HW - Malik (Reproducible Reporting)

Laavanya Joshi Malik

2023-10-10

**Introduction**

For my reproducible reporting project, we use the Lead-IQ data set which has been derived as a result of the original study published, Neuro-Psychological dysfunction in children with chronic low-level lead absorption in 1975 by Landrigan PJ, Baloh RW, Barthelt WF, Whitworth RH, Staehlinh NW, and Rosenblum BF.

**Uploading data set and creating data frame**

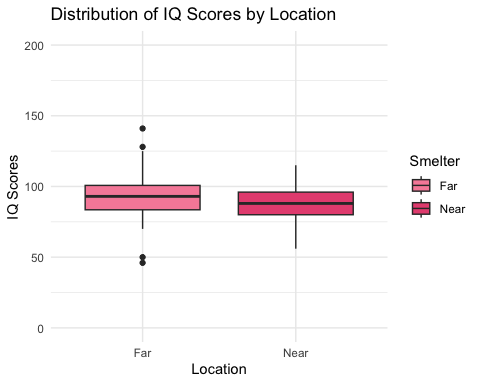
I decided to utilize tidyverse library to upload and create my data frame. This organization will help me with my exploratory analysis.

file\_path <- "/Users/Laavanya/Malik\_Project\_01/DataRaw/lead-iq-01.csv"  
lead <- read\_csv(file\_path)  
  
#head(lead)  
#tail(lead)

**For part A:**

Graph showcases mean IQ levels by location status:

ggplot(lead, aes(x = Smelter, y = IQ, fill = Smelter)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("Far" = "#F78DA7", "Near" = "#E75480")) +  
 labs(title = "Distribution of IQ Scores by Location",  
 x = "Location",  
 y = "IQ Scores") +  
 ylim(0, 200) +  
 theme\_minimal()



**For part B:**

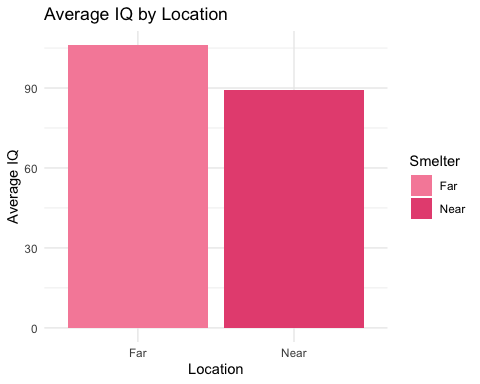
lead\_table <- head(lead)  
  
knitr::kable(lead\_table, "simple")

| Smelter | IQ |
| --- | --- |
| Far | 70 |
| Far | 85 |
| Far | 86 |
| Far | 76 |
| Far | 96 |
| Far | 94 |

After graphing a box plot displaying the ID scores

**For part C:**

lead %>%  
 group\_by(Smelter) %>%  
 summarise(MeanIQ = mean(IQ)) %>%  
 ggplot(aes(x = Smelter, y = MeanIQ, fill = Smelter)) +  
 geom\_bar(stat = "identity") +  
 scale\_fill\_manual(values=c("Far" = "#F78DA7", "Near" = "#E75480")) +  
 labs(title = "Average IQ by Location",  
 x = "Location",  
 y = "Average IQ") +  
 theme\_minimal()



**For part D:**

**For part E:**