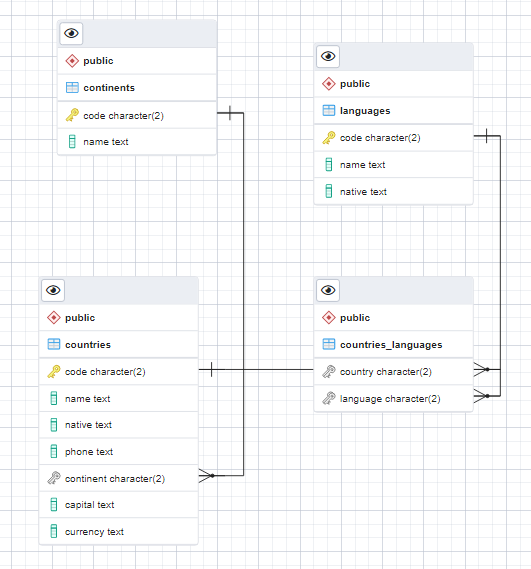
**Your test:**

**Task 1:** Use the country data from the below Github repository to complete the following task.

<https://github.com/annexare/Countries/tree/master/data> (Data source)

1. Create an ETL workflow that brings the data into an SQL database. You can use Postgres or MySQL. We plan to have a database that contains all continents, countries, currencies, and languages. (continents, countries, currencies, languages should all be in a different table)

**Data Model (ERD)**



**Github URL:** <https://github.com/prociple/decagon_etl>

**Github Repository:** [**https://github.com/prociple/decagon\_etl.git**](https://github.com/prociple/decagon_etl.git)

(Including the database creation sql scripts, data model diagram)

1. With the data in the database, create the following reports using SQL.
   1. **List all the continents and the total number of countries in each—for example, Africa 100, Europe 10, etc. The continent's name and country count should be in a different column.**

|  |  |
| --- | --- |
| SQL Query | Query result |
| SELECT continents.name "Continent",  count(countries.continent) "Country Count"  FROM public.countries countries  JOIN continents continents  ON continents.code = countries.continent  GROUP BY continents.name  ORDER BY 1; |  |
|  |

1. **List all the languages and commas separated countries that speak the language.**

|  |
| --- |
| SQL Query |
| SELECT lang.name "Language",  STRING\_AGG(coun.name, ', ') "Countries"  FROM countries\_languages con\_lang  join languages lang  ON con\_lang.language = lang.code  join countries coun  ON con\_lang.country = coun.code  GROUP BY lang.name |
|  |

|  |
| --- |
| Query result |
| \ |
|

**Note**: Languages spoken in many countries like (Arabic, English, Spanish, and French) were also captured separately to show the full view. Hence, they are highlighted with yellow color.

            c. **List all the countries and the total number of languages spoken.**

|  |  |
| --- | --- |
| SQL Query | Query result |
| SELECT countries.name Country,  count(con\_lang.language) "Language Count"  FROM public.countries countries  JOIN countries\_languages con\_lang  ON con\_lang.country = countries.code  GROUP BY countries.name  ORDER BY 1; |  |
|  |

Your submission:

1. Include a readme file that documents your approach.
2. Write your script in Python or R.
3. For each of the results, provide a screenshot.
4. Include all your scripts, screenshots, documentation, and anything else in a folder, and submit it to Github.

The task should take at most 24 hours from the time you receive this email.