

Intuitive Solution for Best-Fitting Line

An Informal Gradient Descent

David Gerbing

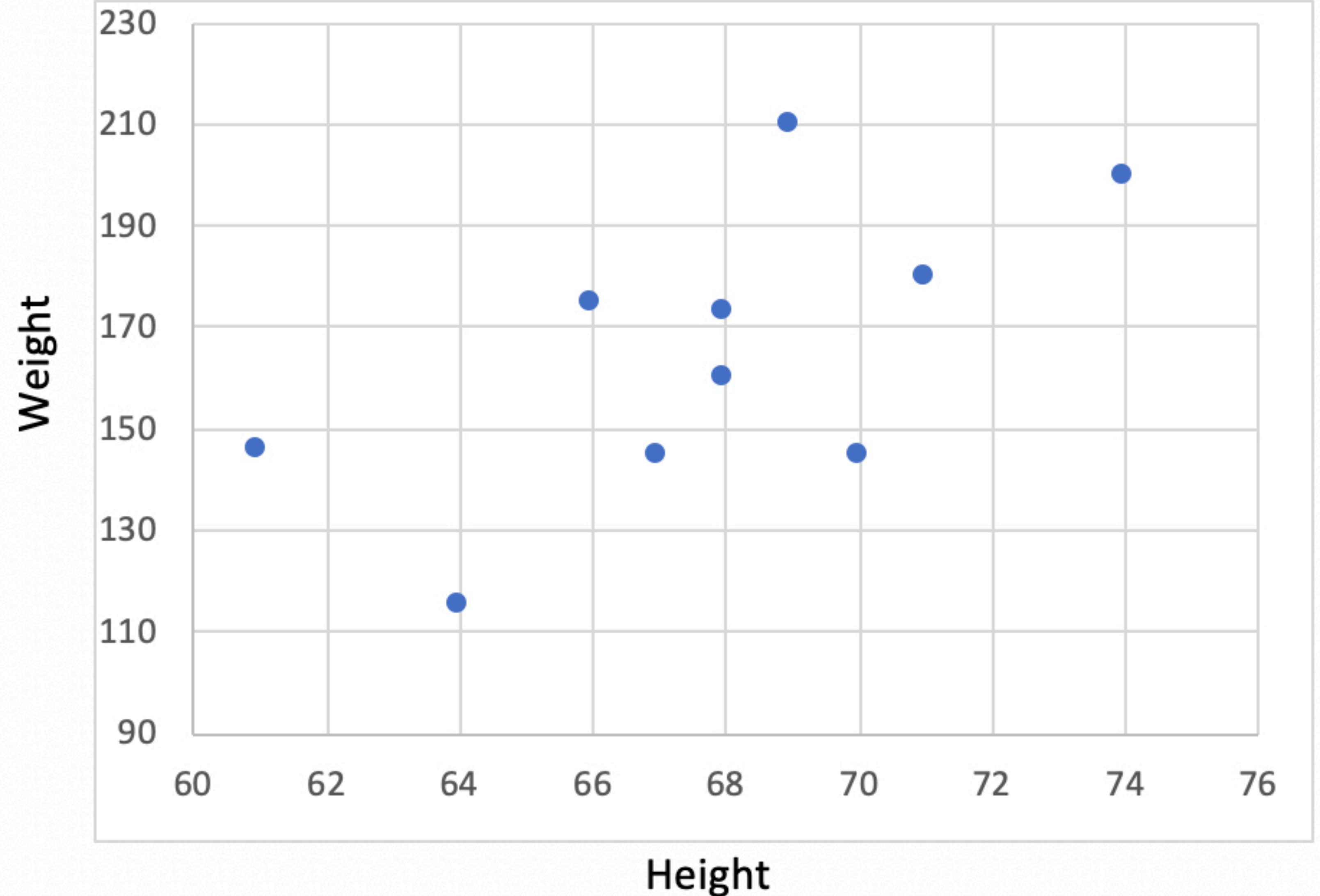
x	y
Height	Weight
64	115
68	160
67	145
74	200
61	146
66	175
71	180
70	145
68	173
69	210

Data

Forecast Weight from Height

Data and Scatterplot

x	y
Height	Weight
64	115
68	160
67	145
74	200
61	146
66	175
71	180
70	145
68	173
69	210



**Approximate the best-fitting line
by trial-and-error**

**Define "best-fit" by the minimization of the
sum of the squared errors**

Equations of the Worksheet

		b_0	0	
		b_1	2	
x	y			
Height	Weight	Fitted	e	e_sq
64	115	=b_0 + b_1*Height		169.00

Fitted Score, y^{\wedge} from x (Height)

		b_0	0	
		b_1	2	
x	y			
Height	Weight	Fitted	e	e_sq
64	115	128.00	=Weight-Fitted	

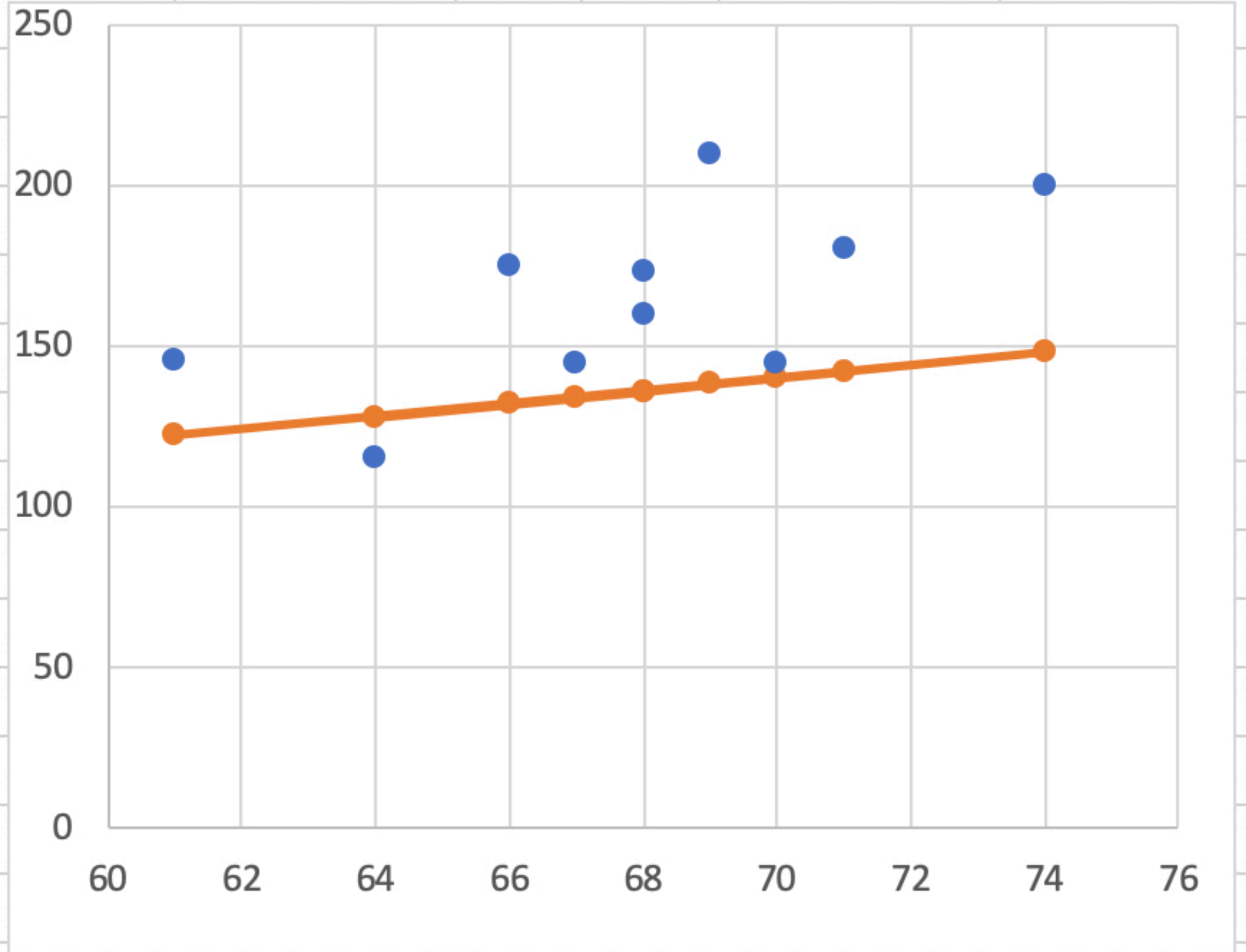
Residual

		b_0	0	
		b_1	2	
x	y			
Height	Weight	Fitted	e	e_sq
64	115	128.00	-13.00	=e^2

Residual Squared

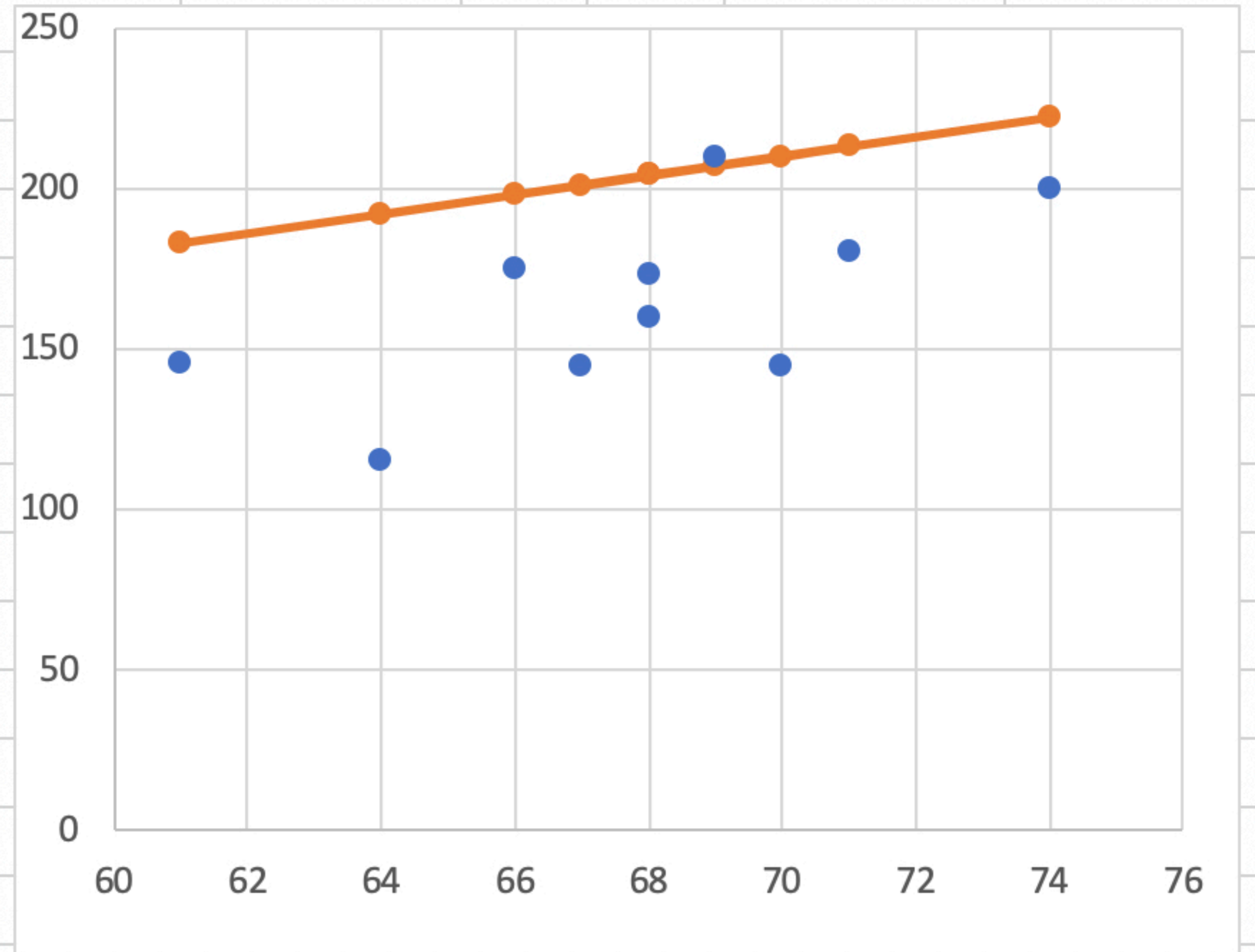
Model: $b_0=0$, $b_1=2$; SSE=14,017

x	y	Fitted	e	e_sq	
64	115	128.00	-13.00	169.00	
68	160	136.00	24.00	576.00	
67	145	134.00	11.00	121.00	
74	200	148.00	52.00	2,704.00	
61	146	122.00	24.00	576.00	
66	175	132.00	43.00	1,849.00	
71	180	142.00	38.00	1,444.00	
70	145	140.00	5.00	25.00	
68	173	136.00	37.00	1,369.00	
69	210	138.00	72.00	5,184.00	
				14,017.00	SSE



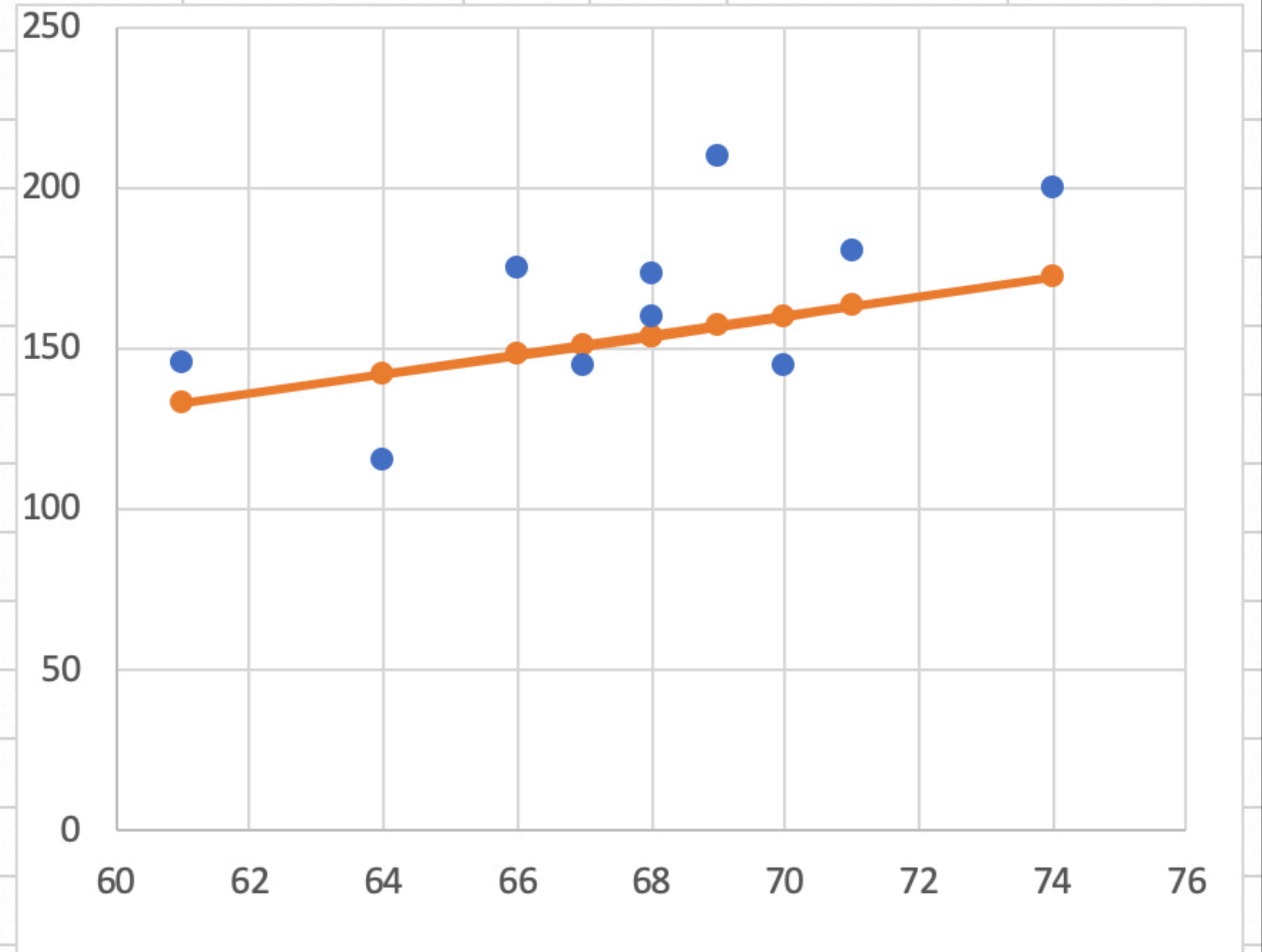
Model: $b_0=0$, $b_1=3$; SSE=19,667

x	y	Fitted	e	e_sq	
Height	Weight				
64	115	192.00	-77.00	5,929.00	
68	160	204.00	-44.00	1,936.00	
67	145	201.00	-56.00	3,136.00	
74	200	222.00	-22.00	484.00	
61	146	183.00	-37.00	1,369.00	
66	175	198.00	-23.00	529.00	
71	180	213.00	-33.00	1,089.00	
70	145	210.00	-65.00	4,225.00	
68	173	204.00	-31.00	961.00	
69	210	207.00	3.00	9.00	
				19,667.00	SSE



Model: $b_0 = -50$, $b_1 = 3$; $SSE = 6,167$

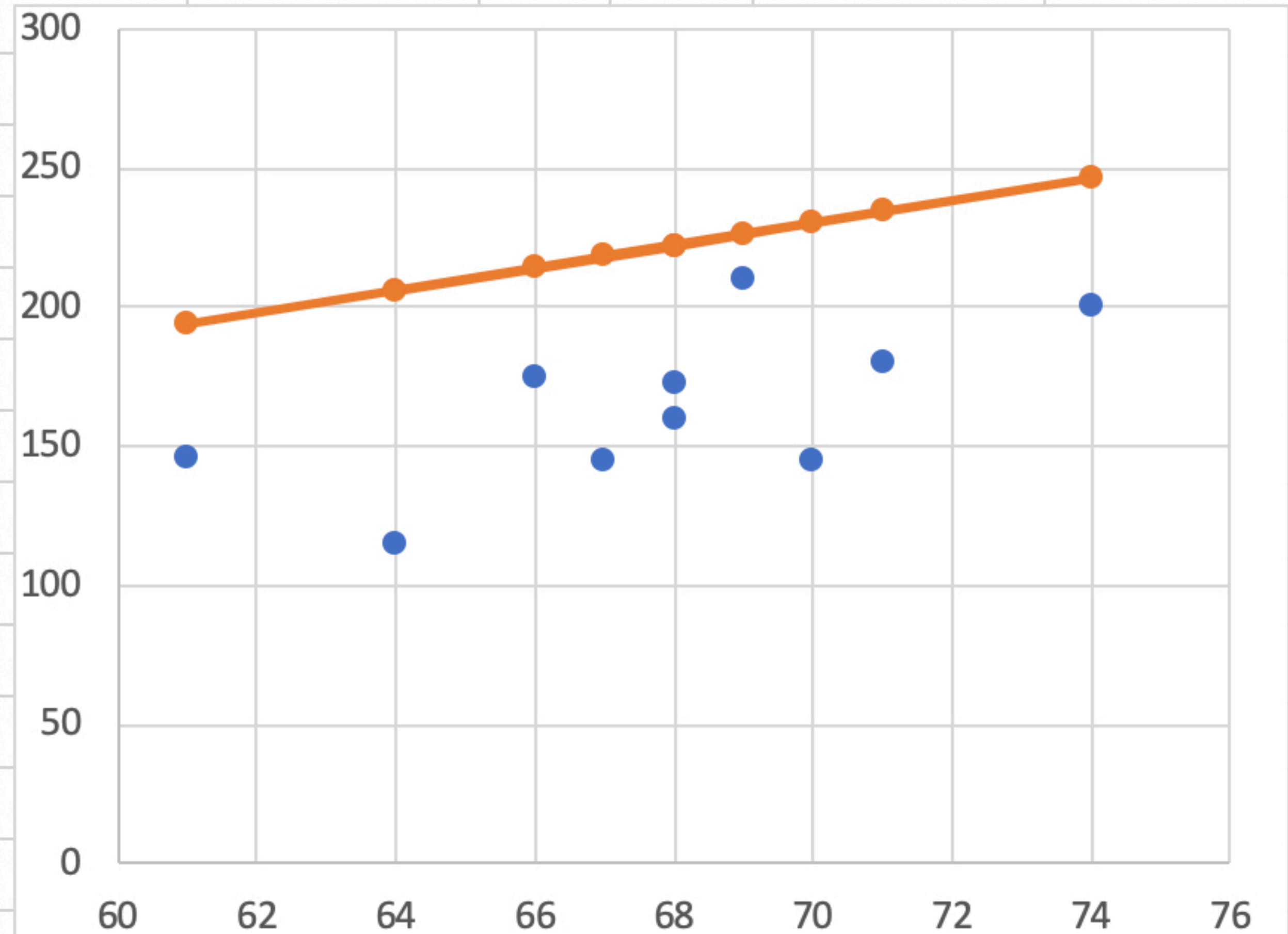
x	y	Fitted	e	e_sq
64	115	142.00	-27.00	729.00
68	160	154.00	6.00	36.00
67	145	151.00	-6.00	36.00
74	200	172.00	28.00	784.00
61	146	133.00	13.00	169.00
66	175	148.00	27.00	729.00
71	180	163.00	17.00	289.00
70	145	160.00	-15.00	225.00
68	173	154.00	19.00	361.00
69	210	157.00	53.00	2,809.00
				6,167.00 SSE



Model: $b_0 = -50$, $b_1 = 4$; SSE = 36,193

x	y	Fitted	e	e_sq
64	115	206.00	-91.00	8,281.00
68	160	222.00	-62.00	3,844.00
67	145	218.00	-73.00	5,329.00
74	200	246.00	-46.00	2,116.00
61	146	194.00	-48.00	2,304.00
66	175	214.00	-39.00	1,521.00
71	180	234.00	-54.00	2,916.00
70	145	230.00	-85.00	7,225.00
68	173	222.00	-49.00	2,401.00
69	210	226.00	-16.00	256.00

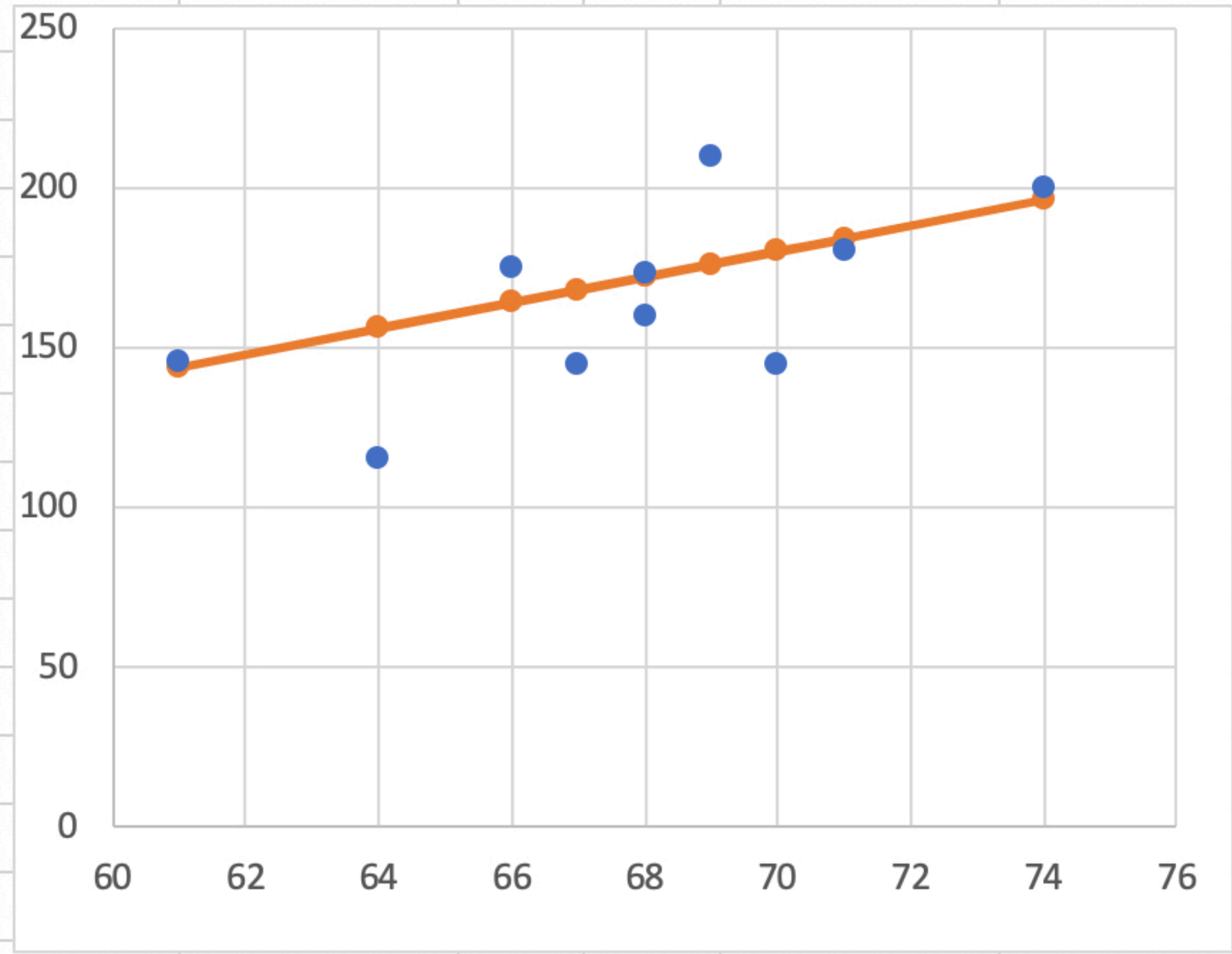
36,193.00 SSE



Model: $b_0 = -100$, $b_1 = 4$; $SSE = 4,893$

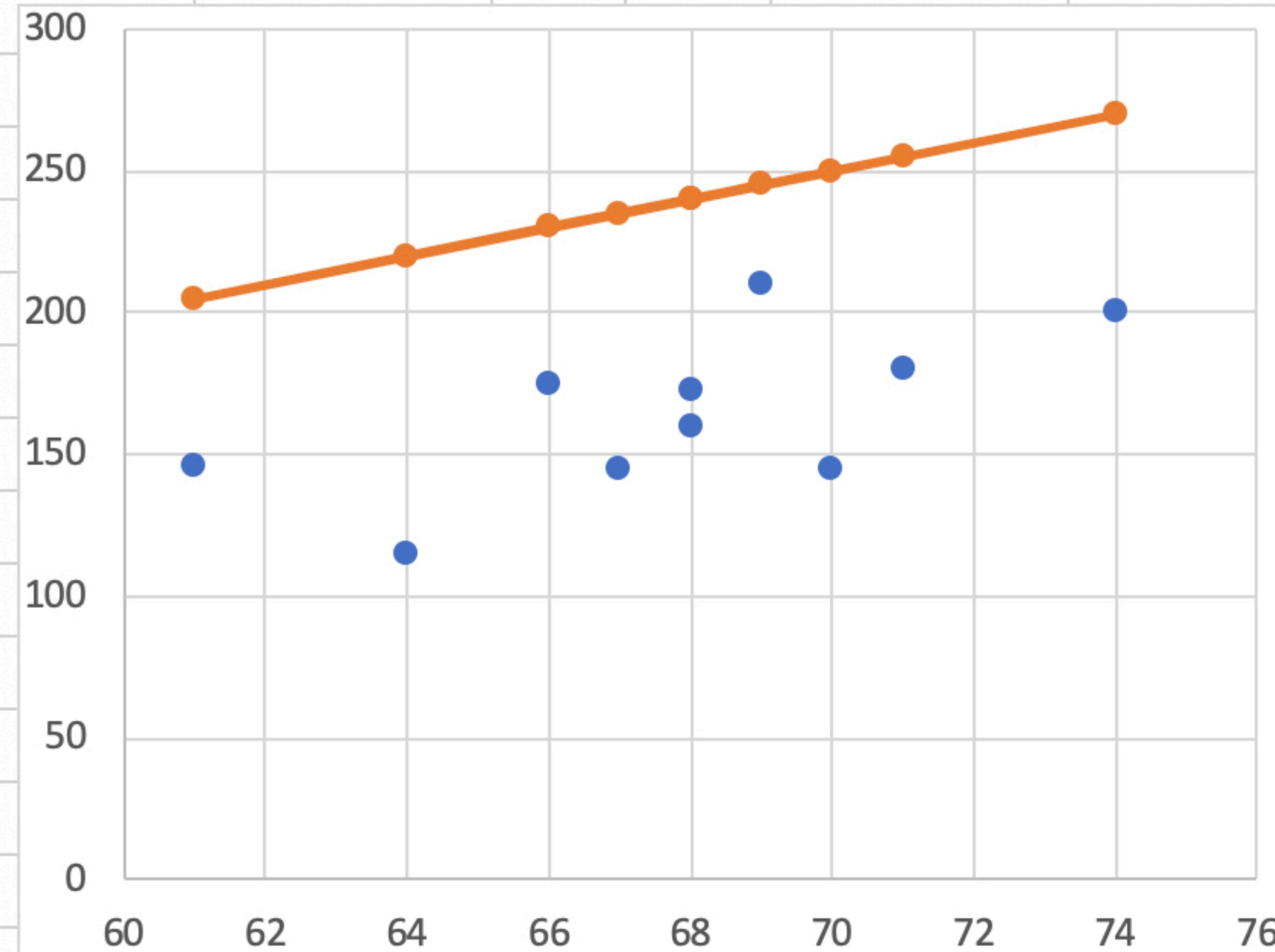
		b_0	-100	
		b_1	4	
x	y	Fitted	e	e_sq
Height	Weight			
64	115	156.00	-41.00	1,681.00
68	160	172.00	-12.00	144.00
67	145	168.00	-23.00	529.00
74	200	196.00	4.00	16.00
61	146	144.00	2.00	4.00
66	175	164.00	11.00	121.00
71	180	184.00	-4.00	16.00
70	145	180.00	-35.00	1,225.00
68	173	172.00	1.00	1.00
69	210	176.00	34.00	1,156.00

4,893.00 SSE



Model: $b_0 = -100$, $b_1 = 5$; $SSE = 59,295$

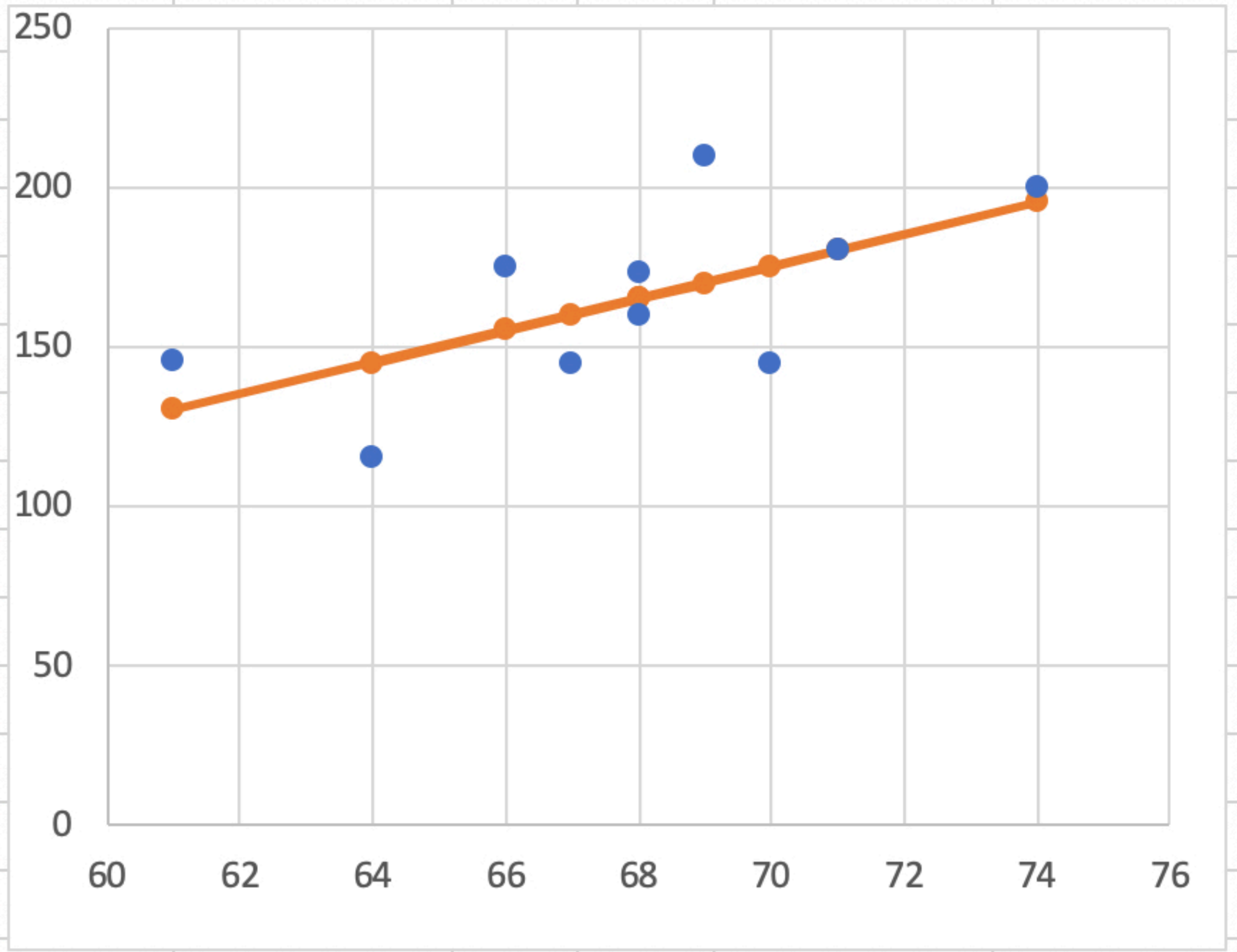
		b_0	-100		
		b_1	5		
x	y				
Height	Weight	Fitted	e	e_{sq}	
64	115	220.00	-105.00	11,025.00	
68	160	240.00	-80.00	6,400.00	
67	145	235.00	-90.00	8,100.00	
74	200	270.00	-70.00	4,900.00	
61	146	205.00	-59.00	3,481.00	
66	175	230.00	-55.00	3,025.00	
71	180	255.00	-75.00	5,625.00	
70	145	250.00	-105.00	11,025.00	
68	173	240.00	-67.00	4,489.00	
69	210	245.00	-35.00	1,225.00	
				59,295.00	SSE



Model: $b_0 = -175$, $b_1 = 5$; $SSE = 4,395$

		b_0	-175	
		b_1	5	
x	y	Fitted	e	e_sq
Height	Weight			
64	115	145.00	-30.00	900.00
68	160	165.00	-5.00	25.00
67	145	160.00	-15.00	225.00
74	200	195.00	5.00	25.00
61	146	130.00	16.00	256.00
66	175	155.00	20.00	400.00
71	180	180.00	0.00	0.00
70	145	175.00	-30.00	900.00
68	173	165.00	8.00	64.00
69	210	170.00	40.00	1,600.00

4,395.00 SSE



Excel Least-Squares Functions

b_0	=INTERCEPT(Weight, Height)		
b_1	4.957		

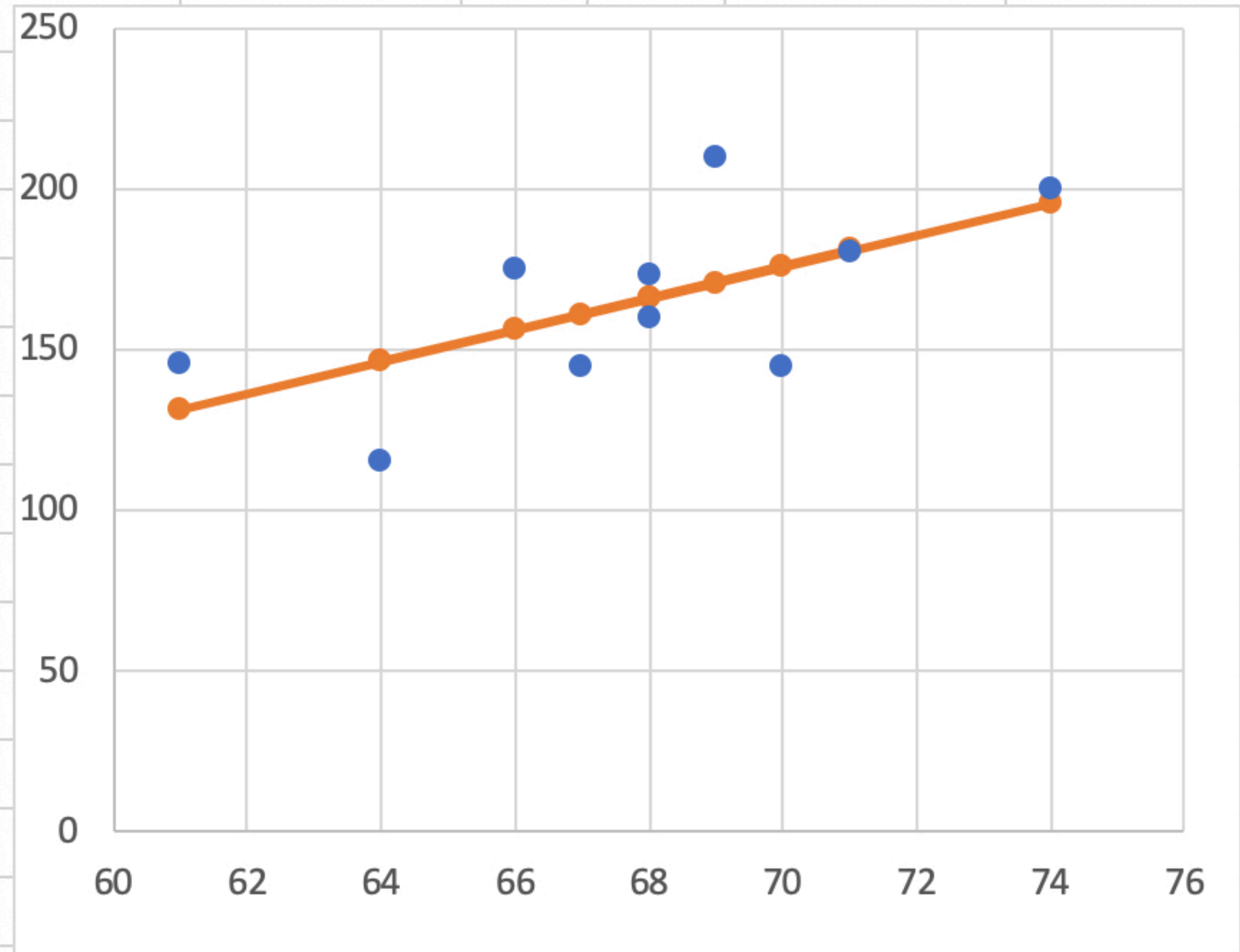
Unlike most machine learning methods, least-Squares solution can be solved analytically

b_0	-171.152		
b_1	=SLOPE(Weight, Height)		

Model: Least-Squares; SSE=4,387

		b_0	-171.152	
		b_1	4.957	
x	y	Fitted	e	e_sq
Height	Weight			
64	115	146.07	-31.07	965.05
68	160	165.89	-5.89	34.71
67	145	160.93	-15.93	253.92
74	200	195.63	4.37	19.09
61	146	131.20	14.80	219.17
66	175	155.98	19.02	361.83
71	180	180.76	-0.76	0.58
70	145	175.80	-30.80	948.91
68	173	165.89	7.11	50.53
69	210	170.85	39.15	1,532.89

4,386.67 SSE



This is roughly how the "gradient descent" solution method works for machine learning

Though the actual method simultaneously changes all parameter estimates on each step to get close to the minimization

The End