

Final RQ

Group_B_158

DS082: water_portability.csv

Data set ID: DS082

Sample of our Dataset:

Variables used in our RQ:

Dependent variable: Solids

Independent variables: Portability

| | ph | Hardness | Solids | Chloramines | Potability |
|---|-----------|----------|----------|-------------|------------|
| 1 | NA | 204.8905 | 20792.32 | 7.300212 | 0 |
| 2 | 3.716080 | 129.4229 | 18630.06 | 6.635246 | 0 |
| 3 | 8.099124 | 224.2363 | 19909.53 | 9.275884 | 0 |
| 4 | 8.316766 | 214.3734 | 22018.42 | 8.059332 | 0 |
| 5 | 9.0922233 | 181.1015 | 17978.99 | 6.546600 | 0 |

Number of rows = 3276

Our Research Question is: “Is there a difference in mean Total Dissolved Solids (TDS) between drinkable and non-drinkable water samples?”

Null Hypothesis H0: There is no difference in the mean Solids between drinkable and non-drinkable water samples.

Alternate hypothesis H1: There is a difference in the mean Solids between drinkable and non-drinkable water samples.