



INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact factor: 6.078

(Volume 6, Issue 2)

Available online at: www.ijariit.com

Impact of a drug on cancer

Reet Bhansali

reetbhansali27@gmail.com

NMIMS Anil Surendra Modi School of Commerce, Mumbai, Maharashtra

ABSTRACT

We investigate whether there is a direct co-relation between impact of a drug on cancer occurrence by taking into consideration various samples of drug. We examine the process by taking into consideration two independent factors related to a drug which are amount of dosage (in grams) and No. of Times (Dosage Per Week). We perform regression, distribution and correlation analysis on all the 3 variables viz. impact, amount of dosage and no. of times (dosage in a week). We have done the analysis with the help of 530 samples of drugs that helped us determine our objectives.

Keywords— Impact of drug on cancer, Dosage and no. of times of dosage

1. OBJECTIVES

- To determine the correlation between the two variables.
- To find the regression analysis.
- To summarize the data in the form of statistics.
- To interpret the results received.
- To interpret the data with the help of graph.

2. INTRODUCTION

There has been a growing impact of various medicines (drugs) on creation of cancer cells in the body. Cancer cells are the cells that the human body has lost control of, and since they are transformed body cells, it makes it more difficult to combat them effectively. Usually, what harms cancer cells, also harms the healthy cells in the body. Now, a new study investigates the effects of drugs on cancer cells and how to treat them. The following research was carried out by University of Zurich, Switzerland.

For this project we got the opportunity to apply and understand the concepts and theories taught to us in class regarding statistics in an effective manner. I made the use of secondary data available on the internet in order to apply the concepts. The liberty to choose as many variables as we wanted was given to us by our faculty Ms. Tejaswini Angre and I went ahead with three variables namely the Impact(in %)-Dependent, Dosage(in grams)-Independent, Number of Times(Dosage Per week)-Independent. I have been successful in doing an interpretation of the relationship between the above-mentioned variables to the best of my abilities with the help of MS-Excel. Quantitative tools and techniques such as Graphs, Correlation, Regression and Descriptive Statistics is used to analyze the inter-relation between the variables.

3. DATA

I have taken help of a secondary data.

<https://www.cancer.gov/about-cancer/treatment/drugs>

NOTE- I have attached a glimpse of my data as a screenshot and the source has been mentioned above.

SR NO.	Impact(in %) (Dependent)	Dosage(in Grams) (Independent)	No. of Times(Dosage Per Week) (Independent)
1	20	366	16
2	25	326	14
3	8	153	9
4	13	339	12
5	5	144	10
6	19	327	11
7	5	284	8

8	22	352	16
9	14	350	20
10	12	239	1
11	16	389	14
12	15	298	13
13	11	224	3
14	24	349	13
15	13	273	21
16	5	146	6
17	17	440	18
18	1	206	1
19	10	249	12
20	4	104	3
21	18	423	14
22	7	240	4
23	17	409	11

4. HISTOGRAM

4.1 Histogram for IMPACT

Class Intervals	Bin Values	Bin Values	Frequency
0-5	4	4	105
5-10	9	9	97
10-15	14	14	105
15-20	19	19	104
20-25	24	24	102
25-30	29	29	17
		More	0

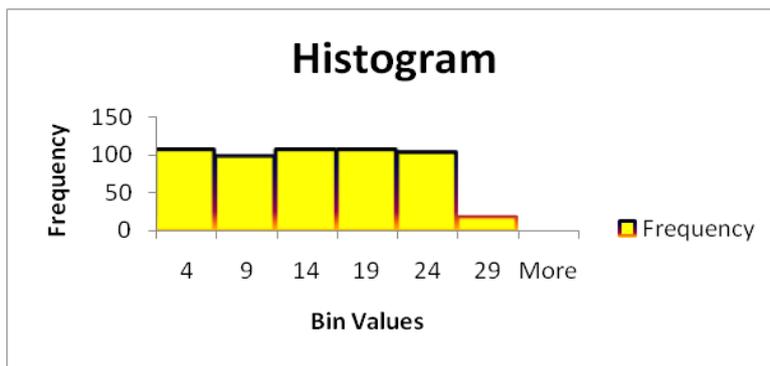


Fig. 1: Histogram

Interpretation for Impact (in %): The given histogram shows that the maximum frequency of impact lies in the 0-5 and 10-15 region. Out of all the given data, 105 of them lie in each region. However, the bars of the Histogram are almost even with each other showing that the range of impact is somewhat similar.

4.2 Pie Chart for DOSAGE

Class Intervals	Bin Values	Bin Values	Frequency
100-200	199	199	141
200-300	299	299	182
300-400	399	399	125
400-500	499	499	82
		More	0

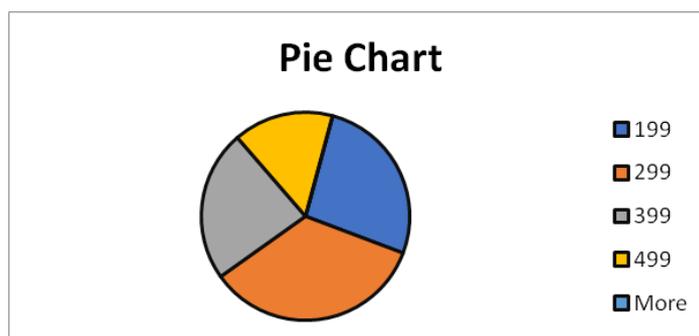


Fig. 2: Pie Chart

Interpretation of Dosage: The given pie chart shows that the maximum frequency of dosage lies in the 200-300 region. Out of all the data, 182 of them lie in this region.

4.3 Histogram for no. of times

Class Intervals	Bin Values	Bin Values	Frequency
0-5	4	4	81
5-10	9	9	117
10-15	14	14	176
15-20	19	19	106
20-25	24	24	50
		More	0

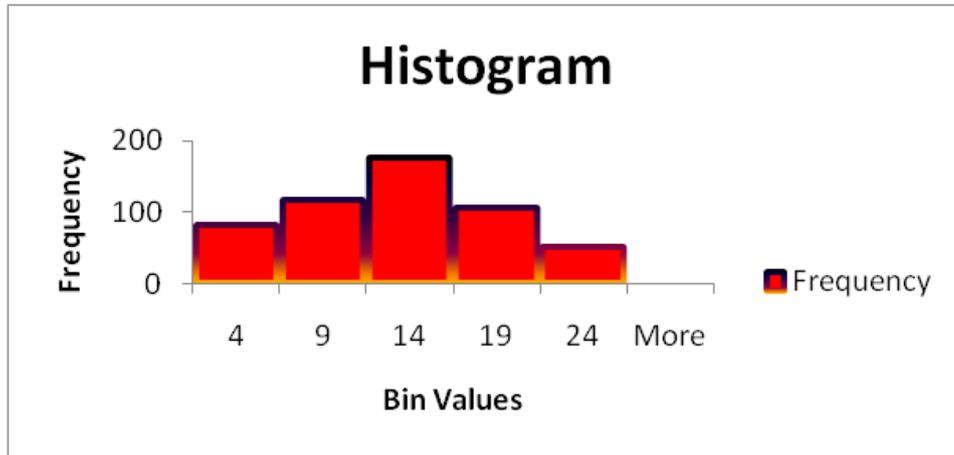


Fig. 3: Histogram

Interpretation of No. Of times (Dosage Per Week): The given histogram shows that the maximum frequency of No. Of times lies in the 10-15 region. Out of all the given data, 176 of them lie in this region.

5. DESCRIPTIVE STATISTICS

<i>Impact(in %)</i>	
Mean	12.34717
Standard Error	0.326675
Median	12
Mode	2
Standard Deviation	7.520623
Sample Variance	56.55977
Kurtosis	-1.20511
Skewness	-0.01677
Range	25
Minimum	0
Maximum	25
Sum	6544
Count	530

Interpretation:

- The average % of impact is 12.34
- The central value of the Impact Data is 12
- The maximum number of impacts is 2%.
- On an average, the deviation of the average impact from the other impact is 7.52
- It is Platykurtic (since it is less than 0)
- It is negatively skewed.
- The difference between the maximum impact and minimum impact is 25.
- The minimum impact is 0.
- The maximum impact is 25.
- The sum of impact by all drugs is 6544.
- A total of 530 data records were collected.

<u>Dosage(in Grams)</u>	
Mean	276.3377
Standard Error	4.377904
Median	276
Mode	250
Standard Deviation	100.7869
Sample Variance	10158
Kurtosis	-1.0934
Skewness	-0.02956
Range	349
Minimum	100
Maximum	449
Sum	146459
Count	530

Interpretation:

- The average grams of dosage are 276.3377
- The central value of the dosage Data is 276
- The maximum number of dosages are 250.
- On an average, the deviation of the average dosage from the other dosage is 100.7869
- It is Platykurtic (since it is less than 0)
- It is negatively skewed.
- The difference between the maximum impact and minimum dosage is 349.
- The minimum dosage is 100.
- The maximum dosage is 449.
- The sum of all dosage of all drugs is 146459.
- A total of 530 data records were collected.

<u>No. of Times(Per Week)</u>	
Mean	11.15094
Standard Error	0.247612
Median	11
Mode	12
Standard Deviation	5.70045
Sample Variance	32.49513
Kurtosis	-0.95422
Skewness	-0.03542
Range	20
Minimum	1
Maximum	21
Sum	5910
Count	530

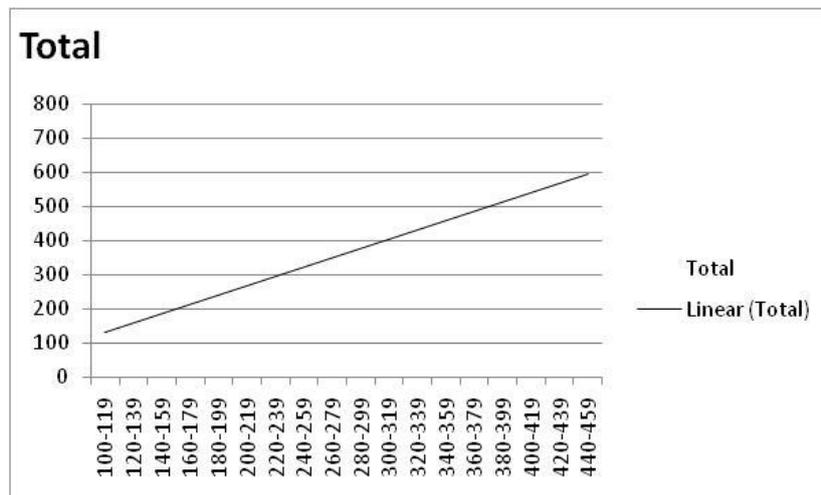
Interpretation:

- The average times of dosage is 11.15
- The central value of the Data is no. Of times is 11.
- The maximum number of times are 12.
- On an average, the deviation of the average no. Of times from the other no. Of times is 32.49
- It is Platykurtic (since it is less than 0)
- It is negatively skewed.
- The difference between the maximum impact and minimum no. Of times is 20.
- The minimum no. Of times is 1.
- The maximum no. Of times is 21.
- The sum of all no. Of times of all dosage is 5910.
- A total of 530 data records were collected.

6. CORRELATION

	Impact (in %)	Dosage (in Grams)	No. of Times (Per Week)
Impact (in %)	1		
Dosage (in Grams)	0.708515992	1	
No. of Times (Per Week)	0.667728264	0.654555528	1
Interpretation of impact and dosage: There exists positive but imperfect correlation of 0.7 between impact and dosage. Interpretation of impact and No. Of times: There exists positive but imperfect correlation of 0.66 between impact and no. of times.			

REGRESSION EQUATION



SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.75814359
R Square	0.574781702
Adjusted R Square	0.573167971
Standard Error	4.913402422
Observations	530

ANOVA					
	df	SS	MS	F	Significance F
Regression	2	17197.53794	8598.768972	356.1817059	1.37799E-98
Residual	527	12722.58281	24.14152336		
Total	529	29920.12075			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	2.69587644	0.625060748	-4.312983134	1.92331E-05	-3.923793017	-1.467959912	-3.923793017	-1.467959912
Dosage (in Grams)	0.035438971	0.002803629	12.64039392	3.55249E-32	0.029931311	0.040946631	0.029931311	0.040946631
No. of Times (Per Week)	0.470805121	0.049569616	9.49785696	7.44551E-20	0.373426823	0.568183419	0.373426823	0.568183419

$$Y = a + bX_1 + bX_2$$

$$Y = -2.695 + 0.035X_1 + 0.407X_2$$

Interpretation: If X1 increases by 1, Y increases by 0.035 and if X2 increases by 1, Y increases by 0.407.

7. CONCLUSION

There has been a growing impact of various medicines (drugs) on creation of cancer cells in the body. Cancer cells are the cells that the human body has lost control of, and since they are transformed body cells, it makes it more difficult to combat them effectively. Usually, what harms cancer cells, also harms the healthy cells in the body. Quantitative tools and techniques such as Graphs, Correlation, Regression and Descriptive Statistics is used to analyze the inter-relation between the variables.