

Part I. (10 points) Feedback for homework.

(10^{pts}) **1. Score distribution and comments.**

Solution: Numerical and visual summaries of the crowdgrader.org scores are below. Four variables are of interest:

Submission.Grade Your grade from how others graded your assignment submission.

Reviews.Completed The number of reviews you completed of others in the class (should be 5).

Review.Grade Your grade of how well you rated others based on how consistent your reviews were of others' reviews of other peoples' assignments.

Crowd.Grade Your final combined grade for the assignment ($0.75 * \text{Submission.Grade} + 0.25 * \text{Review.Grade}$).

(Final.Grade) This is a field I'm not using; it's a manual option if I wanted to assign a grade.

```
fn <- "data/ADA1_01a_Medical_records.csv"
dat <- read.csv(fn)
colnames(dat)

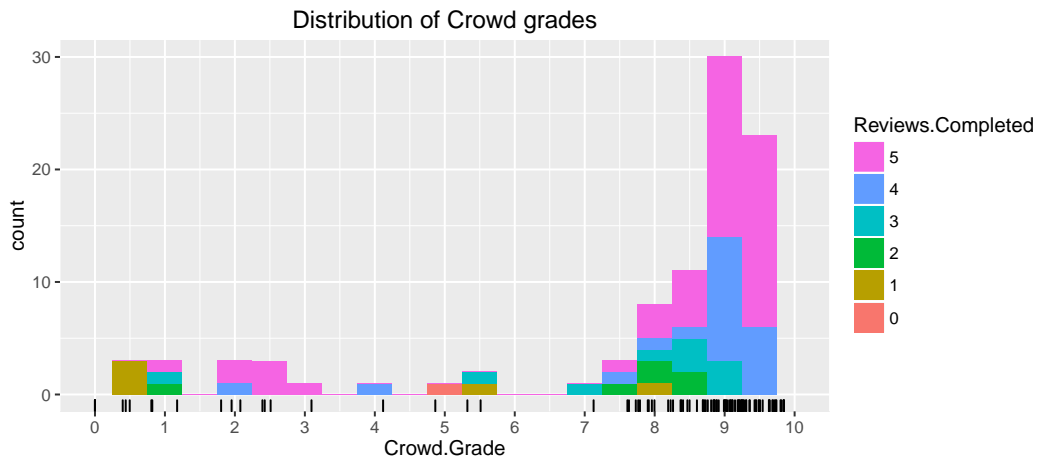
## [1] "Student"           "First.Name"       "Last.Name"
## [4] "Submission.Grade" "Reviews.Completed" "Review.Grade"
## [7] "Crowd.Grade"      "Final.Grade"

dat$Reviews.Completed <- factor(dat$Reviews.Completed)
total.pts <- 10
```

```
# numerical summary
summary(dat$Crowd.Grade)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000  7.756   8.887   7.467   9.284   9.849
```

```
library(ggplot2)
p <- ggplot(dat, aes(x = Crowd.Grade, fill = Reviews.Completed))
#p <- p + geom_vline(xintercept=c(total.pts * seq(0, 1, 0.1)), colour="red", alpha=1/2)
p <- p + geom_histogram(binwidth=total.pts/100*5)
p <- p + geom_rug()
p <- p + scale_x_continuous(limits = c(0, total.pts), breaks=seq(0, total.pts, 1))
p <- p + guides(fill = guide_legend(reverse=TRUE))
p <- p + labs(title = "Distribution of Crowd grades")
print(p)
```



```

library(ggplot2)
p <- ggplot(dat, aes(x = Submission.Grade, y = Review.Grade, colour = Reviews.Completed))
p <- p + geom_abline(intercept = 0, slope = 1, linetype = "dotted")#, colour = "gray75")
p <- p + geom_point()
p <- p + geom_rug()
p <- p + scale_y_continuous(limits = c(0, total.pts), breaks=seq(0, total.pts, 1))
p <- p + scale_x_continuous(limits = c(0, total.pts), breaks=seq(0, total.pts, 1))
p <- p + guides(fill = guide_legend(reverse=TRUE), colour = guide_legend(reverse=TRUE))
p <- p + coord_fixed(ratio = 1)
p <- p + labs(title = "Bivariate distribution of Submission and Review grades")
print(p)
## Warning: Removed 6 rows containing missing values (geom.point).

```

