Data were munged here (https://github.com/TheOregonian/long-term-care-db/blob/master/notebooks/transformation/mung-3-29-scrape.ipynb)

In [33]:
```
import pandas as pd
import numpy as np
from IPython.core.display import display, HTML
display(HTML("<style>.container { width:100% !important; }</style>"))
```

In [34]:
```
df = pd.read_csv('../../data/processed/complaints-3-29-scrape.csv')
```

How many facility ids are associated with multiple online names?

In [35]:
```
id_name = df[['facility_id','online_fac_name']][df['online_fac_name'].notnull()].drop_duplicates().rename(columns={'online_fac_name':'online_name'})
```

In [36]:
```
id_name.head()
```

Out[36]:

	facility_id	online_name
1	385008	PRESBYTERIAN COMMUNITY CARE CENTER
24	385010	LAURELHURST VILLAGE REHABILITATION CENTER
30	385010	LAURELHURST VILLAGE
76	385015	REGENCY GRESHAM NURSING & REHABILITATION CENTER
82	385015	REST HARBOR REHABILITATION & EXTENDED CARE

In [37]:
```
names_per_id = id_name.groupby('facility_id').count().reset_index()
```

In [38]:
```
names_per_id.head()
```

Out[38]:

	facility_id	online_name
0	385008	1
1	385010	2
2	385015	2
3	385018	1
4	385024	2

In [39]:
```
names_per_id[names_per_id['online_name']>1].count()[0]
```

Out[39]: 168

Checking if multiple facilities have the same name