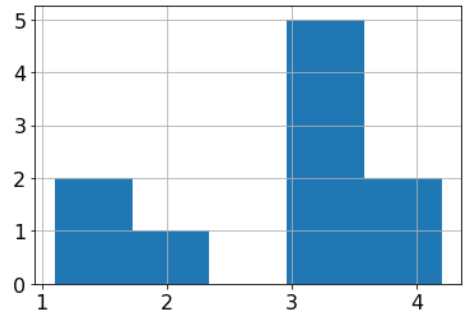
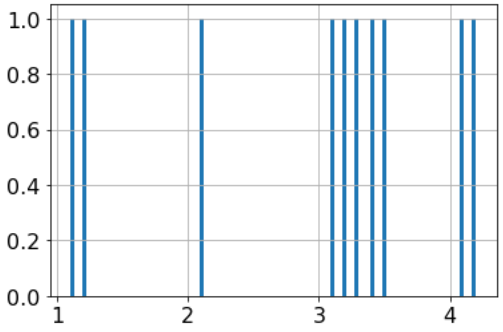
Custom Plots

Density Plots

Problem 1: these are the same data





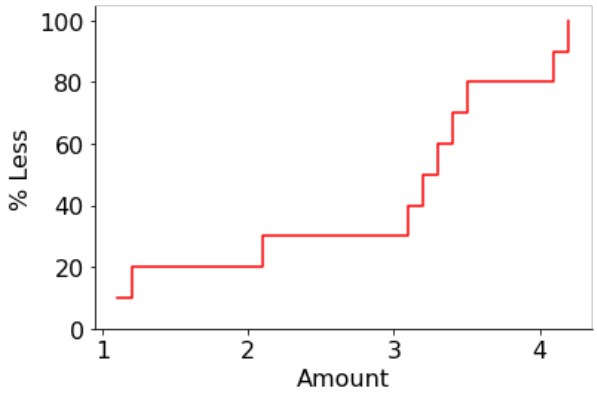
Problem 2: many distributions can't fit in the same area

CDF (Cumulative Distribution Function) Plots

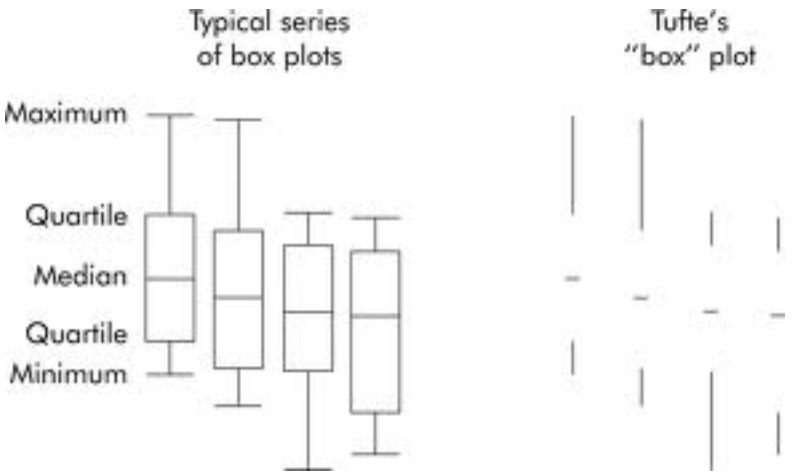
```
def make_cdf(vals):
    s = pd.Series(sorted(vals))
    s = s.sort_values()
    return pd.Series(100*(s.index+1)/len(s), index=s.values)
cdf = make_cdf(vals)

ax=None
#ax = cdf.plot.line(ylim=0, color="red", ax=ax) # OK FOR LOTS OF DATA
#ax = cdf.plot.line(ylim=0, color="red", drawstyle='steps-pre', ax=ax)
# WRONG
ax = cdf.plot.line(ylim=0, color="red", drawstyle='steps-post', ax=ax)
ax.set_xlabel("Amount")
ax.set_ylabel("% Less")
ax.spines["right"].set_visible(False)
ax.spines["top"].set_visible(False)
```

```
1.1     10.0
1.2     20.0
2.1     30.0
3.1     40.0
3.2     50.0
3.3     60.0
3.4     70.0
3.5     80.0
4.1     90.0
4.2    100.0
dtype: float64
```



Box Plots



https://stats.stackexchange.com/questions/81451/combining-box-plot-with-frequency-count?noredirect=1&lq=1

Creating Points and Lines:

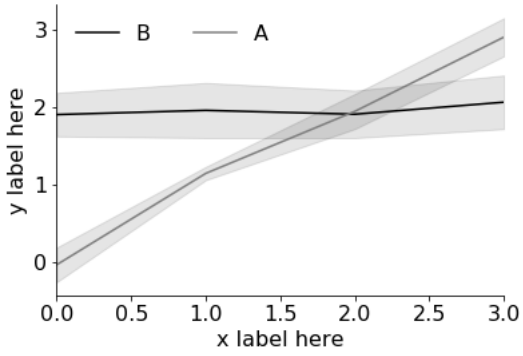
```
- ax.plot(x, y, 'ro')              # plot a red circle at point x,y
- ax.plot((x1, x2), (y1, y2), 'k') # black line from x1,y1 to x2,y2
```

Custom x-ticks:

```
- ax.set_xticklabels(list(df.columns))
- ax.set_xticks(range(1, len(df.columns)+1))
```

Minimalist Plots:

```
- ax.legend(frameon=False, ncol=2)
- ax.spines['right'].set_visible(False)
- ax.spines['top'].set_visible(False)
```



Representing Standard Deviation

Useful snippets:

```
- df.mean(axis=0)
- mean = df.mean(level=1)
- std = df.std(level=1)
```

```
plt.fill_between(mean.index, mean[col]-std[col], mean[col]+std[col],
                 color='black', alpha=0.1)
```