

reviews-csv-gen

April 24, 2019

1 Reviews-to-CSV-Generator

In this Jupyter-Notebook simply transformes the given CSV into the format of the sqlite3-database from django.

Content: - Importing necessary packages - Read the data - Add Author field - Saving to CSV

1.1 Importing necessary packages

```
In [1]: import pandas as pd
import json
import nltk
import random
#import numpy as np
#nltk.download("punkt")
import string
exclude = set(string.punctuation)
import time
from tqdm import tqdm
import re
```

1.2 Read the data

```
In [2]: data = pd.read_csv('Datensatz_Coding_Challenge.csv', delimiter=";")
corpus = data.copy()
```

```
In [6]: corpus.head()
```

```
Out[6]:
```

	StyleID	text	rating	author	\
0	1709054	Die sind okay und dann für den Preis.	5	Unkown	
1	1709054	Qualität und Preis sind gut. Leider sind sie z...	3	Unkown	
2	8623725	lässt schlanker aussehen	5	Unkown	
3	8623725	Material und Farbe gut. Da einige Kundinnen in...	3	Unkown	
4	9743730	Material ist schön zum verdunkel. Leider doch...	5	Unkown	

	date
0	2019-01-28 15:05:07
1	2019-01-28 15:05:07

```
2 2019-01-28 15:05:07
3 2019-01-28 15:05:07
4 2019-01-28 15:05:07
```

1.3 Add Author field

```
In [5]: corpus["author"] = "Unknown"
        corpus["date"] = "2019-01-28 15:05:07"
```

1.4 Saving to CSV

```
In [9]: review = corpus[["author", "text", "date", "rating", "StyleID"]]
        review.to_csv("reviews.csv", sep=";")
```