

1. Load the “polls_us_election_2016” dataset using a relevant command and save it as polls.

Note: This dataset belongs to the library package dslabs

2. Use appropriate R commands to display the summary statistics of the dataset.
3. Use R command to extract the following columns as req_details:
4. Adjpoll_trump, adjpoll_clinton, grade pollster and state.
5. Find the different levels of the factor variables in req_details.
6. Check if any columns in the dataframe polls have missing values.

Hint: is.na() provides values that are NA

7. If there are missing values in numeric columns then extract the column to a dummy variable “columnname.miss”. Convert this variable to a matrix and replace the missing values with mean of the respective columns of the matrix.
8. Extract the name of the states with sample size greater than 1250 and adjpoll_trump is between the values 40.0 and 43.0.
9. Extract the name of the states where clinton has got the adjpoll greater than 42.
10. Extract the name of the states where trump has got the adjpoll greater than 42.
11. What is the average adjpoll of each contestant?