# Example FOUR - Enrich our data using the command line

## Introduction

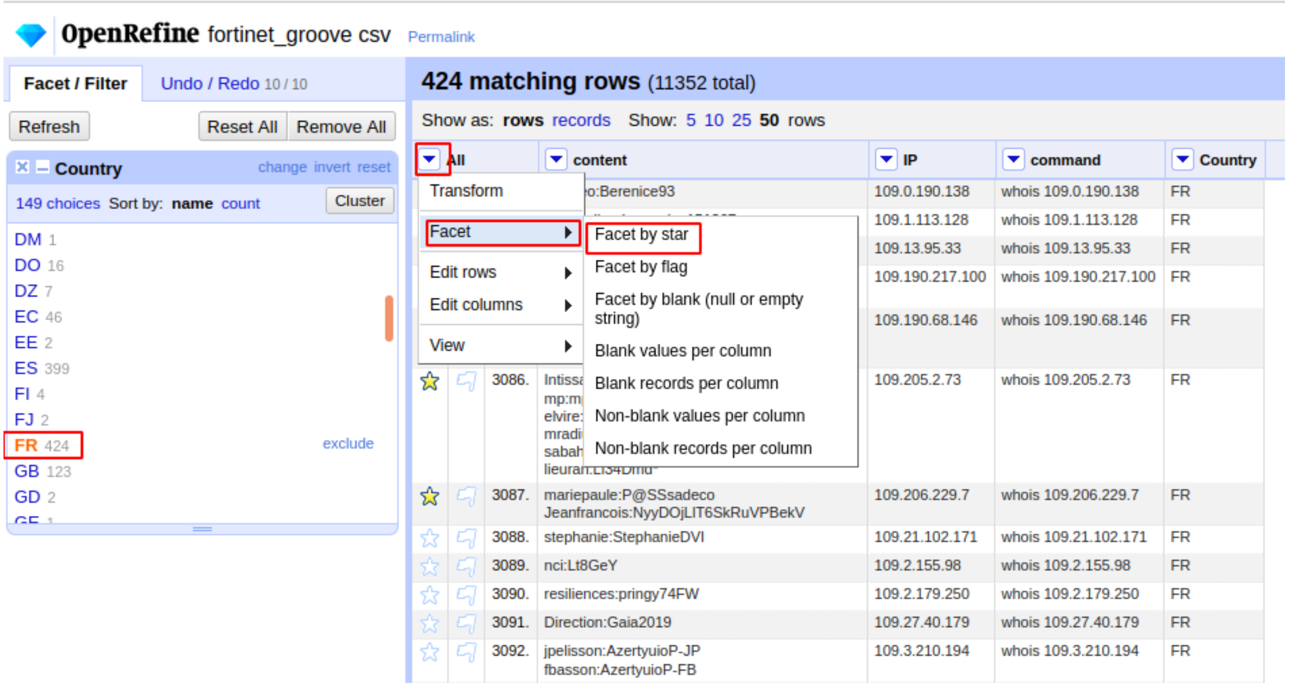
On this example, we will use a dataset proceeding from the Groove ransomware (A Hoax, by the way...) : a list of vulnerable Fortinet IPs.

When you receive such an information, a good move is to quickly qualify such a dataset, to identify key IPs (by country, entities, etc...).

Our suggestion is to automate a *whois* on these IP, and to parse the result and export the scenario as a **recipe**, to replay it on the same dataset.

## Import the data

- create a new project in OpenRefine by importing the csv file. Beware to select the commas separator. The project should content 3 columns and 11352 lines.
- create a new column "command" based on IP, using the formula `"whois "+value`.
- filter the country, by doing a text facet and click on the country of your choice (FR for example).
- randomly select 10 IPs in the dataset for the demo, by starring them.
- in the column "All", select **facet by star** then **true** to only display these IPs.

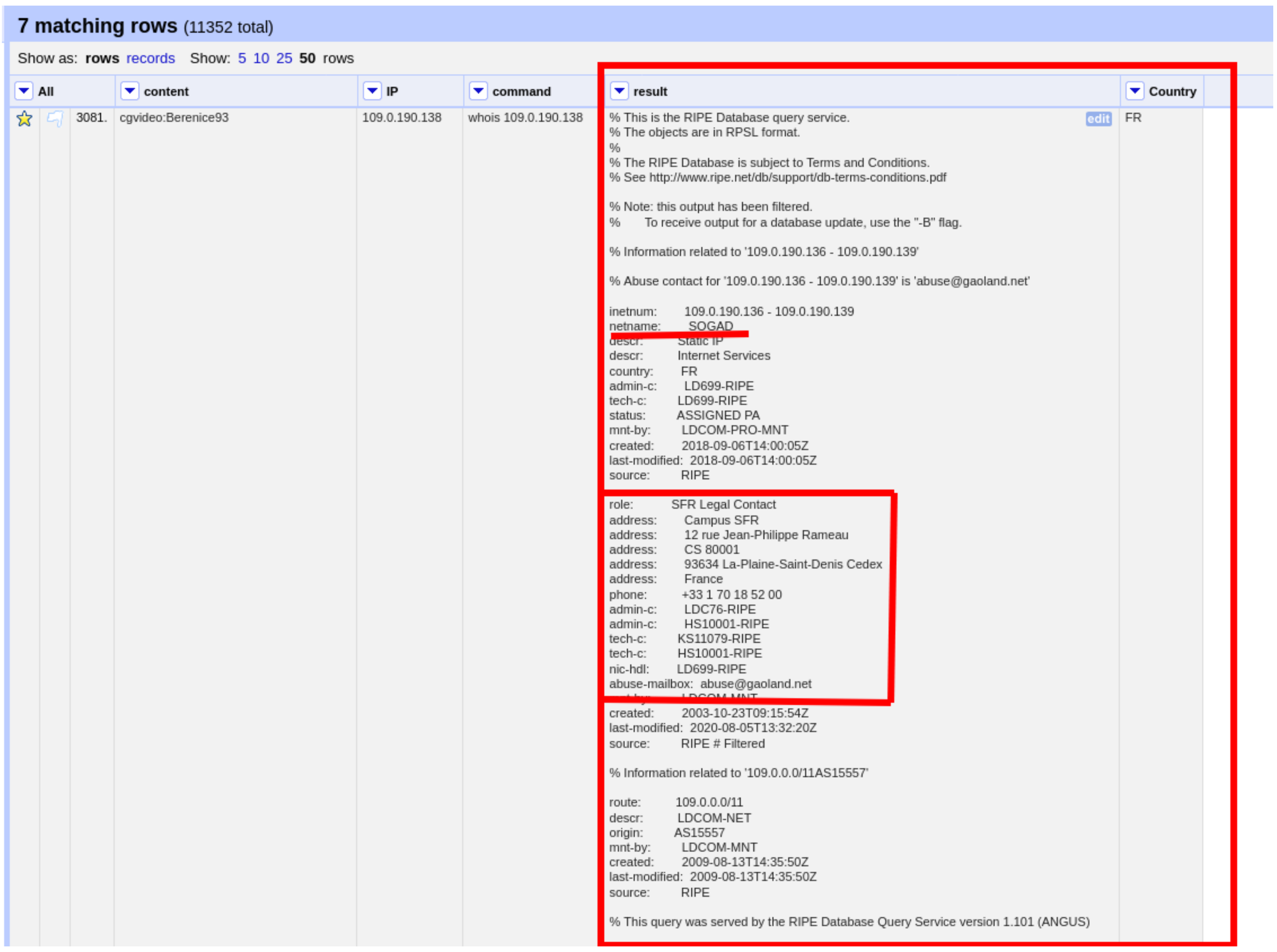


**424 matching rows** (11352 total)

Show as: **rows** records Show: 5 10 25 **50** rows

| All | | | content | IP | command | Country |
|---|---|---|---|---|---|---|
| | | | o:Berenice93 | 109.0.190.138 | whois 109.0.190.138 | FR |
| | | | | 109.1.113.128 | whois 109.1.113.128 | FR |
| | | | | 109.13.95.33 | whois 109.13.95.33 | FR |
| | | | | 109.190.217.100 | whois 109.190.217.100 | FR |
| | | | | 109.190.68.146 | whois 109.190.68.146 | FR |
| ★ | | 3086. | Intissa mp:m elvire: mradi sabah lieuran:Li34Dmd | 109.205.2.73 | whois 109.205.2.73 | FR |
| ★ | | 3087. | mariepaule:P@SSsadeco Jeanfrancois:NyyDOjLIT6SkRuVPBekV | 109.206.229.7 | whois 109.206.229.7 | FR |
| ☆ | | 3088. | stephanie:StephanieDVI | 109.21.102.171 | whois 109.21.102.171 | FR |
| ☆ | | 3089. | nci:Lt8GeY | 109.2.155.98 | whois 109.2.155.98 | FR |
| ☆ | | 3090. | resiliences:pringy74FW | 109.2.179.250 | whois 109.2.179.250 | FR |
| ☆ | | 3091. | Direction:Gaia2019 | 109.27.40.179 | whois 109.27.40.179 | FR |
| ☆ | | 3092. | jpelisson:AzertyuioP-JP fbasson:AzertyuioP-FB | 109.3.210.194 | whois 109.3.210.194 | FR |

- execute the jython script to retrieve the infos...



**7 matching rows** (11352 total)

Show as: **rows** records Show: 5 10 25 **50** rows

| All | | | content | IP | command | result | Country |
|---|---|---|---|---|---|---|---|
| ☆ | | 3081. | cgvideo:Berenice93 | 109.0.190.138 | whois 109.0.190.138 | see below | FR |

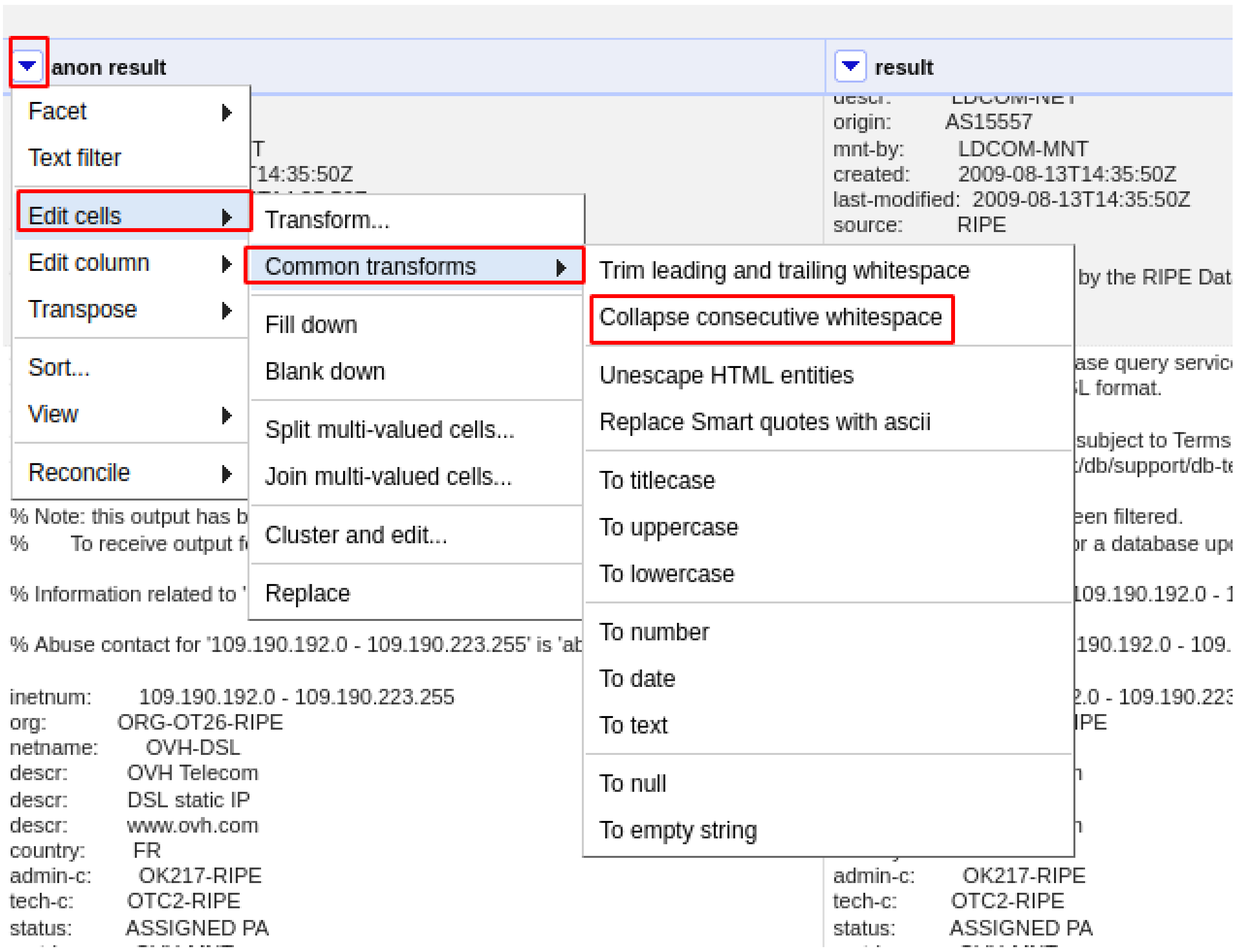
```
% This is the RIPE Database query service.
% The objects are in RPSL format.
%
% The RIPE Database is subject to Terms and Conditions.
% See http://www.ripe.net/db/support/db-terms-conditions.pdf

% Note: this output has been filtered.
%       To receive output for a database update, use the "-B" flag.

% Information related to '109.0.190.136 - 109.0.190.139'

% Abuse contact for '109.0.190.136 - 109.0.190.139' is 'abuse@gaoland.net'

inetnum:        109.0.190.136 - 109.0.190.139
netname:        SOGAD
descr:          Static IP
descr:          Internet Services
country:        FR
admin-c:        LD699-RIPE
tech-c:         LD699-RIPE
status:         ASSIGNED PA
mnt-by:         LDCOM-PRO-MNT
created:        2018-09-06T14:00:05Z
last-modified:  2018-09-06T14:00:05Z
source:         RIPE

role:           SFR Legal Contact
address:        Campus SFR
address:        12 rue Jean-Philippe Rameau
address:        CS 80001
address:        93634 La-Plaine-Saint-Denis Cedex
address:        France
phone:          +33 1 70 18 52 00
admin-c:        LDC76-RIPE
admin-c:        HS10001-RIPE
tech-c:         KS11079-RIPE
tech-c:         HS10001-RIPE
nic-hdl:        LD699-RIPE
abuse-mailbox:  abuse@gaoland.net
mnt-by:         LDCOM-MNT
created:        2003-10-23T09:15:54Z
last-modified:  2020-08-05T13:32:20Z
source:         RIPE # Filtered

% Information related to '109.0.0.0/11AS15557'

route:          109.0.0.0/11
descr:          LDCOM-NET
origin:         AS15557
mnt-by:         LDCOM-MNT
created:        2009-08-13T14:35:50Z
last-modified:  2009-08-13T14:35:50Z
source:         RIPE

% This query was served by the RIPE Database Query Service version 1.101 (ANGUS)
```

**Important**_ You can add a layer of anonymization, by using the *torify* command, which will wrap the command into TOR. Instead of `"curl "+value`, simply use `"torify curl "+value`.

## Parse the data

A good insight is the *netname:* value that can lead you to the company that owns the IP. We will use this expression : `value.partition("netname:")[2].partition("descr")[0]`

But first, let's remove the useless whitespaces in the column.



```
origin:        AS15557
mnt-by:        LDCOM-MNT
created:       2009-08-13T14:35:50Z
last-modified: 2009-08-13T14:35:50Z
source:        RIPE
```

```
% Abuse contact for '109.190.192.0 - 109.190.223.255' is 'ab

inetnum:       109.190.192.0 - 109.190.223.255
org:           ORG-OT26-RIPE
netname:       OVH-DSL
descr:         OVH Telecom
descr:         DSL static IP
descr:         www.ovh.com
country:       FR
admin-c:       OK217-RIPE
tech-c:        OTC2-RIPE
status:        ASSIGNED PA
```

Now create a new column called "netname", based on the column, using the above expression :

```
value.partition("netname:")[2].partition("descr")[0]
```

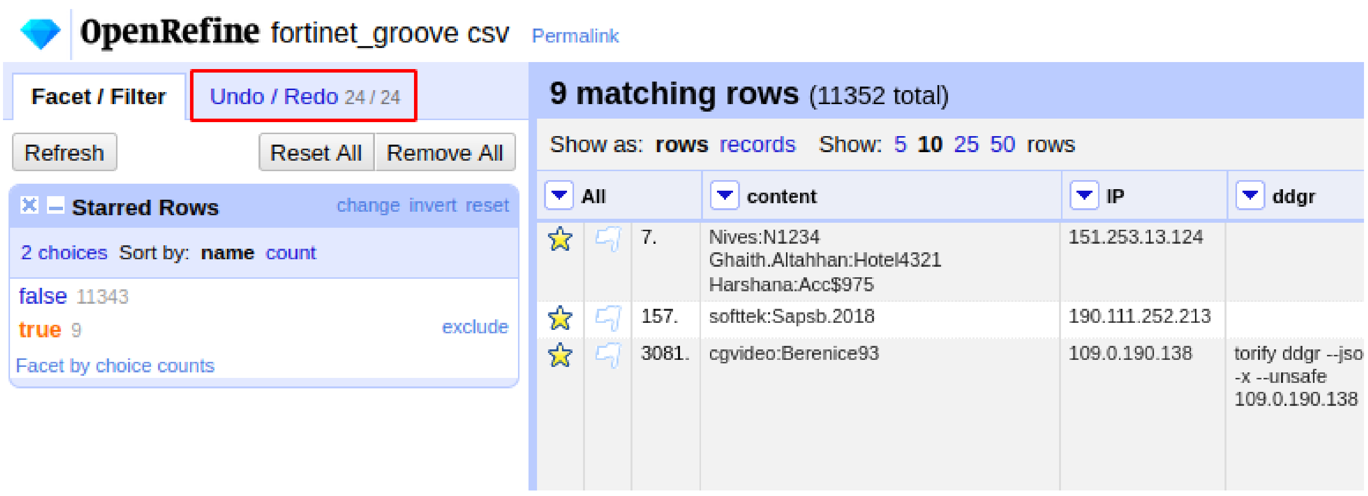
## Check the reputation of an IP on DuckDuckGo

We can check the reputation of these IP using DuckDuckGo, and ddgr. The results are often less precise but ddgr can be torified!

We can create a new column based on the IP column using a request like this : `"torify ddgr --n 20 --json -x --unsafe "+value`

- *-x* means show complete url in result
- *--unsafe* means disable safe search
- *--json* means as json format

## Export your actions as a script



OpenRefine fortinet_groove csv Permalink

Facet / Filter | Undo / Redo 24 / 24

Refresh | Reset All | Remove All

Starred Rows change invert reset

2 choices Sort by: name count

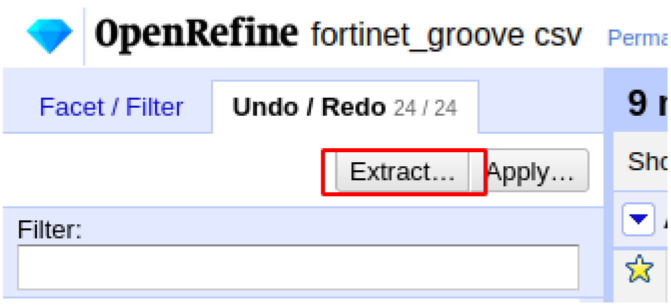
false 11343
true 9 exclude

Facet by choice counts

**9 matching rows** (11352 total)

Show as: rows records Show: 5 10 25 50 rows

| All | | | content | IP | ddgr |
|---|---|---|---|---|---|
| ☆ | | 7. | Nives:N1234 Ghaith.Altahhan:Hotel4321 Harshana:Acc$975 | 151.253.13.124 | |
| ☆ | | 157. | softtek:Sapsb.2018 | 190.111.252.213 | |
| ☆ | | 3081. | cgvideo:Berenice93 | 109.0.190.138 | torify ddgr --jso -x --unsafe 109.0.190.138 |

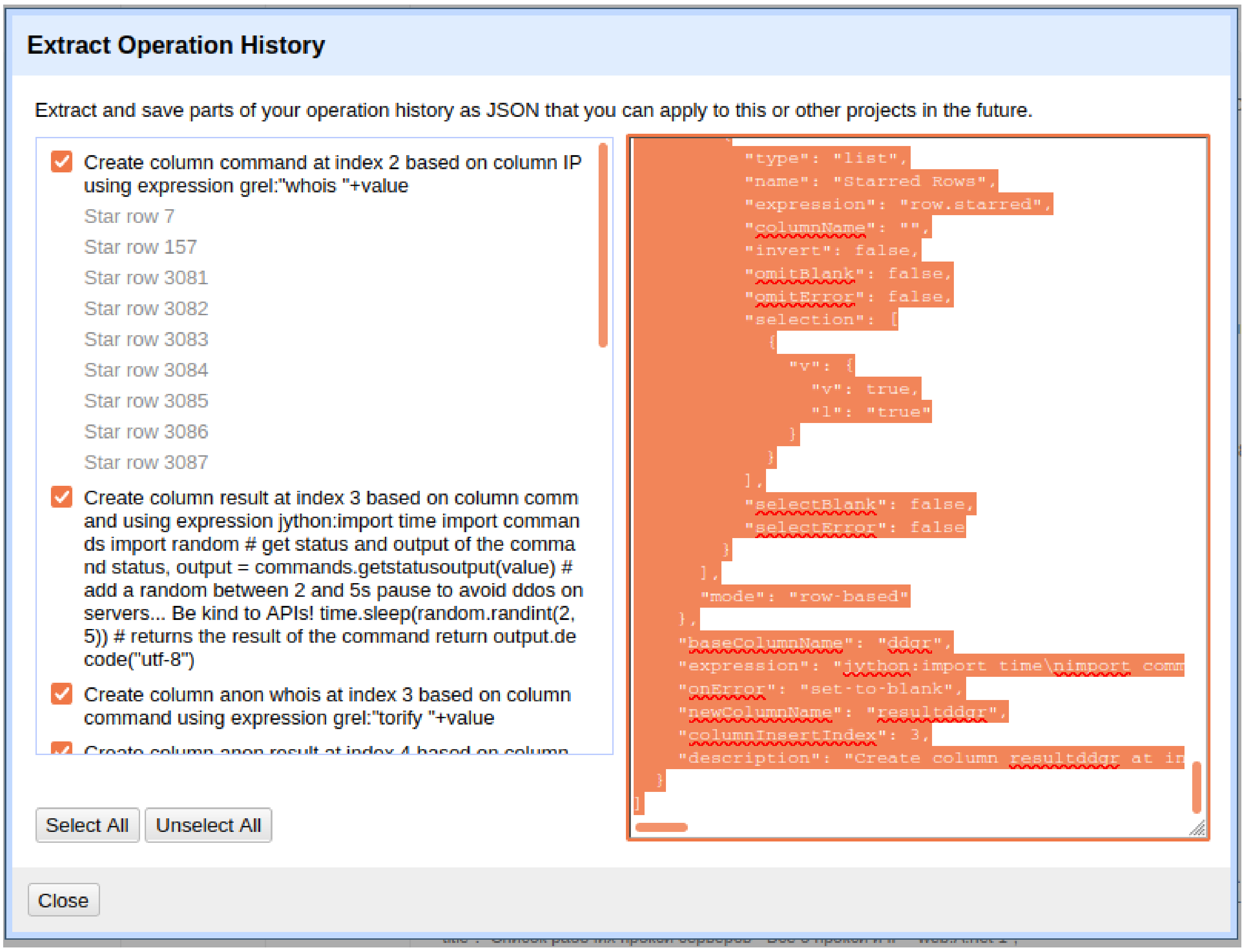


OpenRefine fortinet_groove csv Perma

Facet / Filter | Undo / Redo 24 / 24

Extract... | Apply...

Filter:



**Extract Operation History**

Extract and save parts of your operation history as JSON that you can apply to this or other projects in the future.

- ✓ Create column command at index 2 based on column IP using expression grel:"whois "+value
- Star row 7
- Star row 157
- Star row 3081
- Star row 3082
- Star row 3083
- Star row 3084
- Star row 3085
- Star row 3086
- Star row 3087
- ✓ Create column result at index 3 based on column comm and using expression jython:import time import comman ds import random # get status and output of the comma nd status, output = commands.getstatusoutput(value) # add a random between 2 and 5s pause to avoid ddos on servers... Be kind to APIs! time.sleep(random.randint(2, 5)) # returns the result of the command return output.de code("utf-8")
- ✓ Create column anon whois at index 3 based on column command using expression grel:"torify "+value

```
          "type": "list",
          "name": "Starred Rows",
          "expression": "row.starred",
          "columnName": "",
          "invert": false,
          "omitBlank": false,
          "omitError": false,
          "selection": [
            {
              "v": {
                "v": true,
                "l": "true"
              }
            }
          ],
          "selectBlank": false,
          "selectError": false
        }
      ],
      "mode": "row-based"
    },
    "baseColumnName": "ddgr",
    "expression": "jython:import time\nimport comm
    "onError": "set-to-blank",
    "newColumnName": "resultddgr",
    "columnInsertIndex": 3,
    "description": "Create column resultddgr at in
  }
]
```

Select All | Unselect All

Close

- save this recipe as a text.
- You will use this recipe with the button APPLY.