

## Data Engineer code exercise

The goal of this code challenge is to give an idea on how you approach a typical data engineering task, organize code, communicate and solve problems.

### \*\* Requirements:

Write a script that would import itunes subscriptions into a database in order to allow data scientists to analyze the data. Data.json includes the raw data from which you can derive/calculate fields. Data.js has a **list of itunes subscriptions**, and each subscription has a **list of transactions**. Hint: load data.json in <http://jsonviewer.stack.hu/> for better readability

- use python, and handle exceptions if a file cannot be read.
- Use postgres database for storing data.

### Part 1:

- Create a postgres **database called subscription**
- Add **itunes\_subscription** table with the following fields:

Field	Description
Id	Unique identifier for a subscription, shouldn't change for the lifetime of subscription
<b>transactions</b>	<b>Usually json</b>
trial_start_date	optional
Subscription_start_date	optional
Expiration_date	required
current_status	Free text

### Part 2:

Write a python script that reads data from data.json and inserts it in the database.

- Store the list of transactions for a subscription in the **transactions** field
- Calculate and store **trial\_start\_date**. Rule: trial\_start\_date is equal to purchase\_date where is\_trial\_period = true
- Calculate and store **subscription\_start\_date**. Rule: subscription\_start\_date is equal to purchase\_date for the transaction that immediately follows a trial transaction

- Calculate and store **expiration\_date**. Rule: expiration\_date is equal to expires\_date of the last transaction
- Calculate and store **current\_status**. Rules for calculating current\_status
  - If the subscription only has a trial transaction and the expires\_date is in the future, then the value is "Active Trial"
  - If the subscription only has a trial transaction and the expires\_date is in the past, then the value is "Expired Trial"
  - If the subscription has non-trial (paid) transactions and the expires\_date of the last transaction is in the future, then the value is "Active Subscription"
  - If the subscription has non-trial (paid) transactions and the expires\_date of the last transaction is in the past, then the value is "Expired Subscription"

**Part 3:**

Write an sql query that lists the number of Active Itunes Subscriptions, Expired Itunes Subscriptions, Active trials and expired trials.