BIOS 6621 HW Report

Yizhou Fei

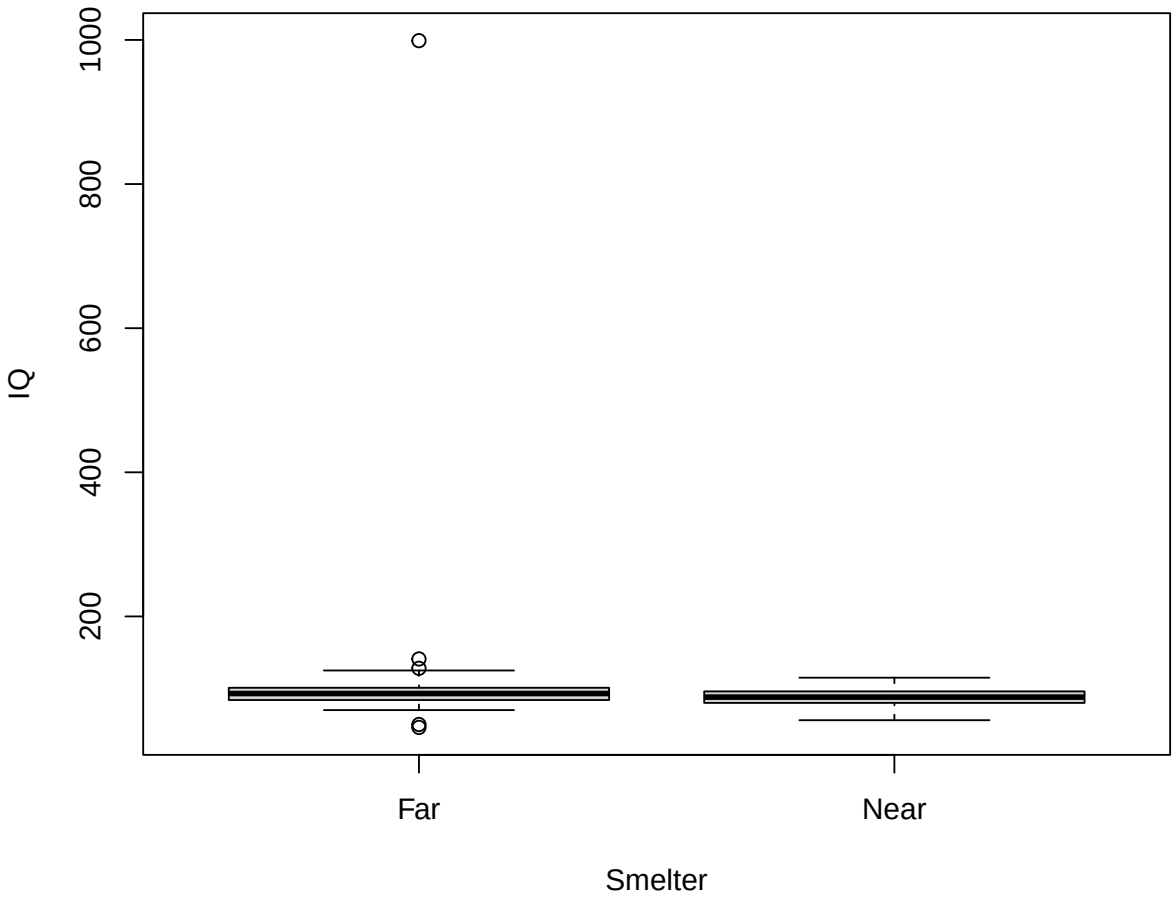
Contents

QA. A graph showing the IQ levels by location status

```
IQ <- read.csv(file = 'D:/yizhou/CU/BIOS 6621/W5/lead-iq-01.csv')
head(IQ)
```

```
##   Smelter IQ
## 1     Far 70
## 2     Far 85
## 3     Far 86
## 4     Far 76
## 5     Far 96
## 6     Far 94
```

```
boxplot(IQ~Smelter,data=IQ)
```



QB. At least one nicely formatted table

```
library(knitr)
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.1.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 4.1.3
```

```
summary <-IQ  %>% group_by(Smelter) %>%
  summarise_at(vars(IQ), list(Mean = mean,Sd = sd, Median = median, Min = min, Max = max))
kable(summary,digits = 2, caption = "Summary statistics of IQ by Location")
```

Table 1: Summary statistics of IQ by Location

Smelter	Mean	Sd	Median	Min	Max
Far	106.12	111.88	93	46	999
Near	89.19	12.17	88	56	115

QC. A couple sentences of text (not on the graph, but in the body of the Rmarkdown document) describing the graph and the table

The boxplot shows an outlier in group Far. As shown in the summary table, the mean in Far group is larger than Near group and the variability is also much larger.

QD. In-line calculations giving the values of the means, that will be updated if the data are changed, and

```
mean <- mean(IQ$IQ)
print(mean)
```

```
## [1] 98.33871
```