In [2]:
```
import pandas as pd
import numpy as np
```

In [6]:
```
df = pd.read_csv("lego/sets.csv")
```

In [8]:
```
df.head(20)
```

Out[8]:

	set_num	name	year	theme_id	num_parts
0	00-1	Weetabix Castle	1970	414	471
1	0011-2	Town Mini-Figures	1978	84	12
2	0011-3	Castle 2 for 1 Bonus Offer	1987	199	2
3	0012-1	Space Mini-Figures	1979	143	12
4	0013-1	Space Mini-Figures	1979	143	12
5	0014-1	Space Mini-Figures	1979	143	12
6	0015-1	Space Mini-Figures	1979	143	18
7	0016-1	Castle Mini Figures	1978	186	15
8	00-2	Weetabix Promotional House 1	1976	413	147
9	00-3	Weetabix Promotional House 2	1976	413	149
10	00-4	Weetabix Promotional Windmill	1976	413	126
11	005-1	Basic Building Set in Cardboard	1965	366	35
12	00-6	Special Offer	1985	67	3
13	00-7	Weetabix Promotional Lego Village	1976	413	3
14	010-1	Basic Building Set in Cardboard	1965	366	57
15	010-3	Basic Building Set	1968	366	77
16	011-1	Basic Building Set	1968	366	145
17	022-1	Basic Building Set	1968	366	110
18	03093-1	The Race to Build It Board Game	1999	502	70
19	033-2	Basic Building Set	1968	366	177

Make DB File with DF

In [16]:
```
from sqlalchemy import create_engine
engine = create_engine('sqlite:///db.db', echo=False)

with engine.begin() as connection:
    df.to_sql('lego', con=connection, if_exists='append')
```

Make SQLite3 File with DF

In [14]:
```
from sqlalchemy import create_engine
engine = create_engine('sqlite:///lego_sqlite_file.sqlite3', echo=False)

with engine.begin() as connection:
    df.to_sql('lego_sqlite3_file', con=connection, if_exists='append')
```

Read SQLite3 File

In [33]:
```
import sqlite3

con = sqlite3.connect("lego_sqlite3_file.sqlite3")

# Load the data into a DataFrame
lego_df = pd.read_sql_query("SELECT * FROM lego_sqlite3_file", con)
```

In [37]:
```
lego_df = pd.read_csv("lego/sets.csv")
```

Find Common Columns

In [38]:
```
# Finding Common columns
a = np.intersect1d(df.columns, lego_df.columns)

# Printing common columns
print ("Common Columns:",a)
```

```
Common Columns: ['name' 'num_parts' 'set_num' 'theme_id' 'year']
```

Nth Percentile (50% Used)

In [39]:
```
df.head()
```

Out[39]:

	set_num	name	year	theme_id	num_parts
0	00-1	Weetabix Castle	1970	414	471
1	0011-2	Town Mini-Figures	1978	84	12
2	0011-3	Castle 2 for 1 Bonus Offer	1987	199	2
3	0012-1	Space Mini-Figures	1979	143	12
4	0013-1	Space Mini-Figures	1979	143	12

In [42]:
```
df.quantile(.5)
```

Out[42]:
```
year         2005.0
theme_id      324.0
num_parts      45.0
Name: 0.5, dtype: float64
```

In [ ]: