Available online at http://www.journalijdr.com

ISSN: 2230-9926

IJDR

International Journal of Development Research
Vol. 09, Issue, 08, pp. 29448-29456, August, 2019

RESEARCH ARTICLE OPEN ACCESS

USING REGRESSION ANALYSIS METHODS IN BIOSTATISTICS AN APPLIED STUDY ON A SAMPLE OF DIABETICS' PATIENTS

***[1]Şakir Işleyen, [2]Amar Yahya Zebari and [2]Rogash Younis Masiha**

[1]Department of Econometry, Van Yüzüncü Yıl Üniversitesi, Van, Turkey
[2]Department of Statistics, Van Yüzüncü Yıl Üniversitesi, Van, Turkey

ARTICLE INFO

Article History:
Received 13th May, 2019
Received in revised form
10th June, 2019
Accepted 09th July, 2019
Published online 30th August, 2019

Key Words:
Biostatistics, Statistical Analysis, Diabetes, Regression Analysis, Mathematical Modeling, Statistical Software Packages for Social Sciences (SPSS).

ABSTRACT

Biostatistics is one of the significant methodologies for researchers in the Medical Sciences for scientific explanations and expectations. The picking of a theme of diabetes to be connected in this research because of the significance of finding a cure of this desease, which is increasing in the most recent years. The reason of this increase and the sorts are explored by the analysts, to represent how a few factors such as weight and age can impact the diabetes. The research was conduct on a sample of 1385 patients with diabetes, has been chose from the network information diabetics in the Diabetics Center of Duhok/ Iraq, and applies the regression analysis methods on this data to create a mathematical equation helps us to anticipate future injury rates. The utilization of using the Statistical Software Packages for Social Sciences (SPSS) in this research was to acquire precise outcomes and decrease the time and voltage. An exponential model fitted the current Duhokinformation.

Citation: Şakir Işleyen, Amar Yahya Zebari and Rogash Younis Masiha. 2019. "Using regression analysis methods in biostatistics an applied study on a sample of diabetics' patients", *International Journal of Development Research*, 09, (08), 29448-29456.

INTRODUCTION

Regression modeling refers to a mathematical clarification of a procedure as far as a lot of related factors. The estimation of a needy variable depends on the degree of numerous free factors. For Example; the yield of a specific creation procedure might rely on the weight, throughput, temperature. The eco-friendliness vehicle might rely on the heaviness of the vehicle, particular motor and body. The available showed product may rely on the value that the client expects to pay. In these circumstances we are interesting to get a "model" for the connection between the dependent variable (regularly symbolized as y), and the independent variables (symbolized as x). Regression analysis deals with the modeling functional relationship between the response variable (dependent variable y) with one or more of the explanatory variables (independent variable x).

Statistical Analysis for Duhok Diabetics Data: In this chapter we will calculate some important descriptive statistics[i] for (1385 diabetes patients) that we choose them randomly from Diabetes Center/ Duhok/ Kurdistan Region/ Republic of Iraq, using the Statistical Package for the Social Science (SPSS).

Descriptive: (weight, length and age): In the Descriptive Statistics table[ii], we found (valid number, range, minimum values, maximum values, mean, standard deviation and variance) for every variable in our study, we can see in the N column that we have a few missed values for every variable, the total number of missed values in our sample is equal to 38. If we draw a curve in the above histogram we can note that the data does not follow the normal distribution, and we will see that more clearly later in this study by the test of normality. Histogram graph shows that the highest levels of the frequencies are between 1954-1956 years of birth. In the pie graph, using an (age binned) variables, it is clear that there is a great affinity between three categories (1948-1954, 1955-1961 and 1962-1968), or we can say that in general the peoplewho aged between (45-65) years oldin 2013. Next, the red space of the pie figure, that is showing the age group (1969-1976). Thus, we can say that the group in which the fewest number of patients that is painted by light brown color which is

***Corresponding author:* Şakir Işleyen,**
Department of Econometry, Van Yüzüncü Yıl Üniversitesi, Van, Turkey

Table 1. Descriptive statistics table for age, weight and length

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Age	1378	84	1927	2011	1962.04	12.812	164.156
Weight	1365	150	11	161	78.17	16.931	286.654
Length	1359	179	15	194	157.49	11.005	121.115
Valid N (listwise)	1347						

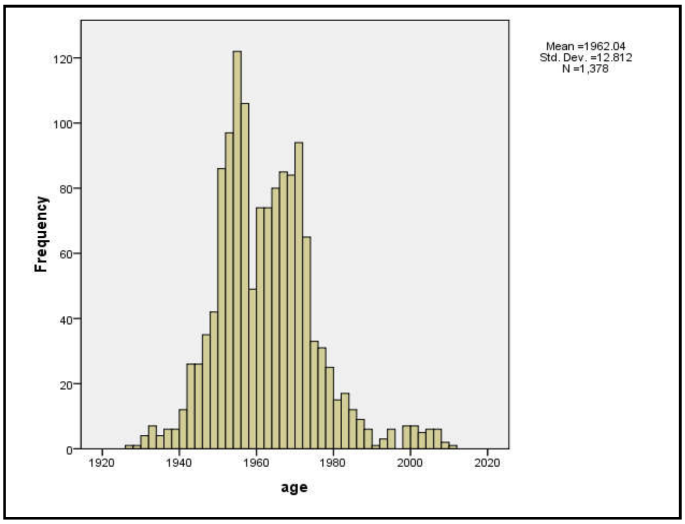

Figure 1. Histogram for Age

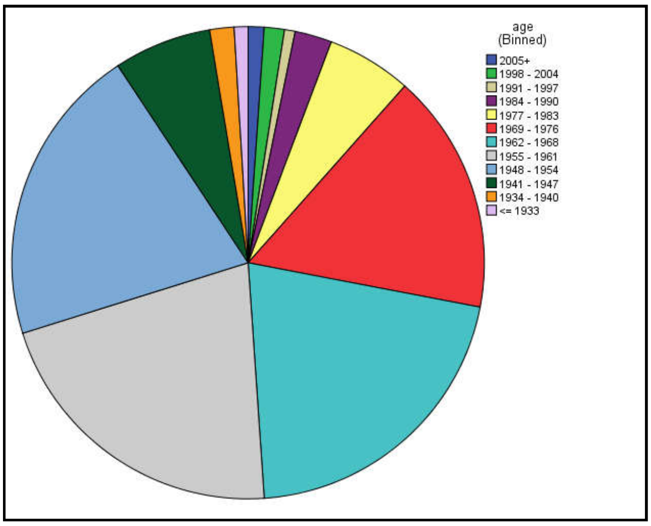

Figure 2. Pie Graph for Age

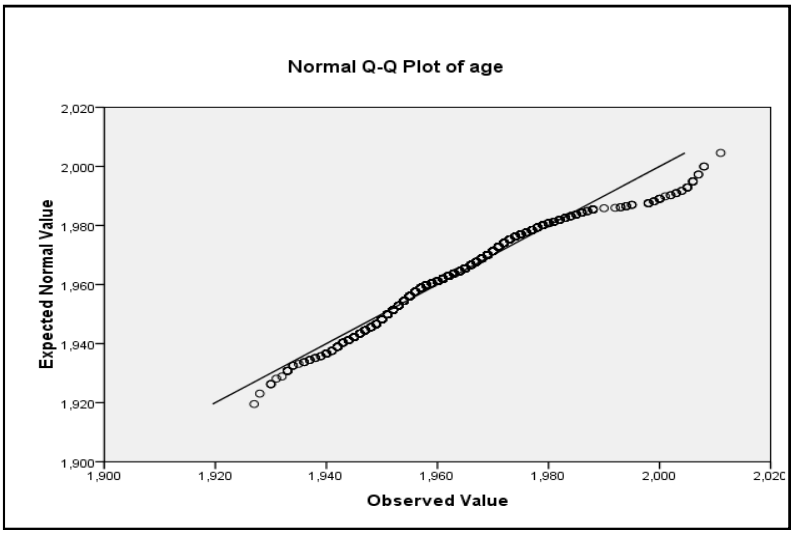

Figure 3. Normal Q-Q Plot for Age

representative the group (1991-1997), or in another words, those people whom ages between 16-22 years oldin 2013. Overview, we can say that the number of diabetic patients in a significant decline in the place under study.

The Quantile - quantile or Q-Q plot[iii]: The Q-Q plot (that called quantile - quantile) is a graphic tool used to show us if we assume the right distribution for our data. In general, this graphic tool works by computing the expected value for every data on the distribution. If the data follow the distribution, then the points should approximately fall on a straight line in Q-Q plot.

Then, in the bellow Q-Q plot that represent our sample, we can see many points does not fall approximately on a straight line, and this means that the data does not follow our assumed distribution –The Normal Distribution.

Normal Distribution Test[iv]

Now, we will test if our data follows the normal distribution, with confidence =95%.

H_0: the data follows the normal distribution.

Ha: the data does not follow the normal distribution.

In our example we assumed the confidence interval is 95%. Then click ok.

We can see that in table 3, the age mean is 1962, and we are sure with confidence 95% that the sample mean lies between 1961.36 to 1962.72 and the median is 1961 in this sample. The birth year for the first patient in this sample is 1927, and the last patients' birth year is 2011. The range is equal to 84.

Table 2. Case processing summary table

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Age	1378	99.5%	7	.5%	1385	100.0%

Table 3. Descriptive statistics table

			Statistic	Std. Error
Age	Mean		1962.04	.345
	95% Confidence Interval for Mean	Lower Bound	1961.36	
		Upper Bound	1962.72	
	5% Trimmed Mean		1961.49	
	Median		1961.00	
	Variance		164.156	
	Std. Deviation		12.812	
	Minimum		1927	
	Maximum		2011	
	Range		84	
	Interquartile Range		17	
	Skewness		.721	.066
	Kurtosis		1.359	.132

Table 4. Tests of Normality table

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
Age	.075	1378	.000	.965	1378	.000

a. Lilliefors Significance Correction

Table 5. Diabetes type statistics table

N	Valid	1385
	Missing	0

Table 6. Diabetes type frequency table

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		24	1.7	1.7	1.7
	T1	72	5.2	5.2	6.9
	T2	1289	93.1	93.1	100.0
	Total	1385	100.0	100.0	

Table 7. Age group information

N	Valid	1378
	Missing	7
Mean		7.42
Median		8.00
Mode		8
Range		11
Minimum		1
Maximum		12

Table 8. Age group statistics

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2005+	15	1.1	1.1	1.1
	1998 – 2004	19	1.4	1.4	2.5
	1991 – 1997	10	.7	.7	3.2
	1984 – 1990	35	2.5	2.5	5.7
	1977 – 1983	80	5.8	5.8	11.5
	1969 – 1976	227	16.4	16.5	28.0
	1962 – 1968	288	20.8	20.9	48.9
	1955 – 1961	293	21.2	21.3	70.2
	1948 – 1954	283	20.4	20.5	90.7
	1941 – 1947	92	6.6	6.7	97.4
	1934 – 1940	23	1.7	1.7	99.1
	<= 1933	13	.9	.9	100.0
	Total	1378	99.5	100.0	
Missing	System	7	.5		
Total		1385	100.0		

It is clear that from both Kolmogrov-Smirnov and Shapiro-wilk tests that significance value=0 < confidence interval $\propto$ = 0.05. So, we will reject null hypothesis H_0 and this means that there is a significant difference and this reassurance our note above in Q-Q plot.

Frequencies: In the diabetes type we have sample size = 1385, and we have 24 missed values which represent 1.7% of the total sample number, then the total number of valid values is 1361. There are 72 persons suffer from Type1 diabetes and this is representing 5.2%, and 1289 person suffer from Type2 diabetes and this is representing 93.1%. It is clearly that Type2 of diabetes is the most widespread.

Age Groups[v]

$$\text{Number of cut point} = 1 + 3.322\log(1378) \approx 12.$$

$$\text{width} = \frac{range}{k} = \frac{2011-1927}{12} = 7.$$

In the age (Binned) table, we found that the most ages affected with diabetes are bounded in the following groups: 1948-1954, 1955-1961, 1962-1968 and 1969-1976, with 283, 293, 288, 227 respectively. There is a small increase in the group 1955-1961. This means that those people whose aged between 65 to 37 years old in 2013. They are 1091 patients and they are representing 79.2% of our sample, contained in the mentioned groups. It is clear that the number of females with diabetes is greater than the number of males in our random sample, there are 868 females with diabetes and this represent 62.7 % from all the sample, and 517 males with diabetes and they are representing 37.3% in the same sample.

Correlations: In this part, we try to find the correlation coefficient between our variables, to illustrate what is the importance of each variable in this study[vi].

Gender Correlation: For both of them (male and female), we will try to find the correlation coefficient each separately[vii].

Gender = Female: Now, the correlation coefficient for female with age, weight, length and date of diabetes variables will be shown in Table 18. First, we want to mention that the total number of females with diabetes in our sample is 868. We note that the age variable has a very high positive linear relationship with female equal to +1. We can see also that we have n= 863 females, this means that there are 5 missed values.

Table 9. Gender Frequency Table

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	F	868	62.7	62.7	62.7
	M	517	37.3	37.3	100.0
	Total	1385	100.0	100.0	

Table 10. Correlations table for females with age, weight, length and date of diabetes

		Age	Weight	Length	Dateofdiabetes
Age	Pearson Correlation	1	-.007-	-.076-*	.278**
	Sig. (2-tailed)		.834	.027	.000
	N	863	850	845	811
Weight	Pearson Correlation	-.007-	1	.304**	.156**
	Sig. (2-tailed)	.834		.000	.000
	N	850	855	847	805
Length	Pearson Correlation	-.076-*	.304**	1	-.028-
	Sig. (2-tailed)	.027	.000		.431
	N	845	847	850	800
Dateofdiabetes	Pearson Correlation	.278**	.156**	-.028-	1
	Sig. (2-tailed)	.000	.000	.431	
	N	811	805	800	816

*. Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level [2-tailed]
a. gender = 1

Table 11. Correlation table for male patients' age, weight, length, and date of diabetes

		Age	weight	length	Dateofdiabetes
Age	Pearson Correlation	1	-.323-**	-.309-**	.212**
	Sig. (2-tailed)		.000	.000	.000
	N	515	508	507	482
Weight	Pearson Correlation	-.323-**	1	.618**	.025
	Sig. (2-tailed)	.000		.000	.589
	N	508	510	507	477
Length	Pearson Correlation	-.309-**	.618**	1	-.055-
	Sig. (2-tailed)	.000	.000		.234
	N	507	507	509	476
Dateofdiabetes	Pearson Correlation	.212**	.025	-.055-	1
	Sig. (2-tailed)	.000	.589	.234	
	N	482	477	476	484

** Correlation is significant at the 0.01 level [2-tailed] a. gender = 2

Table 11. Correlation table for male patients' age, weight, length, and date of diabetes

		Age	weight	length	Dateofdiabetes
Age	Pearson Correlation	1	-.323-**	-.309-**	.212**
	Sig. (2-tailed)		.000	.000	.000
	N	515	508	507	482
Weight	Pearson Correlation	-.323-**	1	.618**	.025
	Sig. (2-tailed)	.000		.000	.589
	N	508	510	507	477
Length	Pearson Correlation	-.309-**	.618**	1	-.055-
	Sig. (2-tailed)	.000	.000		.234
	N	507	507	509	476
Dateofdiabetes	Pearson Correlation	.212**	.025	-.055-	1
	Sig. (2-tailed)	.000	.589	.234	
	N	482	477	476	484

** Correlation is significant at the 0.01 level [2-tailed] a. gender = 2

Table 12. Correlation table for diabetes type with age, weight, length and date of diabetes

		Age	weight	length	Dateofdiabetes
Age	Pearson Correlation	1	-.803-**	-.633-**	.521**
	Sig. (2-tailed)		.000	.000	.000
	N	72	71	69	68
Weight	Pearson Correlation	-.803-**	1	.739**	-.404-**
	Sig. (2-tailed)	.000		.000	.001
	N	71	71	68	67
Length	Pearson Correlation	-.633-**	.739**	1	-.234-
	Sig. (2-tailed)	.000	.000		.061
	N	69	68	69	65
Dateofdiabetes	Pearson Correlation	.521**	-.404-**	-.234-	1
	Sig. (2-tailed)	.000	.001	.061	
	N	68	67	65	68

** Correlation is significant at the 0.01 level [2-tailed]
a. diabetestype = 1

Table 13. Correlation table for diabetes type2 with age, weight, length and date of diabetes

		age	weight	length	Dateofdiabetes
Age	Pearson Correlation	1	.215**	.149**	.255**
	Sig. (2-tailed)		.000	.000	.000
	N	1282	1263	1259	1209
Weight	Pearson Correlation	.215**	1	.281**	.182**
	Sig. (2-tailed)	.000		.000	.000
	N	1263	1270	1262	1199
Length	Pearson Correlation	.149**	.281**	1	.042
	Sig. (2-tailed)	.000	.000		.144
	N	1259	1262	1266	1195
Dateofdiabetes	Pearson Correlation	.255**	.182**	.042	1
	Sig. (2-tailed)	.000	.000	.144	
	N	1209	1199	1195	1216

** Correlation is significant at the 0.01 level [2-tailed]
a. diabetestype = 2

The correlation coefficient between females and their weights is -0.007, this means that almost there is no linear relationship between them. We have n= 850, this means that there are 18 missed values. The correlation coefficient between female and their lengths equal to -0.076 and it is means that almost there is no relationship. For the linear relationship between female and date of diabetes acquire, the correlation coefficient is 0.278 and it is a weak positive relationship.

Correlation Gender = Male: In this part, the correlation coefficient of male patients' age, weights, lengths and date of diabetes variables will be analysed in this part. From Table 11, that shows us the correlation coefficient between male and their ages, weight, length and date of diabetes. At first we should mention that the total number of males with diabetes in our sample is 517. The correlation coefficient between male and their ages is equal to +1, and this means that there is strong linear relationship between males with diabetes and their ages. We have n= 515 patients, and this means that there are just two missed values. The correlation coefficient between males with diabetes and their weights is -0.323, and it is weak negative linear relationship, n=508, so we have 9 missed values. Also, between males with diabetes and their lengths, the correlation coefficient is -0.309, and it is weak negative linear relationship, and we have 10 missed values. The correlation coefficient between male with diabetes and their acquire diabetes is 0.212, and it is also weak positive linear relationship, n=482, so we have 35 missed values.

Correlation Coefficient for Diabetes Type: We try to understand the relationship between both types of diabetes (diabetes type1, diabetes type2) with their ages, weights, lengths and date of diabetes acquire.

Diabetestype = Type1: First, we will find the correlation coefficient for patients with diabetes typ1 and their ages,

Table 14. Paired samples statistics table

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Dfemale	4.25	20	2.900	.648
	Dmale	5.10	20	2.269	.507

Table 15. Paired samples correlations table

		N	Correlation	Sig.
Pair 1	Dfemale&Dmale	20	.716	.000

Table 10. Paired samples test table

	Paired Differences					T	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% CI Lower	Upper			
Dfemale – Dmale	-.850-	2.033	.455	-1.802-	.102	-1.870-	19	.077

Table 16. Model Summary and parameter estimates table

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Linear	.469	18.581	1	21	.000	-48.518-	8.355
Logarithmic	.270	7.782	1	21	.011	-65.016-	52.035
Inverse	.097	2.256	1	21	.148	71.393	-121.051-
Exponential	.783	75.687	1	21	.000	1.597	.206

Table 17. Linear Model Summary and Parameter Estimates table

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Linear	.469	18.581	1	21	.000	-48.518-	8.355

weights, lengths and date of diabetes acquire. We also want to mention that the total number of diabetics with type1 is 72 and with type2 is 1289 patients in our sample. Patients with diabetes type1 have a very strong positive relationship with their ages, that is the correlation coefficient equal to +1, we have n=72 and this means that there are no missed values. The correlation coefficient between patients with diabetes and their weights is -0.803 it is strong negatively linear relationship; we have only one missed value. Also, the correlation coefficient between patients with diabetes type1 and their lengths is -0.633, this is a moderate negative linear relationship, and we have 3 missed values.There is a moderate positive relationship equal to 0.521 between patients with diabetes and their date of diabetes acquires, we have 4 missed values.

Diabetes Type = Type 2: In this step, we will try to illustrate the relationship between patients with diabetes type2 and their ages, weights, lengths and date of diabetes. We should mention that the total number of patients with diabetes type2 in our sample is 1289. We see that there is a strong positive linear relationship equal to +1 between patients with diabetes type2 and their ages, we have 7 missed values. The correlation coefficient between patients with type2 diabetes equal to 0.215 and it is weak positive linear relationship, and there are 26 missed values. For patients with type2 diabetes and their lengths, there is also positive weak relationship equal to 0.149, and there are 30 missed values. The correlation coefficient between patients with diabetes type2 and their date of diabetes acquire is 0.255, it is weak positive linear relationship and we have 80 missed values.

T-Test: In this part, we will apply T-test on our sample, and we used T-test because we do not know the standard deviation of the population. Paired Sample Statistics for Duhokmalesand Duhokfemales. Test hypothesis mean to find that if Duhok male mean equal or not to Duhok female mean.

$$H_0: \mu_{DF} = \mu_{DM}$$
$$H_0: \mu_{DF} \neq \mu_{DM}$$

Where;

μ_{DF}: The mean of Duhok female
μ_{DM}: The mean of Duhok male.

There is a high positive correlation between Duhok male and Duhok female equal to 0.716 and significant value approximately equal to 0.000 < 0.05[viii]. We cannot reject H_0, this means that we accept that mean of D_f and mean of D_m are equal with 95% confidence.

Curve Fit for Linear, Logarithmic, Inverse and Exponential Equations[ix]: In this part, we will try to find the fit mathematical model regression of diabetes incidence in Duhok/ Kurdistan region of Iraq, using the Statistical Package for the Social Science (SPSS)[x]. We assume that the dependent variable is the number of patients and the independent variable is the date of diabetes. The independent variable is yearrank. From R square column[xi], we see that the first three models (Linear, Logarithmic and Inverse) have a normal positive correlation approximately equal to 0.41, but the exponential model has a higher positive correlation approximately equal to 0.83.

Table 18. Logarithmic model summary and parameters estimates table

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Logarithmic	.270	7.782	1	21	.011	-65.016-	52.035

The independent variable is year rank.

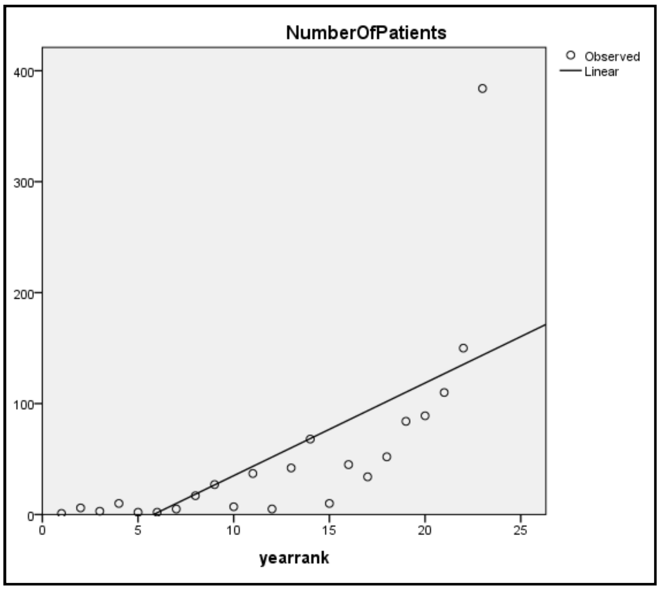

Figure 4. Linear Equation Plots

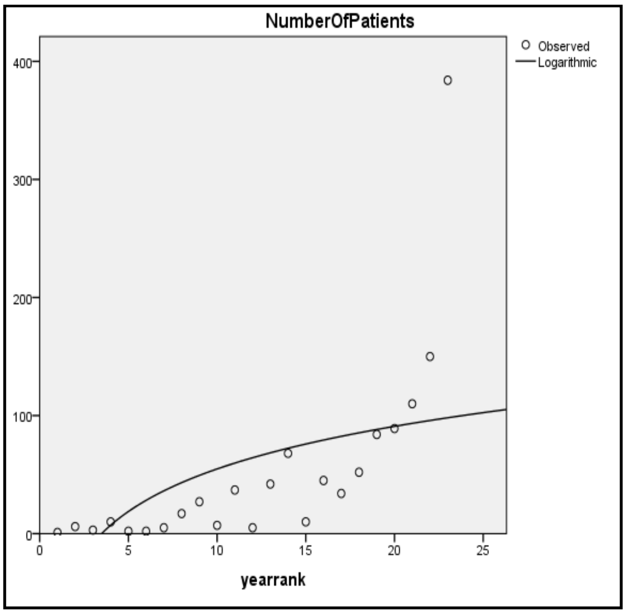

Figure 5. Logarithmic Equation Plots

Table 19. Inverse model summary and parameters estimates table

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Inverse	.097	2.256	1	21	.148	71.393	-121.051-

The independent variable is year rank.

Table 20. Inverse model summary and parameters estimates table

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Exponential	.783	75.687	1	21	.000	1.597	.206

The independent variable is yearrank. Dependent Variable: Number Of Patients.

Curve Fit for Linear Equation[xii]: There is one dependent variable that is the number of patients and one independent variable that is the date of diabetes[xiii]. In the table 17, we note that there is a positive correlation coefficient equal to 0.409. Significant value approximately to 0.000 and it is less than 0.05 this means that there is a significant linear correlation. Our linear equation model is:

$$\hat{Y} = -48.518 + 8.355(x)$$

Curve Fit Logarithmic[xiv]: Also, in the Logarithmic curve fit we have positive correlation coefficient equal to 0.408. Also, we have significant value approximately equal to 0.011 and it is less than 0.05, this means that there is a significant logarithmic correlation. Our logarithmic equation model is:

$$\hat{Y} = -65.016 + 52.035\, Log(x)$$

Curve Fit for Inverse Equation[xv]

In table 43 correlation coefficient are positive and equal to 0.406. Significant value is approximately equal to 0.148 and more than 0.05 this means that there is no significant Inverse regression. Our Inverse equation model is:

$$\hat{Y} = 71.393 + \frac{-121.051}{x}$$

Curve Fit for Exponential Equation: We note that from table 43 that there is a high positive correlation coefficient equal to 0.826 and it is the highest one among other three equations models. Significant value is approximately equal to 0.000 and it is less than 0.05 this means that there is a significant exponential regression. The exponential equation model is the best equation among other three equations (linear, logarithmic and Inverse).

Where;

Date: means the real date of diabetes
Year: refers to the rank for every year
Fit-1: the expected number for patients for every year
ERR-1: the expected error between expected number of patients and the real number.

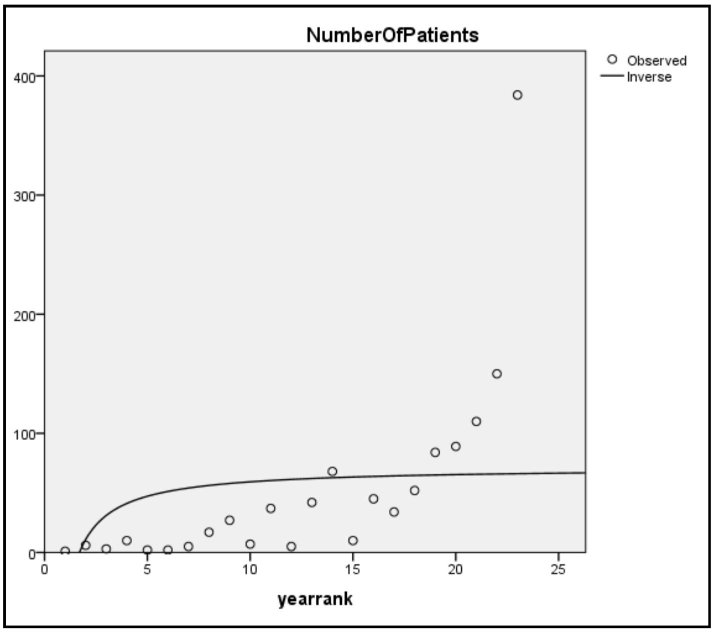

Figure 6. Inverse Equation Plots

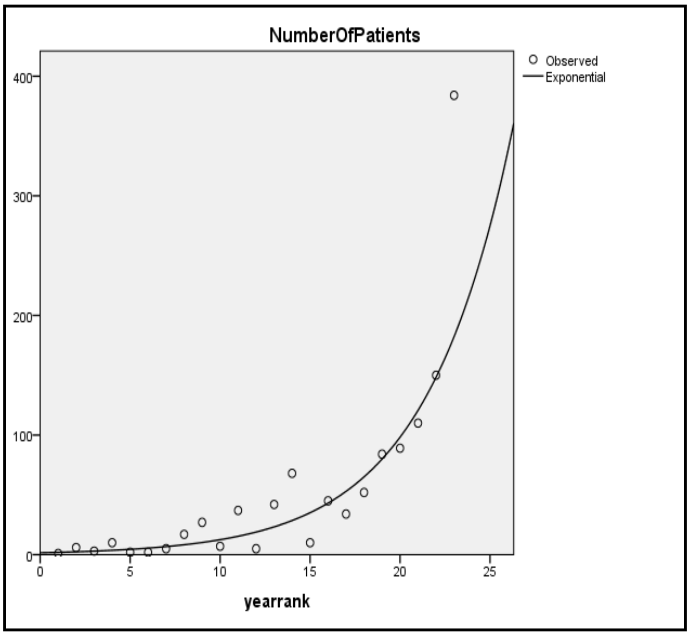

Figure 7. Exponential Equation Plots

Table 21. Expected number of patients and error between expecting and original values

Date	Year	Num	Fit-1	ERR-1
1990	1	1	1.96192	-.96192-
1991	2	6	2.41063	3.58937
1992	3	3	2.96197	.03803
1993	4	10	3.63940	6.36060
1994	5	2	4.47176	-2.47176-
1995	6	2	5.49449	-3.49449-
1996	7	5	6.75114	-1.75114-
1997	8	17	8.29518	8.70482
1998	9	27	10.19237	16.80763
1999	10	7	12.52346	-5.52346-
2000	11	37	15.38769	21.61231
2001	12	5	18.90700	-13.90700-
2002	13	42	23.23121	18.76879
2003	14	68	28.54441	39.45559
2004	15	10	35.07278	-25.07278-
2005	16	45	43.09426	1.90574
2006	17	34	52.95032	-18.95032-
2007	18	52	65.06056	-13.06056-
2008	19	84	79.94052	4.05948
2009	20	89	98.22367	-9.22367-
2010	21	110	120.68834	-10.68834-
2011	22	150	148.29090	1.70910
2012	23	384	182.20642	201.79358

Table 22. ANOVA table for parameter

	Sum of Squares	d.f.	Mean Square	F	Sig.
Regression	42.930	1	42.930	75.687	.000
Residual	11.911	21	.567		
Total	54.842	22			

The independent variable is yearrank.

Exponential Curve Fit[xvi]: We want to test wither $\beta_1 = \beta_2 = 0$ or not, so we will use ANOVA table to proof the following hypothesis testing.

$$H_0 = \beta_1 = \beta_2 = 0$$
$$H_1 = \beta_1 \neq \beta_2 \neq 0$$

We can see that the significant value P-value is approximately $= 0.000 < 0.05$ so we will reject H_0. This means that our regression model is significant.

Conclusion

- The more people with diabetes in our sample are those aged between 45 to 65 years old.
- Number of female with diabetes is more than number of male in the sample, the proportion of females in the sample reaches to 62.7%.
- Most number of people with diabetes is those people whose weights among 73-85 kg with proportion 35.8%.

- Most number of people with diabetes are those people whose lengths among 154-163 cm with proportion 30.7%.
- In the recent years the number of people is in continuous increase, where the highest rates of infection in the period 2008-2013 by up to 66.4% of the sample size.
- People with genetic diabetes represents 58.7% of our sample, and people with other diabetes causes by up to 41.3%.
- The number of male with diabetes type1 by up to %59 is more than the number of female. At the same time, the number of female by up to %64 with diabetes type2 is more than the number of male.
- In both, genetic diabetes cause and other diabetes cause, the number of females are more than the number of males.
- We note that from correlation coefficient analysis parts, that age variable has strong positive linear relationship with both types of diabetes, weights, lengths and date of diabetes variables equal to +1.
- In the correlation coefficient part, we also see that diabetes type1 have strong negative relationship with weight equal to -%80.
- Regression modelling for number of patients in Duhok per year with 95% confidence interval is: $\hat{\square} = 1.597 * \square^{0.206(\square)}$
- Maybe the reason of increasing in the number of patients with diabetes in Duhok city in recent years is because the disturbed changes in the region economy starting from 1980 to 1988 through the Iraq-Iran war, and the economy siege from 1990 to 2003. Then, détente this economic crisis after 2003 and speeding citizens. All of these reasons may be effect on the high increasing the number of people with diabetes at recent years.
- We recommend researchers in their future studies to take into consideration the effects of economic changes, climate changes and social situation of people with diabetes.
- We recommend that there should be more statistical studies for row data of people with diabetes, and the comparisons that shows the differences between the number of diabetics in different regions, to determine the causes of acquiring the disease or the causes that leads to an increasing in the incidence of the diabetes.

REFERENCES

[i]Comparative Statistical Analysis of Inpatients with Diabetic Myocardial, Priscilla O.OkunjiPhd, Afrooz Afghani.

[ii]Zealure C. Holcomb, Fundamental of Descriptive Statistics.

[iii]ArisSpanos, Probability Theory and Statistical Inference, p243.

[iv]Wlodzimierz Bryc, The Normal Distribution Characterizations with Applications, p24.

[v]RachadAntonios, Interpreting Quantitative Data in SPSS, p51.

[vi]Thomas A. Lang, Michelle Secic, How to Report Statistics in Medicine, p82.

[vii]Sharon Lawner, Sarah K Napp, Statistics Using SPSS an integrative approach, p138

[viii]Barbara Hazard Munro, Statistical Methods for Health Care Research, p137.

[ix]BovasAbrahamas, Johannes Ledolter, Introduction to Regression Modeling, p82

[x]Ronald H. Heck, Scott L. Thomas, Multilevel Modelling of Categorical Outcomes Using IBM SPSS, p232.

[xi]Rudolf J. Fruend, William J. Wilson, Statistical Modelling of a Response Variable, p102

[xii]Frederick J. Gravetter, Larry B. Wallnau, Statistics For The Behavioural Sciences, p563

[xiii]Leona S. Aiken, Stephen G. West, Multiple Regression: Testing and Interpreting Interactions, p29

[xiv]Deburah Rumsey, Intermediate Statistics for Dummies, p144.

[xv]P. G. Guest, Numerical Methods of Curve Fitting, p211.

[xvi]Sandra L. Arlinghaus, Practical Handbook of Curve Fitting, p36.

****