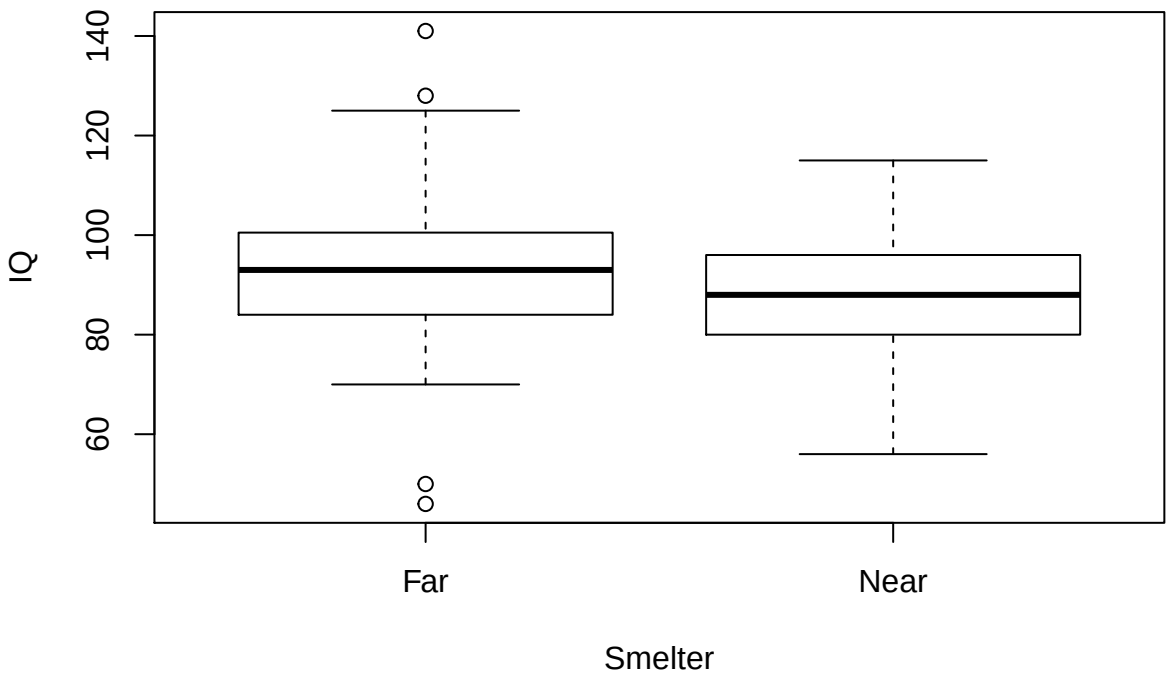
Lead IQ data set description

Jiayuan Shi

2022-10-07

```
library(knitr)
dataset=read.csv('C:/UCHealth/FALL 2022/BIOS 6621 Statistical Consulting I/Week 6/lead-iq-01.csv')
dataset$IQ[which(dataset$IQ == 999)] <- 99
boxplot(IQ ~ Smelter, data = dataset)
```




```
kable(table(dataset$IQ, dataset$Smelter))
```

	Far	Near
46	1	0
50	1	0
56	0	1

	Far	Near
70	1	0
71	0	1
72	2	0
73	0	1
74	0	1
75	1	2
76	3	3
77	1	2
78	1	1
79	2	0
80	2	4
82	1	1
83	1	1
84	0	1
85	3	4
86	3	2
87	1	1
88	1	5
89	2	2
90	1	0
91	1	3
92	3	1
93	2	1
94	3	1
95	0	1
96	4	4
97	3	0
98	1	1
99	3	1
100	2	0
101	3	1
102	1	0
104	2	2
105	1	1
106	0	1
107	2	2
108	1	0
111	1	1
112	0	1
114	0	1
115	1	1
118	1	0
120	1	0
125	1	0
128	1	0
141	1	0

```
mean(dataset$IQ)
```

```
## [1] 91.08065
```

The boxplot shows the IQ levels by location status. We can see that the median of IQ scores in the Far

group is higher than that in the Near group. The table displays the frequency of the IQ levels by location status. It seems subjects with the highest 5 IQ scores (118-141) in the dataset are in the Far group, i.e., live far from the 1 mile of the lead smelter.