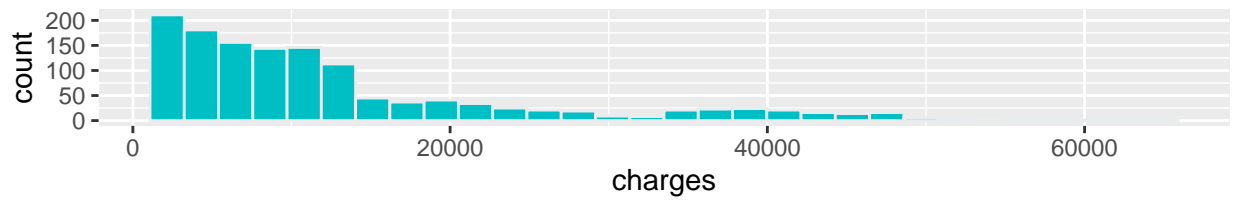
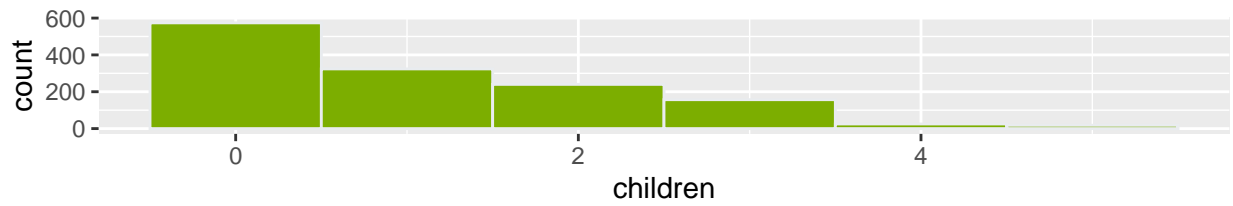
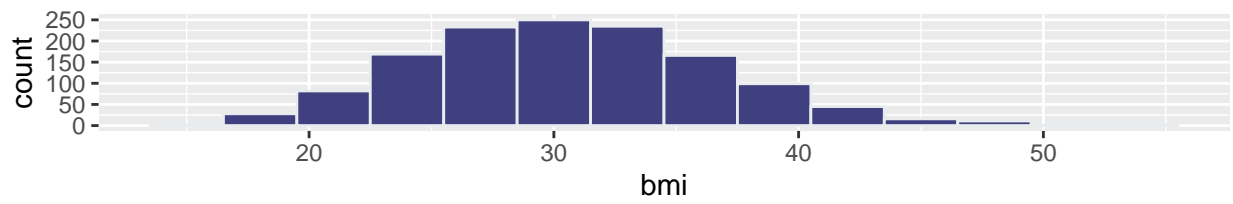
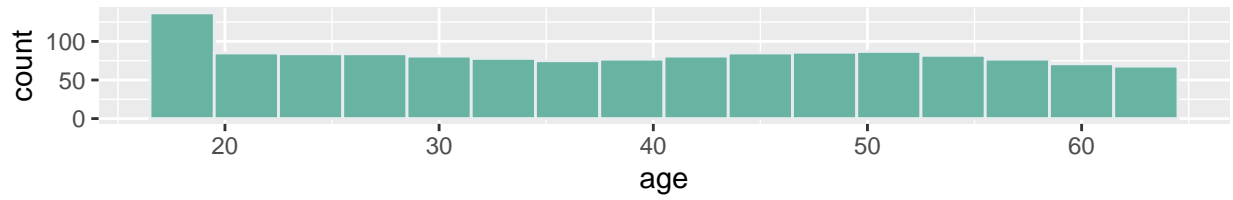
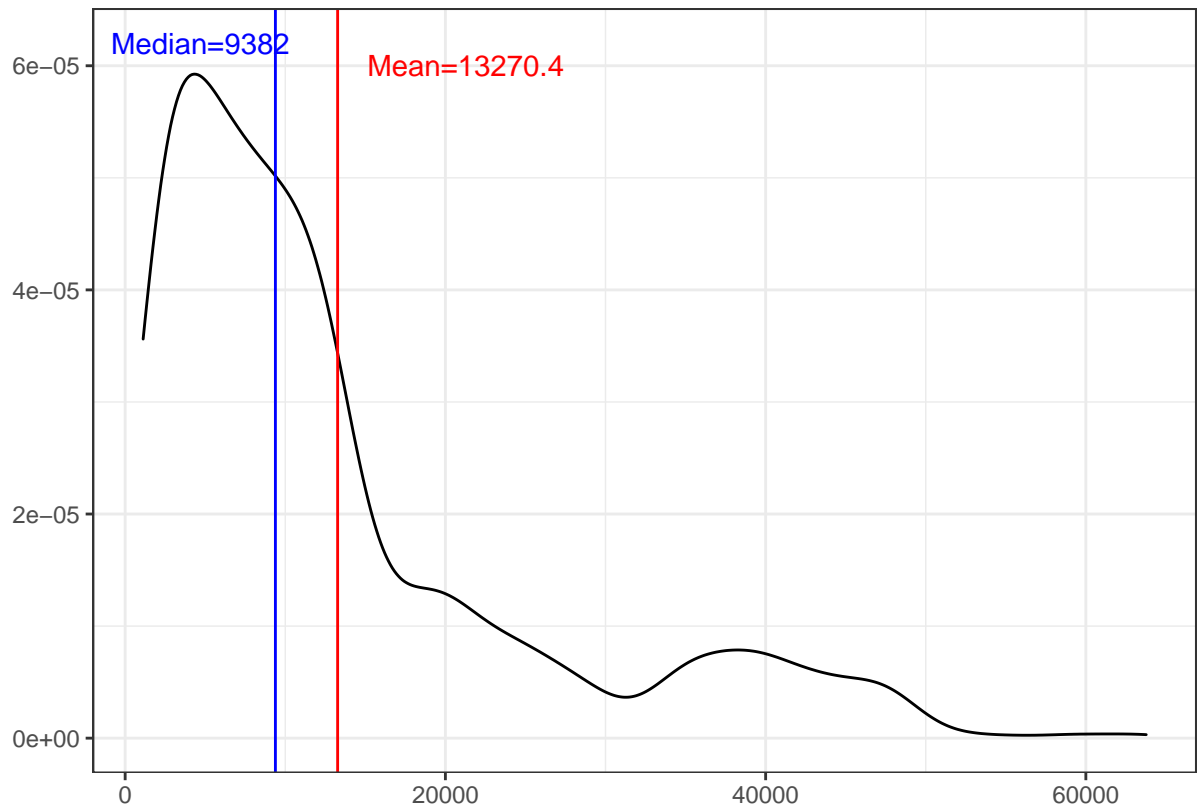


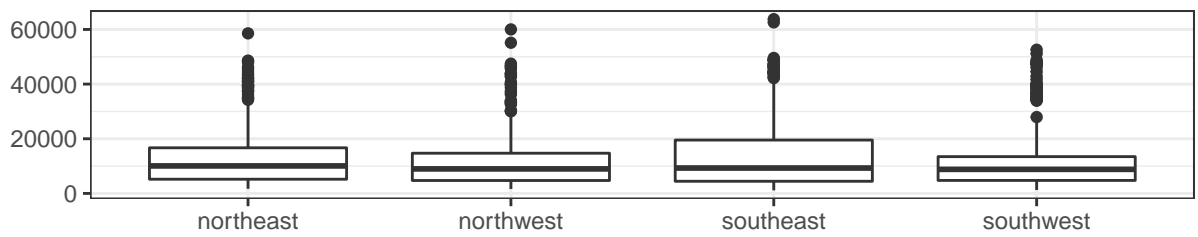
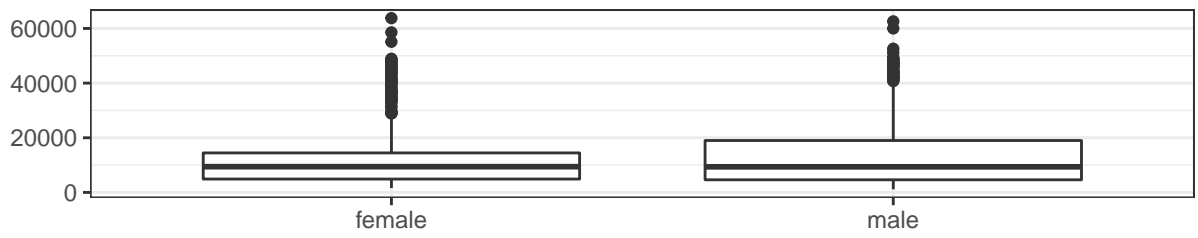
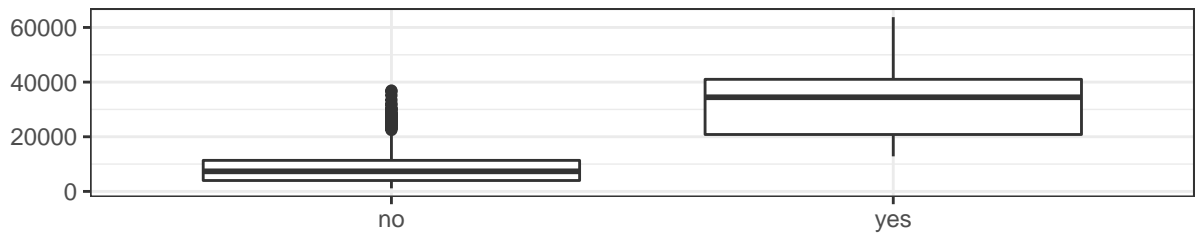
HW_Visualization

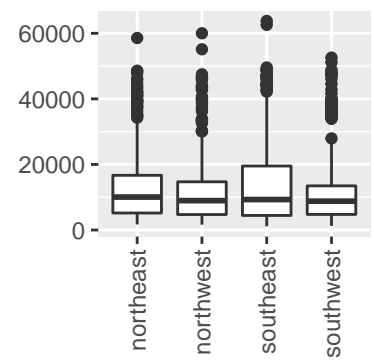
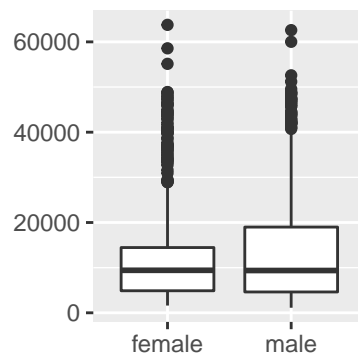
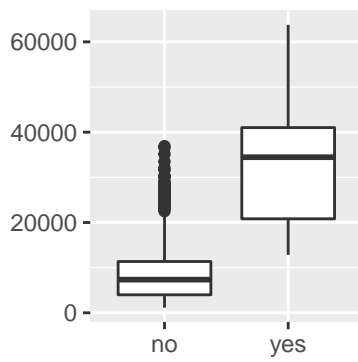
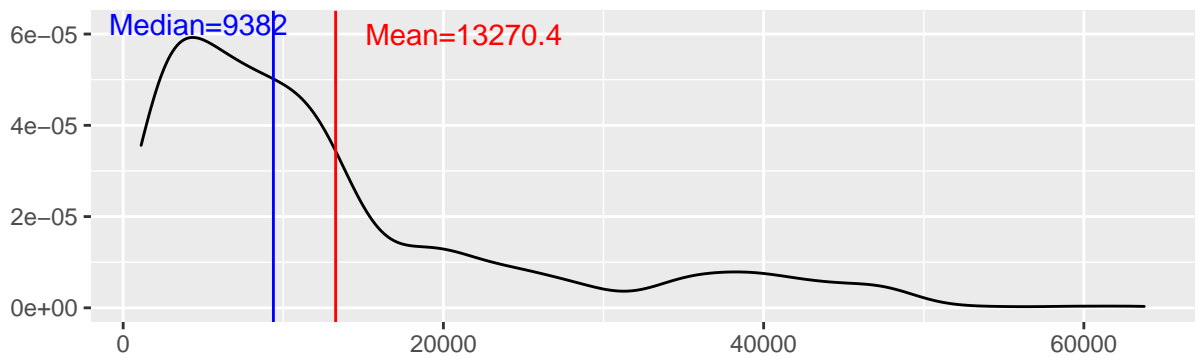
Ivanova Ekaterina

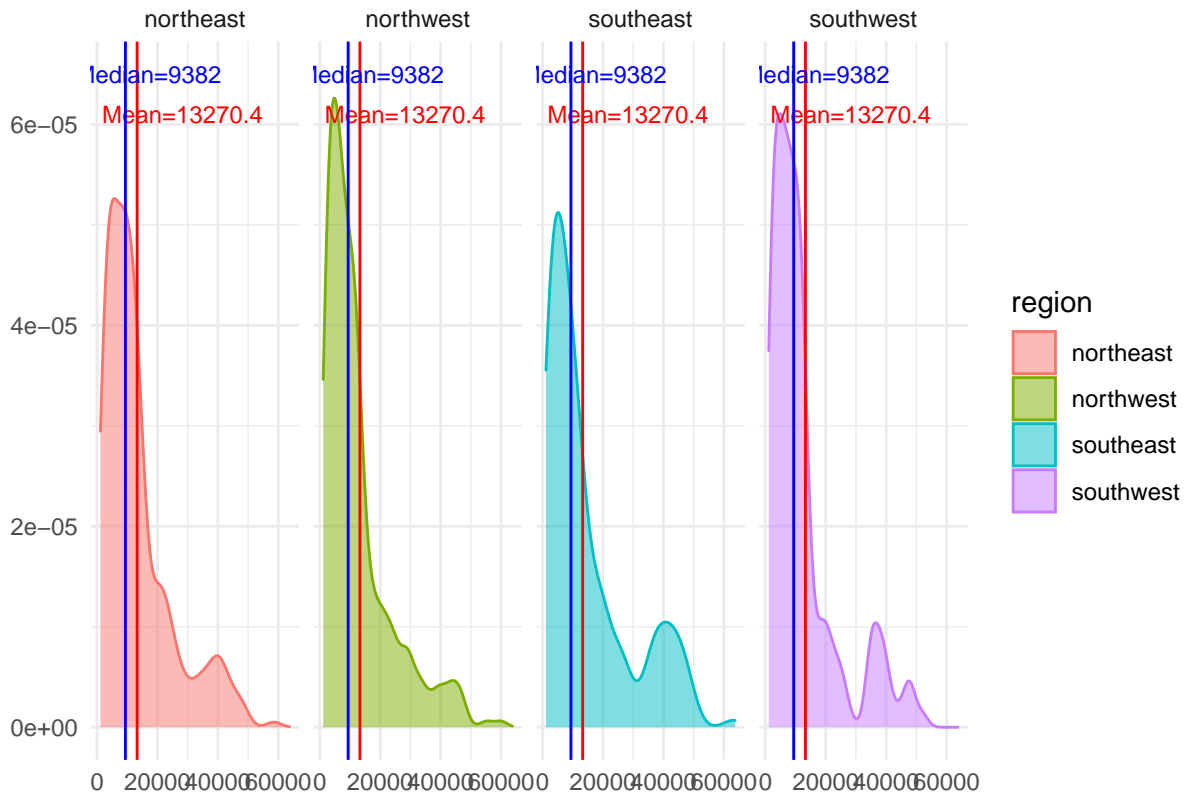
2022-10-20

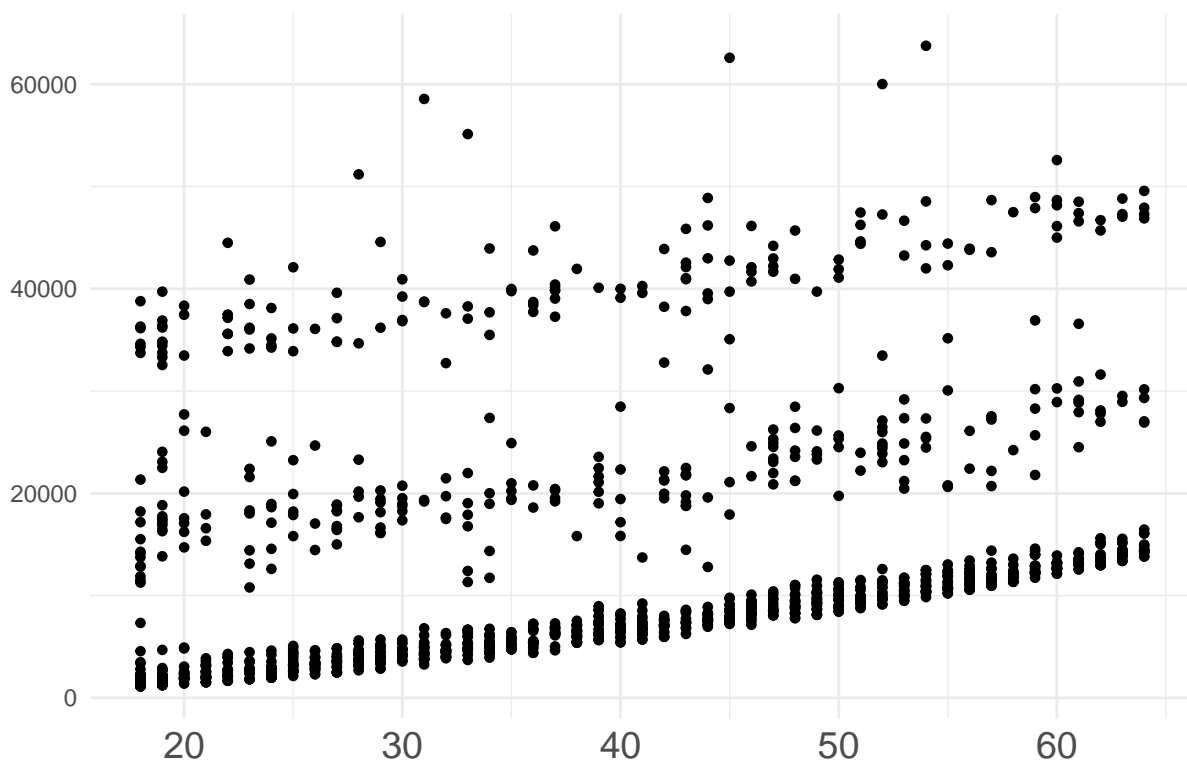




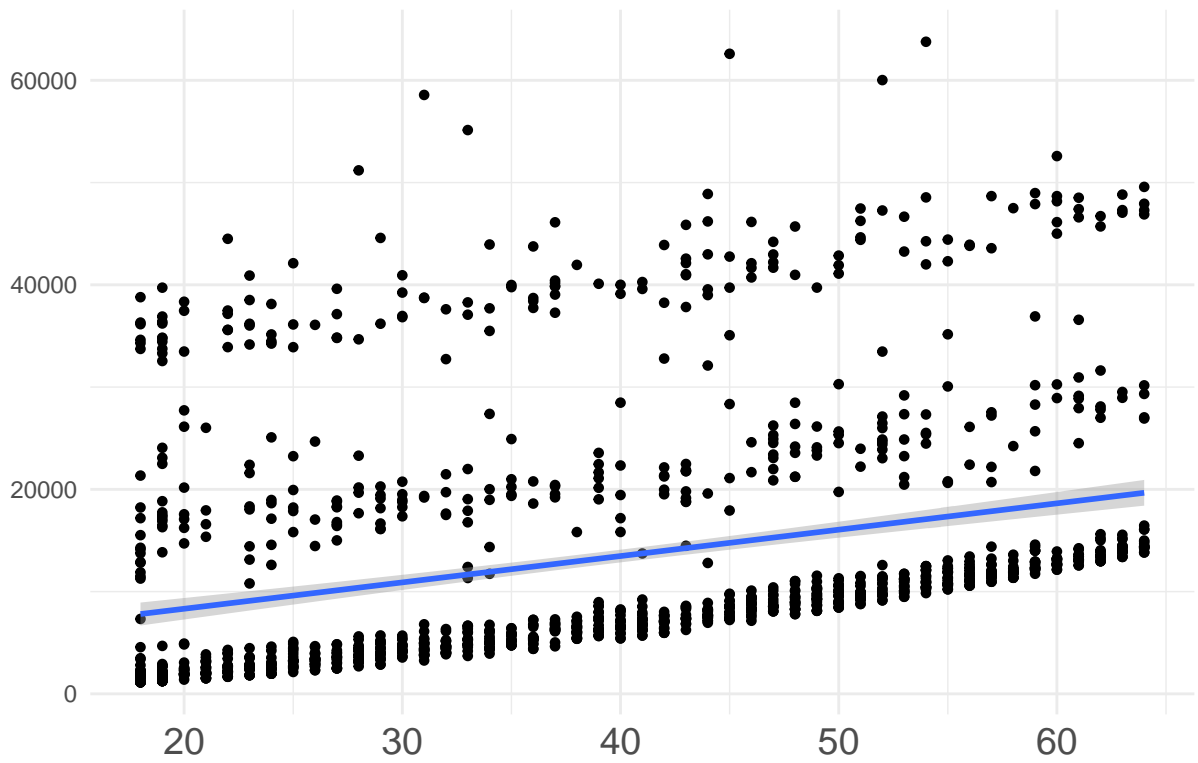






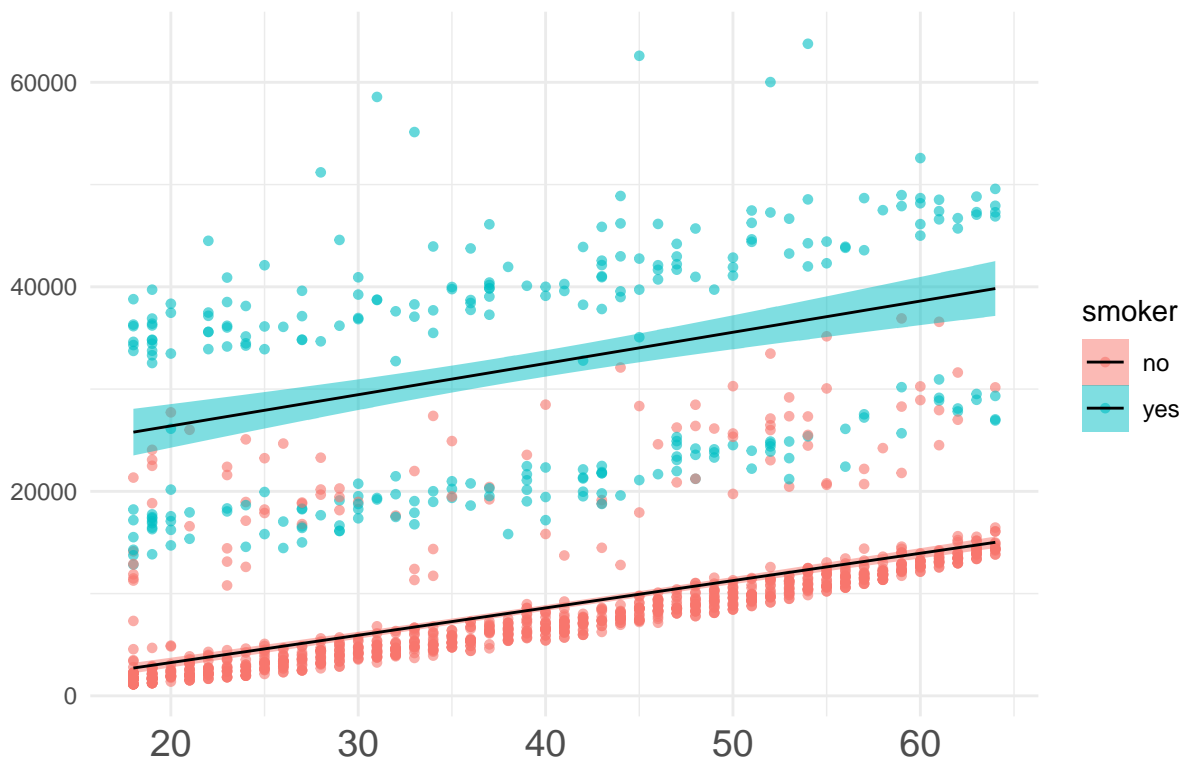


FALSE `geom_smooth()` using formula 'y ~ x'

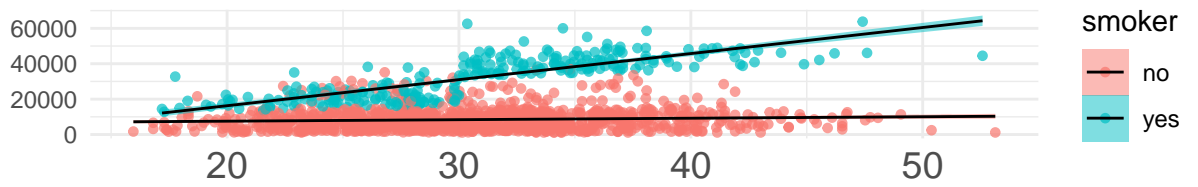
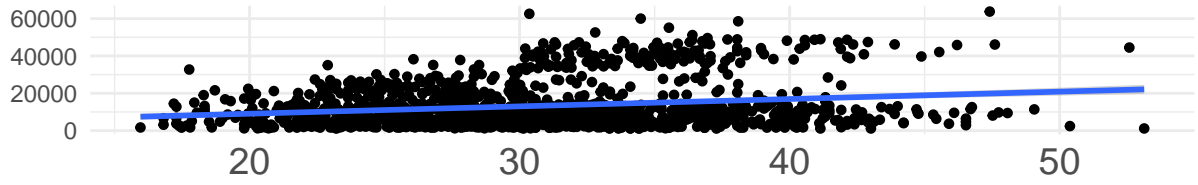
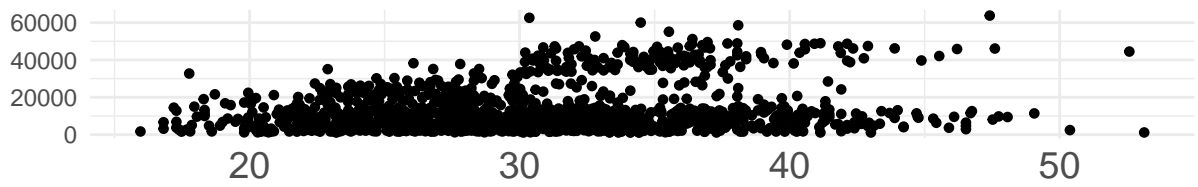


FALSE `geom_smooth()` using formula 'y ~ x'

FALSE `geom_smooth()` using formula 'y ~ x'



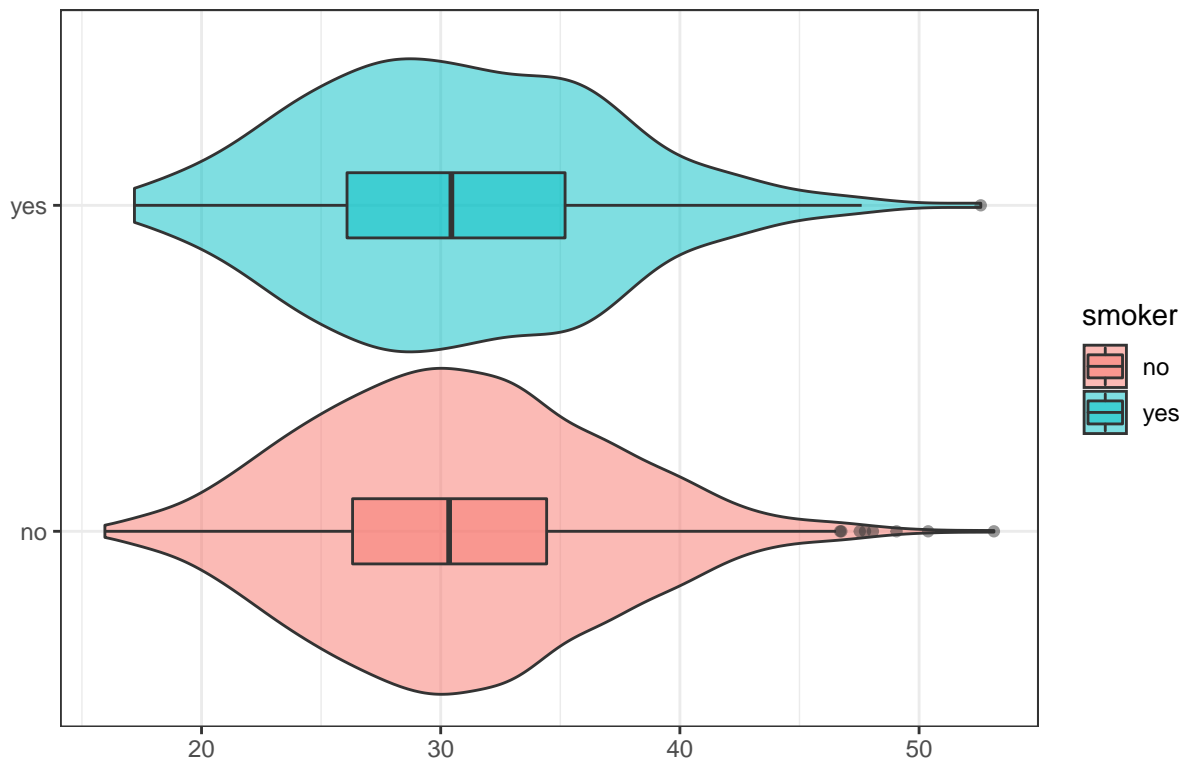
```
FALSE `geom_smooth()` using formula 'y ~ x'
FALSE `geom_smooth()` using formula 'y ~ x'
FALSE `geom_smooth()` using formula 'y ~ x'
```

Task 11

Does BMI differ between smokers and non-smokers?

It can be assumed that smokers are slimmer than non-smokers. Let's build boxplots and look at the distributions.



```
FALSE # A tibble: 2 x 3
FALSE   smoker statistic  p.value
FALSE   <chr>      <dbl>    <dbl>
FALSE 1 no          0.994 0.000422
FALSE 2 yes          0.991 0.0852

FALSE
FALSE   Shapiro-Wilk normality test
FALSE
FALSE data:  cost$bmi
FALSE W = 0.99389, p-value = 2.605e-05

FALSE
FALSE   Wilcoxon rank sum test with continuity correction
FALSE
FALSE data:  bmi by smoker
FALSE W = 145309, p-value = 0.9359
FALSE alternative hypothesis: true location shift is not equal to 0
```

Based on the data obtained, it can be concluded that, BMI is similar in smokers and non-smokers. Medians were about 30, Q1 and Q3 are also close. At the same time, in the group of smokers in the form of distribution, the second peak can be noted at the BMI value of 35.

The Wilcoxon test confirmed the conclusion based on the resulting visualization, the p-value was 0.936, the smoking status groups did not differ in BMI.