Access Datenimport

Viele Daten existieren bereits irgendwo in einer anderen Datenbank.

Um diese Daten nicht noch einmal manuell eingeben zu müssen, sondern zwischen IT Systemen auszutauschen, hat jedes Datenbankprogramm Möglichkeiten zum Export und Import von Daten. Das erspart das Abschreiben der Daten, wenn man sie in einem anderen Computer benötigt. Damit der Austausch klappt, müssen exportierende und importierende Datenbank sich auf eine gemeinsame Struktur der Übertragungsdaten einigen.

Der Exportierende ist also gefragt, die Daten so zur Verfügung zu stellen, dass er dem Importierenden die Sache möglichst leicht macht. Bevor man den Export der Daten vornimmt, ist es empfehlenswert, sich gemeinsam zu einigen, welche Importformate das Zielsystem unterstützt.

Das am weitesten verbreitete Übertragungsformat ist vermutlich „csv“, die dafür verwendeten Dateien fallen unter den Oberbegriff "Textdatei". Die Datensätze werden zeilenweise geschrieben, jede Zeile ist also ein kompletter Datensatz. Die Datenfelder werden getrennt durch ein vorab vereinbartes Trennzeichen. Üblich sind Komma (,), Semikolon (;) und Tabulator (Tab).

Das .csv Format kennt zahllose Varianten und Einstellungen, die beim Export und Import gleich eingestellt werden müssen. Access versucht zwar, die Einstellungen an Hand der zu importierenden Daten zu „erraten“, das klappt aber in der Praxis selten reibungslos.

Daher bietet Access einen Import-Assistenten, der eine "Vorschau" der Daten bietet. Vor dem eigentlichen Import kann man überprüfen, ob die gewählten Einstellungen zum gewünschten Ergebnis führen. Man korrigiert die Einstellungen so lange, bis die Vorschau richtig aussieht.

Beim eigentlichen Import schreibt Access ein Fehlerprotokoll mit, in dem man nachlesen kann, ob es beim Import Probleme gab. Meistens steht auch ein Hinweis dabei, was das Problem war (z.B. der Versuch, einen Text in ein Datenfeld zu importieren, das nur Zahlen oder Datumswerte aufnehmen kann).

Häufige Fehlerquellen sind:

- Falscher Zeichensatz. Die Darstellung von Sonderzeichen und Umlauten wurde nie abschließend festgelegt, der Export- und der Importrechner müssen von Hand auf denselben Zeichensatz ("Codepage") eingestellt werden. Zu erkennen ist dieses Problem daran, dass nach dem Import statt den deutschen Umlauten (äöüß) irgendwelche anderen Zeichen erscheinen.
- Falsche Zahlenwerte. Die Darstellung von Zahlen wurde ebenfalls nie abschließend festgelegt. Im deutschsprachigen Raum wird als Dezimaltrennzeichen das Komma "," z.B. 123,45) verwendet, auf der übrigen Welt ist der Punkt "." z.B. 123.45) gebräuchlich. Darüber hinaus gibt es bei gewöhnlichen Zahlen ("Zahl") eine maximale Zahlengröße von ungefähr 2.000.000.000 (eine zwei mit zusätzlich 9 Nullen hinten dran). Ab da wird es gefährlich, dass die Zahl zu groß wird für einen reibungslosen Import. Für extrem große Zahlen gibt es daher einen eigenen Zahlentyp ("Große Zahl").

  Eine weitere Fehlerquelle sind Felder, die allgemein als "Zahlen" bezeichnet werden, die aber eigentlich keine sind, weil sie Buchstaben enthalten können. Beispiele sind

© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0

Hausnummern ("11c" ist keine gültige Zahl), Telefonnummern, Kontonummern, Postleitzahlen mit Länderkennzeichen (z.B. Wien = "AT-1010"), usw. Solche Zahlen kann man besser als "Kurzer Text" speichern.

- Falsche Datumswerte. Auch die Darstellung von Datumsangaben wurde nie verbindlich festgelegt. Im deutschsprachigen Raum sind sowohl Tag.Monat.Jahr (3.4.2002) als auch Jahr-Monat-Tag (2002-4-3) gebräuchlich. In den USA ist das Standardformat Monat/Tag/Jahr üblich (4/3/2002).

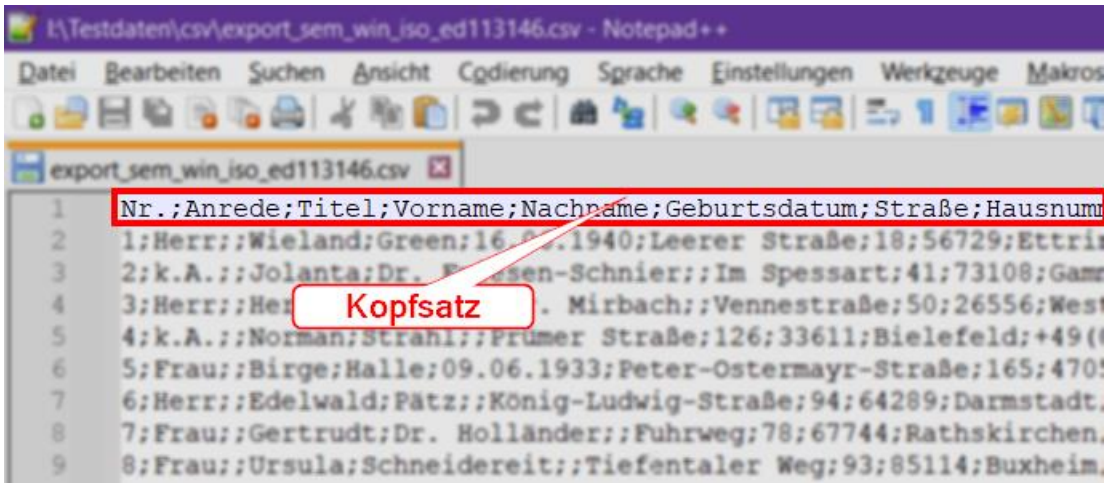
  Falsch importierte Datumswerte sind erst einmal schwer zu sehen, weil sie auf dem Bildschirm richtig aussehen. Man bemerkt das Problem erst dann, wenn man nach Datum filtern möchte, und statt den mächtigen Datumsfiltern nur einfache Textfilter verwenden kann. Mit Textfiltern ist keine vernünftige Verarbeitung von Datumswerten möglich.

- Zu lange Texte. Bei Texten gibt es ein ähnliches Problem wie bei Zahlen. Standardtexte ("Kurzer Text") können maximal 255 Zeichen lang werden. Für extrem lange Texte gibt es ein eigenes Datenformat ("Langer Text", "Memo").

Der Kopfsatz

Derjenige, der die Datei erstellt, kann sich entscheiden, einen "Kopfsatz" mitzugeben. Dabei wird in der ersten Zeile der Name des Datenfeldes mit übertragen.

Für den Import hat das den Vorteil, dass die Felder korrekte Feldnamen bekommen.



Der Kopfsatz steht immer in der ersten Zeile, darunter folgen dann die eigentlichen Daten, also, im Bild oben, die Adressen von Herrn Wieland, Frau Dr. Jolanta Friesen-Schnier, und so weiter.

Ein Kopfsatz muss nicht unbedingt mitgegeben werden, wenn er fehlt, muss man die Datenbankstruktur von Hand anlegen. Fehlt der Kopfsatz, nummeriert Access die Felder einfach durch, sie heißen dann "Feld1", "Feld2", "Feld3" usw. Das ist lästig, wenn man die Daten später weiterbearbeiten möchte, z.B. ein Formular daraus machen, oder einen Filter definieren.

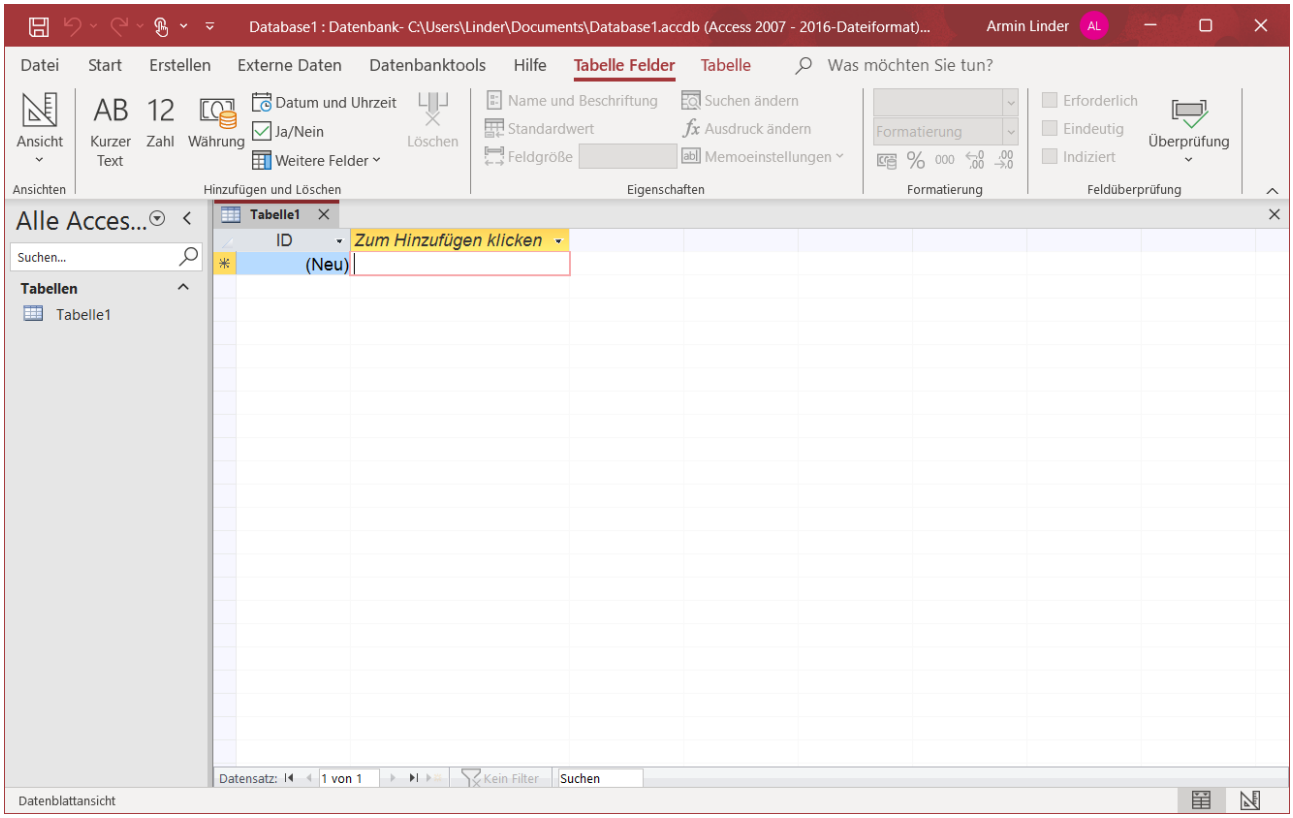
Tipp: man kann die Felder auch noch nach dem Import in der "Entwurfsansicht" der Tabelle umbenennen.

© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0

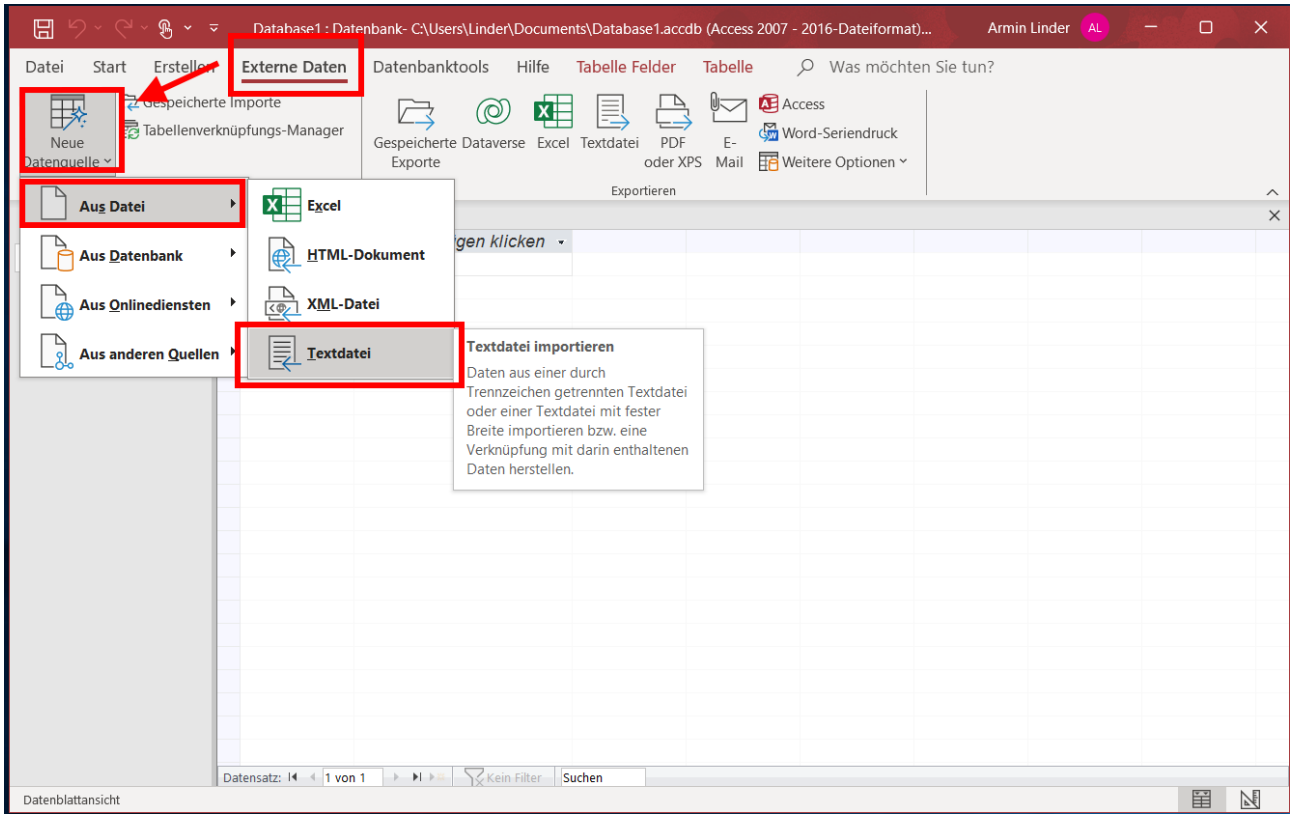
Zur Praxis: ein Import-Beispiel

Wir haben also eine .csv Datei erhalten. Nun wollen wir sie in eine Datenbank importieren. Geh mit mir durch das Beispiel, und lass uns Schritt für Schritt machen und dabei mitdenken, wozu er gut ist.

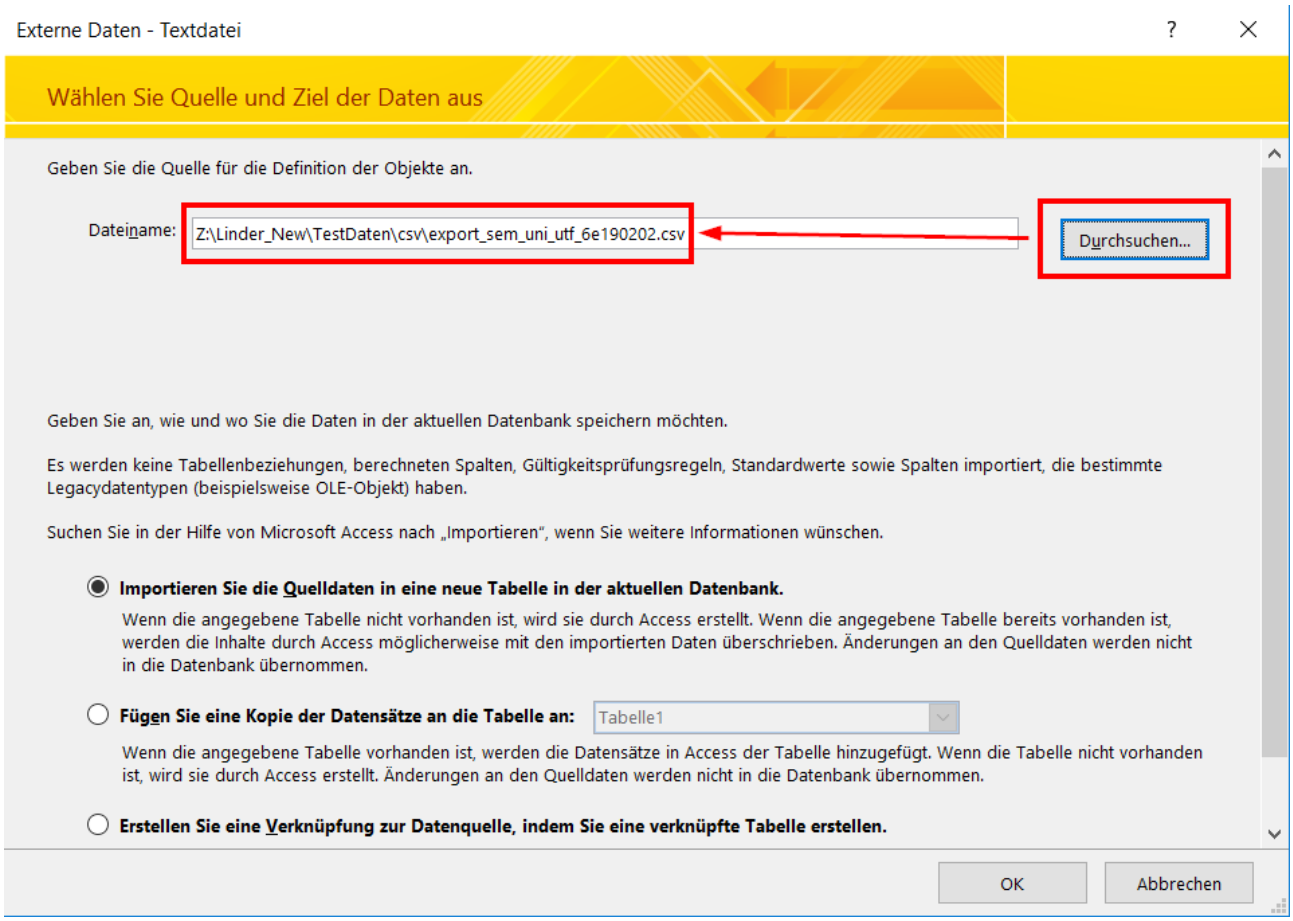
Ausgangssituation. Du hast eine leere Datenbank erstellt. In die wollen wir nun die csv Datei importieren.



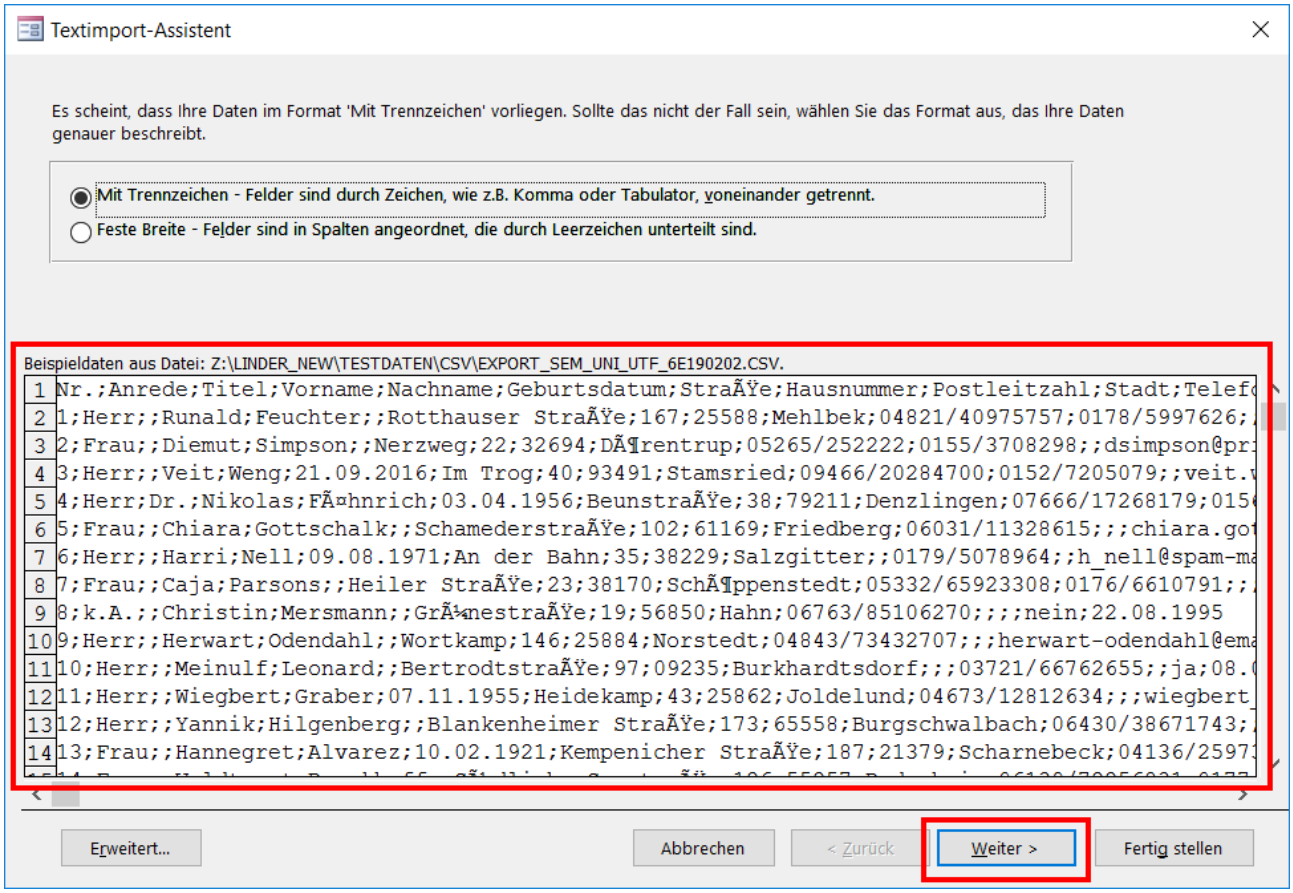
Wähle dazu im Menü "Externe Daten" ganz links "Neue Datenquelle" und danach "Aus Datei", "Textdatei".



Gib den Namen der Importdatei an, oder wähle ihn mit der Schaltfläche "Durchsuchen" aus.



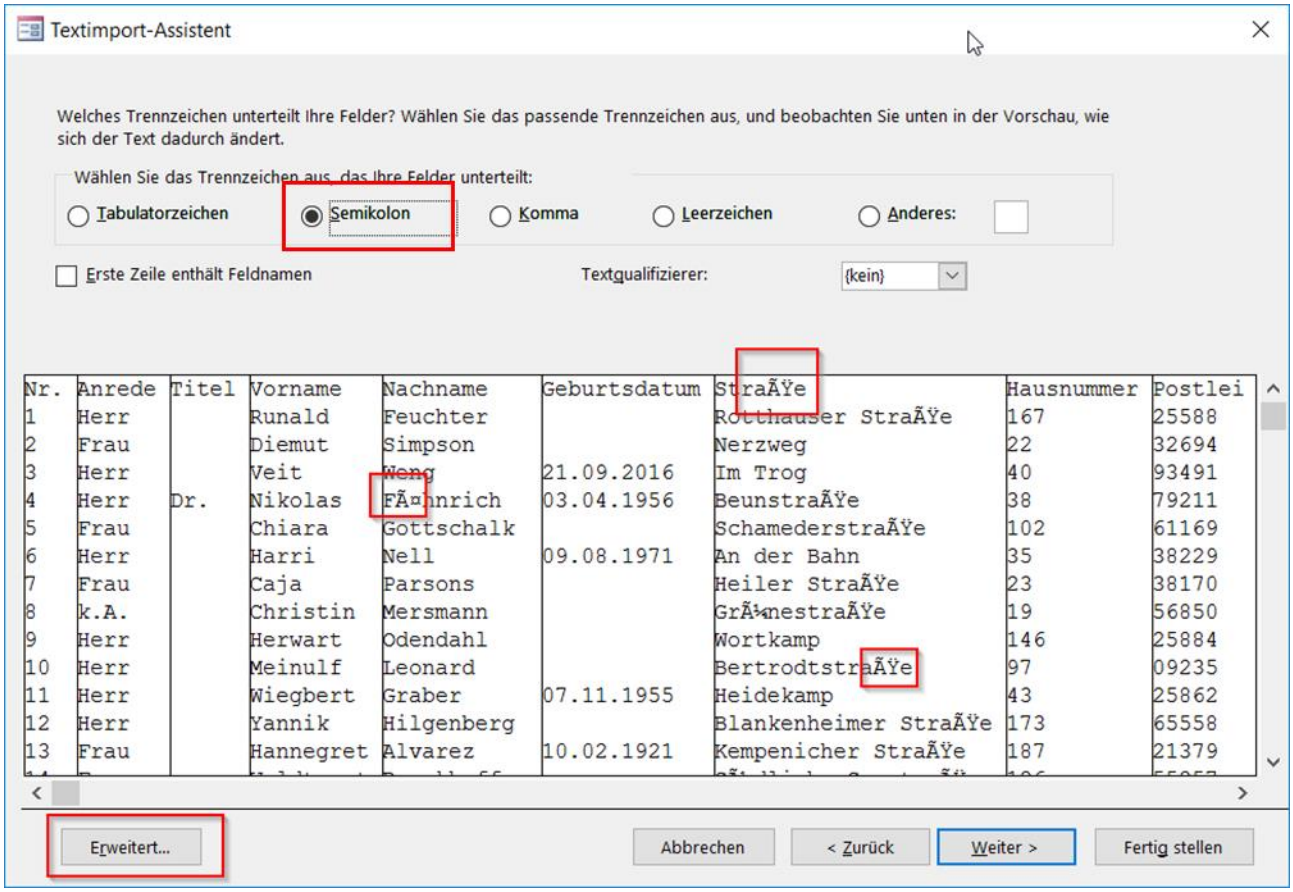
Klicke auf "Ok". Jetzt wird die Datei geöffnet, und die Vorschau-Ansicht für den Import erscheint. Hier müssen wir je nach Datei noch Einstellungen machen. Dabei hilft uns das Vorschaufenster.



Erst mal verschaffen wir uns eine grobe Übersicht. Ok, es ist offenbar eine csv Datei. Check. In der ersten Zeile stehen Feldnamen im Kopfsatz, super, wir müssen die Feldnamen also nicht von Hand eingeben. Check. Das Trennzeichen ist offenbar das Semikolon (";"). Check. Und wir sehen auffällige Feldnamen, die uns warnen, dass da vielleicht Fallen lauern könnten: Zahlenfelder ("Hausnummer") und Datumsfelder ("Geburtsdatum") sind immer gut für Probleme. Bei denen passen wir beim nächsten Schritt besonders auf. Also klicken wir "Weiter".

Die Vorschau zeigt: da stimmt wohl noch einiges nicht. Lass es uns Schritt für Schritt korrigieren. Schau immer wieder die Vorschau an, du wirst sehen, dass sie immer besser wird.

© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0

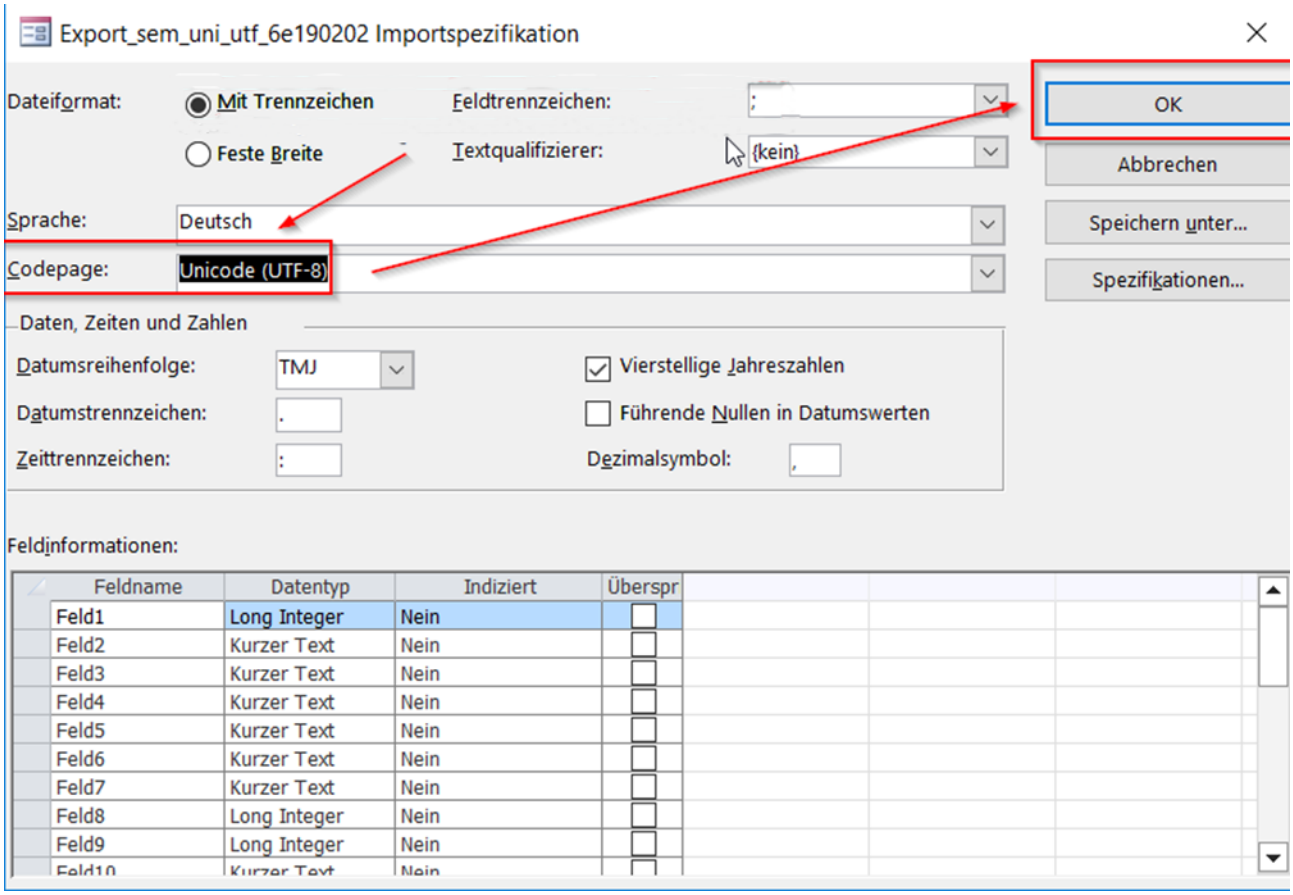


Wir checken schnell ganz oben das Trennzeichen … Semikolon (";"), ok, das passt. Sonst wären die Spaltengrenzen völlig durcheiniander, aber die scheinen zus timmen.

Zuallererst schauen wir nach dem Zeichensatz. Man sieht sofort: die deutschen Umlaute sind alle verhunzt. Im konkreten Fall stammt die Datei aus dem Internet, und ist wie vieles was aus dem Internet stammt mit dem Zeichensatz "Unicode (UTF-8)" codiert. Access vermutet offenbar irgend einen anderen Zeichensatz, daher die verhunzten Umlaute.

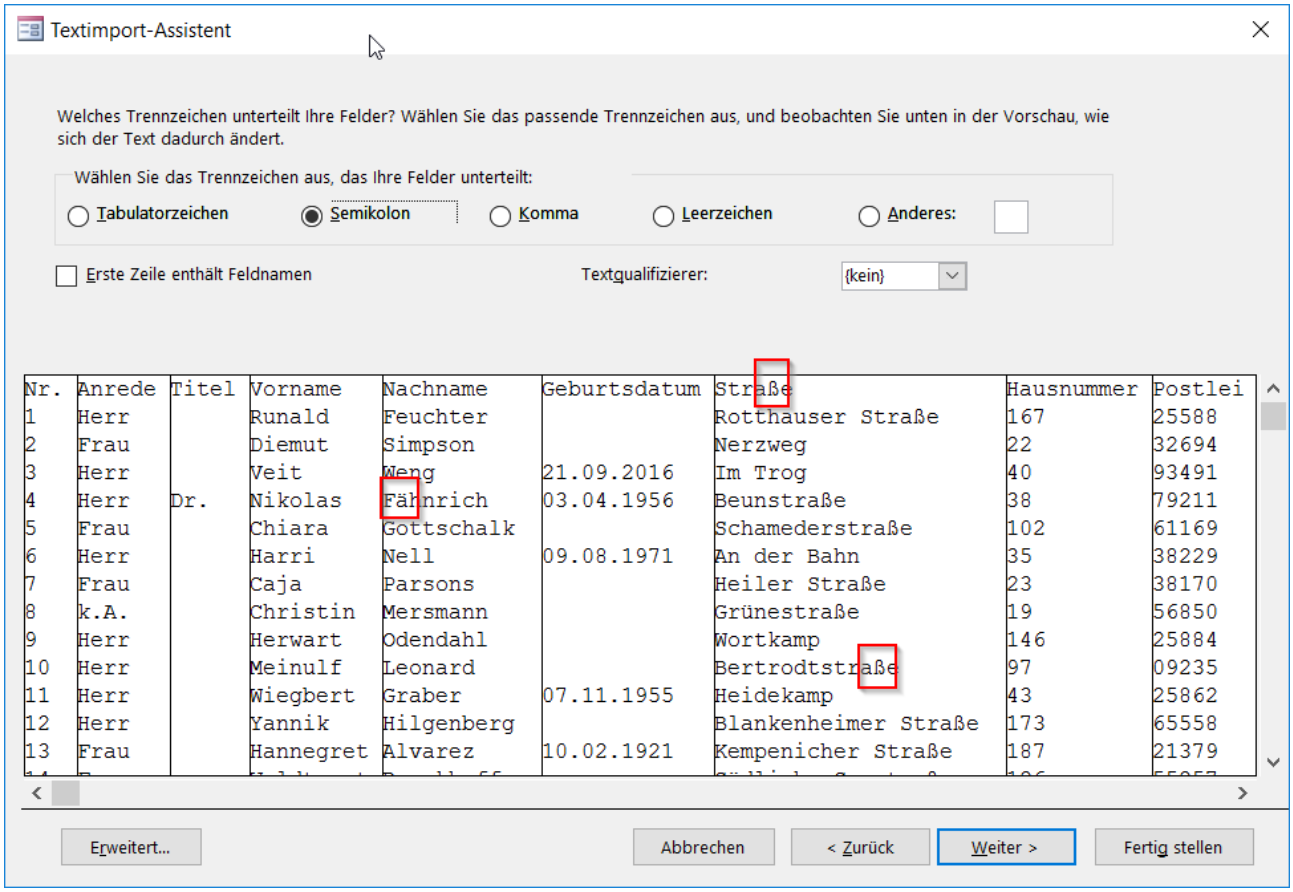
Klicke auf die Schaltfläche "Erweitert", um den koprrekten Zeichensatz einzustellen.

© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0



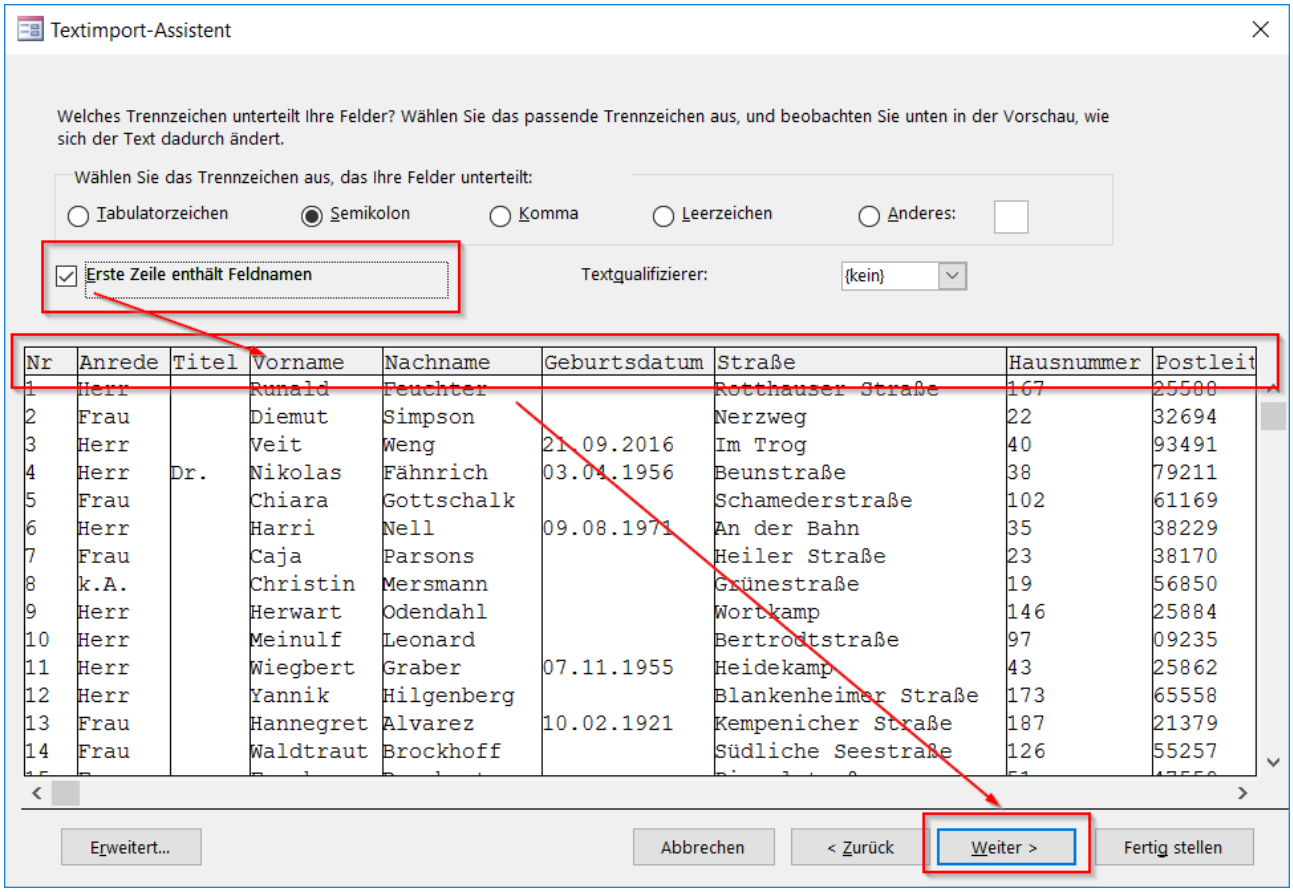
Darunter ändern wir die Codepage in "Unicode (UTF-8)". Man könnte jetzt noch mehr machen, was ich später in eigenen Schritten zeige, aber lass uns Schritt für Schritt arbeiten. Mehr gibt es hier im Moment nicht zu tun, also klicken wir auf "Ok."

© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0

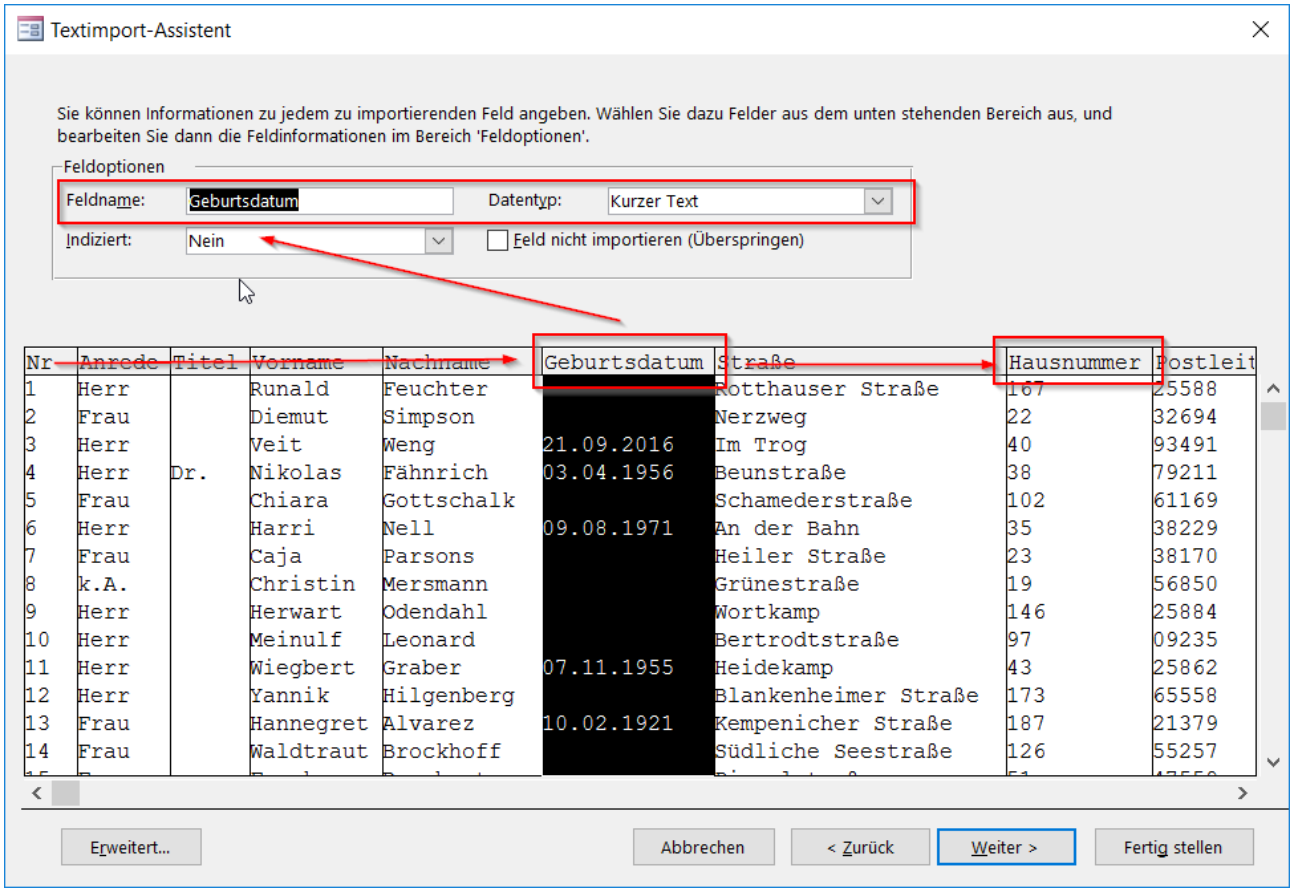


Jetzt schaut die Vorschau schon besser aus: die Umlaute sind korrekt, und die Aufteilung in Spalten auch. Also weiter, Access hat auch den Kopfsatz nicht richtig erkannt. Aktiviere die Schaltfläche "Erste Zeile enthält Feldnamen", und die Vorschau zeigt den Kopfsatz korrekt (grau hinterlegt) an.

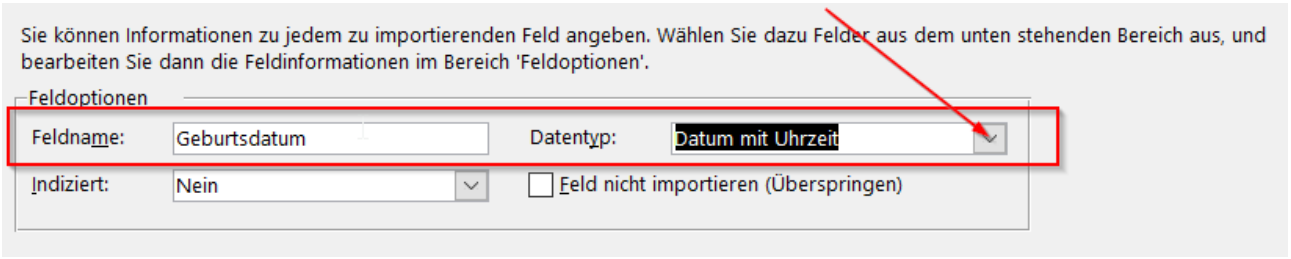
© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0



Der folgende Dialog gibt dir die Möglichkeit, Feld für Feld zu prüfen, ob der richtige Datentyp erkannt wurde.



Wir werden beim Geburtsdatum fündig: Access vermutet "kurzer Text", es ist aber ein Datum. Klicke auf den Pfeil rechts von "kurzer Text" und stelle den Datentyp auf "Datum mit Uhrzeit" um.

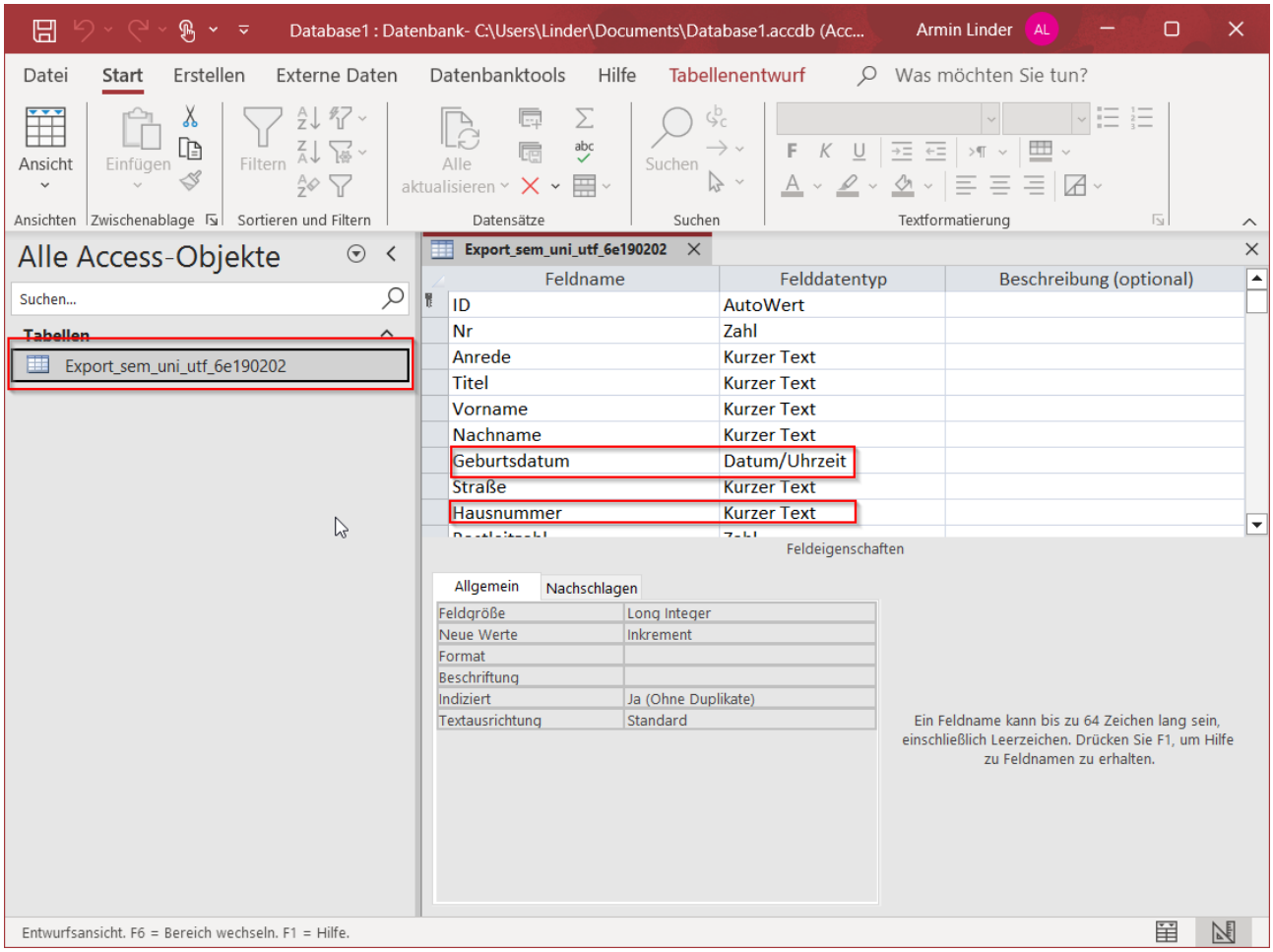


Auch beim Feld "Hausnummer" solltest du nachbessern. Access vermutet "kurze Zahl", aber da es Hausnummern mit Buchstaben gibt (z.B. 11a), würden dann alle Datensätze, in denen so ein Buchstabe vorkommt, beim Import als fehlerhaft betrachtet und übersprungen.

Es folgen nun noch weitere Dialoge, die für größere Datenbanken nützlich sein können, aber für unser erstes Importbeispiel haben wir alles eingegeben was notwendig ist. Wir kürzen daher ab, und klicken ganz rechts unten auf "Fertig stellen".

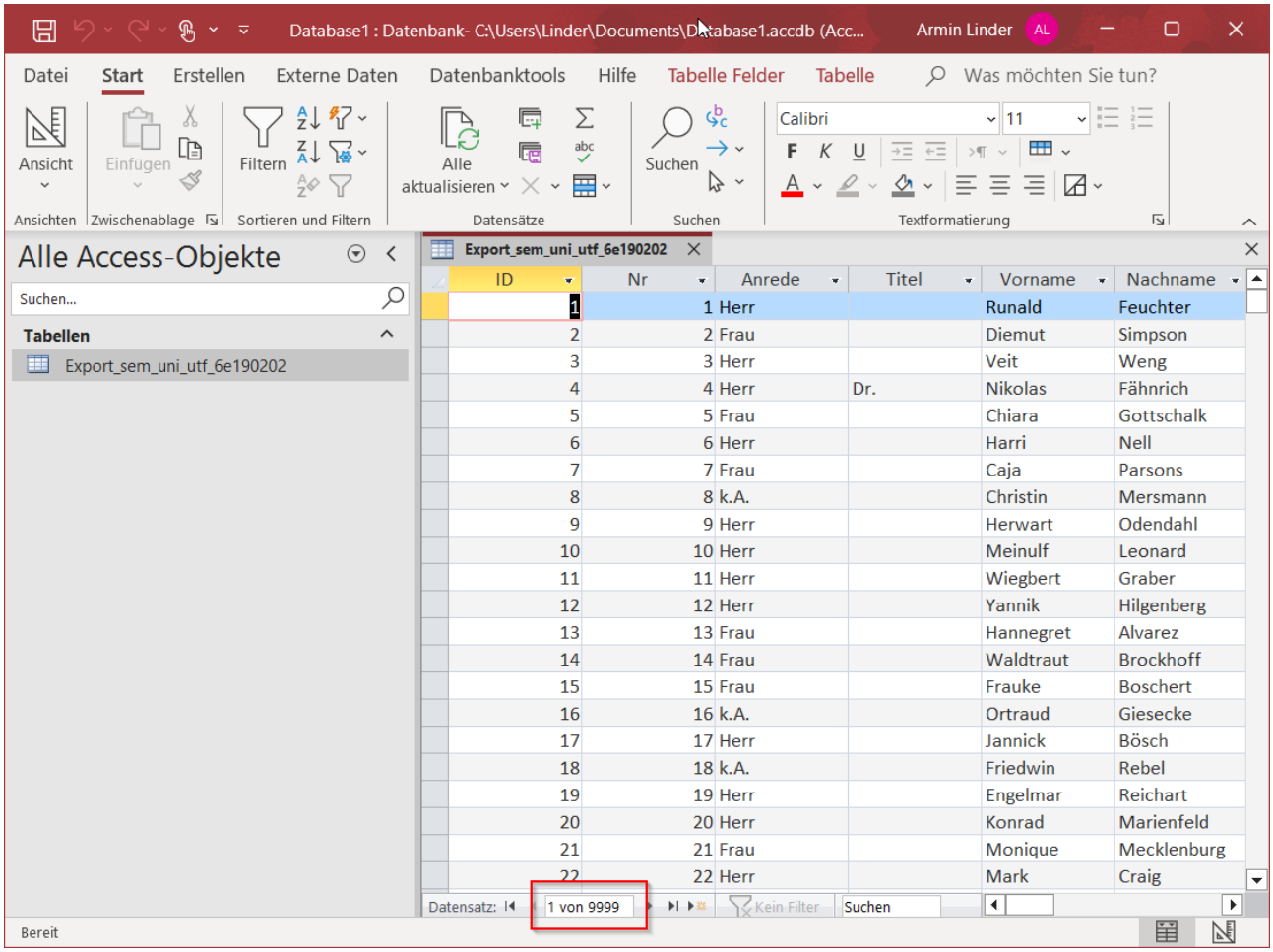
Letzter Check: wir doppelklicken auf die importierte Tabelle, und wählen "Start" – "Ansicht" – Entwurfsansicht. Prima, alle Feldnamen sind korrekt, und die Datentypen auch. Man könnte hier ggf. auch noch kleinere Fehler korrigieren.

© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0



Und ein Blick auf die Datenbank mit Hilfe von "Start" – "Ansicht" – "Datenblattansicht" zeigt uns eine wunderschöne Datenbank mit nicht weniger als 9999 Datensätzen.

© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0



Import gelungen, fertig! Jetzt können wir wie schon im letzten Satz Arbeitsblätter gelernt weitermachen und weitere Daten eingeben, Daten suchen und filtern, Formulare zeichnen und so weiter.

© 2023 Comeniusschulen Töging
18.11.2023/lia Lizenz: CC BY-NC-SA 4.0