

# " Allocating Sales Target Using Multiple Linear Regression Technique."

## - *Setting Right Sales Targets-Dilemma of a National Sales Manager.*

<sup>1</sup>Dr.Aniruddha Bodhankar,

Associate Professor,

Dr.Ambedkar Institute of Management Studies & Research, Nagpur( India).

### **Abstract :**

Mr.Suvikas Banerjee aged, about 45 years, a down to earth Sales Person, has recently taken up the charge of National Sales Manager with a Mumbai (India) based Selestar Pharmaceuticals Ltd\*. Which is recently being launched in Indian market. Mr.Banerjee arose from various levels in Pharmaceuticals Industry to be a National Sales Manager by the age of his 45.Through out his career of 25 years in Sales, Mr.Suvikas Banerjee always was under the thought that the sales targets designed for the Sales Staff of Indian Pharmaceuticals Industry were irrational and were purely designed out of hunches. During his official journey from Mumbai to Nagpur, Suvikas met with Dr.Aniruddha Bodhankar, who also was from FMCG & Marketing Research Sector. En-route the discussions went on the attrition happening in Sales Industry due to non-achievement of targets. Dr.Aniruddha was also aware of these happenings due to irrational target setting by the line managers in the industry. Mr.Suvikas requested Dr.Aniruddha to develop a mechanism through which a rational and achievable sales targets can be designed for the field staff for a newly launching company. Dr.Aniruddha, although knowing fully the difficulty and challenges involved in the task, primarily while collecting of data from the retailers & distributors, agreed for the task. After reaching Nagpur, Dr.Aniruddha conducted a pilot survey to finalize the research variables. Basis the class of variables, he decided to develop a multiple linear regression model first for the sales prediction followed by the process of target setting.

**(Key words : Case Study On Multiple Linear Regression, Right Sales Target, Sales Target Setting, Sales Manager.)**

### **Objective of study :**

Developing a model for Setting right & achievable Sales Target for Sales Force using Analytics.

### **Research Setting:**

Basis the discussions held with Mr.Suvikas Banerjee, a pilot research was conducted in Vidarbha zone of Maharashtra State. This pilot research was conducted mainly to understand the psychology of the pharmaceutical retailers, the distribution pattern of the pharmaceuticals industry, the inventorying capacity of the retailers and the economics of the market. The pilot research was conducted with 157 pharmaceutical retailers mainly from the Semi-Urban and Rural Parts. The rural and semi urban areas were chosen primarily due to the fact that, those are the areas highly affected due to seasonality, schemes & offers and product substitutions. The doctors prescribing the products are from within the same areas and their prescriptions are honored next door pharmaceutical shop. The market was also selling the similar product range as that of Selestar Pharmaceuticals Co. Ltd.

(Selestar Pharmaceuticals Co Ltd. was having all me too range of products mainly targeted to GPs and Physicians as was told by Mr.Suvikas Banerjee). This pilot research revealed the various variables which can be used during the study. One important fact which was revealed was that the retailers do not maintain a huge inventory as they buy on daily basis from the wholesalers. Hence the Sales of the retailers (Cash Memos prepared towards the effected sale to the patients are same as their purchases from the wholesalers and the number of cash memos generated per day is equal to the footfall of the shop on that day). These variables were

1. Average daily billing by retailers - Outcome Variable - Continuous Variable
2. Average daily footfall - Response Variable - Continuous Variable
3. Average approximate cost of each prescription - Response variable -Continuous
4. Market Classification (Rural/Semi Urban/Urban)-Response variable - Nominal
5. Prescription Status (Floating/Area Rx/Next door Dr's Rx) - Response - Nominal
6. Availability of full range of Medicines - Response - Nominal
7. Type of prescribing doctor (GP/Consultant/Hospital) - Response - Nominal

The rationale behind selecting Vidarbha as area under study lies in the fact that Vidarbha and especially Nagpur is the centre of India and is a cosmopolitan area. The patients from entire Central Zone, M.P; Chhattisgarh, Marathwada visit Nagpur for medical treatment. The cost of treatment is quite economical here compared to rest of India. And more appropriately the psychology of the patients, doctors and retailers is the replica of that of entire India so the results drawn from Central India can be applied to entire India. There are around 20000 pharmaceutical retailers in Central India and almost every pharmaceutical company operates in Central India. We planned to draw a sample of about 1000 pharmaceutical retailers from Central India, which was 5% of the total area census.

The samples were drawn predominantly from the areas where the stand alone prescribers, such as GPs and Physicians were present, instead of a bunch of Corporate Hospitals to avoid the possible outliers. Specialists and Super Specialists doctors areas were also purposefully avoided as their prescription habits did not match with the product mix of Selestar. Finally we landed up at 708 pharmaceutical retailers which constituted our samples for study.

Selestar Pharmaceuticals Ltd; is a newly launching pharmaceutical company promoted by Mr.Jagjit Singh Ahaluwalia, who diversified into pharmaceuticals manufacturing and marketing from Transportation and Shipping line business. Currently Selestar is in the process of one manufacturing unit at Baddi in Himachal Pradesh and is planning for all India operation with a field force of 300 people. Selestar plans to invest Rs.500 Crore in this business in next five years. Selestar has currently a plan to introduce all me too products in the market to gain the initial 3% market share of Rs.60000 Crore (OPPI Data) domestic ethical pharmaceutical market in the very first year. Selestar is poised to gain a market share of 15% during next 5 years with their aggressive marketing strategy.

Looking into the types of variables, we decided to build a model based on Multiple Linear Regression Technique as the case was a fit for MLR. The model was build on SPSS Version 20.

This study began in the month of April 2018 and was continued till September 2018. Total number of students deployed on this project was 15 along with 2 faculty members as their mentors. The students undertook this study as a live project. Out of these 15 students, 12 were deployed for the fieldwork where as 3 students were handling the back end operations including data entry.

The data was collected from the pharmaceutical retailers for the period of April-2018 till September-2018 on the following heads.

1. Demographic details of the retailer, such as name of the shop, Drug License Number, Area (Location) etc.
2. Average daily sales to patients (000'Rs) this was equivalent to average daily purchases by the same retailer. Out of planned 1000 retailers 706 retailers only provided the data (n=706)
3. Average daily effective footfall (Captured as Number of sales bills generated per day)
4. Average cost of each prescription (daily Sales divided by daily footfall)
5. Market Classification (With dummy variables as If Rural= 0, If Semi-Urban= 1 & If Urban=2)
6. Prescription Status (With dummy Variables as If Prescription is Floating = 0, If Prescription is from doctor from Same Area = 1 & If the Prescription is from Next door doctor = 2)

7. Availability of full range of Medicines (With dummy variables as Never = 0, Sometimes = 1 & Always = 2)
8. Type of prescribing doctor (with dummy Variables as If Big Hospital = 0, If Exclusive Consultant = 1 and If General Practitioner = 2)
- The data so collected along with dummy variables was recorded in SPSS Version 20 for further analysis.

### Assumptions for Regression Model Building

For operating a Multiple Linear Regression on the data, the data so collected must satisfy the following eight conditions.

1. There should be only one outcome (Dependent) variable which should be as continuous variable. (Criterion Satisfied)
2. There should be more than one response (Independent) variables which can be either continuous or nominal variables. (Criterion Satisfied)
3. Observations should be independent of each other i.e. there should be no auto correlation amongst the observations.
4. There should exist a linear relationship between the outcome variable and each of the response variable individually as well as collectively.
5. The data should have homoscedasticity.
6. There should be no multicollinearity in the data.
7. The data should not possess significant outliers or influential parameters.
8. Errors should be normally distributed.

### Estimating Regression Parameters

Out of 706 samples, 607 samples were used for regression model building where as 99 samples were used for regression model validation. And out of total eight variables, only four variables (Daily Purchase, Daily Footfall, Market Classification and prescription status were used for regression model building. The other four variables were removed from the study as they were violating the assumptions of multiple linear regression which were destabilizing the regression model). While building the regression model, we used the direct method of model building instead of step wise regression model building.

**Table 1: Descriptive Statistics**

	Mean	Std. Deviation	N
Retailer's Daily Purchases	4736.97	2098.441	607
Average Daily Footfall	30.96	19.314	607
Market Classification	1.37	.833	607
Prescription Status	.94	.701	607

Since we have used direct method for regression model building, upon checking Pearson Correlation, we observed that NONE of the response variable or group of response variables show multicollinearity with the outcome variable (Retailer's daily purchase). Thus criterion six of no multicollinearity in the variables is satisfied.

**Table 3: Model Summary<sup>b</sup> (Test For checking Auto-Correlation)**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.545 <sup>a</sup>	.297	.293	1763.993	1.904

a. Predictors: (Constant), Classification of Rx, Average Daily Footfall, Market Classification

b. Dependent Variable: Retailer's Daily Purchases

**Interpretation** : From the Table 3, it is evident that the Coefficient of Determination ( $R^2$ ) is 0.297 indicates that 29.7% variations in the model are explained by these variables. The Durbin-Watson Index of 1.904 indicates that there is no auto-correlation among the variables. Thus the observations are independent of each other. This satisfies the third assumption of regression - No auto-Correlation among the observations.

**Table 4: ANOVA <sup>a</sup> (Test For checking Regression Model Fitness)**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	792156161.419	3	264052053.806	84.859	.000 <sup>b</sup>
	Residual	1876337967.105	603	3111671.587		
	Total	2668494128.524	606			

a. Dependent Variable: Retailer's Daily Purchases

b. Predictors: (Constant), Classification of Rx, Average Daily Footfall, Market Classification

**Interpretation:** Table 4 shows the significance level of 0.000 which indicates the entire regression model is significant at  $F(603, 3) = 84.86$ .

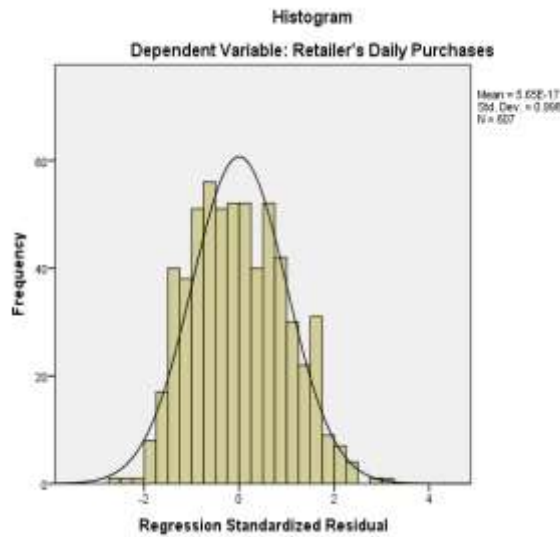
**Table 5 : Regression Coefficients <sup>a</sup> ( Test For Model Significance & Multicollinearity)**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics		
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero - order	Partial	Part	Tolerance	VIF	
1	(Constant)	1489.624	320.061		4.654	.000	861.055	2118.192						
	Average Daily Footfall	59.427	3.832	.547	15.509	.000	51.902	66.953	.506	.534	.530	.937	1.067	
	Market Classification	681.116	114.901	.270	5.928	.000	455.462	906.770	.128	.235	.202	.560	1.785	
	Prescription Status	509.308	137.874	.170	3.694	.000	238.536	780.080	-.085	.149	.126	.550	1.818	

a. Dependent Variable: Retailer's Daily Purchases

**Interpretation** : From the table 5, it is evident that the regression constant  $B_0$  , and all the other response variables such as Average Daily Footfall, Market Classification and Prescription Status are significant (Significance Level  $<0.005$ ). The table further opines that at 95% level of confidence, the regression constant  $B_0$  will fall between 861.05 and 2118.192 similarly the average daily foot fall would be between 51.90 (Practically 52) and 66.95 (Practically 67). The tolerance values ( $>0.1$ ) shows that there is no multicollinearity between the variables.

To check assumption - The Residuals are almost Normally Distributed



Histogram

**Interpretation :** The Histogram above shows the errors are almost equally distributed. This satisfies the regression assumption number 8.

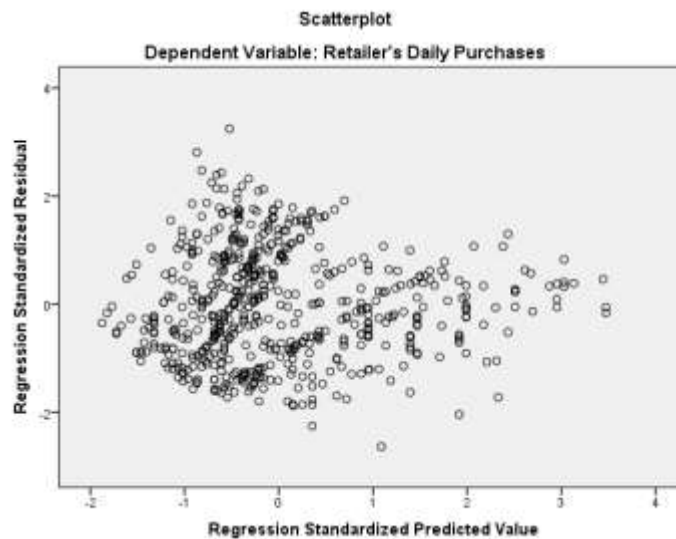
To check assumption of any possible outlier in the data



Normal P-P Plot

**Interpretation :** The normal P-P Plot shows there are no outliers in the data, which satisfies the assumption for no outliers. Thus the assumption no.7 is also satisfied.

To detect the assumption of linearity between the variables



Scatter Plot

**Interpretation :** The above scatter plot shows that the relationship between the outcome variable and response variable is near to linear as the points are equally spread on both the sides of the 45° angle bisector line. This satisfies the assumption 4 of linear multiple regression.

Thus all the eight assumptions of multiple linear regression are completely satisfied in this case of Selestar Pharmaceuticals Ltd.

### Regression Model Building and Model Validation For Selestar Pharmaceuticals Ltd.

The regression coefficients are derived in Table 5. With reference to these coefficients, we built a multiple linear regression model for Selestar Pharmaceuticals Ltd. The model so constructed reads as below -

$$\text{Predicted Purchase (Y) by each retailer} = 1481.624 + (59.427 * \text{Average Daily Footfall}) + (681.116 * \text{Market Classification}) + (509.308 * \text{Prescription Status})$$

We have kept the data of 99 samples reserved for validating the regression model. For validating the above developed model, we used these 99 samples. During model validation, we took the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) of the data (n=99). The calculated figure for MAPE are 16.66% whereas the same for RMSE are 1959. Depending on the context of the research where a lot of human thought process from the pharmaceutical retailers is involved, I considered the MAPE of 16.66% as fair enough and hence the model is validated and to be accepted.

### Recommendation :

From the above research, we would like to suggest Mr. Suvikas Banerjee, that the exact sales targets for every sales person can be arrived by replacing the appropriate values for the variables in the above model and solving it & then by summing up these values for each sales person for the number of stores this sales person is expected to visit. The same method could be applied to every new counter/area the company would like to visit/begin operations so as to estimate the possible purchase by the retailers therein.