Practice 2

Elena Tuzhilina

January 17, 2023

For the following data `score` represents the test score, `prep_time` represents the preparation time (in hours) and `attend` represents if a person attended the lectures.

score	prep_time	attend
1	0	none
5	5	some
10	10	all
9	14	all
4	3	some
7	5	all
11	14	all
8	8	all
3	6	some
2	5	none

1. Draw a histogram for `score` using cutoffs 0,3,6,9,12.
2. What is the score range?
3. What are the means and standard deviations of `score` and `prep_time`?
4. You got 13 score on the test. Find z-score.
5. If the histogram was symmetric and bell-shaped, would it be a good z-score?
6. Draw `score` vs `prep_time` scatteplot. Do you think there is a relationship?
7. What is the correlation between `score` and `prep_time`? Does this confirm your scatterplot findings.
8. Compute the distribution table (relative frequencies) for the `attend` variable.
9. Draw a stacked diagram for `attend`.
10. Find the average value of `score` and `prep_time` for each category of `attend`. (Try to interpret these values :)