| CSC 124 Project 1 | |
| --- | --- |
| Project Number | Project Number 1 |
| Project Name | Dataframes |
| Project Filenames | CSC-124\_Project\_1.docx proj1\_data.csv |
| Points | 100 points |
| Assistance | This is an individual assignment. You should not request or provide assistance from/to others. |
| Project Description | Students will use the concepts learned in this course to analyze data about COVID-19. Be sure you have reviewed the material in Brightspace in Study Unit 3 about filter(), install.packages() and library(). |
| Tasks | 1. **Read the proj1\_data.csv file into R. You can use read.csv() to do this.  Double check to make sure there is no garbage in the column names. You can use** colnames(myDataFrame) **to do this. If there is garbage, you can clean it up as follows:** colnames(myDataFrame) <- c('date', 'state', 'fips', 'cases', 'deaths') 2. **Clean the datset by removing all rows before March 1, 2020 and after December 31, 2021. You may notice some days with negative cases or deaths. What might account for this?** 3. **Remove the "fips" column.** 4. **What is the average daily case count?** 5. **What is the average daily case count for 2020? For 2021? You may need to create additional dataframes to do this.** 6. **What was the highest number of cases in a day, and what was the date and state associate with it? (Note: if you want to sort the dataframe in order to do this, it's okay.)** 7. **Create a new dataframe that only includes data from North Carolina.** 8. **Repeat Step 5 for the NC dataframe** 9. **Repeat Step 6 for the NC dataframe** 10. **For 2020 and 2021, what is the overall mortality rate for COVID-19? (deaths divided by cases)** 11. **If you wanted to determine how widespread cases are in a particular state, what additional information would you want, and why?** |
| Deliverable 1 of 2 | **A Notepad file:**  The file should include only the R code to get the answers to items 1 through 10. |
| Deliverable 2 of 2 | **A Document (docx or pdf)** which contains the answers to items 2 through 11, excluding #3. |
| Upload Instructions | Upload to Brightspace. |