

Warning: Though it is pretty easy to "do" regression analysis using Excel, there are many pitfalls to getting statistically sound results. I strongly suggest that you work with a skilled statistician prior to engaging in real-world regression analysis.

The Concept behind multiple regression analysis:

The value of one variable depends on the values of more than one other variable.

Example:

A worker's salary

*← dependent variable
Y*

depends on

Years of education X_1

Years of work experience X_2

*} independent variables
 X_3*

Pitfalls to proper multiple regression analysis

1. Omitted independent variables:

Do not exclude any important thing that might influence the value of the dependent variable.

Example:

New home sales

← dependent variable

depend on

Average price of new home

Average income of people

Mortgage interest rates

) Independent variables

Pitfalls to proper regression analysis, II:

2. Related independent variables

Do not include independent variables that are related. (multicollinearity)

Example:

Number of restaurant visits per year

— Y

depends on

Weekly income of the person

~~Yearly income of the person~~

) X_s

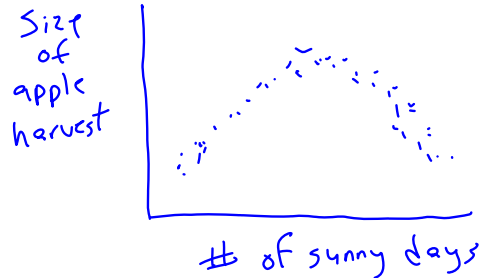
Pitfalls to proper regression analysis, III

3. Nonlinearity

Do not use standard regression analysis of the relationship between a dependent variable and an independent variable is nonlinear

Example:

Size of apple harvest
depends on
Number of sunny days



Advanced statistical techniques can overcome this pitfall.

Unit 12 Video 1 Page 5

Pitfalls to regression analysis, IV

4. Outliers

Data to run a regression comes from the past. If the data occurred due to an extraordinary circumstance that you aren't interested in, then it should be excluded.

Example: When a meteor crashed into Galveston beach on a nice summer day, few people visited the beach. (This does not provide evidence that nice summer days keep people away from the beach.)

Unit 12 Video 1 Page 6

There are more pitfalls. Please consult a skilled statistician if you will do regression analysis in the real world.

The Excel example that I will now present would be really bad if it were done for a real-world purpose outside of the classroom.

Unit 12 Video 1 Page 7

<u>Full time adult worker</u>	<u>Weekly salary</u>	<u>Years of education</u>	<u>Years of work experience</u>	<u>Worker's weight</u>	<u>comment</u>
1	200	10	1	154	
2	220	11	0	175	
3	240	10	1	109	
4	250	10	2	189	
5	300	9	3	201	
6	320	11	2	105	
7	468	12	3	201	
8	546	14	5	143	
9	678	16	6	122	
10	876	14	3	132	
11	876	15	4	287	
12	1123	16	7	108	
13	1234	18	10	165	
14	1321	18	2	145	
15	1454	16	12	189	
16	1587	18	10	123	
17	1672	18	3	150	
18	1754	20	12	165	
19	1888	16	10	120	
20	2011	18	15	212	
21	4003	2	1	150	works at mom's business

Unit 12 Video 1 Page 8

Full time adult worker	Weekly salary	Years of education $\times 1$	Years of work experience $\times 2$	Worker's weight $\times 3$	comment
1	200	10	1	154	
2	220	11	0	175	
3	240	10	1	109	
4	250	10	2	189	
5	300	9	3	201	
6	320	11	2	105	
7	468	12	3	201	
8	546	14	5	143	
9	678	16	6	122	
10	876	14	3	132	
11	876	15	4	287	
12	1123	16	7	108	
13	✓ 1234	✓ 18	✓ 10	✓ 165	
14	1321	18	2	145	
15	1454	16	12	189	
16	1587	18	10	123	
17	1672	18	3	150	
18	1754	20	12	165	
19	1888	16	10	120	
20	2011	18	15	212	
21	4003	2	1	150	works at mom's business