ESCO data:

1. Isco job
2. List of corresponding esco skills
3. For each esco skill – mark whether it is skill (essential or optional) or knowledge
4. Broader category

Notebook 01 – gets esco data through API calls - the result is 2 files:

* escoOcc2escoessentialSkills\_all\_with\_mark.csv (132 of skills)
* escoOcc2escooptionalSkills\_all\_with\_mark.csv (318 of skills)

They look like:

A screenshot of a computer

Description automatically generated

Code – is isco job, then 132 or 318 columns for skills and then for marks.

Notebook 2 - similarity table for isco

given the results of notebook 1

It makes the files:

* 'iscoEssentialSkills\_ALL.csv'
* 'iscoOptionalSkills\_ALL.csv'

which look like this:A table with numbers and numbers

Description automatically generated

And same for optional.

It is a similarity table between isco jobs based on how many esco skills these jobs have in common. Entries in the table - frequencies.

Second, this notebook also scrapes broader categories for each of the esco skill from the file in notebook 1. The result is files named:

* 'iscoSkillsCategoriesV1.1.0\_%s.csv'

Such files are created for each of the isco codes.

So in other words: each isco job gets 1 file with all the info about essential skills and 1 file with all the info about optional skills. Each files looks like this: number of rows correpodns to number of skills for that isco job and number of columns – number of skills \* 3 ( skills themselves, categories, and marks)

A screenshot of a computer

Description automatically generated

PS This V1.1.0 is a version and it is a parameter in API call – it was the latest version as of 2022, could be different now.

Notebook 3 - constructs a concise table of the requested type: Isco job – skil – broader category code

It used the tables from above and restructures them the result is one single table named 'iscoEssentialSkills\_mapped\_all.csv'. This is how the result looks like:

A screenshot of a computer

Description automatically generated