In [25]:
```
import pandas as pd
import json

with open('tbs-prov-std/Canadian_prov_terr_du_Canada.json') as f:
    prov_json = json.loads(f.read())

# json contains $schema and data, we're interested in the latter.
prov_df = pd.json_normalize(prov_json, record_path='data').set_index(
'code')
prov_df
```

Out[25]:

code	nm_en	nm_fr	ab_en	ab_fr
NL	Newfoundland and Labrador	Terre-Neuve-et-Labrador	N.L.	T.-N.-L.
PE	Prince Edward Island	Île-du-Prince-Édouard	P.E.I.	Î.-P.-É.
NS	Nova Scotia	Nouvelle-Écosse	N.S.	N.-É.
NB	New Brunswick	Nouveau-Brunswick	N.B.	N.-B.
QC	Quebec	Québec	Que.	Qc
ON	Ontario	Ontario	Ont.	Ont.
MB	Manitoba	Manitoba	Man.	Man.
SK	Saskatchewan	Saskatchewan	Sask.	Sask.
AB	Alberta	Alberta	Alta.	Alb.
BC	British Columbia	Colombie-Britannique	B.C.	C.-B.
YT	Yukon	Yukon	Y.T.	Yn
NT	Northwest Territories	Territoires du Nord-Ouest	N.W.T.	T.N.-O.
NU	Nunavut	Nunavut	Nvt.	Nt

In [26]:
```
# print the French name for QC
prov_df.loc['QC']['nm_fr']
```

Out[26]:
```
'Québec'
```

In [27]:
```
# load the crosswalk, and merge with provinces
with open('tbs-prov-std/crosswalk_prov_terr_correspondance.json') as f
:
    prov_crosswalk_json = json.loads(f.read())

prov_crosswalk_df = pd.json_normalize(prov_crosswalk_json, record_path
='data').set_index('code')
prov_df = prov_df.join(prov_crosswalk_df)
prov_df
```

Out[27]:

code	nm_en	nm_fr	ab_en	ab_fr	code_statcan	code_iso	cp_ab_pc
NL	Newfoundland and Labrador	Terre-Neuve-et-Labrador	N.L.	T.-N.-L.	10	CA-NL	NL
PE	Prince Edward Island	Île-du-Prince-Édouard	P.E.I.	Î.-P.-É.	11	CA-PE	PE
NS	Nova Scotia	Nouvelle-Écosse	N.S.	N.-É.	12	CA-NS	NS
NB	New Brunswick	Nouveau-Brunswick	N.B.	N.-B.	13	CA-NB	NB
QC	Quebec	Québec	Que.	Qc	24	CA-QC	QC
ON	Ontario	Ontario	Ont.	Ont.	35	CA-ON	ON
MB	Manitoba	Manitoba	Man.	Man.	46	CA-MB	MB
SK	Saskatchewan	Saskatchewan	Sask.	Sask.	47	CA-SK	SK
AB	Alberta	Alberta	Alta.	Alb.	48	CA-AB	AB
BC	British Columbia	Colombie-Britannique	B.C.	C.-B.	59	CA-BC	BC
YT	Yukon	Yukon	Y.T.	Yn	60	CA-YT	YT
NT	Northwest Territories	Territoires du Nord-Ouest	N.W.T.	T.N.-O.	61	CA-NT	NT
NU	Nunavut	Nunavut	Nvt.	Nt	62	CA-NU	NU

In [28]:
```
# load a dataset on murder rates. it has columns:
#    geo_code StatCan province code.
#    UOM unit of measure description
#    value murder rate per 100k, with 2 decimal points (e.g. 1.71)
murder_df = pd.read_csv('sample_dataset_homicides.csv')
murder_df.columns = murder_df.columns.str.lower() # no need to SCREAM
murder_df = murder_df.set_index('geo_code')
report = murder_df.join(prov_df.set_index('code_statcan'))
report[['nm_en','nm_fr','value','uom']]
```

Out[28]:

geo_code	nm_en	nm_fr	value	uom
10	Newfoundland and Labrador	Terre-Neuve-et-Labrador	0.96	Rate per 100,000 population
11	Prince Edward Island	Île-du-Prince-Édouard	1.27	Rate per 100,000 population
12	Nova Scotia	Nouvelle-Écosse	0.62	Rate per 100,000 population
13	New Brunswick	Nouveau-Brunswick	1.93	Rate per 100,000 population
24	Quebec	Québec	0.91	Rate per 100,000 population
35	Ontario	Ontario	1.69	Rate per 100,000 population
46	Manitoba	Manitoba	5.26	Rate per 100,000 population
47	Saskatchewan	Saskatchewan	4.68	Rate per 100,000 population
48	Alberta	Alberta	2.29	Rate per 100,000 population
59	British Columbia	Colombie-Britannique	1.77	Rate per 100,000 population
60	Yukon	Yukon	2.45	Rate per 100,000 population
61	Northwest Territories	Territoires du Nord-Ouest	4.46	Rate per 100,000 population
62	Nunavut	Nunavut	18.05	Rate per 100,000 population