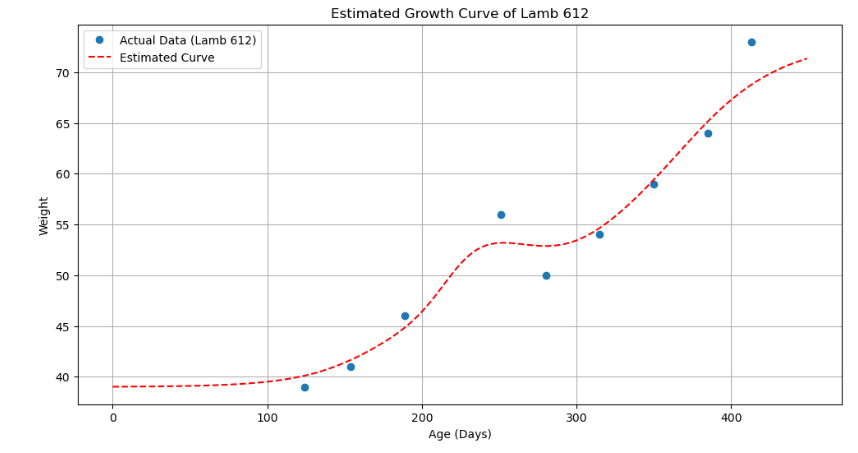
**Report 4**

Nadraya-Watson Estimated curve for Lamb 612 is shown below in the figure.



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 non- parametric regression method this doesn’t depend on any variable rather it generates the curve based on the data points itself.

Now, here we want to see the impact of different variables such as litter size, gender on the growth curve.

Let’s first move forward using the Ordinary Least Square method to overcome the dependency issue.

**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.

* Define **the Local Linear Model**: For a point x, fit a linear model to the points in its neighbourhood:

yi≈β0+β1(xi−x)y\_i \approx \beta\_0 + \beta\_1 (x\_i - x)yi​≈β0​+β1​(xi​−x)

* Minimize **the Weighted Least Squares Error**:

minimize∑i=1nK(xi−xh)(yi−β0−β1(xi−x))2\text{minimize} \sum\_{i=1}^n K \left(\frac{x\_i - x}{h}\right) \left(y\_i - \beta\_0 - \beta\_1 (x\_i - x)\right)^2minimizei=1∑n​K(hxi​−x​)(yi​−β0​−β1​(xi​−x))2

### OLS Equation for Local Linear Regression

For each point xxx, we need to solve the following weighted least squares problem:

1. **Weighted Matrix and Vectors**:

W(x)=diag(K(x1−xh),K(x2−xh),…,K(xn−xh))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)W(x)=diag(K(hx1​−x​),K(hx2​−x​),…,K(hxn​−x​)) X=(1(x1−x)1(x2−x)⋮⋮1(xn−x)),Y=(y1y2⋮yn)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}X=​11⋮1​(x1​−x)(x2​−x)⋮(xn​−x)​​,Y=​y1​y2​⋮yn​​​

1. **Parameter Estimation**:

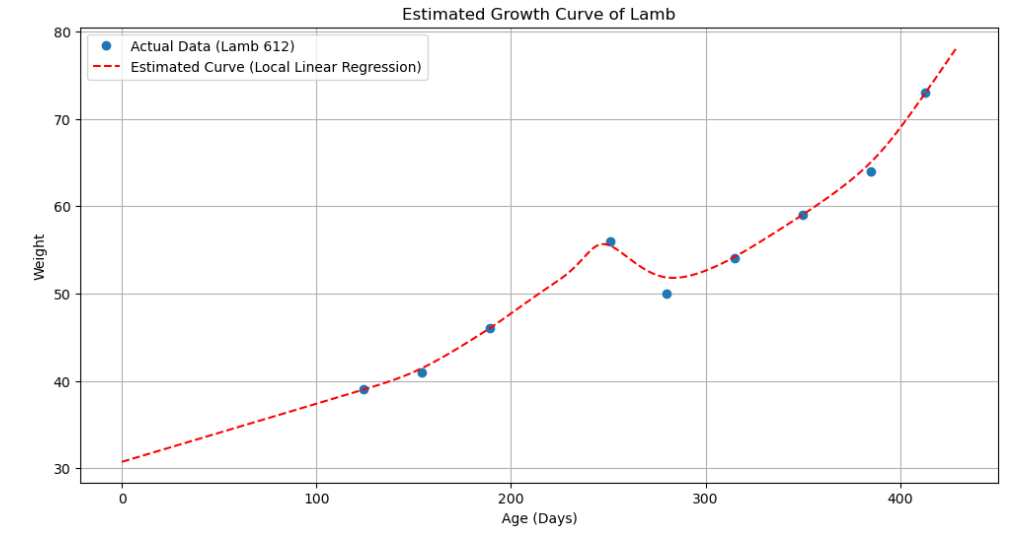
β^(x)=(XTW(x)X)−1XTW(x)Y\hat{\beta}(x) = (X^T W(x) X)^{-1} X^T W(x) Yβ^​(x)=(XTW(x)X)−1XTW(x)Y

where β^(x)=(β^0(x)β^1(x))\hat{\beta}(x) = \begin{pmatrix} \hat{\beta}\_0(x) \\ \hat{\beta}\_1(x) \end{pmatrix}β^​(x)=(β^​0​(x)β^​1​(x)​).

1. **Prediction**:

y^(x)=β^0(x)\hat{y}(x) = \hat{\beta}\_0(x)y^​(x)=β^​0​(x)

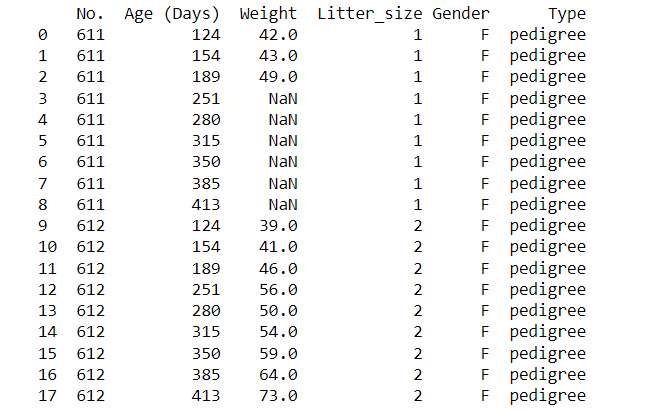
Bandwidth Takes = 20



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 analyse the impact on the growth curves.

Now this is how the data looks like:

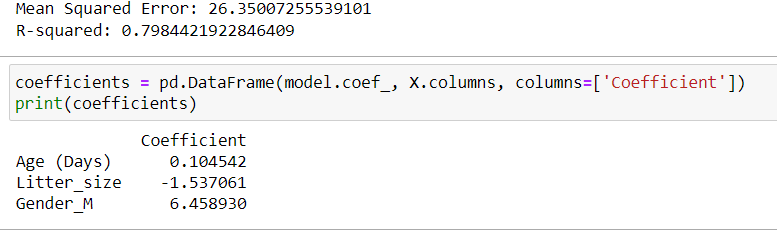


We have now added the columns Litter\_Size, Gender and Type of lamb to the data. Next step is to analyse the impact of these new variables to the growth curve and compare how this might affect the outcomes.

**Here we are going to use the 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**)



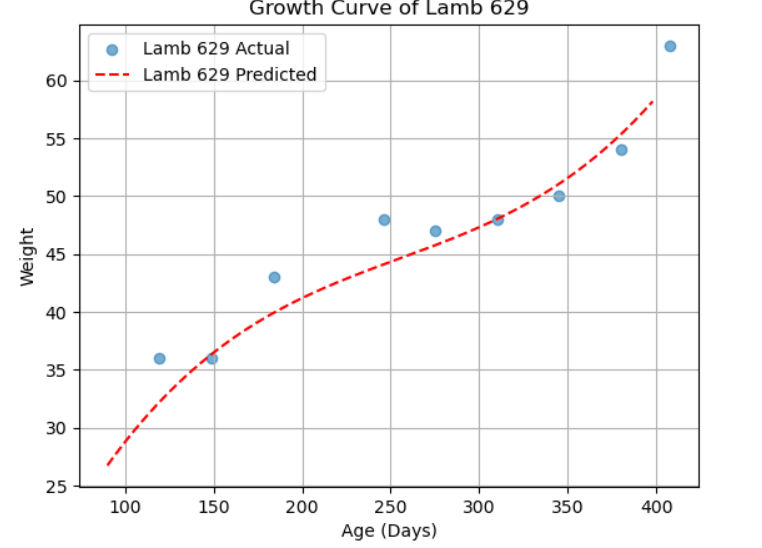
Based on the outcome we can say that the model An R² value of 0.7984 means that approximately 79.84% of the variance in lamb weights can be explained by the model. This is a relatively high value, indicating a good fit.

Weight increases as age increases is indicated by the positive 0.1 coefficient value.

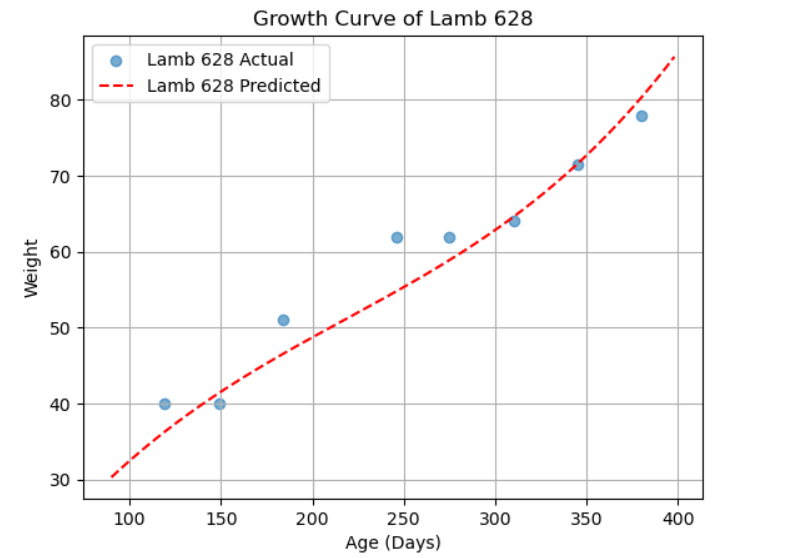
* For each additional lamb in the litter, the weight of a lamb is expected to decrease by approximately 1.537061 units, assuming other variables (age and gender) remain constant. This negative coefficient suggests that lambs from larger litters tend to weigh less.
* Male lambs (indicated by Gender\_M) are expected to weigh approximately 6.458930 unit’s more than female lambs, assuming other variables (age and litter size) remain constant.

Comparing with the actual data points:

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



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



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

