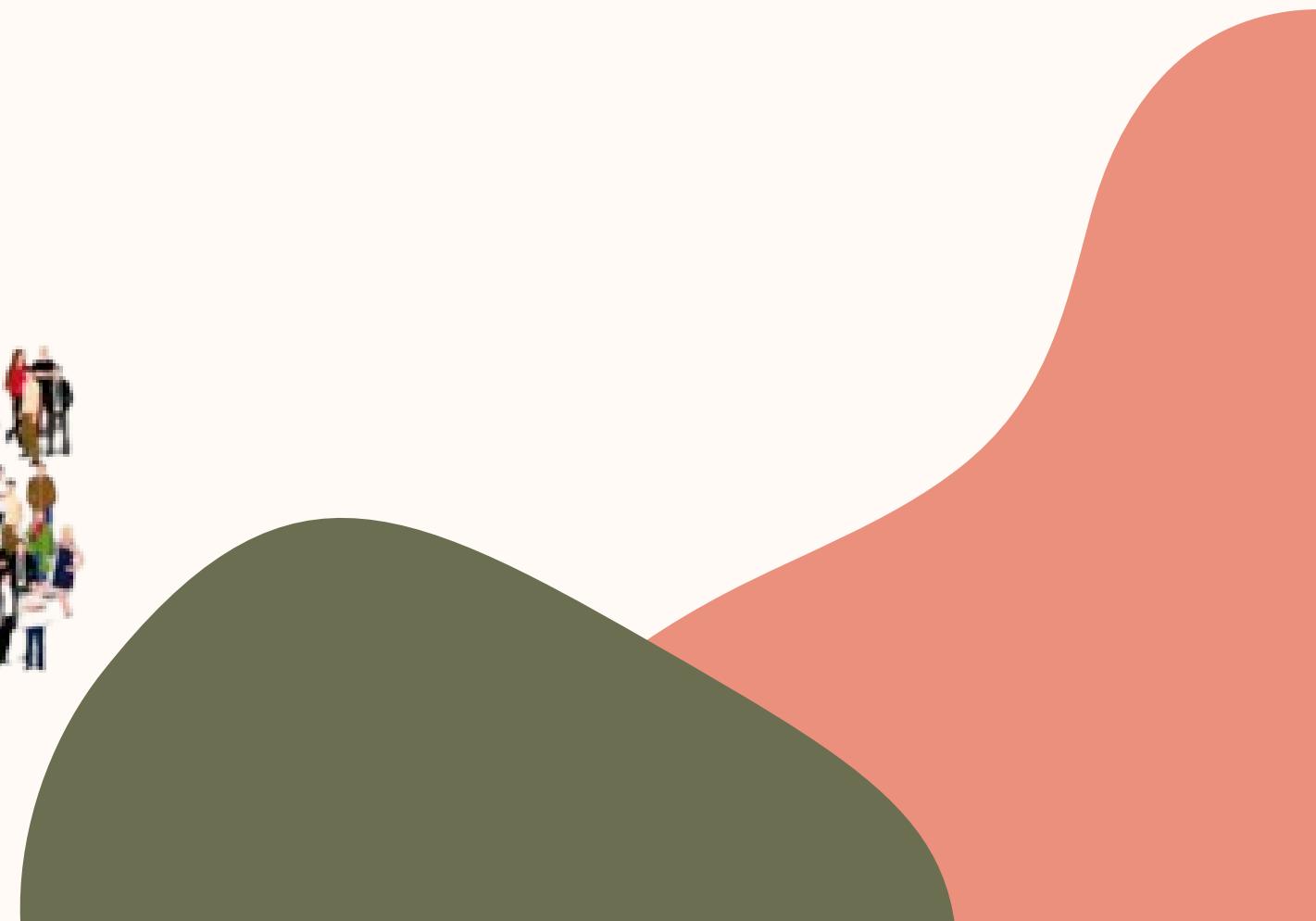

Projet Dana - Gr 6.



DATA SET.

Lien : <https://www.kaggle.com/datasets/iamsouravbanerjee/world-population-dataset>

Description : Évolution de la population des capitales de tout les pays de 1970 a 2022.

Le data set contient les informations suivantes : La colonne Rank indique le classement des pays selon leur population, CCA3 fournit un code à trois lettres associé à chaque pays. Le nom du pays dans Country/Territory, Capital et Continent. Plusieurs colonnes retracent l'évolution de la population à différentes périodes. La superficie du pays est donnée dans Area (km^2), la densité de population dans Density (per km^2). Enfin, Growth Rate exprime le taux de croissance de la population, et World Population Percentage indique la part que représente chaque pays dans la population mondiale totale.

Source: <http://www.kaggle.com>

ÉTAPE 1.

Choix du data set.

Un membre du groupe avait trouver l'idée, pour ce relier facilement aux autres, et comme nous n'avons pas l'information sur le choix des autres groupes. De prendre un data set large pouvant facilement être relier et la première idée a été d'utiliser un data set contenant les pays, villes. C'est comme cela que nous avons trouver notre data set.

# Rank	Rank	Δ CCA3	CCA3	Δ Country/Territory	Country/Territory	Δ Capital	Capital	Δ Continent	Continent	# 2022 Population	2022 Population	# Area (km²)	Area (km²)	# Density (per km²)	Density (per km²)	# Growth Rate	Growth Rate
1	234	234	unique values	234	unique values	234	unique values	Africa	24%	510	1.43b	1	17.1m	0.03	23.2k	0.91	
36	AFG	Afghanistan	Kabul	Asia	41128771	652230	63.8587	1.0257									
138	ALB	Albania	Tirana	Europe	2842321	28748	98.8702	0.9957									
34	DZA	Algeria	Algiers	Africa	44903225	2381741	18.8531	1.0164									
213	ASM	American Samoa	Pago Pago	Oceania	44273	199	222.4774	0.9831									
203	AND	Andorra	Andorra la Vella	Europe	79824	468	170.5641	1.01									
42	AGO	Angola	Luanda	Africa	35588987	1246700	28.5466	1.0315									
224	AIA	Anguilla	The Valley	North America	15857	91	174.2527	1.0066									
201	ATG	Antigua and Barbuda	Saint John's	North America	93763	442	212.1335	1.0058									
33	ARG	Argentina	Buenos Aires	South America	45510318	2780400	16.3683	1.0052									
140	ARM	Armenia	Yerevan	Asia	2788469	29743	93.4831	0.9962									
198	ABW	Aruba	Oranjestad	North America	106445	180	591.3611	0.9991									
55	AUS	Australia	Canberra	Oceania	26177413	7692024	3.4032	1.0099									
99	AUT	Austria	Vienna	Europe	8939617	83871	106.5877	1.002									
91	AZE	Azerbaijan	Baku	Asia	10358074	86600	119.6082	1.0044									
176	BHS	Bahamas	Nassau	North America	409984	13943	29.4843	1.0051									
154	BHR	Bahrain	Manama	Asia	1472233	765	1924.4876	1.0061									

data set simplifié

ÉTAPE 1.

Transformation en RDF.

via open refine

plusieur version avec diff prefixx

oublie du type xml xsd

RDF Schema alignment

The RDF schema alignment skeleton below specifies how the RDF data that will get generated from your grid-shaped data. The cells in each record of your data will get placed into nodes within the skeleton. Configure the skeleton by specifying which column to substitute into which node.

Base URI: <http://127.0.0.1:3333/> [Edit](#)

[RDF skeleton](#) [RDF Preview](#)

Available prefixes: ex dbo schema xsd rdfs [+ Add](#) [Manage](#)

(Row index) URI [Add type](#)

[Add property](#)

[Rank Cell](#)
[CCA3 Cell](#)
[Country/Territory Cell](#)
[Capital Cell](#)
[Continent Cell](#)
[2022 Population Cell](#)
[2020 Population Cell](#)
[2015 Population Cell](#)
[2010 Population Cell](#)
[2000 Population Cell](#)
[1990 Population Cell](#)
[1980 Population Cell](#)
[1970 Population Cell](#)
[Area \(km²\) Cell](#)
[Density \(per km²\) Cell](#)
[Growth Rate Cell](#)
[World Population Percentage Cell](#)

[dbo:rank](#)
[dbo:countryCode](#)
[rdfs:label](#)
[dbo:capital](#)
[dbo:continent](#)
[dbo:populationTotal](#)
[ex:pop2020](#)
[ex:pop2015](#)
[ex:pop2010](#)
[ex:pop2000](#)
[ex:pop1990](#)
[ex:pop1980](#)
[ex:pop1970](#)
[dbo:areaTotal](#)
[dbo:populationDensity](#)
[schema:growthRate](#)
[schema:populationShare](#)

RDF skeleton OpenRefine

ÉTAPE 1.

Requête 1.

dire que fait la requête

PREFIX dbo: <<http://dbpedia.org/ontology>>

PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

```
SELECT ?country ?popDensity
WHERE {
  ?s rdfs:label ?country ;
    dbo:populationDensity ?popDensity ;
  FILTER (regex(?country,"^A"))
}
ORDER BY (?popDensity)
LIMIT 10
```

Résultat de la requête 1 Fuseki

country	popDensity
Australia	3.4032
Argentina	16.3683
Algeria	18.8531
Angola	28.5466
Afghanistan	63.0587
Armenia	93.4831
Albania	98.8702
Austria	106.5877
Azerbaijan	119.6082
Andorra	170.5641
Anguilla	174.2527
Antigua and Barbuda	212.1335
American Samoa	222.4774
Aruba	591.3611

ÉTAPE 1.

Requête 2.

dire que fait la requête

PREFIX schema: <http://schema.org/>

PREFIX dbo: <http://dbpedia.org/ontology/>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

```
SELECT ?country ?area
WHERE {
  ?s rdfs:label ?country ;
    dbo:areaTotal ?area ;
    schema:growthRate ?gr ;
    dbo:populationTotal ?pop .
  FILTER (?gr < 1 && ?area > 300000)
}
ORDER BY (?pop)
LIMIT 10
```

Résultat de la requête 2 Fuseki

country	area
Ukraine	603500
Italy	301336
Germany	357114
Japan	377930
Russia	17098242

Merci pour votre écoute !.



Lien GitHub : https://github.com/BraKann/Projet_RDF_WorldPopulation.git

