

## Convert geochemical «Porphyry Cu Geochemistry Module» dataset to 3NF Database «sqlite» format.

### 1. Abstract

It's good practice to use the data in relational database form, because it's the way to use powerful SQL language and python math libraries in any combinations. The source dataset Porphyry Cu Geochemistry Module.xls was transformed to a SQLite database in third normal form (3NF).

Source data was transferred as is to sqlite database to table «src». Each table has integer primary key id with autoincrement. Foreign key constraints are virtual and realize without SQLite supports.

### 2. Table description

2.1. Table «holes» contains 14 rows with drill holes

2.2. Table «samples» contains data of each sample with link to the hole by  
samples.hole\_ref=holes.id

2.3. Table «samples» contains data of each sample with link to the hole by  
samples.hole\_ref=holes.id

2.4. Table «elements» contains data rows of 48 elements

2.5. Table «data» contains 183291 values of concentration to each chemical element in each sample.

- Link to the «sample» «data.sample\_ref»= «sample.id»
- Link to the «element» «data.element\_ref» = «elements.id»
- Expo = -2 for «ppm» and -6 for «pct»
- Mant – is the value of concentration, the «concentration» = «mant» \* 10<sup>«expo»</sup>
- Link to the data exclusion reason is «data.exclusion\_ref»= «exclusion\_reasons.id». This field is used to exclude wrong data due the statistical ambiguity. If «data.exclusion\_ref»=0 it means no exclusion normal data. If 1,2,3 etc... so look to «exclusion\_reasons» table about the reasons list.

Additional segment of data model contains next tables:

2.6. Table «exclusion\_reasons» contains list of reasons.

id	name
0	normal
1	apparent outlier
2	contradicts to base physics
3	results below the limit of detection

2.7. Table «geochem\_methods» contains list of methods 4 acid Digestion

id	name
1	1F2 – Near Total Digestion – ICP
2	UT-4M – Near Total Digestion ICP-MS
3	Ultratrace 4 – Near Total Digestion – ICP and ICP-MS
4	Ultratrace 6 – Near Total Digestion – ICP and ICP-MS
5	UT-6M – Near Total Digestion – ICP and ICP-MS
6	8 - 4 Acid ICP Assay

2.8. Table « detection\_limits» contains limits of detection with lower and upper bounds for each method and each element contains 228 rows with laboratory instruments detection limits  
From <https://actlabs.com/geochemistry/exploration-geochemistry/4-acid-near-total-digestion/>

- Link to the «sample» « detection\_limits. element\_ref »= «elements.id»
- Link to the «element» « detection\_limits. method\_ref» = « geochem\_methods.id»
- Detection limit is «detection\_limit»\*10^ «detection\_limit\_expo»

2.9. Table «assumptions» contains rows with calculation hypothesis

2.10. Table « distrib\_properties » contains rows with calculations of statistical parameters for each hypothesis.

- M1 is the Mean.
- M2 is the Variance.
- M3 and M4 is the 3rd and 4th order moment
- Skew reflects symmetry of distribution:  $M3 / (M2^{3/2})$
- Kurtosis reflects flatness of distribution:  $M4 / \text{Sqr}(M2)$

# Normalization of a geological dataset into a standart 3NF database.

