A screenshot of a cell phone

Description automatically generated

1. How many individuals are in the data set?

124 individuals.

2. How many variables are in this data set?

40 variables.

3. Can you tell if any of the variables are categorical (i.e. qualitative)? Identify specific ones.

In this data frame, such as the first column called “ID” is a categorical variable because ID is fixed, each row has its ID respectively. We cannot see the same ID in this data frame. Also, although ID is a number, it cannot be calculated so it is not a qualitative variable.

Two important variables that were studied were (1) MAXFT = the number of finger-wrist taps in the dominant hand (a measure of neurological function) and (2) IQF = the Wechsler full-scale IQ score. You will explore the relationship of lead exposure to one of these two outcome variables.

4. Is this an observational study or a randomized experiment? Explain why.

5. How many individuals have MAXFT scores measured? How many have IQF scores measured?

99 individuals have MAXFT scores measured, 124 individuals have IQF scores measured.

6. Pick one of MAXFT or IQF of interest to you. We are primarily interested in comparing the

distribution of the outcome of interest (MAXFT or IQF) for the two different groups of children (GROUPS 1 and 2, those children with elevated blood-lead levels > 40 g/ml and those with lower levels, < 40 g/ml, respectively.)

a. What are the mean and median of the outcome of interest (MAXFT or IQF) for each GROUP?

Mean: Group 1 of MAXFT is 47.42857

Median: Group 1 of MAXFT is 48

Mean: Group 2 of MAXFT is 54.4375

Median: Group 2 of MAXFT is 53.5

b. Describe the shape of the distribution (i.e. histogram) of the outcome for each GROUP.

c. What information can we get from the Boxplot of the outcome for each GROUP?

d. Based on these summaries, what is your assessment of the differences between the two groups

of children on the outcome of interest? Discuss the role of randomization in this study.

References:

• Rosner B, (2005) Fundamentals of Biostatistics, Thomson.

• Landrigan PJ, Whitworth RH, Baloh RW, Staehling NW, Barthel WF, Rosenblum BF (1975)

Neuropsychological dysfunction in children with chronic low-level lead absorption. Lancet, 1, 708-715.