Final Exam Review - R Part

1. In assessing the efficiency of a ground-loop heat pump system for heating in the winter with changing ground temperature the data from the table on the left was gathered.

Temperature (°C)	Efficiency (%)
5.6	85
4.5	87
3.4	93
1.5	95
-0.3	99
-2.3	103
-4.5	111
-8.7	210
-10.2	350
-13.5	650

   a. Create a scatterplot of efficiency v.s. ground temperature. Give your scatterplot an appropriate main title and axis labels. **Give the command(s) you used and the scatterplot as your answer to this question.**

   b. Compute the linear correlation coefficient between efficiency and ground temperature. **Give the numerical value and the command(s) you used as your answer to this question.**

   c. Determine the best fit line to the data. **Give the linear equation and the command(s) you used as your answer to this question.**

   d. Find the 95% prediction interval for the efficiency at a ground temperature of 2°C. **Give the interval output and the command(s) you used as your answer to this question.**

2. This question uses the **RailTrail** dataset, which is part of the **mosaicData** package in R.

   a. Does the set of trail use volumes appear to be symmetric, skewed left, or skewed right? Justify your answer with the appropriate command(s). **Give a sentence answer, the command(s) you used, and their output(s) as your answer to this question**.

b. Create a dataset consisting of the RailTrail data for weekdays, and another consisting of the RailTrail data for weekends. **Include only your command(s) as an answer to this question.**

c. At α=0.05, test the hypothesis that mean trail use volume on weekdays is less than the mean trail use volume on weekends. **Give your command(s), the output, your conclusion, and a sentence interpreting your output to explain how you drew your conclusion.**

3. This question uses the **crabs** dataset, which is part of the **MASS** package.

   a. Find the median frontal lobe size of the crabs. **Include both your command(s) and the median.**

   b. First, create a dataset consisting of the male crabs. Then determine the proportion of the male crabs that have frontal lobes larger than the median frontal lobe size you found in part (a)? **Include both your command(s) and the proportion.**

   c. Create side by side boxplots of the male crabs' frontal lobe size, grouped by species. Give your boxplots descriptive titles and axis labels. **Include both your command(s) and the boxplot.**

4. A standard deck of cards contains 52 cards. There are 13 ranks of card (ace, 2,3, ...,10,J,Q,K), and 12 cards (the 4 jacks, 4 queens, and 4 kings) are "face card.

   a. Write a command or a series of commands that simulates 10000 cards with replacement, and returns a bar plot that gives the frequency of each rank of cards. Don't worry about labelling your barplot accurately.

   b. Write a command or series of commands that simulates selecting 10 cards without replacement, and returns the number of face cards. Run your command(s) 10 times and include the output.

5. Batteries have lifetimes that are exponentially distributed with a mean of 50h.

   a. Estimate the probability that a battery will last longer than 20h by simulating selecting 100000 batteries and returning the proportion that last longer than 20h.

   b. A device runs on four batteries. Estimate the probability that the device will last longer than 20h by simulating running 100000 devices on four batteries each and returning the proportion of devices that last longer than 20h. (Note that all four batteries need to work in order for the device to function.)