Growth Curves of Lambs: Impact of Litter Size and Gender

Shrimani Tundurwar

July 2024

1 Nadaraya-Watson Estimated Curve for Lamb 612

The Nadaraya-Watson estimated curve for Lamb 612 is shown below in the figure.

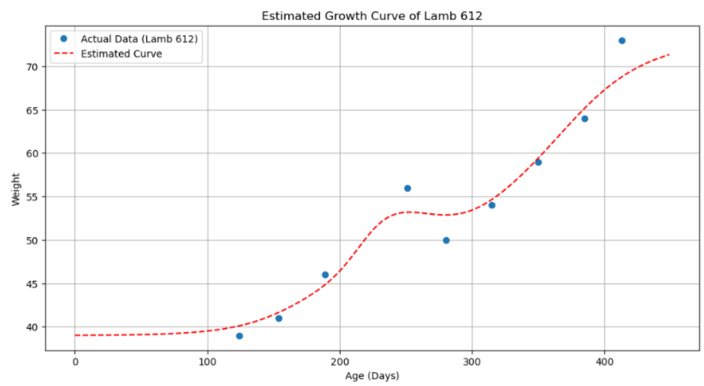

Figure 1: Nadaraya-Watson Estimated Curve for Lamb 612

Here we can see the blue points are the actual data points while the red line represents the Nadaraya-Watson estimated curve. Based on the figure, we can say that the estimated curve follows the trend successfully.

Till now, we have seen that the Nadaraya-Watson estimator has helped us to create a growth curve of lambs without any dependent variables. As this method is a non-parametric regression method, it doesn't depend on any variable; rather, it generates the curve based on the data points themselves.

Now, here we want to see the impact of different variables such as litter size and gender on the growth curve. Let's first move forward using the Ordinary Least Squares method to overcome the dependency issue.

2 Local Linear Regression

Local linear regression is a method that fits a linear model to a subset of the data near the point of interest. This approach combines the flexibility of non-parametric methods with the interpretability of linear models.

2.1 Define the Local Linear Model

For a point x, fit a linear model to the points in its neighborhood:

$$y_i \approx \beta_0 + \beta_1(x_i - x)$$

2.2 Minimize the Weighted Least Squares Error

$$\text{minimize} \sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right)(y_i - \beta_0 - \beta_1(x_i - x))^2$$

2.3 OLS Equation for Local Linear Regression

For each point x, we need to solve the following weighted least squares problem.

2.3.1 Weighted Matrix and Vectors

$$W(x) = \text{diag}\left(K\left(\frac{x_1 - x}{h}\right) K\left(\frac{x_2 - x}{h}\right) \ldots K\left(\frac{x_n - x}{h}\right)\right)$$

$$X = \begin{pmatrix} 1 & (x_1 - x) \\ 1 & (x_2 - x) \\ \vdots & \vdots \\ 1 & (x_n - x) \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

2.3.2 Parameter Estimation

$$\hat{\beta}(x) = (X^T W(x) X)^{-1} X^T W(x) Y$$

where

$$\hat{\beta}(x) = \begin{pmatrix} \hat{\beta}_0(x) \\ \hat{\beta}_1(x) \end{pmatrix}$$

2.3.3 Prediction

$$\hat{y}(x) = \hat{\beta}_0(x)$$

Bandwidth Taken = 20

Based on the Local Linear Model and minimizing the error, we can say that the growth curve is better than the Nadaraya-Watson estimation technique. We are transforming the data to add the litter size and gender column to analyze the impact on the growth curves.

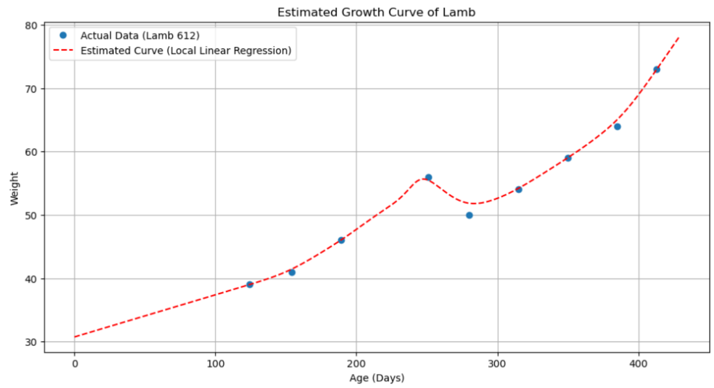

Figure 2: OLS Estimated Curve for Lamb 612

Transformed Data: We have now added the columns *Litter_Size*, *Gender*, and *Type of lamb* to the data. The next step is to analyze the impact of these new variables on the growth curve and compare how this might affect the outcomes.

	No.	Age (Days)	Weight	Litter_size	Gender	Type
0	611	124	42.0	1	F	pedigree
1	611	154	43.0	1	F	pedigree
2	611	189	49.0	1	F	pedigree
3	611	251	NaN	1	F	pedigree
4	611	280	NaN	1	F	pedigree
5	611	315	NaN	1	F	pedigree
6	611	350	NaN	1	F	pedigree
7	611	385	NaN	1	F	pedigree
8	611	413	NaN	1	F	pedigree
9	612	124	39.0	2	F	pedigree
10	612	154	41.0	2	F	pedigree
11	612	189	46.0	2	F	pedigree
12	612	251	56.0	2	F	pedigree
13	612	280	50.0	2	F	pedigree
14	612	315	54.0	2	F	pedigree
15	612	350	59.0	2	F	pedigree
16	612	385	64.0	2	F	pedigree
17	612	413	73.0	2	F	pedigree

Figure 3: Transformed Data with Litter Size, Gender, and Type of Lamb

3 Multivariate Regression Model

To measure the impact of additional factors such as litter size or gender on the growth curve of lambs, multivariate regression models are more appropriate than a simple polynomial regression model. Multivariate regression allows you to include multiple independent variables (predictors) to explain the dependent variable (weight).

For this model, we are excluding types of lambs and only considering the factors of litter size and gender. (Polynomial 3rd Order)

```
Mean Squared Error: 26.35007255539101
R-squared: 0.7984421922846409
```

```
coefficients = pd.DataFrame(model.coef_, X.columns, columns=['Coefficient'])
print(coefficients)
```

```
            Coefficient
Age (Days)     0.104542
Litter_size   -1.537061
Gender_M       6.458930
```

Figure 4: Transformed Data with Litter Size, Gender, and Type of Lamb

3.1 Model Outcome

Based on the outcome, we can say that the model has an R^2 value of 0.7984, which means that approximately 79.84% of the variance in lamb weights can be explained by the model. This is a relatively high value, indicating a good fit.

Weight increases as age increases is indicated by the positive 0.1 coefficient value. For each additional lamb in the litter, the weight of a lamb is expected to decrease by approximately 1.537061 units, assuming other variables (age and gender) remain constant. This negative coefficient suggests that lambs from larger litters tend to weigh less.

Male lambs (indicated by *Gender_M*) are expected to weigh approximately 6.458930 units more than female lambs, assuming other variables (age and litter size) remain constant.

3.2 Comparing with Actual Data Points

Lamb no. 629: Litter Size = 3, Gender = Female

This lamb was taken from the 2018 data having sex as female and was a part of a triplet. Choosing this lamb will give us the idea of how the growth curve predicts for a triplet lamb and with gender as female.

Lamb no. 628: Litter Size = 3, Gender = Male

Lamb no. 691: Litter Size = 1, Gender = Female

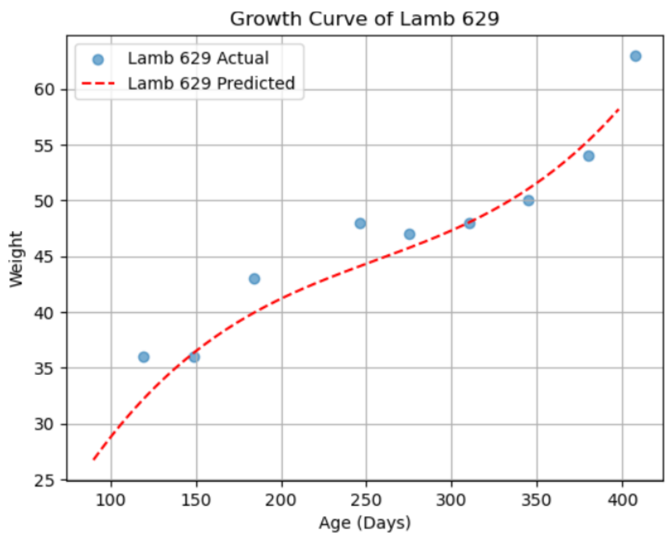

Figure 5: Comparison of lamb 629 with Actual Data Points

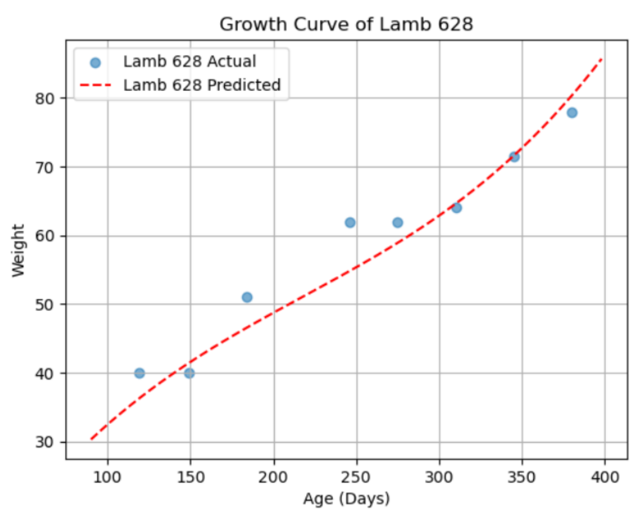

Figure 6: Comparison of lamb 628 with Actual Data Points

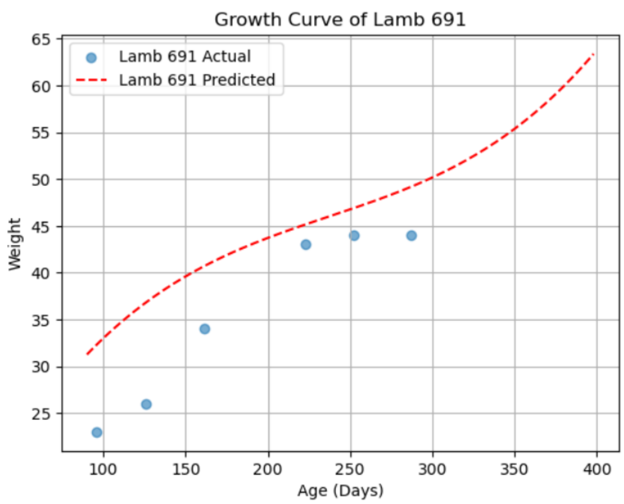

Figure 7: Comparison of lamb 691 with Actual Data Points