A PROJECT REPORT

on

"To Study Effect Of Various Factors On Body Weight Of Kids & Dam Milk Production Of Goats"

Submitted By

Miss.Gargade Shweta maruti

M.Sc. II Statistics (2016-2017)

Under the guidance of,

Dr. D. M. Sakate

Department Of Statistics

Shivaji University, Kolhapur

CERTIFICATE

This is certify that the project report entitled *"To Study Effect Of Various Factors On Body Weight Of Kids &Dam Milk Production Of Goats*"being submitted by***Miss Gargade Shweta Maruti***, as a partial fulfillment for the award of degree of M.Sc. in Statistics of Shivaji University, Kolhapur, is a record of confidework carried out by her under my supervision and guidance.

To the best of my knowledge the matter presented in the project in the project has not been submitted earlier.

Dr. D. M. Sakate
Project Guide
Department Of Statistics
Shivaji University, Kolhapur

Dr. D. N. Kashid
Head of the Department
Department Of Statistics
Shivaji University, Kolhapur

Place: Kolhapur

Date:

ACKNOWLEDGEMENT

We wish to thank Dept. of Statistics (Shivaji University, Kolhapur) for giving us an opportunity to do a project.

We are thankful to Dr. ChandaNimbkar, Director Nimbkar Agricultural Research Institute, Phaltan, for helping us by giving data and guidance throughout. We express a deep sense of gratitude towards her.

This report has been prepared under the guidance of Dr. D.M. Sakate. We would like to express our profound gratitude towards him for his guidance we constructive throughout this project.

Also we would like to thank Dr. D. N. Kashid and all staff members for their support, suggestions and guidance for this project.

Finally we would like to thank HOD, teachers, non-teaching staff and Research students of our Department for their valuable co-operation in this project.

Contents

1. Introduction :-

Being students of M.Sc.II (Statistics), we were more interested in knowing how the data are handled in Animal husbandry sector. We were really interested in analysis of Body measures of Goats and their fodder species. In this regard, Dr. D. N. Kashid, Head of Department requested to Nimbkar Agriculture Research Institute, Phaltan for giving an opportunity to work with them.

As, India is an agricultural country, large part of the population is of farmers. But one should not forget that there are many factors which influence the prices of agricultural product, cattles etc. So being aware of this fact in the life, the farmers should start some additional occupation with the agriculture. In order for social and economic welfare of farmers many organizations of government and also private institutes offer some scheme. One of such private institute is "Nimbkar Agricultural Research Institue"(NARI).

NARI has working lot for the welfare of farmers. Therefore with the help of statistical tools and figures we can bring the scenario of agricultural facts in front of farmers and ultimately make then awarenof additional business to agriculture in order to get secured against the unfavourable events like abrupt change in market prices of agricultural products, uncertainty involved in mansoon changes in climatic condition.

Animal Huabandary Division of NARI is working now on various aspects in this concern. Mainly rearing of goats cultivation of the nutritional fodder required by them and breeding activities are performed. So with the help of some statistical tools we are trying to bring the awareness in people about the objectives of NARI for the welfare of farmers.

2. Technical Information :-

Background History of NARI

The Osmanabadi Goat Field Unit is a new unit under the AICRP on Goat Improvement sanctioned to the Non-Govt. Organization Nimbkar Agricultural Research Institute located at Phaltan town, 110 km south-east of Pune in rural Maharashtra State. The unit has been sanctioned for the years 2009-10 to 2011-12.

NARI is an NGO and non-profit research and development institute. The Nimbkar Agricultural Research Institute (NARI) was established by Mr. B.V. Nimbkar on 12 March 1968 and is registered under the Societies Registration Act XII of 1860 and the Bombay Public Trust Act of 1950. NARI undertakes research, development and extension in agriculture, renewable energy, animal husbandry and sustainable development. NARI is recognized by the Department of Scientific and Industrial Research of the Government of India as a scientific research organization. NARI has carried out village goat improvement programs successfully since 1990.

Weight measurement of Goats: -

Dam milk slukling kids

Osmanabadi Breed **Gaoran Breed**+++

3. Project Problem :-

India being agricultural country, most of the people depend on agriculture for their daily bread and butter. But now-a-days, due to drastic change in the environmental conditions and natural calamities, dependency on agriculture for daily needs has become risky. This project is based on the idea of bringing the cattle among different rearing as joint business to agriculture. Our focus is mainly weight(kg) of kids and production of Dammilkbreeds of goats (cross-breeding).The breed for kids for particular region and season is also the important factor in rearing. Here, the Gaoran and Osmanabadigaots are studied and this is achieved by using statistical techniques.

The prime objective of this project is to check the effect of breed, Wt age, Dam parity, Birth season, Number born on body weight of kids. And also to check the effect of Breed , Wt age, Weight , Dam parity on Number of born kids. The comparisons can be among different breeds of goat, seasons etc.To check the behavior of Dam Morning milk, Dam Night milk and Dam total milk of different types of breed over the period march 2013 to march 2017.

4. Methodology :-

a. Data collection and Data Description :-

The entire data used for the analysis is obtained from Nimbkar Agricultural Research Institute (NARI) Animal Husbandry Division Phaltan. For this I visited NARI. After getting introduced to Director, Dr. ChandaNimbkar she gave us data on goats and she described the entire scenario to us.

The data used for the analysis is of secondary type. It was collected periodically by the workers in NARI as a routine work. It comprised of body weights of kids (Osmanabadi and Gaoran ,Osmanabadi-X), and also different types of Dam milk production(osmanabadi , Gaoran).In present data of body weights of kids, we have 1282instance and 11 variables.

As discussed earlier, many types of goats are reared in NARI. We managed to get data on only three types of kids (osmanabadi and Gaoran and Osmanabadi-X). It is comprised of different variables like body weight(KG), Age at weight, Bith date, Birth season, City, Sex, dam parity and Number Born. We also collected data of milk production of different Breeds namely Osmanabadi and Gaoran.The data in tabular form is given in appendix.

4.2 Statistical Tools :-

1. Exploratory Data Analysis
2. Non-parametric one way ANOVA
3. Ordinal Logistic Regression
4. Artificial Neural Network (Regression)
5. Time Series Analysis

4.3 Softwaresused :-

1. MINITAB
2. SAS

3. R
4. SPSS
5. MS-Excel
6. STATISTICA

5. Some Statistical Techniques and Concepts to be used :-

In this topic, we have given description about statistical concept which is helpful to us analyzing the kids and Dam milk data. We also discuss some terms of Kids and Dam milk related to our project work.

5.1 Non-Parametric one way ANOVA : - (Kruskal–Wallis test)

The measurement variable does not have the normality assumption of a one-way ANOVA in this situation we use Non-parametric ANOVA.
The **Kruskal–Wallis test** by ranks, or **One-way ANOVA on ranks** is a non-parametric method for testing whether samples originate from the same distribution It is used for comparing two or more independent samples of equal or different sample sizes. The parametric equivalent of the Kruskal-Wallis test is the one-way analysis of variance (ANOVA).

The test statistic is given by:

$$H = (N-1)\frac{12}{N(N+1)}\sum_{i=1}^{K}\frac{R_i^2}{n_i} - 3(N+1)$$

Where,
n_i is the no. of observations in group i.
N Total no. of observations across all group.
R_i is sum of rank of the ranks of the n_i observations in the i^{th} sample
.

5.2Ordinal logistic Regression :-

The logistic regression model for ordinal response is a Propotional odds model. If in our data set response is binary then wecan usethe logistic regression model. When response have more than two categories and are ordered we could run multinominal regression model.The disadvantage is that we are throwing

away information about the ordering. An ordinal ogistic regression model preserves that information.

Cumulative Logits :

Consider a case where there are J possible categorical outcomes which are ordinal. The usual approach to model this type of response, data is to use logits of cumulative probabilities.

The cumulative logits are defined as,

$$logit[P(y \leq j|x)] = log\frac{P(Y \leq j|x)}{1 - P(Y \leq j|x)}$$

$$= log\frac{\pi_1(x)+\cdots\ldots\ldots\ldots+\pi_j(x)}{\pi_{j+1}(x)+\cdots\ldots\ldots\pi_J(x)} \quad \text{j=1,2,… J-1}$$

Each cumulative logits use all J response categories.

Proportional Odds Model:

A model that simultaneously uses all cumulative logits is

$logit[P(y \leq j|x)] = \propto_j + \beta' x$ j=1,2,…….J-1.

Each cumulative logit has its own intercept. The $\{\propto_j\}$ are increasing in j, since $logit[P(y \leq j|x)]$ increases in j for fixed x, and the logit is an increasing function of this probability.

This model has the same effects β for each logit. For a continuous predictor x.

Data mining techniques:

Confusion Matrix:

A confusion matrix contains information actual and predicted classification done by a classification system.

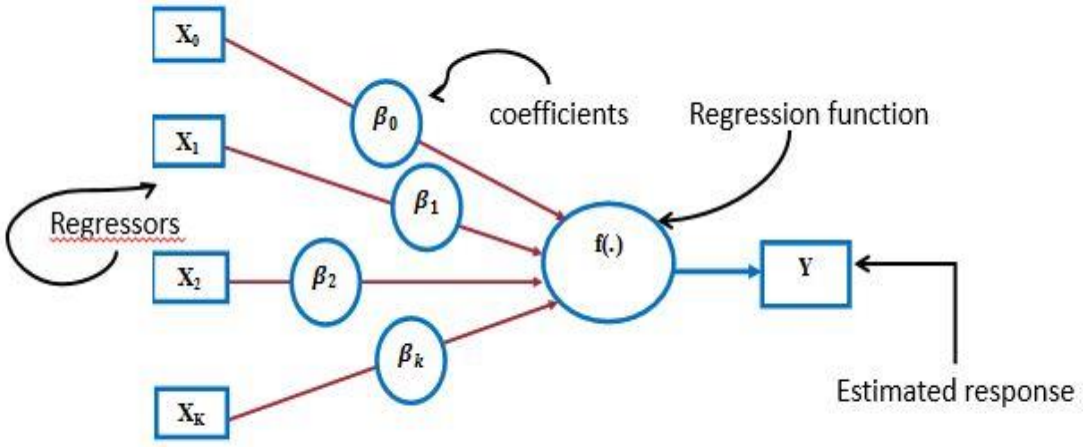
	Predicted	
Actual	Negative	Positive
Negative	TN	FP
Positive	FN	TP

Accuracy = (TP+TN)/ (TN+TP+FN+FP)

5.3 ANN (Artificial Neural Network): - Regression

A neural network is a network of massive interconnections of simple computing units called as neurons or processing units. An artificial neuron is a mathematical function conceived as a simple model of a real neuron.Neuron model has four components: input values, weight vector,Output values, threshold values.

Regression Model :A regression model with k independent variables is similarto a single layer feed-forward neural network as shown below:



Sigmoid function:

The non-linear curved S- shape function is called the sigmoid function. This is most common type of activation used to construct the neural networks. It is differential and strictly increasing function.

$$Y= \frac{1}{1+e^{(-\lambda * net)}} 0 <= Y <= 1$$

Single layer feed-forword network consists of a single layer of weights, where inputs are directly connected to outputs.

5.4 Time Series Analysis:

For analyzing this I will first identify patterns in time series data (such as smoothing and curve fitting techniques and autocorrelations), then I will use models that can be used to represent time series data and generate predictions (autoregressive and moving average models). Finally, I will apply some simple but commonly used modeling and forecasting techniques based on given model.

The Different types of model are,

a) Linear Trend Model

$$Y_t = a + b X_t$$

b)ARIMA(0,0,1)

$$X_t = Z_t + b Z_{t-1} \qquad \text{where, } Z_t \sim WN(0, \sigma^2)$$

6. Objectives :-

1) **To investigate the affect of Breed on Body weight of kids of Goat.**
2) **To investigate which factors affected on the Number Born kids.**
3) **To investigate which factors affect on the body weight of kids.**
4) **To forecast the milk production of dams for next 5 months.**

7.Data Pre-processing :-

Data which is obtained from the NARI. We have two data sets for the analysis. In the data set of body weight of goats, which includes the goats "age at weight" from 1 day to 650 days. In this whole scenario we consider only birth state weight of kids (i.e. age at weight from 1 day to 30 days). In this data set

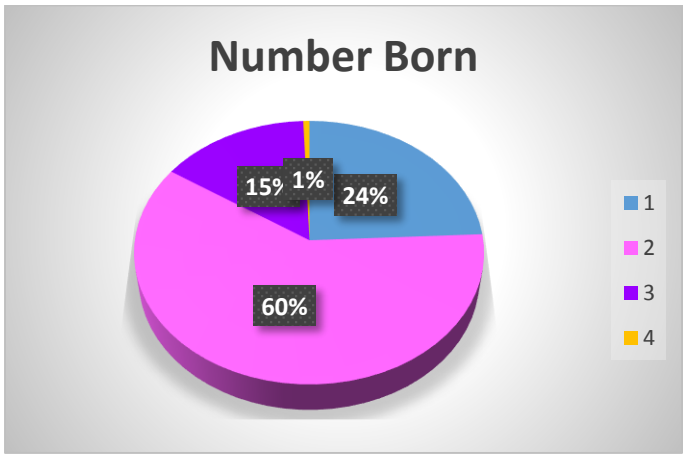
there was some outliers in the variable weight(kg) we deleted these observations form the data.

In second data set of Dam milk of production of Osmanabadi and Gaoran breed, there were some missing values in daily dam milk production. In order to deal with these missing values , the daily milk production is converted into monthly production, by taking average of these values.

8.Analysis:-

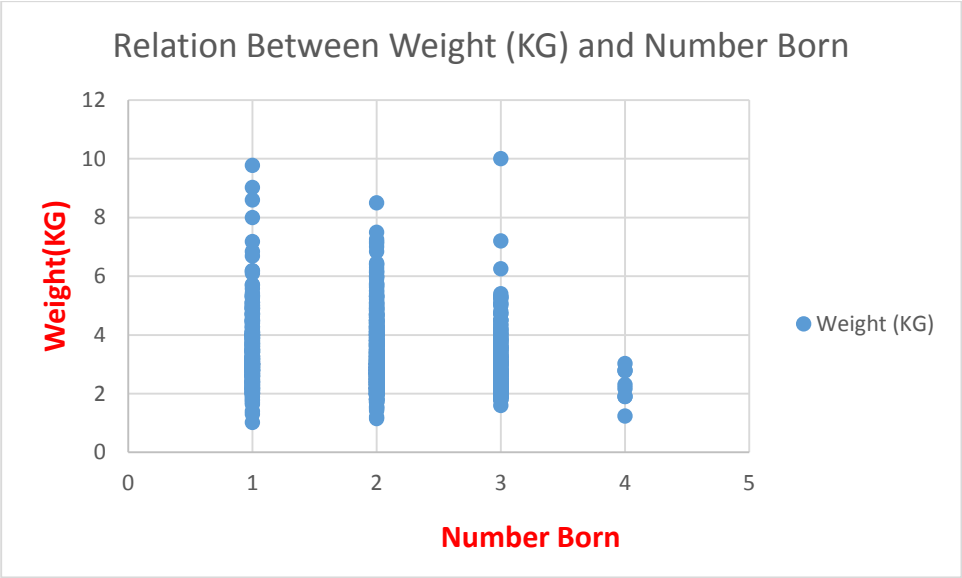
8.1 EXPLORATORY DATA ANALYSIS:

8.1.1 Number Born :-



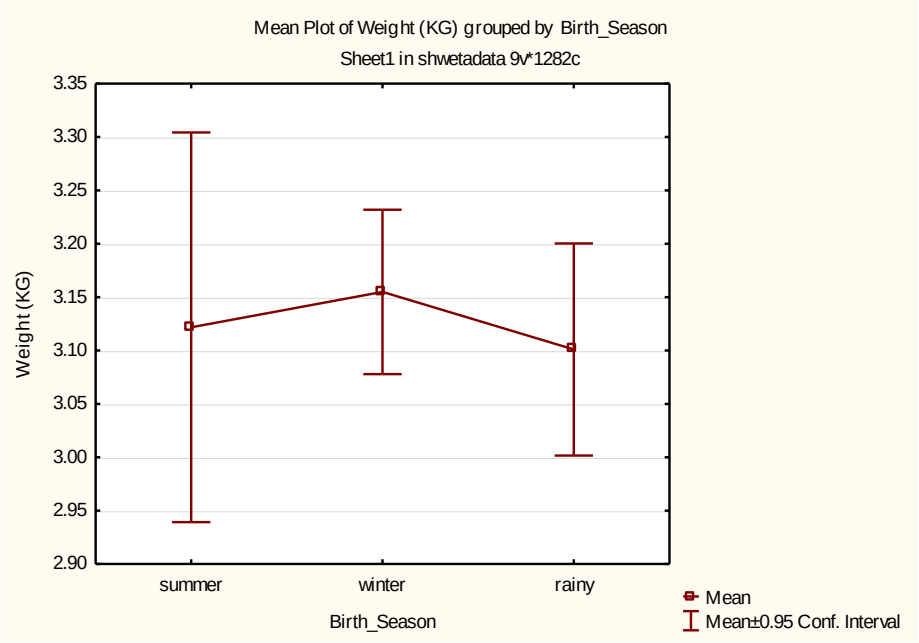
Conclusion :From the Above Pie chart we conclude that the percentage of giving birth to ***single*** kid is 24%, the percentage of giving birth to ***two***kids is 60%. The percentage of giving birth to ***three*** kids is 15% and percentage of giving birth to ***four*** kids is negligible (i.e. 1%).

8.1.2 Relationship between Number Born and weight(KG) of kid at birth :-



Conclusion :-From the above scatter diagram we conclude that as Number born increases the average birth weight decreases.

8.1.3 Birth_Season - Weight(KG) :



Conclusion: From the above mean plot of weight(KG) by birth season we conclude that, weight of the kids is ***larger*** in ***winter season*** than the other two seasons.

8.1.4 OnwerID Group – Average weight :

a) Total – OwnerID group wise Count and Average wt(KG) of kids :

OwnerID Group	Count	Average wt
1 to 5	261	3.259923372
6 to 10	253	3.13055336
11 to 15	244	2.970040984
>15	524	3.149465649

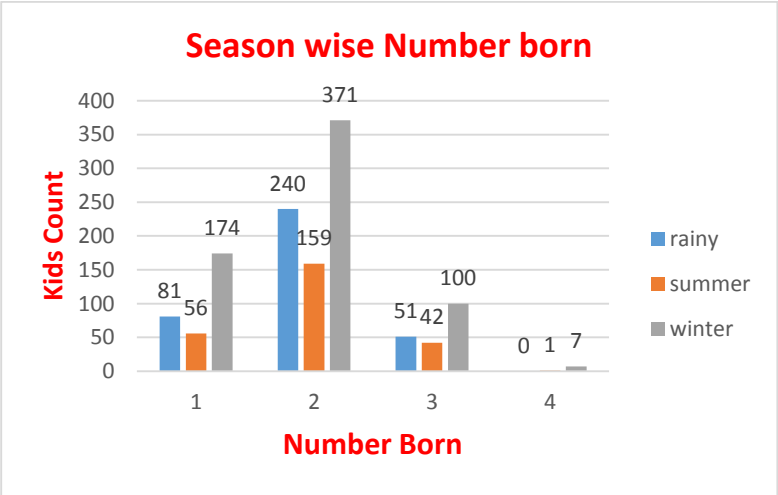
Conclusion: From the table we conclude that, OwnerID group ***"1 to 5"*** have ***larger*** weight than the other OwnerID groups. Also it can be seen that among all the OnwerIDgroupsaverage weight of the kids are nearly same.

b) Per Breed- OwnerID group wise Count and Average wt(KG) of kids :

	Breed count			Breed Average wt(kg)		
OwnerID Group	**Gaoran**	**Osmanabadi**	**Osmanabadi-X**	**Gaoran**	**Osmanabadi**	**Osmanabadi-X**
1 to 5	21	218	22	3.4128571	3.27105504	3.2436363
6 to 10	11	209	33	3.1990909	3.12923445	3.1160606
11 to 15	18	204	22	3.1011111	2.9972549	2.610454
>15	62	415	47	3.2235483	3.1275180	3.0055319

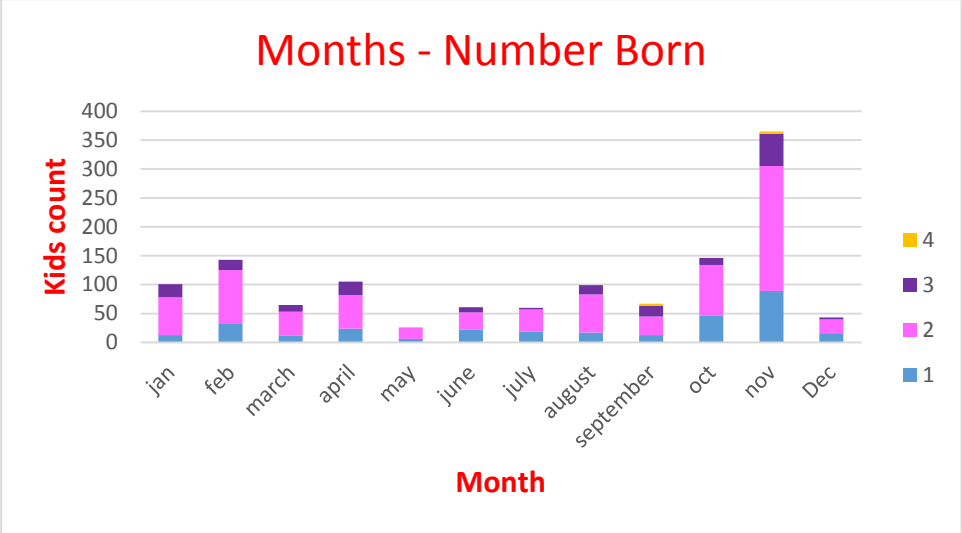
Conclusion: From the table we conclude that, OwnerID group ***"1 to 5"*** have ***larger*** weight than the other OwnerID groups, and also seen that, kids of ***Gaoran*** breed have ***larger*** weight than other two breeds.

8.1.5 Season wise Number Born of kids:-



Conclusion: From the above Multiple bar diagram we conclude that, for each Number born category kids count in ***winter*** season is ***high*** than other seasons.

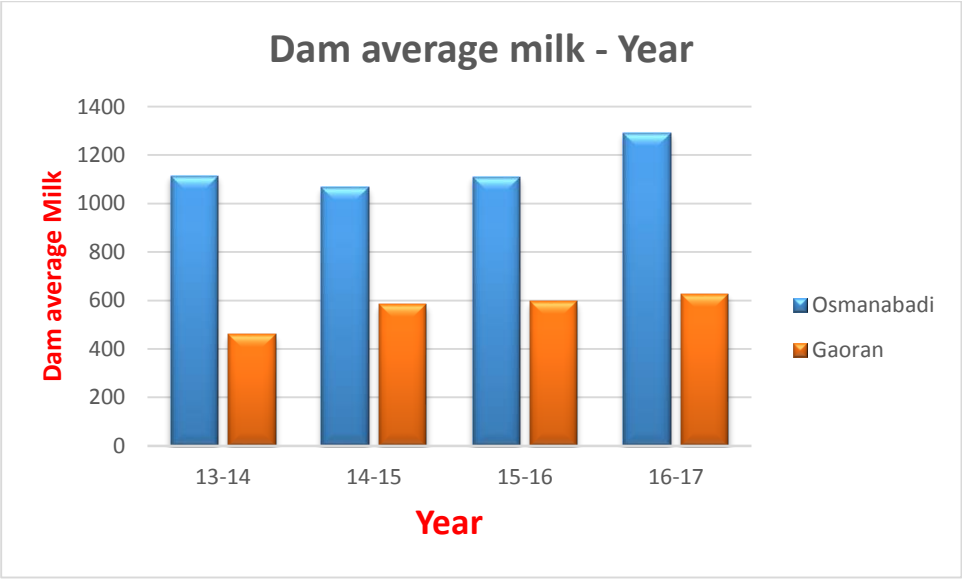
8.1.6 Month – Number Born :-



Conclusion : Form the above sub-divided bar diagram we conclude that, among all the months Number born ***"2"*** kids is ***high*** in ***November*** month.

Analysis for Dam milk:

8.1.7 Dam average milk – Year :



Conclusion: we conclude that, ***Osmanabadi*** breed have ***maximum*** Dam average as compare to Gaoran breed. And also it can be seen that, from the year 2013 to 2017 there is slightly ***increasing trend***.

8.2 ANALYSIS FOR KIDS OF GOATS :

8.2.1 To investigate the Breed affect on weight of kids of age less than one month:-

Data used is given in Table 1.1(Appendix)

In order to determine, the effect of Breed category on Body weight(KG) of kids.

We can analyze the data by using Non-Parametric ANOVA.

For this test hypothesis is,

H_0 :The distribution of weight of kids is the same across the category Breed.

V/s

H_1 :The distribution of weight of kids is not the same across the category Breed

Using SPSS,

Descriptive Statistics:

	N	Mean	Std. Deviation	Minimum	Maximum
Weight	1282	3.1341	1.07047	1.02	10.00

Ranks

	Breed	N	Mean Rank
Weight	Gaoran	112	693.85
	Osmanabadi	1046	639.58
	Osmanabadi-X	124	610.44
	Total	1282	

Test Statistics

Chi-Square	3.1409
Df	2
Asymp. Sig.	0.2080

Asymptotic significance ie. P-value is greater than 0.05. There fore , we **accept** the null hypothesis. And we conclude that The distribution of weight of kids is the **same** across the category Breed.

Conclusion: There is no effect of Breed on the Weight(KG) of kids age less than one month.

8.2.2 To investigate which factors affected on the Number Born kids:

Here all 1282 observations in our data set were used in the analysis.

Variable selection:

Applying variable selection method for building a model we have get those variables which are appropriate. Here variables are selected in the model by using R software.

By using **backward selection** method we get following Significant Variables

Breed Weight_KG_ Wt_AgeDam_parity

Response :No_born

Response level : 4

By using SAS software ordinal logistic regression model is,

Now we see that the model convergence :

Convergence criterion (GCONV=1E-8) **satisfied.**

The Model Fit Statistics :

Residual Deviance (-2 log likelihood): 2324.319 AIC: 2338.319

Analysis of maximum likelihood estimate :

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	**1**	1	-1.1255	0.3517	10.2392	0.0014
Intercept	**2**	1	1.9901	0.3572	31.0404	<.0001
Intercept	**3**	1	5.4837	0.4986	120.9785	<.0001
Breed		1	-0.3329	0.1311	6.4415	0.0111
Weight__KG_		1	0.5406	0.0750	52.0036	<.0001
Wt_Age		1	-0.0558	0.0109	26.3141	<.0001
Dam_Parity		1	-0.1789	0.0185	93.6587	<.0001

From the P-value we can say that all variables are statistically significant.

The proportional odds assumption :

Test hypothesis is,

H_0: proportional odds are non-significant

H_1: proportional odds are significant

Score Test for the Proportional Odds Assumption		
Chi-Square	**DF**	**Pr > ChiSq**
11.4608	8	0.1769

we also see that the test of the proportional odds assumption is **non-significant**. One of the assumptions underlying ordinal logistic regression is that the relationship between each pair of outcome groups is the **same.** In other words, ordinal logistic regression assumes that the coefficients that describe the relationship between, say, the lowest versus all higher categories

of the response variable are the same as those that describe the relationship between the next lowest category and all higher categories, etc.Because the relationship between all pairs of groups is the same, there is only one set of coefficients (only one model).

Proportional odds ratios and their 95% confidence intervals :

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
Breed	0.717	0.554	0.927
Weight__KG_	1.717	1.482	1.989
Wt_Age	0.946	0.926	0.966
Dam_Parity	0.836	0.806	0.867

Conclusion :

For **Breed**, we would say that for a one unit increase in **Breed**, (i.e., going from 0 to 1), the odds of " 4 " No_born versus the combined " 3 " , "2 " , " 1 " categories of No_born are 0.717 greater, given that all of the other variables in the model are held constant.

Because of the proportional odds assumption (see below for more explanation), the same increase, 0.717 times, is found between " 1 " No_born and the combined categories of " 2 " , " 3 " , " 4 " No_born.

And other variable of interest is Dam_parity. For **Dam_parity**, we would say that for a one unit increase in **Dam_parity**, (i.e., going from 0 to 1), the odds of " 4 " No_born versus the combined " 3 " , "2 " , " 1 " categories of No_born and low are 0.836 greater, given that all of the other variables in the model are held constant.

The models are,

$logit[P(No\_born \leq 1|x)]$= -1.1255 – (0.3329*Breed) + (0.5406*Weight_KG_) – (0.0558*Wt_Age)-(0.1789*Dam_parity)

$logit[P(No\_born \leq 2|x)]$= 1.9901–(0.3329*Breed) + (0.5406*Weight_KG_)– (0.0558*Wt_Age)–(0.1789*Dam_parity)

$logit[P(No\_born \leq 3|x)]$= =5.4837 –(0.3329*Breed) + (0.5406*Weight_KG_)– (0.0558*Wt_Age)–(0.1789*Dam_parity)

Now check the accuracy of the model we devide data into traning and testing part.Training data have 381 observations and test data have 901 observations.

Confusion matrix :

	Actual				
pred	1	2	3	4	Total
1	11	2	0	0	13
2	77	240	40	1	358
3	2	4	4	0	10
4	0	0	0	0	0
Total	90	246	44	1	381

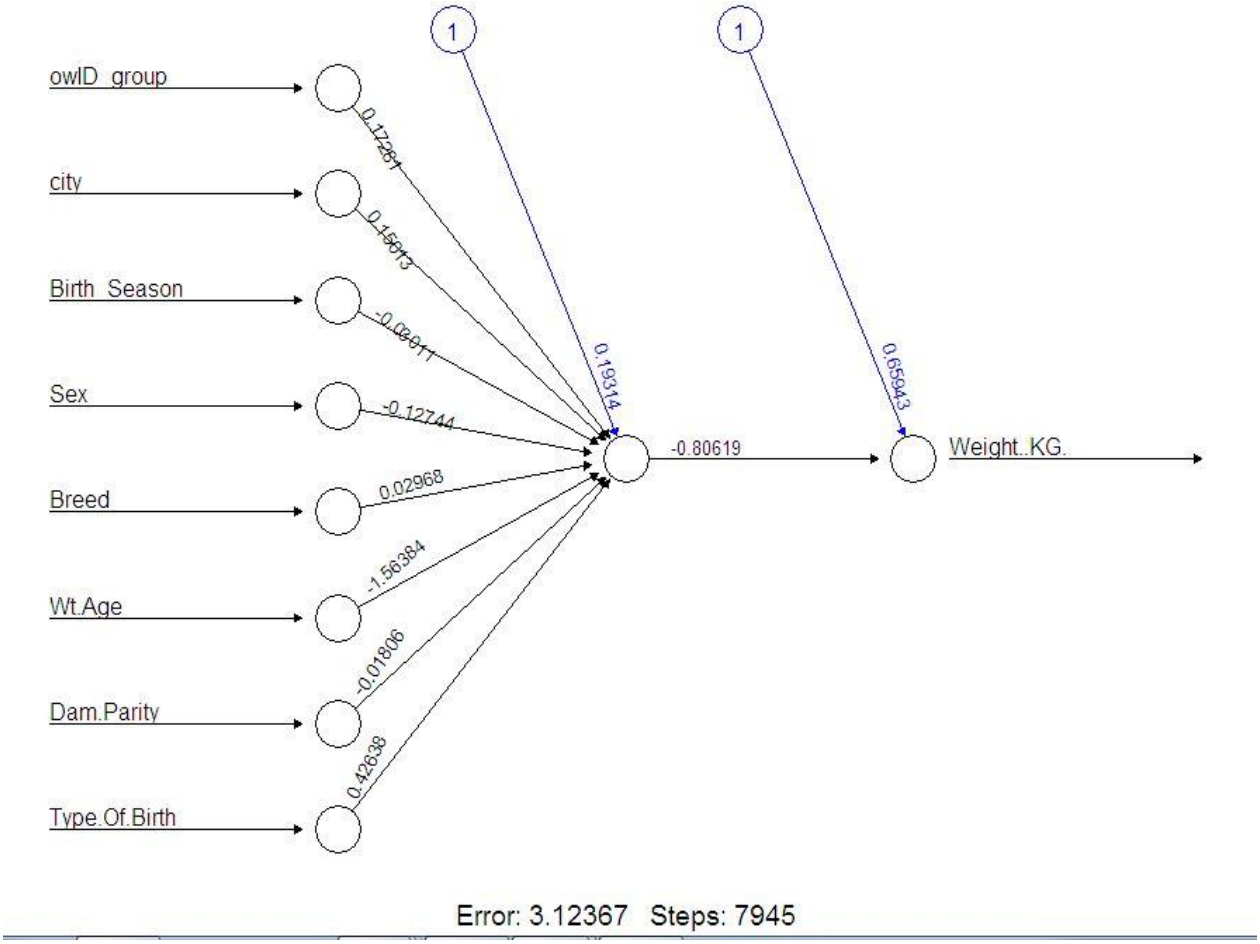
Accuracy = 66.92913 %

8.2.3 To investigate which factors affected on the Body weight of kids.

Here all 1282 observations in our data set were used in the analysis.

In order to determine, the effect of group of goats as per Owner, Dams wt, sex, wt age (in days), Number born and Dam parity City Breed on body weight of kids. We analyzed data by using Artificial neural network(Regression). By using R-software, the ANN model is given by,

error	3.12367
reached.threshold	0.009198
steps	7945
Intercept.to.1layhid1	0.193144
owID_group.to.1layhid1	0.172814
city.to.1layhid1	0.150127
Birth_Season.to.1layhid1	-0.03011
Sex.to.1layhid1	-0.12744
Breed.to.1layhid1	0.029675
Wt.Age.to.1layhid1	-1.56384
Dam.Parity.to.1layhid1	-0.01806
No_born.to.1layhid1	0.426377
Intercept.to.Weight..KG.	0.65943
1layhid.1.to.Weight..KG.	-0.80619



Conclusion: From the above analysis we can conclude that the variables Type of birth (Number born) , OwID group and city are more effective to the Weight of kids.

8.3 ANALYSIS FOR DAM MILK PRODUCTION :

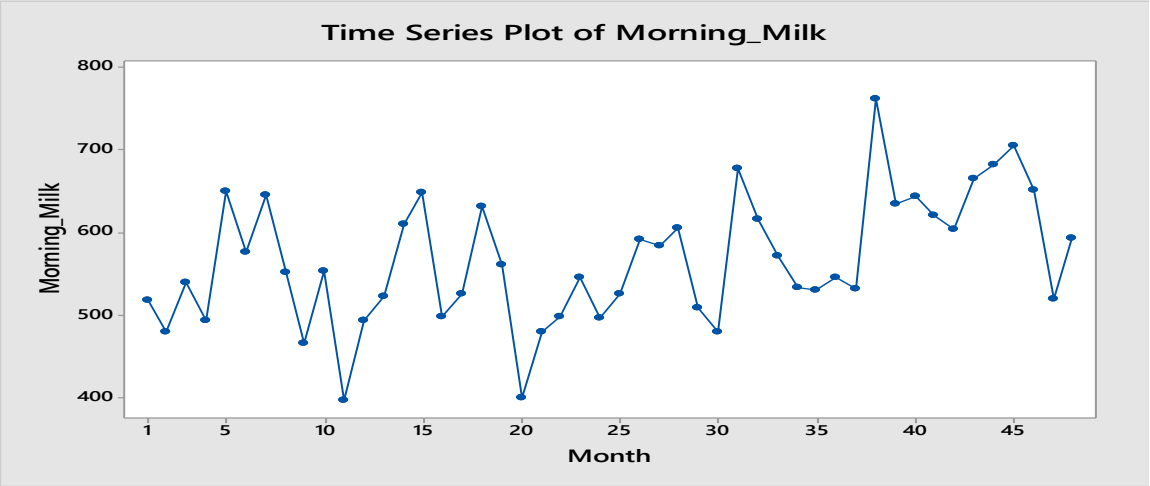
We want to study the behavior **Milk production of Goats** (namely Osmanabadi and Gaoran) on the basis of Satistical tools. For this purpose I used Time Series Analysis. I predict the future value and check the behavior of milk production.

For this problem I used **monthly**Milk production of Goats data from March 2013 to march 2017. This analysis is done by using MINITAB.

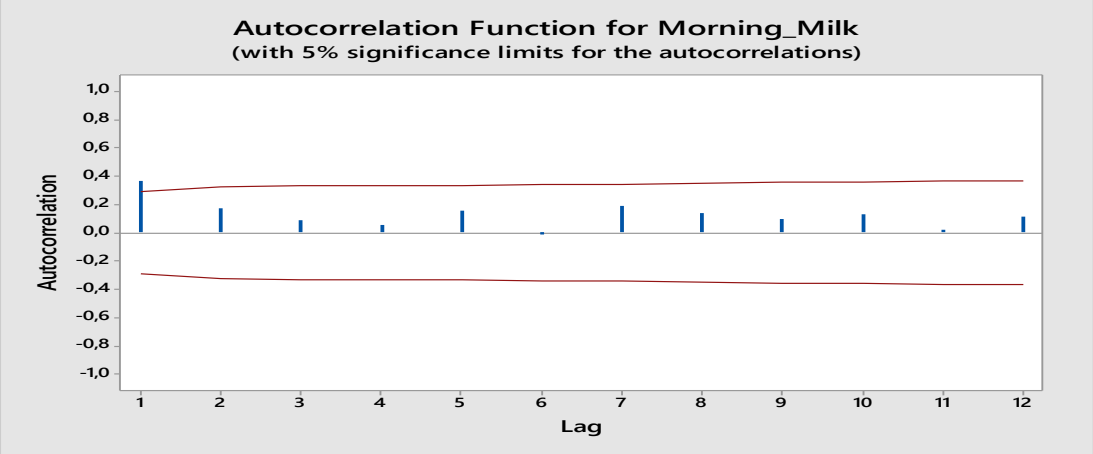
a) OsmanabadiBreed :-

8.3.1 To forecast the morning Dam milk production of Osmanabadi breed for next 5 days.

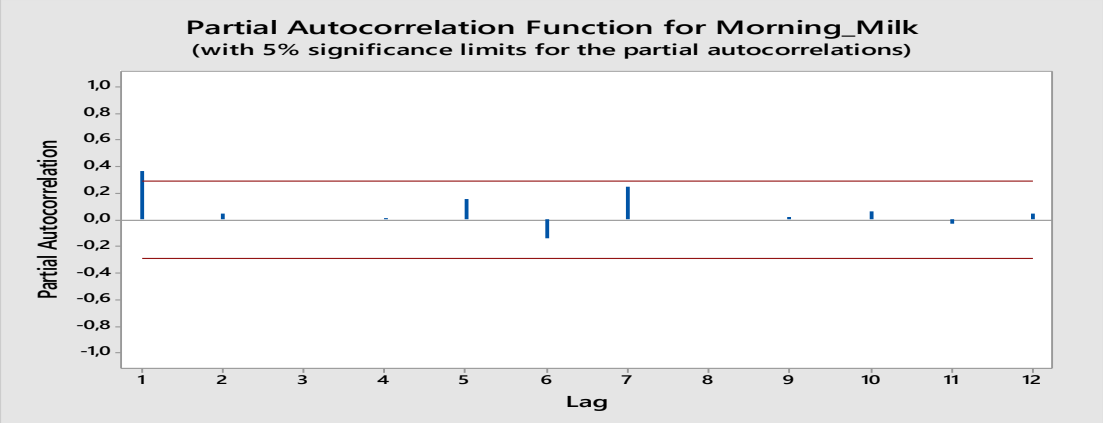
Time series plot for morning milk of Osmanabadi Goats:



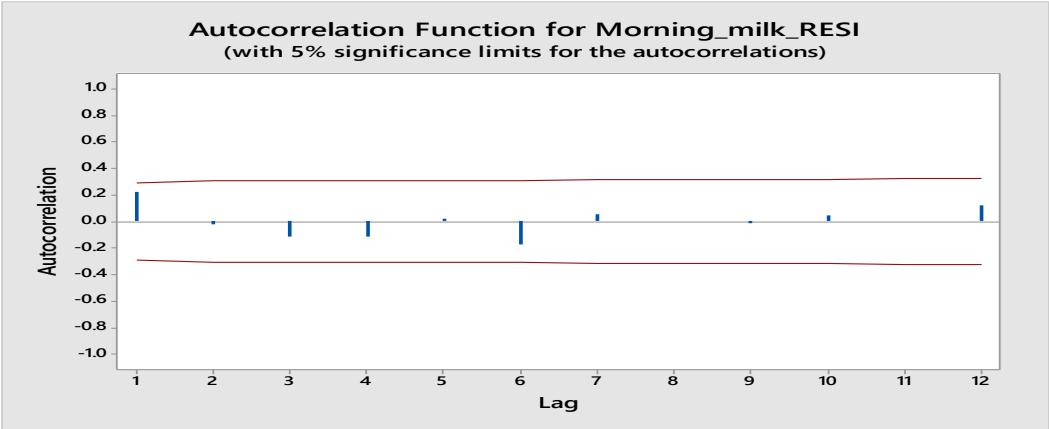
Autocorrelation Function for Motning milk



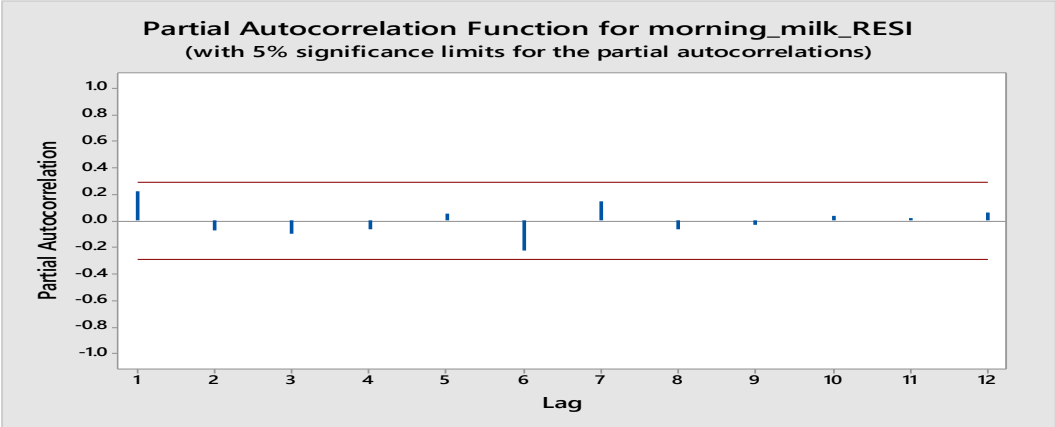
Partial Autocorrelation Function for morning milk



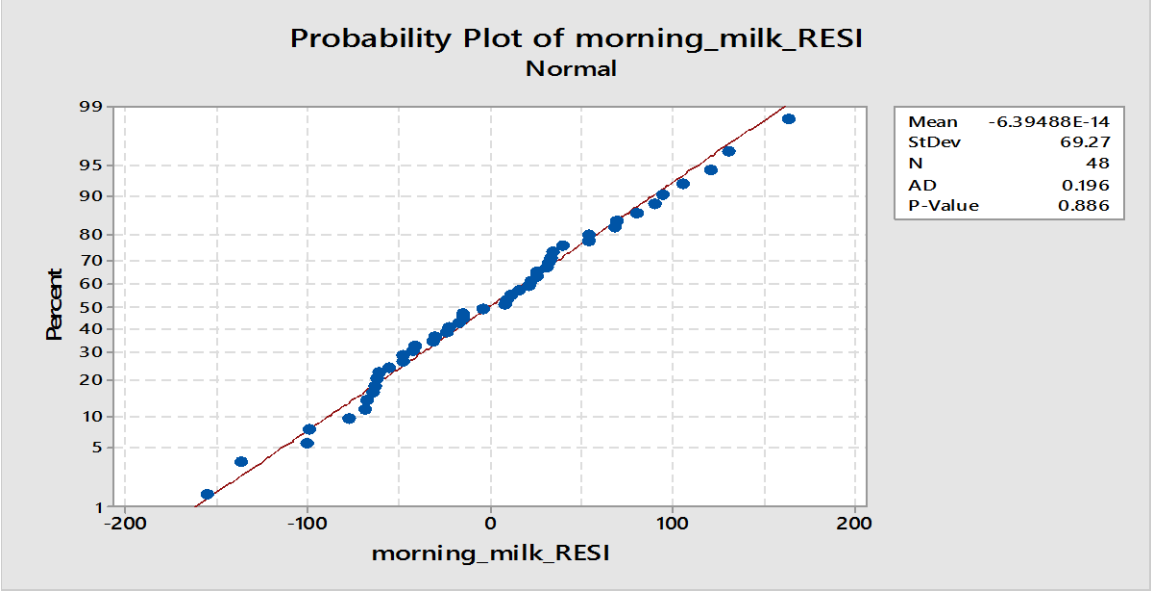
Autocorrelation Function for morning milk Residuals



Partial Autocorrelation Function for morning milk Residuals

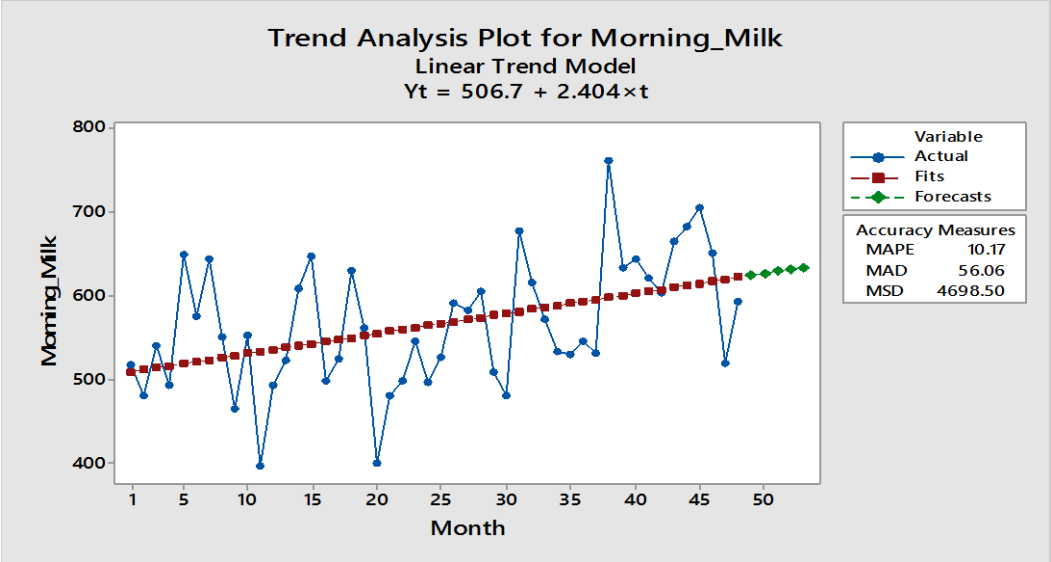


Now we check normality for resulting series by using normal probability plot as follows,



Here the P-value is **0.866**; therefore the given series is satisfying the

normality assumption.
And also, residuals has white noise with mean approxiately zero and variance 69.27.
Therefore by using Trend analysis we predict the next five values.

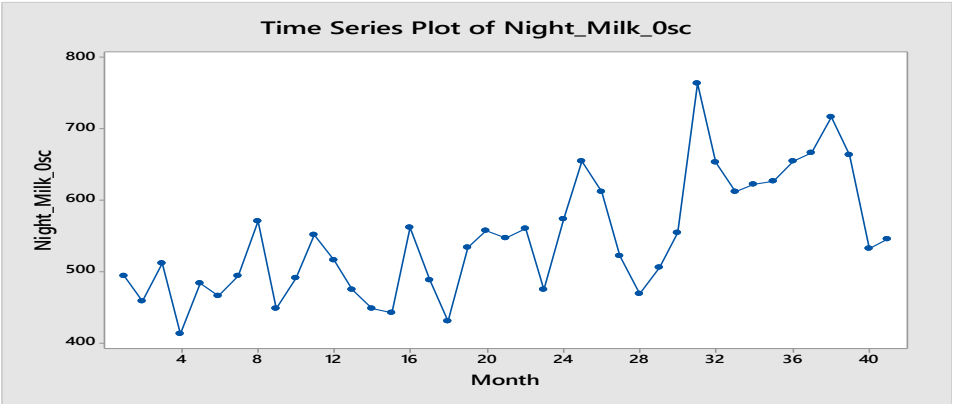


Forecasts:

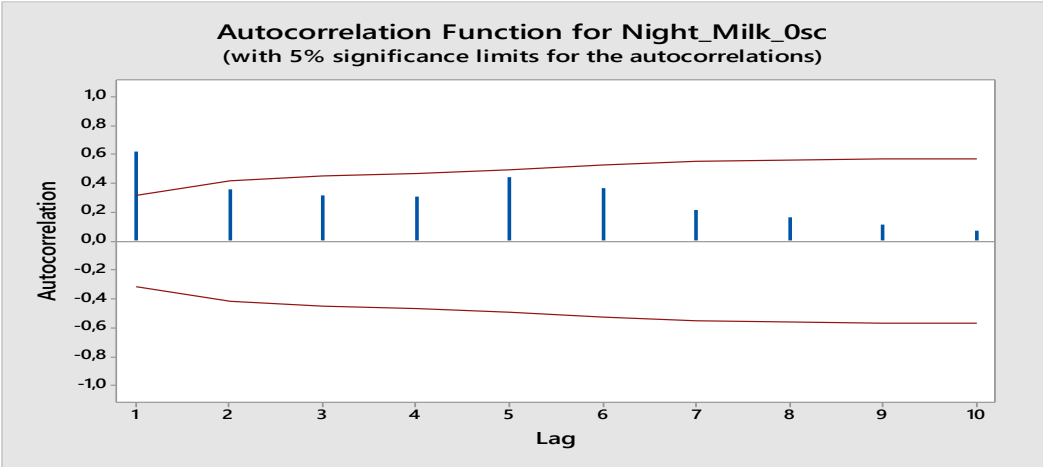
Period	Forecast
49	624.520
50	626.924
51	629.327
52	631.731
53	634.134

8.3.2 To forecast the Night Dam milkproductionofOsmanabadi Breed for next 5 days :

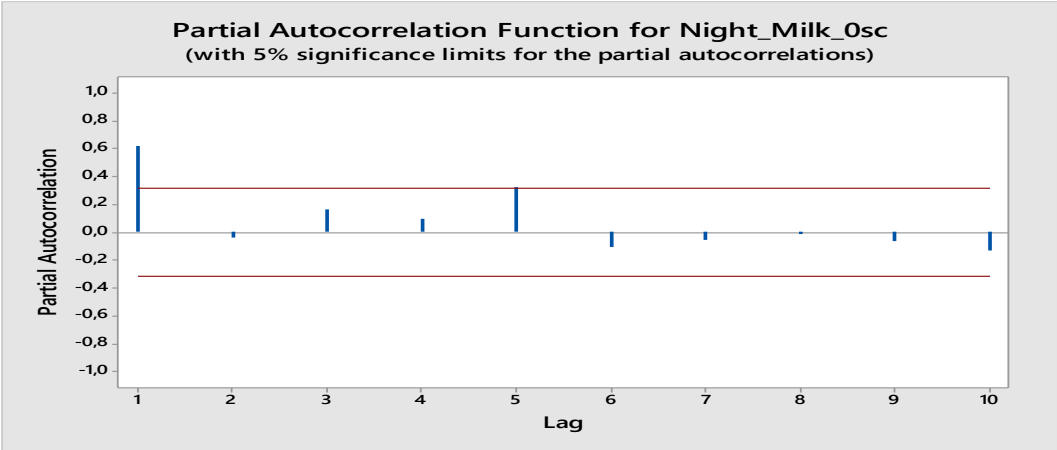
Time series plot for night milk of Osmanabadi Goats:



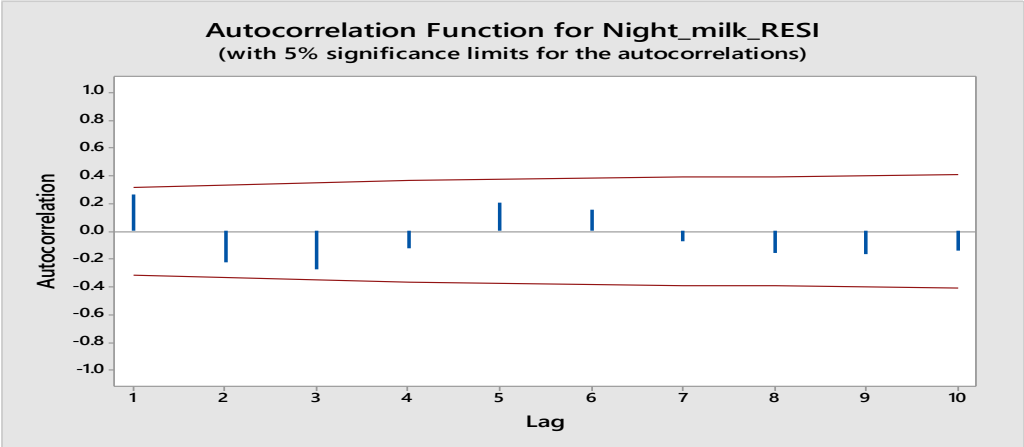
Autocorrelation Function for Motning milk



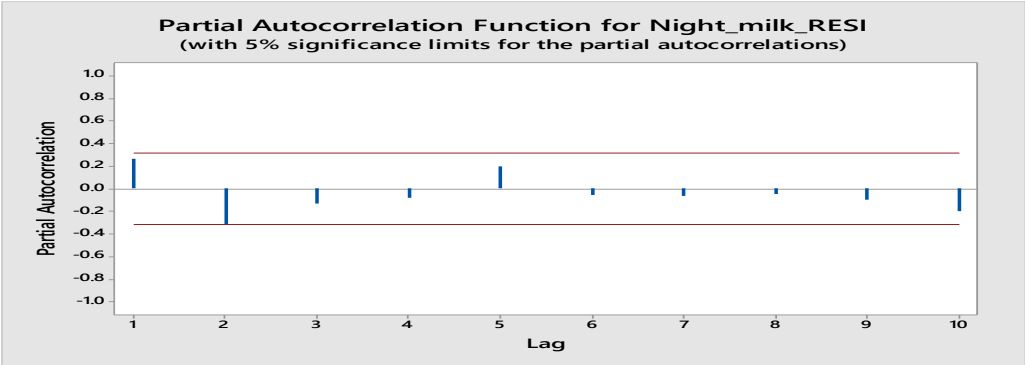
Partial Autocorrelation Function for Motning milk



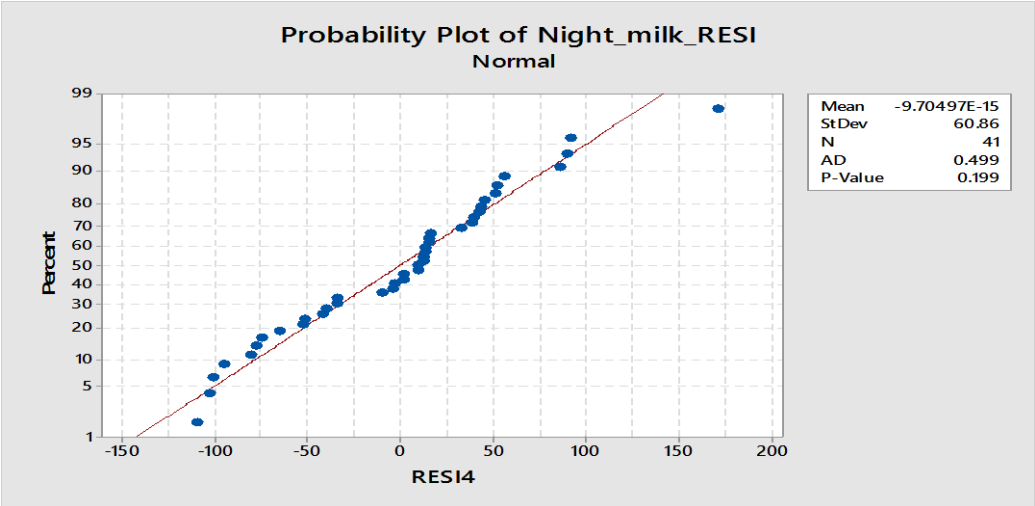
Autocorrelation Function for Motning milk Residuals



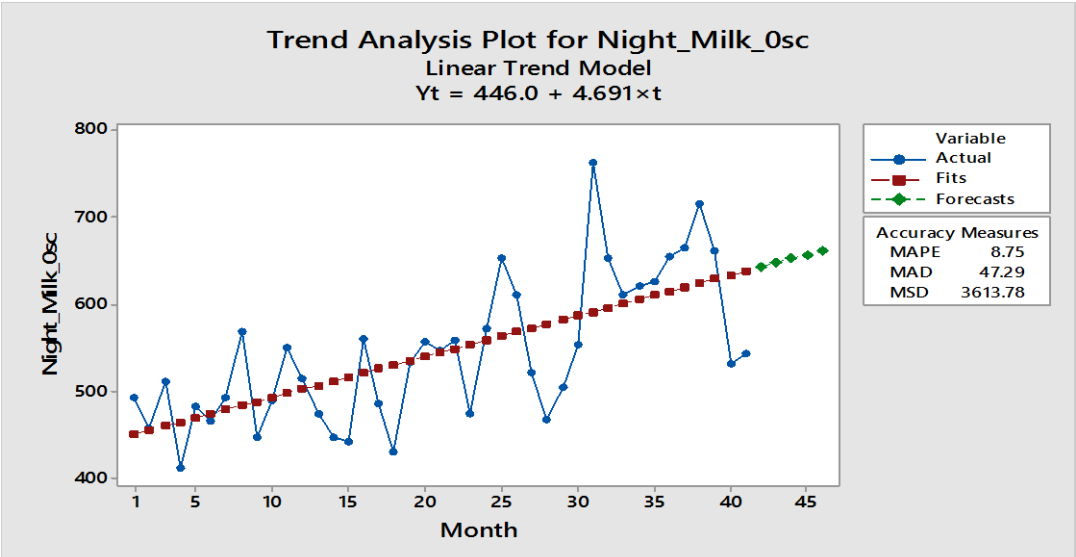
Partial Autocorrelation Function for Motning milk Residuals



Now we check normality for resulting series by using normal probability plot as follows,



Here the P-value is **0.199**; therefore the given series is satisfying the normality assumption.
And also, residuals has white noise with mean approxiately zero and variance 60.86.
Therefore by using Trend analysis we predict the next five values.

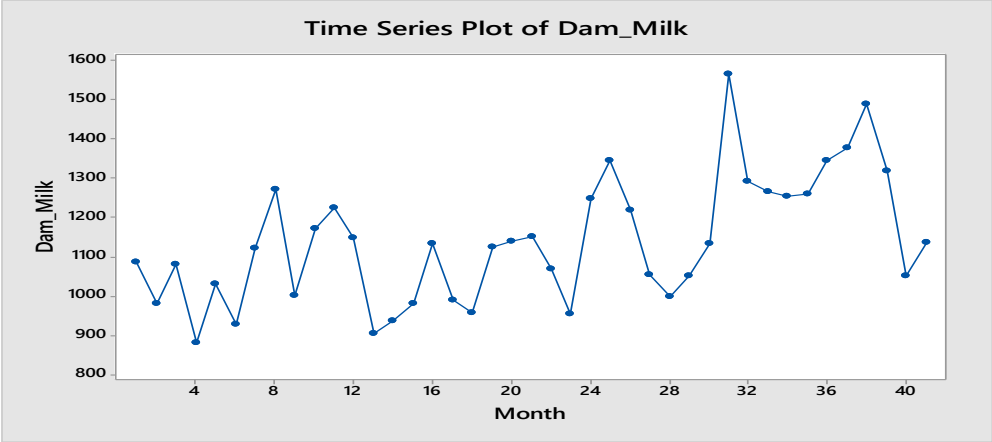


Forecasts:

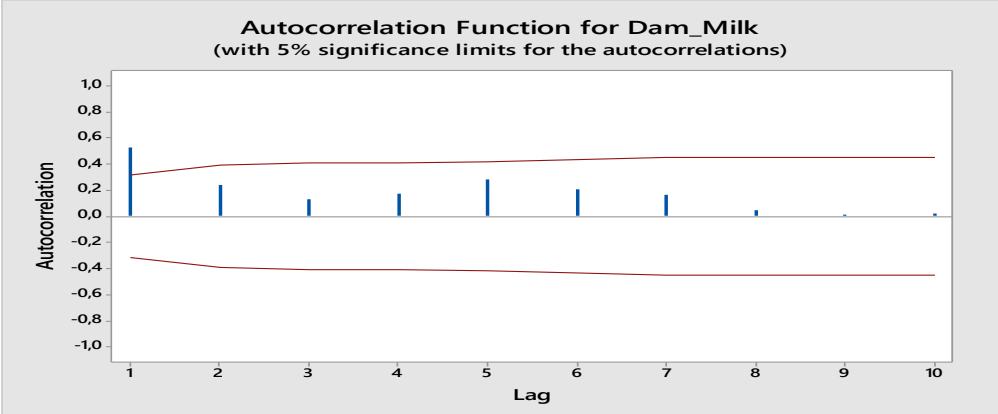
Period	Forecast
42	643.031
43	647.722
44	652.414
45	657.105
46	661.797

8.3.3 To forecast the Dam milk production of Osmanabadi Breed for next 5 days :

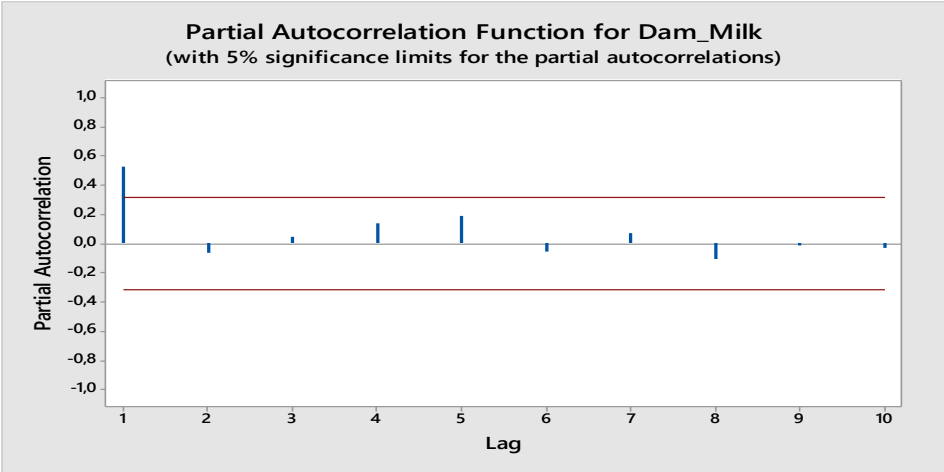
Time series plot for Dam milk



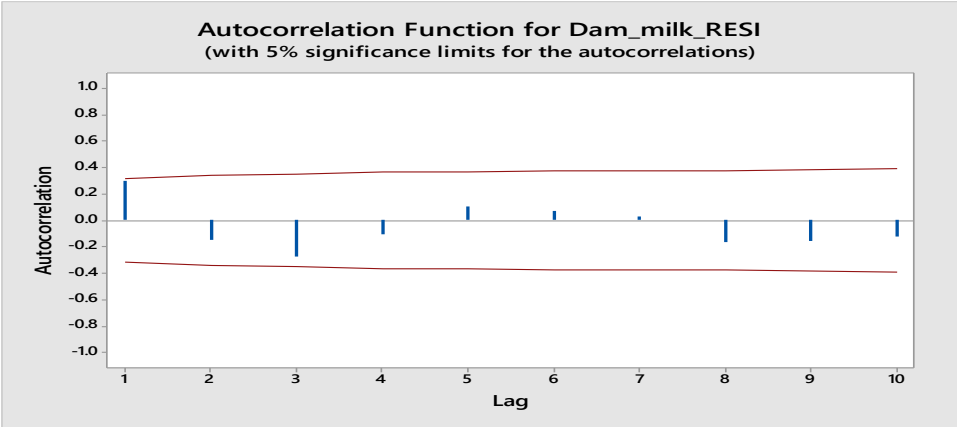
Autocorrelation Fumction for Dam milk



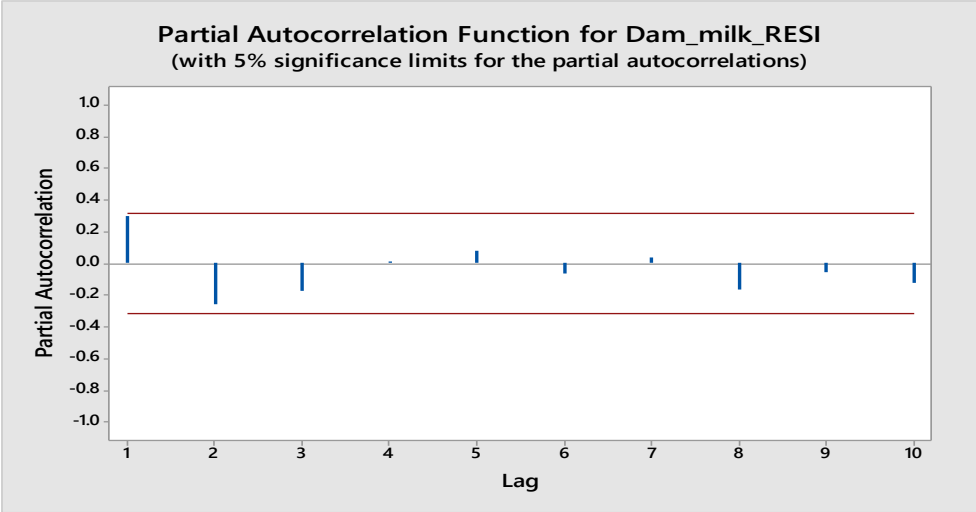
Partial Autocorrelation Fumction for Dam milk



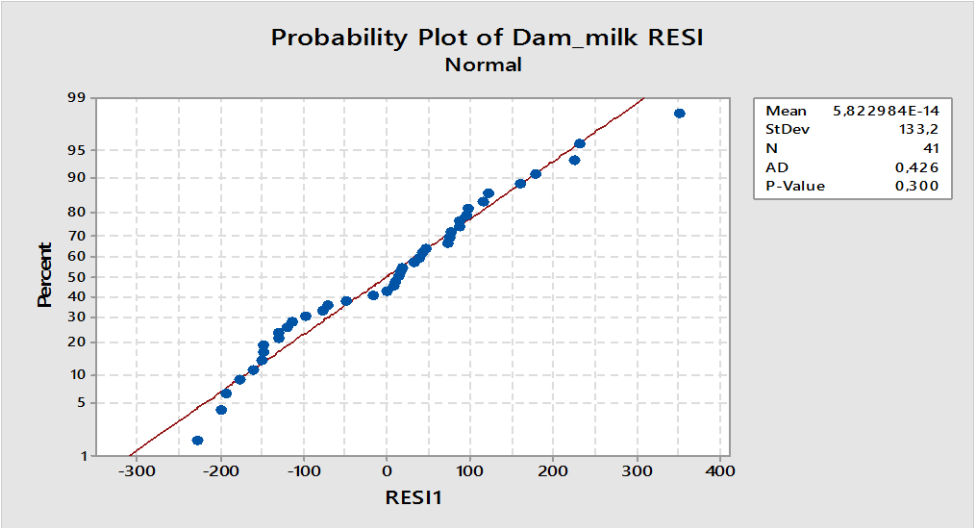
Autocorrelation Fumction for Dam milk Residuals



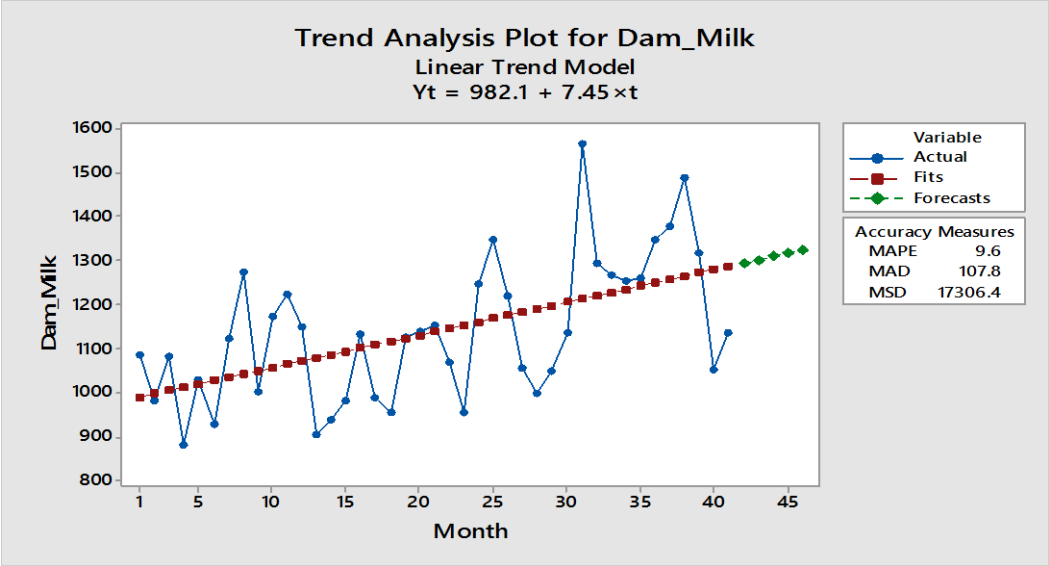
Partial Autocorrelation Fumction for Dam milk Residuals



Now we check normality for resulting series by using normal probability plot as follows,



Here the P-value is **0.3**; therefore the given series is satisfying thenormality assumption.
And also, residuals has white noise with mean approxiately zero and variance 132.2
Therefore by using Trend analysis we predict the next five values.



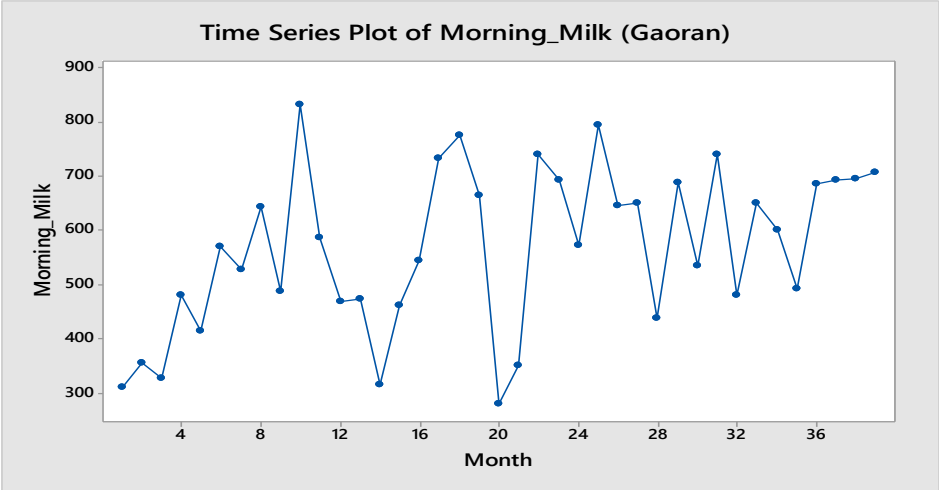
Forecasts:

Period	Forecast
42	1294.83
43	1302.28
44	1309.72
45	1317.17
46	1324.61

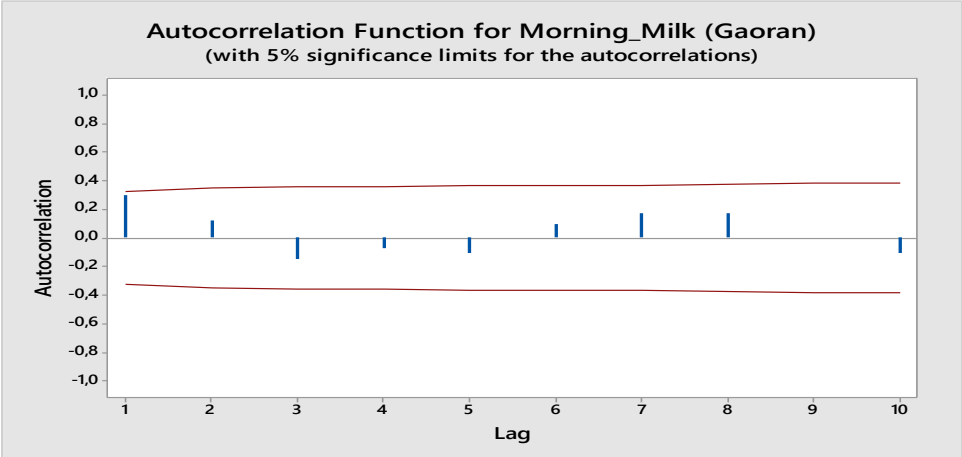
b) GaoranBreed :-

8.3.4 To forecast the morning Dam milk production of Gaoran breed for next 5 days.

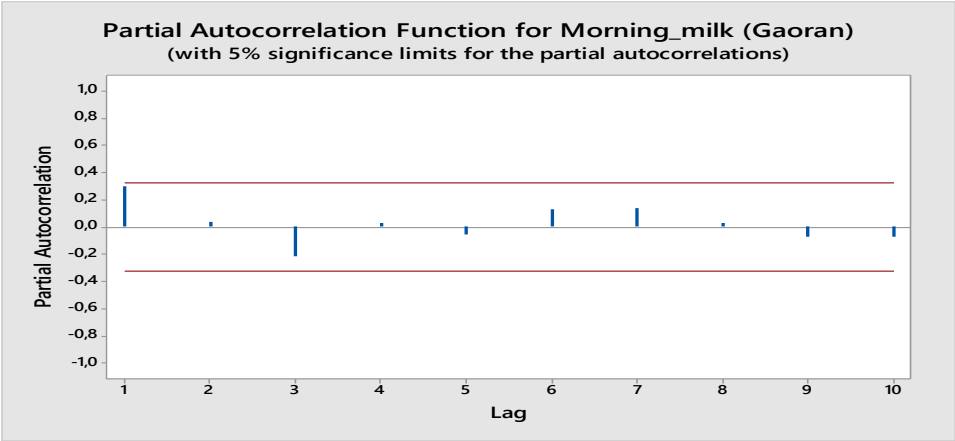
Time Series plot for Morning milk of Gaoran Goat:



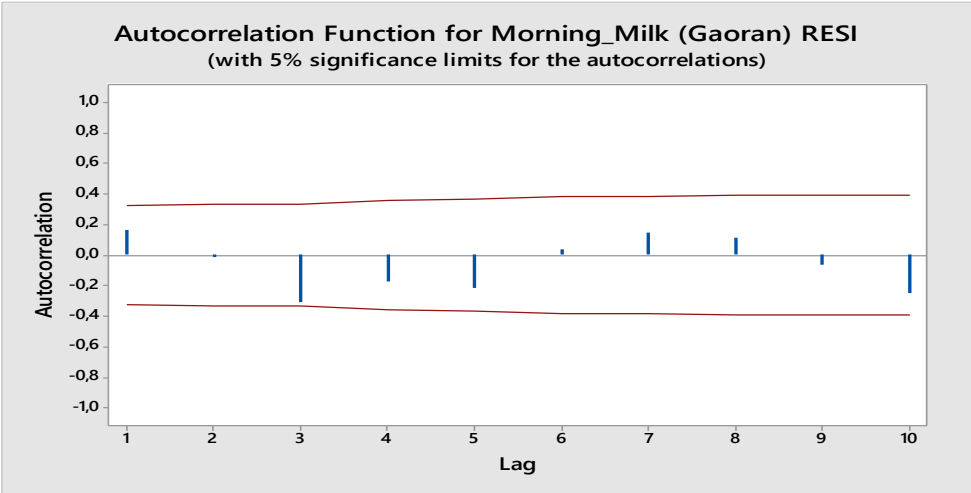
Autocorrelation Function for morning milk



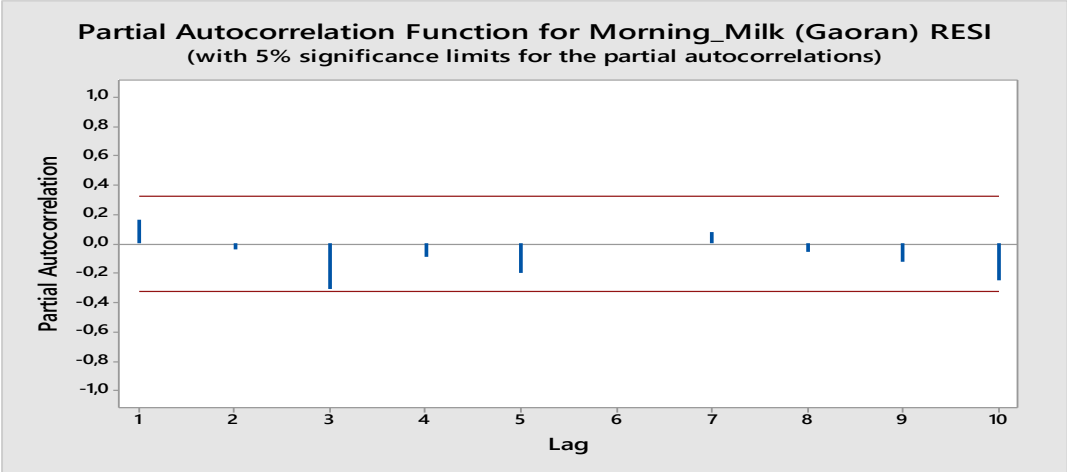
Partial Autocorrelation Function for morning milk



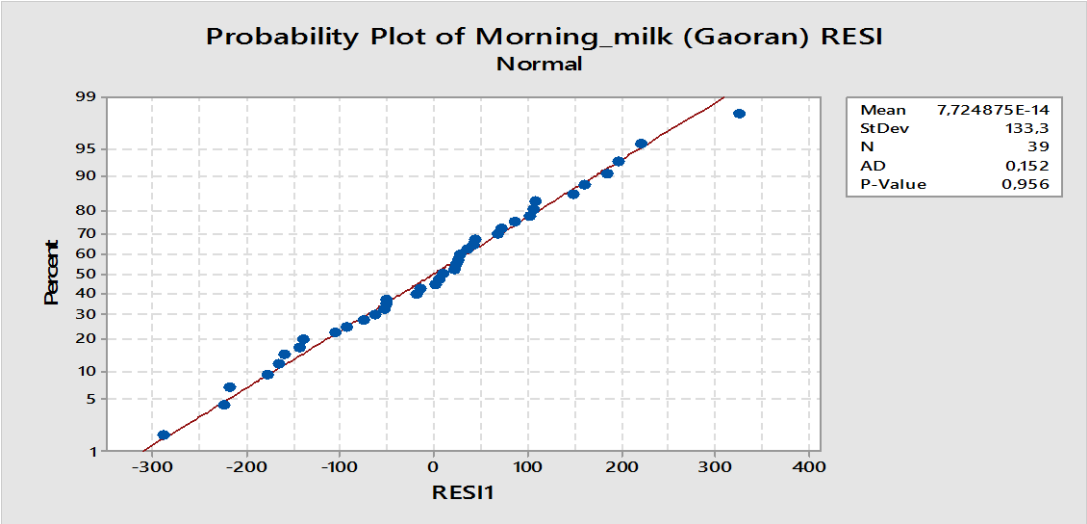
Autocorrelation Function for morning milk Residuals



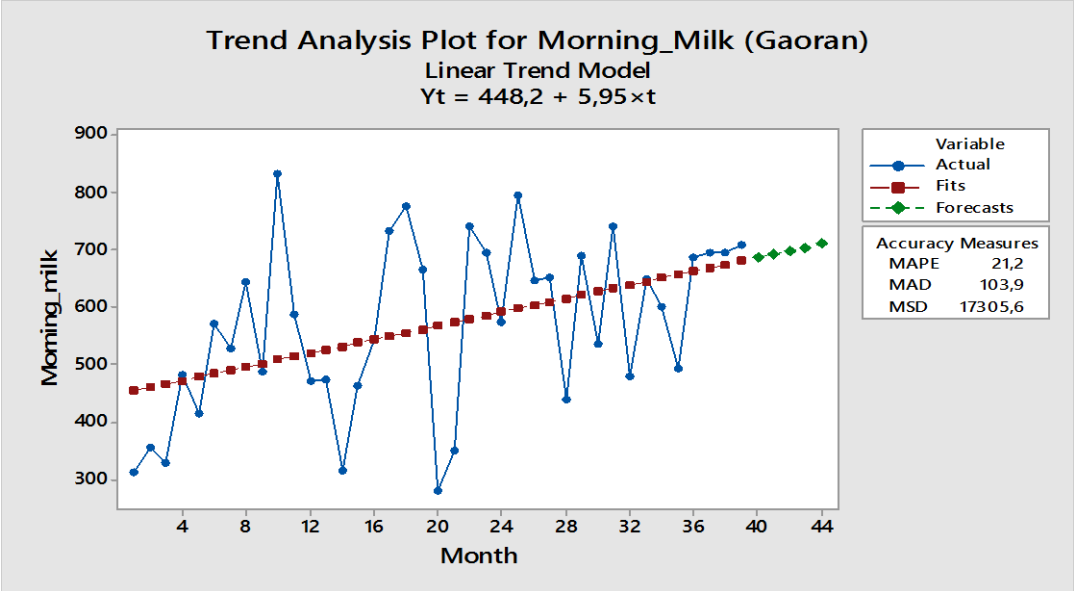
Partial Autocorrelation Function for morning milk Residuals



Now we check normality for resulting series by using normal probability plot as follows,



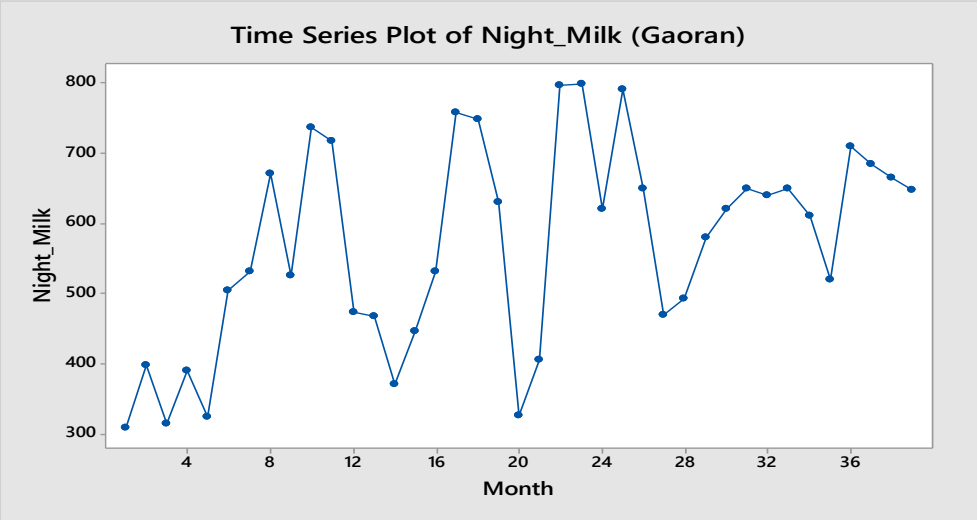
Here the P-value is **0.956**; therefore the given series is satisfying the normality assumption.
And also, residuals has white noise with mean approxiately zero and variance 133.3.
Therefore by using Trend analysis we predict the next five values.



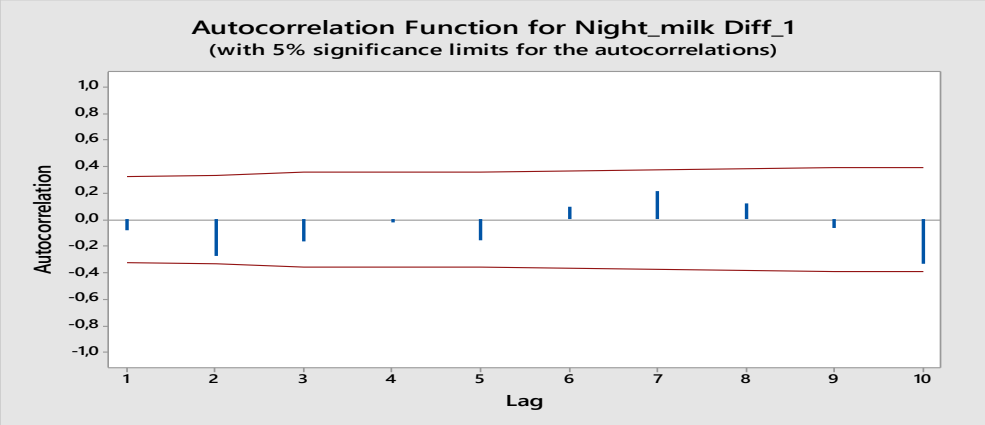
Period	Forecast
40	686.359
41	692.313
42	698.266
43	704.22
44	710.173

8.3.5 To forecast the Night Dam milk production of Gaoran breed for next 5 days.

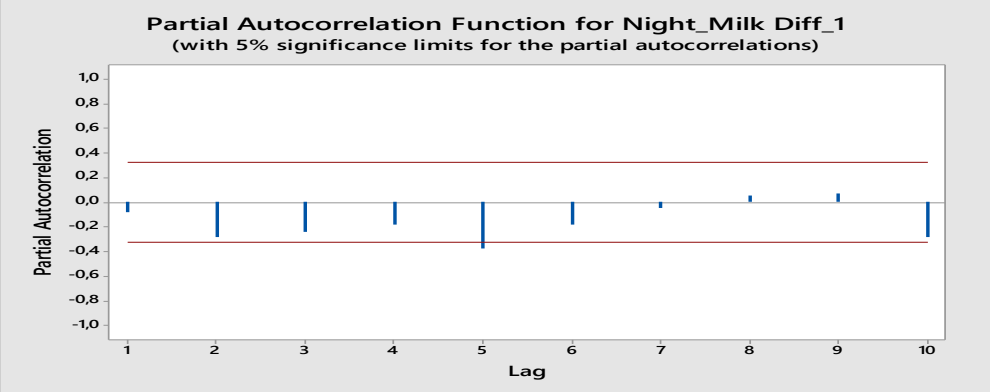
Time series Plot for Night milk of Gaoran Goats



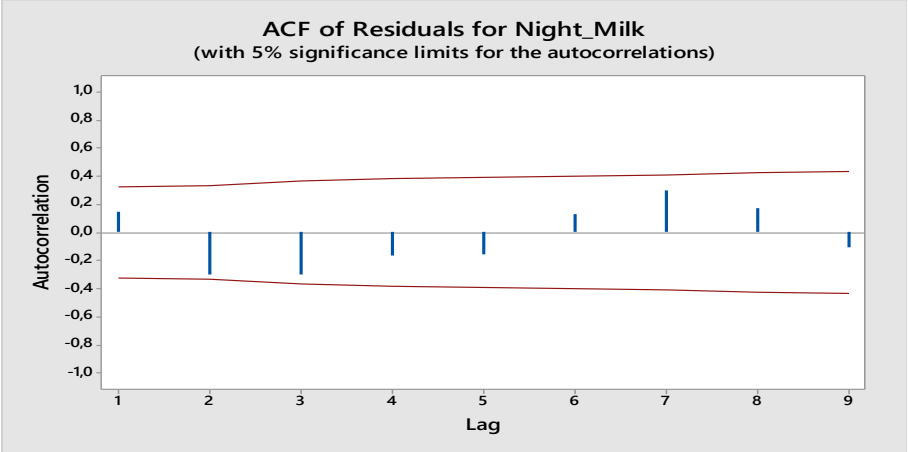
Autocorrelation Function for Night milk Diff_1



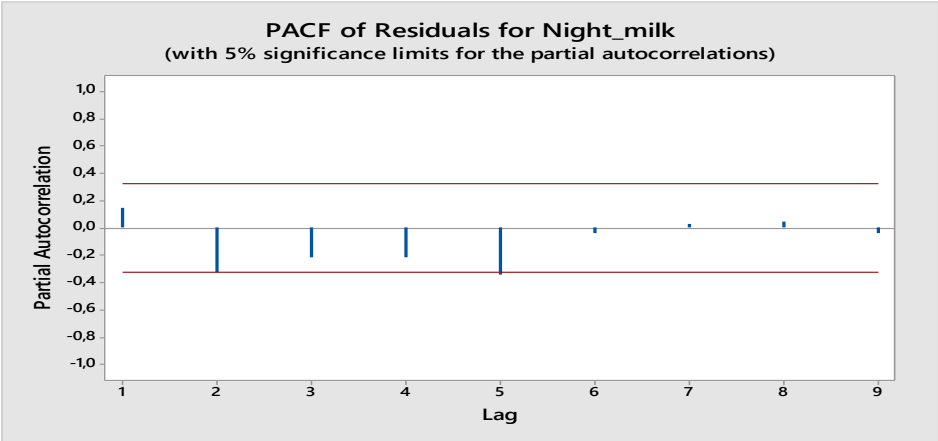
Partial Autocorrelation Function for night milk



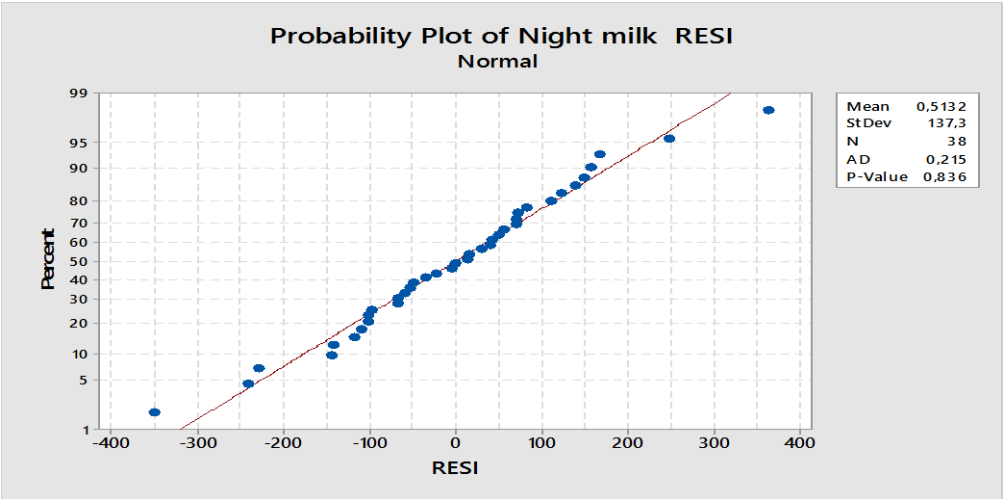
Autocorrelation Function for night milk Residuals



Partial Autocorrelation Function for night milk



Now we check normality for resulting series by using normal probability plot as follows,



Here the P-value is **0.836**; therefore the given series is satisfying the normality assumption.

And also, residuals has white noise with mean approxiately zero and variance 137.2.

The fitted model is ARIMA(0,1,1) which is given as below,

$$X_t = 0.3565Z_{t-1} + 8.71$$

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

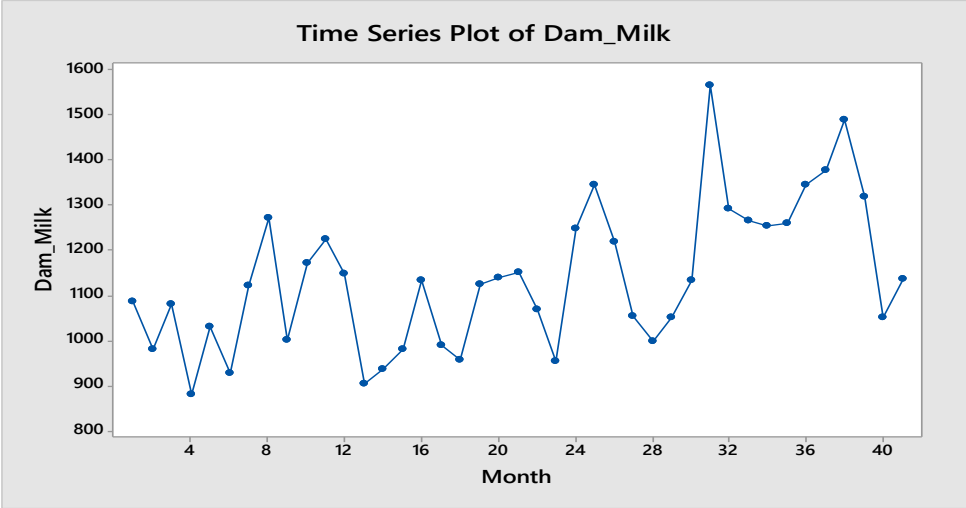
Lag	12	24	36
Chi-Square	23,9	34	35,2
DF	10	22	34
P-Value	0.08	049	0.412

Forecasts:

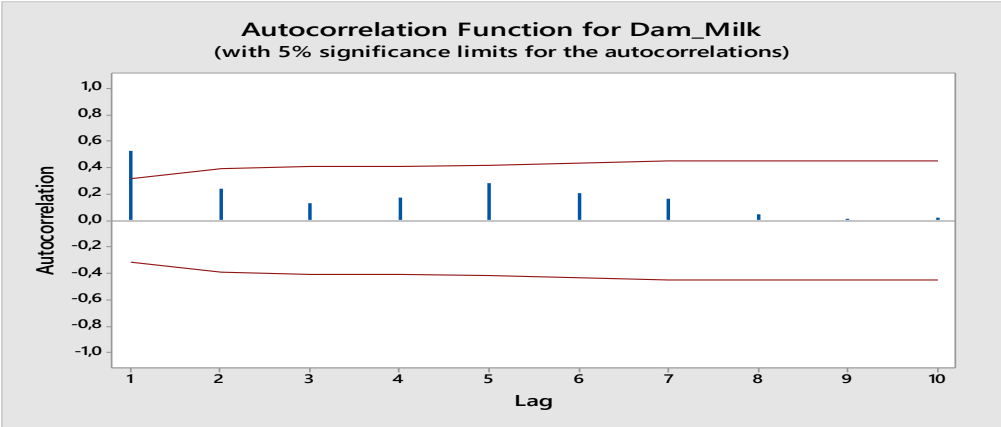
Period	Forecast
40	667.76
41	676.47
42	685.18
43	693.9
44	702.61

8.3.6 To forecast the Dam milk production of Gaoran breed for next 5 days.

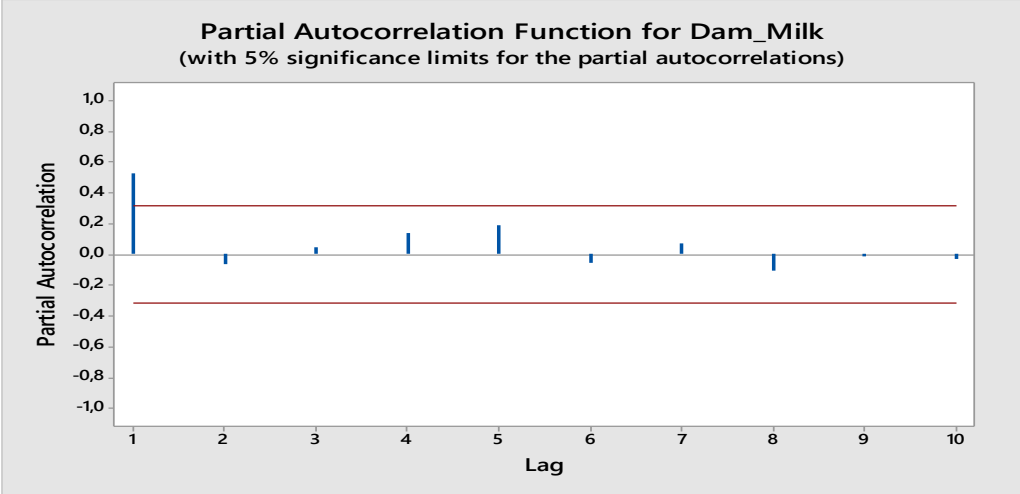
Time series plot of Dam milk of Gaoran goats



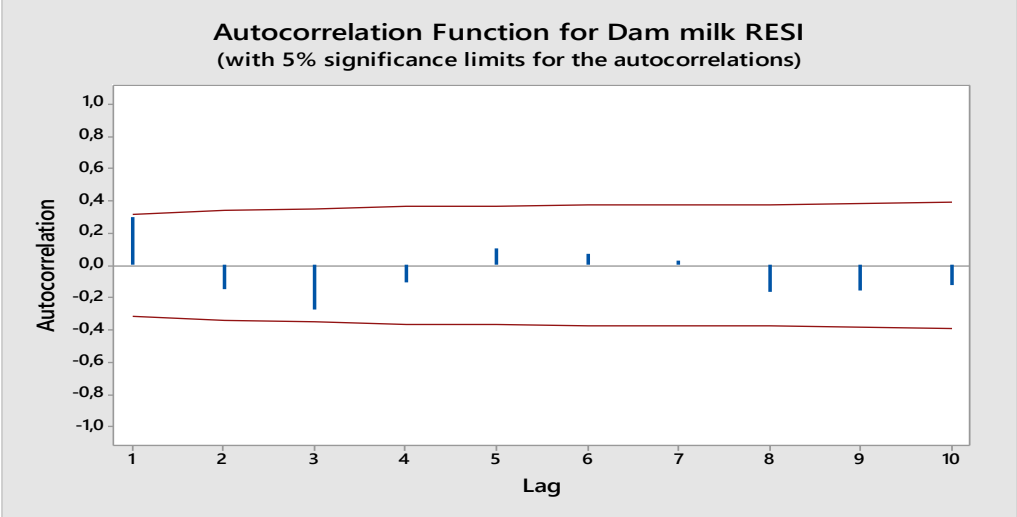
Autocorrelation Function for Dam milk



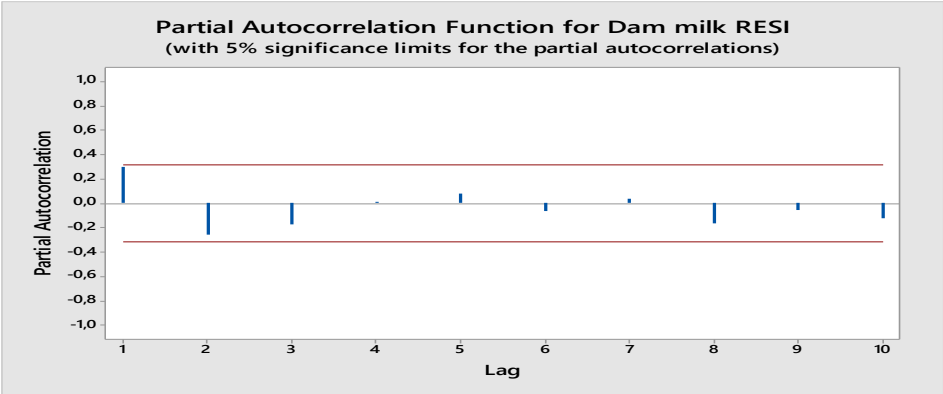
Partial Autocorrelation Function for Dam milk



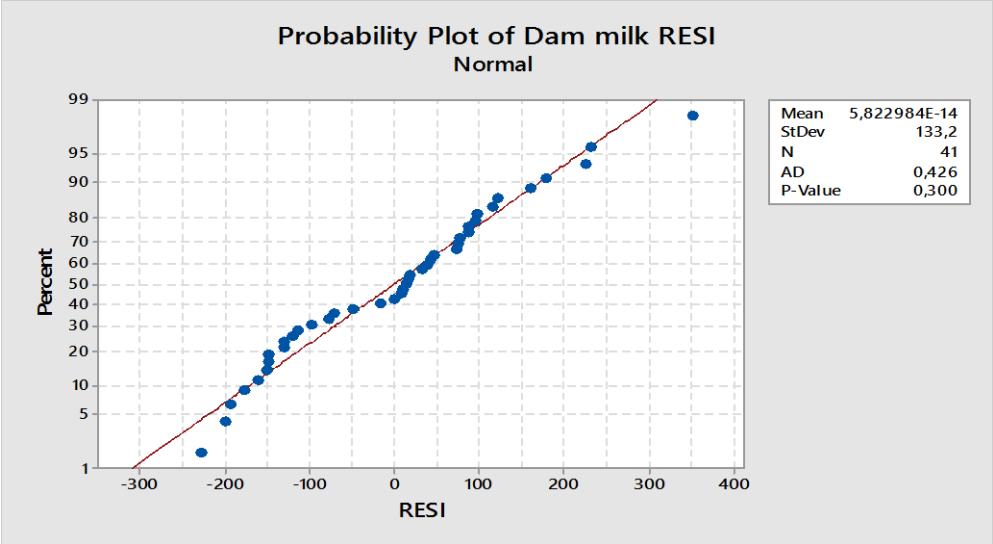
Autocorrelation Function for Dam milk Residuals



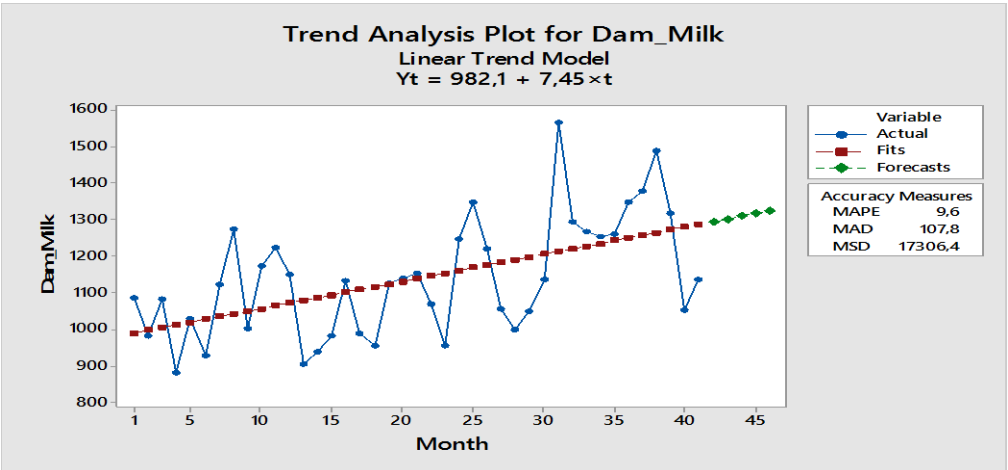
Partial Autocorrelation Function for Dam milk Residuals



Now we check normality for resulting series by using normal probability plot as follows,



Here the P-value is **0.3**; therefore the given series is satisfying the normality assumption.
And also, residuals has white noise with mean approxiately zero and variance 133.2.
Therefore by using Trend analysis we predict the next five values.



Forecats:

Period	Forecast
42	1294.83
43	1302.28
44	1309.72
45	1317.17
46	1324.61

9.Major Findings :-

1. From the above analysis we can say that there is high chance of giving birth to two kids at time.
2. Also birth season has significant impact on number born i.e. winter and rainy season is good for number born as compare to summer.
3. Also we can say that number born has significant impact on birth weight and weaning weight of kid. i.e. as number born increases both weights decreases.
4. Breed has no significant impact on body weight of kids.
5. Breed, weight, Age at weight and Dam parity has significant impact on Number born.
6. Type of birth (Number born) ,OwID group and city are more effective to the Weight of kids.

10.Limitations:-

1) Body weight analysis techniques are used only for NARI organization not on other organization.

2) Body weight analysis is used only for Age group 0 to 30 days kids of goat.

3) Focasting techniques is used only for NARI organization.

4) Forecasts are not very close to actual value.

5) Forecasting is under the assumptions that the future time is independent on any externalities.

Scope :

In similar manner we can study the body weight of kids from other area. We can do the region wise analysis of other types of goats as Damaskas, Boar etc. and study the effect of Dams wt, sex , wt age (in days),type of birth and Dam parity on body weight of these goats.And In similar manner we can forecast the dam milk productionfrom other area. We can do the region wise other types of goats as Damaskas, Boar, sirohi etc.

Refferences :

- Effect of Season on Kidding and Birth Weight in Osmanabadi Goats Reared in an Organized Farm
  Ch. Harikrishna, T. Raghunandan and M. Gnana Prakash
- Data Mining: Concepts and the Techniques. MorganGaufmann.
  Han, J. and Kamber, M. and Pei, J. (2012)
- Categorical Data Analysis. Wiley, New York:
  Agesti A. (1990)
- Time Series Theory and Methods:
  Peter J.Brockwell and Richard A. Davis

Apendix:

Ordinal logistic Regression :

#Variable Selection R Code:

```
s=read.csv("H:/Shweta project/data.csv",header=T)
s
names(s)
y=as.ordered(s$No_born)
s1=s[,-9]
library(MASS)
m=polr(y~.,data = s1)
summary(m)
```

#Model Selection : we use this model for analysis

```
step(m,direction ="both")    # Backword , Forword


#ordinal logistic model for variable selction model:
model=polr(y~ + Breed + Weight..KG. + Wt.Age + Dam.Parity,s1,Hess=TRUE)
summary(model)

# Calculate and store p values:
(ctable=coef(summary(model)))
p=pnorm(abs(ctable[,"t value"]),lower.tail = FALSE)*2
(ctable=cbind(ctable,"p value"=p))

#confidence interval for parameter estimate :
exp(cbind(coef(model),confint(model)))


#odds ratio:
e=exp(cbind(OR=coef(model),confint(model)))
e

#prediction
data=s[,c(-1,-2,-3,-4)]
data
names(data)
data$NO_born=as.ordered(data$No_born)
ind=sample(2,nrow(s),replace = TRUE,prob = c(0.7,0.3))
train=data[ind==1,]
test=data[ind==2,]
testd=data[ind==2,-5]

model1=polr(No_born~ Breed + Weight..KG. + Wt.Age +
Dam.Parity,train,Hess=TRUE)
pred=predict(model1,testd)
pred
t=table(pred,test$No_born)
t
miscler=1-sum(diag(t))/sum(t)
miscler
```

SAS Code :

```
proc means data = Shweta.data;
var Weight__KG_;
run;
proc print data=Shweta.data1;
run;

*/ ordinal logistic regression*/
proc logistic data=Shweta.data desc;
class Breed(ref='1')  Dam_parity(ref='1')/ param = reference;
model No_born= Breed Weight__KG_ Wt_Age Dam_Parity;
run;

*/Estimate probabilities*/
estimate "Pr prob No_born=4 at Breed=1" intercept 1 Weight__KG_ 1 Wt_Age 1
Dam_Parity  3.134 / ilink category='4';
estimate "Pr prob No_born=4 at Breed=2" intercept 1 Breed 1 Weight__KG_ 1
Wt_Age 1 Dam_Parity  3.134 / ilink category='4';
estimate "Pr prob No_born=4 at Breed=3" intercept 1 Breed 1 Weight__KG_ 1
Wt_Age 1 Dam_Parity  3.134 / ilink category='4';
estimate "Pr prob No_born=2 or 3 or 4 at Breed=1" intercept  1 Weight__KG_
1 Wt_Age 1 Dam_Parity  3.134 / ilink category='2';
estimate "Pr prob No_born=2 or 3 or 4 at Breed=2" intercept 1  Breed 1
Weight__KG_ 1 Wt_Age 1 Dam_Parity  3.134 / ilink category='2';
estimate "Pr prob No_born=2 or 3 or 4 at Breed=3" intercept 1  Breed 1
Weight__KG_ 1 Wt_Age 1 Dam_Parity  3.134 / ilink category='2';
run;
```

Artificial Neural Network:

R Code :

```
s=read.csv("F:/Shweta project/data.csv",header=T)

s

maxs=apply(s,2,max)

mins=apply(s,2,min)

scaled=as.data.frame(scale(s,center=mins,scale=maxs-mins))

train=scaled[1:897,]

test1=scaled[898:1282,]

testd=scaled[898:1282,-6]

library(neuralnet)

n=names(train)
```

```
f=as.formula(paste("Weight..KG.~",paste(n[!n%in%"Weight..KG."],collapse="+")))

nn=neuralnet(f,data=train,hidden=1,linear.output=T)

nn

summary(nn)

plot(nn)

pnn=compute(nn,testd)

pnn1=pnn$net.result*(max(s$Weight..KG.)-min(s$Weight..KG.))+min(s$Weight..KG.)

test.r=(testd)*(max(s$Weight..KG.)-min(s$Weight..KG.))+min(s$Weight..KG.)

mse.nn=sum((test.r-pnn1)^2)/nrow(test)
```