie MS Word!).
- Falls Sie das öfters machen müssen: Lernen Sie eine Skriptsprache wie *Perl* und automatisieren Sie die Vorverarbeitung damit.

Feature-Liste

- Umgang mit verschiedenen Textkodierungen
- Suchen und Ersetzen mit Mustern
- Flüssiges Arbeiten auch mit umfangreichen Daten
- Makro-Fähigkeit

Vorschlag: Notepad++ für Windows

- Löschen Sie überflüssige Zeilen.
- Geben Sie Ihren Daten einen Namen, falls noch nicht geschehen.
- Fixieren Sie die erste Zeile:
So können Sie die Daten durchblättern.
- Schalten Sie den Filter ein:
Damit können Sie Teilmengen untersuchen.
- Vor dem Sortieren alle Daten markieren, ankreuzen, dass die Daten Überschriften haben.

Rohdaten

```
Art;Text;Betrag
K;1009;02.04.2019
E;Pommes;2,20
E;Pizza Marg.;4,95
E;Döner veget.;3,95
E;Pizza Marg.;4,95
S;4;16,05
K;1010;02.04.2019
E;Döner groß;4,50
T;Cola;1,95
E;Döner groß;4,50
S;3;10,95
...
```

- Feldtrenner Strichpunkt, Dezimaltrenner Komma, Umlaute vorhanden
 - Vier Arten von Zeilen:
 - K Belegkopf: Belegnummer und Datum
 - E Lebensmittel: Name und Preis
 - T Getränk: Name und Preis
 - S Summenzeile: Artikelzahl, Summe
- Die Zeilenarten müssen getrennt ausgewertet werden.
- Spalte *Text* enthält Artikelnamen, Belegnummern und Artikelanzahl.
 - Spalte *Betrag* enthält Einzelpreis, Summe und Belegdatum.

- Stellen Sie die Zeichenkodierung so ein, dass die Umlaute richtig angezeigt werden.
- Stellen Sie den Feldtrenner auf Strichpunkt (Semikolon) und den Dezimaltrenner auf Komma.
- Falls die dritte Spalte Ärger macht:
Datentyp von Spalte 3 auf Text setzen, nochmals importieren.

Rohdaten

```
Art;Text;Betrag
K;1009;02.04.2019
E;Pommes;2,20
E;Pizza Marg.;4,95
E;Döner veget.;3,95
E;Pizza Marg.;4,95
S;4;16,05
K;1010;02.04.2019
E;Döner groß;4,50
T;Cola;1,95
E;Döner groß;4,50
S;3;10,95
...
```

Importierte Daten

Art	Text	Betrag
K	1009	02.04.2019
E	Pommes	2,20
E	Pizza Marg.	4,95
E	Döner veget.	3,95
E	Pizza Marg.	4,95
S	4	16,05
K	1010	02.04.2019
E	Döner groß	4,50
T	Cola	1,95
E	Döner groß	4,50
S	3	10,95
...

Falls Sie die dritte Spalte als Text importieren mussten:

- DATWERT() wandelt eine Datumsangabe in ein Datum.
- ZAHLENWERT() wandelt einen Text in eine Zahl.
- Mit WENN() können Sie entscheiden, welche Funktion angewandt werden soll.

```
=WENN(A2="K";Datwert(C2);Zahlenwert(C2;""))
```

- Alternative: Zeilen sortieren, einzelne Formeln in die korrekten Zeilen kopieren.
- Falls die Zahlen an den ursprünglichen Platz sollen: Ergebnisse markieren und mit *Inhalte einfügen* in Spalte C kopieren.

- Filtern *oder*
- Daten sortieren, Blöcke getrennt bearbeiten
 - Fügen Sie vor dem Sortieren Zeilennummern ein (z.B. Spalte D).
 - Sortieren Sie die Zeilennummern mit.
 - Wenn Sie die Original-Reihenfolge nochmals brauchen:
Sortieren Sie nach den Zeilennummern.

oder

- Mit Funktionen wie `ZählenWenn()` und `SummeWenn()` arbeiten.
- Pivot-Tabelle benutzen.

Tabellenkalkulation wird manchmal als Zwischenstation verwendet, um z.B. Daten aus dem alten Warenwirtschaftssystem für das neue aufzubereiten.

Schwierigkeiten

- Auswahl geeigneter Trennzeichen
- Auswahl der richtigen Zeichenkodierung
- Verlust von leeren Zellen am Zeilenende

Problem

- Leere Zellen am Zeilenende werden nicht exportiert.
- Datenbanksoftware bricht das Einlesen ab, weil die Zeilen als unvollständig gelten.

Lösung

- Fügen Sie eine Spalte mit einem festen Text an, der sonst in den Daten nicht vorkommt.
- Öffnen Sie die Datei nach dem Export in einem Editor und entfernen Sie die Dummy-Spalte mittels *Suchen und Ersetzen*