



# Программирование в среде R

Шевцов Василий Викторович,  
директор ДИТ РУДН, [shevtsov\\_vv@rudn.university](mailto:shevtsov_vv@rudn.university)

# ggplot2

# install.packages("ggplot2")

downloaded 1.9 MB

package 'stringi' successfully unpacked and MD5 sums checked  
package 'magrittr' successfully unpacked and MD5 sums checked  
package 'colorspace' successfully unpacked and MD5 sums checked  
package 'Rcpp' successfully unpacked and MD5 sums checked  
package 'stringr' successfully unpacked and MD5 sums checked  
package 'RColorBrewer' successfully unpacked and MD5 sums checked  
package 'dichromat' successfully unpacked and MD5 sums checked  
package 'munsell' successfully unpacked and MD5 sums checked  
package 'labeling' successfully unpacked and MD5 sums checked  
package 'digest' successfully unpacked and MD5 sums checked  
package 'gtable' successfully unpacked and MD5 sums checked  
package 'plyr' successfully unpacked and MD5 sums checked  
package 'reshape2' successfully unpacked and MD5 sums checked  
package 'scales' successfully unpacked and MD5 sums checked  
package 'ggplot2' successfully unpacked and MD5 sums checked

<https://www.rdocumentation.org/packages/ggplot2/versions/3.1.0>

# ggplot2

В ggplot2 график является результатом взаимодействия ряда элементов:

- Массив данных (**data**)
- Схема соответствия переменных из массива визуальным средствам (**aesthetic**)
- Геометрический объект (**geom**)
- Статистическое преобразование (**stat**)
- Координатная система (**coord**)
- Ориентиры (**guide**)
- Панели (**facet**)
- Художественное оформление (**theme**)

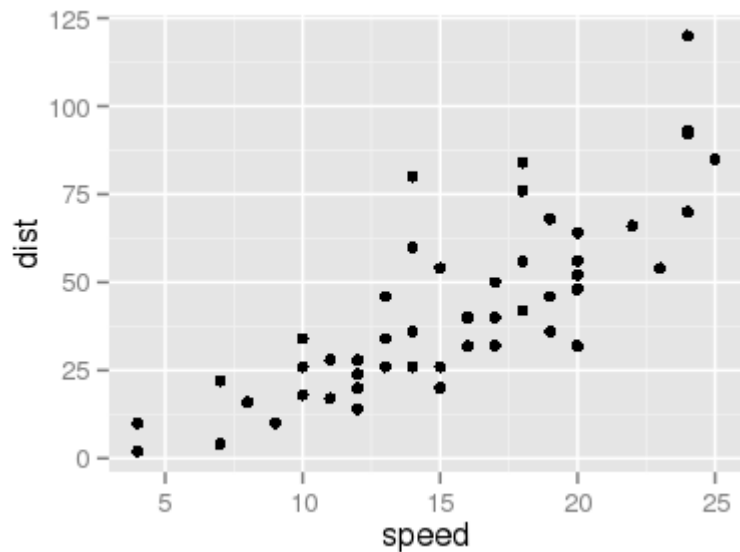
## ggplot2 пример

```
ggplot( data = cars, aes(x = speed, y = dist) ) + geom_point()
```

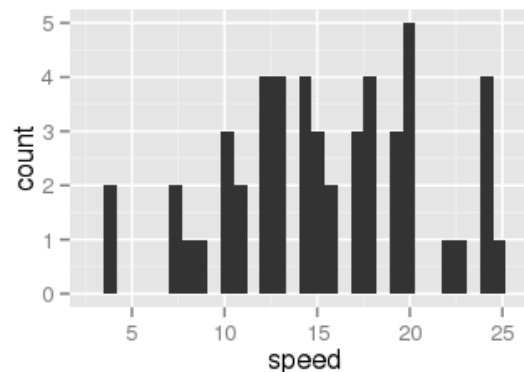
Массив  
данных

Схема соответствия  
переменных из  
массива визуальным  
средствам

Геометрический  
объект



```
ggplot (  
  data = cars, aes(x = speed)  
)  
+ geom_bar()
```

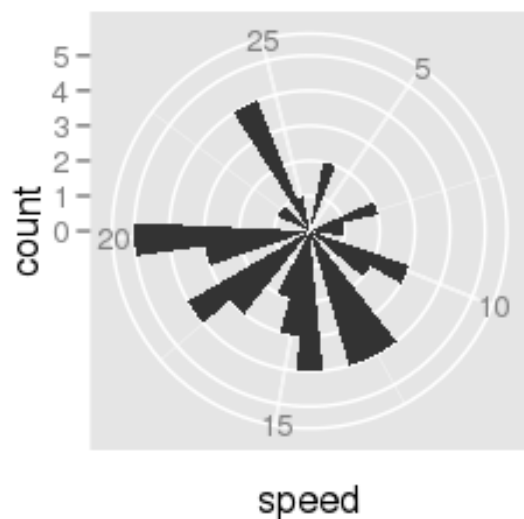


# ggplot2

Координатная система по умолчанию – декартова (coord\_cartesian).

Полярные координаты (coord\_polar):

```
ggplot(data = cars, aes(x = speed)) + geom_bar() + coord_polar()
```



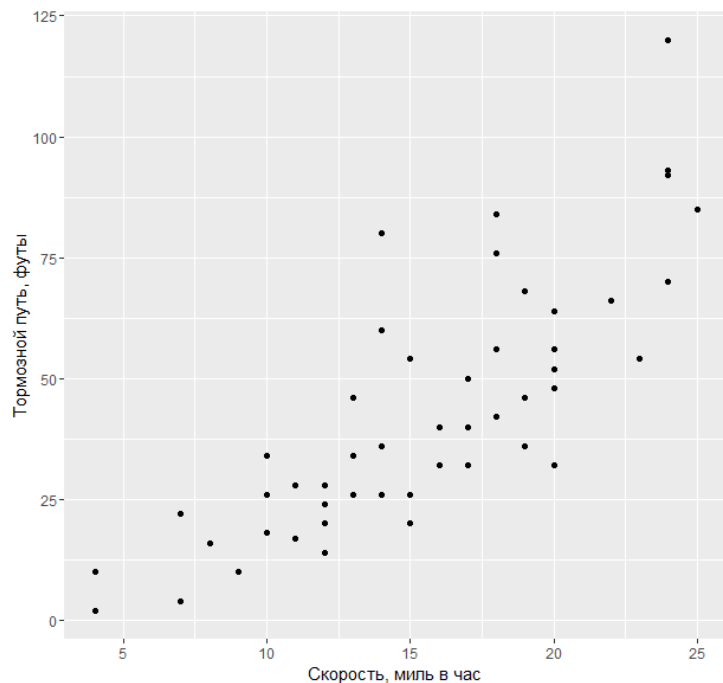
## ggplot2. Подписи осей графика

```
fig <- ggplot(data = cars, aes(x = speed, y = dist)) + geom_point()
```

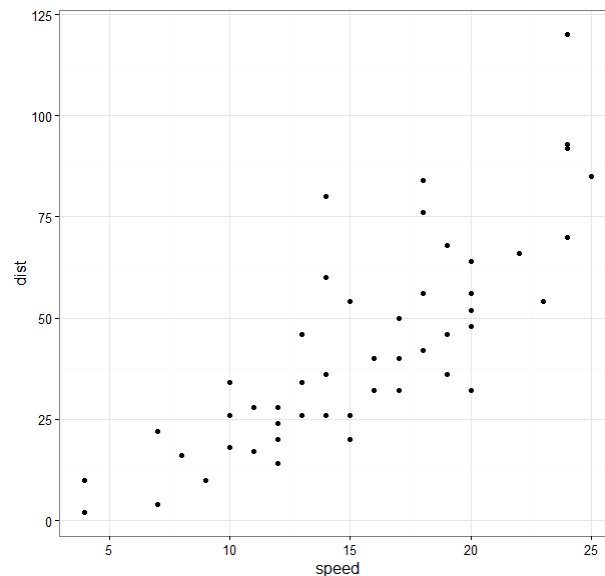
```
fig + xlab("Скорость, миль в час") + ylab("Тормозной путь, футы")
```

```
fig
```

```
fig + labs(x = "Скорость, миль в час", y = "Тормозной путь, футы")
```



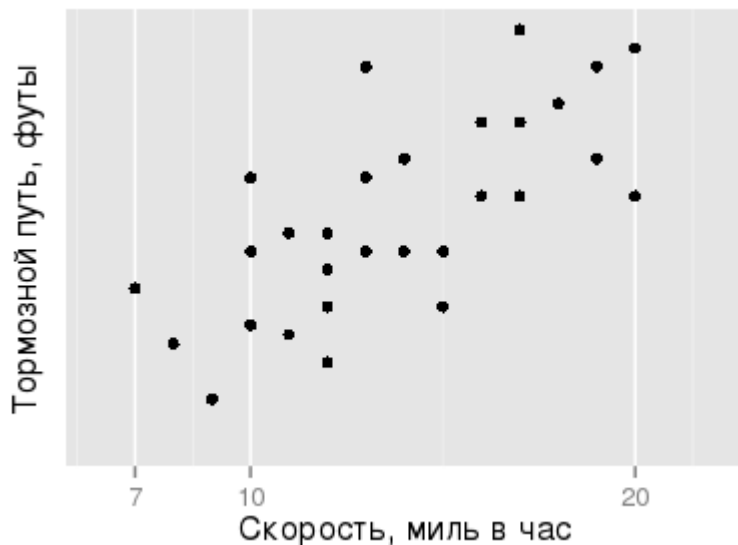
```
fig + theme_bw()
```



## ggplot2. координатные оси

```
fig + scale_x_continuous(name = "Скорость, миль в час") +  
  scale_y_continuous(name = "Тормозной путь, футы")
```

```
fig + scale_x_continuous(name = "Скорость, миль в час",  
  breaks = c(7, 10, 20),  
  limits = c(6, 22)) +  
  scale_y_continuous(name = "Тормозной путь, футы",  
    limits = c(5, 50),  
    breaks = NULL)
```

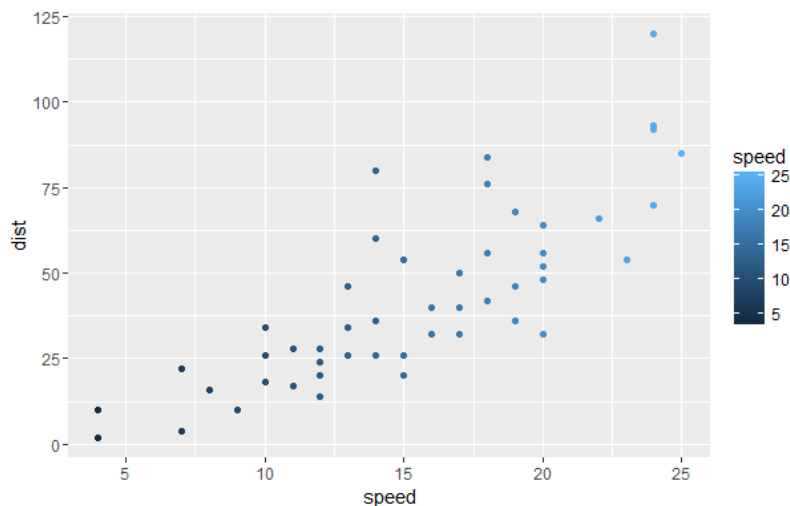


по оси X «обрезали» края графика, выходящие за пределы менее 6 и более 22 миль в час, указали, чтобы линии сетки проходили только по точкам 7, 10 и 20 миль в час. для оси Y убраны горизонтальные линии сетки и соответствующие подписи.

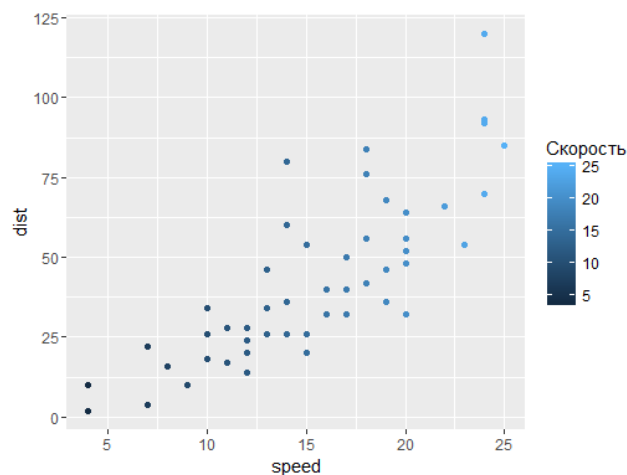


## ggplot2. легенды

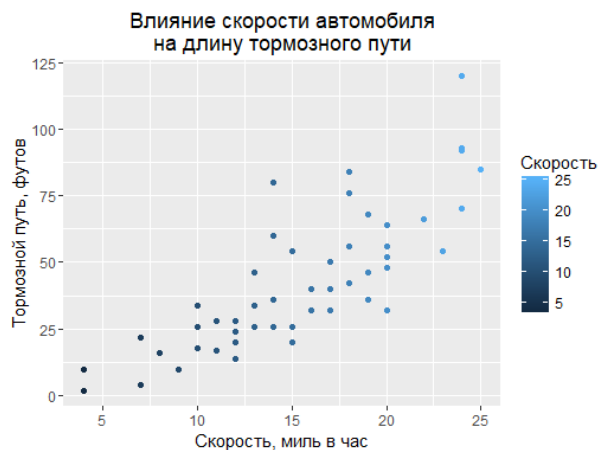
```
fig2 <- ggplot(data = cars, aes(speed, dist, color = speed))+geom_point()
```



```
fig2 + scale_color_continuous("Скорость")
```



```
fig2 <- ggplot(data = cars, aes(x = speed, y =  
dist, color = speed)) + geom_point()  
fig2 + scale_color_continuous("Скорость") +  
ggtitle("Влияние скорости автомобиля\nна  
длину тормозного пути") +  
xlab("Скорость, миль в час") +  
ylab("Тормозной путь, футов")
```

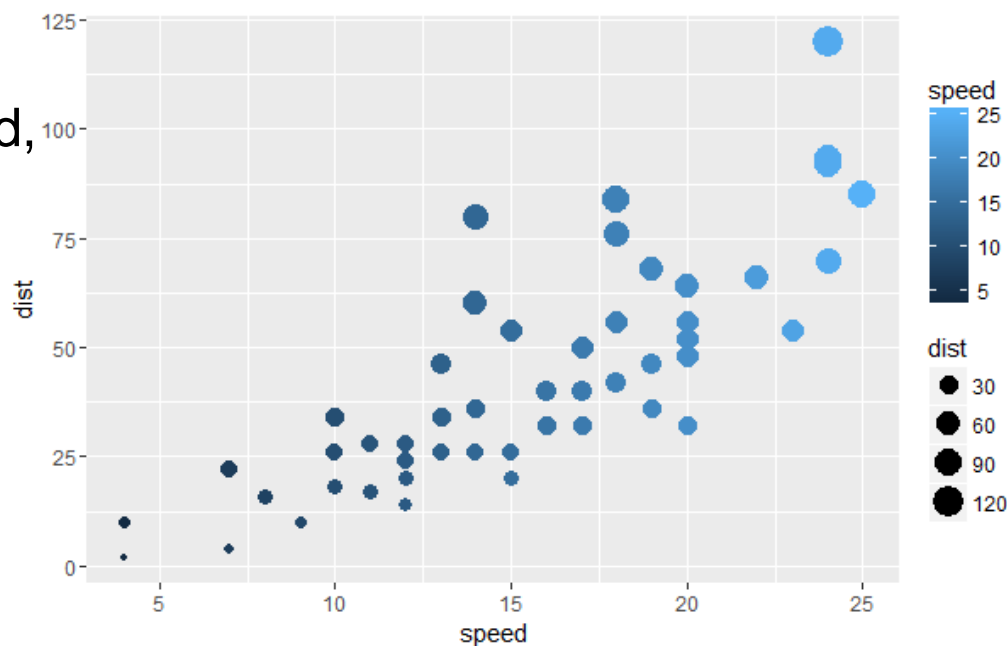


## ggplot2. визуальные средства и геометрические объекты

Визуальные средства для отрисовки геометрических объектов типа «точка»

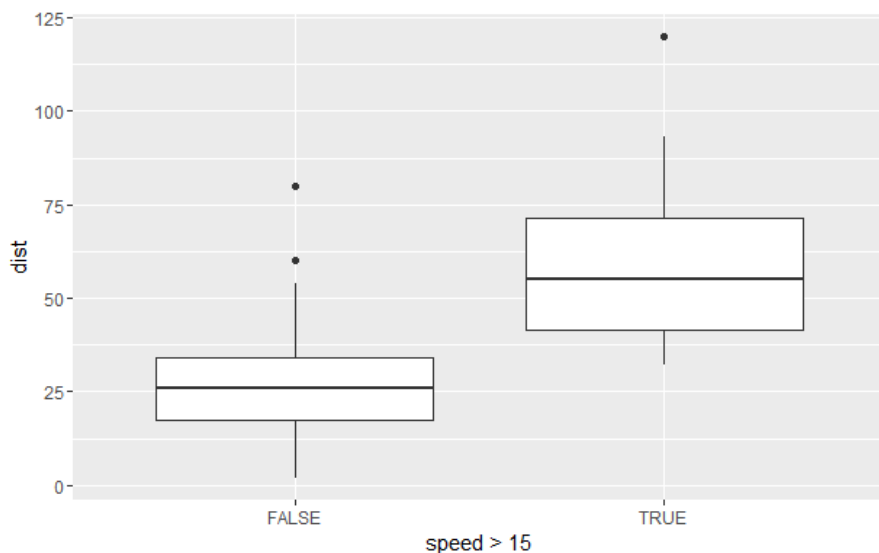
- (point) — это расположение по горизонтали (x) и вертикали (y),
- цвет границы (colour)
- цвет заливки (fill)
- размер (size)
- форма (shape)
- прозрачность (alpha)

```
ggplot(data = cars, aes(x = speed,  
                        y = dist,  
                        color = speed,  
                        size = dist))  
+ geom_point()
```



## ggplot2. boxplot

```
ggplot(data = cars, aes(x = speed > 15, y = dist)) + geom_boxplot()
```



### Описание

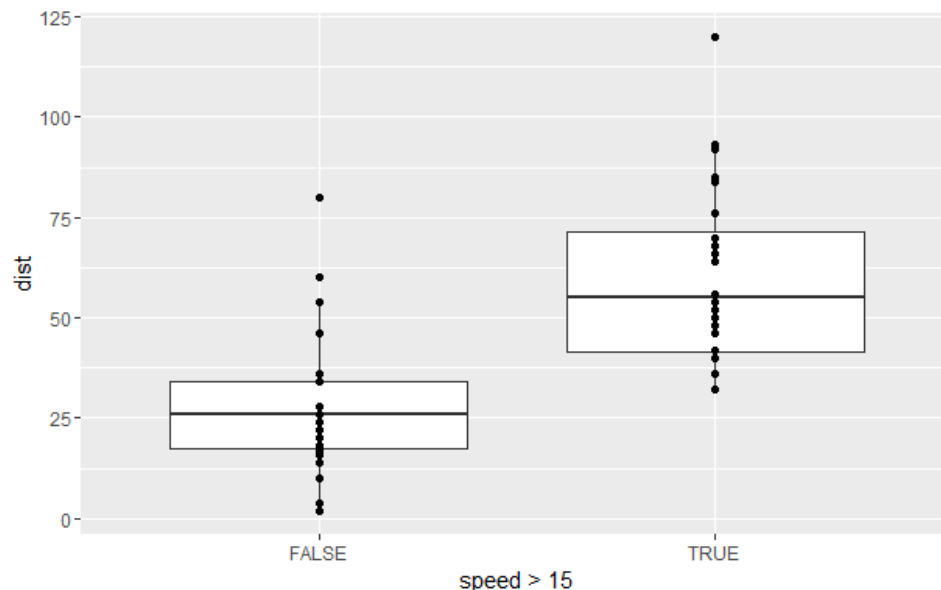
Берётся массив данных cars.

Результат выполнения выражения `cars$speed > 15` сопоставляется абсциссе, переменная `dist` — ординате. При этом абсцисса (x) из числовой переменной (мили в час) превратилась в категориальную (логическую: больше 15 миль в час или нет), поэтому все наблюдения разбились на эти две категории. Данные передаются статистическому преобразованию `stat_boxplot()`, которое из предыдущего шага знает, что `cars$speed > 15` нужно использовать для деления массива на группы, а `dist` — для расчёта статистик, необходимых для отрисовки «ящичков с усами».

Результаты работы `stat_boxplot()`, а это для каждой группы медиана, квартили, границы нижнего и верхнего «усов», передаются алгоритму отрисовки.

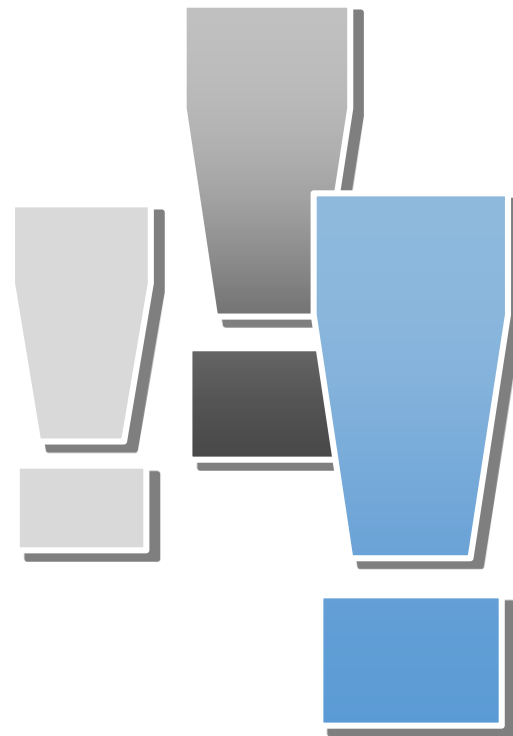
## ggplot2. слои

```
ggplot(data = cars, aes(x = speed > 15, y = dist)) +  
  geom_boxplot() +  
  geom_point()
```



добавление к ggplot2-объекту новых  
geom\_\*() или stat\_\*() приводит к  
появлению на графике нового слоя

# Спасибо за внимание!



Шевцов Василий Викторович

shevtsov\_vv@rudn.university  
+7(903)144-53-57