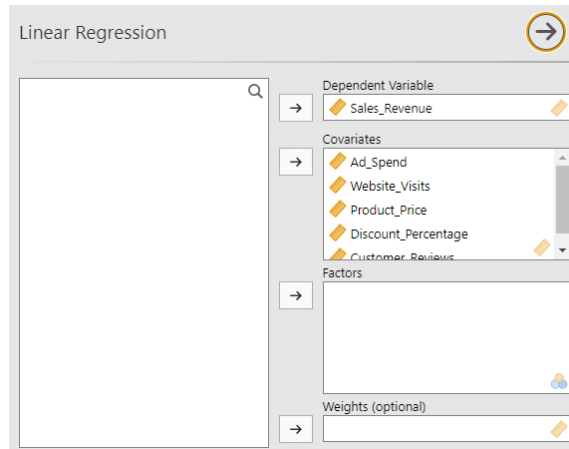


## Homework-5

วิเคราะห์ Multiple linear regression จากข้อมูล sale\_train

Ad_Spend	Website_...	Product_P...	Discount_...	Customer...	Sales_Rev...
2931	3751	44	48	3	28625
2664	4230	96	12	4	22337
2081	2989	90	22	1	40307
1512	2131	99	24	4	12056
8805	3316	17	34	3	16516
6237	1515	35	40	1	49261
1559	1848	83	29	4	6542
4672	3587	99	16	1	26677
8404	3339	43	48	3	44081
4343	835	16	19	4	32569
5070	1395	77	47	1	28322
1580	4046	67	24	3	24065
8004	1351	84	21	1	22924
9130	2187	38	12	1	27556
4147	2431	45	18	2	31962
5249	4246	98	48	2	33295
2679	2927	30	35	3	18807
4506	2068	45	11	2	47348
8253	4500	19	40	3	31432
3744	3238	82	18	2	22749
6188	3998	33	11	1	46459
3141	4413	73	8	3	31256
7617	1566	58	6	3	45262
3479	4880	45	27	1	42080
7484	2944	91	13	3	6324
8421	4607	33	30	3	30731
5452	3873	32	18	2	40909
6881	4469	71	46	3	28803
7665	3046	46	15	4	24870
3849	1253	21	4	1	39578

ข้อมูลไฟล์ Sale\_train



เลือกตัวแปรสำหรับคำนวณ

Dependent

- Sale\_Revenue

Covariates

- Ad\_Spend
- Website\_Visits
- Product\_Price
- Discount\_Percentage
- Customer\_Reviews

### Assumption Checks

Collinearity Statistics		
	VIF	Tolerance
Ad_Spend	1.19	0.843
Website_Visits	1.07	0.936
Product_Price	1.16	0.859
Discount_Percentage	1.06	0.940
Customer_Reviews	1.03	0.966

[3]

ผล Assumption Checks – ตรวจสอบ Collinearity (ตัวแปรอิสระมีความสัมพันธ์กัน)

จากผลการวิเคราะห์พบว่าค่า VIF สูงสุดมีค่า 1.19 ซึ่งไม่ควรเกิน 5

และค่า Tolerance ที่มีค่าน้อยที่สุด .843 ซึ่งไม่ต่ำกว่า .2

แสดงว่า ตัวแปรอิสระไม่มีความสัมพันธ์กัน (ไม่เกิด Multicollinearity)

Model Fit Measures			
Model	R	R <sup>2</sup>	RMSE
1	0.648	0.419	8629

Note. Models estimated using sample size of N=30

มีค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ (R) เท่ากับ **.648**

สามารถพยากรณ์ความพึงพอใจในชีวิตสมรส (R<sup>2</sup>) ได้ร้อยละ **41.9%**

และมีค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) เท่ากับ **8629**

Model Coefficients - Sales_Revenue				
Predictor	Estimate	SE	t	p
Intercept	51576.583	9194.482	5.610	< .001
Ad_Spend	-0.239	0.807	-0.297	0.769
Website_Visits	1.125	1.561	0.721	0.478
Product_Price	-177.872	69.803	-2.548	0.018
Discount_Percentage	-36.230	133.674	-0.271	0.789
Customer_Reviews	-5243.529	1654.214	-3.170	0.004

ค่า Estimate

- Ad\_Spend = -0.239
- Website\_Visits = 1.125
- Product\_Price = -177.872
- Discount\_Percentage = -36.230
- Customer\_Reviews = -5243.529

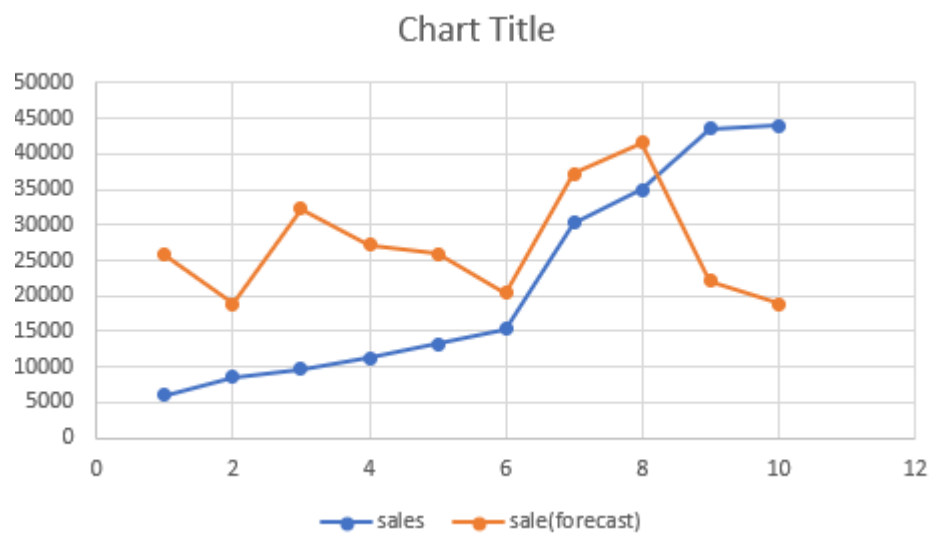
โมเดลหรือสมการ Multiple linear regression

$$\text{Sale\_Revenue} = 51576.583 - 0.239(\text{Ad\_Spend}) + 1.125(\text{Website\_Visits}) - 177.872(\text{Product\_Price}) - 36.23(\text{Discount\_Percentage}) - 5243.529(\text{Customer\_Reviews})$$

ผลการทดสอบโมเดลกับข้อมูล sale\_test

Ad_Spend	Website_Visits	Product_Price	Discount_Percentage	Customer_Reviews	Sales_Revenue	Sales_Revenue(predicted)	predicted-actual	^2
5777	2796	61	23	3	13152	25927.311	12775.311	163208571.1
4900	2761	88	1	3	43518	22092.055	-21425.945	459071119.1
7677	2786	56	44	4	15267	20346.962	5079.962	25806013.92
7981	4702	65	25	1	30289	37247.915	6958.915	48426497.98
7376	3378	95	13	2	6062	25758.081	19696.081	387935606.8
1523	4643	23	6	2	34926	41640.467	6714.467	45084067.09
3393	4144	99	2	2	11229	27258.81	16029.81	256954808.6
3593	4932	37	46	3	9632	32287.925	22655.925	513290937.6
8240	2555	96	22	3	44056	18878.239	-25177.761	633919649
7937	1776	87	45	3	8596	18841.839	10245.839	104977216.8
							Sum/N	263867448.8
						RMSE	sqrt	16243.997

ค่า RMSE เท่ากับ 16243.997



กราฟเปรียบเทียบข้อมูลจริงและผลการทำนาย