



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 90 minutes. You will need to return this document with your name and your signature. You will also need to turn in an electronic copy of the Rmarkdown document that you will create during this exam.

The exam is to be completed using Rstudio and you must upload the completed file into the folder called “Final 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 with the TA 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. All elements of the analysis below are to be completed using base R functions. You should not load any packages.

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: _____

1. Create a new project using the *exam* folder. Save it on your desktop and name it *FinalExam*.
2. Create a new R markdown file and save it using your LSU PAWS id (e.g., gpicoli, tsmith34, bgreen21). Select HTML as the default output option.
3. Change the YAML header to reflect:
 - a. “ISDS 3105 Mid-term Exam” as the title
 - b. Your full name as the author
 - c. Today’s date as the date
4. Mute code echoing for the whole document.
5. Delete the default narrative and sample code chunks. Unless the question explicitly tells you to use narratives/inline code, insert a **new chunk for each question**.
6. Load the library *tidyverse* in the first chunk. Mute any warnings or messages.
7. Load “employeesBR.Rdata”, a workspace with:
 - a. *dt*, a dataset with information about the employees of the City of Baton Rouge
 - b. *lkt*, a lookup table with the division names and division codes
8. Use the appropriate join function to merge *dt* and *lkt* on *division_num*. Save the output to *dt*.
9. Create a header titled “Introduction” and write a short but meaningful introduction to your report, using inline code to show the total number of observations in *dt*. In this narrative use at least once:
 - a. bold text
 - b. italics text.
10. Create a header titled: **Data Inspection**.
11. Add the following sentence to the above narrative: “There are a total of X departments in this dataset.” Substitute X with the appropriate value computed dynamically with in-line code.
12. Create a table for the count of employees in each department. Save this object as *employeesCount*. Make sure the table is visible when you knit the report.
13. Write some narrative and use in-line code to list the unique names of the departments in *employeesCount*. Order the names by count of employees (descending).
14. Add the sentence “The total number of observations for MOSQUITO & RODENT CONTROL is X”, where X is the number of observations calculated using inline code.



15. Create a header titled: **Data Preparation.**
16. From *dt*, **filter out** the employees for the department COUNCIL ADMINISTRATOR. Save the output to *dt*.
17. The variable *end_date* indicates the employee's retirement date, and is equal to NA if the employee did not retire. Use *end_date* to create a new variable *retirement* that is FALSE when *end_date* is NA, and TRUE otherwise. Then, use *retirement* to count the total employees who retired/did not retired by department name (using *department_name*).
18. Arrange the *depart_name* in the table from 17 by number of retirements (descending). Save the table into an object *retir*.
19. Add a column to *retir* showing the percent of retirements by department
20. Add the sentence "A total of X employees retired, while Y did not". Calculate X and Y using inline code
21. Create a header titled: **Data Visualization.**
22. Create a bar chart to plot the count of retirements by department, mapping *retirement* TRUE/FALSE to the fill of each bar. Add a title and a subtitle. Rotate the x-axis' labels by 45° to make them readable (hint: you need to tweak the argument *axis.text.x* in *theme()*).
23. Plot a barchart of the average base pay (using *base_pay*) by department. In your narrative report which department pays the most. Use inline code.
24. Create a header titled: **Data Analysis.**
25. Find the 10 department with the highest average salary.
26. Calculate the average cost of overtime hours worked by employees each employee ($\text{costOvertime} = \text{overtime_hourly_rate} \times \text{total_overtime_hours}$), then use *costOvertime* to calculate the average cost of overtime by department. Report the 3 departments with the highest average cost.
27. Once you have completed your work save the .Rmd file and make sure you can render the report to HTML. Login to Moodle and load the **.Rmd** file (*not* the .html) into the folder called "Final Exam Dropbox".