exploration1-questions

Duu Renn

August 29, 2016

Contents

1 Data and Description **1**

2 Plots **1**

3 Regression already? **4**

Questions
As people ask questions, I'll post responses here and push them to the shared folder. You can email me these questions or create an "issue" in the 'uips-stat-share' repository. Note that this is different than the 'explorations' repository – we should probably try to keep that relatively clutter-free.

1 Data and Description

In the previous assignment, there were missing values which corresponded to an "NA" observation. Are there any such values in this dataset? Are you sure?

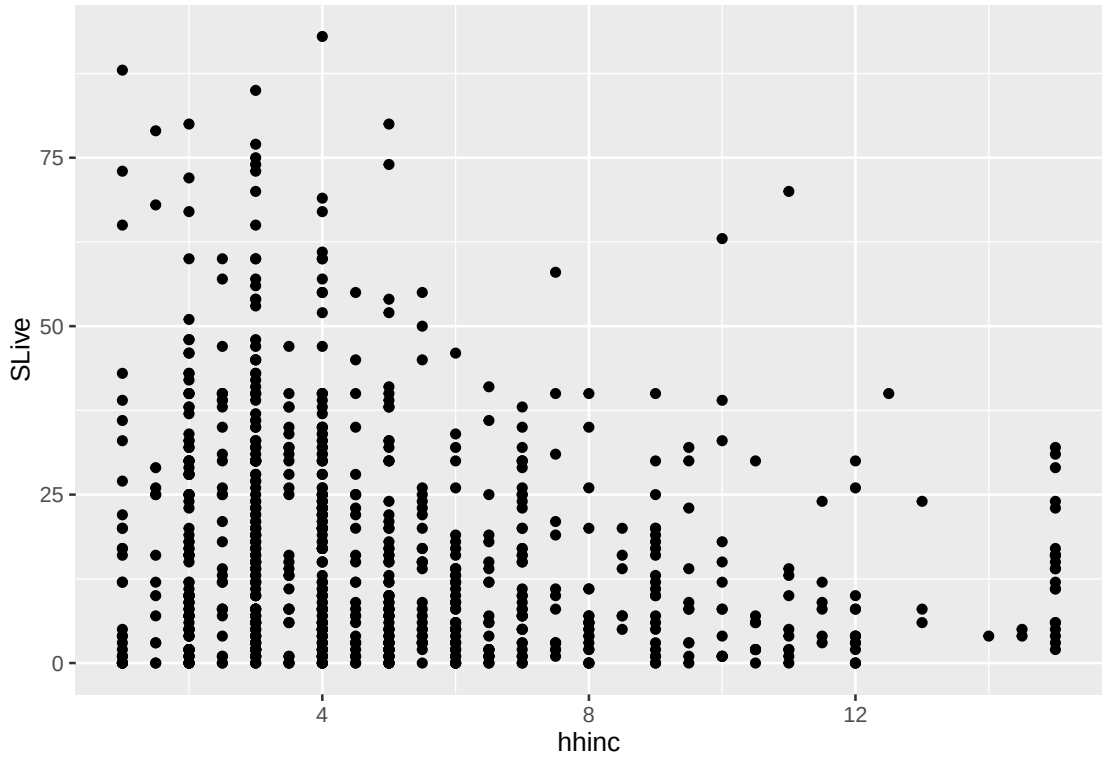
Make sure you explore the codebook!

Note that you're engaging in description of data. While discussion of significance is unnecessary, discussion of the variables is crucial – especially when it comes to coverage and content.

2 Plots

One thing that you'll want to include are plots of the data. It's up to you to interpret these and make decisions about the actual variables (this is a purposely vague statement). I'm just going to include another ggplot example if you decide to use that package instead of R's built in plot() command. Note that this example isn't even the dataset that you're working on this week, so it really is up to you to design your own figure.

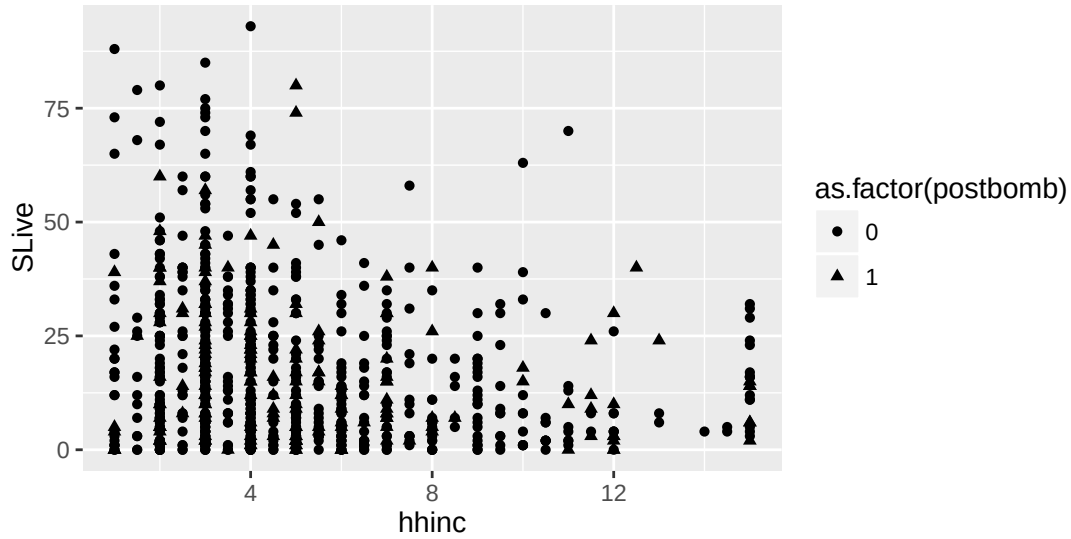
```
require(ggplot2)

ggplot(ho05, aes(hhinc, SLive)) + geom_point()
```



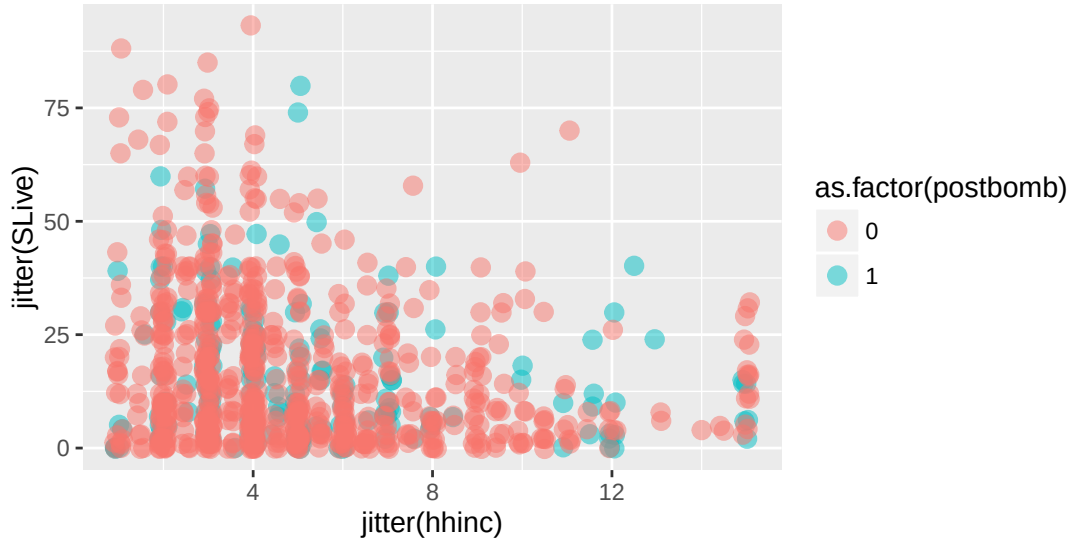

```
# Notice that this is missing some crucial things, like labels for X and Y
# as well as a title
```

You can introduce more "layers" to your visual depictions of the data, even just in X-Y plot. But make sure you choose the right type of variable.

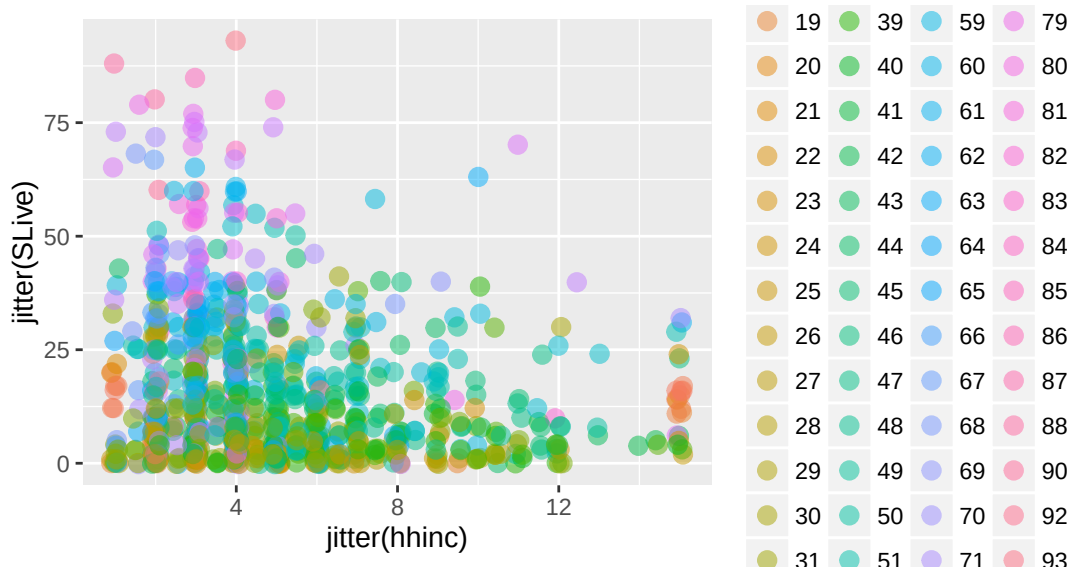
```
ggplot(ho05, aes(hhinc, SLive, shape = as.factor(postbomb))) + geom_point()
```



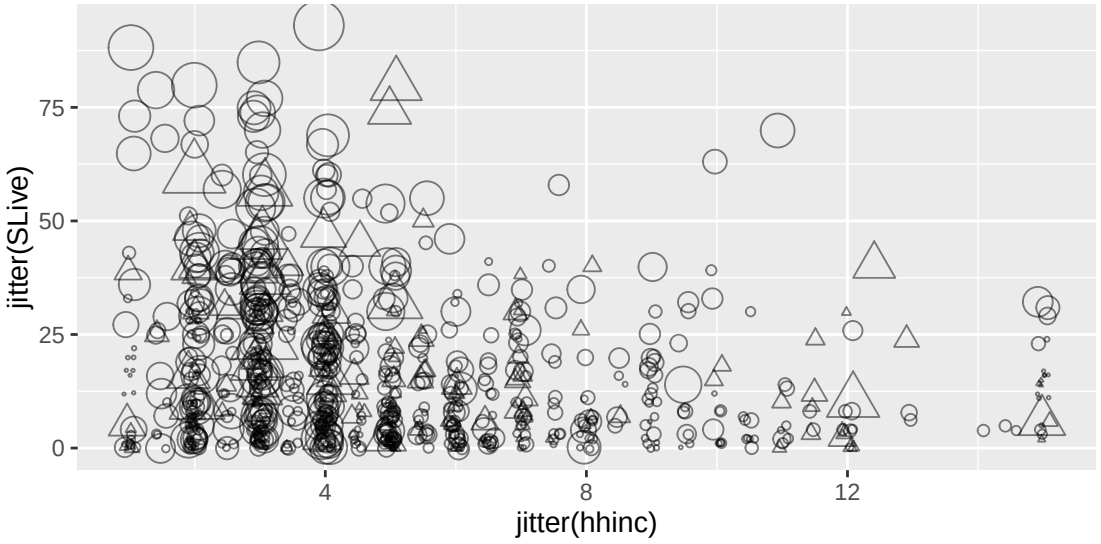
```
# Or perhaps easier to distinguish
ggplot(ho05, aes(jitter(hhinc), jitter(SLive), color = as.factor(postbomb))) +
    geom_point(size = 3, alpha = 0.5)
```




```
# But this would be bad
ggplot(ho05, aes(jitter(hhinc), jitter(SLive), color = as.factor(Rage))) + geom_point(size = 3,
    alpha = 0.5)
```




```
# Maybe a little better, but at some point your graphic is not informative.
# Choose wisely
ggplot(ho05, aes(jitter(hhinc), jitter(SLive))) + geom_point(size = (ho05$Rage^2)/1000,
    alpha = 0.5, shape = as.factor(ho05$postbomb))
```



3 Regression already?

Well, you'll definitely see more of it but if you're having trouble right now remember that both Kaplan and the James et al readings have a lot of material have a lot of examples with code *and* interpretation. Don't negelct the latter.