

One-way ANOVA Practical Report – Baseline Value by Fetal Health

1. Objective

To test whether the mean baseline_value differs among the three fetal_health groups (Normal, Suspect, Pathological).

2. Hypotheses

- **Null Hypothesis (H_0):** Mean baseline_value is the same across all fetal_health groups.
 $H_0: \mu_1 = \mu_2 = \mu_3$
 - **Alternative Hypothesis (H_1):** At least one group has a different mean.
 $H_1: \text{At least one } \mu_i \text{ is different}$
-

3. Assumption Checks

3.1 Homogeneity of Variances (Levene's Test)

Test	F-value	p-value	Decision
Levene	9.3392	9.157×10^{-5}	Variances are not equal

Interpretation: The assumption of equal variances is violated. ANOVA is still robust for large sample sizes, but caution is advised.

3.2 Normality of Residuals (Shapiro-Wilk Test)

Test	W	p-value	Decision
Shapiro-Wilk	0.99624	3.807×10^{-5}	Residuals are not perfectly normal

Interpretation: Residuals deviate slightly from normality. Given the large sample size ($n = 2126$), ANOVA results are considered reliable.

Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J & Sir M.V. college

4. ANOVA Results

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
fetal_health	2	24073	12036	140.6	< 2e-16 ***
Residuals	2123	181717	86	-	-

Interpretation: The p-value < 0.05 indicates that there is a **statistically significant difference** in baseline values between at least two fetal_health groups. Therefore, we reject H_0 .

5. Post-hoc Analysis (Tukey HSD)

Comparison	Mean Difference	95% CI	p-value	Significance
2 – 1	9.7029	8.3316, 11.0742	0.0000	Yes
3 – 1	-0.2944	-2.0147, 1.4260	0.9151	No
3 – 2	-9.9972	-12.0639, -7.9306	0.0000	Yes

Interpretation of post-hoc results:

- Group 2 (Suspect) has significantly higher baseline values than Group 1 (Normal).
 - Group 3 (Pathological) is **not significantly different** from Group 1 (Normal).
 - Group 3 (Pathological) has significantly lower baseline values than Group 2 (Suspect).
-

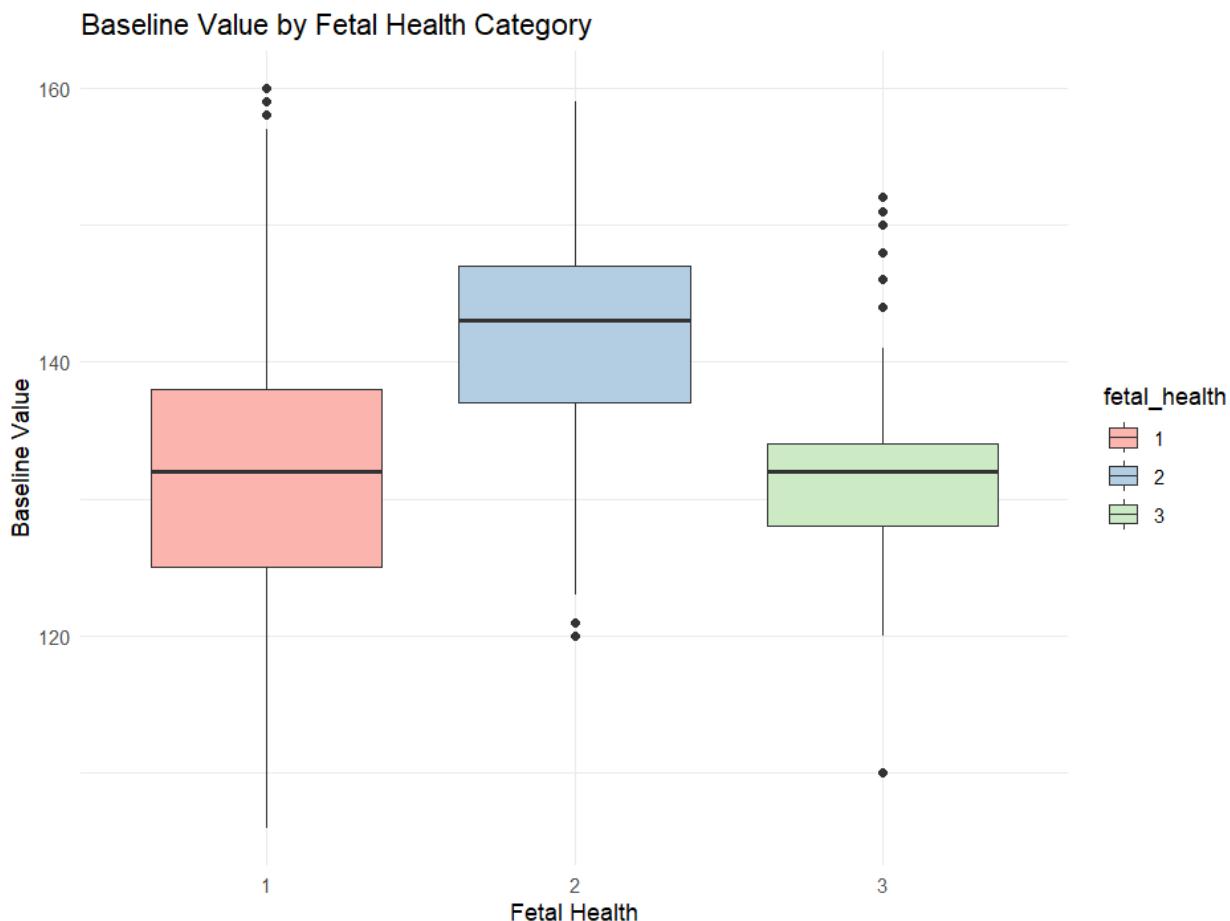
Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J & Sir M.V. college

6. Visualization

- **Boxplot of Baseline Value by Fetal Health**



- **ANOVA and TukeyHSD output screenshots**

```
>
> # Perform ANOVA
> summary(anova_model)
      Df Sum Sq Mean Sq F value Pr(>F)
fetal_health    2 24073   12036   140.6 <2e-16 ***
Residuals     2123 181717       86
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> ... . . . . .
```



```
> # Post-hoc analysis
> TukeyHSD(anova_model)
  Tukey multiple comparisons of means
  95% family-wise confidence Level

Fit: aov(formula = baseline_value ~ fetal_health, data = df)

$fetal_health
        diff      lwr      upr      p adj
2-1  9.7028727  8.331557 11.074188 0.000000
3-1 -0.2943731 -2.014732  1.425986 0.915065
3-2 -9.9972458 -12.063926 -7.930566 0.000000
```

Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J & Sir M.V. college

Interpretation from Boxplot:

- Group 2 (Suspect) shows a higher median baseline value.
- Group 3 (Pathological) overlaps with Group 1 (Normal).
- Visual observation aligns with the TukeyHSD findings.

7. Conclusion

The analysis confirms that **baseline_value** differs significantly across fetal health categories, mainly due to elevated values in the Suspect group (Group 2). The results are consistent across ANOVA and post-hoc TukeyHSD tests, and the boxplot visually demonstrates these differences.

Note: Minor violations of ANOVA assumptions (normality and homogeneity of variances) exist, but given the **large sample size (n = 2126)**, the results are still reliable.

Screenshots

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Console Terminal Background Jobs
R 4.5.2 - ->
> # Enhanced ggplot boxplot
> install.packages("ggplot2")
> # load data
> df <- read.csv("C:/Users/Priya Gupta/Downloads/fetal_health.csv")
> head(df)
baseline_value accelerations fetal_movement uterine_contractions light_decelerations severe_decelerations prolonged_decelerations abnormal_short_term_variability
1 120 0.000 0 0.000 0.000 0 0.000 73
2 132 0.006 0 0.006 0.003 0 0.000 17
3 133 0.003 0 0.008 0.003 0 0.000 16
4 134 0.003 0 0.008 0.003 0 0.000 16
5 132 0.007 0 0.008 0.000 0 0.000 16
6 134 0.001 0 0.010 0.009 0 0.002 26
mean_value_of_short_term_variability percentage_of_time_with_abnormal_long_term_variability mean_value_of_long_term_variability histogram_width histogram_min histogram_max
1 0.5 43 2.4 64 62 126
2 2.1 0 10.4 130 68 198
3 2.1 0 13.4 130 68 198
4 2.4 0 23.0 117 53 170
5 2.4 0 19.9 117 53 170
6 5.9 0 0.0 150 50 200
histogram_number_of_peaks histogram_number_of_zeroes histogram_mode histogram_mean histogram_median histogram_variance histogram_tendency fetal_health
1 2 0 120 137 127 73 1 2
2 6 1 141 136 140 122 0 1
3 5 1 141 135 138 13 0 1
4 11 0 137 134 137 13 1 1
5 9 0 137 136 138 11 1 1
6 5 3 76 107 107 170 0 3
> str(df)
'data.frame': 2126 obs. of 22 variables:
$ baseline_value : int 120 132 133 134 132 134 134 122 122 122 ...
$ accelerations : num 0 0.006 0.005 0.003 0.007 0.007 0.000 0.001 0 0 ...
$ fetal_movement : num 0 0.006 0.005 0.004 0.005 ...
$ uterine_contractions: num 0 0.006 0.008 0.008 0.008 0.008 0.008 0.01 0.013 0 0.002 0.003 ...
$ light_decelerations: num 0 0.003 0.003 0.003 0.003 0.009 0.009 0.009 0 0 0 ...
$ severe_decelerations: num 0 0 0 0 0 0 0 0 0 0 ...
$ prolonged_decelerations: num 0 0 0 0 0 0.002 0.003 0 0 0 ...
$ abnormal_short_term_variability: num 73 17 16 16 26 29 83 84 86 ...
$ mean_value_of_short_term_variability: num 0.5 2.1 2.1 2.4 2.4 5.9 6.3 0.5 0.5 0.3 ...
$ percentage_of_time_with_abnormal_long_term_variability: int 43 0 0 0 0 0 0 6 5 6 ...
$ mean_value_of_long_term_variability: num 2.4 10.4 13.4 23 19.9 0 0 15.6 13.6 10.6 ...
$ histogram_width: int 64 130 130 117 117 150 150 68 68 68 ...
$ histogram_min: int 62 68 68 52 52 56 56 62 62 62 ...

```

Name :- Priya Gupta

Roll no. :- S081

Sheth L.U.J & Sir M.V. college

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Console Terminal Background Jobs
R 4.5.2 ->
\$ histogram_min : int 62 68 68 53 53 50 50 62 62 62 ...
\$ histogram_max : int 126 198 198 170 170 200 200 130 130 130 ...
\$ histogram_number_of_peaks : int 2 6 5 11 9 5 6 0 0 1 ...
\$ histogram_number_of_zeroes : int 0 1 0 0 3 3 0 0 0 ...
\$ histogram_mode : int 120 141 143 137 137 76 71 122 122 122 ...
\$ histogram_mean : int 137 136 134 134 136 107 107 122 122 122 ...
\$ histogram_median : int 121 140 138 137 138 107 106 123 123 123 ...
\$ histogram_variance : int 111 113 117 117 115 70 215 3 3 1 ...
\$ histogram_tendency : int 1 0 0 1 1 0 0 1 1 1 ...
\$ fetal_health : int 2 1 1 1 1 3 3 3 3 3 ...
>
> # Convert fetal_health to factor
> df\$fetal_health <- as.factor(df\$fetal_health)
> str(df\$fetal_health)
Factor w/ 3 levels "1","2","3": 2 1 1 1 1 3 3 3 3 3 ...
>
> # Check ANOVA assumptions
> anova_model <- aov(baseline_value ~ fetal_health, data = df)
>
> # Normality of residuals
> res <- residuals(anova_model)
> shapiro.test(res)
Shapiro-Wilk normality test
data: res
W = 0.99624, p-value = 3.807e-05
>
> # Homogeneity of variance
> #install.packages("car")
> library(car)
> leveneTest(baseline_value ~ fetal_health, data = df)
Levene's Test for Homogeneity of Variance (center = median)
DF F value Pr(>F)
group 2 9.3392 9.157e-05 ***
2123

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> # Perform ANOVA

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Console Terminal Background Jobs
R 4.5.2 ->
DF F value Pr(>F)
group 2 9.3392 9.157e-05 ***
2123

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> # Perform ANOVA
> summary(anova_model)
Fit: aov(formula = baseline_value ~ fetal_health, data = df)
\$fetal_health
diff lwr upr adj
2-1 9.702727 8.331557 11.074188 0.000000
3-1 -0.2943731 -2.014732 1.425986 0.915065
3-2 -9.9972458 -12.063926 -7.930566 0.000000
>
>
> # Enhanced ggplot boxplot
> #install.packages("ggplot2")
> library(ggplot2)
> ggplot(df, aes(x = fetal_health, y = baseline_value, fill = fetal_health)) +
+ geom_boxplot() +
+ theme_minimal() +
+ labs(title = "Baseline Value by Fetal Health Category",
+ x = "Fetal Health", y = "Baseline Value") +
+ scale_fill_brewer(palette = "Pastel1")
> |

Name :- Priya Gupta
Roll no. :- S081