> data<- read.csv("HR\_comma\_sep.csv")

> dim(data)

[1] 14999 10

> summary(data)

satisfaction\_level last\_evaluation number\_project average\_montly\_hours

Min. :0.0900 Min. :0.3600 Min. :2.000 Min. : 96.0

1st Qu.:0.4400 1st Qu.:0.5600 1st Qu.:3.000 1st Qu.:156.0

Median :0.6400 Median :0.7200 Median :4.000 Median :200.0

Mean :0.6128 Mean :0.7161 Mean :3.803 Mean :201.1

3rd Qu.:0.8200 3rd Qu.:0.8700 3rd Qu.:5.000 3rd Qu.:245.0

Max. :1.0000 Max. :1.0000 Max. :7.000 Max. :310.0

time\_spend\_company Work\_accident left promotion\_last\_5years

Min. : 2.000 Min. :0.0000 Min. :0.0000 Min. :0.00000

1st Qu.: 3.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000

Median : 3.000 Median :0.0000 Median :0.0000 Median :0.00000

Mean : 3.498 Mean :0.1446 Mean :0.2381 Mean :0.02127

3rd Qu.: 4.000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.00000

Max. :10.000 Max. :1.0000 Max. :1.0000 Max. :1.00000

sales salary

Length:14999 Length:14999

Class :character Class :character

Mode :character Mode :character

> str(data)

'data.frame': 14999 obs. of 10 variables:

$ satisfaction\_level : num 0.38 0.8 0.11 0.72 0.37 0.41 0.1 0.92 0.89 0.42 ...

$ last\_evaluation : num 0.53 0.86 0.88 0.87 0.52 0.5 0.77 0.85 1 0.53 ...

$ number\_project : int 2 5 7 5 2 2 6 5 5 2 ...

$ average\_montly\_hours : int 157 262 272 223 159 153 247 259 224 142 ...

$ time\_spend\_company : int 3 6 4 5 3 3 4 5 5 3 ...

$ Work\_accident : int 0 0 0 0 0 0 0 0 0 0 ...

$ left : int 1 1 1 1 1 1 1 1 1 1 ...

$ promotion\_last\_5years: int 0 0 0 0 0 0 0 0 0 0 ...

$ sales : chr "sales" "sales" "sales" "sales" ...

$ salary : chr "low" "medium" "medium" "low" ...

> mean(data$satisfaction\_level)

[1] 0.6128335

> median(data$satisfaction\_level)

[1] 0.64

> mode(data$satisfaction\_level)

[1] "numeric"

> min(data$satisfaction\_level)

[1] 0.09

> max(data$satisfaction\_level)

[1] 1

> quantile(data$satisfaction\_level)

0% 25% 50% 75% 100%

0.09 0.44 0.64 0.82 1.00