Stat 5014 HW4

Bob Settlage

2017-09-25

Problem 3

Roger Peng's book is a nice summary. According to Roger, EDA has as it's focus:

Problem 4

First, we will read in and create a tidy dataset. After tidying, a summary is in Table 1 with a boxplot in Figure 1. I will put this code in an Appendix.

Table 1: Datasaurus data summary

block	mean_phos	mean_depth	block_cor	observations
Min. : 1	Min. :47.8	Min. :54.3	Min. :-0.0694	Min. :142
1st Qu.: 4	1st Qu.:47.8	1st Qu.:54.3	1st Qu.:-0.0686	1st Qu.:142
Median : 7	Median :47.8	Median :54.3	Median :-0.0666	Median :142
Mean : 7	Mean :47.8	Mean :54.3	Mean :-0.0660	Mean :142
3rd Qu.:10	3rd Qu.:47.8	3rd Qu.:54.3	3rd Qu.:-0.0641	3rd Qu.:142
Max. :13	Max. :47.8	Max. :54.3	Max. :-0.0603	Max. :142

[your caption here][your caption here]: datasaurus.gif

Appendix 1: R code

```
summary_by_block <- HW4_tidy %>% group_by(block) %>% summarize(mean_phos = mean(phosphate),
    mean_depth = mean(depth), block_cor = cor(phosphate,
        depth), observations = n())
knitr::kable(summary(summary_by_block), caption = "Datasaurus data summary")
```

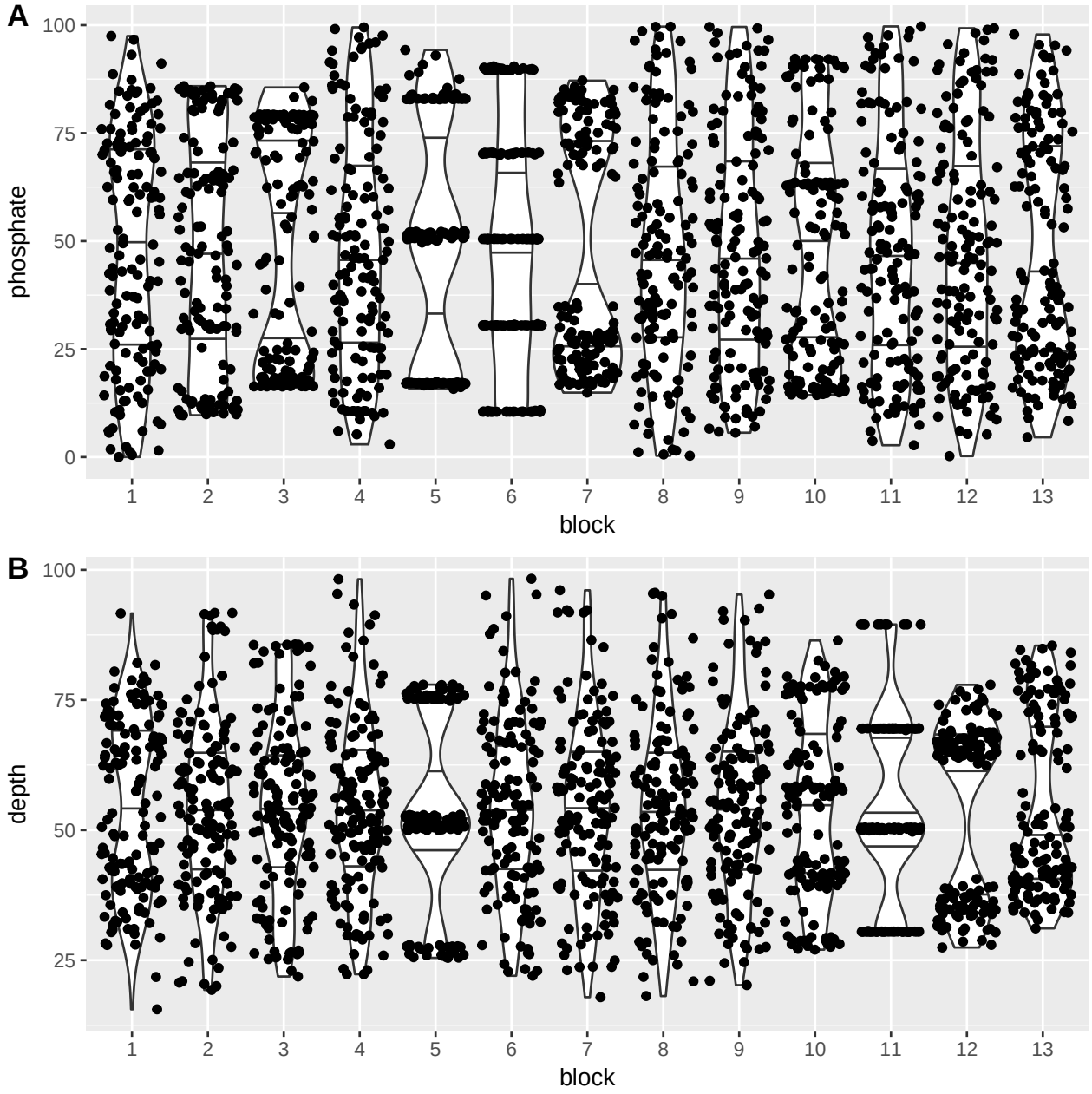

Figure 1: Datasauras data as violins.