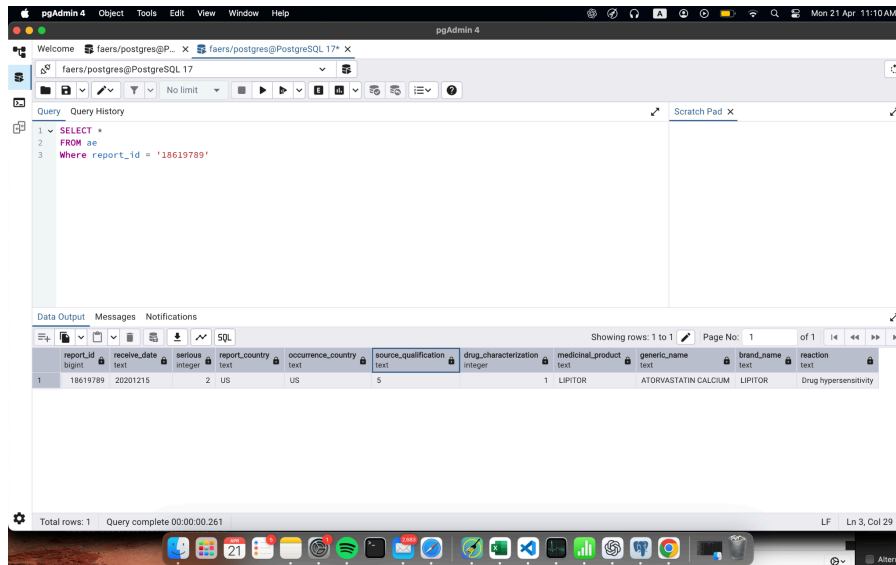


1. Test Case: API Query Verification

Select a random report ID, search for it using the API query, and verify that the API result matches the downloaded data.

SQL Query:



The screenshot shows the pgAdmin 4 interface. The query editor contains the following SQL query:

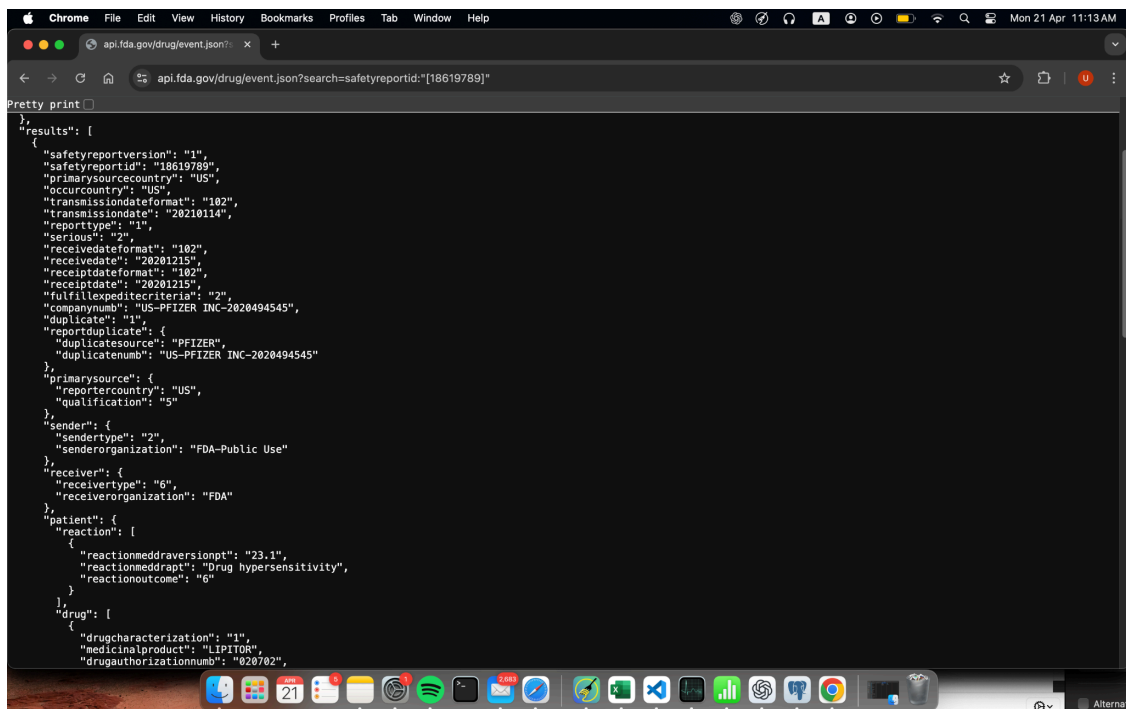
```
1 SELECT
2 FROM ae
3 Where report_id = '18619789'
```

The results pane displays a single row of data with the following columns and values:

report_id	report_date	serious	report_country	occurrence_country	source_qualification	drug_characterization	medicinal_product	generic_name	brand_name	reaction
18619789	20201215	2	US	US	5	1	LIPITOR	ATORVASTATIN CALCIUM	LIPITOR	Drug hypersensitivity

API query:

- Verifying Report ID, serious, received date, report country, occurrence country, drug characterization, medicinal product:



The screenshot shows a Chrome browser window with the URL `api.fda.gov/drug/event.json?search=safetyreportid:"[18619789]"`. The response is displayed in a dark-themed JSON viewer, showing the following structure:

```
{
  "results": [
    {
      "safetyreportversion": "1",
      "safetyreportid": "18619789",
      "primarysourcecountry": "US",
      "occurcountry": "US",
      "transmissiondateformat": "102",
      "transmissiondate": "20210114",
      "reporttype": "1",
      "serious": "2",
      "receiveddateformat": "102",
      "receiveddate": "20201215",
      "receiptdateformat": "102",
      "receiptdate": "20201215",
      "fulfillspeditercriteria": "2",
      "companynumb": "US-PFIZER INC-2020494545",
      "duplicate": "1",
      "reportduplicate": {
        "duplicatesource": "PFIZER",
        "duplicatenumb": "US-PFIZER INC-2020494545"
      },
      "primarysource": {
        "reportercountry": "US",
        "qualification": "5"
      },
      "sender": {
        "sendertype": "2",
        "senderorganization": "FDA-Public Use"
      },
      "receiver": {
        "receiververtype": "6",
        "receiverorganization": "FDA"
      },
      "patient": {
        "reaction": [
          {
            "reactionmeddraversion": "23.1",
            "reactionmeddrapt": "Drug hypersensitivity",
            "reactionoutcome": "6"
          }
        ]
      },
      "drug": [
        {
          "drugcharacterization": "1",
          "medicinalproduct": "LIPITOR",
          "drugauthorizationnumb": "020702"
        }
      ]
    }
  ]
}
```

b. Verifying Reaction, generic_name

```
{
  "sender": {
    "sendertype": "2",
    "senderorganization": "FDA-Public Use"
  },
  "receiver": {
    "receivertype": "6",
    "receiverorganization": "FDA"
  },
  "patient": {
    "reaction": [
      {
        "reactionmodrversionpt": "23.1",
        "reactionmoddrapt": "Drug hypersensitivity",
        "reactionoutcome": "6"
      }
    ]
  },
  "drug": [
    {
      "drugcharacterization": "1",
      "medicinalproduct": "LIPITOR",
      "drugauthorizationnumb": "020702",
      "drugdosagetext": "UNK",
      "drugdosageform": "FILM-COATED TABLET",
      "actiondrug": "5",
      "drugadditional": "3",
      "activesubstance": {
        "activesubstancename": "ATORVASTATIN CALCIUM"
      },
      "openfda": {
        "application_number": [
          "NDA020702"
        ],
        "brand_name": [
          "LIPITOR"
        ],
        "generic_name": [
          "ATORVASTATIN CALCIUM"
        ],
        "manufacturer_name": {
          "Parke-Davis Div of Pfizer Inc",
          "Viatris Specialty LLC"
        },
        "product_ndc": [
          "0071-0155",
          "0071-0156",
          "0071-0157"
        ]
      }
    }
  ]
}
```

2. Test case: Verifying row-count of the downloaded data

Total no. of rows: 1790374

Query Query History

```
1 select count(*) from ae
```

Data Output Messages Notifications

	count bigint
1	1790374

Unique record count: 1536804

Query

Query History

1

2

SELECT COUNT(DISTINCT ROW(report_id, receive_date, serious, report_country, occurrence_country, source_qualification, drug_characterization, medicinal_product, generic_name, brand_name, reaction)) FROM ae;

Data Output

Messages

Notifications

Showing

count
bigint

1536804

3. Test case: Cleaning the downloaded data.

- Merging the NULL values: 'source_qualification' column has both NULL and 'NULL' (as a string) values.

Query

Query History

1

2

3

4

5

UPDATE ae
SET source_qualification = NULL
WHERE source_qualification = 'NULL';

Data Output

Messages

Notifications

UPDATE 97742

Query returned successfully in 4 secs 838 msec.

b. Clean and download the data.

Data cleaning summary:

- i. Removed duplicate records
- ii. Filtered the dataset to include only records where receive_date falls within the year 2020
- iii. Retained only those records where generic_name contains **Atorvastatin**, **Rosuvastatin**, or **Simvastatin**.
- iv. Excluded records where drug_characterization is **NULL**, ensuring each drug entry has a defined role in the report.

```
with clean as (  
  select distinct report_id, receive_date, serious, report_country, occurrence_country, source_qualification,  
                 drug_characterization, medicinal_product, generic_name, brand_name, reaction  
  from ae  
  where  
    (receive_date >= '20200101' and receive_date < '20210101') and  
    (generic_name like '%SIMVASTATIN%' or  
     generic_name like '%ATORVASTATIN%' or  
     generic_name like '%ROSUVASTATIN%') and  
    (drug_characterization is not NULL)  
    order by report_id  
)  
select * from clean
```

Output Messages Notifications

report_id	receive_date	serious	report_country	occurrence_country	source_qualification	drug_characterization	medicinal_product
bigint	text	integer	text	text	text	integer	text
17221363	20200101	1	US	US	[null]	2	CRESTOR
17221363	20200101	1	US	US	[null]	2	CRESTOR
17221363	20200101	1	US	US	[null]	2	SIMVASTATIN.
17221363	20200101	1	US	US	[null]	2	SIMVASTATIN.
17221371	20200101	1	US	US	5	2	ROSUVASTATIN CALCIUM.
17221371	20200101	1	US	US	5	2	ROSUVASTATIN CALCIUM.
17221377	20200101	1	US	US	5	2	LIPITOR

4. Test case: Connecting data to Tableau

Importing the cleaned data (fact table) and connecting it with the dimension tables

Connections

cleaned_statin_ae

Text file

Add

Files

☐ Use Data Interpreter

Data Interpreter might be able to clean your Text file workbook.

ae_cleaned.csv

cleaned_statin_ae.csv

drug_charact...on_codes.csv

serious_codes.csv

source_quali...on_codes.csv

New Union

New Table Extension

ae_cleaned.csv

13 fields 102356 rows

Name

ae_cleaned.csv

Fields

Type	Field Name	Phys...	Rem...
#	Report Id	ae_cle...	report...
📅	Receive Date	ae_cle...	receiv...

ae_cleaned.csv

ae_cleaned.csv

ae_cleaned.csv

ae_cleaned.csv

Report Id	Receive Date	Serious	Report Country
17221363	01/01/2020	1	US
17221363	01/01/2020	1	US
17221363	01/01/2020	1	US
17221363	01/01/2020	1	US
17221371	01/01/2020	1	US
17221371	01/01/2020	1	US