1. **How many of these bonds were approved by voters, and how many were defeated? Are there any differences in the rates of approved bonds across the four different government types? Calculate the appropriate descriptive statistics to answer these questions.**

   7210 bonds were approved by voters and 1638 bonds were defeated.

   When comparing the results with proportions, we can see that overall, more bonds were approved than they were defeated. Across all governments, the 250 COUNTY governments were the ones to approve and defeat the least amount of bonds, coming in at 207 (82.8%) approved and 43 (17.2%) defeated. On the other side, the 4531 ISD governments were the ones to approve and defeat the most amount of bonds, coming in at 3282 (~72.4%) approved and 1249 (~27.6%) defeated. Both the CITY and WD governments were in similar situations, with the 1767 CITY governments being ~87.5% approved and ~12.5% defeated and 2300 WD governments being ~94.5% approved (The highest approval rate of all four types) and ~5.5% defeated. It's interesting to note that even though the WD governments have a lower amount of bond election counts than ISD governments, the WDs had the highest approval rate and the ISDs had the lowest approval rate.

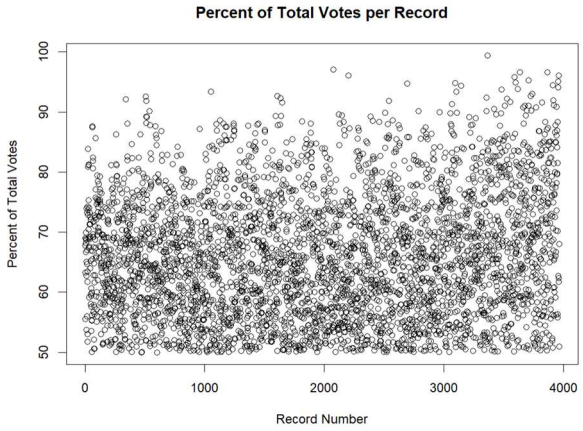
2. **Some of these bonds were on the ballot during presidential elections and therefore had very high voter turnout. Calculate a new variable in the dataframe called "Votes_Total" that is the sum of the votes "for" and "against" the bond measure. When and where did the bond measure with the highest voter turnout occur? What was it for?**

   The highest voter turnout occurred in Harris County on the 8th of September 2022. With the category being Road Utilities, the total votes placed during the bond election was 1,030,414. Of the total votes placed, 711,352 votes (69.04%) were for the bond and 319,062 votes (30.96%) were against the bond.

CONTINUE TO NEXT PAGE

3. **Let's look at the margins by which the carried bonds were approved, ignoring those with very low voter turnout. Create a subset of this dataset that contains the bond measures that were approved <u>and</u> had at least 100 total votes. Next, create a new variable within the subset dataframe that gives the percentage of total votes that were for the bond measure and make a graph of the distribution of this new variable. Describe its distribution with the appropriate statistics.**

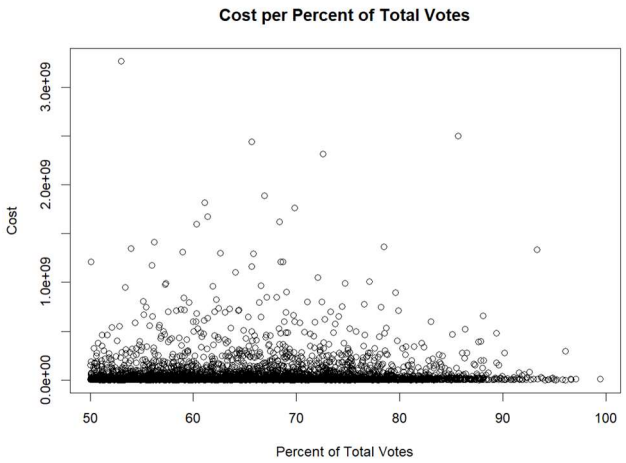
   The image of the graph is below:



   This relationship shows (per bond election) how much of the votes from each bond election where the bond won were votes for the bond. We can see that the scatterplot is very distributed, with nearly every percentage having an equal number of records. The minimum percentage leading a bond to be approved was 50.00404% and the maximum percentage was 99.44444%. The standard deviation of the graph is 10.10567% (which is very large) and the average percentage of votes for the bond in all these bond elections was 66.03232%.

4. **Is the margin a bond was approved by related to its cost? Use your subset from #3 to create a graph to display this relationship. Then, answer this question, citing the appropriate descriptive statistic.**

   The image of the graph is below:



   This graph shows the relationship between the cost of the bond and the percentage of total votes the bond was approved by. We can see there is a huge cluster closer to the cost of the bond, but the actual distribution of the percentages per bond cost is very distributed. In fact, the correlation between the cost and the percentage of total votes the bond was approved by is 0.000559102, which is nearly no correlation at all.

   So, in summary, the cost of the bond made no real impact on how large of a percentage the bond election was won by (aka there is no correlation between them).