Analysis Report

This report is structured as follows.

Contents

Sample Characterization

The sample was composed by a total of 75 individuals, in which 50 opted for non-surgical approaches while 25 undertook surgery. The 'age' variable was categorized in three intervals according to the data distribution: younger than 65 years old; between 65 and 74 and 75 years or older. The table below shows the frequencies of the other variables in the dataset.

Category	Level	Count	Percentage
Age_Groups	<65	16	21.3
	65-74	33	44.0
	>=75	26	34.7
Diabetes	no	50	66.7
	yes	25	33.3
Hypertension	no	32	42.7
	yes	43	57.3
Asthma	no	71	94.7
	yes	4	5.3
hISTORY_OF_PREVIOUS_SURGERY	no	35	46.7
	yes	40	53.3
PSA_LEVELS	high	44	59.5
	low	15	20.3
	medium	15	20.3
GLEASON	high	31	41.3
	low	9	12.0
	medium	35	46.7

The sample is quite balanced when it comes to the presence of Diabetes and Hypertension. The majority of the sample does not have Asthma (95%). 59% of the sample has high levels of PSA and 41% has high levels of GLEASON.

Chi-Square Results

In the Chi-Square Results, we observe significant associations in some categories. Specifically, the age group variable demonstrates a significant chi-square statistic ($\chi^2 = 6.171$, $p = 0.046$), indicating a meaningful association between different age groups and the surgery outcome. Notably, the highest surgery percentages were observed in the 65-74 age group (64.0%), followed by the >=75 group (24.0%), and the <65 group (12.0%). For the PSA_LEVELS, the chi-square statistic was significant ($\chi^2 = 6.669$, $p = 0.036$), suggesting a strong relationship between different PSA levels and the choice of surgery. A particularly high percentage of surgery was noted in the high PSA level group (76.0%). Other variables such as Diabetes, Hypertension, Asthma, History of Previous Surgery, and GLEASON did not show a statistically significant association with surgery outcomes.

Variable	Level	Surgery		Chi_Square	P_Value
		No (%)	Yes (%)		
Age_Groups	<65	26.0	12.0	6.171	0.046
	65-74	34.0	64.0		
	>=75	40.0	24.0		
Diabetes	no	66.0	68.0	0.000	1.000
	yes	34.0	32.0		
Hypertension	no	40.0	48.0	0.170	0.680
	yes	60.0	52.0		
Asthma	no	96.0	92.0	0.033	0.856
	yes	4.0	8.0		
hISTORY_OF_PREVIOUS_SURGERY	no	40.0	60.0	1.935	0.164
	yes	60.0	40.0		
PSA_LEVELS	high	51.0	76.0	6.669	0.036
	low	28.6	4.0		
	medium	20.4	20.0		
GLEASON	high	44.0	36.0	3.687	0.158
	low	16.0	4.0		
	medium	40.0	60.0		

Probit Model Results

Turning to the Probit Model Results, McFadden's R-squared of 0.216 indicates a moderate model fit. The model was found to be statistically significant ($\chi^2(11) = 20.673$, $p = 0.037$), suggesting that the predictors, as a whole, meaningfully contribute to predicting the surgery outcome. This model evaluated the effect of all variables in a multivariate approach. So the observed effect of one variable is calculated whilst controlling for the effect of all other variables. Examining individual predictors, the variable hISTORY_OF_PREVIOUS_SURGERY_yes showed a negative association with the likelihood of undergoing surgery (B = -0.837, OR = 0.433, p = 0.040), indicating that those with a history of previous surgery were less likely to undergo surgery again. Conversely, PSA_LEVELS_high was positively associated with surgery (B = 1.445, OR = 4.243, p = 0.030), suggesting that higher PSA levels increase the likelihood of undergoing surgery. Those with high PSA levels are 4.2 times more likely to choose surgery compared to those with low levels. Other variables, including Age_Groups, Diabetes, Hypertension, Asthma, and GLEASON levels, did not show statistically significant associations with the surgery outcome in this model.

Predictors	Coefficients	Odds_Ratios	P_Values
(Intercept)	-1.892	0.151	0.033
Age_Groups_65_to_74	0.558	1.748	0.259
Age_Groups_olderthan75	0.220	1.246	0.680
Diabetes_yes	-0.035	0.966	0.935
Hypertension_yes	-0.162	0.850	0.693
Asthma_yes	1.235	3.437	0.181
hISTORY_OF_PREVIOUS_SURGERY_yes	-0.837	0.433	0.040
PSA_LEVELS_high	1.445	4.243	0.030
PSA_LEVELS_medium	1.113	3.044	0.132
PSA_LEVELS_NA	-2.517	0.081	0.995
GLEASON_high	0.035	1.036	0.962
GLEASON_medium	0.796	2.216	0.256

Predictors	Coefficients	Odds_Ratios	P_Values
(Intercept)	-1.892	0.151	0.033