

Курсовой проект на тему: «Конвертация данных из DataFrame в таблицу БД Oracle»

Выполнил:

студент группы Т12О-101М-20

Чувашов Владислав Максимович

Руководитель:

Поповкин Александр Викторович

Постановка задачи

- * Продемонстрировать конвертацию данных из файла формата .csv в Dataframe;
- * Сравнить эти данные с другим файлом формата .csv ;
- * Получить Dataframe, который содержит только общие данные;
- * Загрузить полученный Dataframe в таблицу базы данных Oracle.

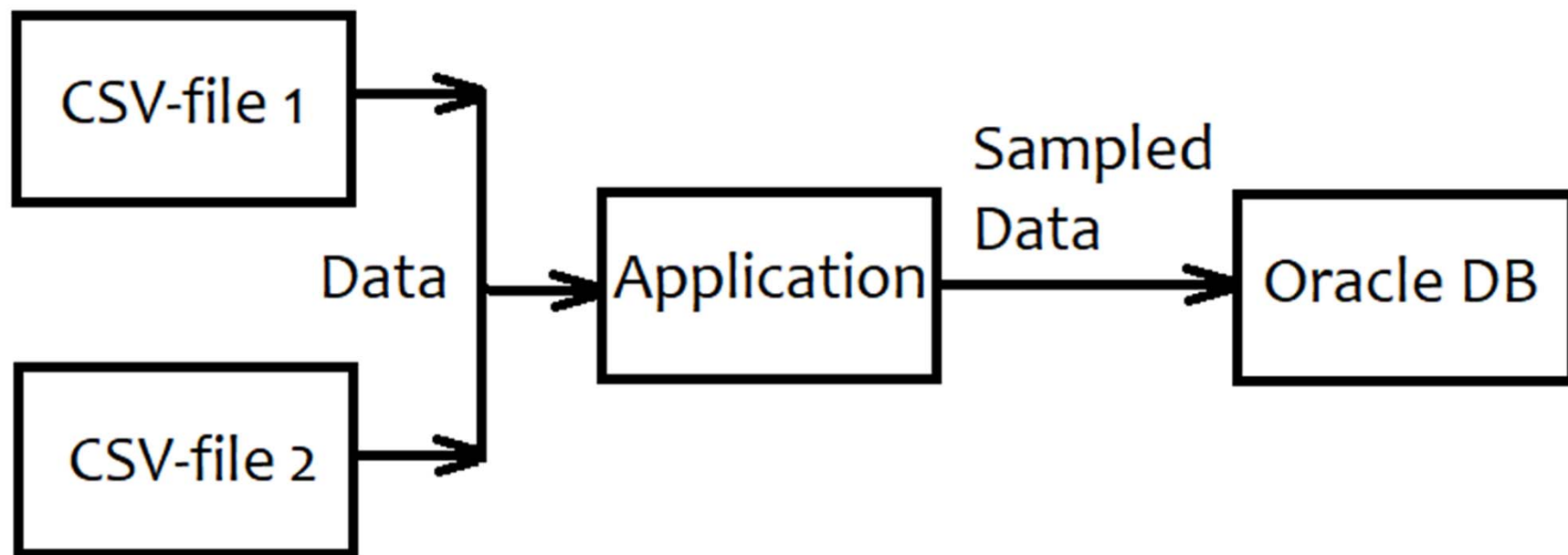
Бизнес - логика

- * Необходимо импортировать файл в .csv формате;
- * После импорта сравнить эти данные с другим файлом формата .csv, провести аутентификацию;
- * Экспортировать данные в удобный для чтения и анализа формат таблицы.

Инструменты разработки

- * Для того, чтобы конвертировать файл формата .csv в Dataframe я решил использовать язык программирования Python, библиотеки Numpy и Pandas.
- * Для загрузки данных в базу Oracle я установил соединение с помощью клиента Oracle.

Архитектура приложения



Содержимое CSV-файлов

```
data – Блокнот
Файл  Правка  Формат  Вид  Справка
GA;GB;G1;G2;G3;G4;G5;G6;G7;G8;G9;G10
Федеральное казенное учреждение «Налог-Сервис» Федеральной налоговой службы;ФКУ ""НАЛОГ-СЕРВИС"" ФНС РОССИИ";125373, г. Москва, Походный проезд, дом 3, корп. 3;Филимошин Роман Влад
Федеральное казенное учреждение «Объединенная дирекция единого заказчика Федеральной налоговой службы»;ФКУ «ОДЕЗ ФНС РОССИИ»;127381, Москва, ул.Неглинная, 23;Мелешкин Михаил Федоро
Федеральное бюджетное лечебно-профилактическое учреждение «Санаторий «Радуга» Федеральной налоговой службы»;ФБЛПУ ""САНАТОРИЙ ""РАДУГА"" ФНС РОССИИ"";354008, Краснодарский край, г
Федеральное бюджетное лечебно-профилактическое учреждение «Санаторий «Эллада» Федеральной налоговой службы»;ФБЛПУ ""САНАТОРИЙ ""ЭЛЛАДА"" ФНС РОССИИ"";353456, Краснодарский край, г
Федеральное бюджетное лечебно-профилактическое учреждение «Лечебно-реабилитационный центр «Подмосковье» Федеральной налоговой службы»;ФБЛПУ «ЛРЦ «Подмосковье» ФНС России»;141052, М
Федеральное бюджетное лечебно-профилактическое учреждение «Санаторий-профилакторий «Золотой берег» Федеральной налоговой службы»;ФБЛПУ ""САНАТОРИЙ-ПРОФИЛАКТОРИЙ ""ЗОЛОТОЙ БЕРЕГ"" Ф
Федеральное бюджетное лечебно-профилактическое учреждение «Санаторий «Днепр» Федеральной налоговой службы»;ФБЛПУ ""САНАТОРИЙ-ПРОФИЛАКТОРИЙ ""Днепр"" ФНС РОССИИ"";298662, Республи
Федеральное бюджетное детское оздоровительное учреждение «Детский оздоровительный лагерь «Березка» Федеральной налоговой службы»;ФБДОУ ""ДОЛ ""БЕРЕЗКА"" ФНС РОССИИ"";143570, Москв
Федеральное государственное бюджетное образовательное учреждение дополнительного профессионального образования «Приволжский институт повышения квалификации Федеральной налоговой сл
Федеральное государственное бюджетное образовательное учреждение дополнительного профессионального образования «Северо-Западный институт повышения квалификации Федеральной налоговой сл
```

Данные CSV-файла 1

```
authentication – Блокнот
Файл  Правка  Формат  Вид  Справка
GA;GB;G1;G2;G3;G4;G5;G6;G7;G8;G9;G10
Федеральное казенное учреждение «Налог-Сервис» Федеральной налоговой службы;ФКУ ""НАЛОГ-СЕРВИС"" ФНС РОССИИ";125373, г. Москва, Походный проезд, дом 3, корп. 3;Филимошин Роман Влад
Федеральное казенное учреждение «Объединенная дирекция единого заказчика Федеральной налоговой службы»;ФКУ «ОДЕЗ ФНС РОССИИ»;127381, Москва, ул.Неглинная, 23;Мелешкин Михаил Федоро
Федеральное бюджетное лечебно-профилактическое учреждение «Санаторий «Днепр» Федеральной налоговой службы»;ФБЛПУ ""САНАТОРИЙ-ПРОФИЛАКТОРИЙ ""Днепр"" ФНС РОССИИ"";298662, Республи
Текст;Для;Примера;Который;Не;Должен;Попасть;В;Базу;Данных
```

Данные CSV-файла 2

Результат работы приложения

```
main.py > ...
12 print('Reading csv-file 1...')
13 df0 = pd.read_csv("data.csv",sep=';')
14 print('Reading csv-file 2...')
15 df1 = pd.read_csv("authentication.csv",sep=';')
16 print('Converting data to DataFrame...')
17 mapped=df0[df0['GA'].isin(df1['GA'])]
18 mapped = mapped.replace(np.nan, '', regex=True)
19
20 # Connect to Oracle
21
22 LOCATION_ORACLE = r"C:\Users\Влад\Downloads\instantclient_19_9"
23 os.environ["PATH"] = LOCATION_ORACLE + ";" + os.environ["PATH"]
24 conn = cx_Oracle.connect(
25     config.user,
26     config.password,
27     config.dsn,
28     encoding=config.encoding)
29 curs = conn.cursor()
30
31 # Export DataFrame to Oracle
32 print('Exporting data to Oracle DB...')
33 for item in mapped.values:
34     curs.execute(
35         "INSERT INTO DATA (GA, GB, G1, G2, G3, G4, G5, G6, G7, G8, G9, G10) VALUES (:GA, :GB, :G1, :G2, :G3, :G4, :G5, :G6, :G7, :G8, :G9, :G10)",
36         GA=item[0], GB=item[1], G1=item[2], G2=item[3], G3=item[4], G4=item[5], G5=item[6], G6=item[7], G7=item[8], G8=item[9], G9=item[10], G10=item[11]
37     )
38
39 conn.commit()
40 print('Export completed successfully!!!')
41 conn.close()
42
```

ПРОБЛЕМЫ Выходные данные КОНСОЛЬ ОТЛАДКИ ТЕРМИНАЛ

2: Python Debug Consc

```
PS C:\Users\Влад\Desktop\TEST> cd 'c:\Users\Влад\Desktop\TEST'; & 'C:\Users\Влад\AppData\Local\Programs\Python\Python39\python.exe' 'c:\Users\Влад\.vscode\extensions\ms-python.python-2020.11.371
526539\pythonFiles\lib\python\debugpy\launcher' '52136' '--' 'c:\Users\Влад\Desktop\TEST\main.py'
Reading csv-file 1...
Reading csv-file 2...
Converting data to DataFrame...
Exporting data to Oracle DB...
Export completed successfully!!!
PS C:\Users\Влад\Desktop\TEST>
```

Результат работы приложения

ADMIN

[Worksheet]*

Consumer Group: LOW

1 Select *from data

Query Result Script Output DBMS Output Explain Plan Autotrace SQL History Data Loading

Download Execution time: 0.005 seconds

	ga	gb	g1	g2	g3	g4	g5	g6
1	Федеральное казе...	ФКУ "НАЛОГ-СЕРВ...	125373, г. Москва, ...	Филимошин Рома...	7 (495) 913-09-65	Обработка данных	55681777	1024000
2	Федеральное казе...	ФКУ «ОДЕЗ ФНС Р...	127381, Москва, ул...	Мелешкин Михаил...	7 (495) 606-71-94	Управление недви...	54939334	1037700
3	Федеральное бюд...	ФБЛПУ "САНАТОР...	298662, Республик...	Грищенко Юрий В...	7 10 (38065) 424-72...	Деятельность сана...	53528847	1026101

Перспективы развития проекта

- * Добавить возможность работы с несколькими типами данных;
- * Добавить возможность экспорта в различные типы данных;
- * Сделать пользовательский интерфейс для работы с приложением.

СПАСИБО ЗА ВНИМАНИЕ!