Problem 1: (15 points)

The published figure for the median systolic blood pressure of middle-aged men is 128. To determine if there has been any change in this value, a random sample of 100 men has been selected. Test the hypothesis that the median is equal to 128, we find the following 3 scenarios.

(a) 60 men have readings above 128.
(b) 70 men have readings above 128.
(c) 80 men have readings above 128.
In each case, determine the p-value.

Problem 2: (20 points)

We would like to compare two online book sellers and see if the prices, they are charging are significantly different or not. For this we choose a small set of hypothetical best sellers for the last year. The table below shows the prices charged by the seller for each of the bestselling books. Note all data below are hypothetical.

Book Title	**Seller A**	**Seller B**
How to be successful	10.39	10.19
That path to wisdom	13.44	13.18
Capitalism	9.99	9.9
Do you know your data	9.97	9.78
Think Fast	16.25	16.17
Lean Methods	18.03	17.68
Study shows	24.08	23.97
How to be better than Big Nate	17.73	17.73
Tales from Peru	10.25	10.25
Organize your life	16.11	16.03

Is there is a systematic difference in book prices between the two online booksellers?

a) **5 points:** Formulate this question in terms of null and alternative hypotheses. (max 5 sentences)
b) **5 points:** Justify your analysis method. (max 5 sentences)
c) **5 points:** Then compute the test statistic.
d) **5 points:** What would you conclude? (max 5 sentences)

Problem 3: (20 points)

Insurance adjusters are concerned about the high estimates they are receiving from Bob's garage. To see if the estimates are unreasonably high, each of the 10 damaged cars was taken to Bob's garage and to another garage, and the estimates (in dollars) were recorded. Here are the results

Car #	**Bob's Garage**	**Other Garage**
1	1410	1250
2	1550	1300
3	1250	1250
4	1300	1200

5	900	950
6	1520	1575
7	1750	1600
8	3600	3380
9	2250	2125
10	2840	2600

You are tasked to determine if Bob's garage is charging higher.

a) **5 points:** Formulate this question in terms of null and alternative hypotheses. (Max 5 sentences)
b) **5 points:** Justify your analysis method. (Max 5 sentences)
c) **5 points:** Then compute the test statistic.
d) **5 points:** What would you conclude? (Max 5 sentences)

Problem 4: (60 points)

This problem concerns defining and solving a problem in real life setting. You will need to go through the following steps to complete this problem.

Identify a real-life problem that you would like to investigate by using techniques of non-parametric methods in module 12, 13, 14. You have complete freedom in choosing the problem. The problem could be related to what you are doing in professional life or something that excites you. Then you will form your hypotheses that you want to test. You need do this first before you go ahead and collect your data. Before collecting that data, you need to think through the process of data collection to make sure you can make good analysis. Then you will use the data you collected to perform analysis and make your conclusions. Collecting your own data for the problem you are studying is extremely important, so budget some time to finish the data collection process.

a) (10 points): Provide a clear motivation why you want to study this problem. (Max 15 sentences).
b) (5 points): Clearly state your null hypothesis and alternate hypothesis. (5 sentences).
c) (10 points): Clearly state how you collected data to test your hypothesis. Please identify clearly all the design choices you made in your data collection process. (Max 15 sentences).
d) (5 points): Tabulate the data you collected with clear explanation and annotation.
e) (10 points): Clearly state the method(s) you used in your analysis and justification (Max 10 sentences)
f) (10 points): Perform your analysis. Clearly show your code with results executed in R-notebook.
g) (10 points): State the conclusion of your analysis (Max 10 sentences).