

Stuttgart AG Tagung 2019 GAVO Puzzler: Solution

You'll want to solve this problem using TAP. And for one-off jobs, the recommended TAP client is TOPCAT (there's nothing wrong with using, say, pyVO or Aladin if that's your usual tooling, of course).

So, start TOPCAT, enter the VO/Table Access Protocol window, type "ngc" into the "Keywords" field, and in the result, select *GAVO DC TAP*, where there's a table with the OpenNGC (again, you can use an NGC on VizieR if you want; you'll have to adapt the queries then, though). Hit "Use Service", and have a look at the table and columns metadata for the `openngc.data` table.

What you want now is a statistics of counts versus HEALPixes. The GAVO DC service kindly provides an example for how to do something quite similar – look for it in the "Service-Provided" submenu under the "Examples" button. There's also some explanation for the example that you can read when you hit the "Info" button before you edit the example query; in particular, this tells you how to make TOPCAT draw HEALPix maps like those on the flyer.

You will have to edit that a bit to solve the first part of the problem to retrieve counts. That would yield a query somewhat like:

```
select
  count(*) as ct,
  ivo_healpix_index(3, raj2000, dej2000) as hpx
from openngc.data
group by hpx
```

If you sort the result, you can just read off the healpix. On level 3, that's HEALPix 433, on level 4, it's 2068.

There's a little catch here if you want to generate the HEALPix maps as per the example recipe, because currently TOPCAT doesn't like NULL HEALPix indices when drawing such maps. OpenNGC has NULL positions for a couple of obscure objects, and you'll have to filter these in order to have a valid index on all rows. That would be:

```
select
  count(*) as ct,
  ivo_healpix_index(3, raj2000, dej2000) as hpx
from openngc.data
where raj2000 is not null
group by hpx
```

But how do you figure out the positions? The trouble is that you can't just select `raj2000` and `dej2000` in your query, because you don't group by them. Think about it: In each group, there are as many *different* RAs and Decs as there are objects. Which of these should the service return?

No, you need to convert your `hpx` column to the center of the corresponding HEALPix. You could use some external tool for that, but really, TOPCAT and TAP are enough. To see how, inspect the "Service" tab in TOPCAT. If you scroll down a little, you see a list of user-defined functions available on the service. One of those is `ivo_healpix_center(hpxOrder INTEGER, hpxIndex BIGINT)` – that's exactly what we need.

You can't however, select that directly, because you can't use `hpx` in the select clause. You could do a table upload with your result tables, perhaps like this:

```
select my.*, ivo_healpix_center(3, hpx) as pos
from tap_upload.t1
```

But it's a lot more elegant to do it all in one query, and that's not hard if you use the magic of subqueries. This would look like this:

```
select top 5
q.*, ivo_healpix_center(3, hpx) as pos
from (
  select
    count(*) as ct,
    ivo_healpix_index(3, raj2000, dej2000) as hpx
  from openngc.data
  where raj2000 is not null
  group by hpx) as q
order by ct desc
```

As you can see, I've also taken the liberty of limiting the query to the five most populated healpixes (I'm not just getting the topmost one as it's always better to see if the runners-up also look plausible lest the top one is the result of a bug).

The results are something like 185.625, 9.5941 on level three, and 83.57143, -69.42255 on level four – so, changing the scale clearly changes the answers a lot. The sky is really inhomogeneous.

But what constellations are these in? Well, turns out nobody has yet made a table with the constellations as TAP-queriable polygons. We will, some day. Promised.

Until then, you could just fire up stellarium and look. Or use ancient VO tech in the form of the venerable old Simple Cone Search. Open TOPCAT's window for that, type constellations in the search field, and you'll see a service that returns constellation names for (ICRS) positions. Cut and paste the positions you've found, and you'll see that on level 3, it's Virgo and its rich galaxy cluster that takes the prize, whereas on level 4, it's Doradus with all the emission nebulae and clusters in the LMC.

To get a feeling of what that means: HEALPix level 7 is about a full moon, and the scale doubles with each level. So, level 3 means bins of about eight degrees diameter, and level 4 is about four degrees.