

Diamond Sizes

```
# setup
import altair as alt
import pandas as pd
import numpy as np
alt.data_transformers.enable('json')

# load and format data
diamonds = pd.read_csv("https://github.com/byuidatascience/data4python4ds/raw/master/diamonds.csv")
smaller = diamonds.query('carat <= 2.5')
```

We have data about diamonds. Only those larger than 2.5 carats. The distribution of the remainder is shown below:

```
# plot data
chart = (alt.Chart(smaller).
    encode(alt.X('carat'), alt.Y('count()')).
    mark_line())

chart.save('markdown/diamonds_25.png')
```

