E.J OURSO COLLEGE OF BUSINESS

ISDS DEPARTMENT
Nicholson Extension
Baton Rouge, LA 70803
USA

ISDS 3105

Spring 2018 – Mid Term Exam

This exam is exclusively focused on practical skills in R. The allocated time is **80 minutes**. You will need to return this document with your name and your signature. You will also need to **turn the Rmarkdown documen**t of the responses.

The exam is to be completed using RStudio and you must upload the completed file into the folder called "Mid Term Exam Dropbox" available on Moodle. It is your responsibility to correctly upload the correct version of the exam file. You should not leave the room until you have verified that the correct file is stored on Moodle.

You can use all the facilities of Rstudio, including the help functionality. **You cannot use any other software or external aid** (e.g., Google or any other website) to complete the exam. You cannot communicate, verbally or through electronic means, with anyone during the exam.

You cannot open any other file other than the exam Rmarkdown file and HTML document. You are not allowed to copy and paste any code from documents stored on your computer.

Any behavior not conforming to the above directions is considered cheating and will be reported.

Name _____

PAWS ID _____

By signing my name, I certify that I have not received nor given any help during the test. I further certify that I did not refer to any material or text during the test other than the built-in help in RStudio.

Signature: _____

The dataset **dt** includes crimes reported in Baton Rouge and handled by the Baton Rouge Police Department (source: https://data.brla.gov/):

filenumber	Crime report identifier
offensedate	Date of the crime
offensetime	Time of the crime
committed	Status of the crime
zip	Crime location – zip number
district	Crime location – district number
zone	Crime location – zone number
subzone	Crime location – subzone number
offensetime_discrete	Part of the day (discretized offensetime)
crimeId	Crime type identifier
week	Week of the crime (1 to 52, starting Jan 1st)
month	Month of the crime (1 to 12)
day	Day of the crime (1 to 31)

The object **lkt** includes crime names and crime codes for the crimes in **dt**:

crime	Crime type name
crimeId	Crime type identifier

1. Create a new folder on your desktop and name it *MidTerm*. Download *dataset.RData* from Moodle and save it in the *MidTerm* folder.

2. Create a new RStudio project in the *MidTerm* folder.

3. Create a new RMarkdown file and save it using your LSU PAWS id (e.g., gpiccoli, tsmith34, bgreen21). Select HTML as the default output option.

4. Change the YAML header to reflect:

   a. "ISDS 3105 Exam" as the title
   b. Your full name as the author
   c. Today's date as the date

5. Delete the default narrative and sample code chunks. Unless the question explicitly tells you to use inline code, insert a **new chunk for each question**.

6. Create the first chunk. In it, load the library *tidiyverse* and *dataset.RData*. Mute any warnings or messages for this chunk.

7. Create a header titled: **Introduction**

8. Write a short but meaningful introduction to your report, including at least one:

   a. bold text

   b. italics text

9. Add the sentence: "There are a total of X reported crimes and Y crime types in this dataset." Substitute X and Y with the appropriate values computed using in-line code.

10. Add the sentence: "The dataset contains records from X to Y". Substitute X and Y with the appropriate max/min values from the column *year* using in-line code.

11. Create a header titled: **Data Analysis**

12. From *dt*, filter out any rows that have missing values on the variable *year*. Save the output to *dt*.

13. Use the appropriate join function to merge *dt* and *lkt* on crimeId. Save the output to *dt*.

14. From *dt*, drop the variables *district, zone, subzone*. Save the output to *dt*.

15. Create a table showing the count of attempted and committed crimes for each year. Save the output into a new object called *yearCount*. Make sure you show the table in the output.

16. Add a variable *percent* to *yearCount* showing the percentage of committed and attempted crimes **by year.** Manipulate *percent* as needed to format the percentage from 1-100 and adding the "%" sign.

17. Add the sentence: "The part of the day with the most thefts is X". Substitute X with the appropriate value calculated on *dt*.

18. Create a table with the 5 most dangerous days in 2017 (those with the most <u>committed</u> crimes).

19. From dt, create a subset of homicides (both attempted and committed) which occurred in 2016 and 2017 only. Assign the output to *recentHomicides*.

20. Create a header titled: **Data Visualization**

21. Use *dt* to plot a barchat of the count of crimes by month. Add a title using the appropriate function.

22. Use *yearCount* to plot a barchart of the count of crimes by year. Map attempted/commit to the fill of each bar.

23. Use *dt* to plot a barchart of the number of crimes by *offensetime_discrete*, facetting by year.

24. The media stressed the sharp increase in homicide rates in 2017 compared to 2016. Plot two line charts with the monthly counts of homicides in 2016 and 2017. Use faceting on the variable *committed* to separate attempted/committed homicides. Each facet should have two lines (one for each year). In your narrative, comment on whether the media are right.

25. Once you have completed your work save the .Rmd file and knit to HTML.

26. Login to Moodle and load the .Rmd file into the folder called "Mid Term Exam Dropbox."

USEFUL FUNCTIONS

Base R	Operators	dplyr	ggplot	tidyr
<- as_* (as_integer, as_double, etc.) factor max mean min names paste paste0 sort sum unname unique which	Extracting/selecting: [; [[; $ Logical Operators: <; >; ==;!=; !, %in%, &, |, <=, >= Mathematical operators *; /; +; -	arrange between count desc distinct filter group_by inner_join left_join mutate n pull right_join slice select summarise top_n ungroup	aes facet_wrap ggplot geom_bar geom_col geom_line geom_point ggtitle theme	gather spread

Question	Points	Score
3	1	
4	1	
5	1	
6	4	
7	1	
8	3	
9	7	
10	9	
11	1	
12	7	
13	7	
14	6	
15	8	
16	8	
17	6	
18	7	
19	5	
20	1	
21	6	
22	6	
23	7	
24	8	
Total	**110**	