

# Assignment 1: Case 1

Jyotsna Dalal, Pallav Walia, Shashank Singhal

04/02/2020

## Data

The data gives us detailed information about 4 different diets given to 50 chickens for a period of 21 days. There are 578 observations over 4 variables. Weight (in gms) and Time (in days) are numeric variables whereas Chicken and Diet are factor variables. The first few rows of the data look like this-

```
head(ChickWeight)
```

```
## Grouped Data: weight ~ Time | Chick
##   weight Time Chick Diet
## 1     42    0     1    1
## 2     51    2     1    1
## 3     59    4     1    1
## 4     64    6     1    1
## 5     76    8     1    1
## 6     93   10     1    1
```

```
summary(ChickWeight)
```

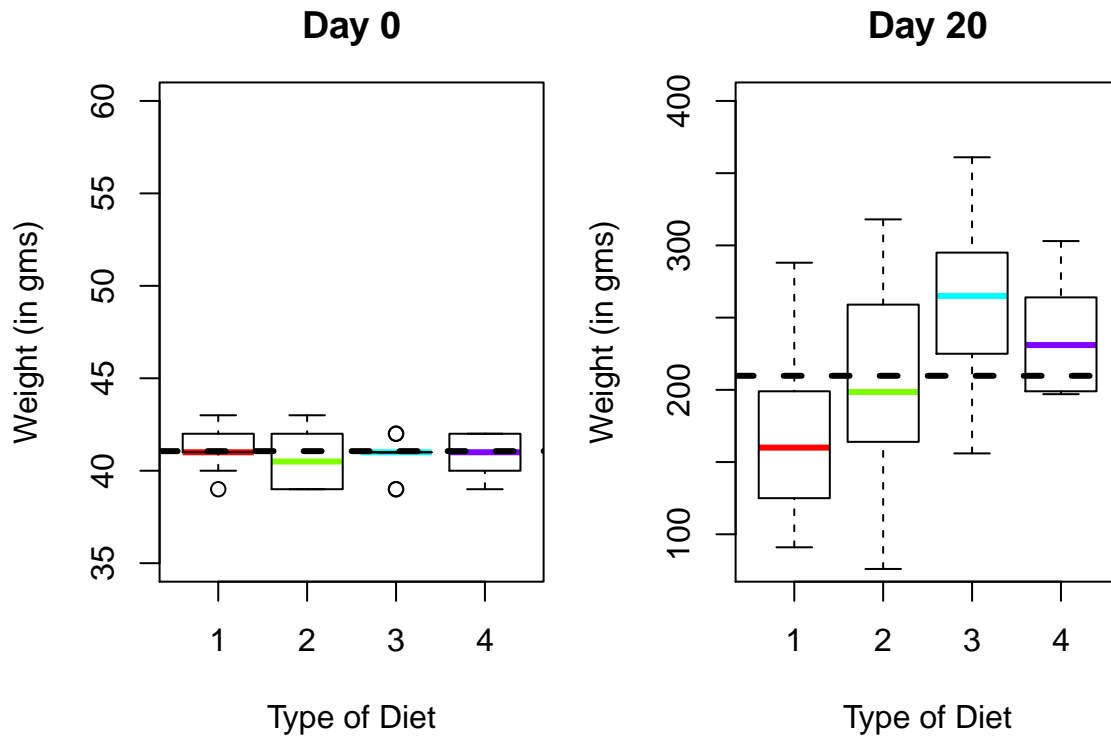
```
##           weight           Time           Chick           Diet
## Min.      : 35.0   Min.      : 0.00   13      : 12   1:220
## 1st Qu.: 63.0   1st Qu.: 4.00    9       : 12   2:120
## Median :103.0   Median :10.00   20      : 12   3:120
## Mean    :121.8   Mean    :10.72   10      : 12   4:118
## 3rd Qu.:163.8   3rd Qu.:16.00   17      : 12
## Max.    :373.0   Max.    :21.00   19      : 12
##                                     (Other):506
```

Summary of the data shows that the weight of a chicken ranges from 35 gms to 373 gms and there are no missing values in data.

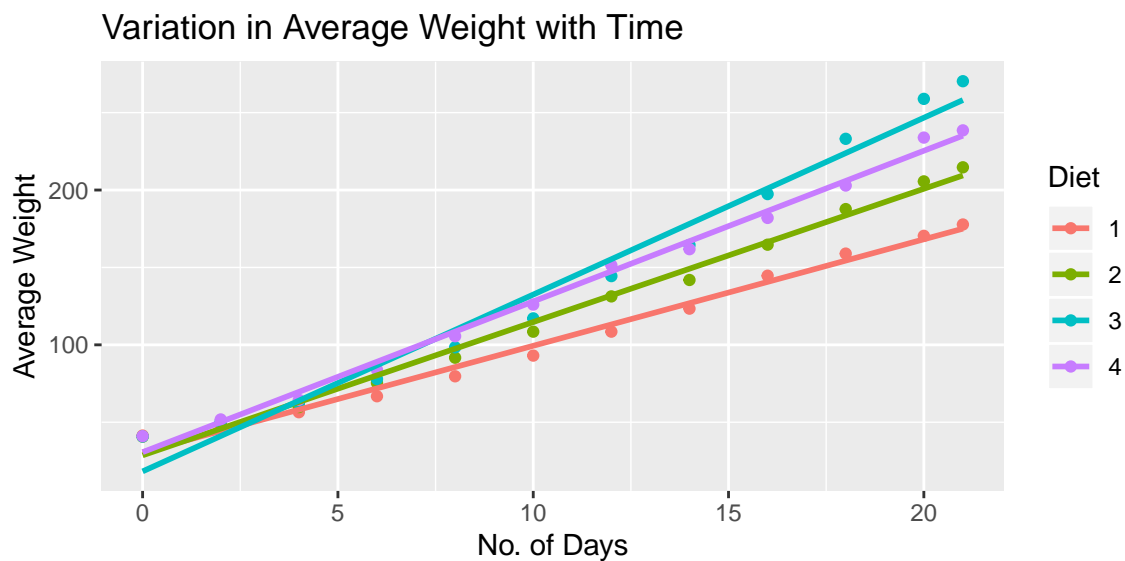
## Analysis

We get a basic overview of the data by plotting a simple box and whiskers of the four diets against weight. The range of weight on day 0 is from 39 gms to 43 gms with mean approx. 41 gms shown by a horizontal dashed line and the spread is highest for Diet 2. On day 20, the median of “Diet 3” is highest followed by “Diet 4”. The mean weight has increased from 43 to 210 gms

## Weight Distribution

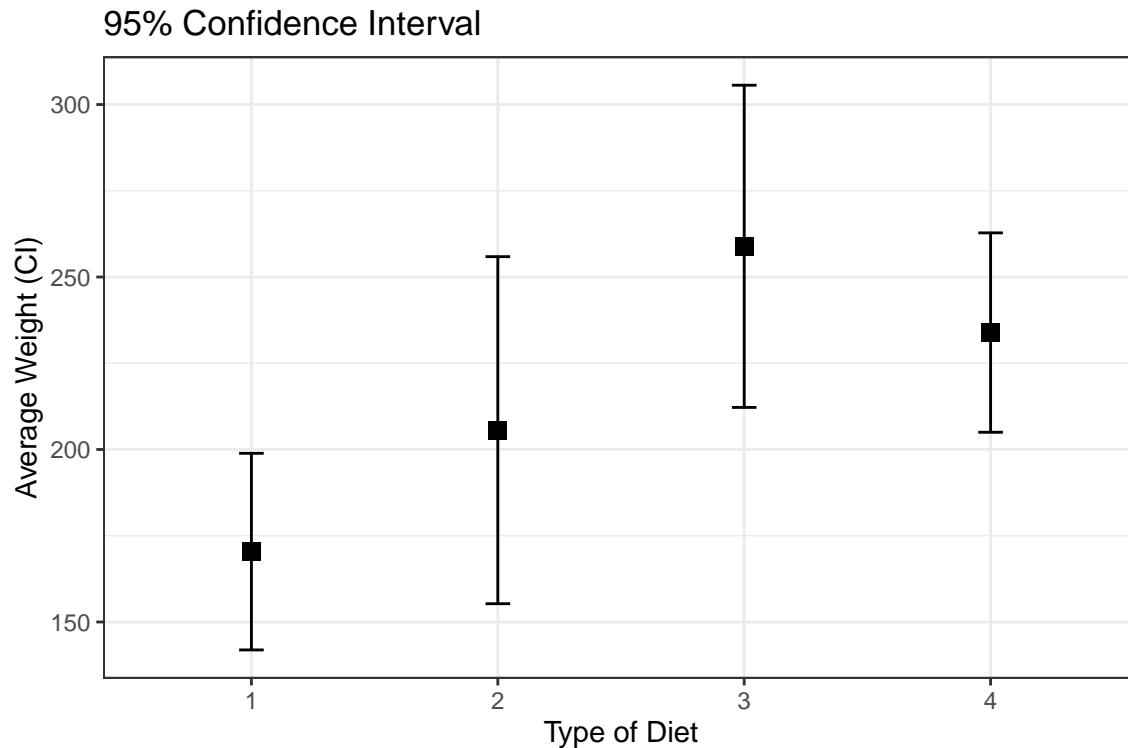


Next we look into the variations of weight with time. If we plot a bar graph of the four diets with average weight plotted against time, it will provide a better understanding of the effects of these diets on chickens over a period of time. For creating this graph, we need to first calculate the day wise mean weight among all diets using pipes function, then with the help of ggplot, we can compare the average weight of all diets using linear regression.



During the first week, “Diet 4” has the most positive impact however after 20 days, “Diet 3” stands out to be most effective among all with average chicken weight above 250 gms. There is a positive correlation between time and average weight for all the diets.

For calculating the confidence interval of mean weight on day 20 for each diet, we will use rcompanion library to summarise the results of mean and confidence interval and store the result in a new data frame called day20stats.



```
day20stats
```

```
##   Diet   n  Mean Conf.level Trad.lower Trad.upper
## 1     1  17 170.4      0.95    141.9    198.9
## 2     2  10 205.6      0.95    155.3    255.9
## 3     3  10 258.9      0.95    212.2    305.6
## 4     4   9 233.9      0.95    205.0    262.8
```

Confidence Interval gives us the level of certainty that the true value of population mean lies in the interval. The graph of CI tells us that “Diet 3” and “Diet 4” definitely increase the average weight more than “Diet 1”. For overlapping confidence intervals, a t-test for unknown population mean can tell us which mean is higher than the others.

## Limitations

Though the data shows the relationship between different diets and weight during a period of time but it does not specify the other factors like sex, feeding practices, vaccination etc. which may have contributed to the growth of chickens during the 20 days. With additional inputs, the exploratory data analysis would be more useful for understanding different patterns in the growth of chickens.

## Conclusion

The conclusion from the this analysis is that each diet has increased the weight of chicken in 20 days. “Diet 3” being the most effective followed by “Diet 4”. “Diet 1” is the least effective and “Diet 2” gives highly spreaded results with lower bound at 76 gms and upper bound at 318 gms on day 20.