Simon Gaul
Jake Zenk
Graeme St. Clair

Project 1

For this assignment, we are going to be a looking at a data set that looks at bank closings in the United States. The data set includes: which bank closed, the date the bank closed, the institution that acquired the bank, and also the city and state that the bank closed in. We will be pulling data from this data set and manipulating it to answer questions that we have and also arrange the data in a way that emphasizes the value we want to show.

Questions to be asked:

1. **Count per year?**
   ○ For this we will be looking at the closing dates of the bank and we will be counting how many banks were closed each year
   ○ This one will be a little more difficult because the closing date gives the month and day as well as the year, so we will have to take the year value from the date in order to get a count
   ○ After we get this count we will once again be able to sort based on the count and have the highest value at the top of the list, which gives us the year with the most bank closings
2. **Most popular state?**
   ○ For this deliverable, we are going to load our data into a Data Frame and then get the count for each state in the data sheet
   ○ We can then use this count and sort based on the highest count to the lowest count, which ultimately will put the most popular state for bank closings to be at the top of the list
3. **Compare states to see missing ones**
   ● This last deliverable will be a little more difficult. Instead of finding a count we will be comparing the list of states to a complete list
   ● It includes finding all unique values in the 'ST' column and comparing them to a second list we creating containing the 50 unique state values.
   ● Any value that is listed in both lists will be removed from the second list and when it's complete only the states not included in our data set will remain.
4. **Most popular city?**
   ○ This second deliverable will be a lot like our first only this time we will be finding the count for the cities instead of the state
   ○ We will use the data to get a count for the cities and then we can use the same sorting method that we used for the first one in order to get the highest value at the top of the list
5. **Most popular acquirer?**

- This deliverable will be finding the count of who acquired the most banks in our data set
- We will be counting the number of times each acquirer acquired a bank and then we will then be able to sort those based on their count
- Sorting based on count will give us the acquirer with the most amount of acquires at the top of the list