

# Python

- Scriptsammlung
  - dump-old-shit.py
- Snippets
  - uniq und sort

# Scriptsammlung

# dump-old-shit.py

Dieses Script war mal nötig, um aus einem SQL-Dump alle INSERTs, die vor einem bestimmten Zeitpunkt (Feld timestamp) eingefügt wurden, zu löschen. Die INSERT-Zeilen beinhalten ein Feld namens „Timestamp“, das der Variable timestamp entspricht. Als Feldtrenner wird das Komma verwendet. Zeilen, die nicht mit INSERT beginnen werden unverändert ausgegeben:

```
#!/usr/bin/python

import sys,string

timestamp=20060707000000

file = open("datenbankdump.sql", "r")
while 1:
    line = file.readline()
    if not line: break

    # Zeilen ausgeben, die nicht mit INSERT beginnen
    if 'INSERT' not in line:
        sys.stdout.write(line)
        continue

    # Zeilen mit Timestamp kleiner Dingens auswerfen
    words = string.split(line,',')
    if len(words) >=9:
        if string.atoi(words[10]) >= timestamp:
            sys.stdout.write(line)
```

# Snippets

# uniq und sort

Um in Python die doppelten Einträge aus einer Liste zu filtern können folgende Snippets verwendet werden:

```
def uniq(alist)  # Fastest order preserving
    set = {}
    return [set.setdefault(e,e) for e in alist if e not in set]
```

```
def uniq(alist)  # Fastest without order preserving
    set = {}
    map(set.__setitem__, alist, [])
    return set.keys()
```