HR Metrics & Analytics
MSB21T2001

Index-Session 24

session Objectives- Topics to be discussed

Regression

Recap of Session 26

1. DATA VISUALIZATION OR HR REPORT VISUALIZATION
2. PERFORMING ROOT CAUSE ANALYSIS
3. DATAFICATION OF HUMAN RESOURCES

Before we come and discuss regression, let us look at a famous quote by Ronald Coase. He said, "If you torture the data long enough, it will confess." And, regression is one of the techniques that is frequently used to make the data confess. Let us look at a few interesting hypotheses, which people have come up with. The first one says, “Good looking couples are more likely to have girl child." Personally, I like this hypotheses because I have a daughter, and at least statistically, I'm good looking. The next is vegetarians miss fewer flights, women use camera phone more than men, lefthanded men earn more money, smokers are better salespeople, and those who whistle at workplace are efficient. Now, organizations use these hypotheses to add value. For example, let us look at women use camera phone more than men. If there is a company which makes cellphones; now they can target women and claim that those phones are great for taking photos. And, similarly, smokers are better salespeople. If a company is looking for salespeople, then in the interview they can ask whether they smoke. So, this is how some of these hypotheses are actually used.

Regression is one of the techniques and also one of the powerful techniques. But, they may have used simple hypotheses testing techniques such as Z-test, or T-test, or F-test, and so on. So, we don't know, really, how they actually prove these hypotheses. Now, let us look at what is regression. Regression is a tool for finding existence of an association relationship between a dependent variable (Y) and one or more independent variables in the study. The relationship can be linear or non-linear. Regression is a statistical relationship. So, we had to understand the difference between a mathematical relationship and a statistical relationship. Let us look at a mathematical relationship. We have $Y = \beta 0 + \beta 1X$. So, what do we mean by that? If we know the value of X, we can predict the value of Y exactly, whereas in a statistical relationship, we'll have the relationship as $Y = \beta 0 + \beta 1X + \epsilon$. So, here with the knowledge of X, we'll not be able to predict the value of Y exactly. There will be always some error.

Let us come back and see the nomenclature used in regression. We call a dependent variable or a response variable that measures outcome of a study, also called outcome variable. In the case of die-another-day case, the total cost of treatment is a response variable or outcome variable. And an independent variable or explanatory variables explains the changes in the response variable. So, if you want to understand why the total treatment cost changes, we may have to look at the variables like the patient's height, weight, and the past medical history. So, with that information, we believe that we may be able to tell the value of outcome variable. Regression often sets the values of explanatory variable to see how it affects the response variable. It is important to understand that regression model establishes existence of an association between two variables, but not causation. This is very, very important for the students to understand.

Now, there are techniques like counter factual models and graphical models that can be used for establishing causal relationship. But, we're not going to look at those techniques in this particular course. Here, I have given a table which basically gives us a list of different names used to represent Dependent Variable and Independent Variable. It is also important to understand that dependent variable doesn't mean that it depends on the values of independent variable. Just name that we use in a regression model. And also, regression is not designed to capture causality. The purpose of regression is to predict the value of dependent variable given the value(s) of independent variable(s).

Regression is one of the most frequently used tool in predictive analytics, and let us see why. Whenever we talk about companies, they use many different key performance indicators. So, key performance indicators such as market share, or if I have to use specific example, in case of dieanother-day hospital, they were interested in finding the total cost of treatment. So, you take any organization, there will be a list of key performance indicator. And I will like to know, how these KPIs are affected because of changes in some other factor. So, that's why regression is very important. Let us look at some specific examples under different functional areas of management. When we talk about finance, Capital Asset Pricing Model is an important model, which everybody uses, and at this nothing but a regression model non-performing assets, probability of default, chance of bankruptcy, and credit risk. Many times, people use regression to find these KPIs. When we talk about marketing, people like to know what will be the sales and what will be the market share, customer satisfaction, customer churn, customer retention, and customer life time value.

These are important measures in marketing and, again, researchers use regression to calculate or predict these values. In Operations, things like inventory requirement, productivity, and efficiency, again, this can be calculated or predicted using regression model. In HR or Human Resource Management, things like job satisfaction and attrition can be understood using regression model. For example, company may like to know which employees likely to churn. So, they can develop regression model to find out what is the risk associated with the employee that he or she may likely to churn in the near future.

You must have read the caselet die-another-day hospital. In that, the protagonist is interested in knowing what will be the treatment cost at the time of admission for a patient. Now, to find the treatment cost or to predict the treatment cost, they used factors like age, weight, past medical history, or even if there are blood reports, they can use the information from the blood report to predict the total treatment cost in that case. Now, there are large number of applications. So, let me give you few more generic examples how regression is used. For example, if we are talking about marketing in detail, questions like how to improve success probability of a new product, what is the impact of food label on purchase decision, which promotion is more effective. So, some of these questions can be answered using regression models. These are few examples from banking and Finance: What is a risk associated with a customer? Which customer is likely to default? What percentage of loans is likely to result in a loss? And, how to identify most profitable customer? There are large number of such questions. One can find answer to those questions using regression model. That's why regression is one of the most popular tools in predictive analytics.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.213802259							
R Square	0.045711406							
Adjusted R Square	0.037624214							
Standard Error	0.397565779							
Observations	120							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	0.893397727	0.893397727	5.652321465	0.019039702			
Residual	118	18.65090874	0.158058549					
Total	119	19.54430647						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	11.80404812	0.180542158	65.38111796	1.55176E-94	11.44652549	12.16157074	11.44652549	12.16157074
Body Weight	0.007477239	0.003145052	2.377461139	0.019039702	0.00124918	0.013705299	0.00124918	0.013705299

School of Business

Course Code : MSB21T2001

Course Name: HR Metrics & Analytics

X-No	Body Weight	Cost of Treatment	Ln(Treatment cost)	(X-Xbar)
1	32	1,48,612	11.9093228	576
2	32	1,28,104	11.7656006	576
3	33	1,33,087	11.79875833	529
4	35	1,24,804	11.73469979	441
5	36	1,99,790	12.20503209	400
6	36	1,47,021	11.89833071	400
7	36	1,35,216	11.81412878	400
8	37	1,66,709	12.02400606	361
9	37	1,23,188	11.72166211	361
10	40	2,33,266	12.35993471	256
11	41	1,80,713	12.10466542	225
12	41	1,77,874	12.08883071	225
13	41	1,84,723	11.84741717	225
14	41	1,35,612	11.81755315	225
15	42	1,19,686	11.69262392	196
16	43	2,32,676	12.3574223	169
17	43	2,01,219	12.21214915	169
18	43	1,97,865	12.19534026	169
19	43	1,59,348	11.6897887	169
20	44	1,89,827	12.8734833	144
21	44	2,03,271	12.6885238	144
22	45	3,72,357	12.8276085	121
23	45	3,34,955	12.7317147	121
24	45	2,94,616	12.59942775	121
25	45	1,56,374	11.9600585	121
26	45	1,59,989	11.9990224	121
27	46	1,82,651	12.11533251	100
28	47	3,62,231	12.8000741	81
29	47	3,43,984	12.7483042	81
30	47	1,75,576	12.07582728	81
31	47	1,31,837	11.78932159	81
32	47	1,02,138	11.5279872	81
33	48	1,64,719	12.0119627	64
34	49	2,10,622	12.2578204	49
35	49	1,29,475	11.7712431	49
36	49	1,29,475	11.7712431	49
37	50	2,75,888	12.5277027	36
38	50	1,80,870	12.1053382	36
39	50	1,32,997	11.79838185	36
40	51	4,95,969	13.1142681	25
41	51	3,41,109	12.7396716	25
42	51	1,09,452	11.6032138	25
43	52	2,09,886	12.2543381	16
44	53	3,61,738	12.79867547	9
45	53	2,67,473	12.4190476	9
46	54	2,60,869	12.4717364	4
47	54	2,48,031	12.4253902	4
48	54	2,20,519	12.3037914	4
49	54	1,54,354	11.94700394	4
50	55	1,78,037	12.0897667	1
51	55	1,62,271	11.9670266	1
52	55	1,44,134	11.8784987	1
53	55	1,34,498	11.8093201	1
54	56	4,20,040	12.9481022	0
55	56	2,14,679	12.2768917	0
56	56	1,99,268	12.2024093	0
57	56	1,91,543	12.1732499	0
58	56	1,76,341	12.0801748	0
59	56	1,63,483	12.0046429	0
60	56	1,38,923	11.8416751	0
61	57	2,19,126	12.2974328	1
62	57	2,00,321	12.2076436	1
63	57	1,62,364	11.9975965	1
64	58	4,44,876	13.0055087	4
65	58	3,41,011	12.73967001	4
66	58	1,91,188	12.1620252	4
67	58	1,63,279	11.8715478?	4
68	59	2,62,582	12.47931869	9
69	59	2,60,036	12.4685736	9
70	59	2,14,716	12.2770715	9
71	59	1,64,962	12.0134704?	9
72	59	1,33,009	11.7981720?	9
73	60	3,49,318	12.7627796	16
74	60	3,45,590	12.7530838	16
75	60	2,76,458	12.5298149	16
76	60	2,29,289	12.3427385	16
77	60	2,04,852	12.230648	16
78	60	1,38,093	11.8356828	16
79	60	1,27,899	11.7589617	16
80	61	3,95,163	12.8870536?	25
81	62	3,05,193	12.6286964	36
82	62	2,41,130	12.3930549	36
83	62	1,59,882	11.9821932	36
84	63	1,51,931	11.9311857?	49
85	63	1,09,086	11.5998503?	49
86	63	1,09,086	11.5998503?	49
87	64	5,14,524	13.1529674?	64
88	64	2,51,471	12.4430947	64
89	64	1,86,450	12.1359183?	64
90	64	1,62,957	12.0052436?	64
91	64	1,39,067	11.8427111	64
92	64	1,38,535	11.8388782	64
93	65	5,01,897	13.1261502	81
94	65	4,59,429	12.9442625	81
95	65	2,53,368	12.4425982	81
96	65	2,02,807	12.2203107	81
97	65	1,78,100	12.0955004?	81
98	67	4,49,395	13.0116275?	121
99	67	1,99,677	12.2045634	121
100	68	3,91,276	12.7806323	144
101	68	1,46,309	11.8934732?	144
102	69	2,95,155	12.5952593?	169
103	70	1,57,763	11.9688491?	196
104	70	1,41,402	11.8786643?	196
105	71	2,48,112	12.4263554	225
106	71	2,12,287	12.2650441	225
107	72	4,57,329	12.9888663	256
108	72	1,44,037	11.8778270?	256
109	73	3,48,687	12.7639295	289
110	73	2,78,314	12.5362449?	289
111	73	1,83,204	12.1183557	289
112	74	2,88,960	12.5749355	324
113	75	2,01,998	12.2352174	361
114	76	1,63,765	12.0061877?	400
115	77	3,64,222	12.8055188?	441
116	77	1,31,130	11.7980813?	441
117	78	2,93,127	12.5893623?	484
118	78	1,88,824	12.1483764?	484
119	78	1,49,462	11.9147674?	484
120	78	1,35,131	11.8413809	484
X-Bar=	56		SSX=	19896

Name of the Faculty: Mamta Gaur

Program Name: MBA

REFERENCES

- Bhattacharya ,Deepak (2019). HR Metrics and Analytics, Pearson
- Batt, R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. Academy of Management Journal, 45(3), 587–597.
- Becker, B. E., & Gerhart, B. (1996). Human resources and organizational performance: Progress and prospects. Academy of Management Journal [Special Issue: Human Resources and Organizational Performance], 39(4): 779–801.
- Becker, Gary, S. (1962). Investment in human capital: A theoretical analysis. Journal of Political Economy, 70(2), 437–448.
- (1964). Human capital. New York: National Bureau for Economic Research.
- Beer, M., Spector, R., Lawrence, P., Quinn Mills, D., & Walton, R. (1984). Human resource management: A general managers perspective. Glencoe, IL: Free Press.
- Begin, J.P. (1993). Identifying patterns in HRM systems: Lessons learned from organizational theory. In J. Shaw, P. Kirkbride and K. Rowlands (Eds.), Research in Personnel and Human Resource Management, 3 (pp. 3–20), Greenwich, CT: JAI Press.
- Bhattacharyya, D. K. (2012). Human resource management (3rd ed.). New Delhi: Excel Books

In case, you find any difficulty in understanding the concepts of lecture, please feel free to contact.

Dr Mamta Gaur
9891121396
mamtagaur@galgotiasuniversity.edu.in

Thanks