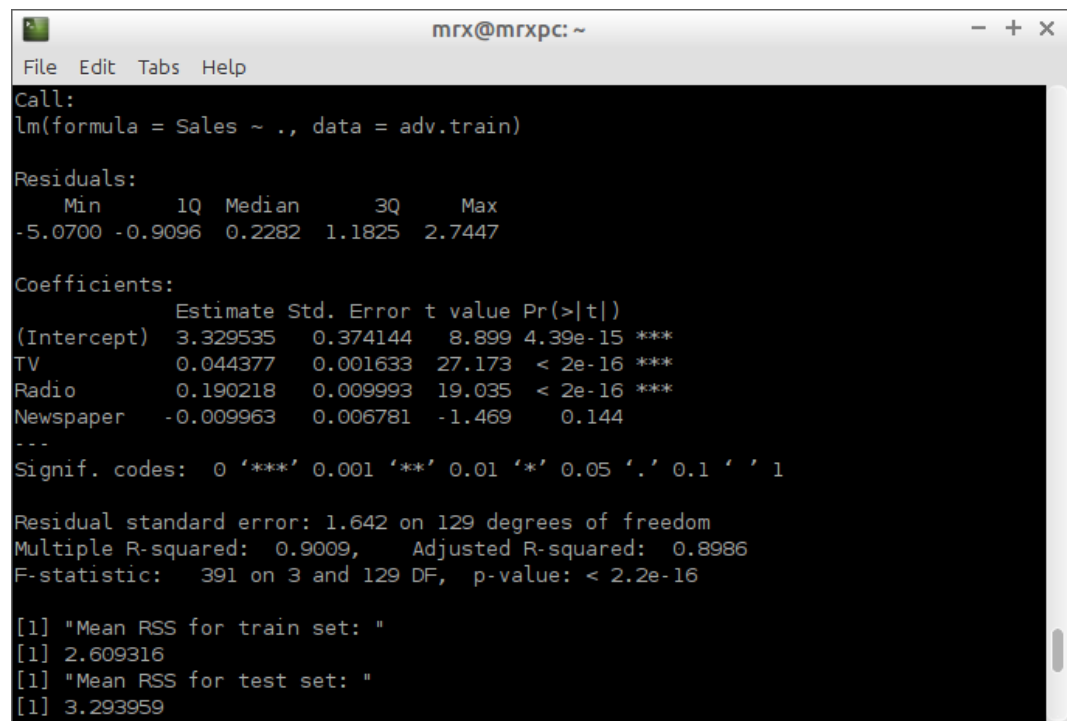


Results analysis for homework 1
Marat Khabibullin

Let's look at the program output shown in Figure 1. One can see (according to p-values) there is a dependency between Sales volume and predictors under consideration and Newspaper predictor is the least significant (because of the high p-value).

Considering mean RSS, we have higher value for the test set and it could be explained by the fact that the model has been created using training data and best fits for this data set.

Scatterplots are shown in Figure 2 and 3 and one can notice clear linear dependence.



```
Call:
lm(formula = Sales ~ ., data = adv.train)

Residuals:
    Min       1Q   Median       3Q      Max
-5.0700 -0.9096  0.2282  1.1825  2.7447

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.329535   0.374144   8.899 4.39e-15 ***
TV           0.044377   0.001633  27.173 < 2e-16 ***
Radio        0.190218   0.009993  19.035 < 2e-16 ***
Newspaper    -0.009963   0.006781  -1.469  0.144
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.642 on 129 degrees of freedom
Multiple R-squared:  0.9009,    Adjusted R-squared:  0.8986
F-statistic: 391 on 3 and 129 DF, p-value: < 2.2e-16

[1] "Mean RSS for train set: "
[1] 2.609316
[1] "Mean RSS for test set: "
[1] 3.293959
```

Figure 1: Model with all predictors

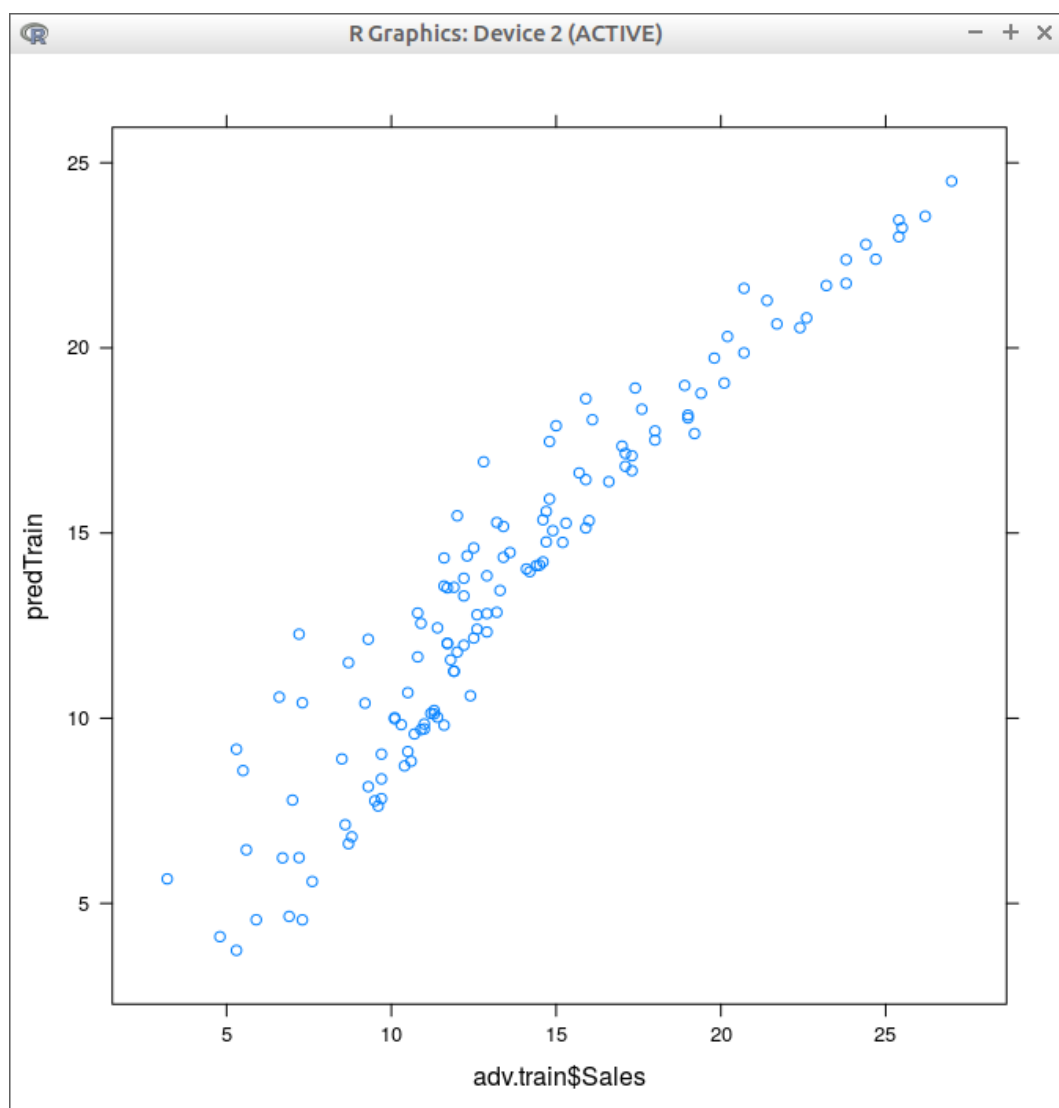


Figure 2: Scatterplot for training set (all predictors)

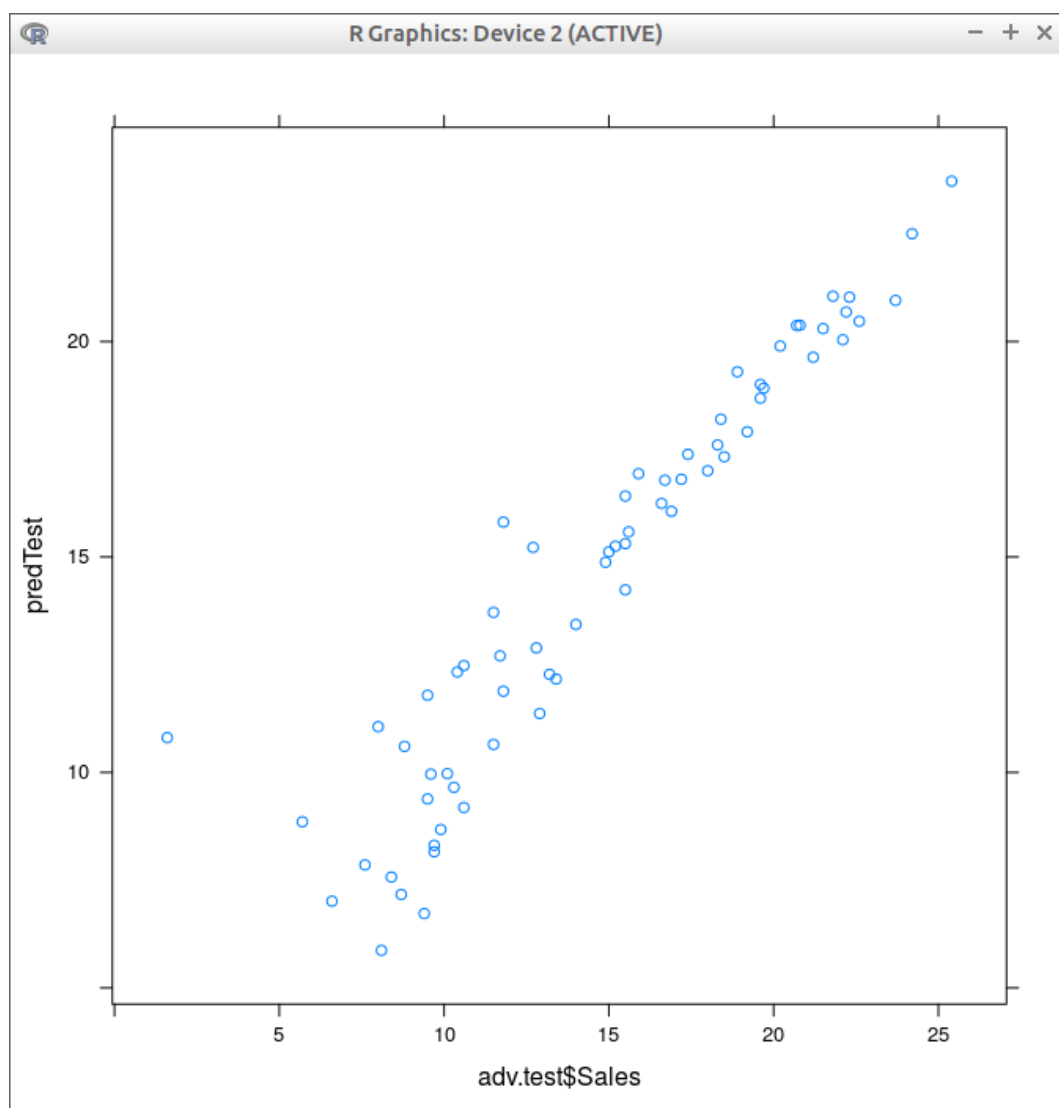
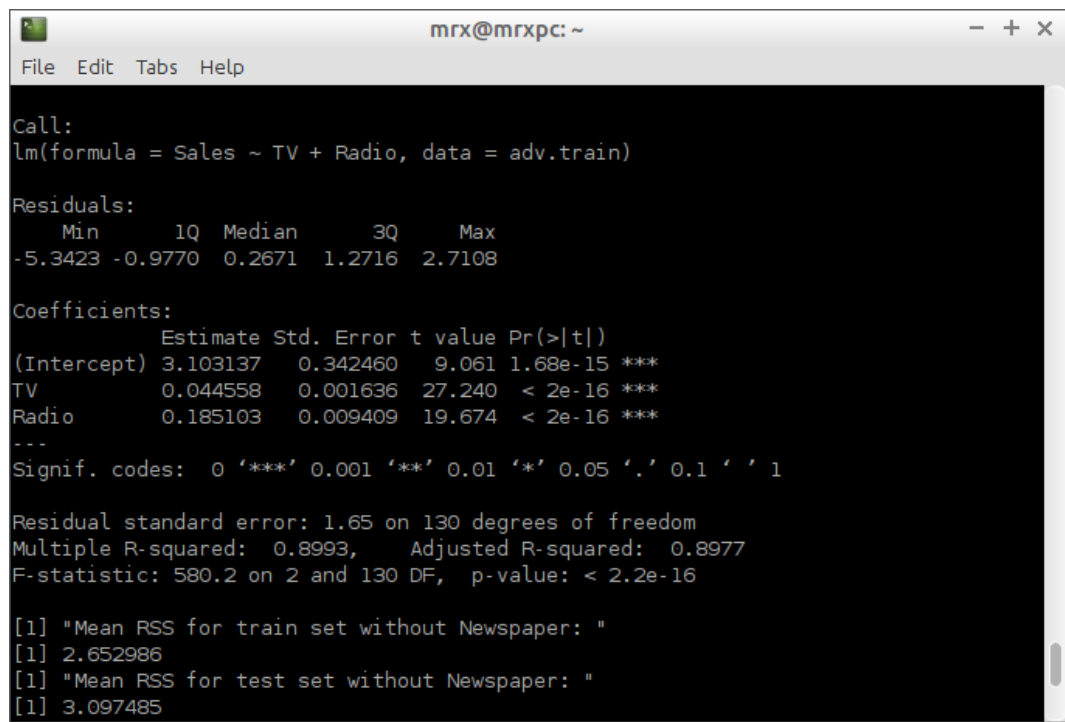


Figure 3: Scatterplot for test set (all predictors)

Now let's consider program output for model with Newspapers predictor removed (Fig. 4). One may notice in comparison with the previous results (see Fig. 1) F-statistics value has grown significantly, so we can conclude model has become better (better describes real data). Moreover, rss values are almost the same as in the model with all predictors. It proves the Newspaper predictor's insignificance. Scatterplots in Figures 5 and 6 still show linear behaviour.



```

Call:
lm(formula = Sales ~ TV + Radio, data = adv.train)

Residuals:
    Min       1Q   Median       3Q      Max
-5.3423 -0.9770  0.2671  1.2716  2.7108

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.103137   0.342460   9.061 1.68e-15 ***
TV           0.044558   0.001636  27.240 < 2e-16 ***
Radio        0.185103   0.009409  19.674 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.65 on 130 degrees of freedom
Multiple R-squared:  0.8993,    Adjusted R-squared:  0.8977
F-statistic: 580.2 on 2 and 130 DF,  p-value: < 2.2e-16

[1] "Mean RSS for train set without Newspaper: "
[1] 2.652986
[1] "Mean RSS for test set without Newspaper: "
[1] 3.097485

```

Figure 4: Model with Newspapers predictor removed

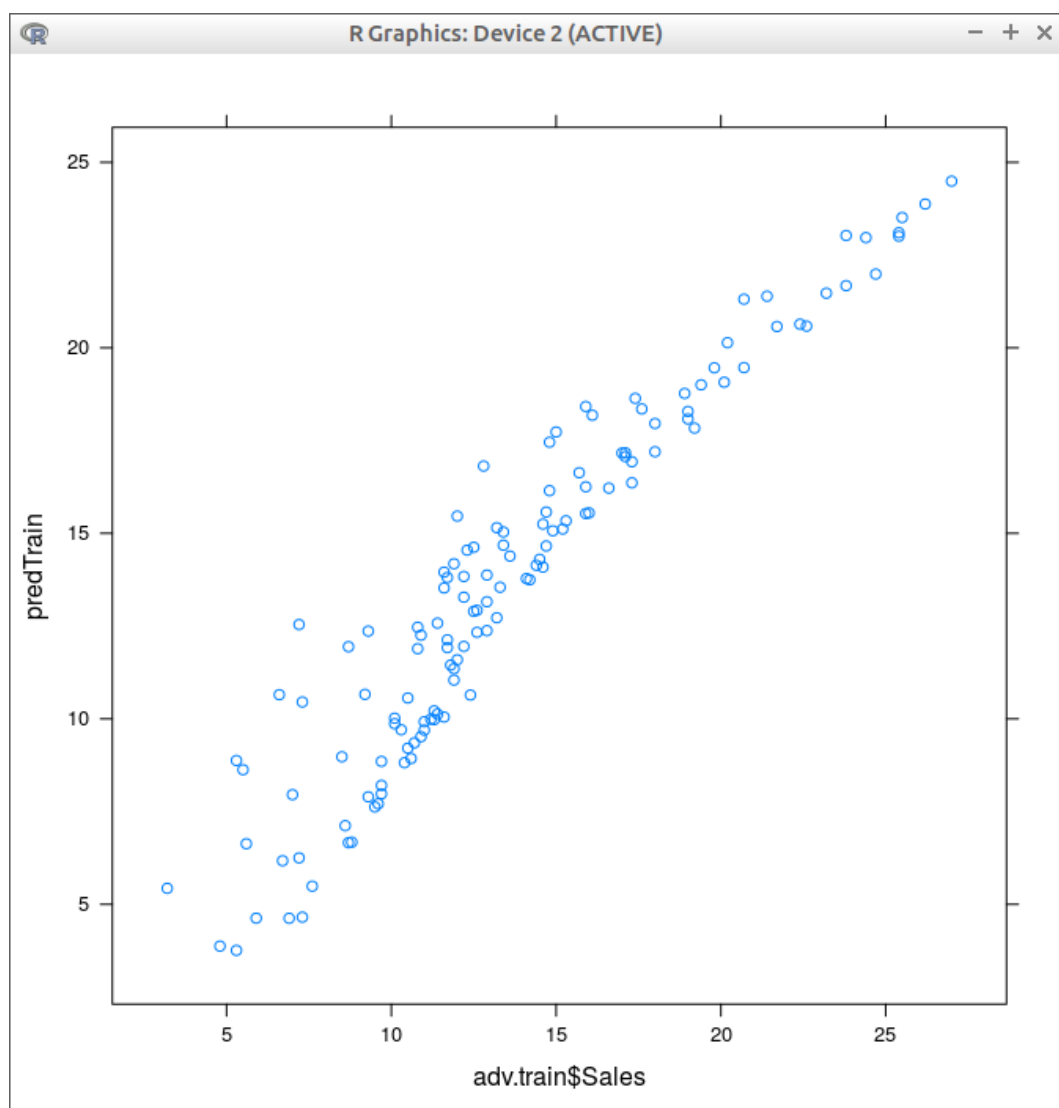


Figure 5: Scatterplot for training set (Newspapers predictor removed)

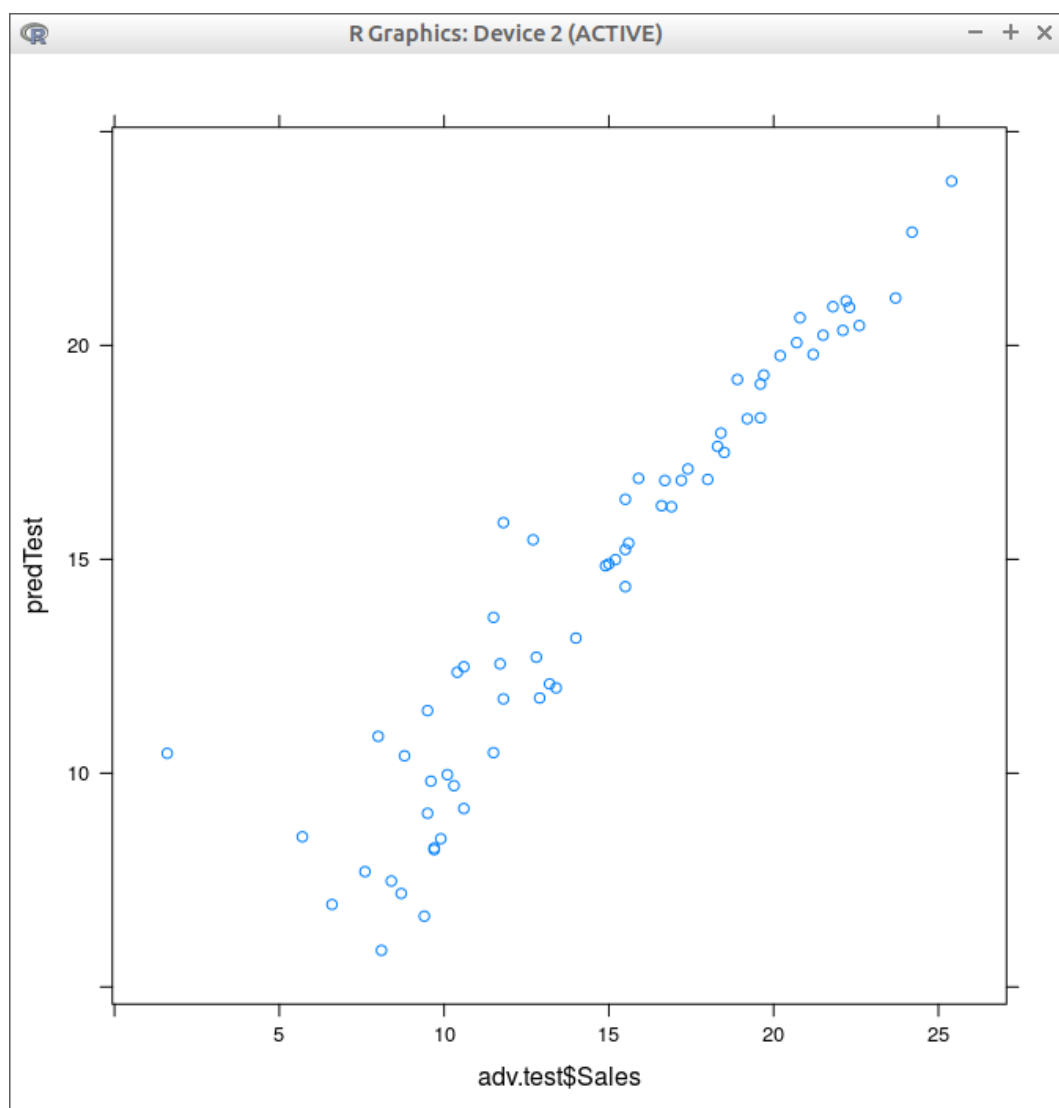
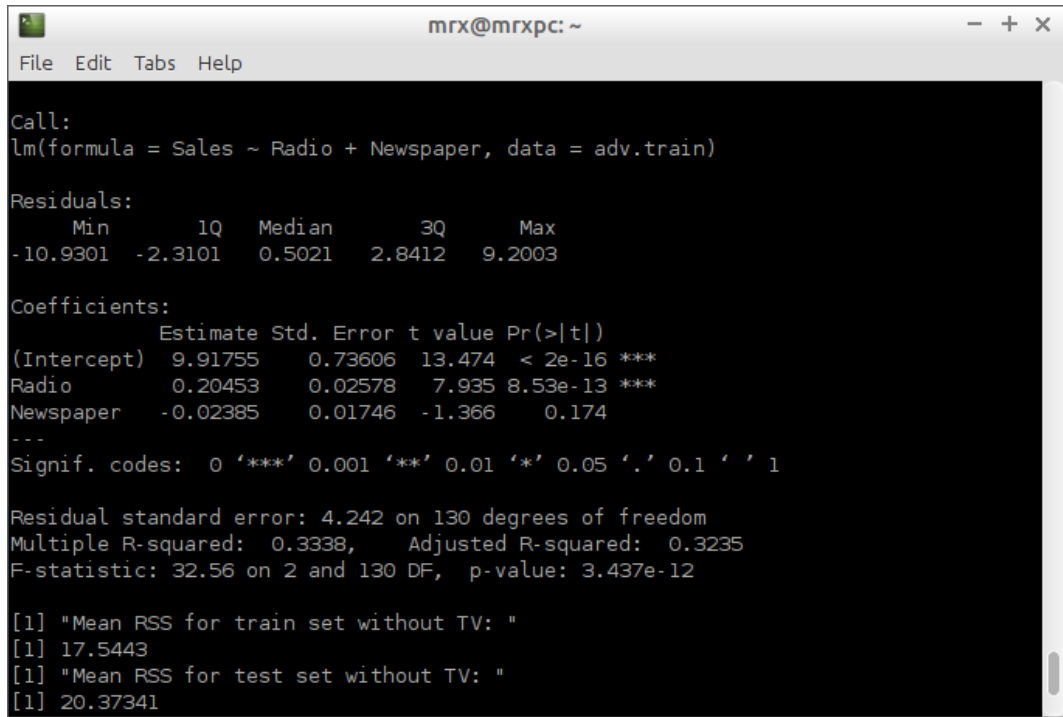


Figure 6: Scatterplot for test set (Newspapers predictor removed)

Next program output (Fig. 7) shows the model with TV predictor removed. F-statistic value has become very low that corresponds to the fact the TV predictor is significant in our model in general. Moreover, removing significant predictor we have increased rss values for both training and test data sets. Scatterplots in Figures 8 and 9 also reflect the fact of removing significant predictor - the dependencies are not linear anymore, predicted data badly corresponds to actual one.



```

mrX@mrXpc: ~
File Edit Tabs Help

Call:
lm(formula = Sales ~ Radio + Newspaper, data = adv.train)

Residuals:
    Min       1Q   Median       3Q      Max
-10.9301  -2.3101   0.5021   2.8412   9.2003

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.91755    0.73606  13.474  < 2e-16 ***
Radio        0.20453    0.02578   7.935 8.53e-13 ***
Newspaper    -0.02385    0.01746  -1.366   0.174
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.242 on 130 degrees of freedom
Multiple R-squared:  0.3338,    Adjusted R-squared:  0.3235
F-statistic: 32.56 on 2 and 130 DF,  p-value: 3.437e-12

[1] "Mean RSS for train set without TV: "
[1] 17.5443
[1] "Mean RSS for test set without TV: "
[1] 20.37341

```

Figure 7: Model with Tv predictor removed

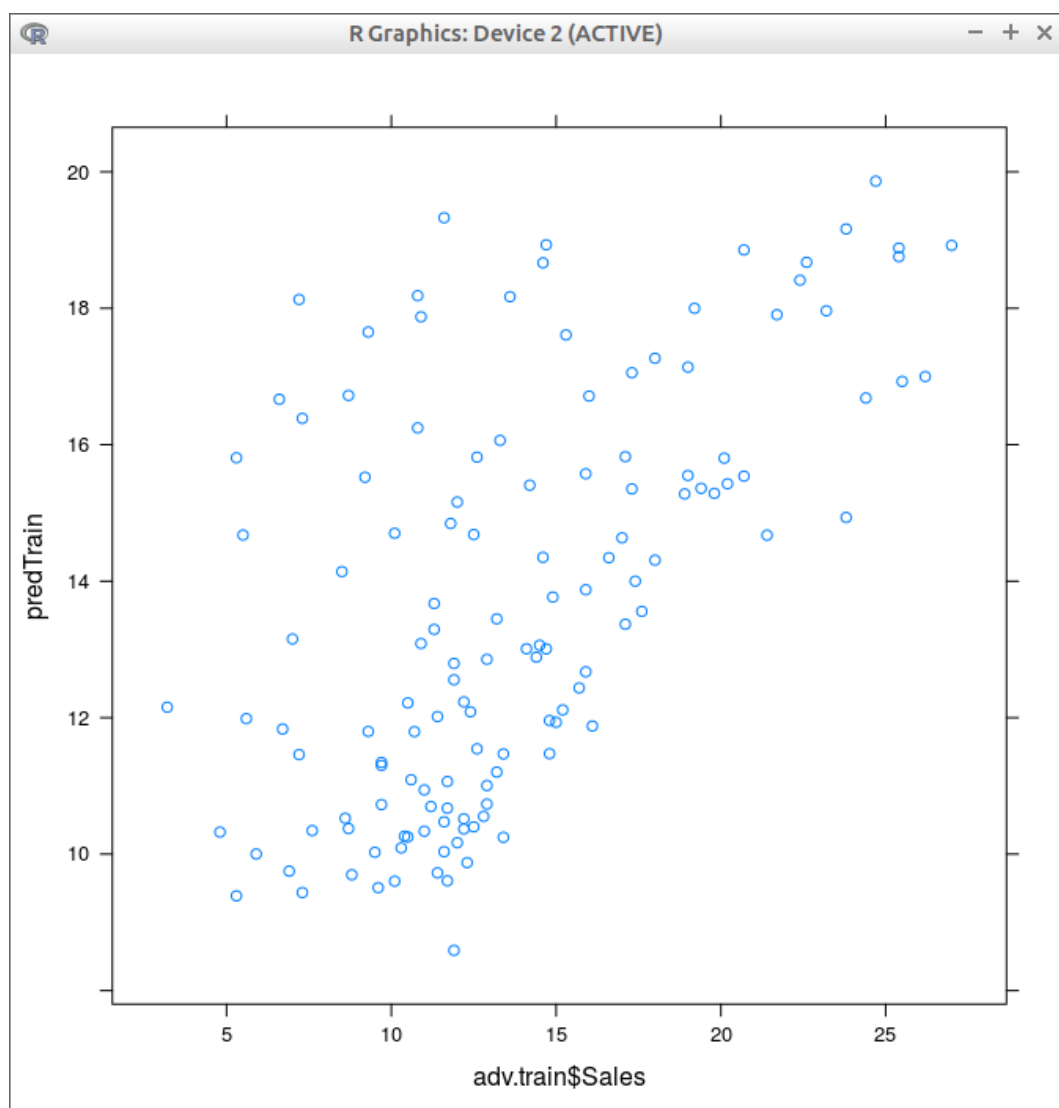


Figure 8: Scatterplot for training set (Tv predictor removed)

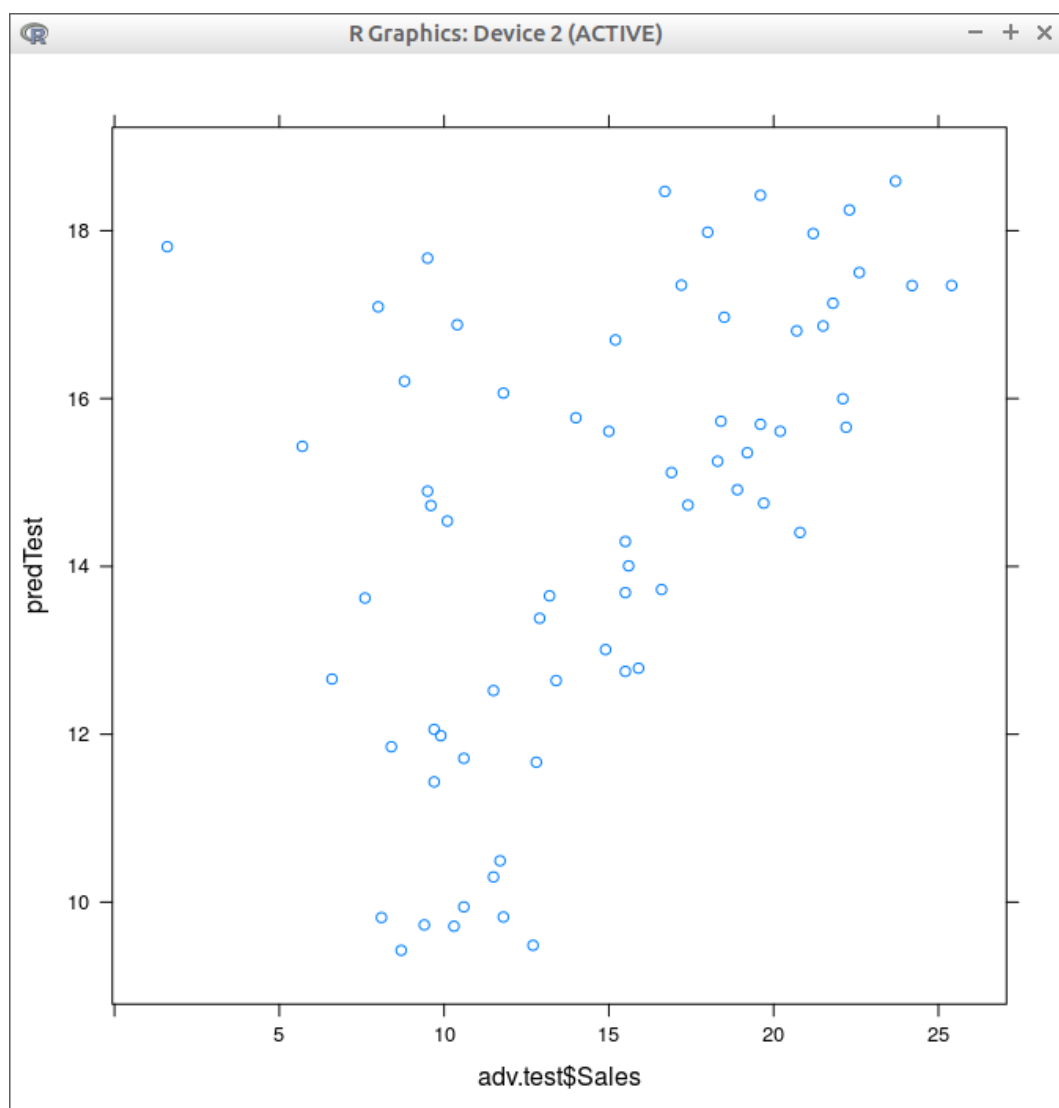
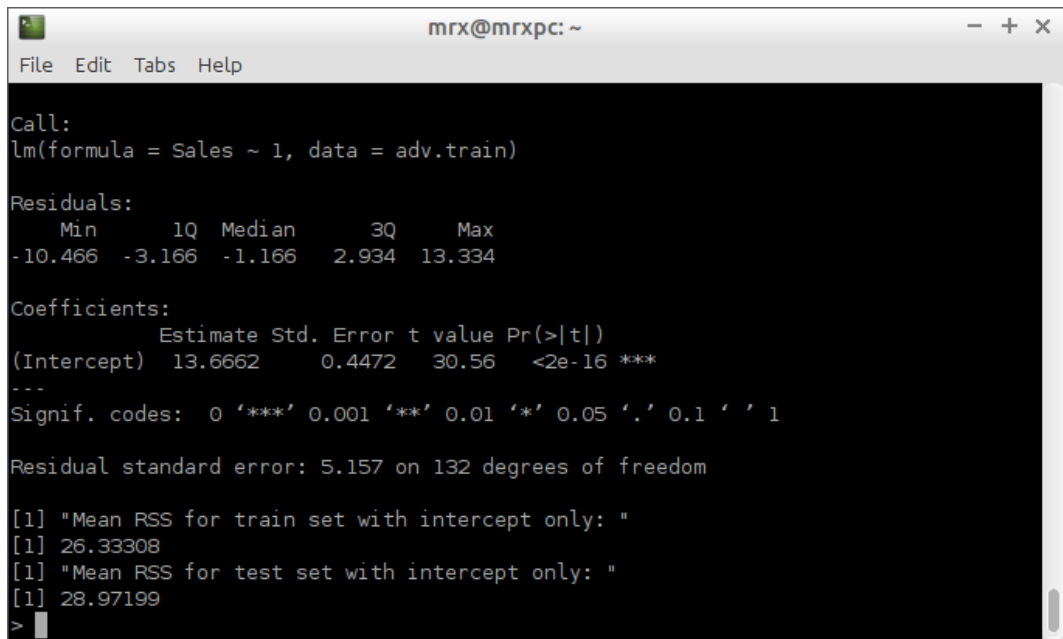


Figure 9: Scatterplot for test set (Tv predictor removed)

The last model to consider is one with all predictor removed (Fig. 10). As we can see rss values have drastically increased in comparison with the initial model values. Also, scatterplots in Figures 11 and 12 show predicted Sales volumes are constant and equals to calculated Intercept value (see Fig. 10). It clearly represents independency of model predictions and input data making such model useless.



```

mrX@mrXpc: ~
File Edit Tabs Help

Call:
lm(formula = Sales ~ 1, data = adv.train)

Residuals:
    Min       1Q   Median       3Q      Max
-10.466  -3.166  -1.166   2.934  13.334

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  13.6662     0.4472   30.56  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.157 on 132 degrees of freedom

[1] "Mean RSS for train set with intercept only: "
[1] 26.33308
[1] "Mean RSS for test set with intercept only: "
[1] 28.97199
>

```

Figure 10: Model with all predictors removed

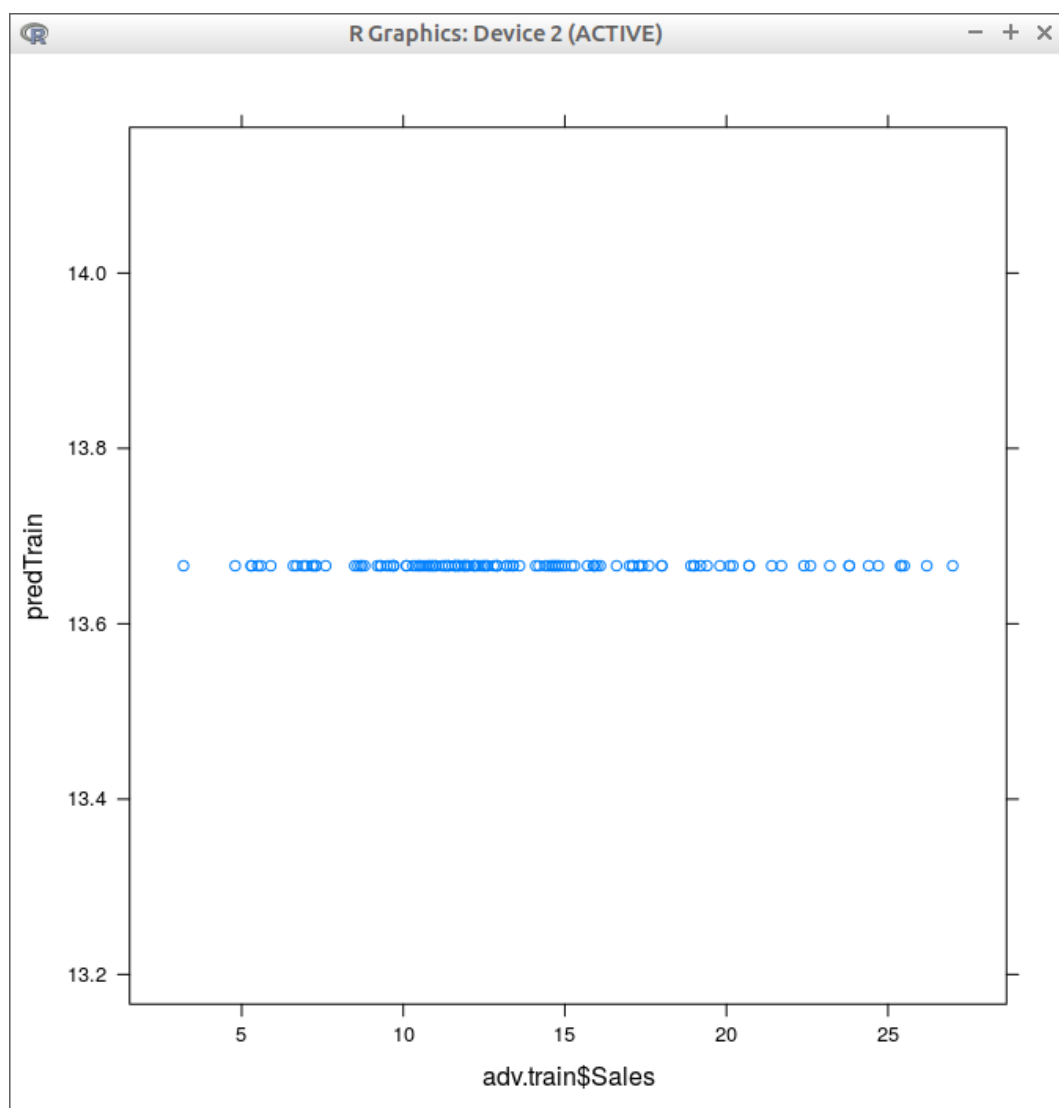


Figure 11: Scatterplot for training set (all predictors removed)

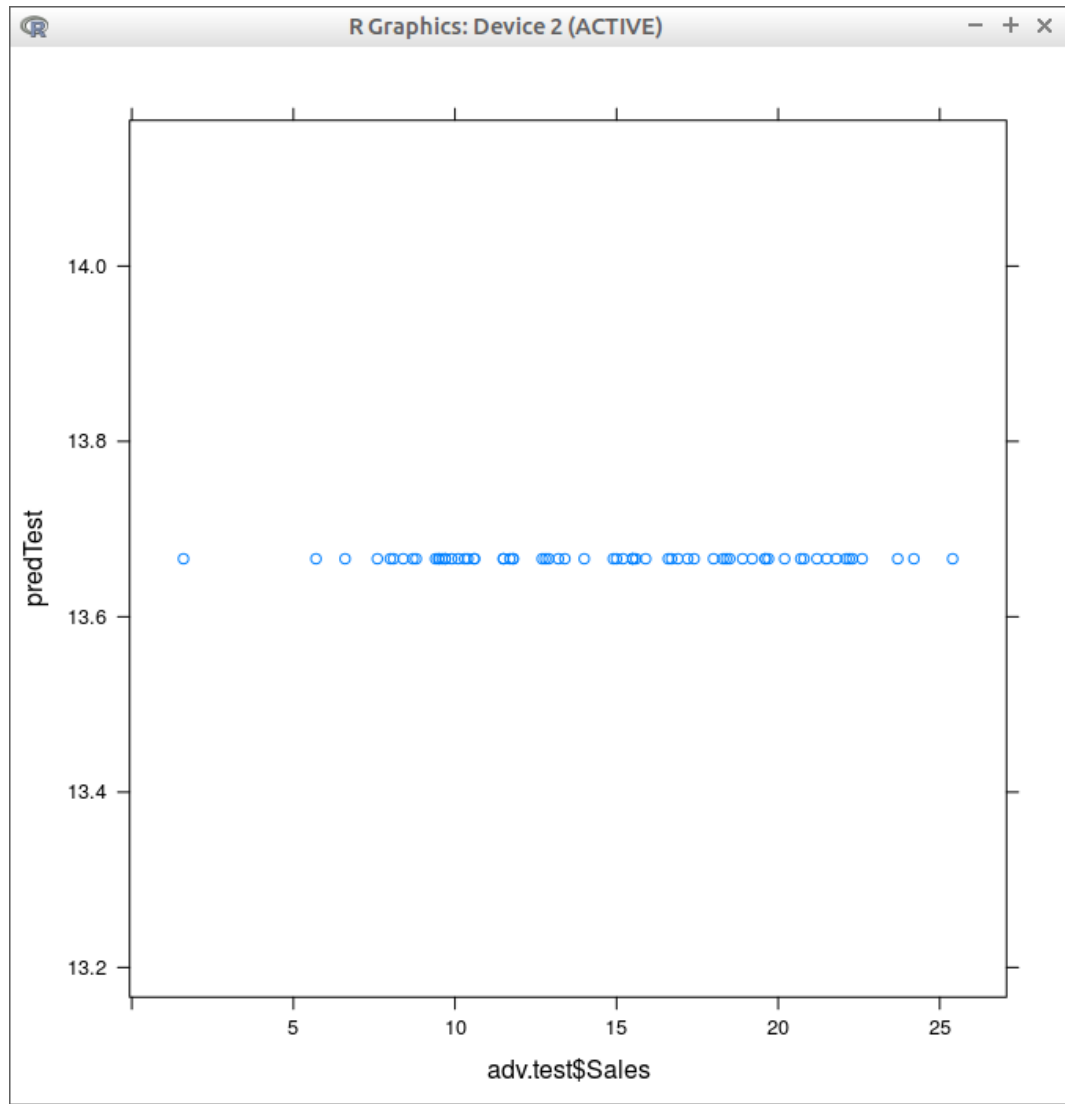


Figure 12: Scatterplot for test set (all predictors removed)

In conclusion it is important to note that on different program runs different results are observed. For example, one program run can give a higher rss value for the test data set in comparison to the training data set. This fact could be explained by the randomness of choosing what part of all the input data will be the test one and what will be the training one.