Growth Analysis of Lambs: Report

Shrimani Tundurwar

July 2024

Data Cleaning

Column name with dates needs to be first converted to date format.

```
In [40]: df.describe()
         df.dtypes

Out[40]: No.                          int64
         Sex                         object
         D.O.B                       object
         Dam                         object
         Sire                         int64
         Weaning Weight               int64
         Slaughter/Sale Date         object
         DW                         float64
         LW                         float64
         Sold as Store or Breeding   object
         3/7/2018                     int64
         2/8/2018                   float64
         6/9/2018                   float64
         7/11/2018                  float64
         6/12/2018                  float64
         10/1/2019                  float64
         14/2/2019                  float64
         21/3/2019                  float64
         18/4/2019                  float64
         type                        object
         dtype: object
```

Figure 1: Data Description of lamb2018 data columns.

Similarly the values in "D.O.B" column are objects, we first need to convert it to string for concatenating the respective year.

We can see from the "figure 1" values doesn't contain the year and we add this by concatenating the year value. For e.g.: 2018.

Now converting the values to date format for calculating the age of lambs by taking the difference between "D.O.B" and column headers.

D.O.B
1/3
1/3
1/3
2/3
2/3

Figure 2: Values in the D.O.B column are in day and month format.

0.1 Code

```
# Convert 'D.O.B' to datetime

df_2018test['D.O.B'] = df_2018test['D.O.B'].astype(str) #+ '/2016'
df_2018test['D.O.B'] = df_2018test['D.O.B'] + '/2018'


df_2018test['D.O.B'] = pd.to_datetime(df_2018test['D.O.B'], format='%d/%m/%Y')
```

Figure 3: Python code for data formatting.

Data Formatting

Now we require the data to be in the format where we get the lamb no. , age (in days), and weight of the lamb. For this we transpose the data for the respective rows.

The figure 3 shows the converted data after doing the transpose on each roe and calculating the age of each lamb.

	No.	Age (Days)	Weight
0	611	124	42.0
1	611	154	43.0
2	611	189	49.0
3	611	251	NaN
4	611	280	NaN
..	...	...	...
715	691	252	44.0
716	691	287	44.0
717	691	322	NaN
718	691	357	NaN
719	691	385	NaN

Figure 4: Data after formatting into Age in days.

Growth Curve Analysis

Now let's try to plot the growth curve of lambs from the data collected in 2018. There are 80 lambs in the dataset for year 2018. "Figure 3" shows the growth curve of all the lambs in the specified year.

Growth Curve of Lambs in 2018

A growth curve is plotted for all 80 lambs in the dataset for the year 2018. Figure 3 depicts the growth curve.

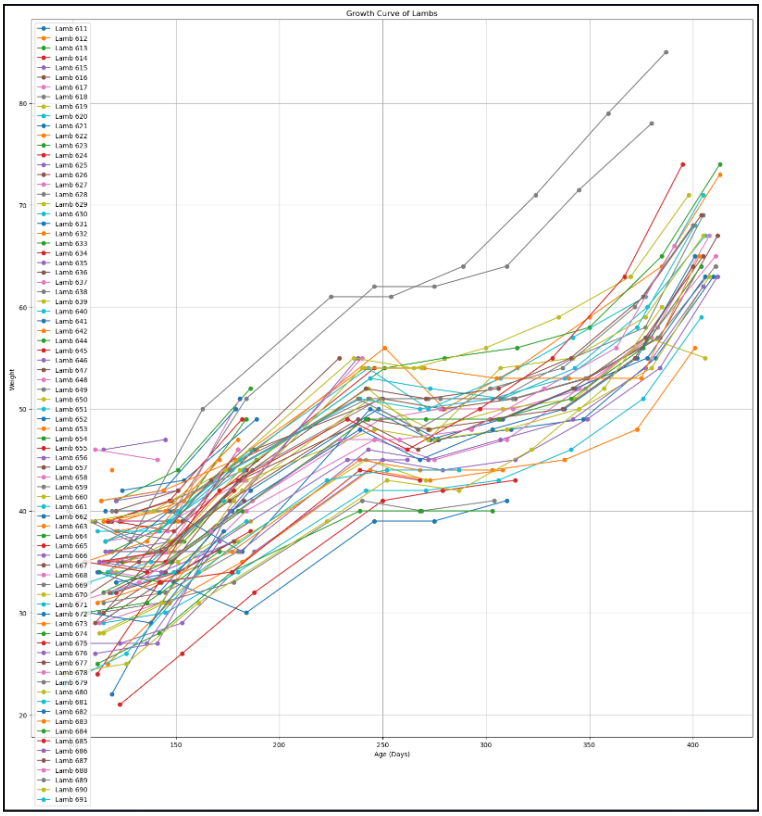

Figure 5: Growth curves of all the lambs in the dataset.

Comments on Growth Curve

- Some lamb data is missing for specific dates.
- The length of growth curves indicates the duration lambs spent on the farm.This shows that some lambs have smaller curve which represents that

either the lamb has been butchered or sold or died of natural/unnatural cause.

- Typical patterns in growth curves are observed, with initial weight increases and potential fluctuations after 250 days.

Nadaraya-Watson Estimator

In statistics, kernel regression is a non-parametric technique to estimate the conditional expectation of a random variable. The objective is to find a non-linear relation between a pair of random variables X and Y.

Testing with Lambs 611 and 612

Data for lambs 611 and 612 are selected for testing the Nadaraya-Watson estimator method. Figure 4 displays the data.

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

Figure 6: Sample data for the lamb no.611 and 612.

Growth Curve and Methodology

A growth curve is plotted for lambs 611 and 612, applying the Nadaraya-Watson method. Figure 5 shows the growth curve.

One of the two lambs in the dataset has fewer values registered as compared to the other. This hows that one lamb was either butchered or sold as pet/spare lamb.

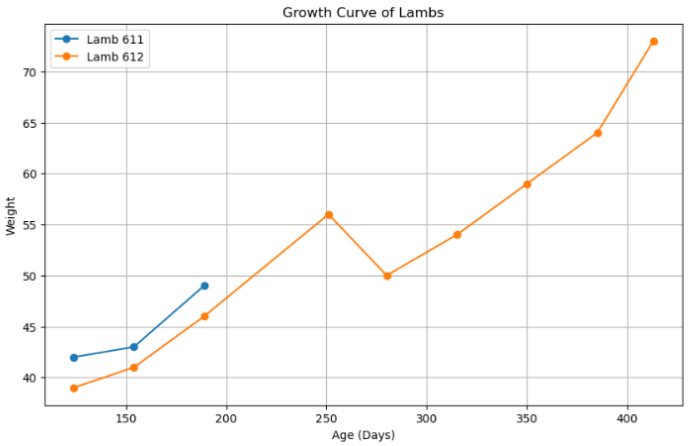

Figure 7: Plot shows the growth of two lambs 611 and 612.

Weight Estimation Methods

Gaussian Kernel Method

Weights are calculated using the Gaussian kernel regression method. Figure 6 illustrates the process. This uses distance between points and a bandwidth for smoothing of curve.

$$\hat{m}_h(x) = \frac{\sum_{i=1}^{n} K_h(x - x_i) y_i}{\sum_{i=1}^{n} K_h(x - x_i)}$$

where

$$K_h(t) = \frac{1}{h} K\left(\frac{t}{h}\right)$$

is a kernel with a bandwidth h such that $K(\cdot)$ is of order at least 1, that is

$$\int_{-\infty}^{\infty} uK(u)\, du = 0.$$

Estimated Weight Calculation

For instance, estimated weight of lamb 612 at 200 days is calculated and visually represented in Figure 8 and 9.

This comes out to be 46.42. Visually representing the estimated value with the original curve.

Inverse Exponential Weighting

An alternative weight calculation method using the inverse exponential function is applied. Results are compared.

```python
# Example: Estimating the weight of lamb 611 at age 200 days
lamb_no = 612
target_age = 200
bandwidth = 30  # Bandwidth Adjustment

estimated_weight = estimate_weight(lamb_no, target_age, bandwidth)
print(f"Estimated weight of lamb {lamb_no} at age {target_age} days: {estimated_weight}")
```

```
Estimated weight of lamb 612 at age 200 days: 46.420193907189756
```

Figure 8: Code for estimating the weight at the age of 200 days for lamb 612.

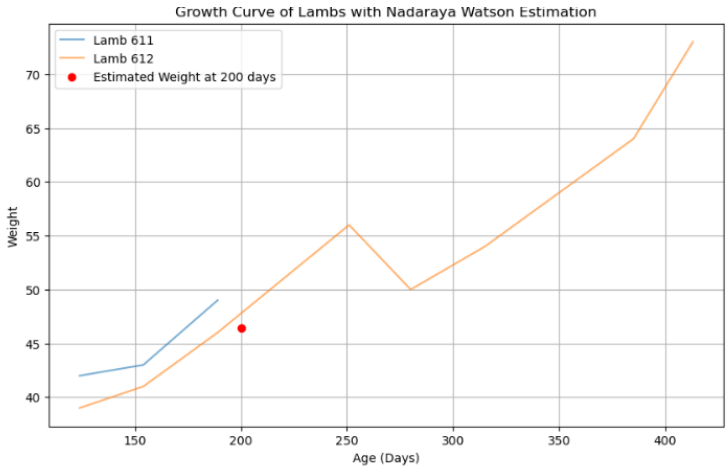

Figure 9: Estimated weight for lamb 612 at the age of 200 days.

$$\text{Weights} = e^{-c(\text{age}-\text{target age})^2}$$

After implementing the Nadaraya estimator we get the following result and curve:

```python
# Example: Estimate the weight of lamb 611 at age 200 days
lamb_no = 612
target_age = 200
c = 0.001  # You can adjust the value of c

estimated_weight = estimate_weight_nadaraya_watson(lamb_no, target_age, c)
print(f"Estimated weight of lamb {lamb_no} at age {target_age} days: {estimated_weight}")
```

```
Estimated weight of lamb 612 at age 200 days: 46.11458465221574
```

Figure 10: Nadaraya Estimator value for lamb 612 using inverse exponential weighting.

Here we got the result as 46.11 kg which is comparatively bad predictor as compared to kernel weighting method.

We can further look into the plot for better visualization for the estimated value.

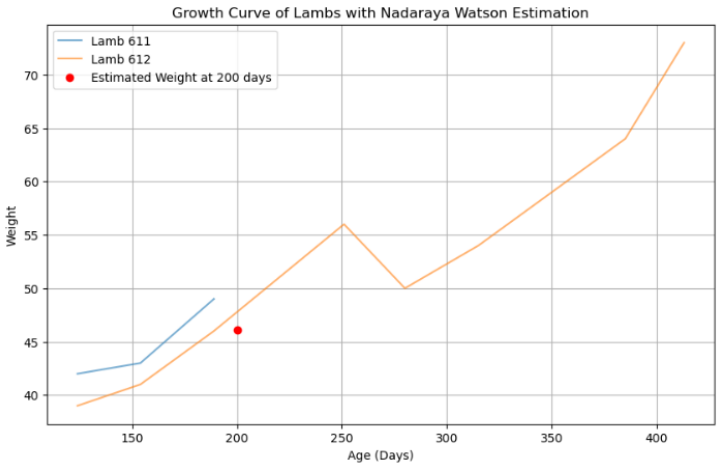

Figure 11: Visual representation of Estimated weight of lamb 612.

Local Linear Fit Model

Using the local linear fit model for estimating the weight of lamb at specified age.

Here are using the same weight function :

$$\text{Weights} = e^{-c(\text{age}-\text{target age})^2}$$

For the same lamb let's calculate the estimated weight at the age of 200 days.

```python
#Fitting the model

def estimate_weight_local_linear(lamb_no, target_age, c):
    lamb_data = final_df2[final_df2['No.'] == lamb_no].dropna(subset=['Weight'])
    X = lamb_data['Age (Days)'].values
    Y = lamb_data['Weight'].values
    estimated_weight = weighted_local_linear_fit(target_age, X, Y, c)
    return estimated_weight

# Example: Estimate the weight of lamb 611 at age 200 days
lamb_no = 612
target_age = 200
c = 0.001  # You can adjust the value of c

estimated_weight = estimate_weight_local_linear(lamb_no, target_age, c)
print(f"Estimated weight of lamb {lamb_no} at age {target_age} days: {estimated_weight}")
```

```
Estimated weight of lamb 612 at age 200 days: 47.72162598850616
```

Figure 12: Estimated weight of lamb 612 using LLF method.

Visual Representation

Figure 13 visually represents the estimated weight using the local linear fit method.

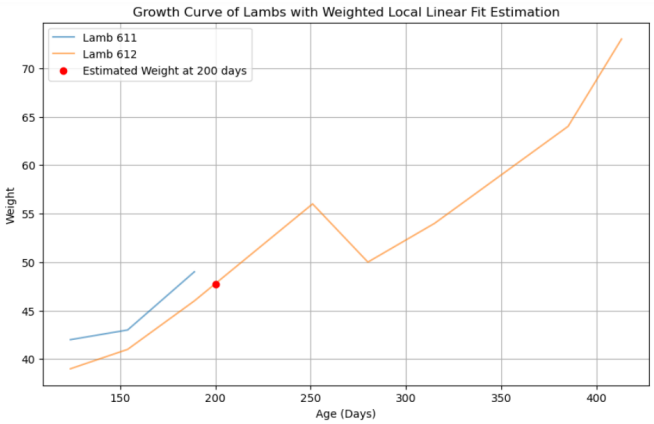

Figure 13: Estimated weight of lamb 612 using LLF method.

Based on the two weighting parameters and method LLF seems to be more precise in estimating the weights of lambs as per the growth curve.