

ØAMET2200 · Fall 2019

Worksheet 5

Instructor: Fenella Carpena

September 20, 2019

Exercise 1 (Real Estate) Skanska Eiendomsutvikling, a residential builder in Norway, has started constructing new apartments in Oslo. Imagine that the apartments are 50-150 square meters. Suppose that you are a real estate analyst for this company. You have been tasked with determining the price that the company should set for apartments of different sizes.

From Finn.no, you gather data on two variables: (1) asking price in millions of kroner (**price**); and (2) size in square meters (**sqm**). Summary statistics for the variables are shown below.

. summarize price sqm ;

Variable	Obs	Mean	Std. Dev.	Min	Max
price	30	8.781667	3.438484	3.71	15.7
sqm	30	89.43333	32.37249	40	151

. correlate price sqm ;

(obs=30)

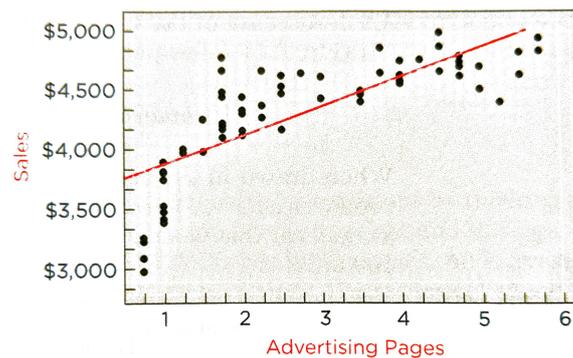
	price	sqm
price	1.0000	
sqm	0.9110	1.0000

- (a) What is the OLS regression line for relating apartment prices to the apartment size in square meters?
- (b) If apartment size increases by 1 square meter, what is our prediction for the increase in price?
- (c) If apartment size increases by 5 square meters, what is our prediction for the increase in price?
- (d) What is our prediction for the price of a 75 square meter apartment?
- (e) What is our prediction for the price of a 100 square meter apartment?

Exercise 2 (True or False) Mark each statement as true or false, and provide a brief explanation for your answer.

- (a) If all of the data lie along a single line with non-zero slope, then the r^2 of the regression is 1 (assume the values of the x variable are not identical).
- (b) If the correlation between x and y is zero, then the slope in the OLS regression line will also be zero.
- (c) The use of a linear equation to describe an association between x and y implies that the change in y when x goes from 10 to 11 is the same as when x goes from 20 to 21.

Exercise 3 (Nille) Imagine that the store Nille recently began selling grocery items (e.g., canned mackerel) in 64 stores located in Oslo. To see how advertising affects the sales of these new items, it varied the number of pages showing grocery items in the advertising circular (*kundeavis*) that it distributes to shoppers. The figure below graphs the average daily sales of grocery items versus the number of pages devoted to grocery items.



- (a) Explain why these data do not satisfy the linear condition.
- (b) Which change in advertising pages appear to have a larger effect on sales: increasing from 1 to 2 pages, or from 4 to 5?
- (c) What transformation does Tukey's bulging rule suggest?

Exercise 4 (Stata) Rema 1000 would like to know what is the best price at which they can sell orange juice (sold in 1.5 liter boxes). Currently, stores are charging kr. 25 per box of orange juice. We want to use data on price and sales history to find the most profitable price.

From our knowledge of microeconomics, we know that the optimal price is equal to $c \cdot \frac{\text{Elasticity}}{\text{Elasticity} + 1}$ where c is the cost of producing one box of orange juice. Let's assume that $c = 9$, which means that one box of orange juice costs kr. 9 to produce for Rema 1000. The datafile `orange_juice.xls` contains data on sales of orange juice over a weekend at 50 locations.

- (a) Read the dataset in Stata using the command `import excel`.
- (b) Create a scatter plot of the variable `sales` (on the y axis) and `price` (on the x axis). Is a linear model appropriate for this relationship? Why or why not?
- (c) Create two new variables: the natural logarithm of price (call it `lnprice`) and the natural logarithm of sales (call it `lnsales`).
- (d) Regress the variable `lnsales` on `lnprice`.
- (e) What is the coefficient on `lnprice`? Interpret this coefficient.
- (f) Create a scatter plot of the variable `lnsales` (on the y axis) and `lnprice` (on the x axis) together with the OLS line.
- (g) Create a scatter plot the residuals from the log-log regression in part (d) on the y axis and `lnprice` on the x axis. Does it show that log-log model is appropriate for the relationship of sales and price?
- (h) Given the results in part (d), what is the price of orange juice that Rema 1000 should set?