```
import pandas as pd

url = 'https://raw.githubusercontent.com/zachparker2014/Zachary-Parker-IntrotoMachineLearning
df = pd.read_csv(url, index_col = 0)
```

```
df.head(7)
```

	carat	cut	color	clarity	depth	table	price	x	y	z
1	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43
2	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31
3	0.23	Good	E	VS1	56.9	65.0	327	4.05	4.07	2.31
4	0.29	Premium	I	VS2	62.4	58.0	334	4.20	4.23	2.63
5	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75
6	0.24	Very Good	J	VVS2	62.8	57.0	336	3.94	3.96	2.48
7	0.24	Very Good	I	VVS1	62.3	57.0	336	3.95	3.98	2.47

```
df.tail(7)
```

	carat	cut	color	clarity	depth	table	price	x	y	z
53934	0.70	Very Good	E	VS2	61.2	59.0	2757	5.69	5.72	3.49
53935	0.72	Premium	D	SI1	62.7	59.0	2757	5.69	5.73	3.58
53936	0.72	Ideal	D	SI1	60.8	57.0	2757	5.75	5.76	3.50
53937	0.72	Good	D	SI1	63.1	55.0	2757	5.69	5.75	3.61
53938	0.70	Very Good	D	SI1	62.8	60.0	2757	5.66	5.68	3.56
53939	0.86	Premium	H	SI2	61.0	58.0	2757	6.15	6.12	3.74
53940	0.75	Ideal	D	SI2	62.2	55.0	2757	5.83	5.87	3.64

```
df.describe()
```

	carat	depth	table	price	x	y
count	53940.000000	53940.000000	53940.000000	53940.000000	53940.000000	53940.000000
mean	0.797940	61.749405	57.457184	3932.799722	5.731157	5.734526
std	0.474011	1.432621	2.234491	3989.439738	1.121761	1.142135
min	0.200000	43.000000	43.000000	326.000000	0.000000	0.000000

```
df.describe(include = 'object')
```

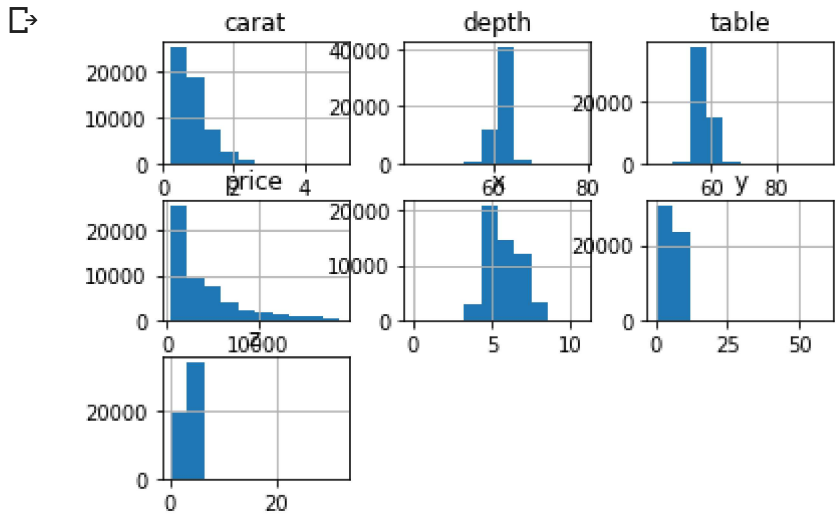
	cut	color	clarity
count	53940	53940	53940
unique	5	7	8
top	Ideal	G	SI1
freq	21551	11292	13065

```
cut_uniques = df['cut'].unique()
print(cut_uniques)
color_uniques = df['color'].unique()
print(color_uniques)
clarity_uniques = df['clarity'].unique()
print(clarity_uniques)
```

```
['Ideal' 'Premium' 'Good' 'Very Good' 'Fair']
['E' 'I' 'J' 'H' 'F' 'G' 'D']
['SI2' 'SI1' 'VS1' 'VS2' 'VVS2' 'VVS1' 'I1' 'IF']
```

```
%matplotlib inline
```

```
hist = df.hist()
```



+ Code + Text

✓ 0s completed at 11:56 AM