In [33]:

```
output:
```

(note, I accidentally deleted one of the blue cells and could not get the highlighter to re-highlight my first plot)1. This strip-plot shows how many uppercase words are in each email. This plot was useful to me because it shows that this feature does not do much to differentiate between spam and ham, which would be harder to determine using a bar plot.

(2). This violin plot was useful because it shows the difference between distributions of spam and ham based on percentage of uppercase words in email subjects. The spam plot is more evenly distributed, while the ham plot has a wide and concentrated base, with a small secondary distribution that reaches up to about 40%. The small cluster of outliers in spam, which tells me that this feature is doing something to differentiate spam from ham.

Out[33]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x113d39128>
```