Demolition and construction

This notebook uses lists of permits for residential demolition and residential construction to compare the total sizes of properties demolished and constructed on the same property.

```
In [1]: import pandas as pd
```

```
In [2]: demo = pd.read_csv(
            '../data-processed/Demo residential.csv',
            dtype=str
        )

        construction = pd.read_csv(
            '../data-processed/New residential.csv',
            dtype=str
        )
```

```
In [3]: dat = pd.DataFrame(demo[
            # Filter for properties in both demolition and construction lists
            demo.TCAD_ID.isin(
                construction.TCAD_ID
            )
        ].assign(
            TotalExistingBldgSQFT = lambda x: x.TotalExistingBldgSQFT.str.replace(
                ',', ''
            ).astype(int)
        ).query(
            # Filter for demolitions where space demolished was more than zero
            'TotalExistingBldgSQFT > 0'
        ).groupby(
            # Sum the total space demolished for each property
            # Some properties had multiple demolition permits
            'TCAD_ID'
        ).TotalExistingBldgSQFT.sum().rename(
            'demolition'
        )).join(
            construction.assign(
                TotalNewAddSQFT = lambda x: x.TotalNewAddSQFT.astype(int)
            ).groupby(
                # Sum the total construction space for each property
                # and join to demolition totals on parcel id.
                'TCAD_ID'
            ).TotalNewAddSQFT.sum().rename('construction')
        ).assign(
            # Compare total construction space to total demolition space
            more_construction = lambda x: x.construction > x.demolition
        )
```

In [4]:
```python
print(
    'Among {:,} properties with sizes given for both demolition and constructi
on, how many had more total space constructed than demolished?\n'.format(
        dat.shape[0]
    )
)

pd.DataFrame(
    dat.more_construction.value_counts().apply(
        '{:,}'.format
    ).rename('number')
).join(
    pd.DataFrame(dat.more_construction.value_counts(normalize=True).apply(
        '{:.2%}'.format
    ).rename('percent'))
)
```

Among 2,370 properties with sizes given for both demolition and construction,
how many had more total space constructed than demolished?

Out[4]:

	number	percent
True	2,291	96.67%
False	79	3.33%

In [ ]: