MULTIPLE REGRESSION ANALYSIS

MULTIPLE REGRESSION

Many practical situations involve analyzing the relationships among three or more variables.

When multiple independent variables are to be included in an analysis simultaneously, multiple regression is very useful.

Multiple Regression Model Population

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

where:

β_0 = Population's regression constant

β_j = Population's regression coefficient for each variable $x_j = 1, 2, \ldots k$

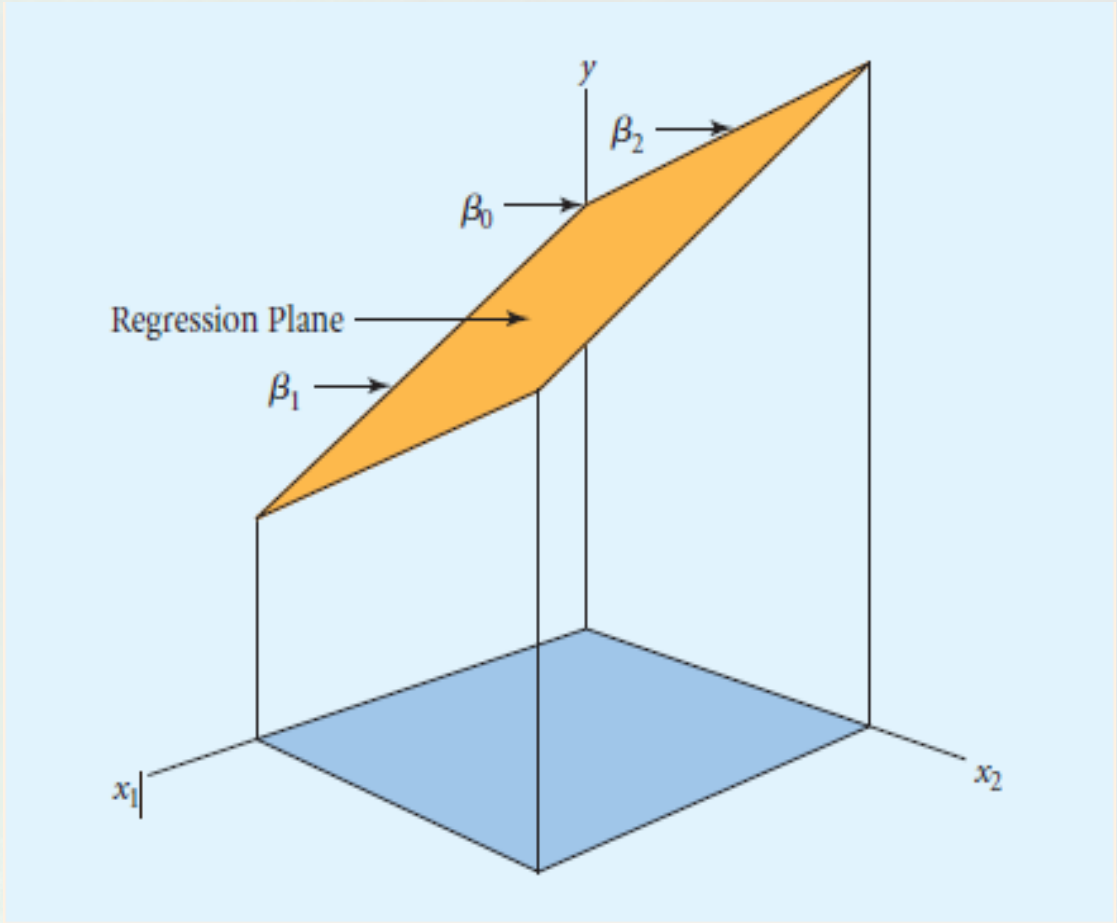
k = Number of independent variables

ε = Model error

Estimated Multiple Regression Model

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_k x_k$$

Figure: Multiple Regression Hyperplane for Population



CONDUCTING MULTIPLE REGRESSION

Hospital readmissions within 30 days of discharge have become a critical quality metric in healthcare systems globally. High readmission rates not only reflect negatively on patient care outcomes but also lead to increased healthcare costs. Hospital administrators seek to understand the key factors that contribute to readmission risks to improve service quality and reduce unnecessary costs.

Given a large dataset of discharged patients—including demographic information, comorbidities, treatment duration, discharge instructions compliance, medication types, lab results, and hospital resource utilization—this study aims to build a multiple linear regression model to predict the likelihood of a 30-day readmission.

However, due to potential interdependencies among predictors (e.g., certain comorbidities often occurring together or overlapping with lab values), multicollinearity is anticipated and must be addressed to ensure model validity and interpretability.

CONDUCTING MULTIPLE REGRESSION

Primary Question:

What are the key predictors of 30-day hospital readmission among discharged patients, and how effectively can a multiple regression model estimate this likelihood based on patient-level and treatment-related variables?

Supporting Questions:

1. Which independent variables have the strongest association with 30-day readmission?
2. Is there evidence of multicollinearity among predictors, and how does it affect the stability and interpretability of the regression model?
3. What strategies can be used to mitigate multicollinearity in the dataset?

📋 Variable Descriptions (Healthcare Readmission Dataset)

Age

- Patient's age in years (whole number).

Length_of_Stay

- Number of days the patient stayed in the hospital during their latest admission.

Comorbidities

- Number of co-existing medical conditions (e.g., diabetes, hypertension).

Lab_Abnormalities

- A derived score indicating abnormalities found in lab tests (often correlated with comorbidities).

Medication_Complexity

- A score reflecting the complexity and number of medications prescribed.

Discharge_Instruction_Score

- Score (1 to 10) evaluating patient understanding of post-discharge instructions.

Resource_Utilization

- A proxy score combining factors such as staff time, equipment, and space used during treatment.

Readmission_Risk

- Predicted risk score (continuous) estimating likelihood of readmission within 30 days.

CONDUCTING MULTIPLE REGRESSION

Business Requirements:

Building a **multiple linear regression model** to predict the likelihood of a 30-day readmission to a hospital.

Data Set:

Readmission_Data_Final.xlsx

Variables Settings: Xs = ?, Y=?

Correlation Analysis: Measuring the strength of the relationship

Multicollinearity Issue: Check which variables

Regression Model Development: Excel, SPSS, R-Programming, Python

Regression Model Diagnostic: Statistical Test, R-Squared, RMSE

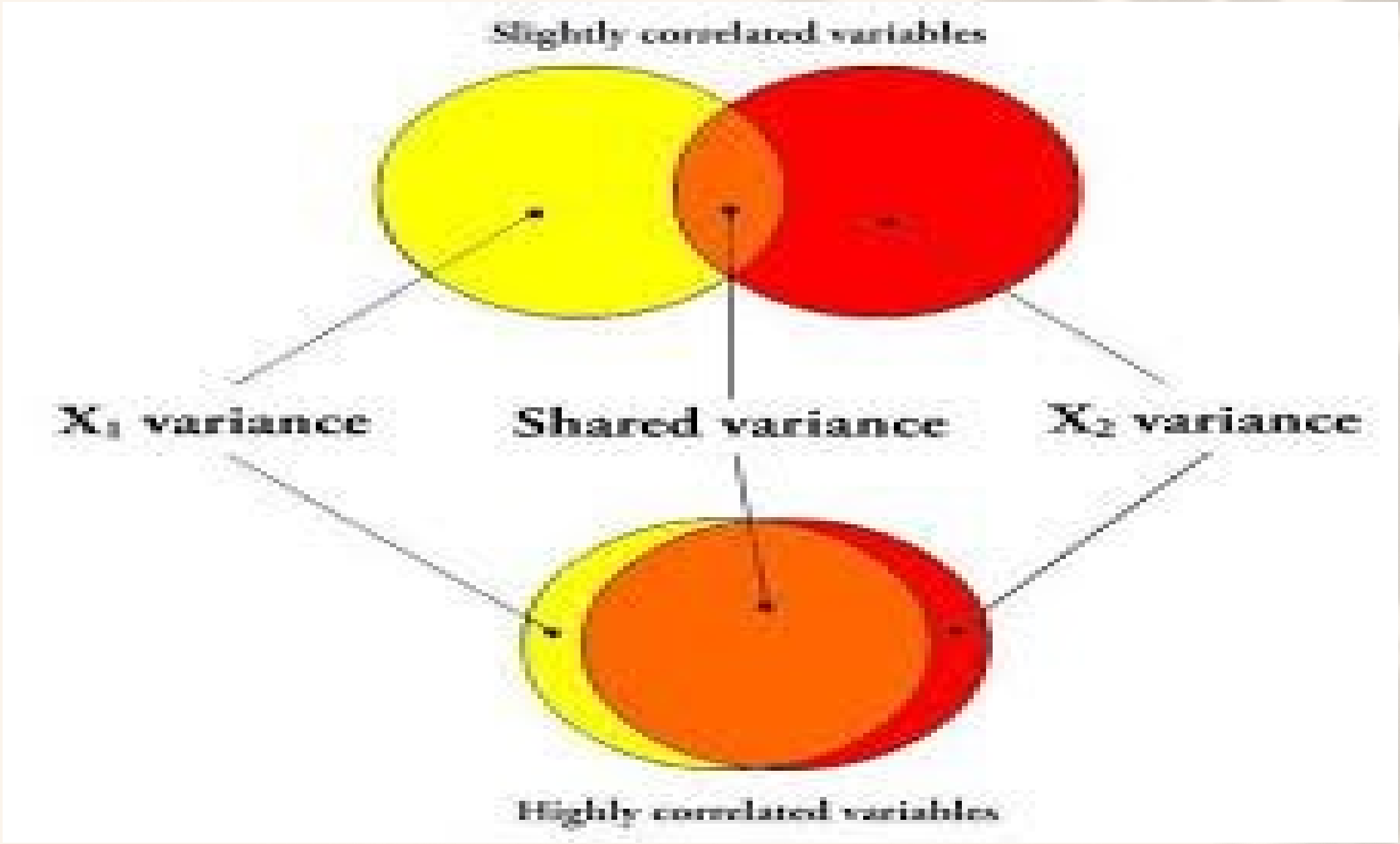
MODEL SPECIFICATION

- Construct correlation matrix
- Checking the relationship between Y and X(s)
- Check the multicollinearity between X(s)

MULTICOLLINEARITY

two explanatory variables that have high correlation.

correlation between two explanatory variables less than -0.7 or greater than 0.7 may be cause for concern


Slightly correlated variables
X_1 variance
Shared variance
X_2 variance
Highly correlated variables

Example of multicollinearity situation:

Problem 1: Prediction analysis on sales of lemonade for XYZ Café. Variables that might help to explain lemonade sales are, outside temperature and air-conditioning bills.

If the researcher includes both explanatory variables in the model, he may get results that are a little strange, because the two explanatory variables are themselves highly correlated.

As temperatures increase, so do air-conditioning bills. It would be meaningless to include both variables in the model because they are both doing the same job when it comes to explaining lemonade sales.

If exists explanatory variables that are highly correlated among themselves watch out for strange results in the regression output.

The strange results:

- Getting estimates of slope coefficients that are the opposite sign of what we would expect.

x_1	x_2	y
11.4	−9.7	14.7
12.5	−11.5	38.8
16.4	−15.9	42.9
14.4	−13.9	45.7
15.3	−14.2	52.3
18	−18.5	55.9
19.5	−21.2	60.1
25.2	−27.2	72.6

Example: Refer to the dataset given and answer the questions:
(a) Find the correlation matrix among all three variables.
(b) Find the least-squares regression model using both x_1 and x_2 as explanatory variables.
(c) Comment on the effect that including both x_1 and x_2 has on the t-test statistics.

Solution:
(a) The correlation matrix is as below.

	x1	x2
x2	-0.99607612	
y	0.89091722	-0.89445738

An extremely high correlation exists between *x*1 and *x*2, so multicollinearity exists between the two variables.

Solution:

(b) Find the least-squares regression model using both x_1 and x_2 as explanatory variables.

Parameter estimates:

Parameter	Estimate	Std. Err.	Alternative	DF	T-Stat	P-value
Intercept	3.1966894	35.467992	≠ 0	5	0.090128851	0.9317
x1	-0.015176864	8.8287255	≠ 0	5	-0.0017190323	0.9987
x2	-2.7209724	6.8440747	≠ 0	5	-0.39756615	0.7074

Analysis of variance table for multiple regression model:

Source	DF	SS	MS	F-stat	P-value
Model	2	1645.8514	822.92568	10.003384	0.0179
Error	5	411.32364	82.264729		
Total	7	2057.175			

Summary of fit:
Root MSE: 9.0699906
R-squared: 0.8001
R-squared (adjusted): 0.7201

The *P*-value for the *F*-test statistic is 0.0179, indicating that the regression model is significant. However, if we look at each individual *t*-test statistic, we see that all regression parameter has a very high *P*-value indicating that neither coefficient is different from zero.

(c) Comment on the effect that including both x_1 and x_2 has on the t-test statistics.

The contradictory results of the regression output occur because both x_1 and x_2 are related to the response variable y, as indicated by the correlation matrix.

However, x_1 and x_2 are also related to each other. So, with x_1 in the model, x_2 adds little explanation. Likewise, with x_2 in the model, x_1 adds little explanation.

The solution is to use only one explanatory variable. Which explanatory variable we choose is up to you. We can choose either the explanatory variable with the lower P value or the explanatory variable that has the higher correlation with the response.

VIF – VARIANCE INFLATION FACTORS

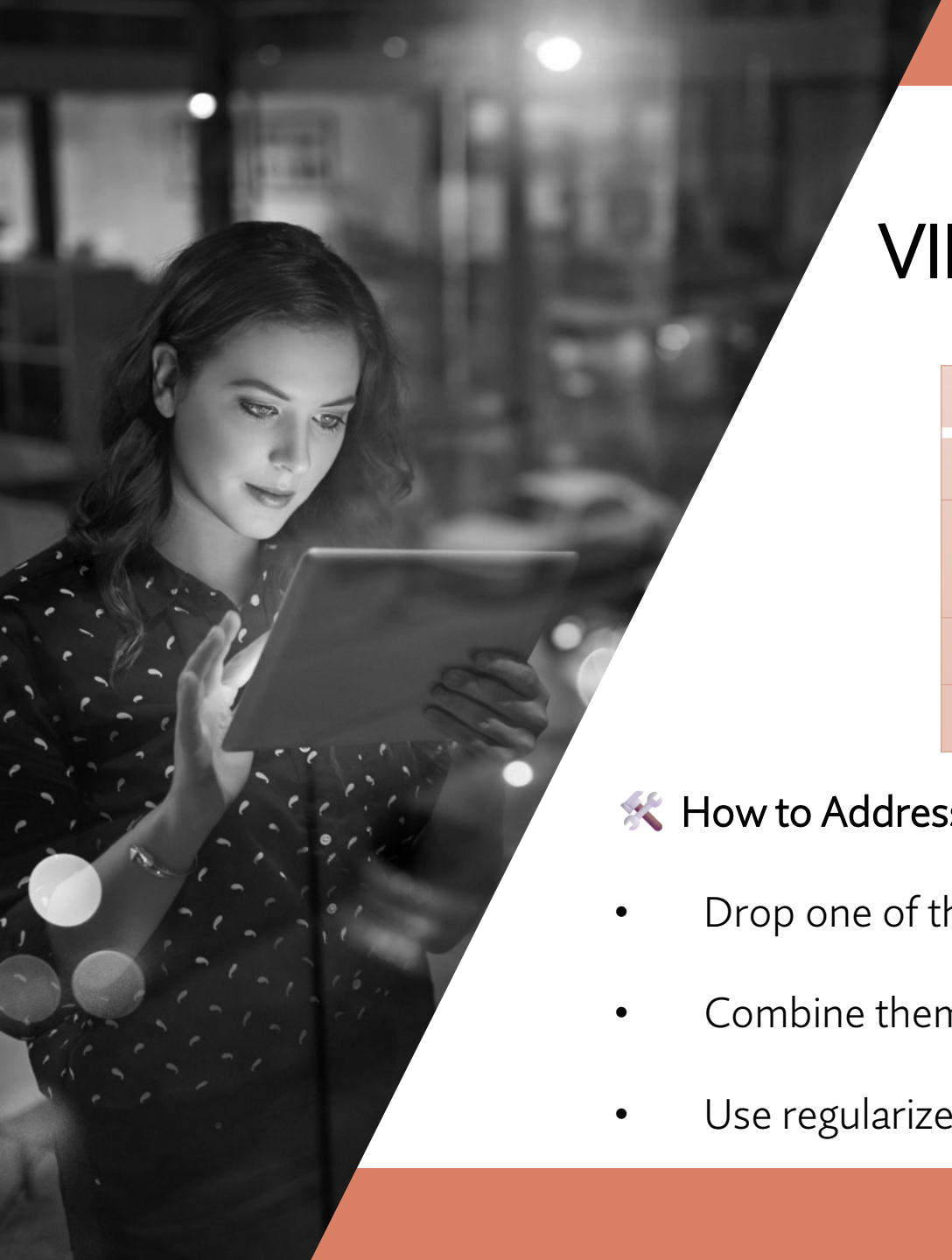
VIF is a statistical measure that helps detect multicollinearity in a multiple regression model.

🔍 How VIF Works:
For each predictor variable, VIF is calculated using this formula:

$$VIF_I = \frac{1}{1-R_i^2}$$

Where:

R_i^2 is the R-squared value from a regression of variable ii against all the other predictors.



VIF – VARIANCE INFLATION FACTORS

VIF Value	Interpretation
< 1	No multicollinearity
1–5	Moderate multicollinearity (acceptable)
>5	High multicollinearity (caution advised)
>10	Serious multicollinearity (problematic)

🛠️ How to Address High VIFs:

- Drop one of the highly correlated variables
- Combine them (e.g., using PCA)
- Use regularized models like Ridge or Lasso regression

✅ How to Run VIF in SPSS

Step 1: Open Your Data

Open your dataset in SPSS.

Step 2: Run Linear Regression

Go to Analyze > Regression > Linear...

Set your dependent variable (e.g., Readmission_Risk)

Move all your independent variables into the Independent(s) box

Step 3: Request Collinearity Diagnostics

Click on the "Statistics..." button

Check Collinearity diagnostics

Click Continue, then OK

Step 4: Interpret Output

In the output table, under Coefficients. you will see two columns:

Tolerance
VIF

MODEL SPECIFICATION

LET CHECK: Multicollinearity

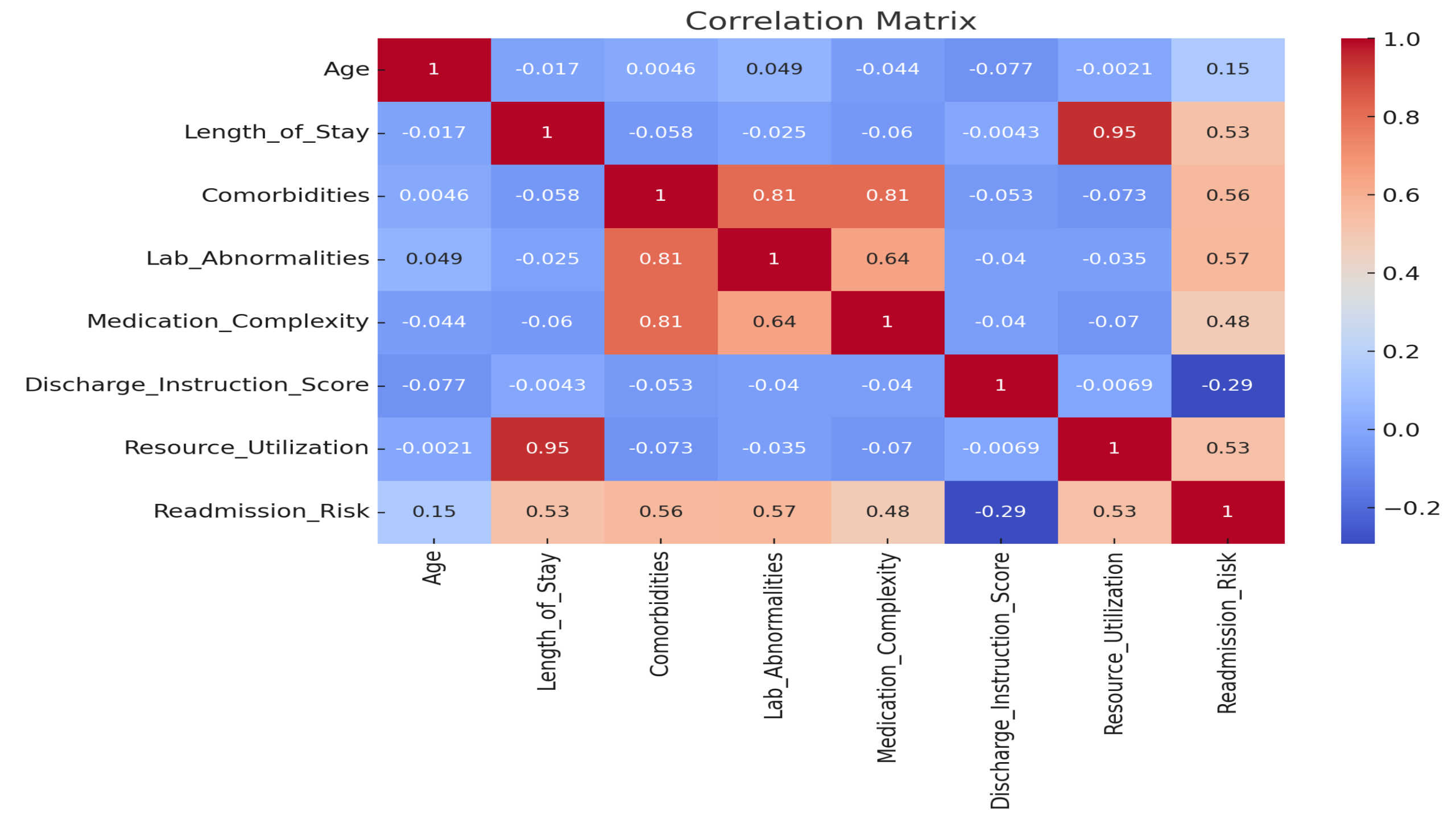
- Construct correlation matrix
- Checking the relationship between Y and X(s)
- Check the multicollinearity between X(s)

Do you find any multicollinearity among the explanatory variable?

Correlations

		Age	Length_of_Stay	Comorbidities	Lab_Abnormalities	Medication_Complexity	Discharge_Instruction_Score	Resource_Utilization	Readmission_Risk
Age	Pearson Correlation	1	-.014	.005	.049	-.042	-.079	-.002	.152**
	Sig. (2-tailed)		.751	.911	.278	.352	.079	.972	<.001
	N	500	500	500	500	500	500	500	500
Length_of_Stay	Pearson Correlation	-.014	1	-.052	-.025	-.052	.000	.947**	.531**
	Sig. (2-tailed)	.751		.241	.570	.249	.992	<.001	<.001
	N	500	500	500	500	500	500	500	500
Comorbidities	Pearson Correlation	.005	-.052	1	.798**	.792**	-.049	-.069	.564**
	Sig. (2-tailed)	.911	.241		<.001	<.001	.277	.123	<.001
	N	500	500	500	500	500	500	500	500
Lab_Abnormalities	Pearson Correlation	.049	-.025	.798**	1	.621**	-.032	-.041	.554**
	Sig. (2-tailed)	.278	.570	<.001		<.001	.472	.364	<.001
	N	500	500	500	500	500	500	500	500
Medication_Complexity	Pearson Correlation	-.042	-.052	.792**	.621**	1	-.048	-.063	.482**
	Sig. (2-tailed)	.352	.249	<.001	<.001		.282	.162	<.001
	N	500	500	500	500	500	500	500	500
Discharge_Instruction_Score	Pearson Correlation	-.079	.000	-.049	-.032	-.048	1	-.001	-.292**
	Sig. (2-tailed)	.079	.992	.277	.472	.282		.987	<.001
	N	500	500	500	500	500	500	500	500
Resource_Utilization	Pearson Correlation	-.002	.947**	-.069	-.041	-.063	-.001	1	.534**
	Sig. (2-tailed)	.972	<.001	.123	.364	.162	.987		<.001
	N	500	500	500	500	500	500	500	500
Readmission_Risk	Pearson Correlation	.152**	.531**	.564**	.554**	.482**	-.292**	.534**	1
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	<.001	
	N	500	500	500	500	500	500	500	500

**. Correlation is significant at the 0.01 level (2-tailed).


Correlation Matrix
Age
Length_of_Stay
Comorbidities
Lab_Abnormalities
Medication_Complexity
Discharge_Instruction_Score
Resource_Utilization
Readmission_Risk
1 -0.017 0.0046 0.049 -0.044 -0.077 -0.0021 0.15
-0.017 1 -0.058 -0.025 -0.06 -0.0043 0.95 0.53
0.0046 -0.058 1 0.81 0.81 -0.053 -0.073 0.56
0.049 -0.025 0.81 1 0.64 -0.04 -0.035 0.57
-0.044 -0.06 0.81 0.64 1 -0.04 -0.07 0.48
-0.077 -0.0043 -0.053 -0.04 -0.04 1 -0.0069 -0.29
-0.0021 0.95 -0.073 -0.035 -0.07 -0.0069 1 0.53
0.15 0.53 0.56 0.57 0.48 -0.29 0.53 1
Age
Length_of_Stay
Comorbidities
Lab_Abnormalities
Medication_Complexity
Discharge_Instruction_Score
Resource_Utilization
Readmission_Risk
1.0
0.8
0.6
0.4
0.2
0.0
−0.2

Do you find any multicollinearity among the explanatory variable?

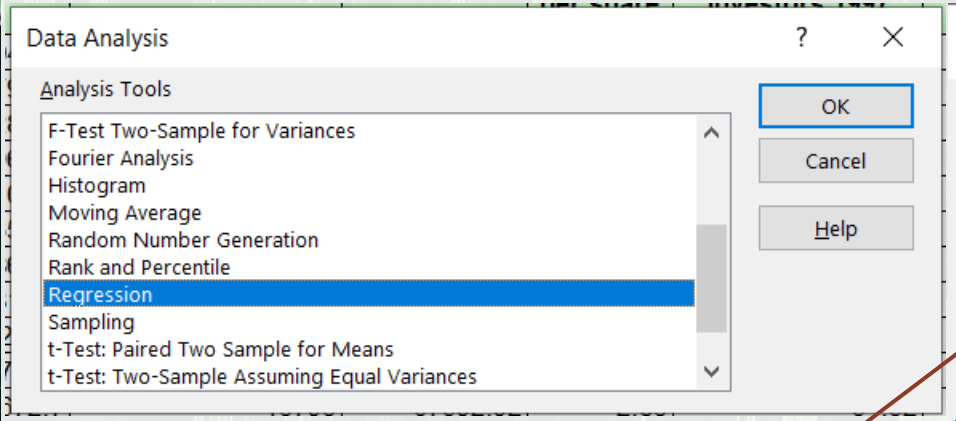
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.113	3.351		.631	.528		
	Age	.265	.046	.126	5.739	<.001	.981	1.019
	Length_of_Stay	1.273	.472	.184	2.696	.007	.102	9.775
	Comorbidities	4.210	.689	.284	6.106	<.001	.220	4.543
	Lab_Abnormalities	3.106	.439	.257	7.076	<.001	.361	2.768
	Medication_Complexity	1.478	.426	.125	3.474	<.001	.371	2.699
	Discharge_Instruction_Score	-1.994	.173	-.253	-11.555	<.001	.991	1.009
	Resource_Utilization	2.592	.445	.397	5.818	<.001	.102	9.789

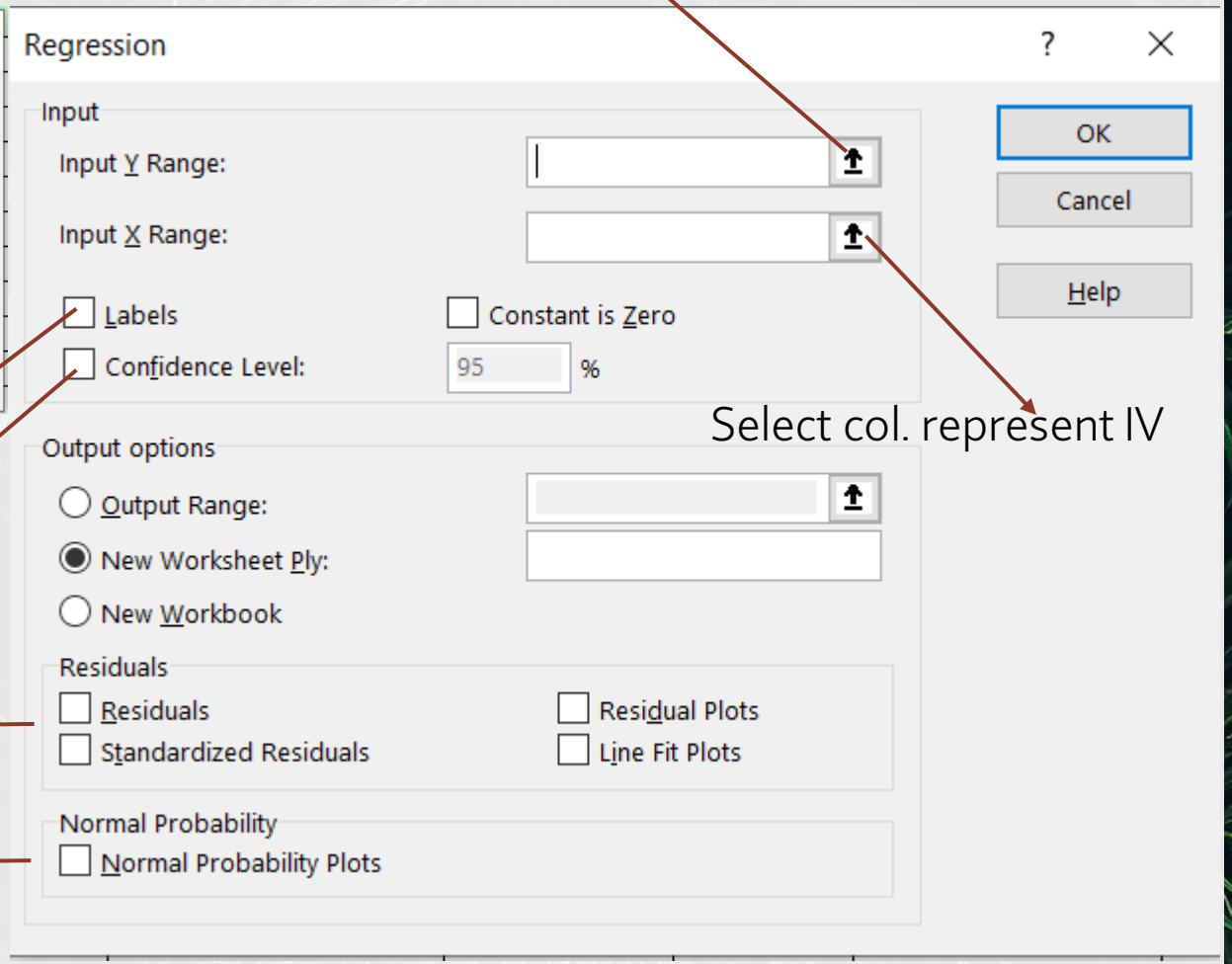
a. Dependent Variable: Readmission_Risk


MICROSOFT
EXCEL

Ms Excel -> Data -> Data Analysis



Select col. represent DV



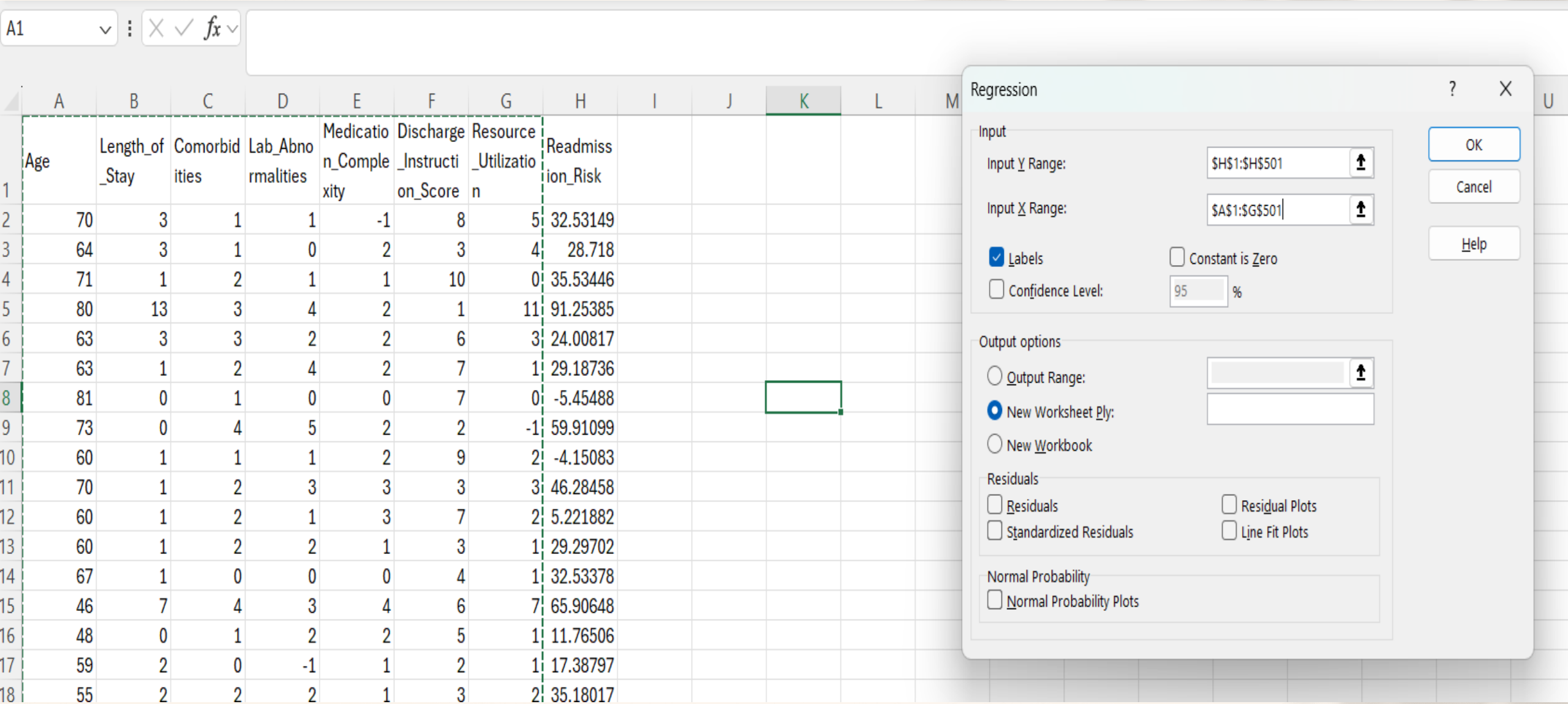
Select col. represent IV

Make 1st row as label

Changing the confidence level

Output for error (residual) diagnosis)

To see either normal or not normal distribution



	A	B	C	D	E	F	G	H
1	Age	Length_of_Stay	Comorbidities	Lab_Abnormalities	Medication_Complexity	Discharge_Instruction_Score	Resource_Utilization	Readmission_Risk
2	70	3	1	1	-1	8	5	32.53149
3	64	3	1	0	2	3	4	28.718
4	71	1	2	1	1	10	0	35.53446
5	80	13	3	4	2	1	11	91.25385
6	63	3	3	2	2	6	3	24.00817
7	63	1	2	4	2	7	1	29.18736
8	81	0	1	0	0	7	0	-5.45488
9	73	0	4	5	2	2	-1	59.91099
10	60	1	1	1	2	9	2	-4.15083
11	70	1	2	3	3	3	3	46.28458
12	60	1	2	1	3	7	2	5.221882
13	60	1	2	2	1	3	1	29.29702
14	67	1	0	0	0	4	1	32.53378
15	46	7	4	3	4	6	7	65.90648
16	48	0	1	2	2	5	1	11.76506
17	59	2	0	-1	1	2	1	17.38797
18	55	2	2	2	1	3	2	35.18017

Regression Model Building and Diagnostic

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.874948681
R Square	0.765535194
Adjusted R Square	0.762199313
Standard Error	10.03168192
Observations	500

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	7	161659.0822	23094.15	229.4851	1.6361E-150
Residual	492	49512.24395	100.6346		
Total	499	211171.3261			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	2.113373435	3.350539192	0.630756	0.528493	-4.46975711	8.696504
Age	0.264513525	0.046088206	5.739289	1.66E-08	0.17395954	0.355068
Length_of_Stay	1.273095669	0.472189717	2.696153	0.007255	0.345338562	2.200853
Comorbidities	4.209721227	0.689416504	6.106209	2.07E-09	2.855157503	5.564285
Lab_Abnormalities	3.10565441	0.438892191	7.076121	5.13E-12	2.243320195	3.967989
Medication_Complexity	1.478398588	0.425511594	3.474403	0.000557	0.642354533	2.314443
Discharge_Instruction_Score	-1.993580038	0.172526552	-11.5552	1.72E-27	-2.332559753	-1.6546
Resource_Utilization	2.592036335	0.44548748	5.818427	1.07E-08	1.716743712	3.467329

Compute the regression equation:

The estimate of the multiple regression model are:

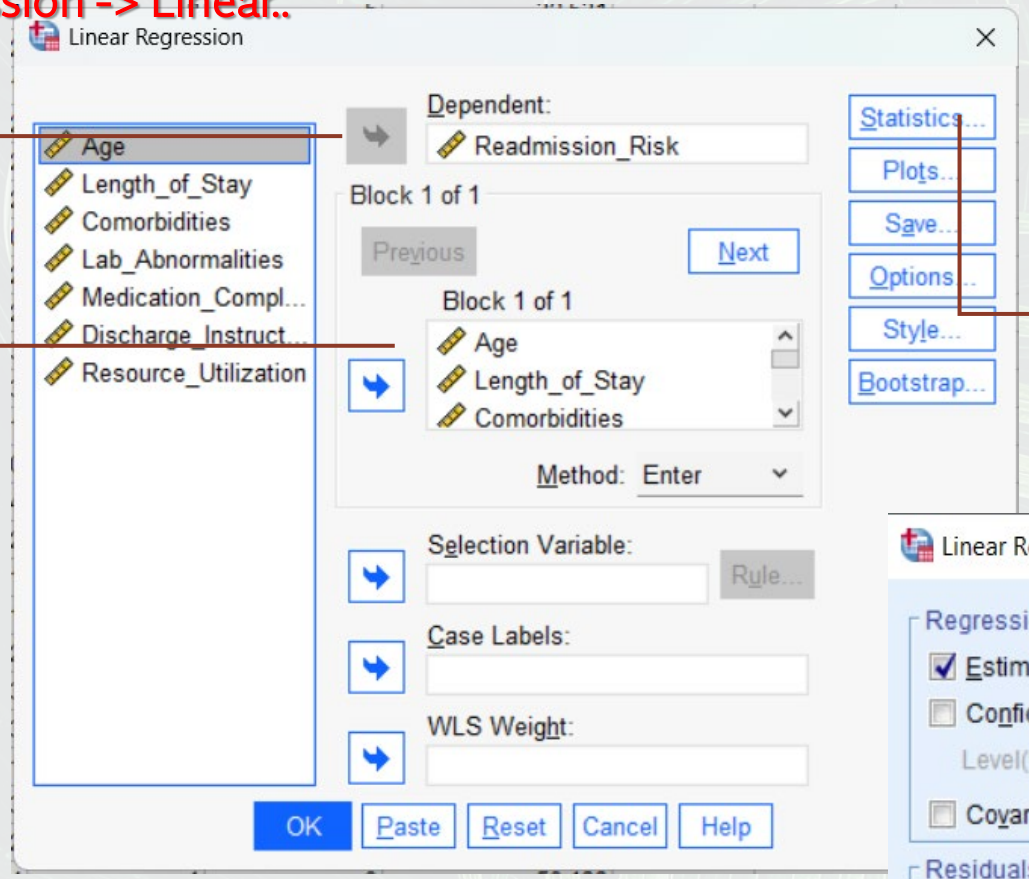
Readmission_Risk=2.11+0.26(Age)+1.27(Length_of_Stay)+4.21(Comorbidities)+3.11(Lab_Abnormalities)+1.48(Medication_Complexity)−1.99(Discharge_Instruction_Score)+2.99(Resource_Utilization)

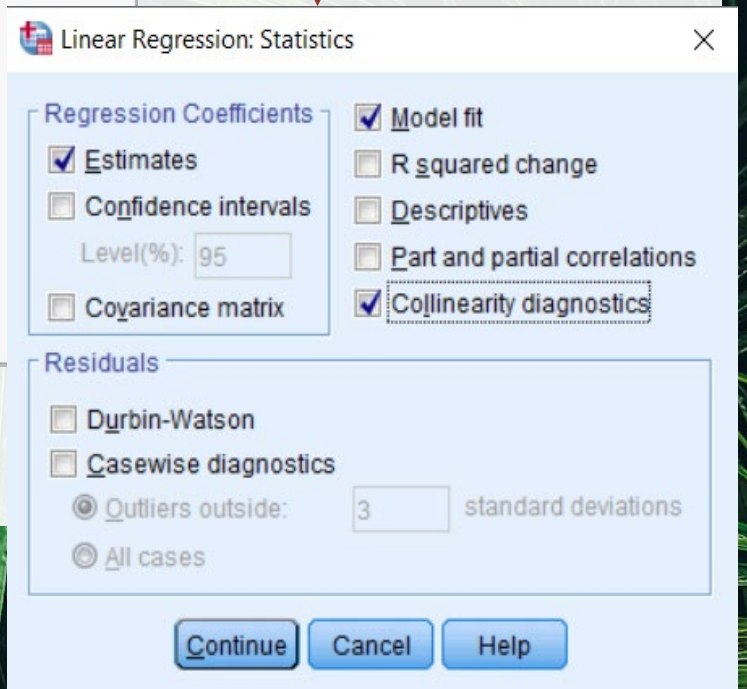

SPSS

SPSS-> Analyze -> Regression -> Linear..

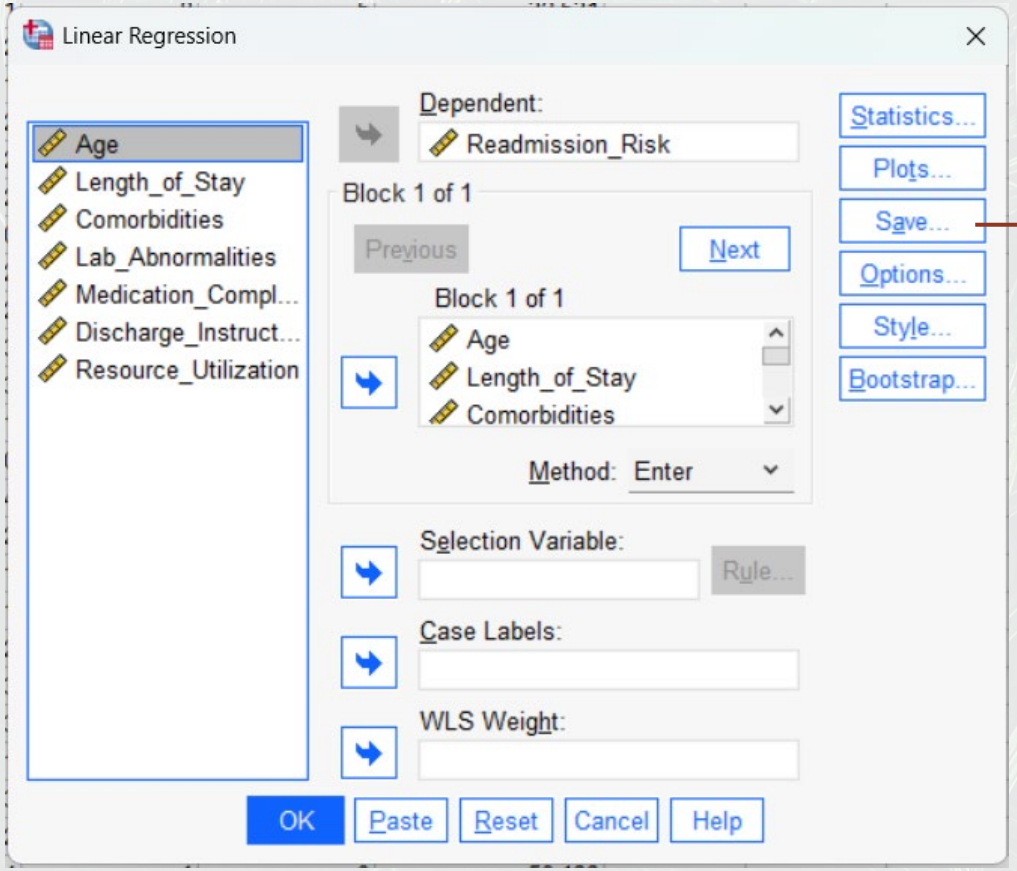
Dependent (Y) variable

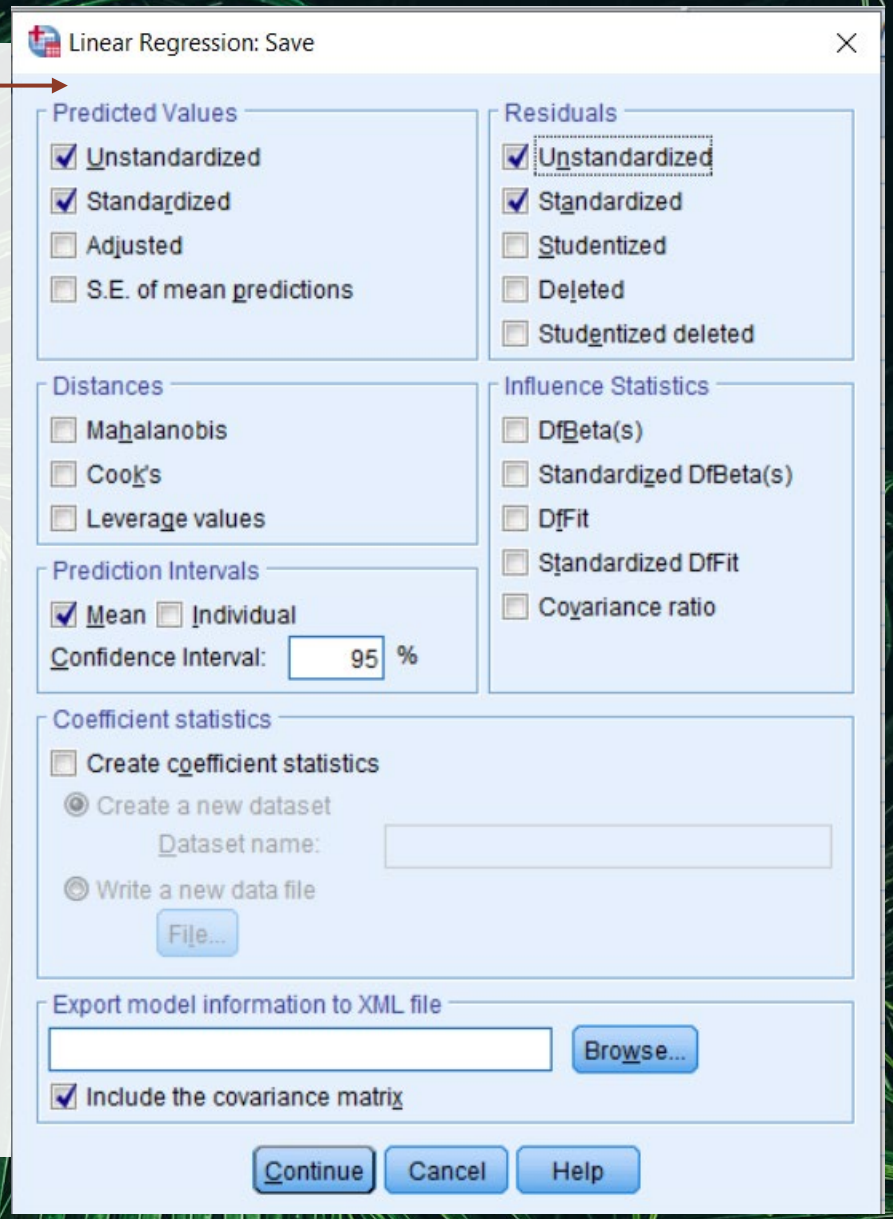
Independent (X) variable





SPSS-> Analyze -> Regression -> Linear..





Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.875[a]	.766	.762	10.031682

a. Predictors: (Constant), Resource_Utilization, Discharge_Instruction_Score, Lab_Abnormalities, Age, Medication_Complexity, Comorbidities, Length_of_Stay

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	161659.082	7	23094.155	229.485	<.001[b]
	Residual	49512.244	492	100.635		
	Total	211171.326	499			

a. Dependent Variable: Readmission_Risk

b. Predictors: (Constant), Resource_Utilization, Discharge_Instruction_Score, Lab_Abnormalities, Age, Medication_Complexity, Comorbidities, Length_of_Stay

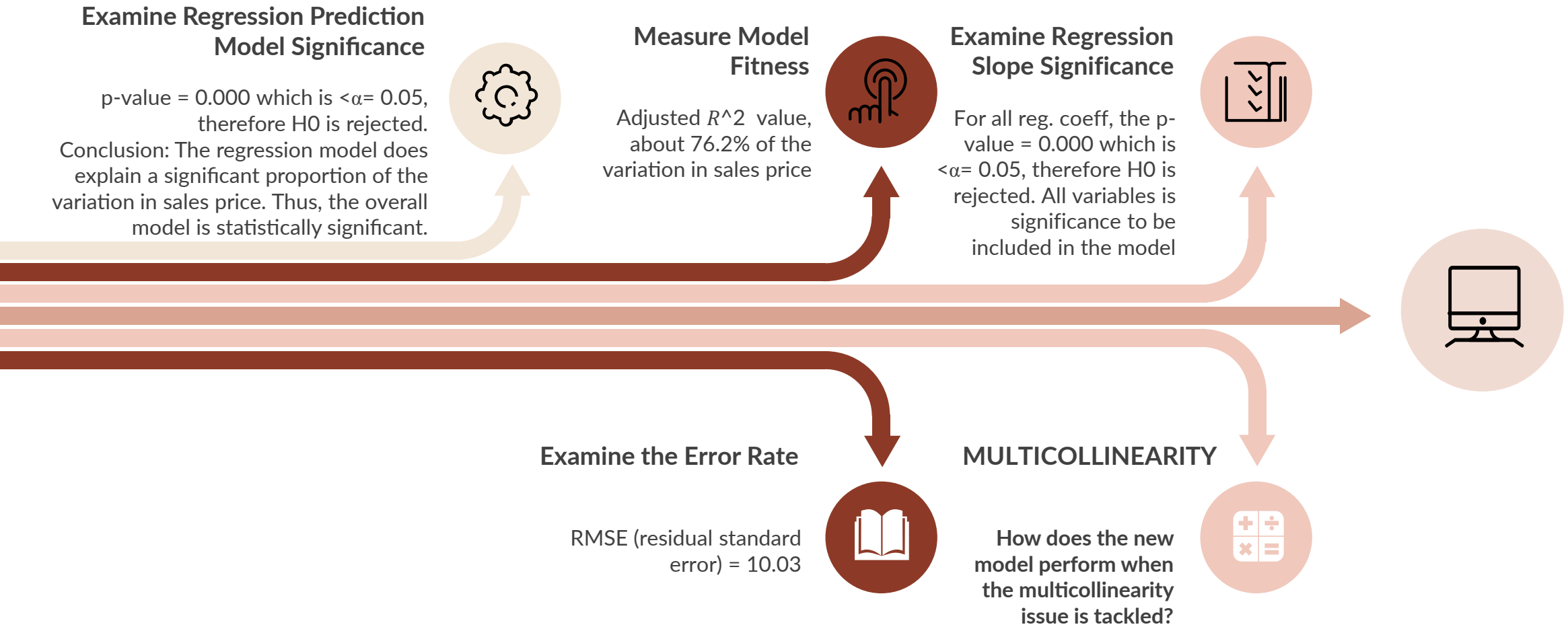
Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	2.113	3.351		.631	.528		
	Age	.265	.046	.126	5.739	<.001	.981	1.019
	Length_of_Stay	1.273	.472	.184	2.696	.007	.102	9.775
	Comorbidities	4.210	.689	.284	6.106	<.001	.220	4.543
	Lab_Abnormalities	3.106	.439	.257	7.076	<.001	.361	2.768
	Medication_Complexity	1.478	.426	.125	3.474	<.001	.371	2.699
	Discharge_Instruction_Score	-1.994	.173	-.253	-11.555	<.001	.991	1.009
	Resource_Utilization	2.592	.445	.397	5.818	<.001	.102	9.789

a. Dependent Variable: Readmission_Risk

Readmission_Risk = 2.11+0.26(Age) +1.27(Length_of_Stay) +4.21(Comorbidities) +3.11(Lab_Abnormalities) +1.48(Medication_Complexity) −1.99(Discharge_Instruction_Score) +2.99(Resource_Utilization)

REGRESSION MODEL DIAGNOSTIC



Variable	Coefficient	Meaning
Age	+0.26	Older patients are more likely to be readmitted.
Length_of_Stay	+1.27	Longer hospital stays suggest more serious conditions = higher risk.
Comorbidities	+4.21	More chronic diseases = higher chance of returning to hospital.
Lab_Abnormalities	+3.11	Abnormal lab results suggest complications = increased risk.
Medication_Complexity	+1.48	Complex medication regimens may lead to non-compliance = higher risk.
Resource_Utilization	+2.99	Higher use of resources implies complex care = increased risk.
Discharge_Instruction_Score	−1.99	Better patient understanding of discharge instructions reduces risk.

ENHANCED REGRESSION MODEL

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.866[a]	.749	.746	10.360553

a. Predictors: (Constant), Discharge_Instruction_Score, Length_of_Stay, Lab_Abnormalities, Age, Medication_Complexity, Comorbidities

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	158252.188	6	26375.365	245.716	<.001[b]
	Residual	52919.138	493	107.341		
	Total	211171.326	499			

a. Dependent Variable: Readmission_Risk

b. Predictors: (Constant), Discharge_Instruction_Score, Length_of_Stay, Lab_Abnormalities, Age, Medication_Complexity, Comorbidities

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1.621	3.459		.469	.640		
	Age	.275	.048	.131	5.778	<.001	.982	1.018
	Length_of_Stay	3.876	.156	.560	24.796	<.001	.996	1.004
	Comorbidities	4.088	.712	.276	5.744	<.001	.220	4.539
	Lab_Abnormalities	3.085	.453	.255	6.806	<.001	.361	2.768
	Medication_Complexity	1.506	.439	.127	3.426	<.001	.371	2.698
	Discharge_Instruction_Score	-1.997	.178	-.254	-11.207	<.001	.991	1.009

a. Dependent Variable: Readmission_Risk

ADDING QUALITATIVE INDEPENDENT VARIABLE

QUALITATIVE INDEPENDENT VARIABLE

There is situation you may wish to use a qualitative (lower level) variable as an explanatory variable in a regression model

Example: use of variable such as; marital status, gender, education level or job performance.

How these variable can be incorporated into a multiple regression analysis?

Dummy (or indicator) variable – a variable that is assigned a value equal to either 0 or 1, depending on whether the observation possesses a given characteristic.

x_1 = 1 if female

x_1 = 0 if male

if more than two mutually exclusive (for example, never married, married, divorced)

x_1 = 1 if never married, 0 if not

x_2 = 1 if married, 0 if not

x_3 = 1 if divorced, 0 if not

Example:

The population from which the sample was selected consists of executives between the ages of 24 and 60 who are working in U.S. manufacturing businesses. Data for annual salary (y) and age (x_1) is describe in the table. The objective of the problem is to determine whether a model can be generated to explain the variation in annual salary for business executives given the explanatory variable of age and the qualitative variable (x_2) of had a master of business administration (MBA) degree. The dummy variable is hold with this indication:

x_2 = 1 if holds MBA degree

x_2 = 0 if did not hold MBA degree

TABLE 15.2 Executive Salary Data Including MBA Variable

Salary($)	Age	MBA
65,000	26	0
85,000	28	1
74,000	36	0
83,000	35	0
110,000	35	1
160,000	40	1
100,000	41	0
122,000	42	1
85,000	45	0
120,000	46	1
105,000	50	0
135,000	51	1
125,000	55	0
175,000	50	1
156,000	61	1
140,000	63	0

From the statistical packages, the estimated regression equation are:

$$\hat{y} = 6{,}974 + 2{,}055x_1 + 35{,}236x_2$$

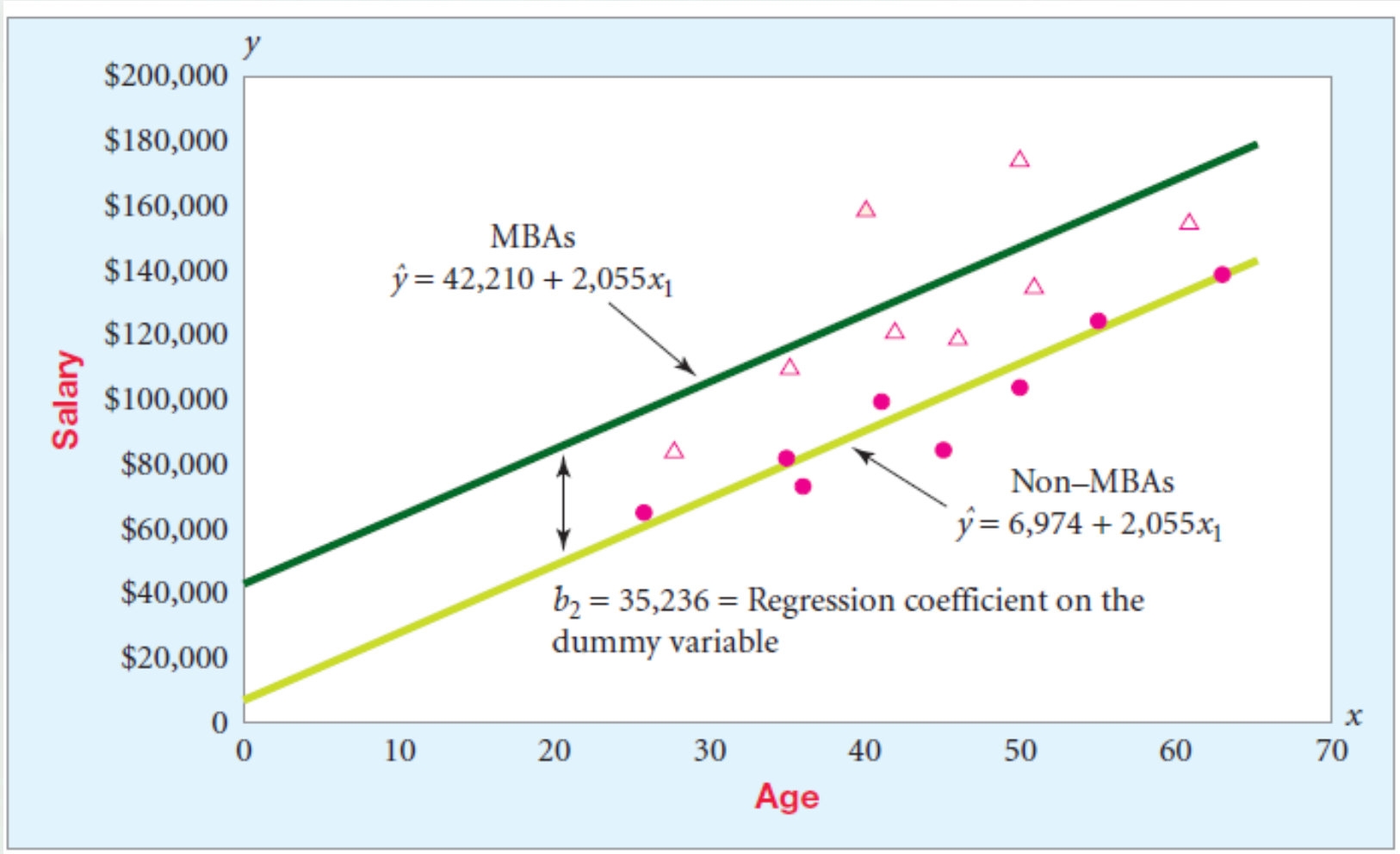
Because the dummy variable, x_2, has been coded 0 or 1 depending on MBA status, incorporating it into the regression model is like having two simple linear regression lines with the same slope, but different intercept.

For instance when x_2= *1* (respondent who holds MBA degree), $\hat{y} = 6{,}974 + 2{,}055x_1 + 35{,}236(1)$

$$\hat{y} = 42{,}210 + 2{,}055x_1$$

and when x_2= *0* (respondent who did not holds MBA degree), $\hat{y} = 6{,}974 + 2{,}055x_1 + 35{,}236(0)$

$$\hat{y} = 6{,}974 + 2{,}055x_1$$


y
$200,000
$180,000
$160,000
$140,000
$120,000
$100,000
$80,000
$60,000
$40,000
$20,000
0
Salary
MBAs
$\hat{y} = 42{,}210 + 2{,}055x_1$
Non–MBAs
$\hat{y} = 6{,}974 + 2{,}055x_1$
$b_2 = 35{,}236$ = Regression coefficient on the dummy variable
0
10
20
30
40
50
60
70
x
Age

NEW DATA SET:

Readmission_data_final_category.xlsx


SPSS

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.885[a]	.782	.779	9.892

a. Predictors: (Constant), Type of Discharge, Medication_Complexity, Age, Length_of_Stay, Discharge_Instruction_Score, Lab_Abnormalities, Comorbidities

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	173064.357	7	24723.480	252.676	<.001[b]
	Residual	48140.561	492	97.847		
	Total	221204.918	499			

a. Dependent Variable: Readmission_Risk

b. Predictors: (Constant), Type of Discharge, Medication_Complexity, Age, Length_of_Stay, Discharge_Instruction_Score, Lab_Abnormalities, Comorbidities

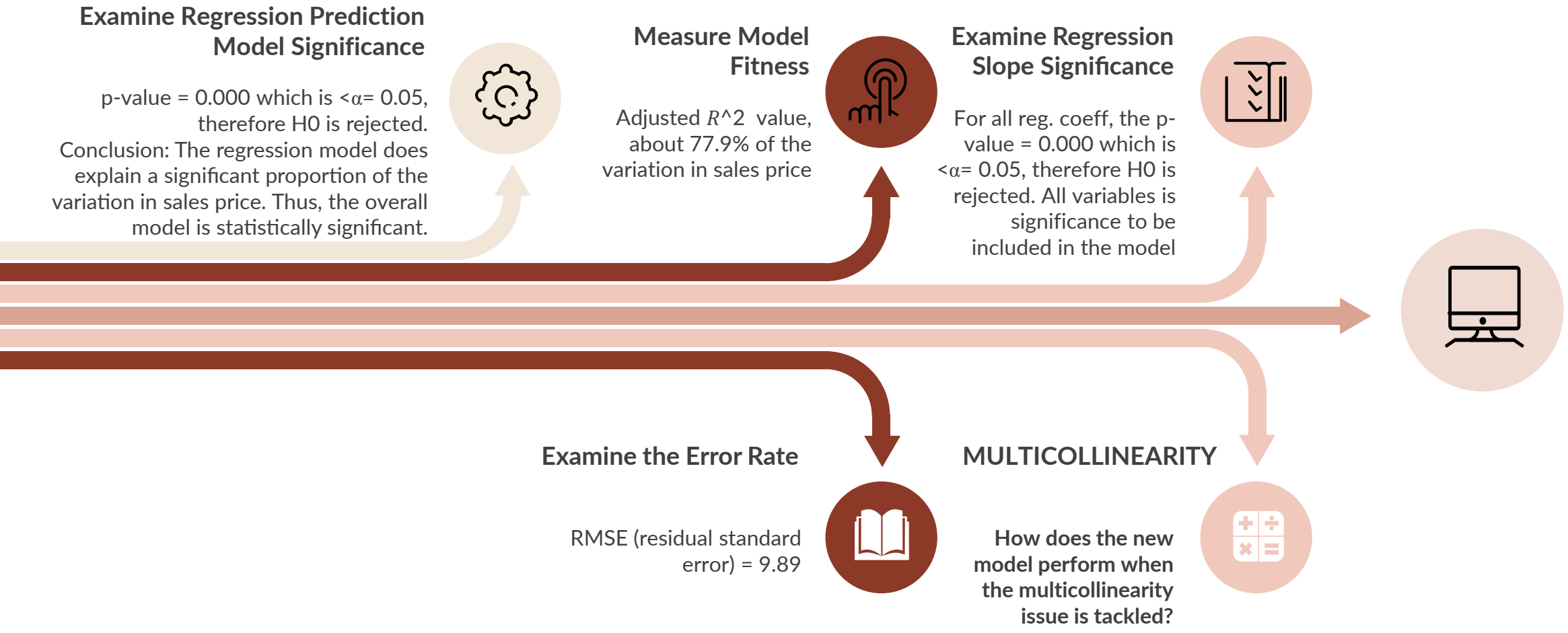
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	4.408	3.321		1.327	.185		
	Age	.232	.045	.108	5.100	<.001	.982	1.018
	Length_of_Stay	4.220	.149	.596	28.252	<.001	.994	1.006
	Comorbidities	2.522	.680	.166	3.712	<.001	.220	4.539
	Lab_Abnormalities	3.922	.433	.317	9.049	<.001	.360	2.777
	Medication_Complexity	2.125	.420	.175	5.063	<.001	.370	2.700
	Discharge_Instruction_Score	-1.997	.170	-.248	-11.738	<.001	.991	1.009
	Type of Discharge	7.584	.991	.162	7.654	<.001	.993	1.007

a. Dependent Variable: Readmission_Risk

Readmission_Risk = 4.41+ 0.23(Age) + 4.22(Length_of_Stay) + 2.52(Comorbidities) +3.92(Lab_Abnormalities) +2.13(Medication_Complexity) −1.99(Discharge_Instruction_Score) + 7.58(Type o Discharge)

REGRESSION MODEL DIAGNOSTIC



ENHANCED REGRESSION MODEL

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.881[a]	.776	.774	10.019

a. Predictors: (Constant), Type of Discharge, Medication_Complexity, Age, Length_of_Stay, Discharge_Instruction_Score, Lab_Abnormalities

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	171716.099	6	28619.350	285.102	<.001[b]
	Residual	49488.819	493	100.383		
	Total	221204.918	499			

a. Dependent Variable: Readmission_Risk

b. Predictors: (Constant), Type of Discharge, Medication_Comple Length_of_Stay, Discharge_Instruction_Score, Lab_Abnormal

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	5.561	3.349		1.660	.098		
	Age	.232	.046	.108	5.036	<.001	.982	1.018
	Length_of_Stay	4.203	.151	.594	27.792	<.001	.995	1.005
	Lab_Abnormalities	4.948	.338	.400	14.629	<.001	.607	1.648
	Medication_Complexity	3.096	.332	.255	9.320	<.001	.606	1.649
	Discharge_Instruction_Score	-2.010	.172	-.250	-11.665	<.001	.991	1.009
	Type of Discharge	7.626	1.004	.162	7.600	<.001	.993	1.007

a. Dependent Variable: Readmission_Risk

THANK YOU

Assoc. Prof. Dr. Nurulhuda Firdaus Bt Mohd Azmi (Huda)

012-5719624

huda@utm.my