

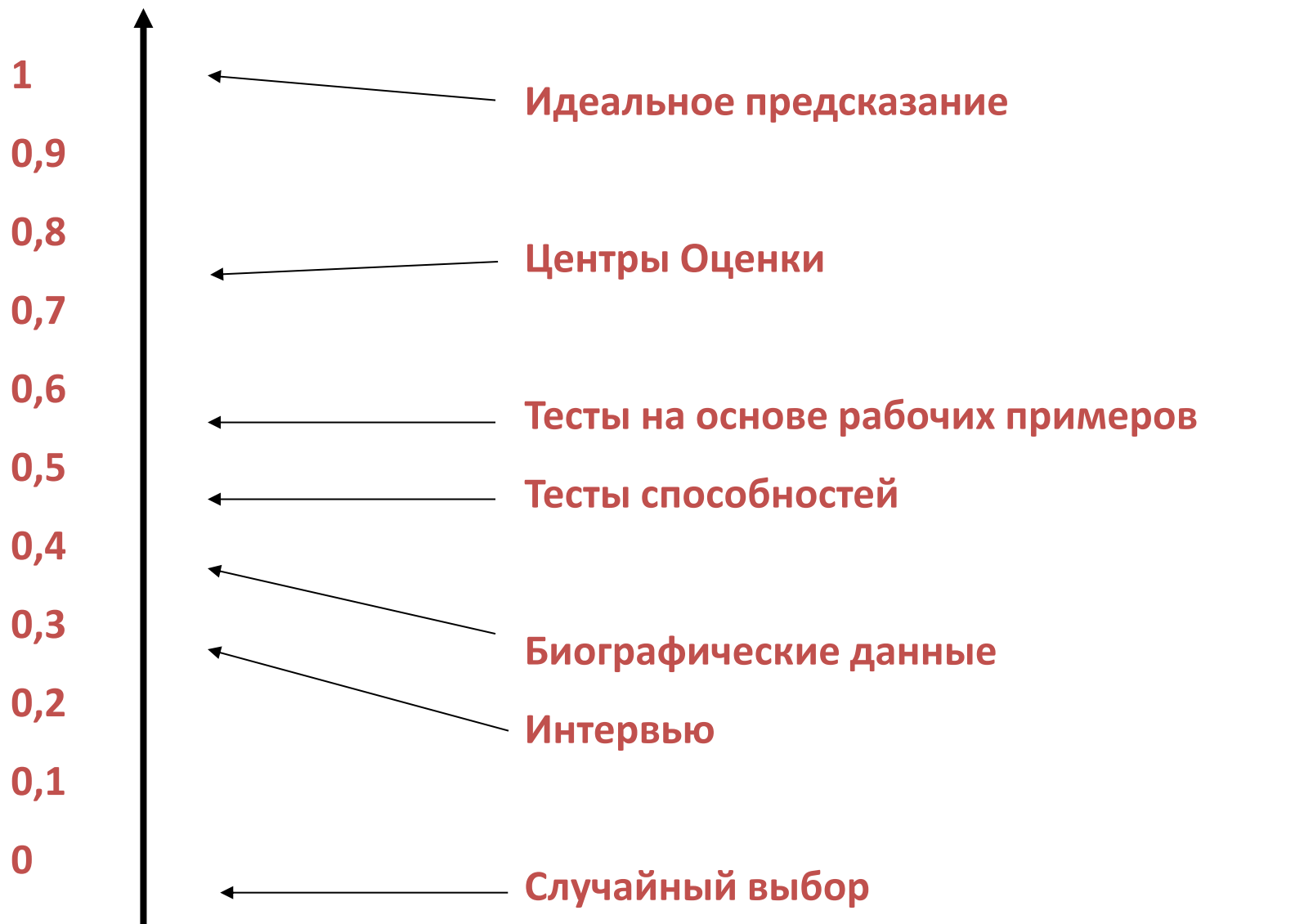
**Статминимум для HR
директора
или
СКОЛЬКО СТОИТ корреляция**

[Бабушкин Эдуард](#)

edvb@yandex.ru

Про что

- ✓ Расскажу конкретный кейс, все необходимые знания будут фоном
- ✓ Начну с предыдущих серий



6 критичных факторов:

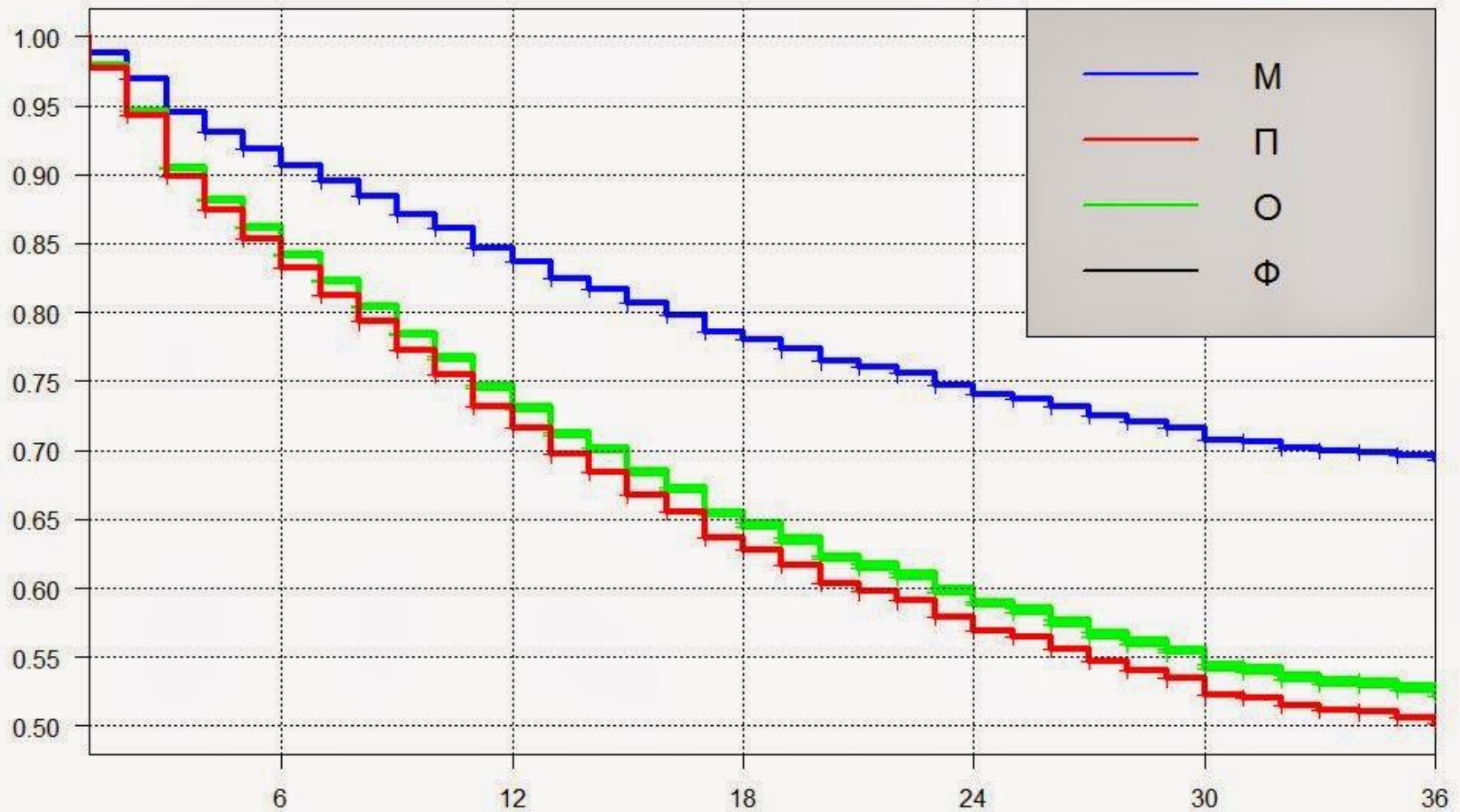
1. Отсутствие грамматических ошибок в резюме
2. Кандидат не бросал школы до окончания цикла обучения
3. Реальный опыт продаж в недвижимости или авто
4. Успешность на предыдущем месте работы
5. Способность работать в условиях неопределенности
6. Работа в условиях многозадачности

Что не имело значения

- Какой ВУЗ закончили
- Какие оценки имели
- Качество рекомендаций

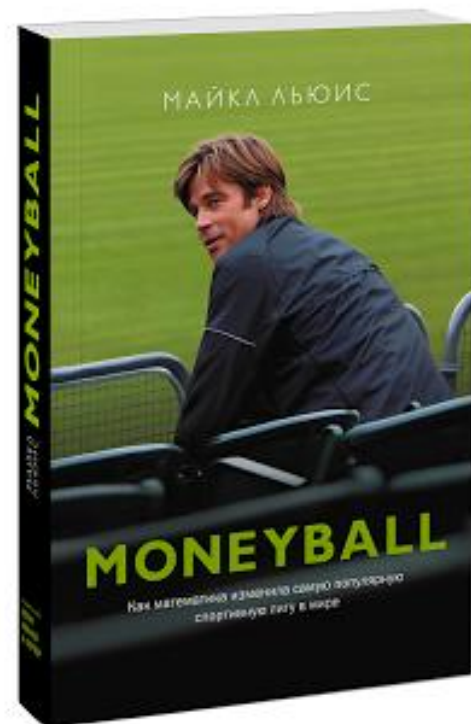
За 6 месяцев
Работы компания
получила
6 миллионов
долларов
дополнительной
прибыли

Домашнее задание

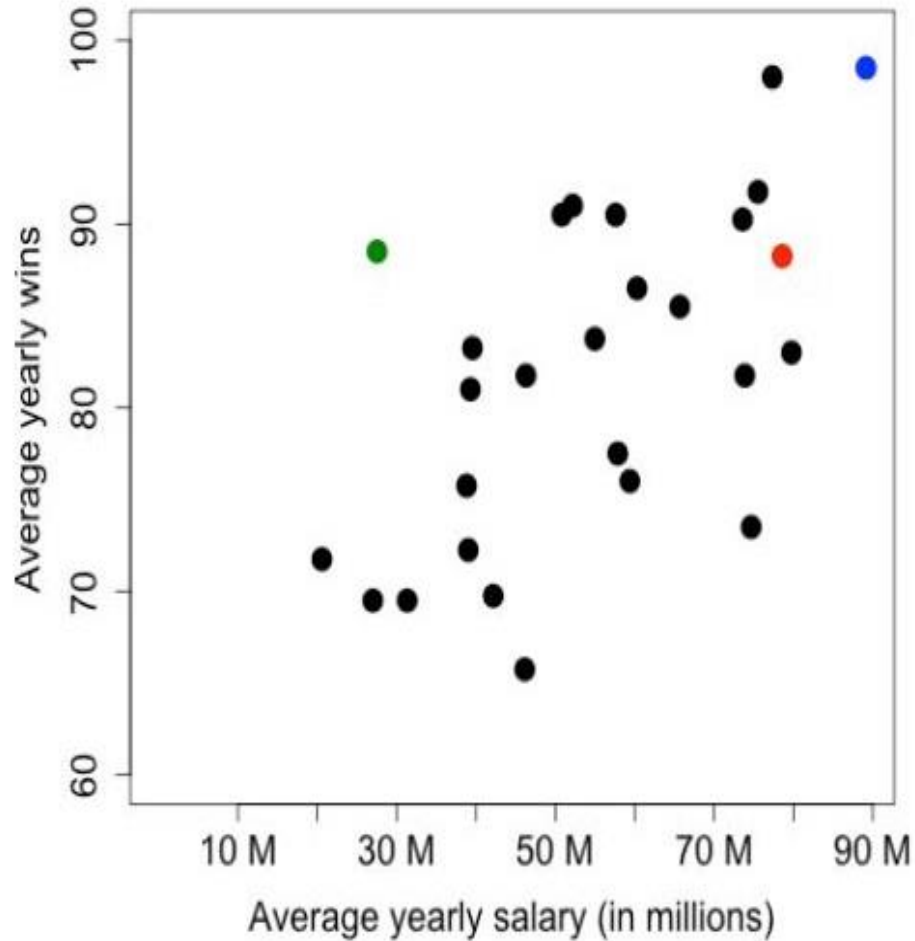


Все началось....

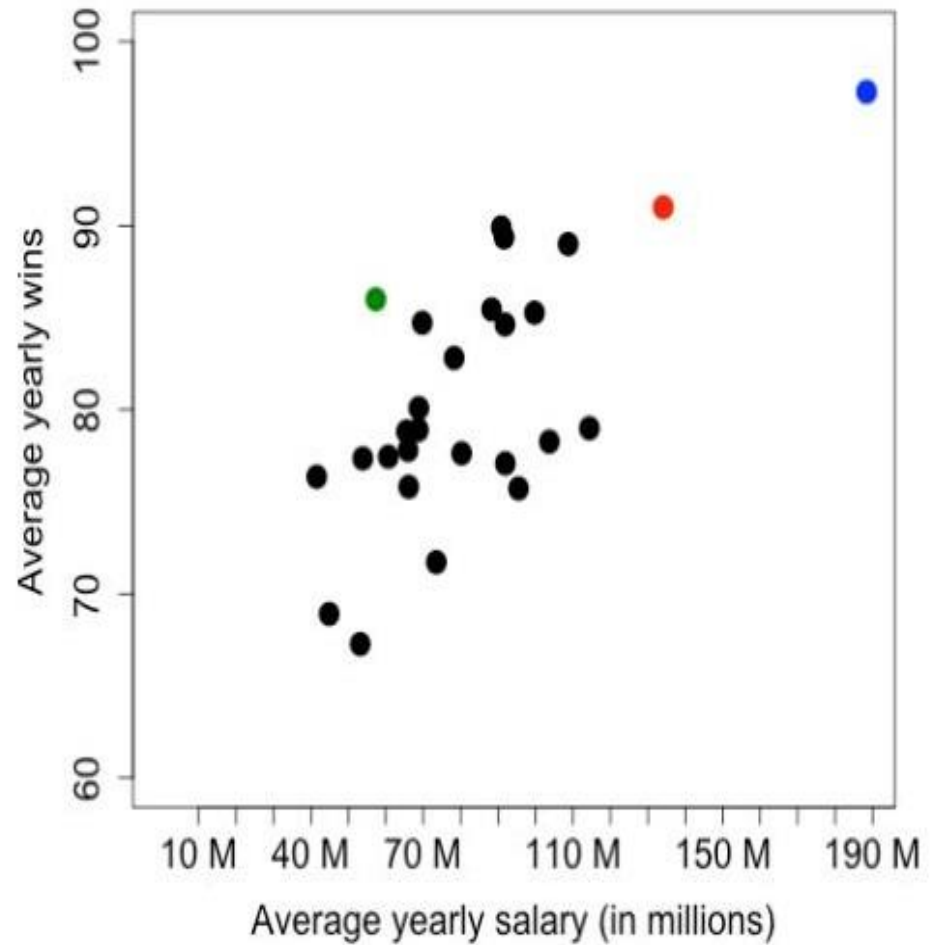
Большие Данные



Salaries and Wins, 1998-2001



Salaries and Wins, 2002-2012



Философский слайд

Цель HR – обеспечивать бизнес персоналом, который 1) эффективен и 2) работает столько, сколько надо бизнесу

Задача аналитика – выявлять факторы / драйверы, которые влияют на эффективность и текучесть.

Задачи Больших Данных

- ✓ **Удержание сотрудников** – что создает высокий уровень вовлечения и удержания?
- ✓ **Показатели продаж** – какие факторы лежат в основе успешности продавцов?
- ✓ **Аварийные ситуации** – какие факторы и какие характеристики людей приводят к аварийным ситуациям?
- ✓ **Анализ трафика руководителей**– кто будет наиболее успешным руководителем, и почему одни могут развивать свои способности, а другие – нет?
- ✓ **Анализ причин воровства** – почему в некоторых филиалах более высокий уровень воровства, и чем он вызывается?
- ✓ **Удержание клиентов** – какие факторы влияют на высокий уровень удовлетворенности и удержания клиентов?
- ✓ **Пробелы в планировании персонала** – где пробелы в наличии необходимого уровня руководителей на определенных должностях, и какие проблемы мы можем предсказать на ближайшие годы?
- ✓ **Анализ трафика кандидатов** – каково качество трафика кандидатов, каковы факторы привлечения и удержания наиболее успешных работников нашей компании?

[Источник](#)

Файл с данными

№	Продажи	IQ	Образование	Опыт
1	94	75	1	1
2	102	104	0	1
3	86	84	1	1
4	78	88	1	0
5	98	89	0	0
6	80	88	1	1
7	84	95	0	1
8	92	91	1	1
9	98	98	1	0
....
81	55	78	1	0

Переменные

✓ Продажи

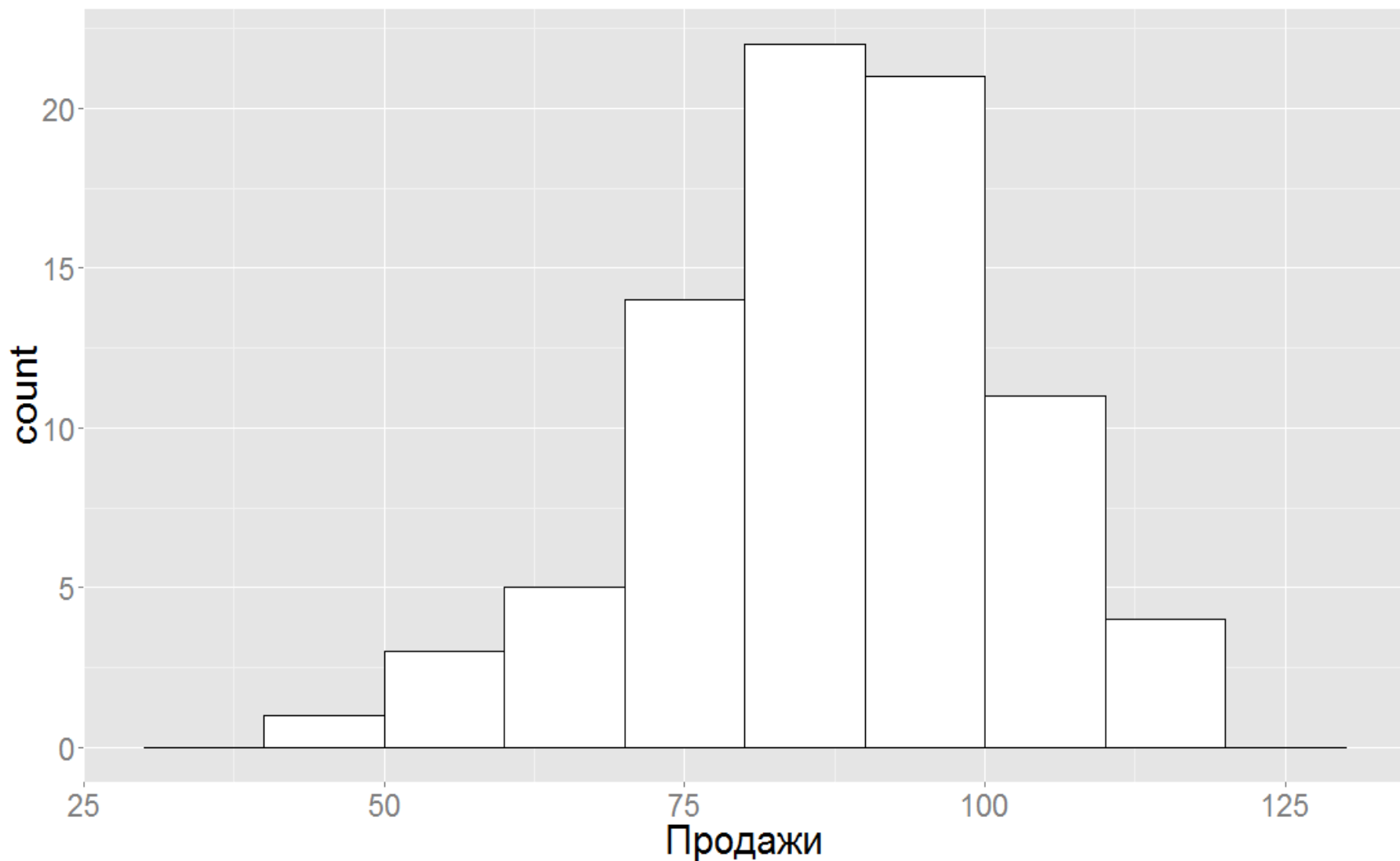
✓ IQ

✓ Образование

✓ Наличие опыта

* Узнаем, что такое тип переменной

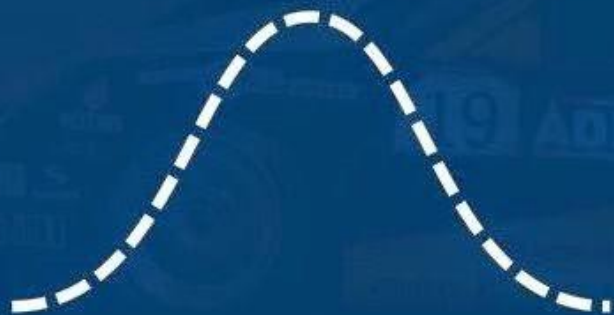
Распределение переменной «Продажи»



Bersin по этому поводу

Driving Hyper Performance

Normal Distribution of Performance



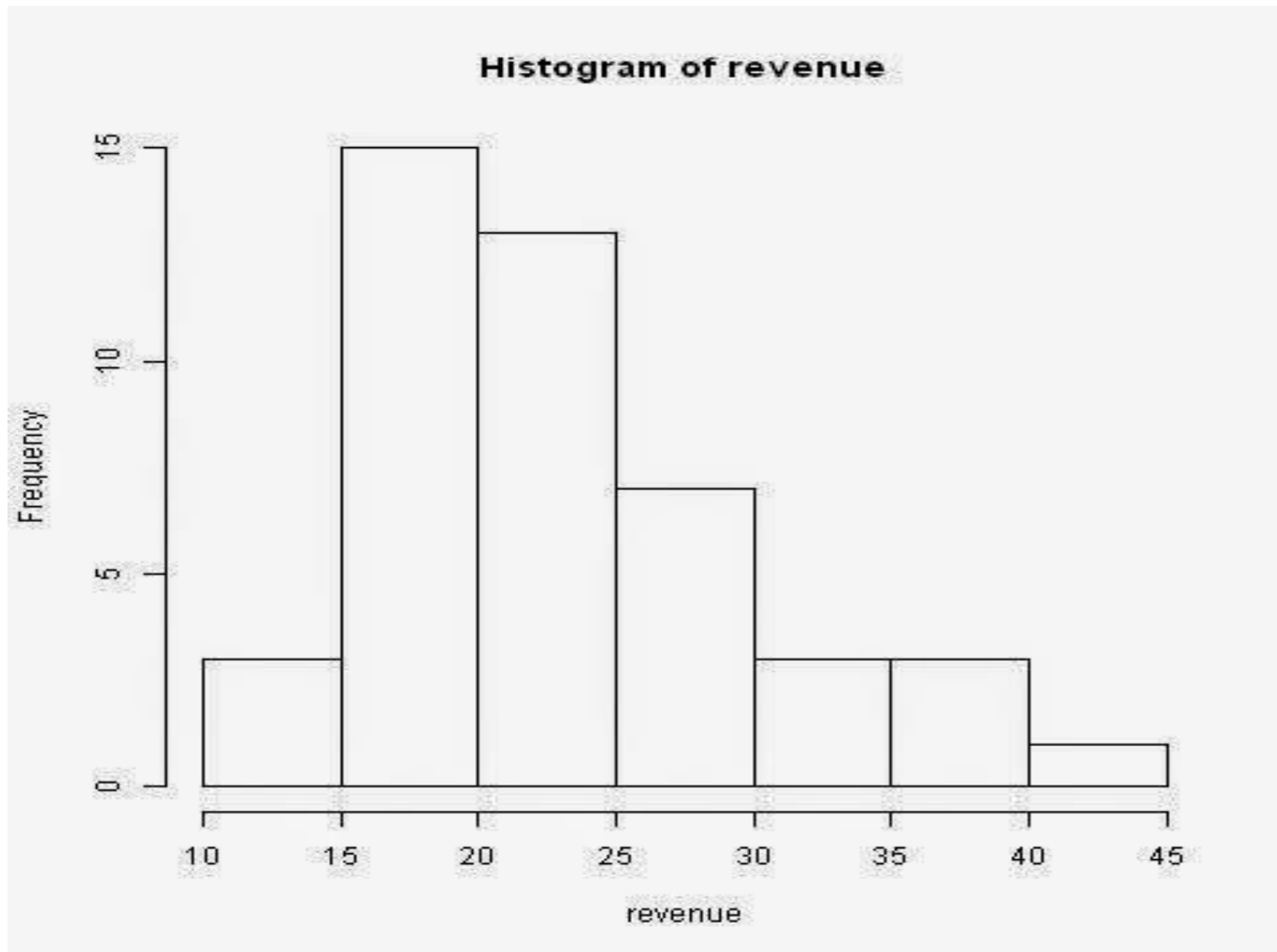
How traditional labor-driven organizations think about talent.

“Power Law” Distribution

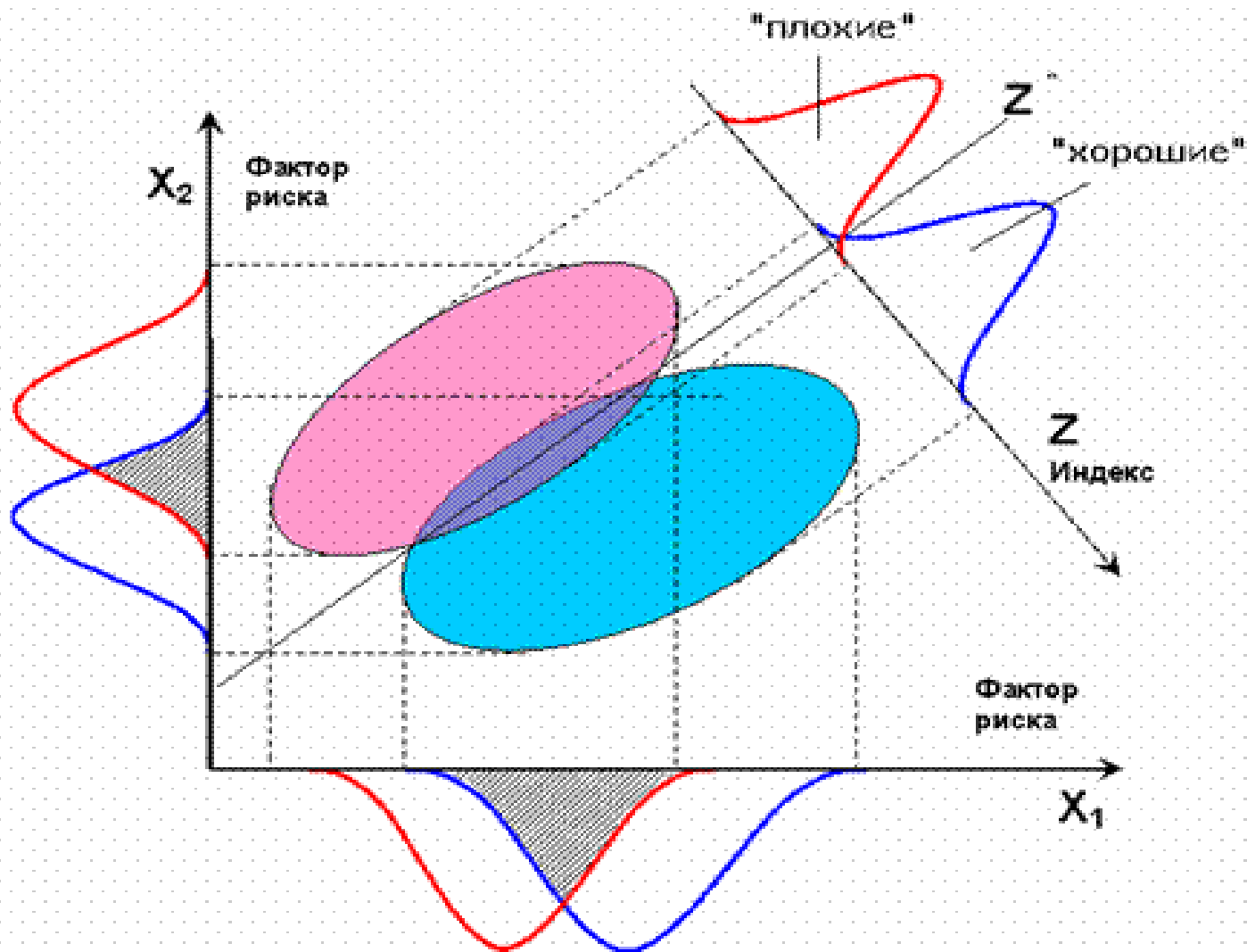


How IP, innovation, and service-driven organizations think about talent.

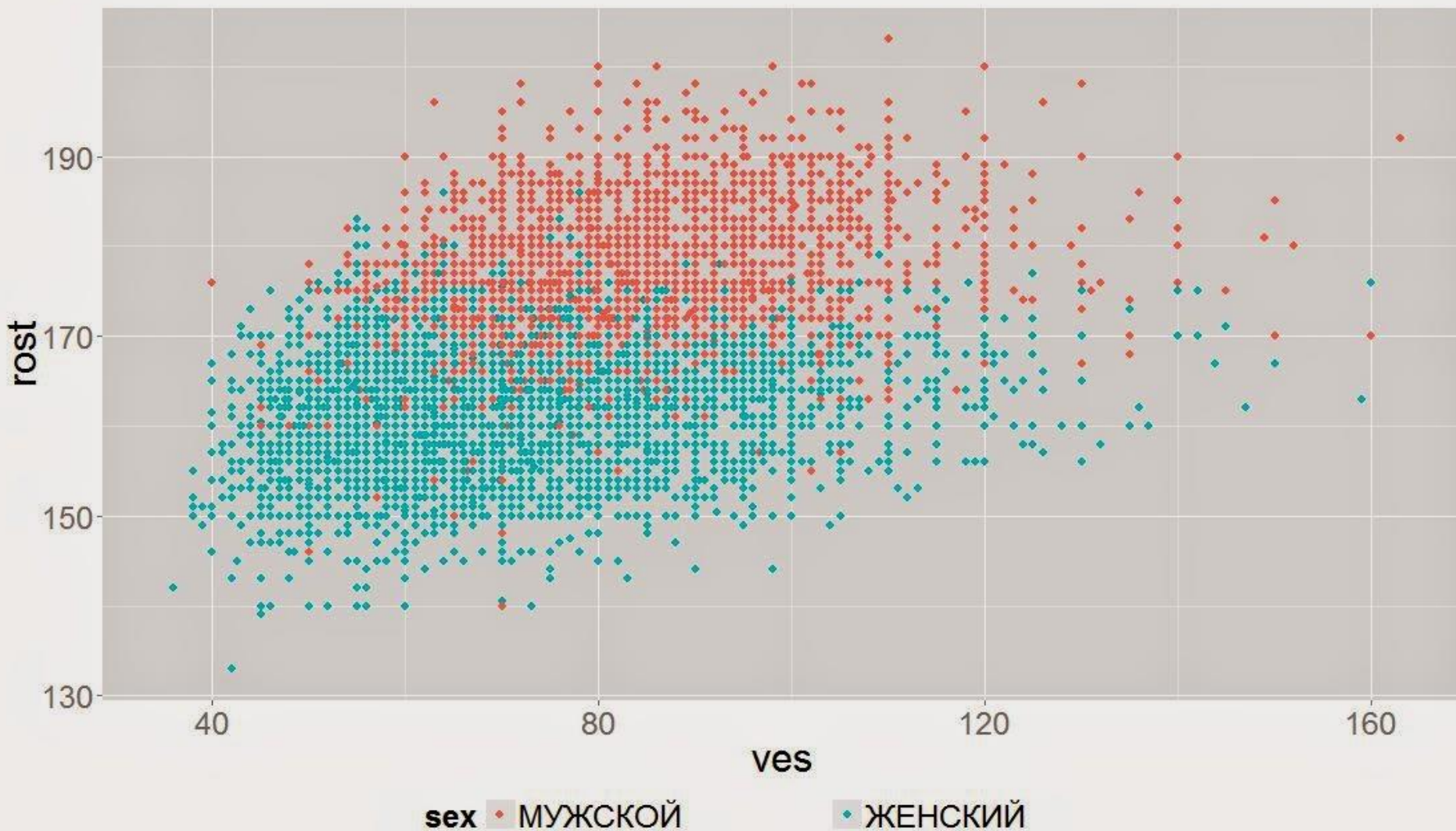
Примеры: распределение продаж



Если эффективность работника не оцифрована

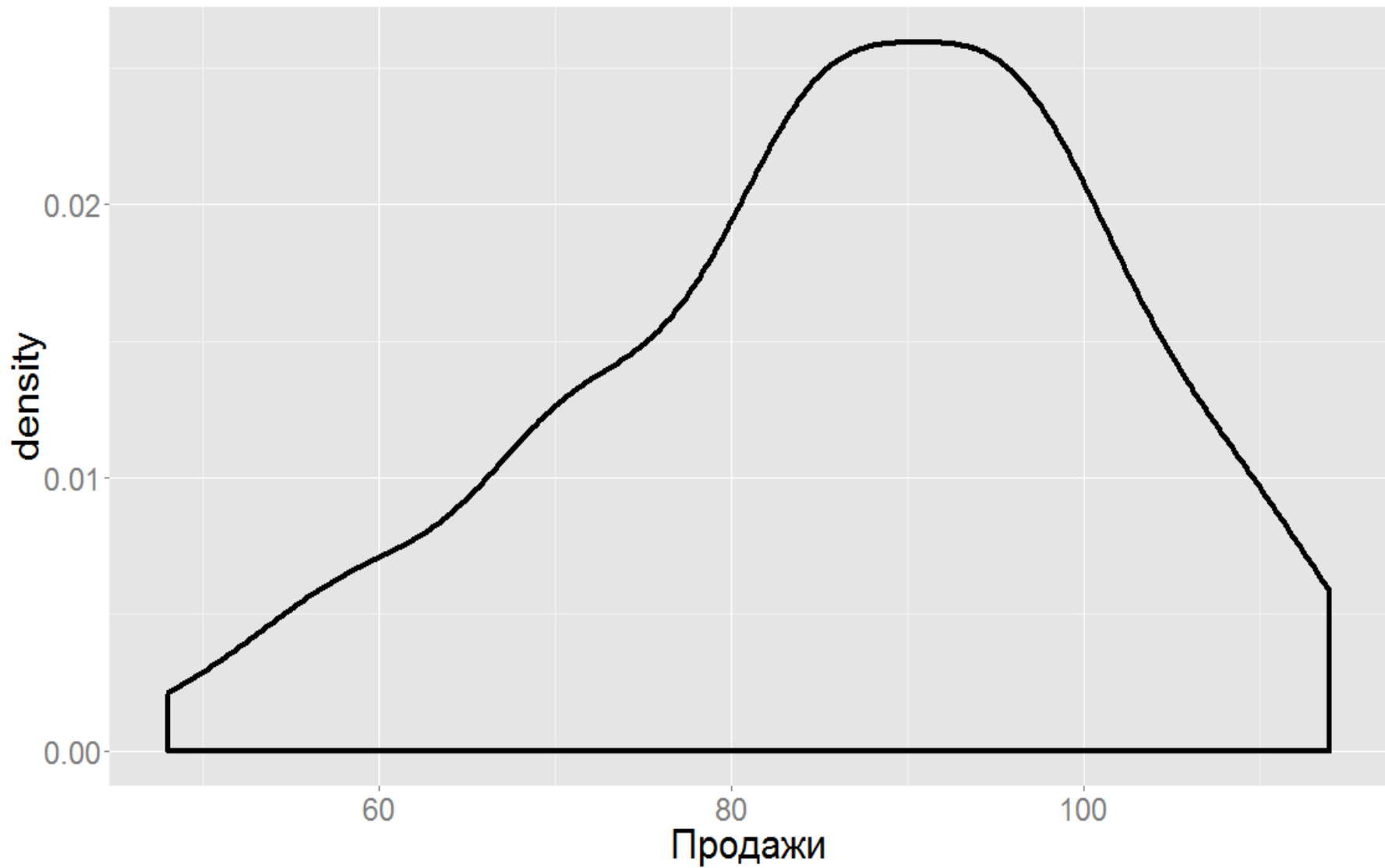


Если эффективность работника не оцифрована

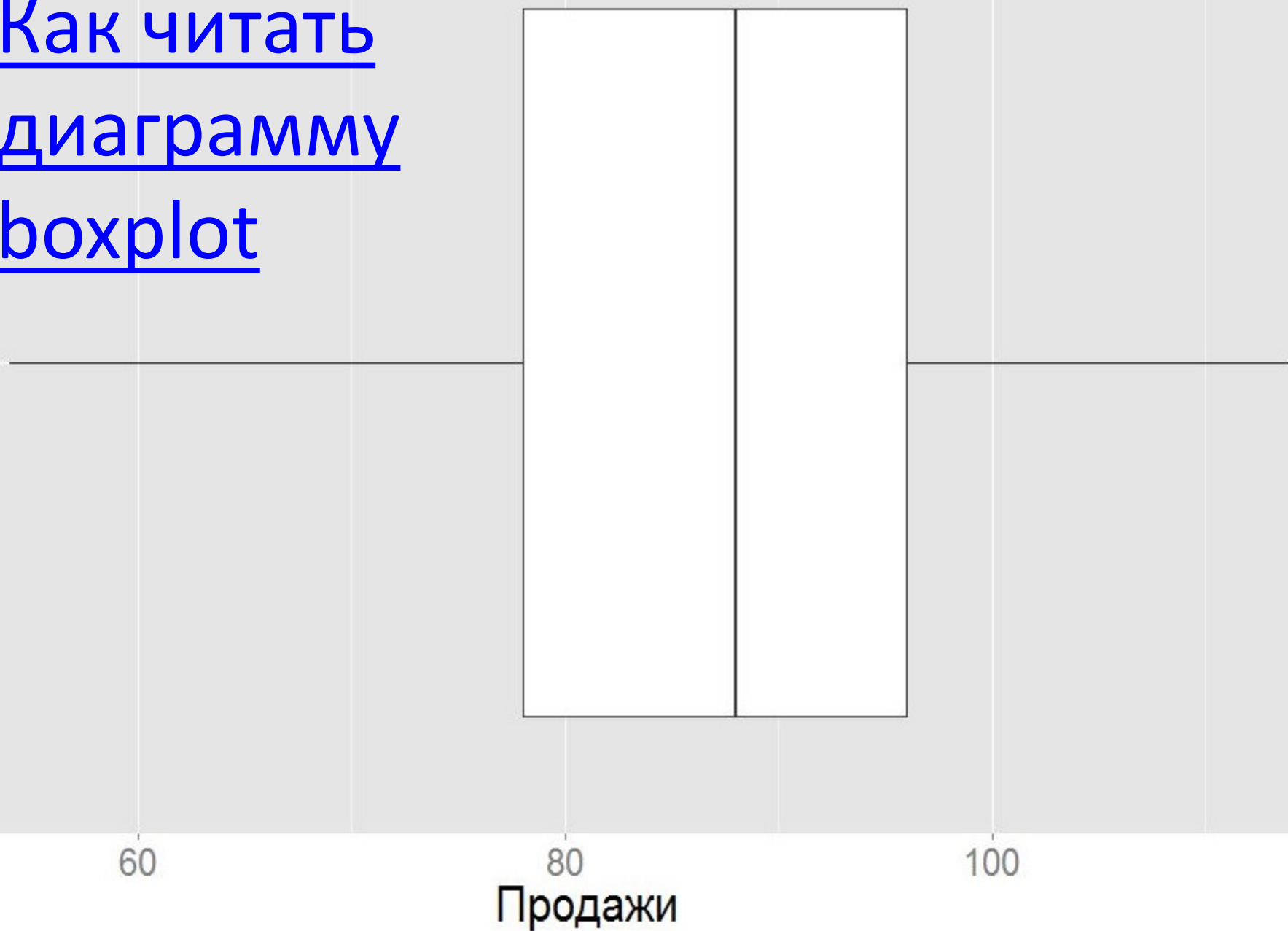


Описательные (дескриптивные) статистики

Минимум	48
1-й квартиль	78
Среднее	86.56
Медиана	88
3-й квартиль	96
Максимум	114
Эксцесс	-0.41
Ассиметрия	-0.42



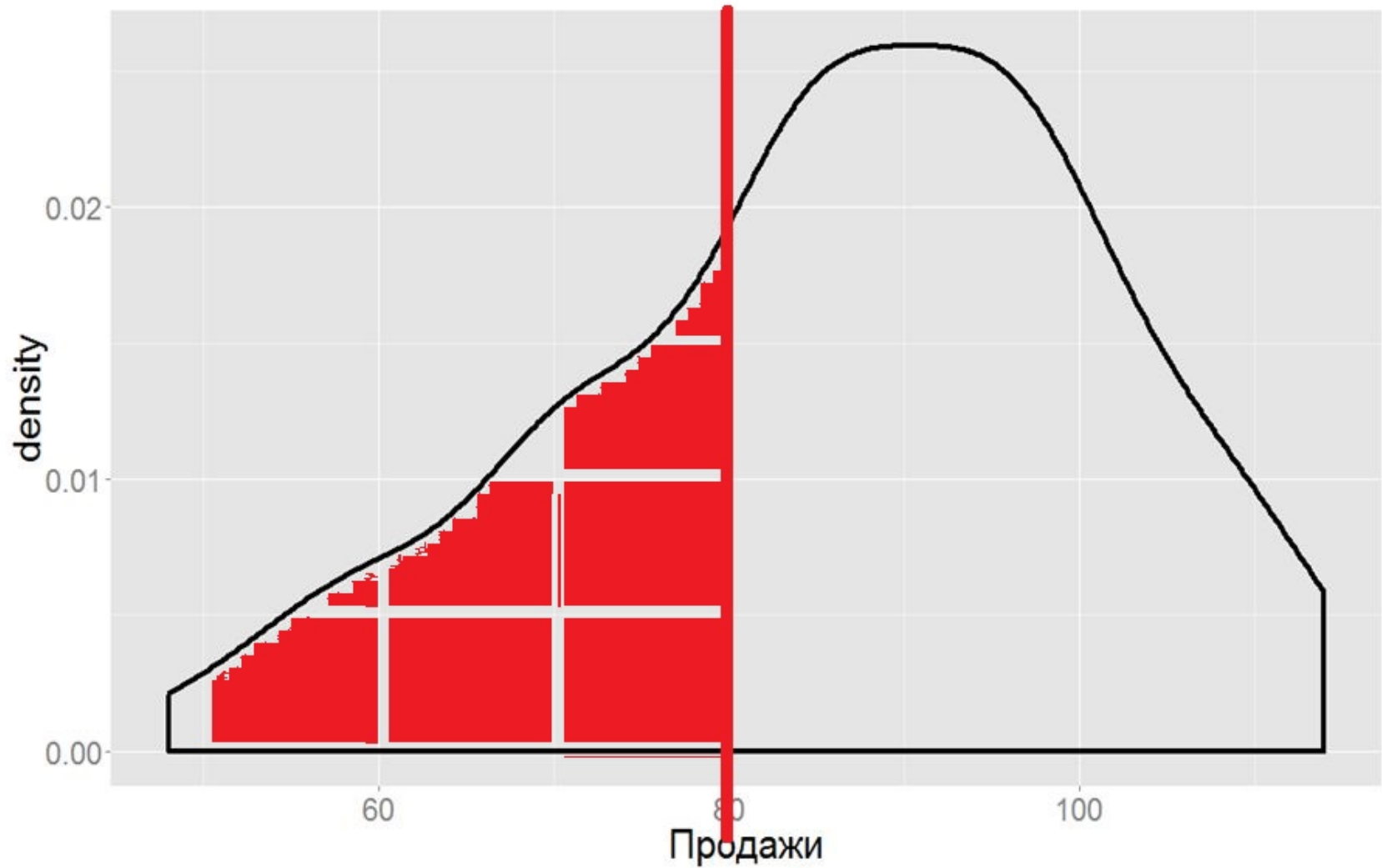
Как читать
диаграмму
boxplot

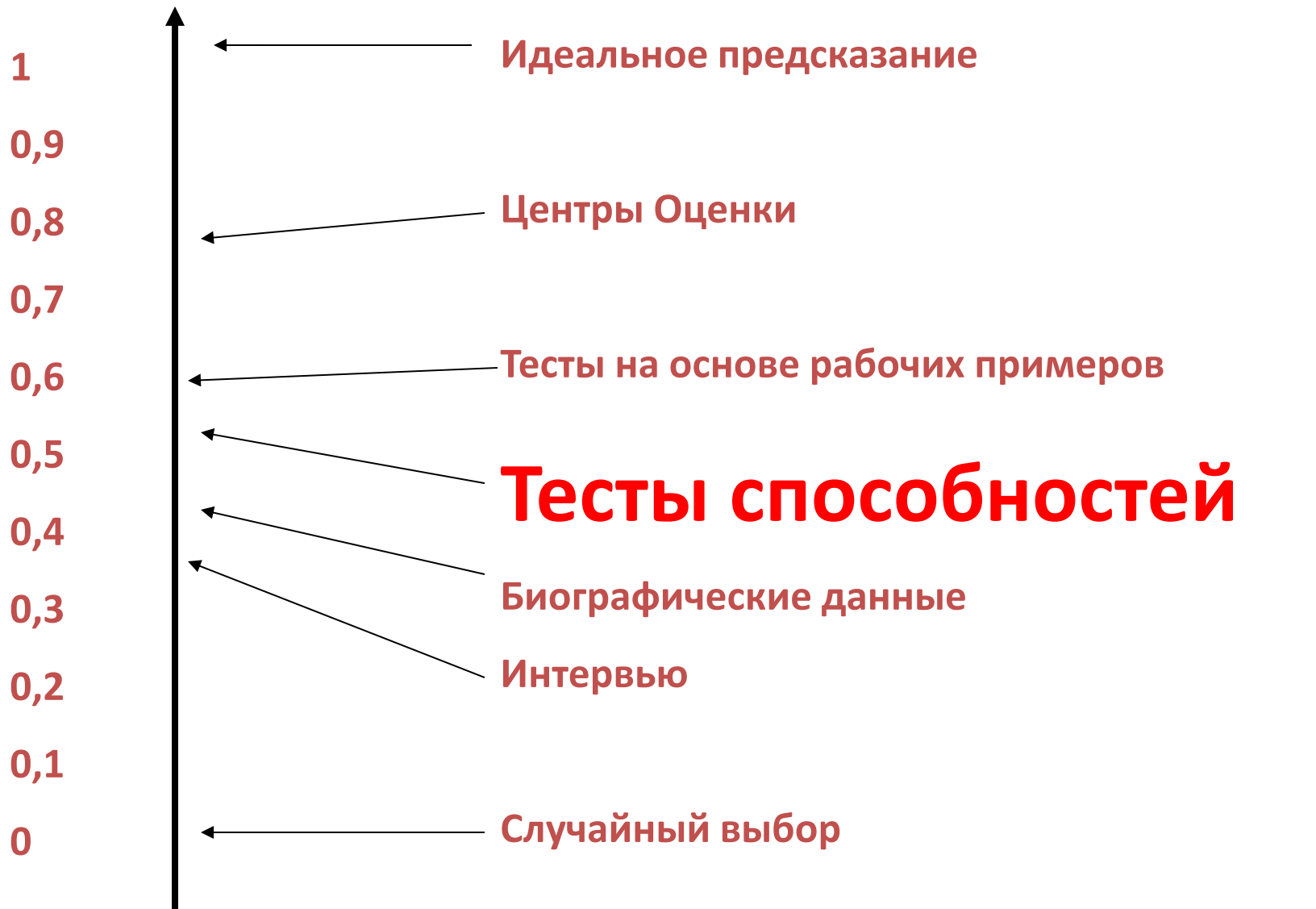


И тут Бизнес нам говорит



И тут Бизнес нам говорит

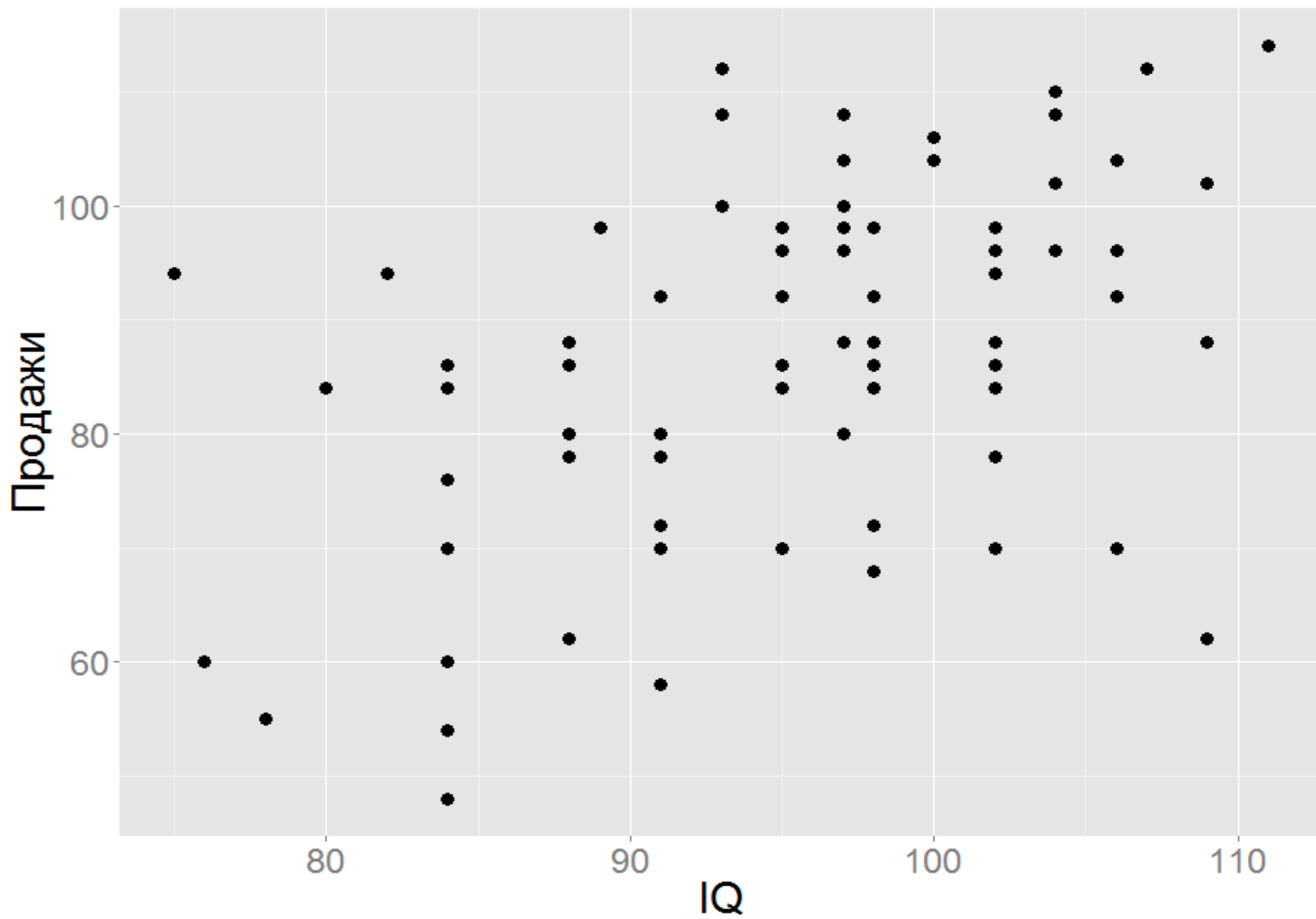




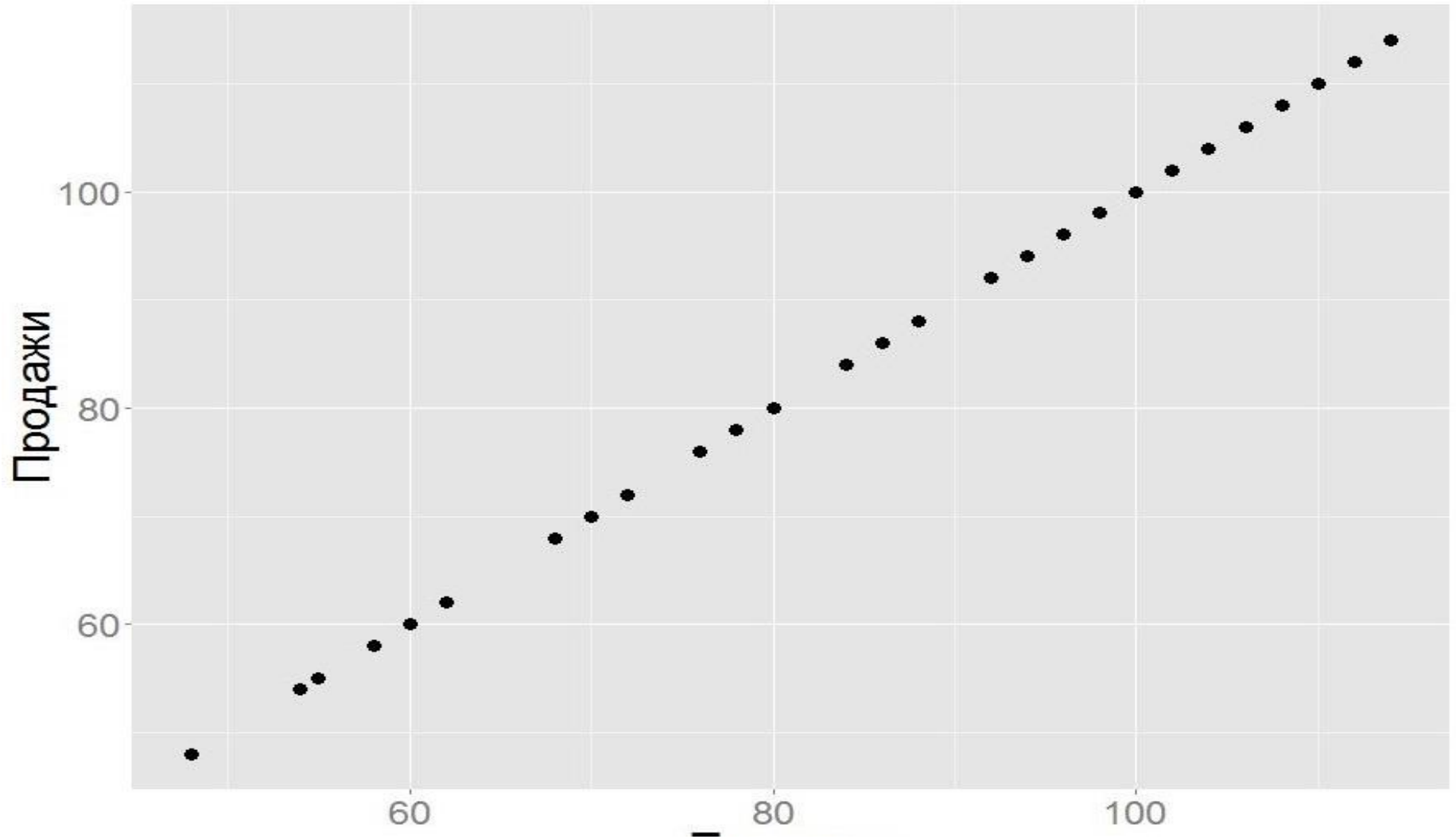
Смотрим в файл

№	Продажи	IQ	Образование	Опыт
1	94	75	1	1
2	102	104	0	1
3	86	84	1	1
4	78	88	1	0
5	98	89	0	0
6	80	88	1	1
7	84	95	0	1
8	92	91	1	1
9	98	98	1	0
....
81	55	78	1	0

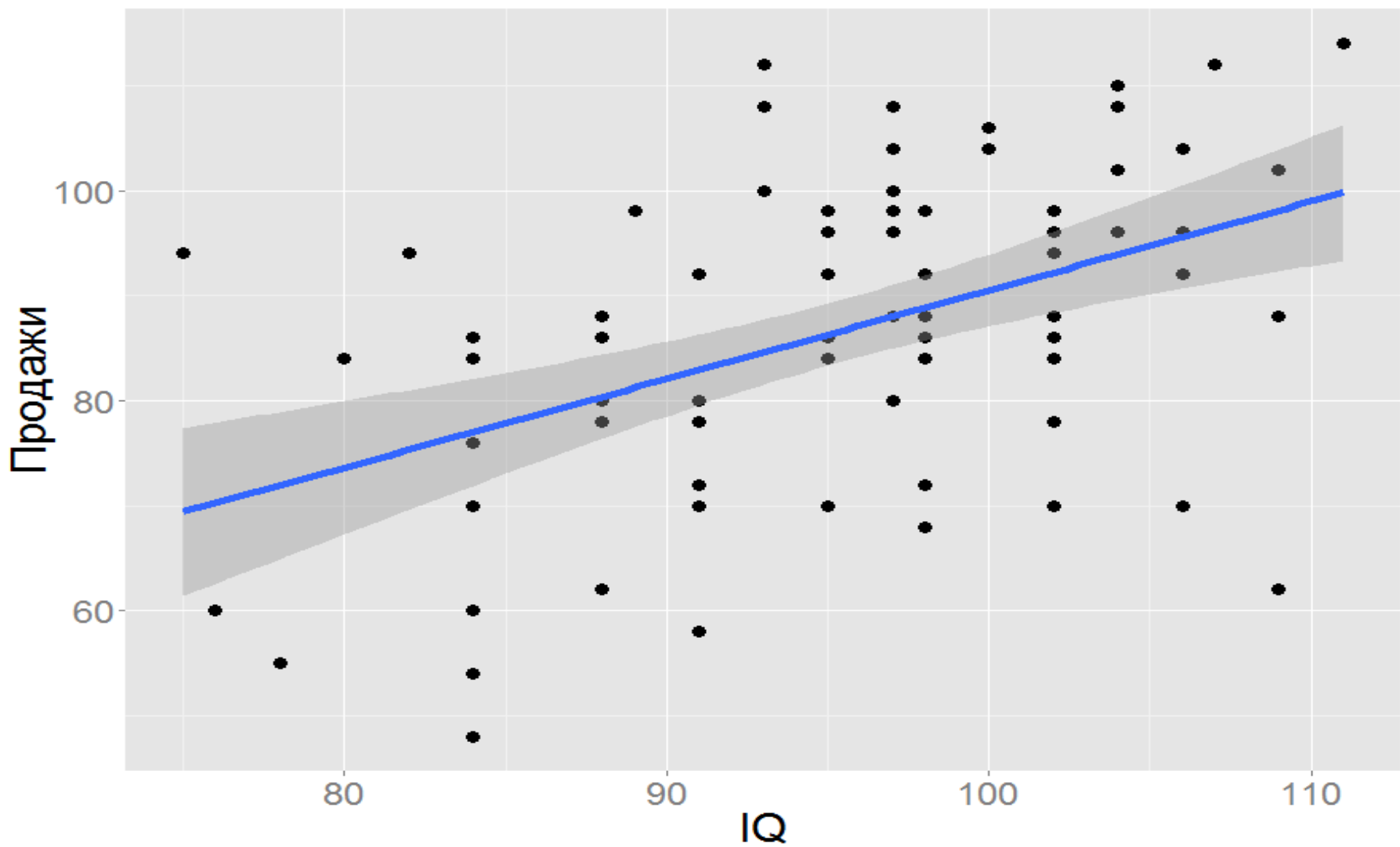
Продажи \sim IQ



Идеал ($R^2 = 1$)



Продажи \sim IQ



Продажи ~ IQ

Call:

lm(formula = Продажи ~ IQ, data = q)

Residuals:

Min	1Q	Median	3Q	Max
-36.093	-7.963	0.440	9.193	27.414

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.0742	17.5997	0.345	0.7309
IQ	0.8442	0.1840	4.589	1.65e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.39 on 79 degrees of freedom

Multiple R-squared: 0.2105, Adjusted R-squared: 0.2005

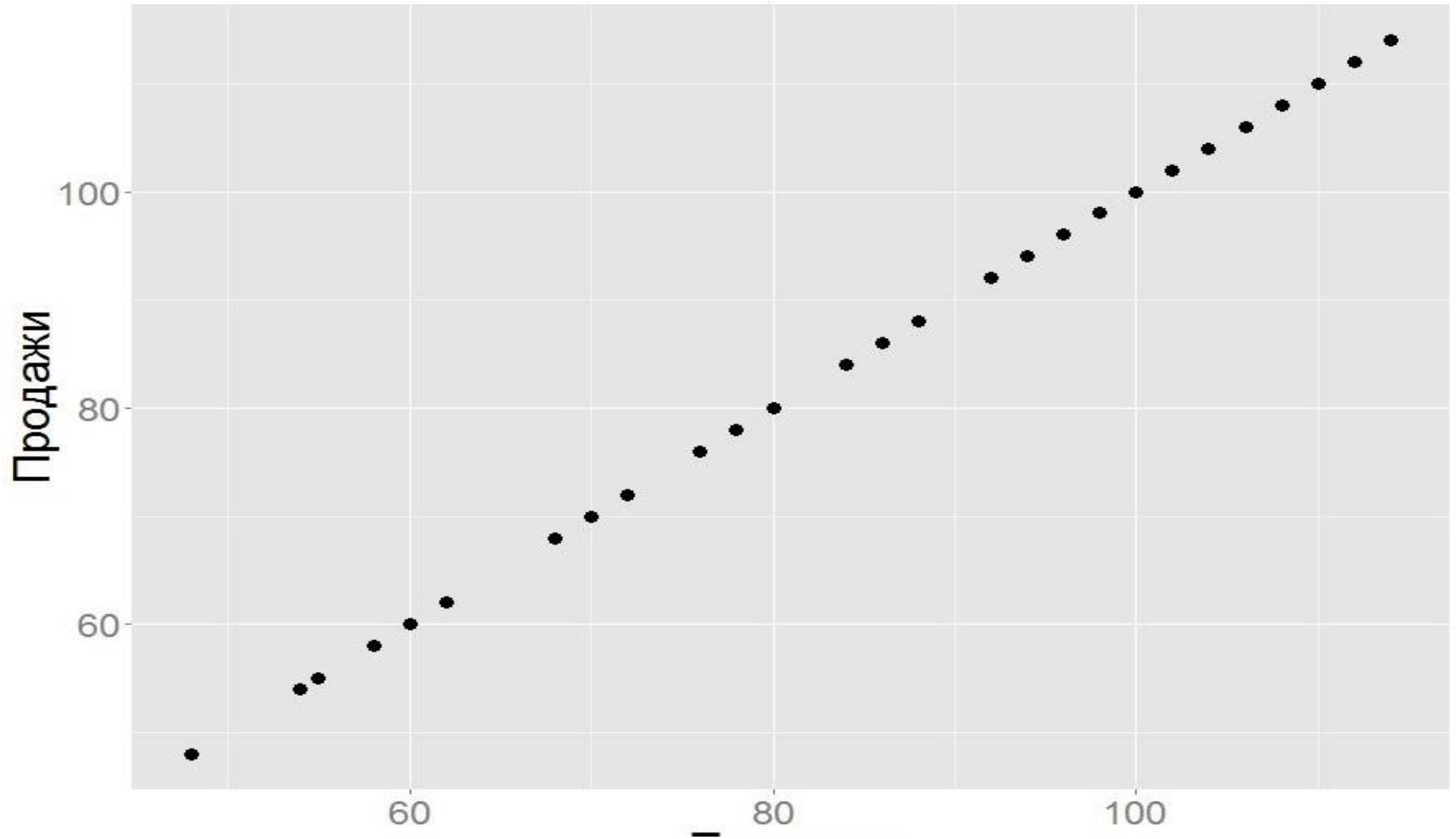
F-statistic: 21.06 on 1 and 79 DF, p-value: 1.65e-05

Продажи \sim IQ

Продажи = 0,8442 * IQ

$R^2 = 0,21$

Идеал ($R^2 = 1$)

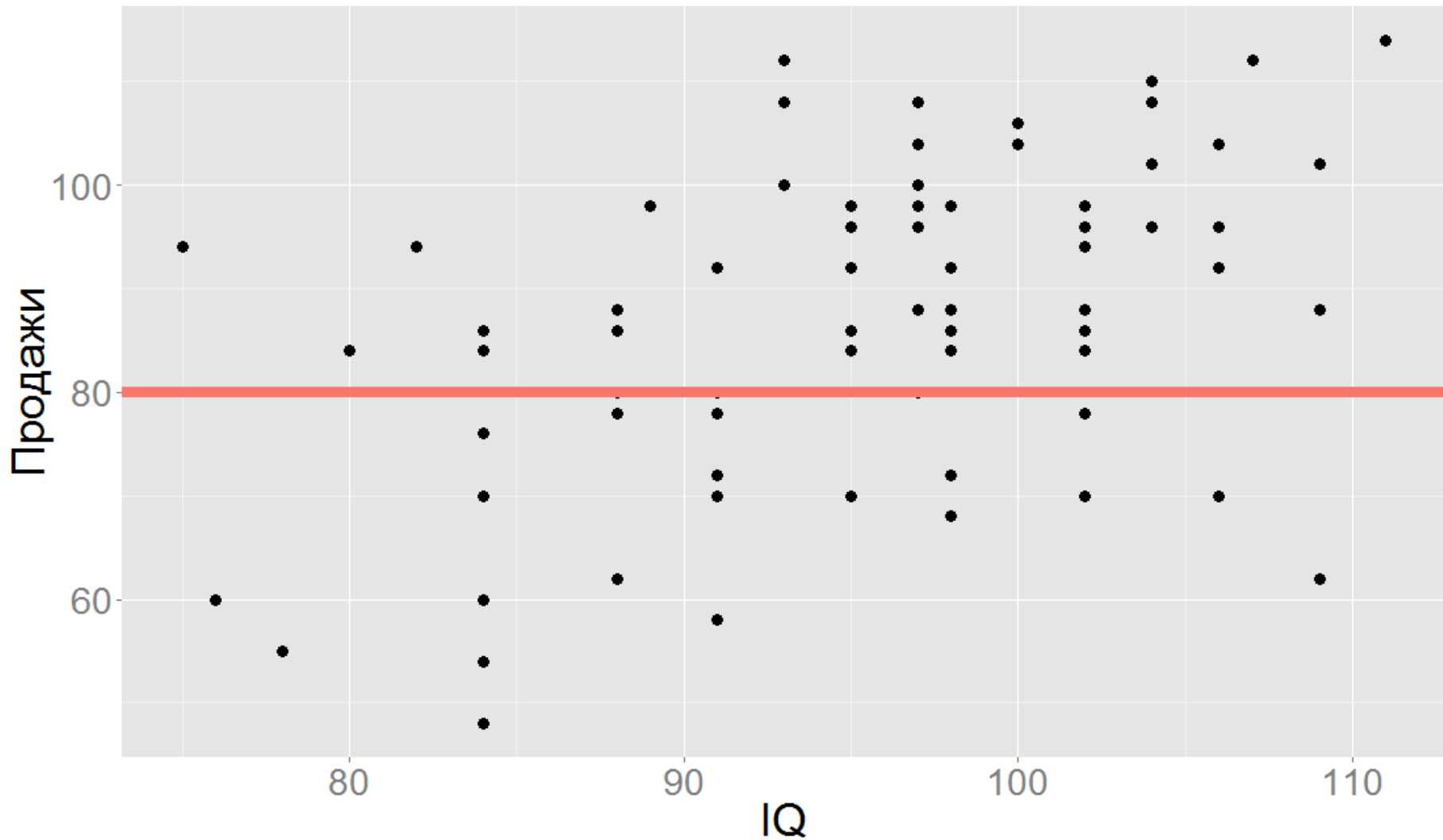


Считаем необходимый уровень IQ для продаж в 80 ед

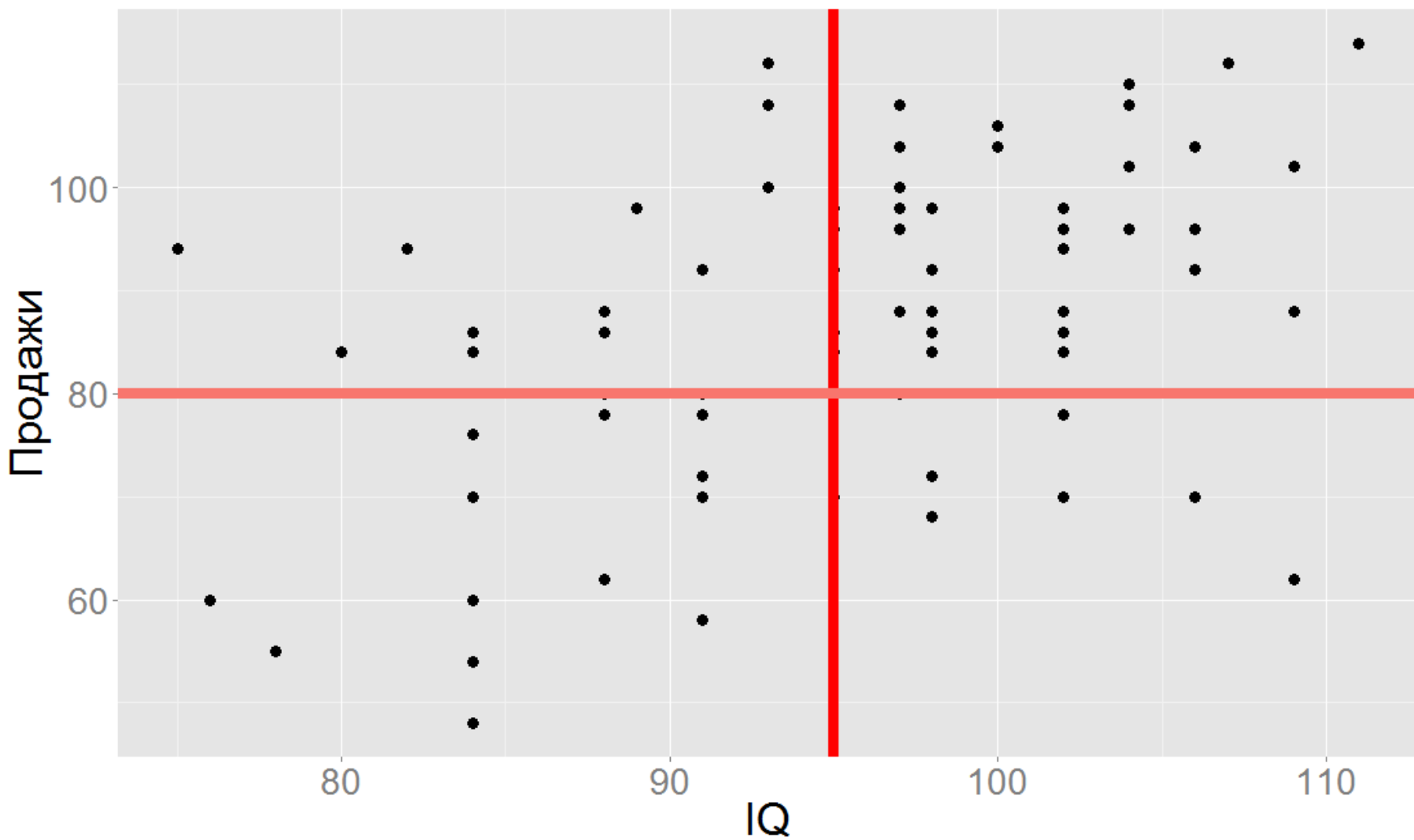
$$\text{Продажи} = 0,8442 * \text{IQ}$$

$$80 / 0,8442 = \text{IQ}$$

$$\text{IQ} = 95$$



Менее 80 = 23, более или равно 80 = 58



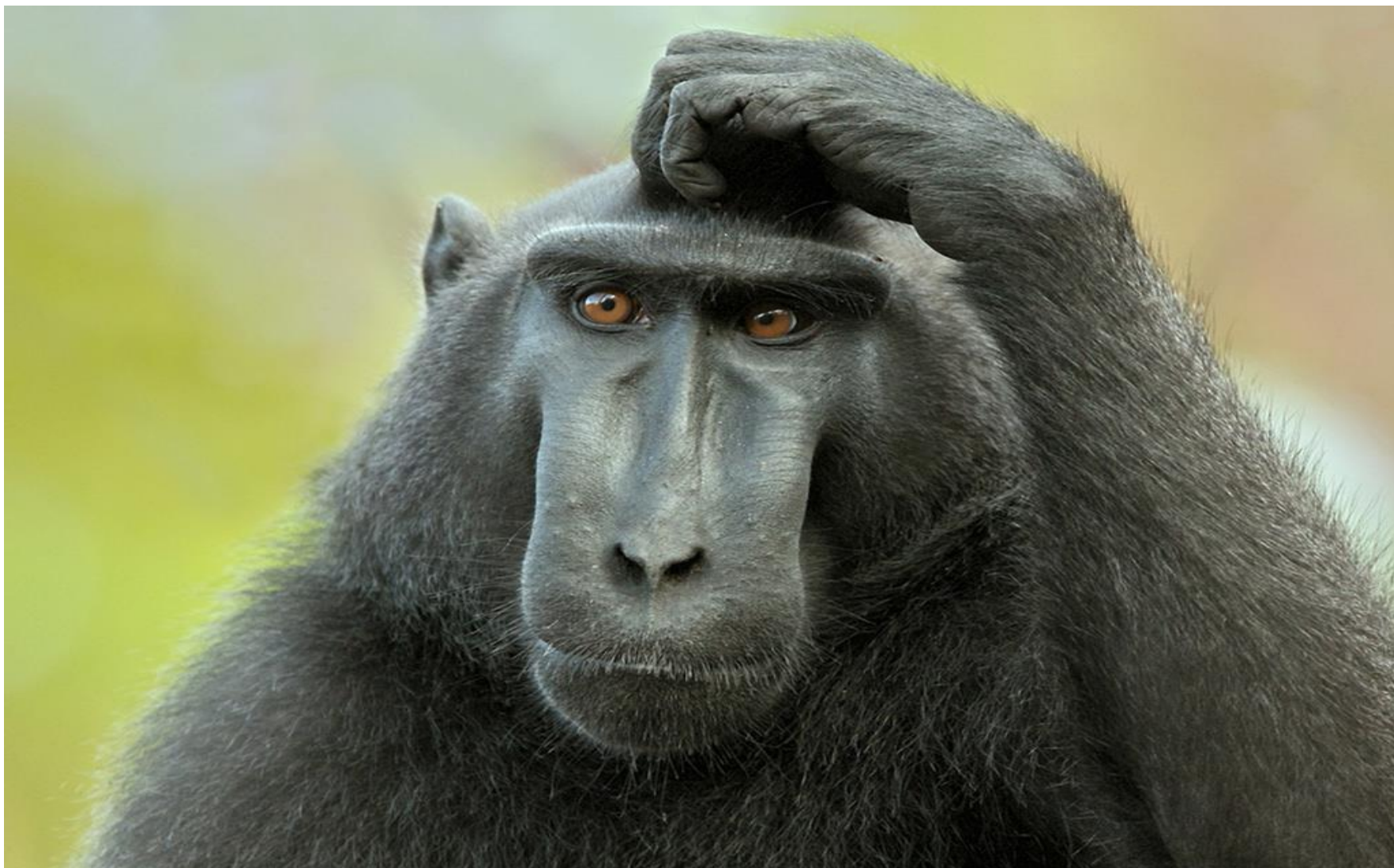
Результат (таблица сопряженности)

	IQ<95	IQ>=95	
Продажи >= 80	15	43	58
Продажи < 80	15	8	23
	30	51	81

Результат (таблица сопряженности)

	IQ<95	IQ>=95
Продажи >= 80	50 %	84 %
Продажи < 80	50%	16 %

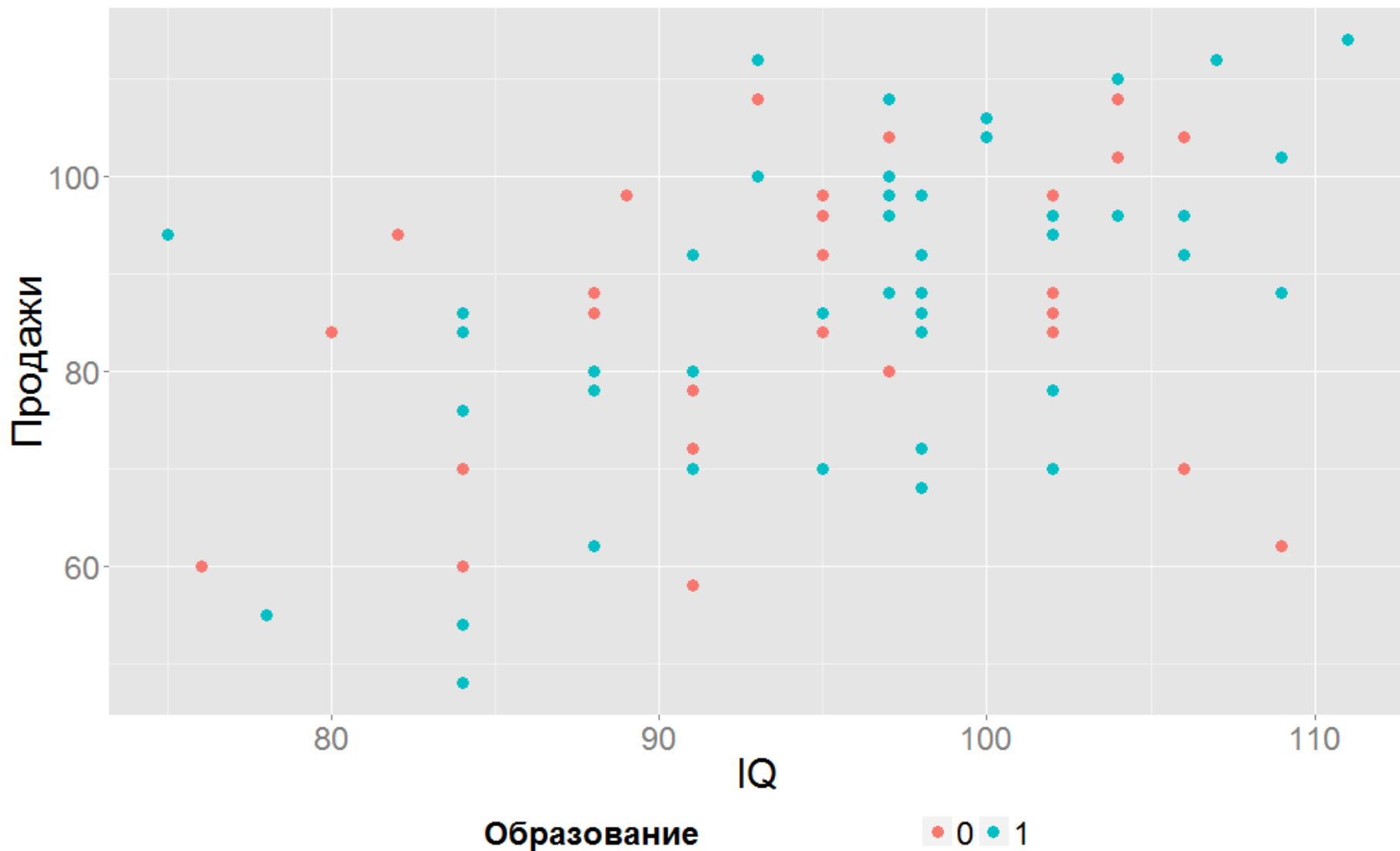
Чё делаем?



Смотрим в файл

№	Продажи	IQ	Образование	Опыт
1	94	75	1	1
2	102	104	0	1
3	86	84	1	1
4	78	88	1	0
5	98	89	0	0
6	80	88	1	1
7	84	95	0	1
8	92	91	1	1
9	98	98	1	0
....
81	55	78	1	0

Образование



Продажи ~ Образование

call:

```
lm(formula = Продажи ~ Образование, data = q)
```

Residuals:

Min	1Q	Median	3Q	Max
-39.367	-9.367	0.687	10.633	26.633

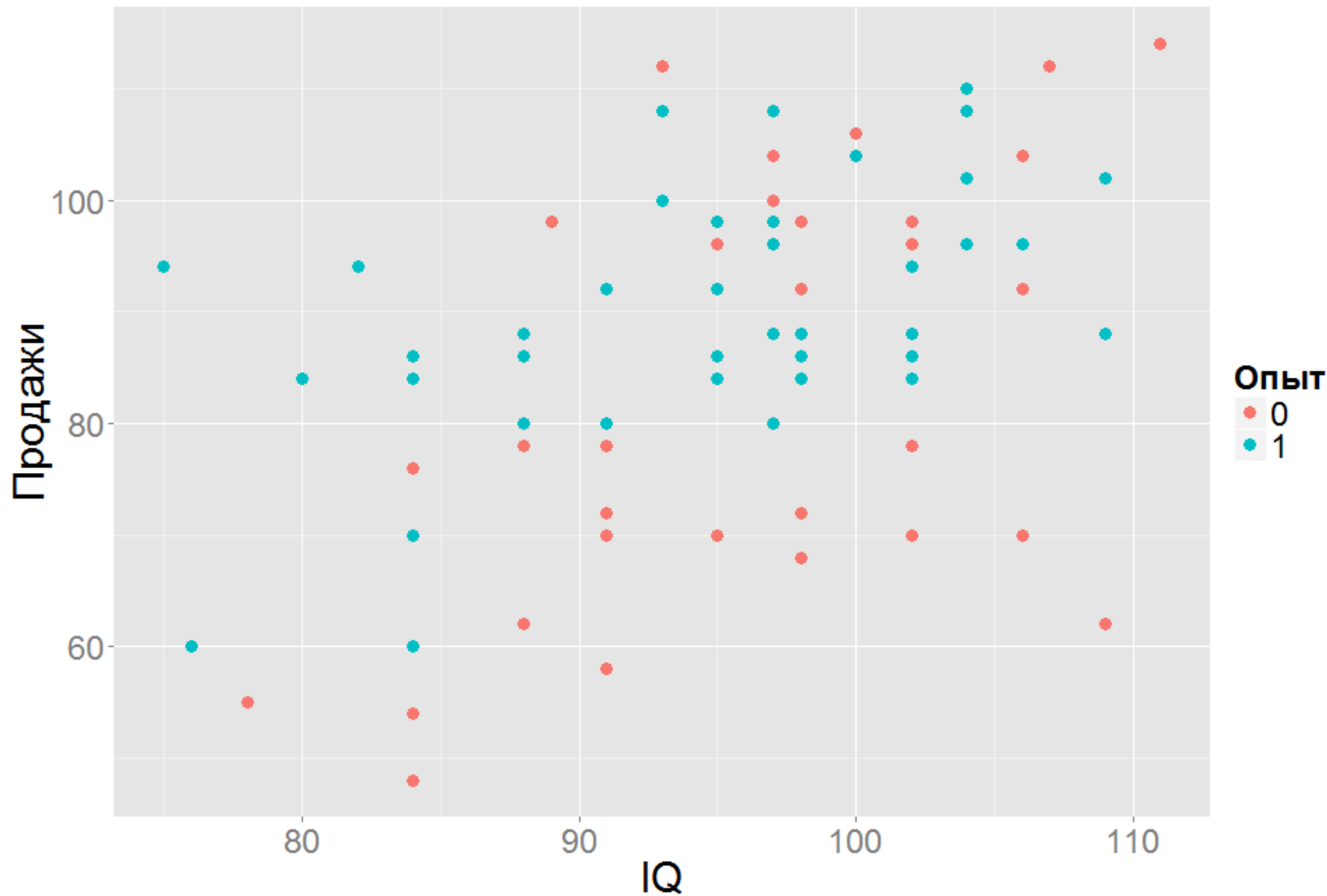
Coefficients:

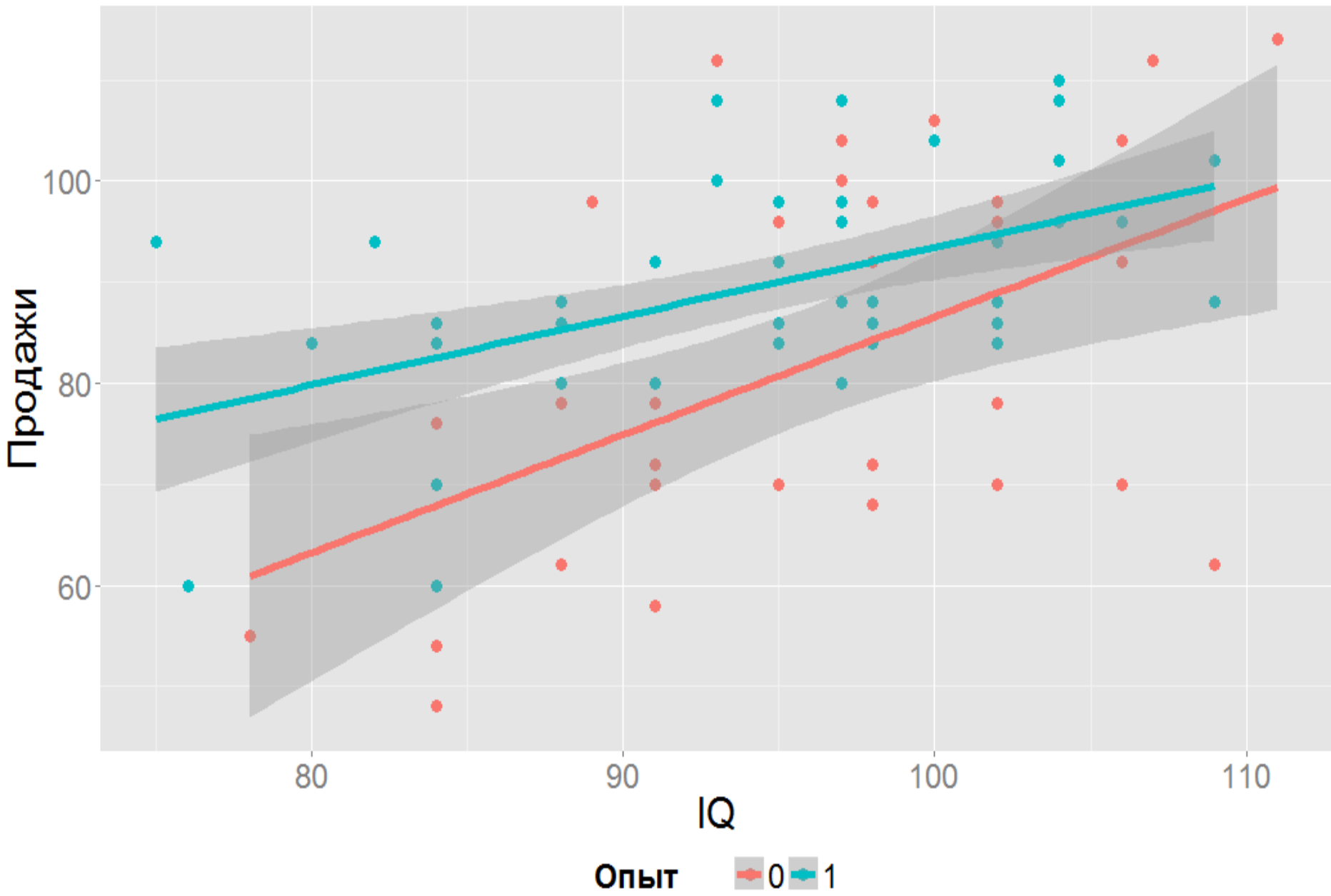
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	85.312	2.659	32.087	<2e-16 ***
Образование1	2.055	3.418	0.601	0.549

signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 17.94 on 79 degrees of freedom
Multiple R-squared: 0.004553, Adjusted R-squared: 0.008047
F-statistic: 0.3613 on 1 Df, p-value: 0.5495

Наличие опыта





Модель

Call:

lm(formula = Продажи ~ Опыт + IQ, data = q)

Residuals:

Min	1Q	Median	3Q	Max
-31.337	-8.689	-0.689	8.337	32.734

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-2.5169	16.9034	-0.149	0.8820
Опыт	9.0257	2.8642	3.151	0.00231 **
IQ	0.8794	0.1747	5.033	3.02e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 12.7 on 78 degrees of freedom

Multiple R-squared: 0.2996, Adjusted R-squared: 0.2917

F-statistic: 16.69 on 2 and 78 DF, p-value: 9.276e-07

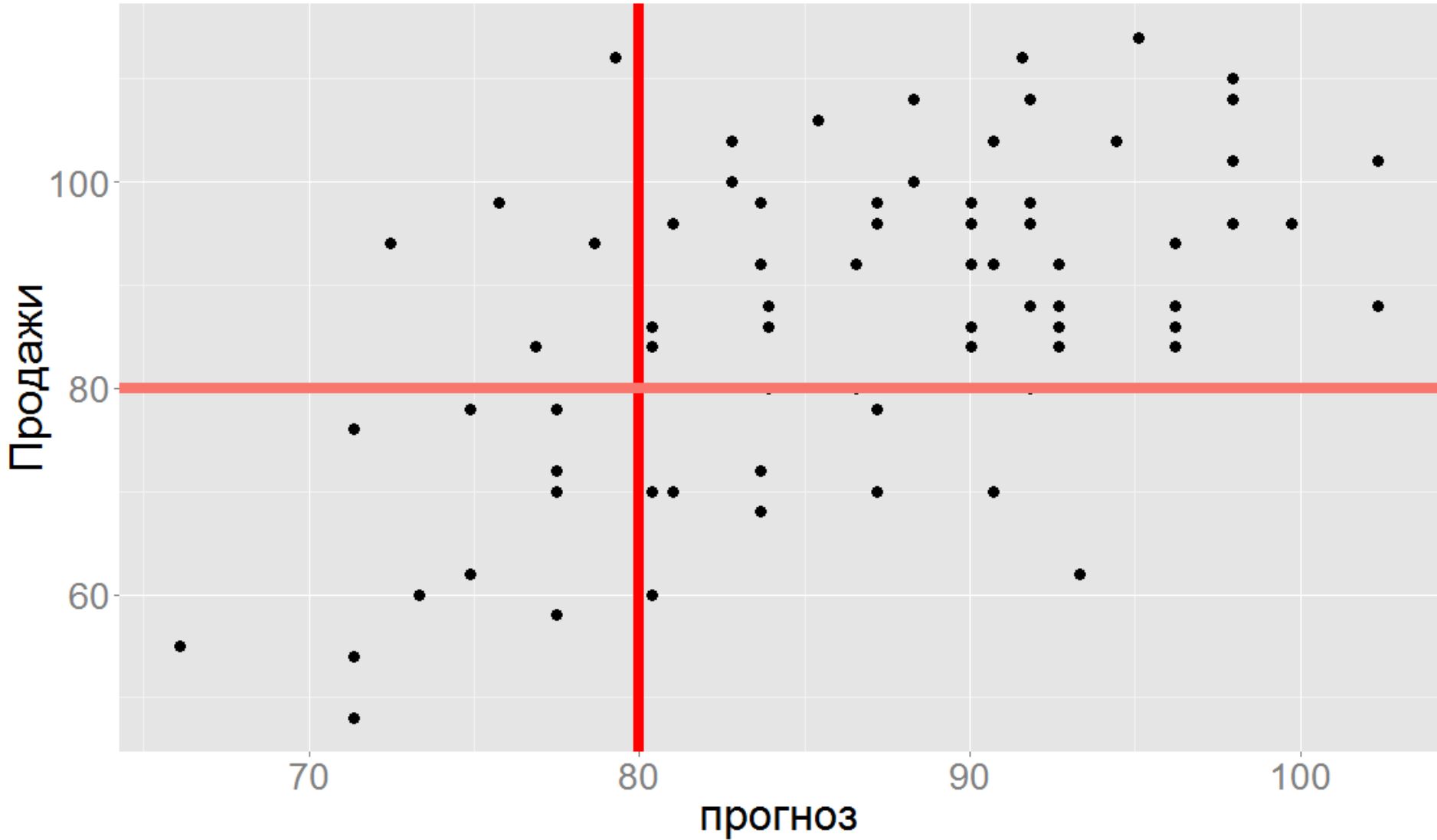
Уравнение с двумя переменными

Продажи = $0,88 \cdot (\text{IQ}) + 9,03 \cdot (\text{есть опыт})$

R^2 с одной переменной IQ = 0,21

R^2 с 2-ми переменными (IQ + опыт) = 0,299

Продажи ~ Прогноз



Прогноз на основе двух переменных

	Прогноз < 80	Прогноз ≥ 80
Продажи ≥ 80	(5) 28 %	(53) 84 %
Продажи < 80	(13) 72 %	(10) 16 %
	100 %	100 %

Дальше чё?

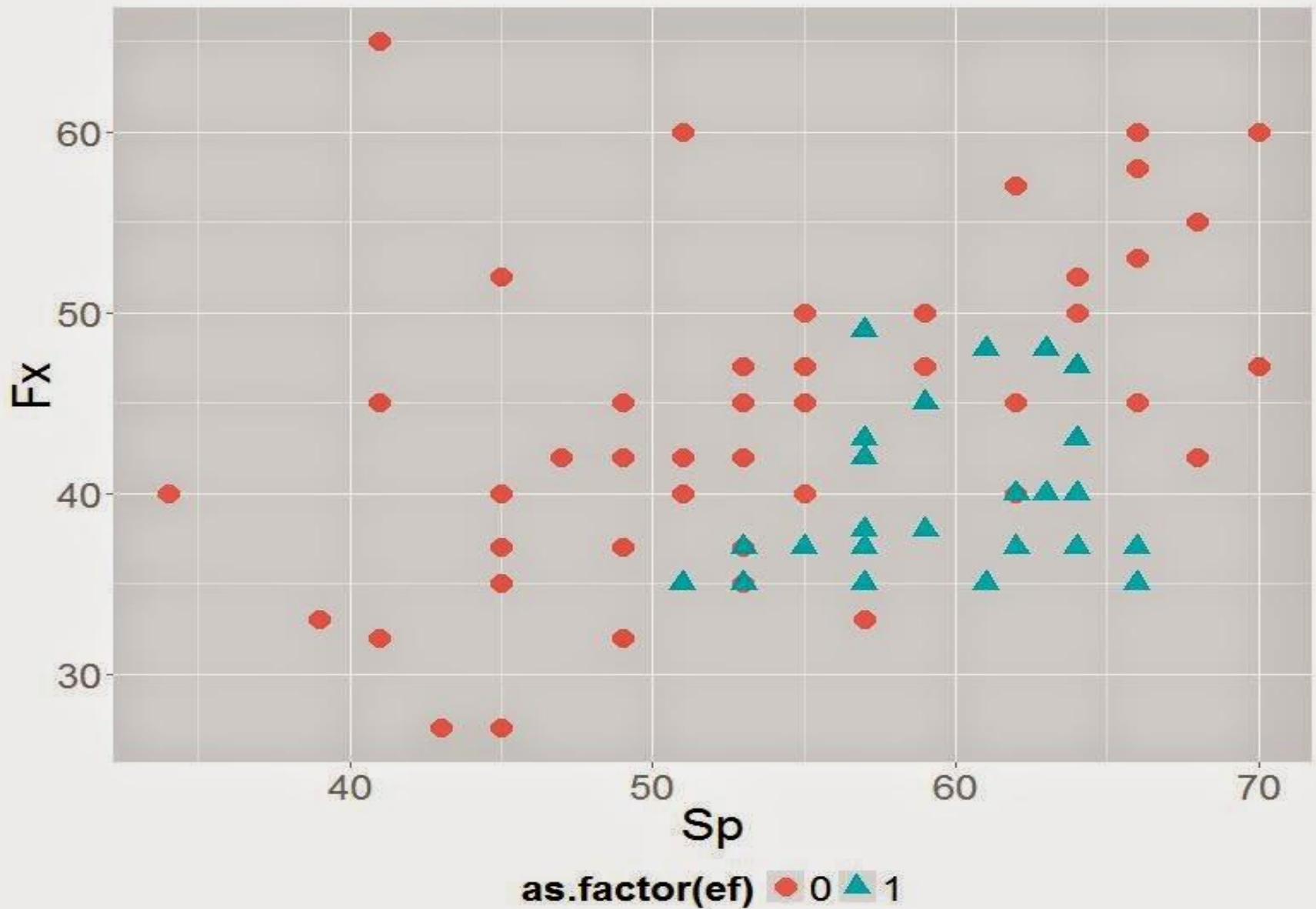
- А данных
больше
нет



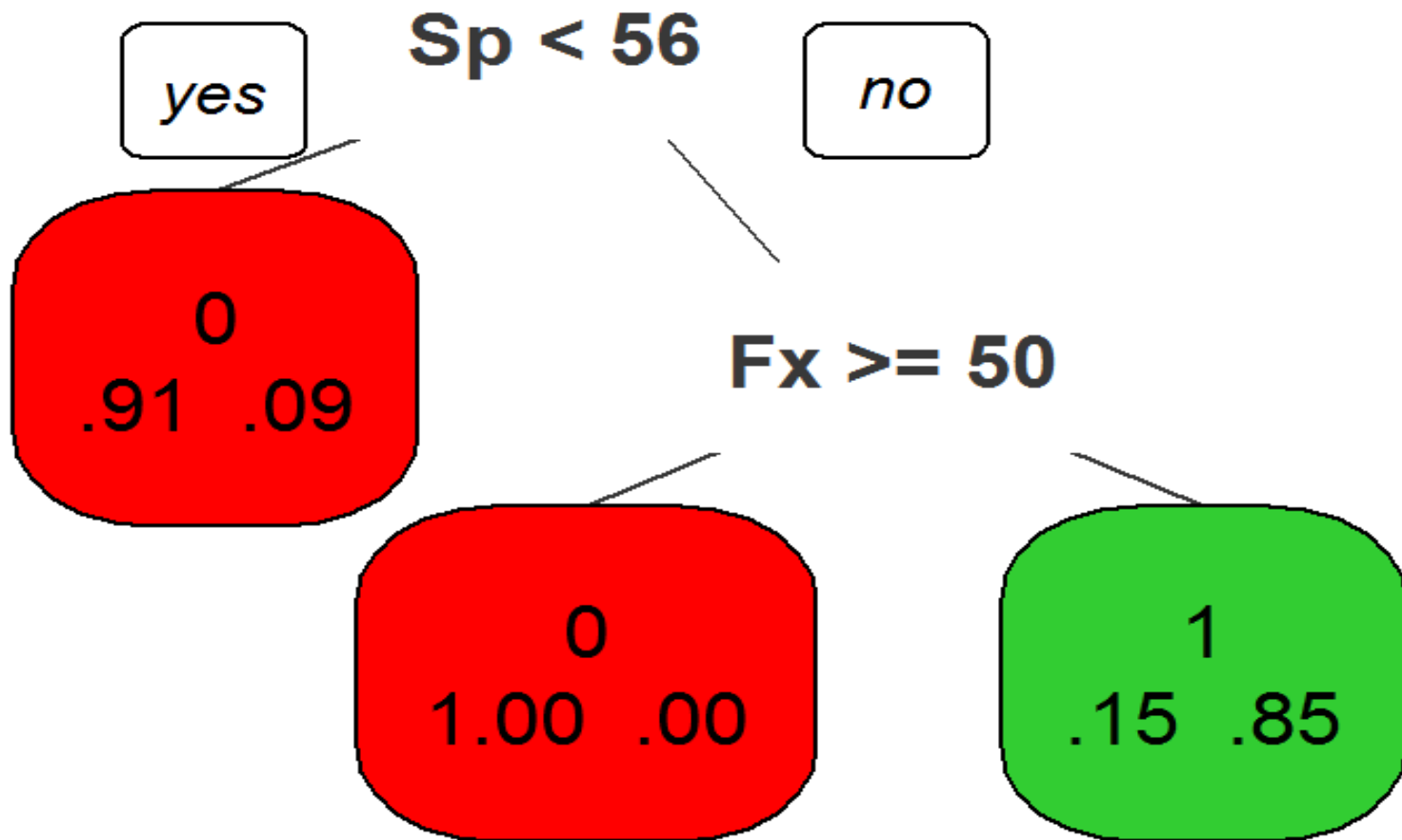
Поэтому



Бонусом



Бонусом



Теперь Ваша очередь



[Бабушкин Эдуард](#)

edvb@yandex.ru