p8130_hw1_yl5508

Yifei LIU

2023/9/23

Problem 1

(a) ordinal.
(b) binary.
(c) nominal.
(d) continuous.
(e) discrete.

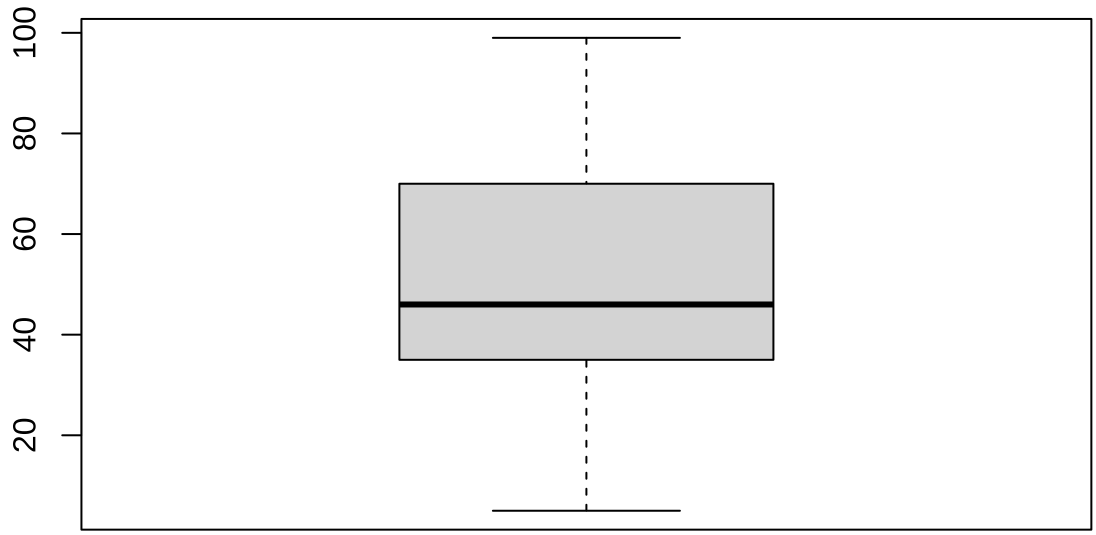
Problem 2

Problem 2.1

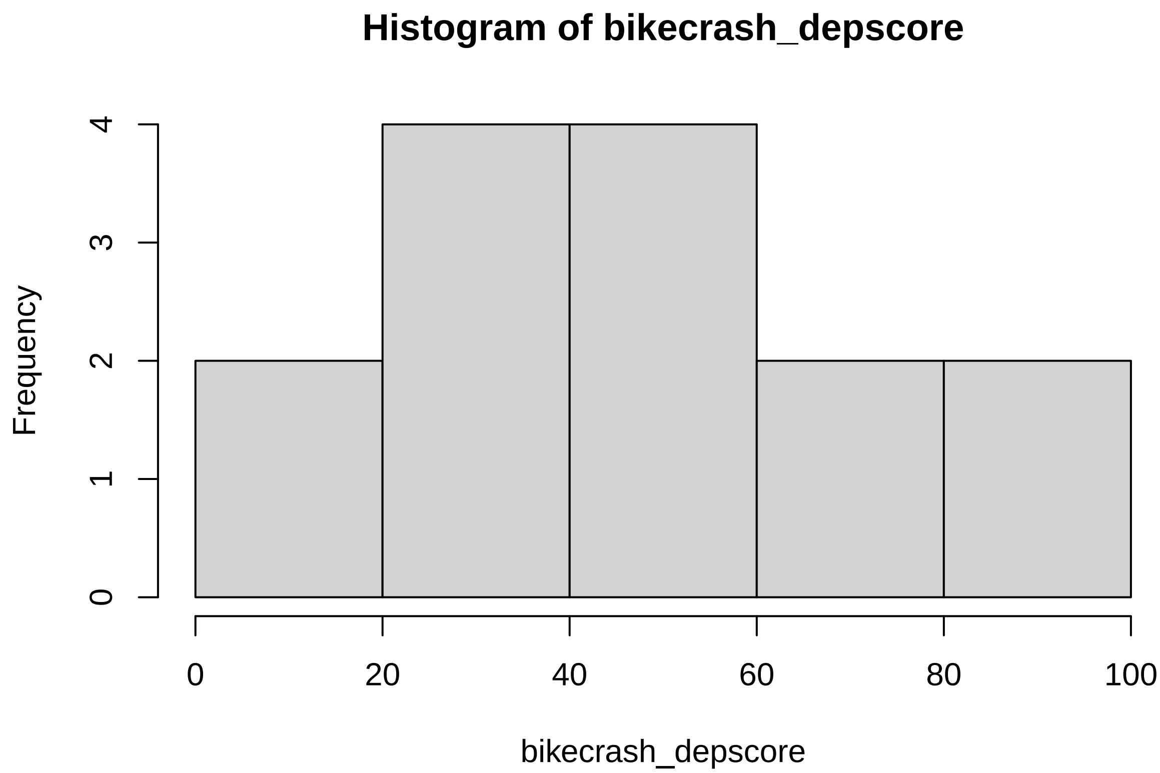
```
bikecrash_depscore = c(45, 39, 25, 47, 49, 5, 70, 99, 74, 37, 99, 35, 8, 59)
mean(bikecrash_depscore)
median(bikecrash_depscore)
range(bikecrash_depscore)
sd(bikecrash_depscore)
#Like var, sd uses denominator n-1.
```

(a) Descriptive summaries: (mean) 49.36, (median) 46, (range) 5, 99, (sd) 28.85.

```
boxplot(bikecrash_depscore)
```




```r
hist(bikecrash_depscore)
```

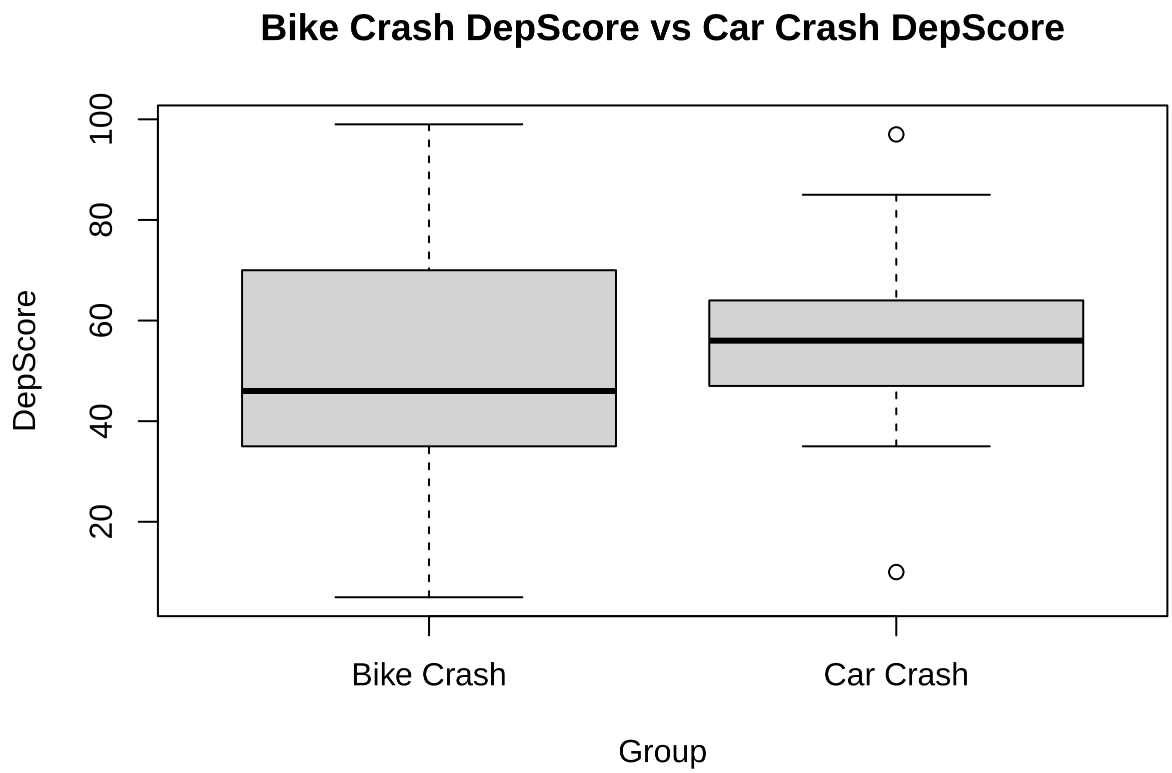


(b) We can see characteristics of underlying distribution: right-skewed, unimodal distribution.

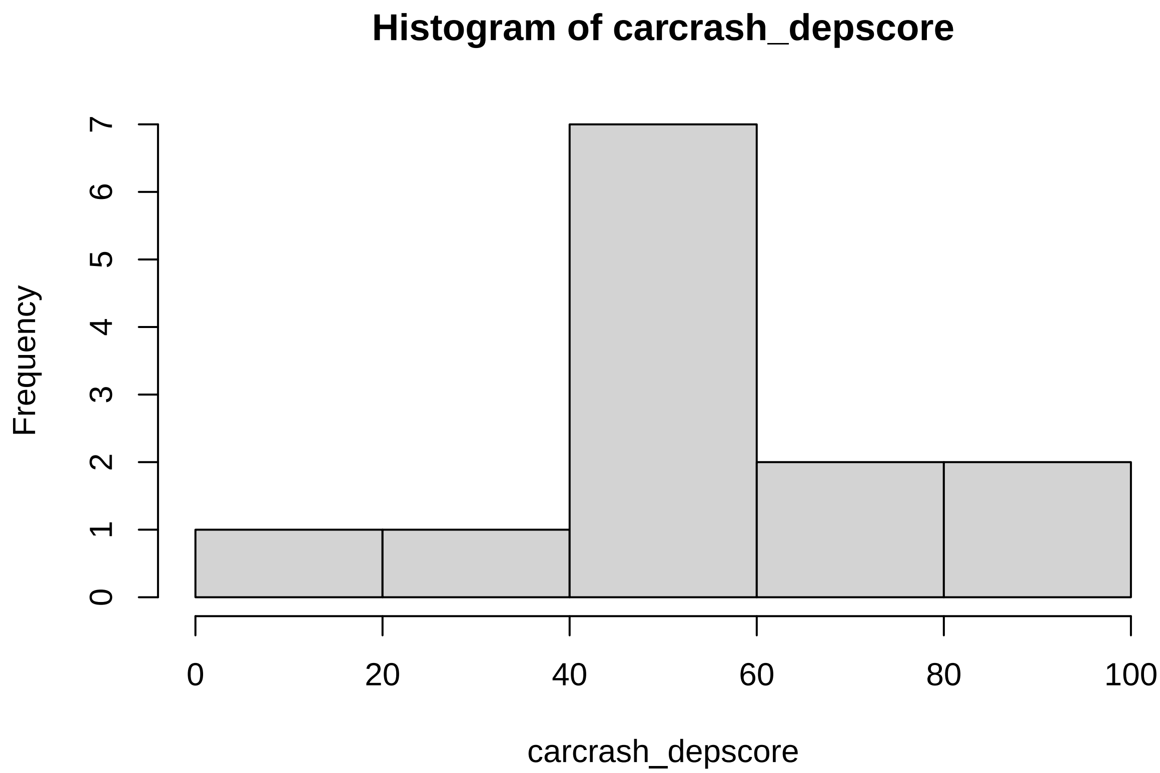
Problem 2.2

(a) side-by-side boxplot (bikecrash vs carcrash)

```
carcrash_depscore = c(67, 50, 85, 43, 64, 35, 47, 97, 58, 58, 10, 56, 50)
boxplot(bikecrash_depscore, carcrash_depscore,
        xlab = "Group", ylab = "DepScore",
        names = c("Bike Crash", "Car Crash"),
        main = "Bike Crash DepScore vs Car Crash DepScore")
```




```r
hist(carcrash_depscore)
```



(b) For car crash DepScore, the characteristic of its distribution is right-skewed and Unimodal.

(c) Group Bike Crash probably have a lower typical depression score. Cause both the median(46) and mean(49.36) of the bike crash depression score are lower than those of car crash (median(56) and mean(55.38)).

Problem 3

(a) $p(A) = 6/12 = 1/2$.
(b) $p(B) = 1/12$.
(c) $\because B \subset A \therefore p(B \cup A) = p(A) = 1/2$.
(d) $\because p(A \cup B) = p(A) + p(B) - p(A \cap B) \therefore p(A \cap B) = p(A) + p(B) - p(A \cup B) = 1/2 + 1/12 - 1/2 = 1/12$
$\because p(A \cap B) \neq p(A) \times p(B) \therefore$ event A and B are not independent$.

Problem 4

$event\ A:\ 75+\ \ y's\ woman\ has\ a\ positive\ scan\ finding$
$event\ B:\ 75+\ \ y's\ woman\ has\ dementia$

$p(B) = 0.05$ $p(A|B) = p(A \cap B)/p(B) = 0.8$ $p(A|B^{\complement}) = p(A \cap B^{\complement})/p(B^{\complement}) = 0.1$

$p(B^{\complement}) = 1 - p(B) = 0.95$ According to $Bayes'\ Rule$,
$p(B|A) = \frac{p(A|B) \times p(B)}{p(A|B) \times p(B) + p(A|B^{\complement}) \times p(B^{\complement})} = \frac{0.8 \times 0.05}{0.8 \times 0.05 + 0.1 \times 0.95} \approx 0.296$
So, the probability that the woman who has a positive scan finding may have dementia is 29.6%.