Let us consider a set of $N = 10$ companies, U, producing tables, let x_i = number of workers in the ith company and y_i = number of tables produced during one particular day by the ith company. The figure below shows a scatter plot of both variables.

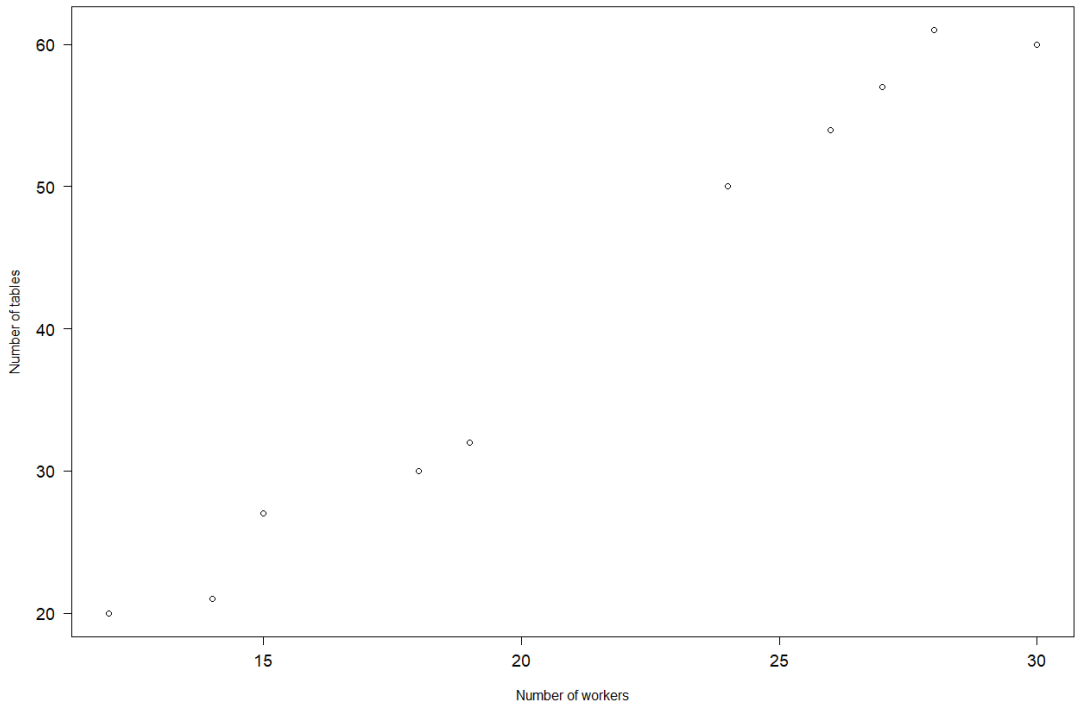


Figure: Scatter plot of number of workers x and number of tables y.

Which of the following lines better describe the set of points?

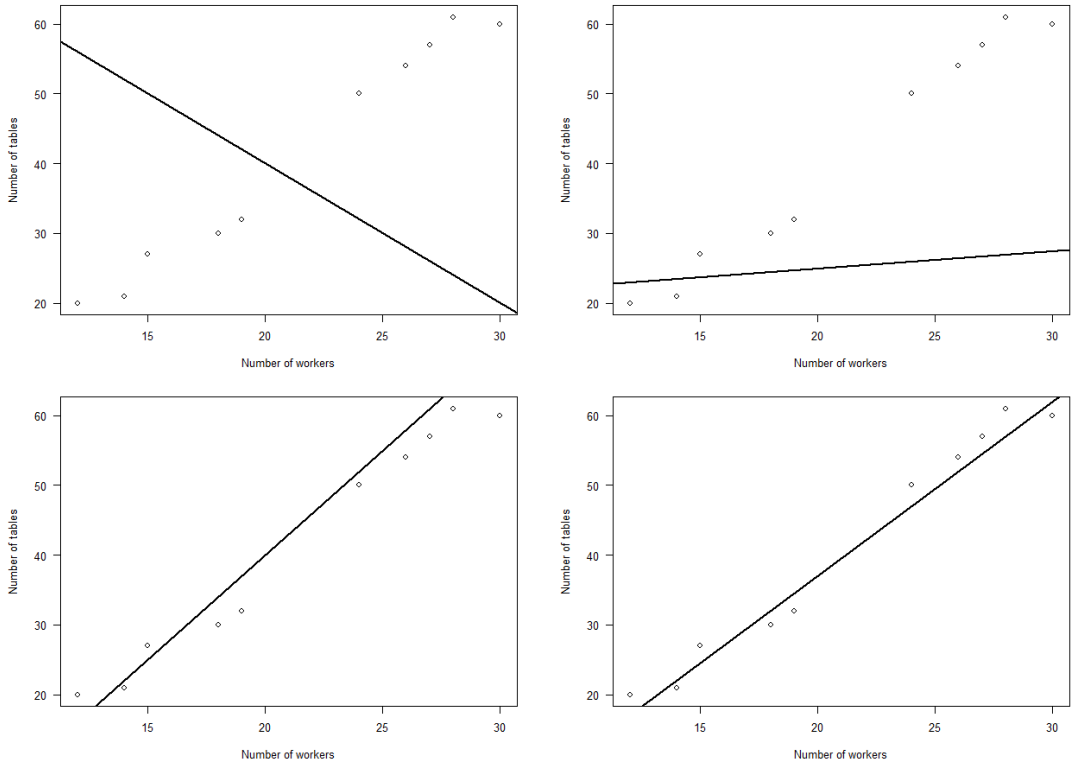
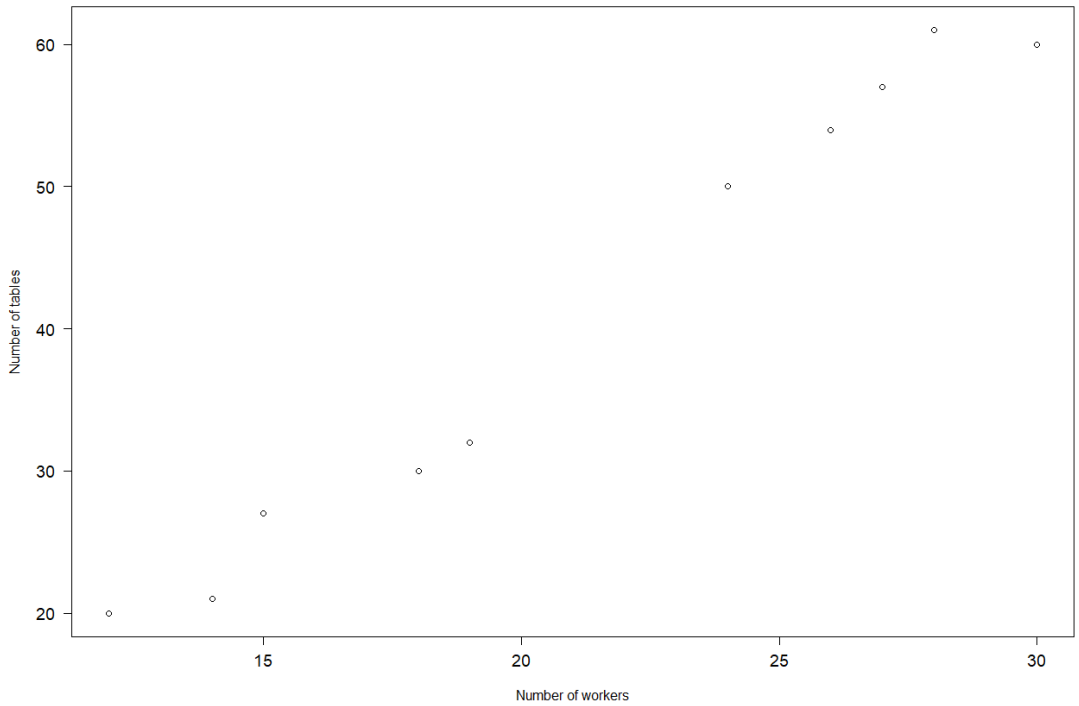


Figure: Four lines fitted to the workers dataset.

Let us consider a set of $N = 10$ companies, U, producing tables, let x_i = number of workers in the ith company and y_i = number of tables produced during one particular day by the ith company. The figure below shows a scatter plot of both variables.



How many tables do you expect to be produced by 20 workers?

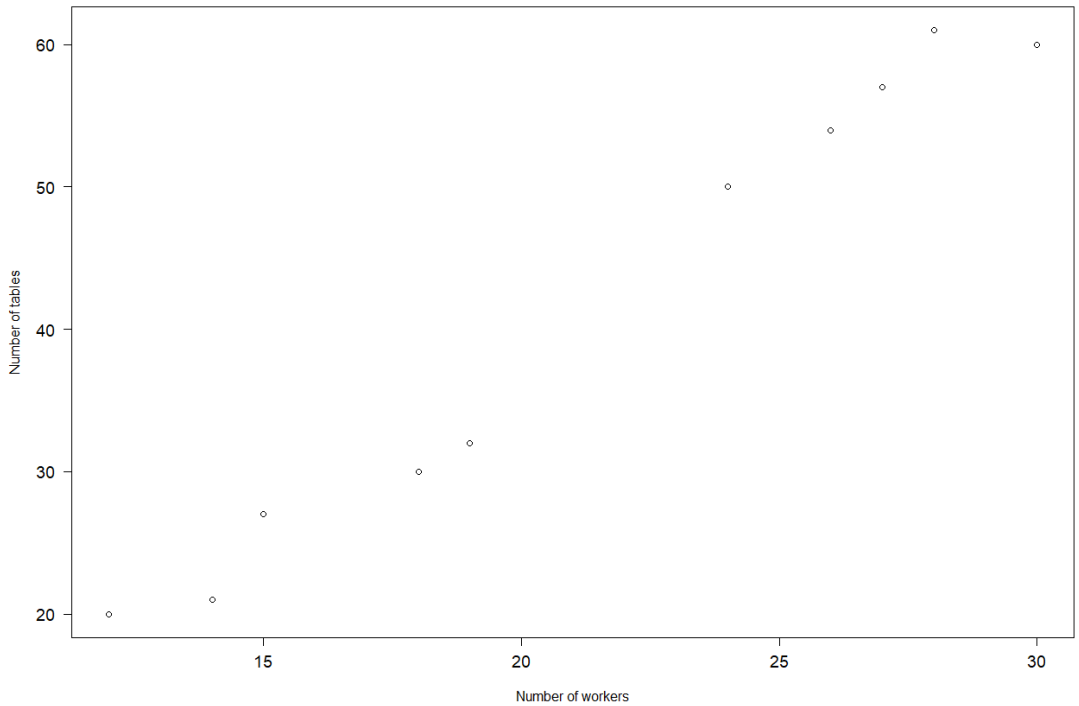


Figure: Scatter plot of number of workers x and number of tables y.

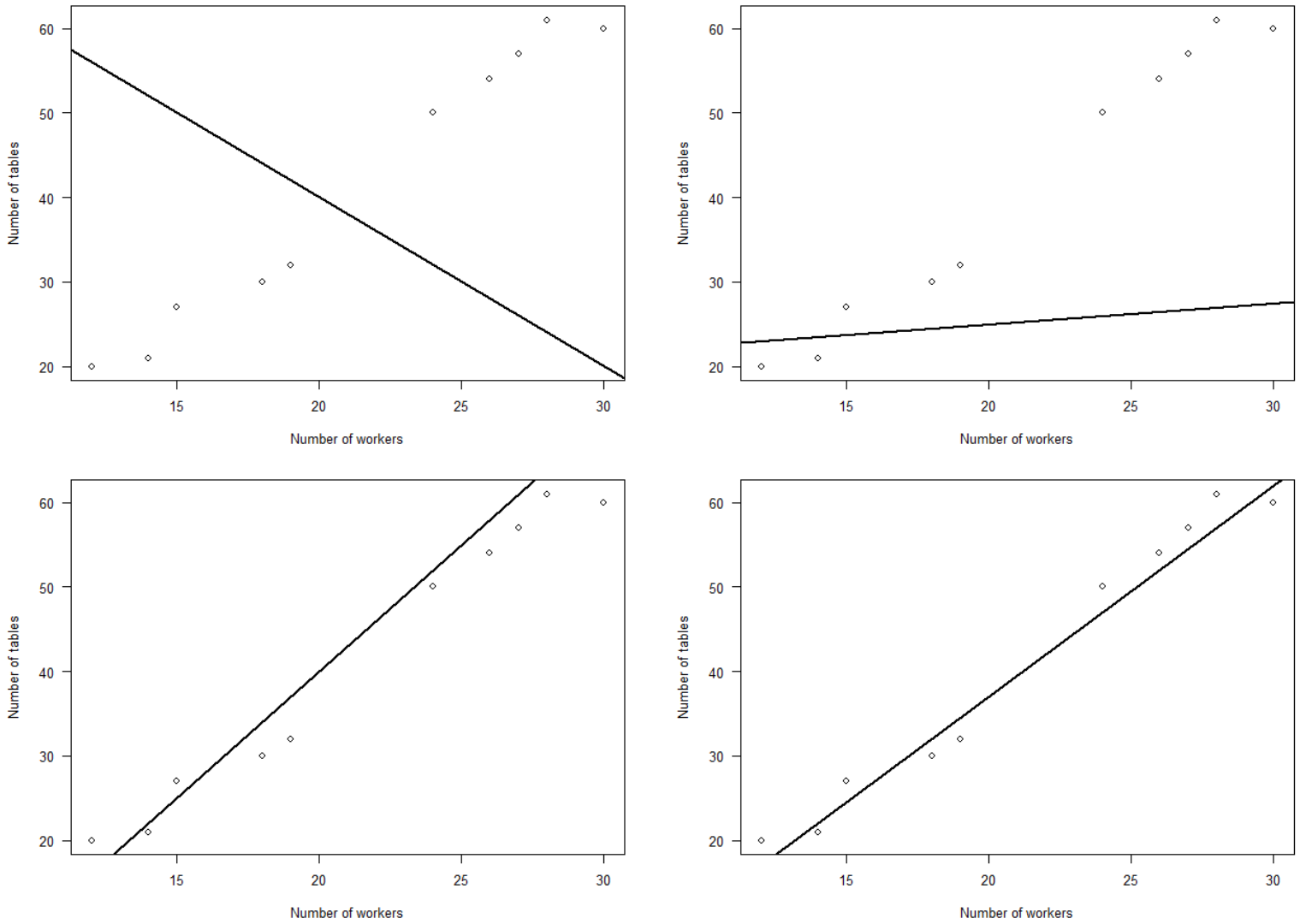


Figure: Four lines fitted to the workers dataset.

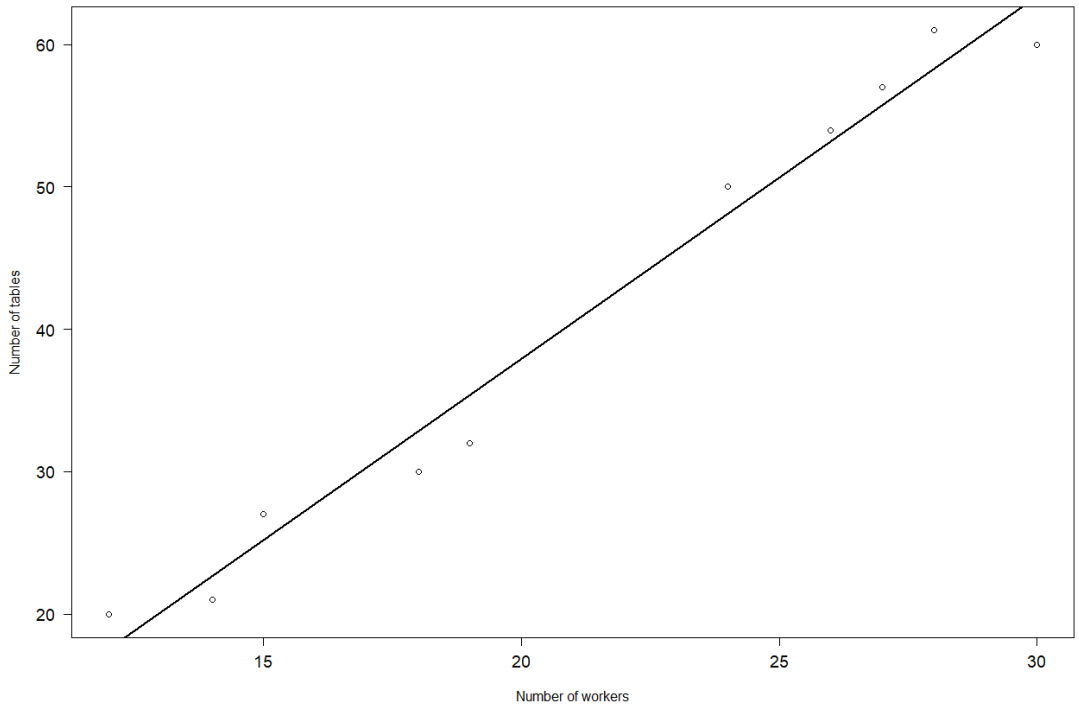


Figure: Least squares regression fitted to the workers dataset.

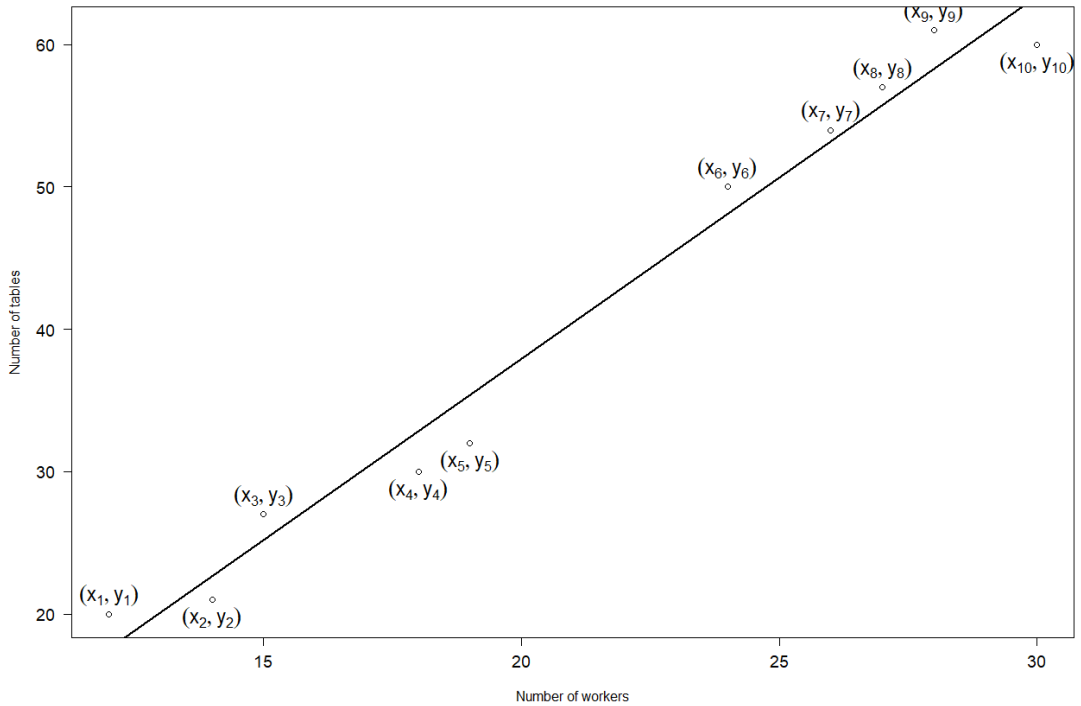


Figure: Least squares regression fitted to the workers dataset.

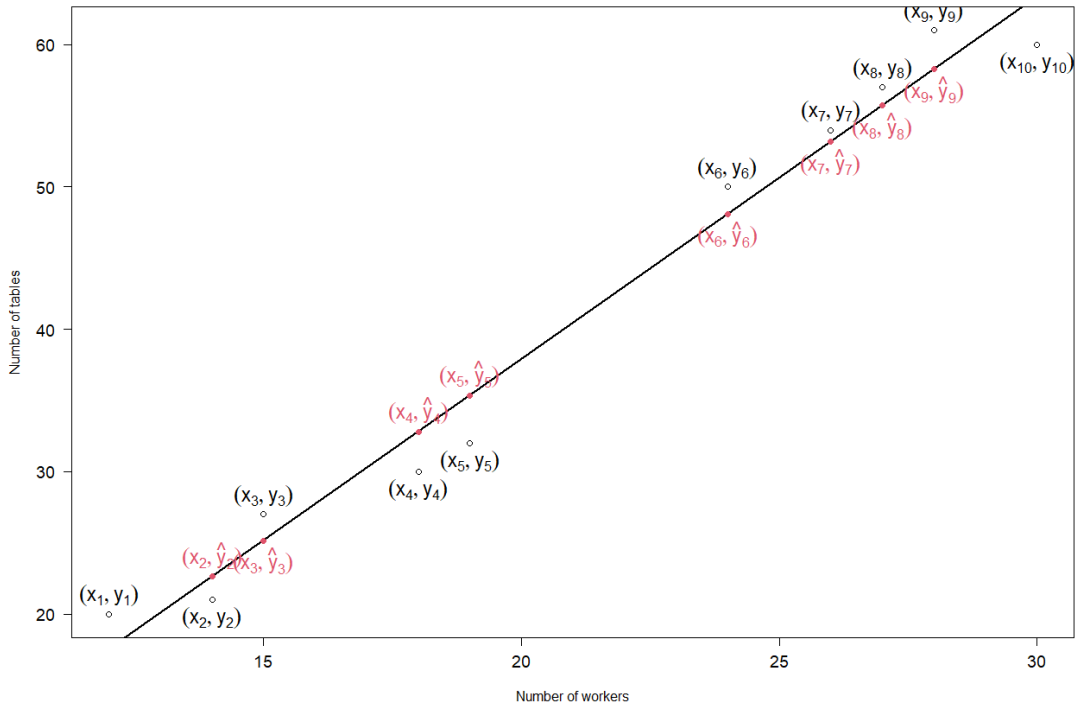


Figure: Least squares regression fitted to the workers dataset.

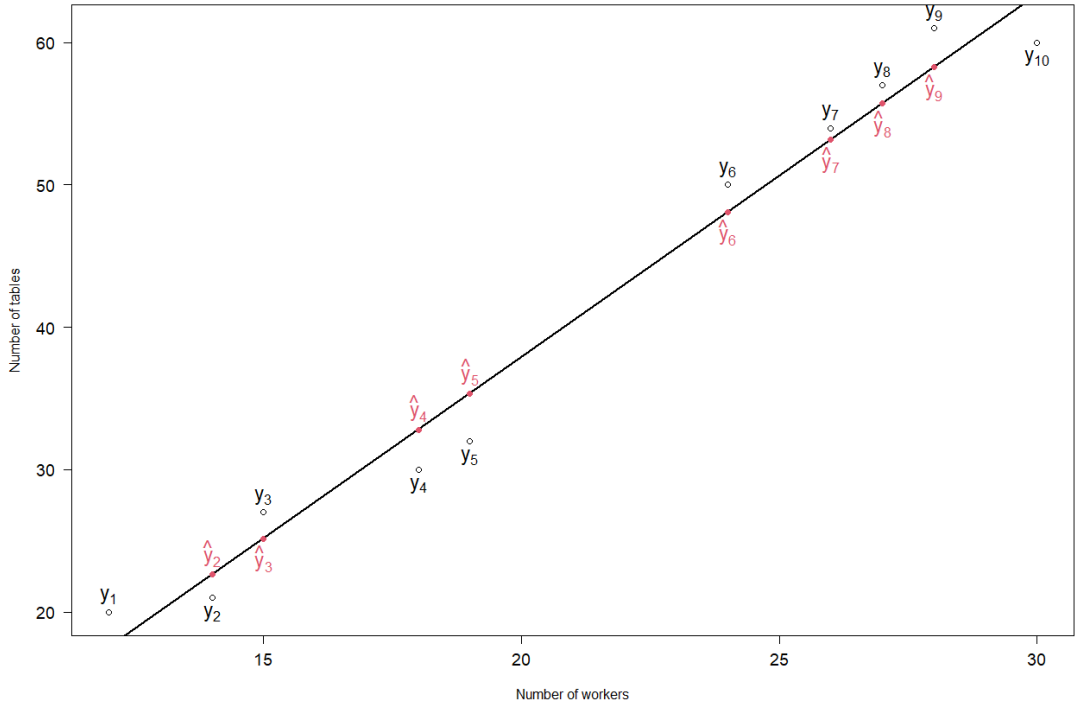


Figure: Least squares regression fitted to the workers dataset.

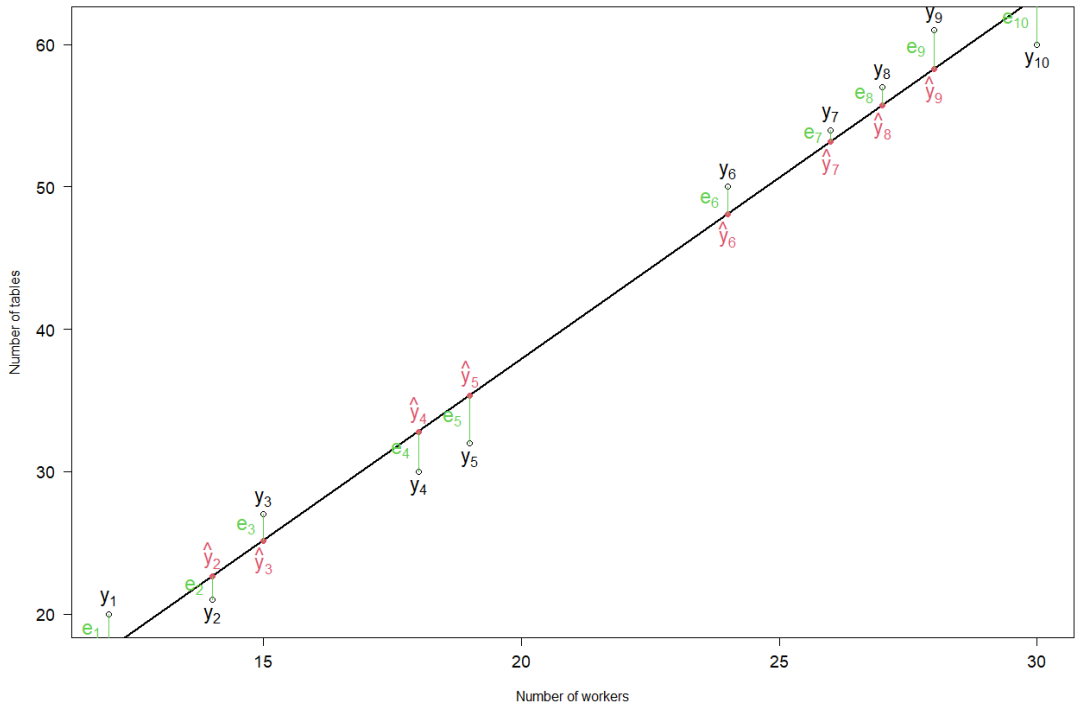


Figure: Least squares regression fitted to the workers dataset.

Definition (Coefficient of determination)

The *coefficient of determination* of a regression, denoted by R^2, is

$$R^2 = \frac{SSR}{SST} \qquad \text{or equivalently} \qquad R^2 = 1 - \frac{SSE}{SST}. \qquad \square$$

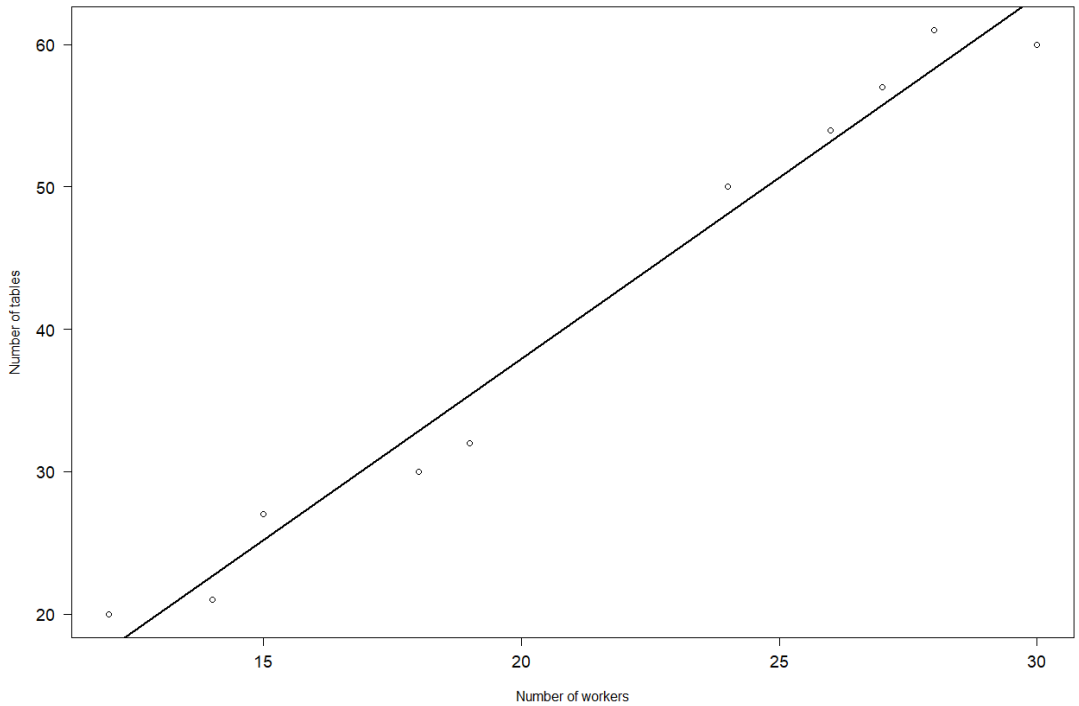


Figure: Least squares regression fitted to the workers dataset.

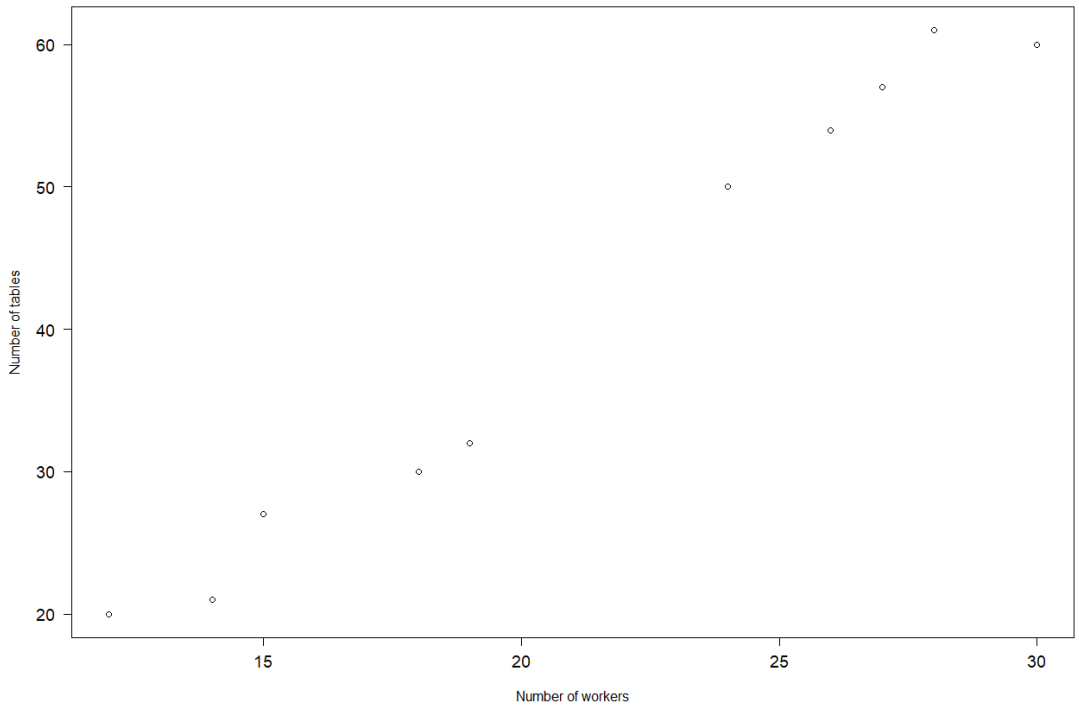


Figure: Scatter plot of number of workers x and number of tables y.

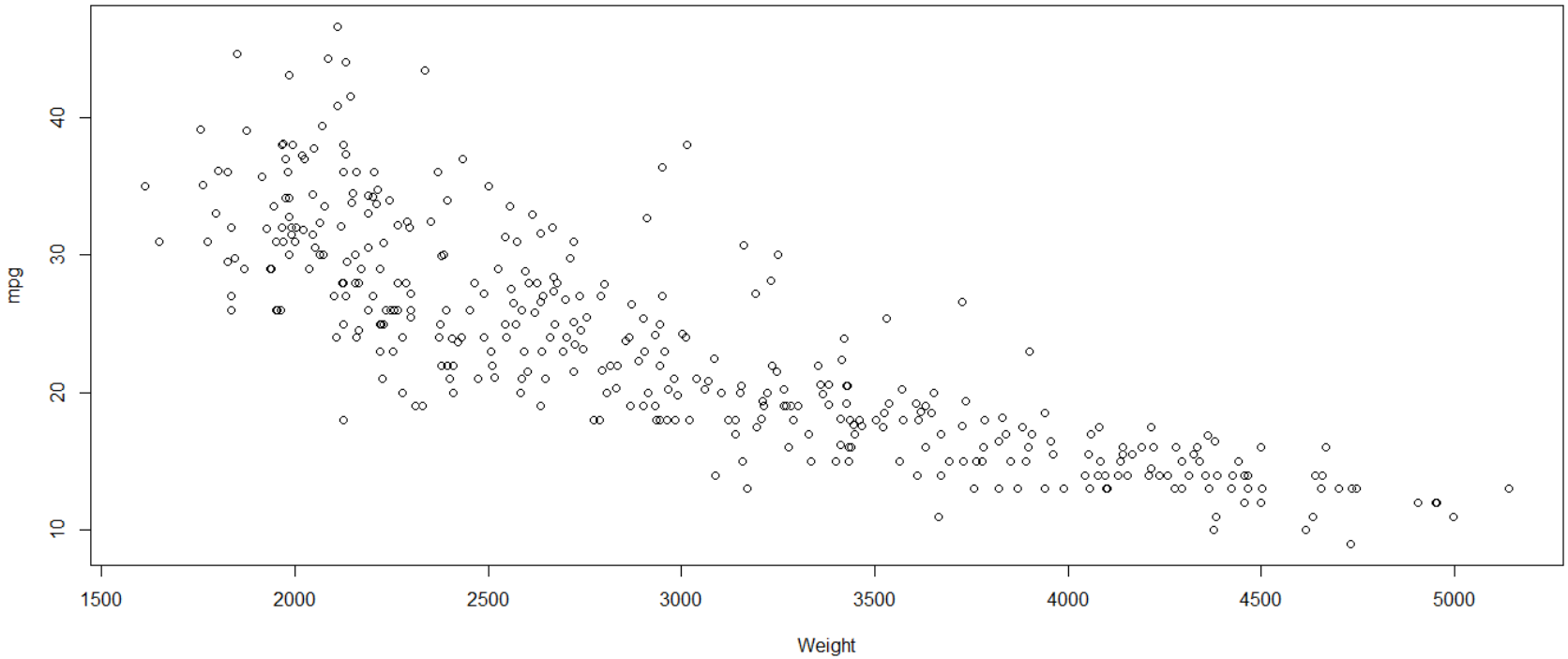


Figure: Scatter plot of weight (pounds) vs. autonomy (miles per gallon) of 392 automobiles.

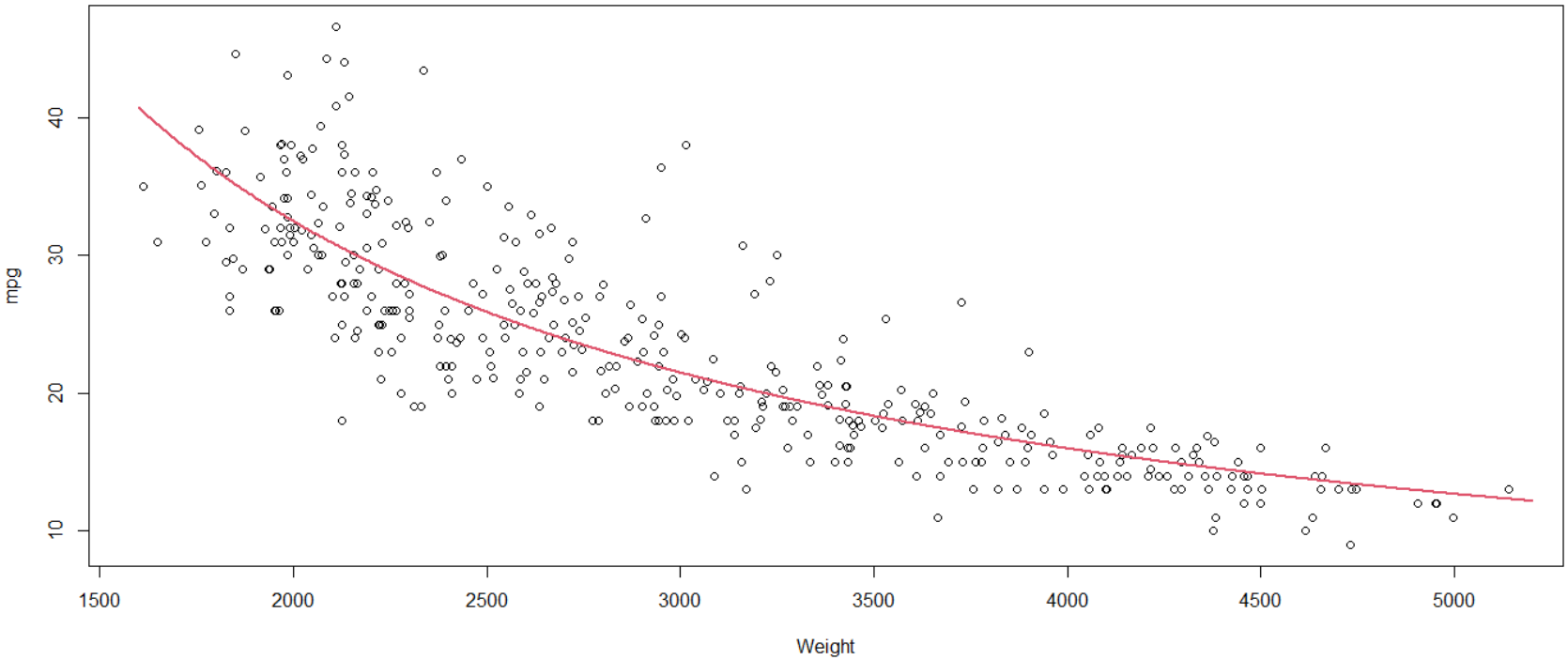
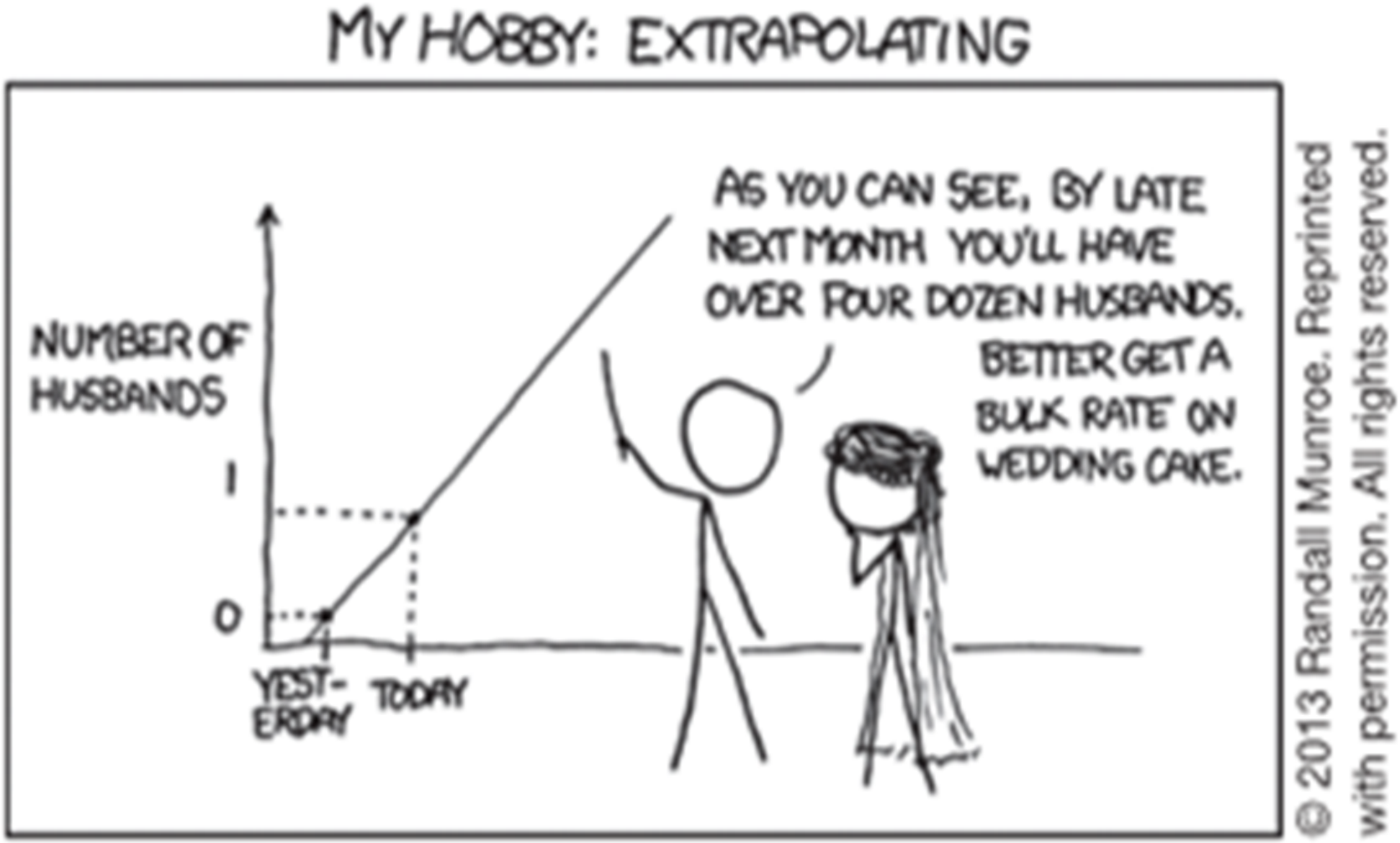


Figure: Scatter plot of weight (pounds) vs. autonomy (miles per gallon) of 392 automobiles.


MY HOBBY: EXTRAPOLATING
AS YOU CAN SEE, BY LATE NEXT MONTH YOU'LL HAVE OVER FOUR DOZEN HUSBANDS.
BETTER GET A BULK RATE ON WEDDING CAKE.
NUMBER OF HUSBANDS
1
0
YEST-ERDAY
TODAY
© 2013 Randall Munroe. Reprinted with permission. All rights reserved.

Set one		Set two	
Starting	Closing	Starting	Closing
3.595	3.91	3.59	3.9
3.595	3.86	2.995	3.17
3.595	3.95	3.495	3.83
3.695	4.04	3.495	3.8
3.695	4.00	3.500	3.77
3.695	4.05	2.195	2.36
3.695	4.06	2.195	2.32
3.695	4.06	2.195	2.32
3.695	4.08	3.450	3.7
3.775	4.11	1.895	1.94
3.795	4.10	2.395	2.54
3.795	4.13	1.699	1.83
3.795	4.16	2.195	2.36
3.795	4.14	2.995	3.17

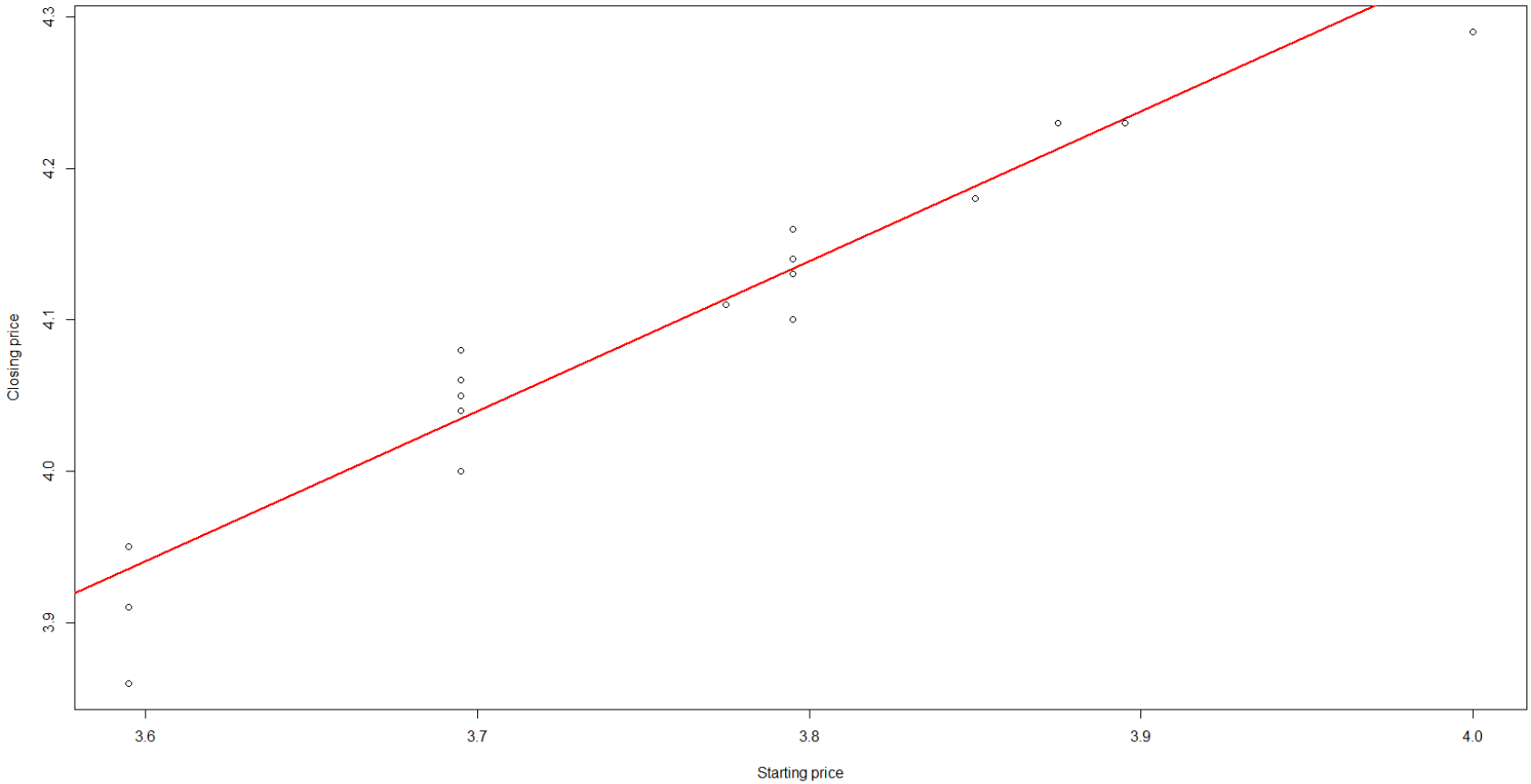


Figure: Scatter plots and fitted regressions of the closing price in terms of starting price for the first set of housing units.

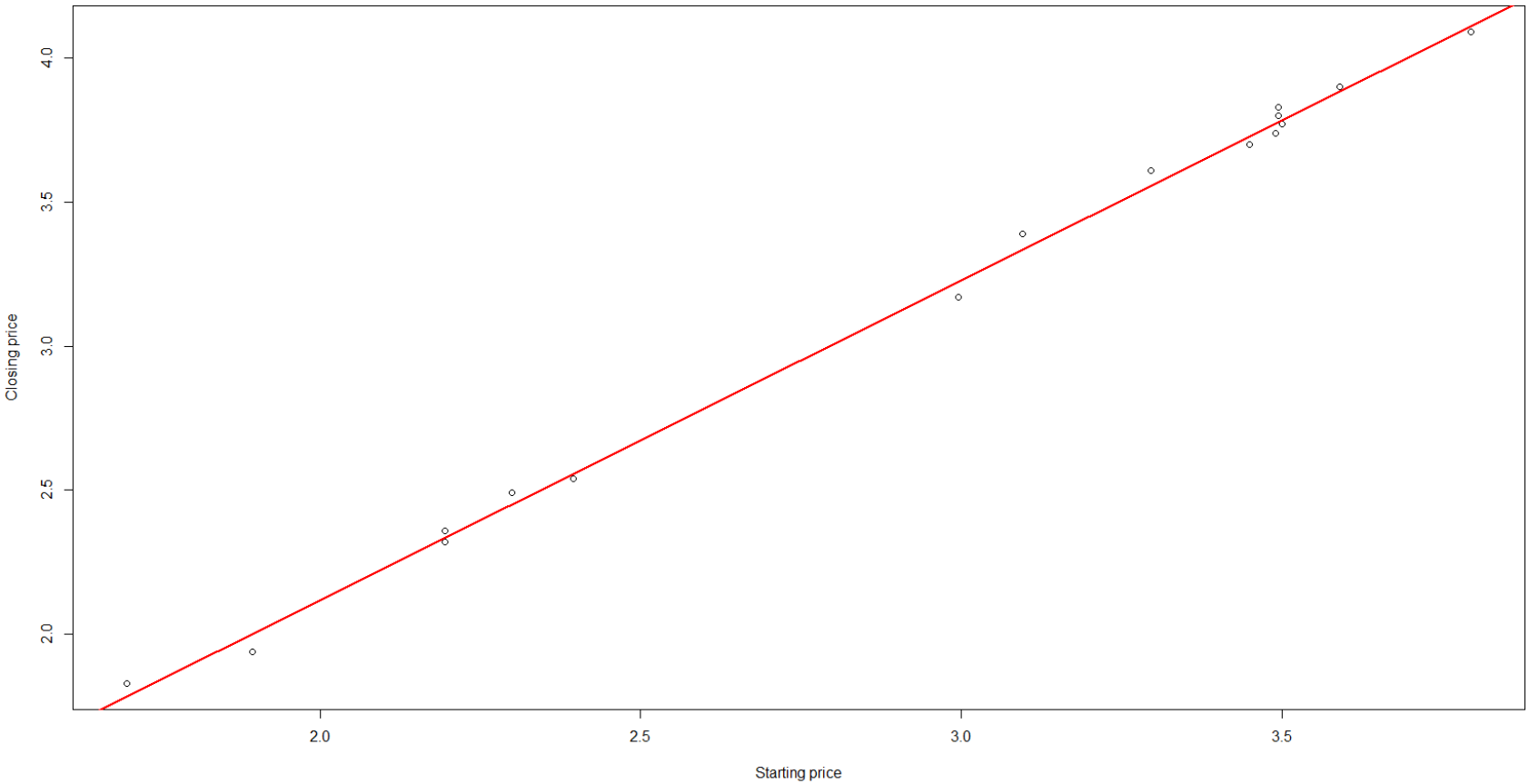


Figure: Scatter plots and fitted regressions of the closing price in terms of starting price for the second set of housing units.

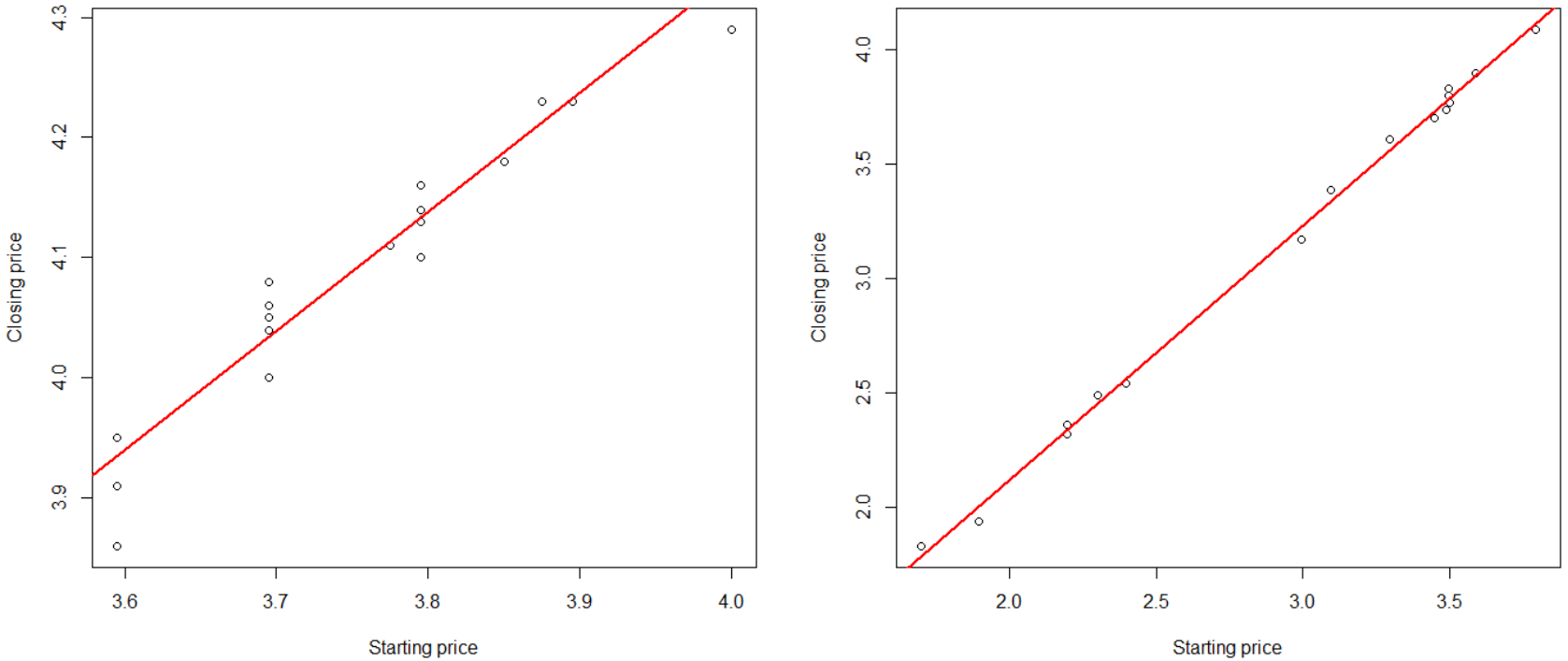
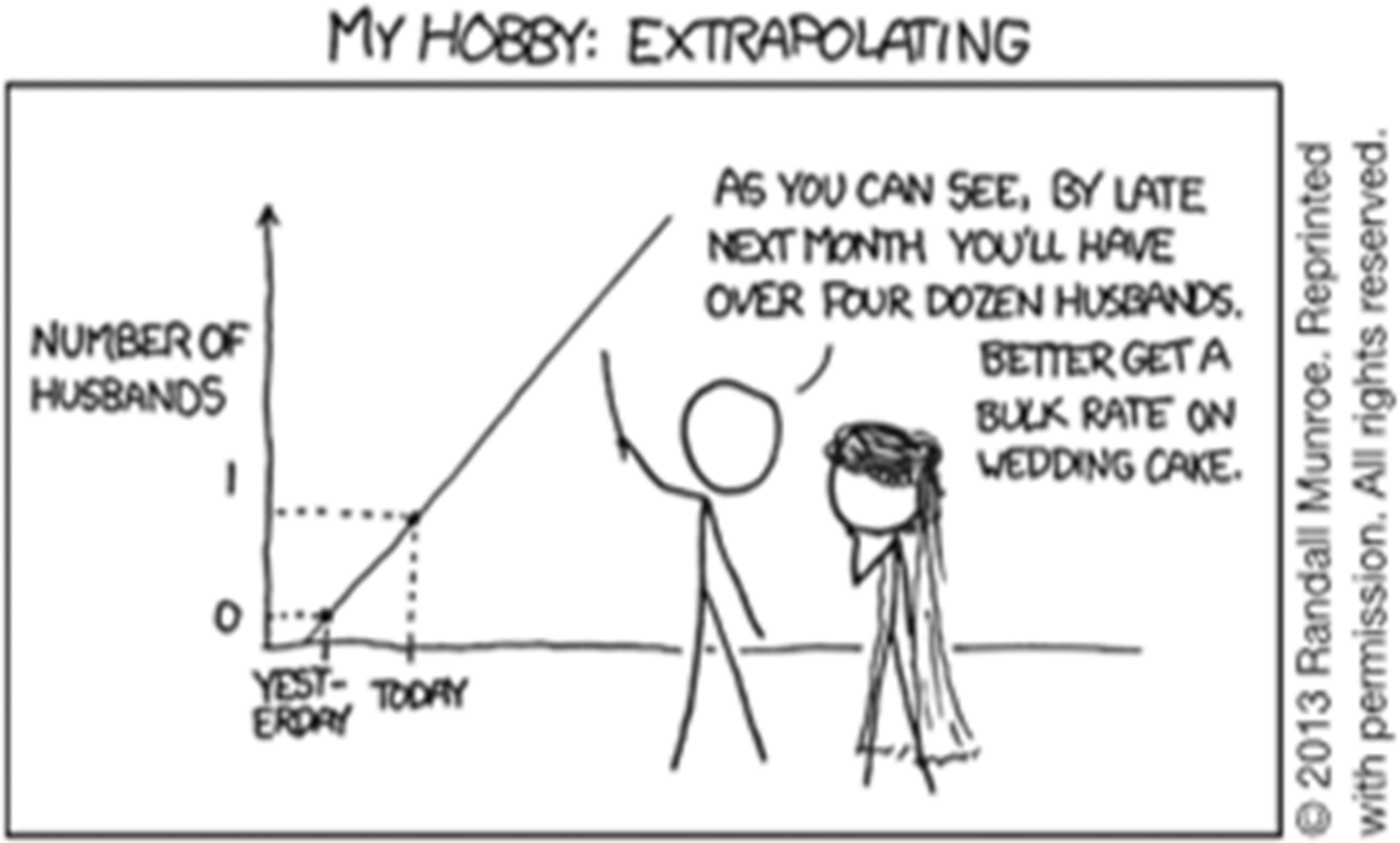


Figure: Scatter plots and fitted regressions of the closing price in terms of starting price for two sets of housing units.


MY HOBBY: EXTRAPOLATING
AS YOU CAN SEE, BY LATE NEXT MONTH YOU'LL HAVE OVER FOUR DOZEN HUSBANDS.
BETTER GET A BULK RATE ON WEDDING CAKE.
NUMBER OF HUSBANDS
1
0
YEST-ERDAY
TODAY
© 2013 Randall Munroe. Reprinted with permission. All rights reserved.