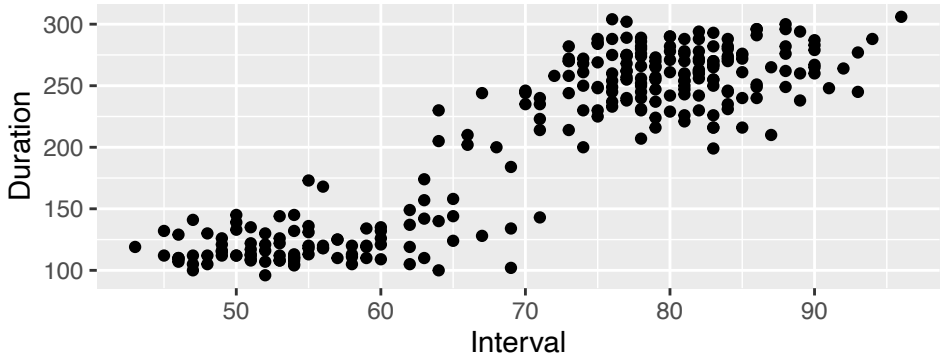
E1.1

The scatterplot shows a definite pattern in which, after a longer interval, an eruption tends to last longer (i.e., have a longer `Duration`).

```
gf_point(Duration ~ Interval, data=oldfaith)
```



E1.2

```
help(CPS85)
nrow(CPS85)
```

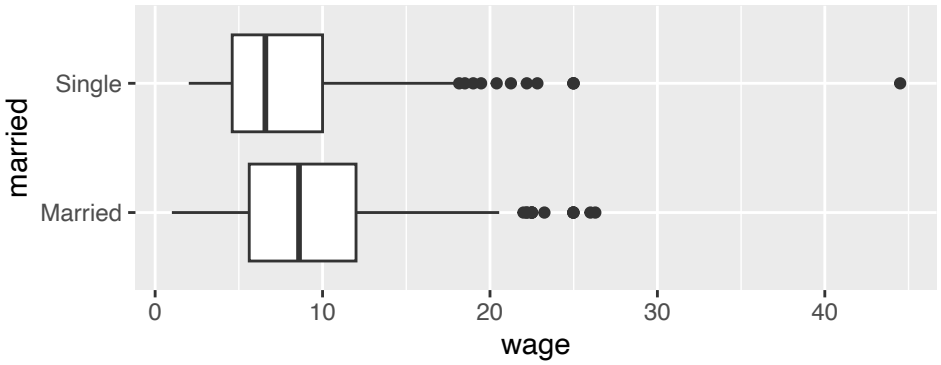
```
[1] 534
```

This data is from the 1985 Current Population Survey, a public survey used to supplement census information. The observational units are random people, most likely adult citizens of the United States. There are 534 observational units in the data set.

E1.5

Below, I look at the breakdown of `wage` based on `married`.

```
gf_boxplot(married ~ wage, data=CPS85)
```

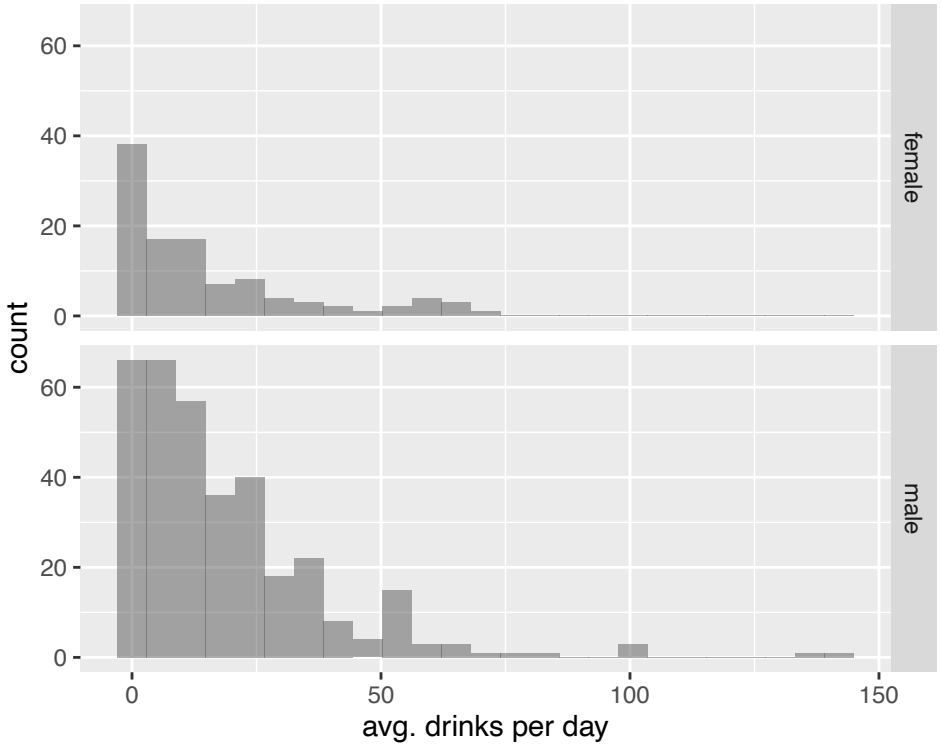


While there is one outlier among single workers with a higher wage than everyone else, there seems to be a slightly elevated distribution of wage among married workers.

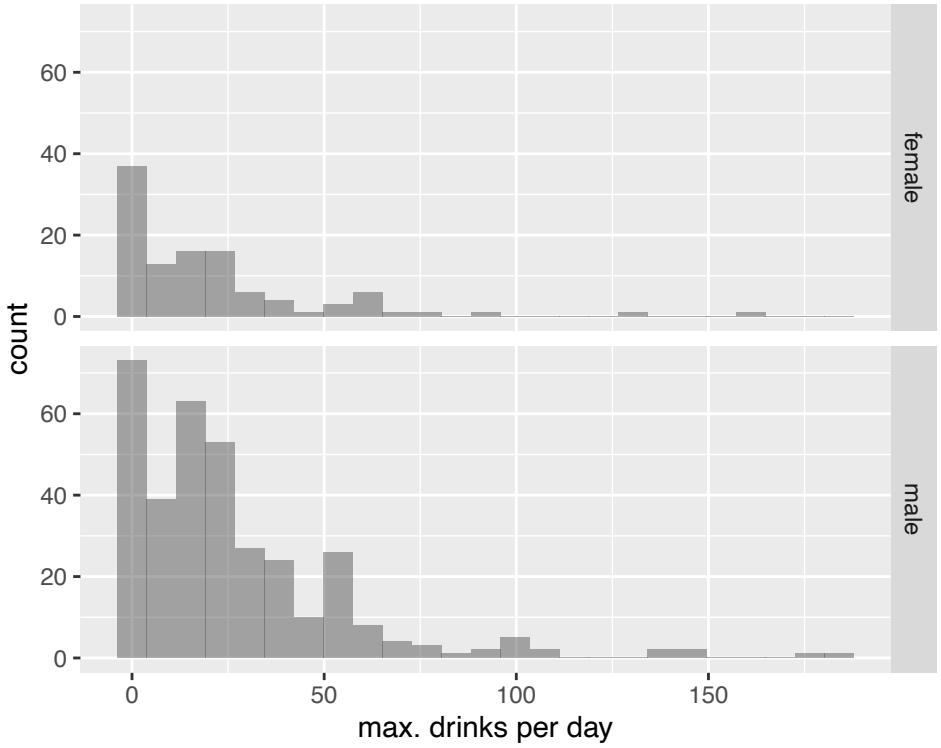
E1.12

The distribution of `i1` from `HELPrct` (average drinks per day) has the same general shape among women and men. However, there are larger outlying values for men than for women. Not surprisingly, this same description applies to maximum drinks per day (`i2`).

```
gf_histogram(~i1 | sex ~ ., data=HELPrct)
```



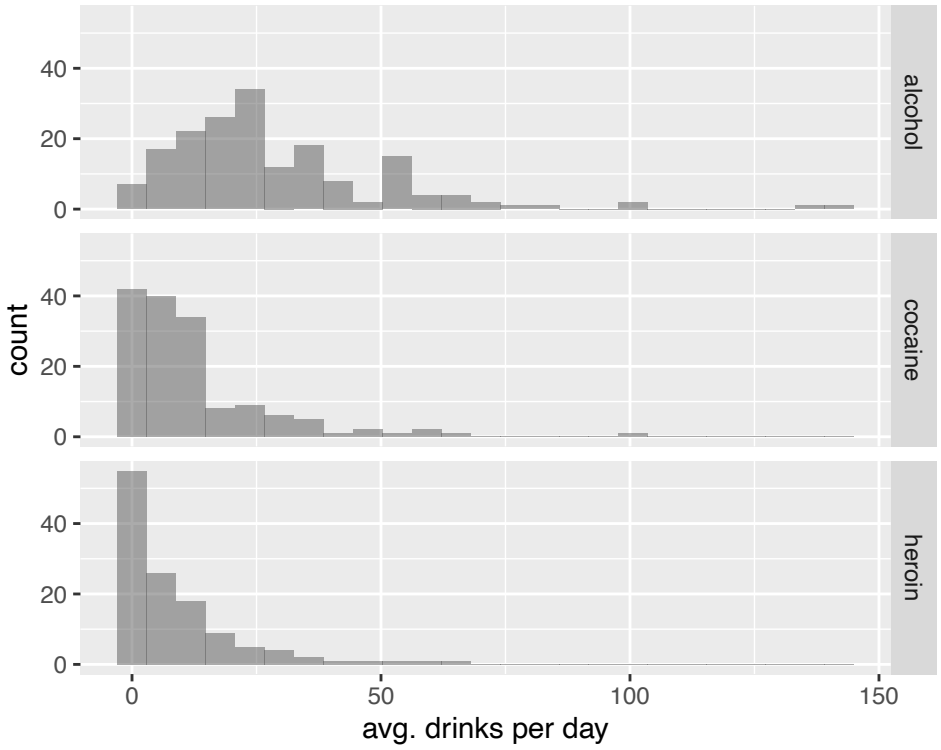
```
gf_histogram(~i2 | sex ~ ., data=HELPrct)
```



E1.13

Given the plots below, it appears drug (cocaine and heroin) addicts have a somewhat lower number of average and maximum count of drinks per day than the alcohol abusers, a fact which does not seem surprising. The `i2` plot may suggest heroin users are less prone to drink high numbers of drinks than cocaine users, which may be somewhat surprising.

```
gf_histogram(~i1 | substance ~ ., data=HELPrct)
```




```
gf_histogram(~i2 | substance ~ ., data=HELPrct)
```