# 2. Маніпуляції SQL базою за допомогою Python 2

Для формування аналітичної інформації щодо орієнтовної вартості хімічної продукції в динаміці (за кілька років) в доларах США був використаний індійський сайт zauba.com. Але для того, щоб отримати вартість товару в дол.США, а не в індійських рупіях (курс рупій до долару не фіксований і спостерігаються значні коливання) був написаний код мовою програмування Python. Була створена таблиця за допомогою простого Python редактора IDLE, який оперую базою даних PostgreSQL, в яку імпортуються дані з двох файлів (.txt). В одному історичний курс валют USD/INR, в іншому цінові дані по хімічним товарам з сайту zauba.com. Після того, як був здійснений імпорт файлів, запускається SQL запит з необхідним переліком стовпчиків, що дають змогу отримати вартість товару в дол.США за окрему одиницю, а також додаткові: опис товару, кількість. PostgreSQL допомагає зв’язати таблиці за однаковим параметром, а саме – дата.

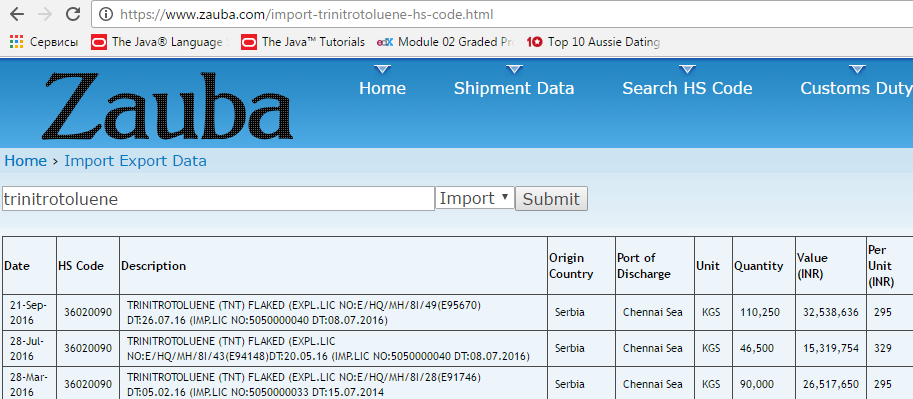
Схема роботи виглядає наступним чином:



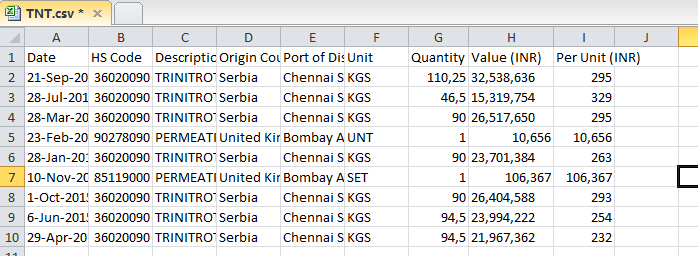
Зеленим кольором відображено роботу, яка виконається безпосередньо на Python.

### 2.1. Отримання безкоштовної інформації, щодо вартості будь-яких товарів

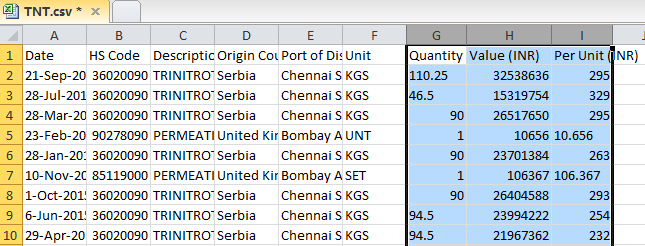
В нашому випадку аналізуємо хімічний продукт – Тринітротолуол чи ТНТ



Копіюємо дані за сайту zauba.com в файл: ТNT**.csv**

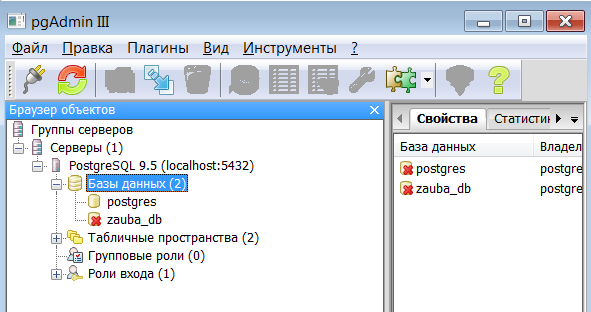


Перероблюємо файл, тобто видаляємо зайві «,»:



### 2.2. Імпорт цінової інформації в PostgreSQL через Python

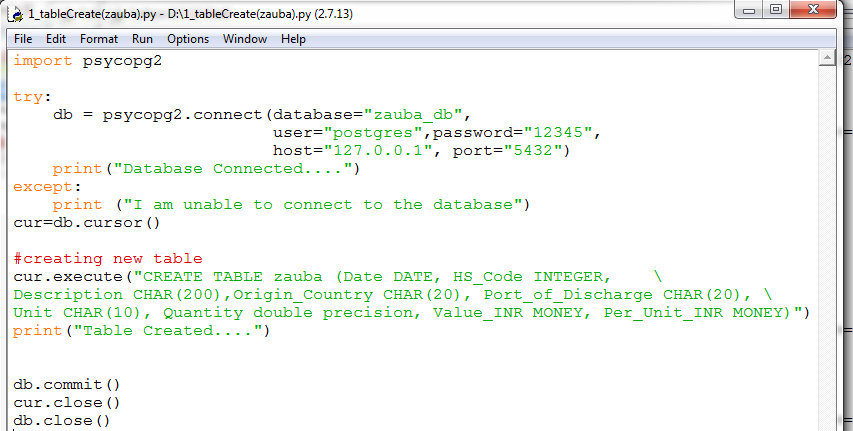
Відкриваємо базу PostgreSQL та створюємо нову базу даних “zauba\_db”.



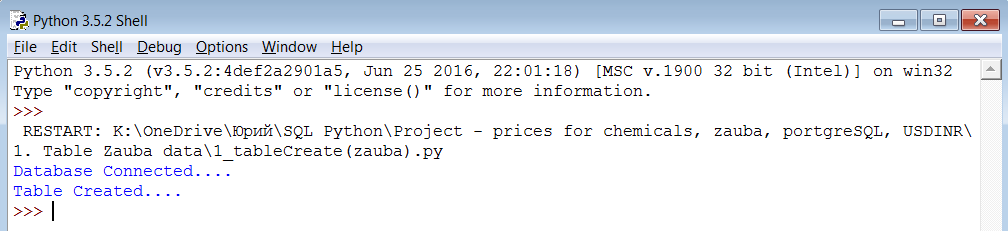
Створення в новій базі zauba\_db нової таблиці zauba з наступними стобчиками:

Date, HS\_Code, Description, Origin\_Country, Port\_of\_Discharge, Unit, Quantity, Value\_INR, Per\_Unit\_INR. Тобто повний перелік стобчиків, як і на сайті zauba.com.

**Код:**

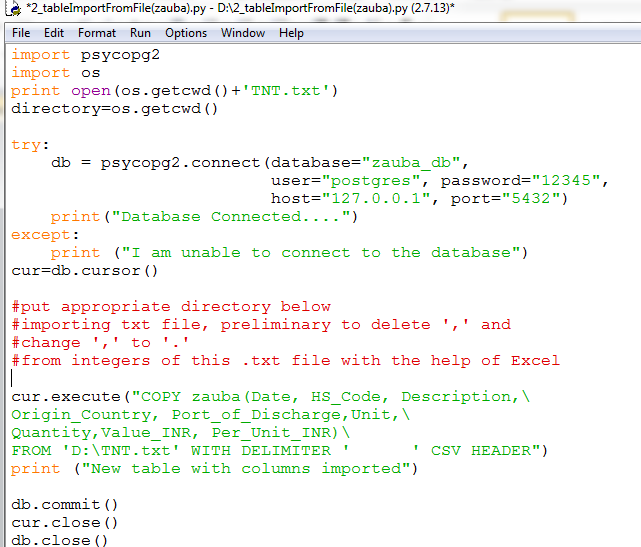
****

**Результат:**

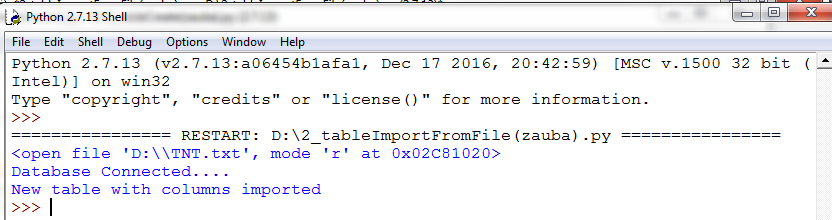


Імпортуємо файл TNT.csv до PostgreSQL за допомогою Python2

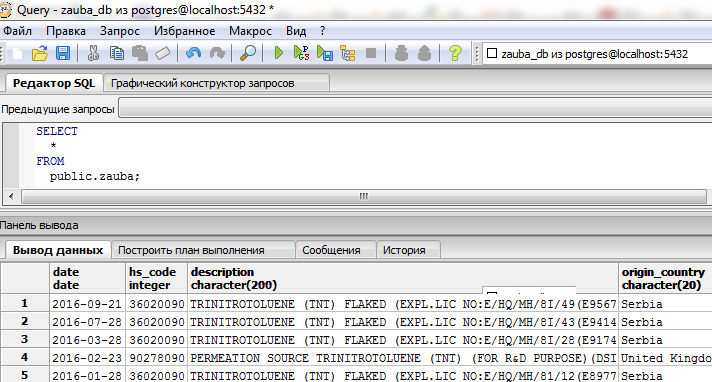
**Код:**



**Результат:**



Перевіряємо, чи дійсно дані були імпортовані

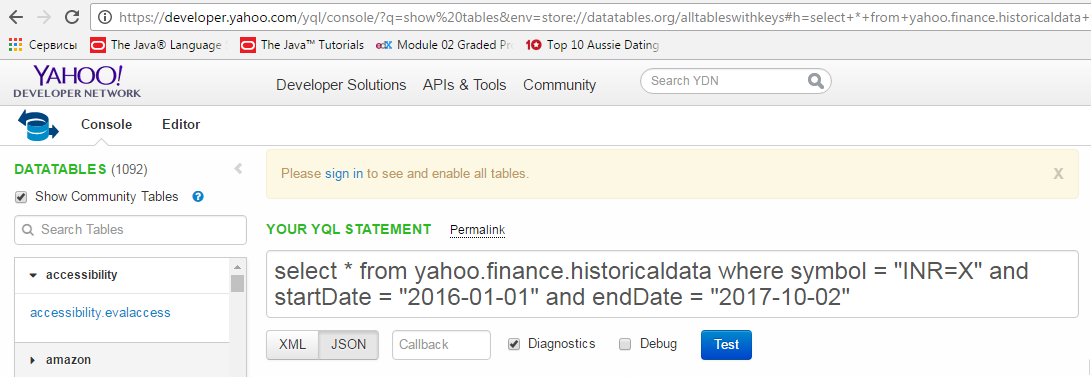


Як бачимо вище SQL запит: SELECT \* FROM public.zauba відобразив інформацію ,що була імпортована за файлу «TNT.csv»

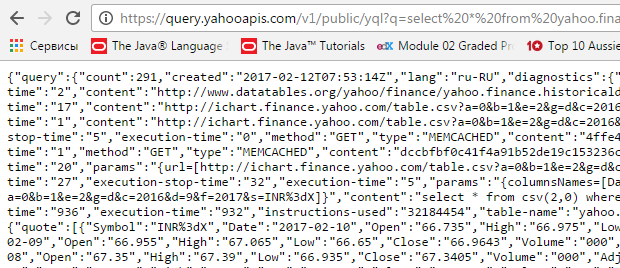
### 2.3. Отримання історичних курсів валют за допомогою Yahoo

На сайті developer.yahoo.com можна отримати історичні курси валют та завантажити їх в форматі «XML» чи «JSON».

Завантажимо курс INR до USD за період з 2016-01-01 по 2017-10-02



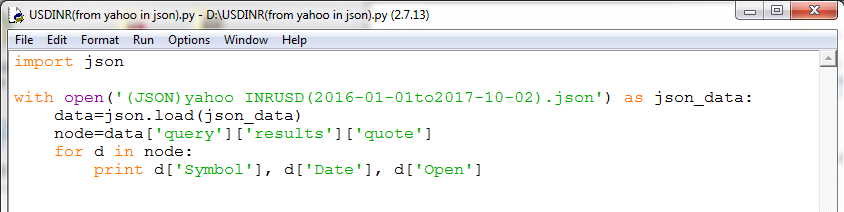
Отримаємо результати



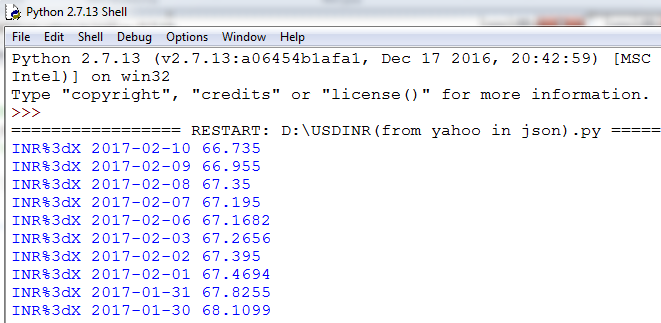
Та збережемо їх у файл: (JSON)yahoo INRUSD(2016-01-01to2017-10-02)**.json**

Відкриємо даний файл у Python, щоб отримати тільки потрібні дані:

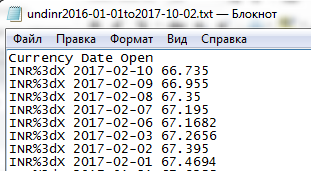
**Код:**



**Результат:**



Щоб не перевантажувати Python код даними, щодо автоматичного запису отриманих результатів в файл, просто скопіюємо отримані результати в текстовий файл: undinr2016-01-01to2017-10-02.**txt**

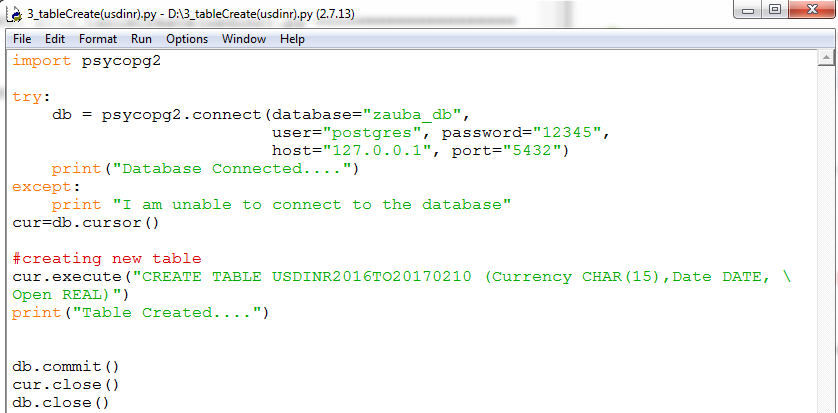


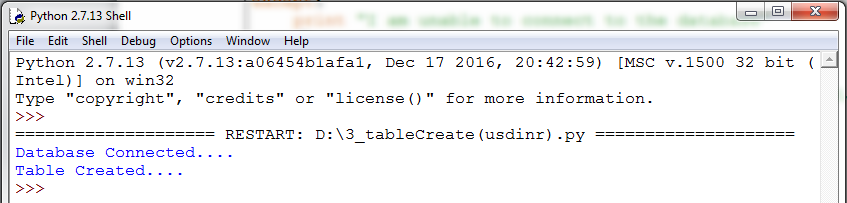
### 2.4. Імпорт курсів валют в PostgreSQL через Python

Створення в новій базі zauba\_db нової таблиці usdinr2016to20170210 з наступними стобчиками:

Currency, Date, Open

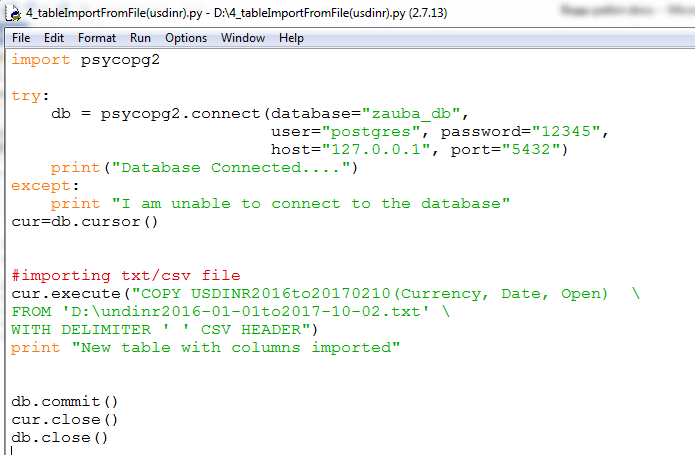
**Код:**

  
**Результат:**

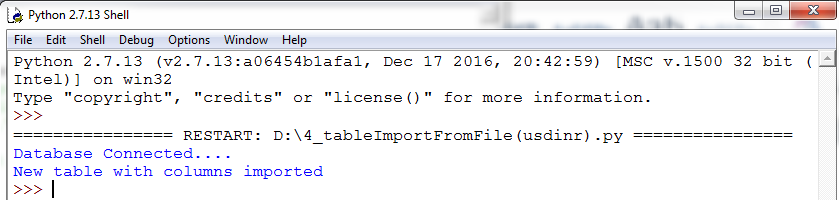


Імпортуємо файл usdinr2016to20170210.txt до PostgreSQL за допомогою Python2

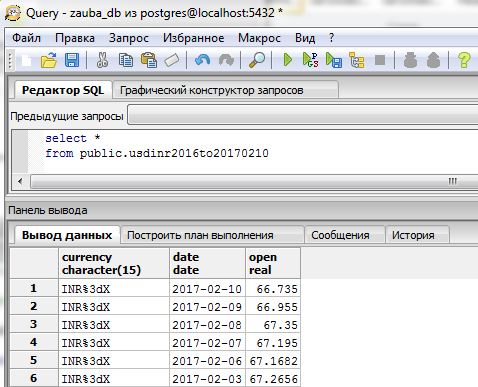
**Код:**



**Результат:**



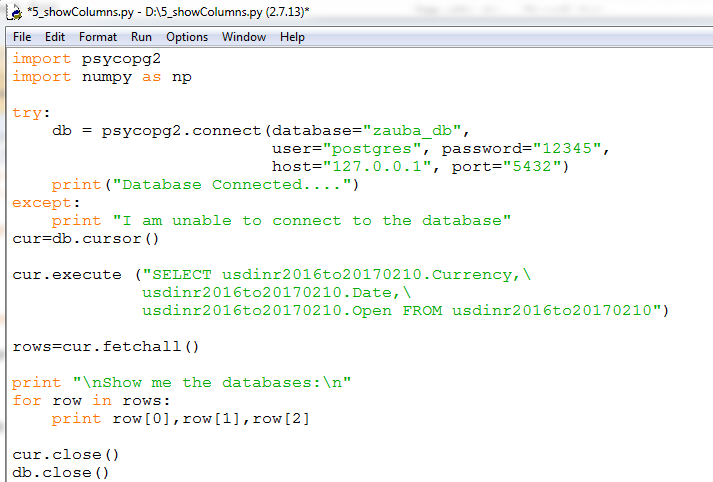
Перевіряємо, чи дійсно дані були імпортовані, безпосередньо відкриваючи PostgreSQL



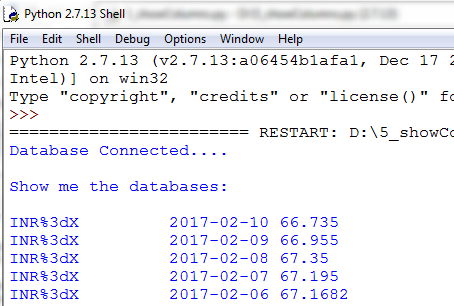
Дані були імпортовані нормально

Також можна теж саме перевірити за допомогою Python:

**Код:**

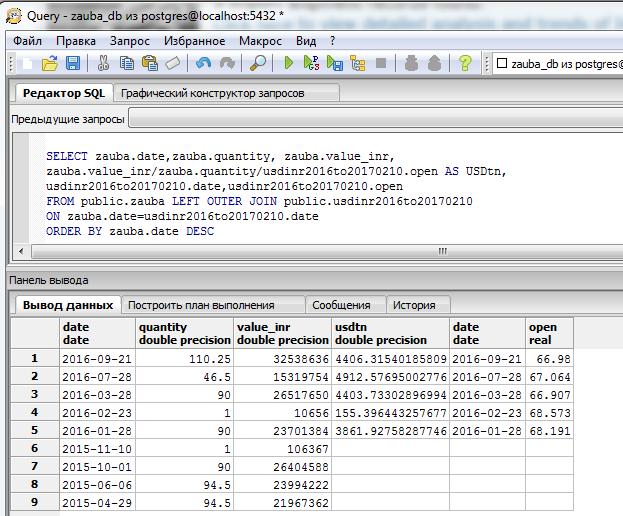


**Результат:**

****

### 2.5. Отримання сводної ціни в дол.США з імпортованих до PostgreSQL двох таблиць

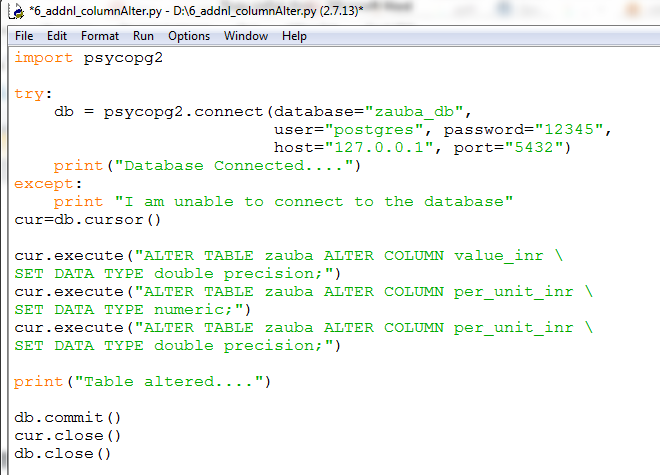
Коли всі дані імпортовані в PostgreSQL, можна отримати вартість одиниці товару в USD за тонну. Отримані нижче пусті рядкі пов’язані з тим, що ми завантажили курс USD/INR починаючи з 2016 р., але дані по імпорту маємо і в 2015 р.

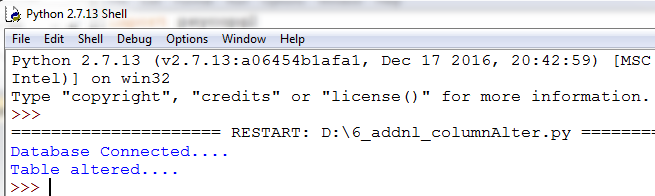


Теж сам можна виконати за допомогою Python:

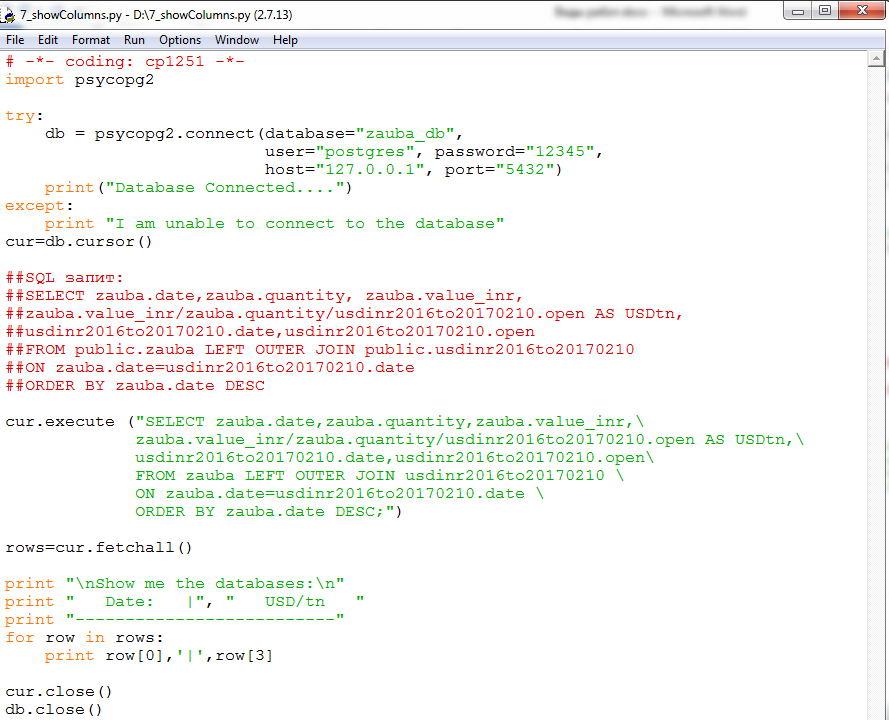
Але з по-перше треба змінити тип даних «MONEY» на «double precision»:

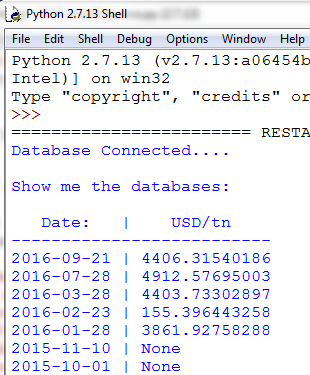
**Код:**

  
**Результат:**



**Код:**

  
**Результат:**



Але більш зручно SQL запити роботи в програмі PostgreSQL, а ніж через Python.

# Висновки:

За допомою мови програмування Python був здійснений парсинг інтернет сторінок, що дозволило автоматизувати процес підготовки до виставки в Дубаї.

В дослідженні були задіяні сучасні та безкоштовні інформаційні джерела по визначенню вартості будь-яких товарів (в нашому випадку хімічної продукції), включаючи період в кілька років, що є важливим та корисним для господарської діяльності підприємства. Тобто на прикладі знайдена орієнтовна вартість товарів, які проходять через територію Індії (експорт чи імпорт). Хоч ця ціна буде відрізнятися наприклад від європейської ціни (можливо до 15-20%) але буде спостерігатися історичний тренд, що добре корелюється з європейськими цінами.

В якості аналітичної достовірної інформації був використаний сайт <https://developer.yahoo.com/>, що дозволяє завантажувати дані в форматах (xml та json) спеціально для розробників програмного забезпечення. В нашому випадку були завантажені та оброблені історичні курси валют.

Код на Python допоміг завантажити аналітичні дані в SQL базу даних та здійснити розрахунок для отримання вартості тринітротолуолу в дол.США. А саме загальна вартість в індійських рупіях була розділа на кількість та розділена на курс USD/INR з іншої таблиці.

# Програмне забезпечення:

**Основні робочі інcтрументи:**

- Python 3.5 та Python 2 та IDLE (для виконання коду)

- модуль BeautifulSoup до Python 3.5 (для парсингу інтернет-сторінок)

- модуль psycopg2 до Python 2 для завантаження даних з txt та csv файлів до PostgreSQL

- модуль numpy до Python

**Додаткові:**

- Notepad++ (для редагування коду)