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

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.
  H0:μ1=μ2=μ3

- **Alternative Hypothesis (H$_1$):** At least one group has a different mean.
  H1:At least one μi 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

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

6. Visualization

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

Baseline Value by Fetal Health Category
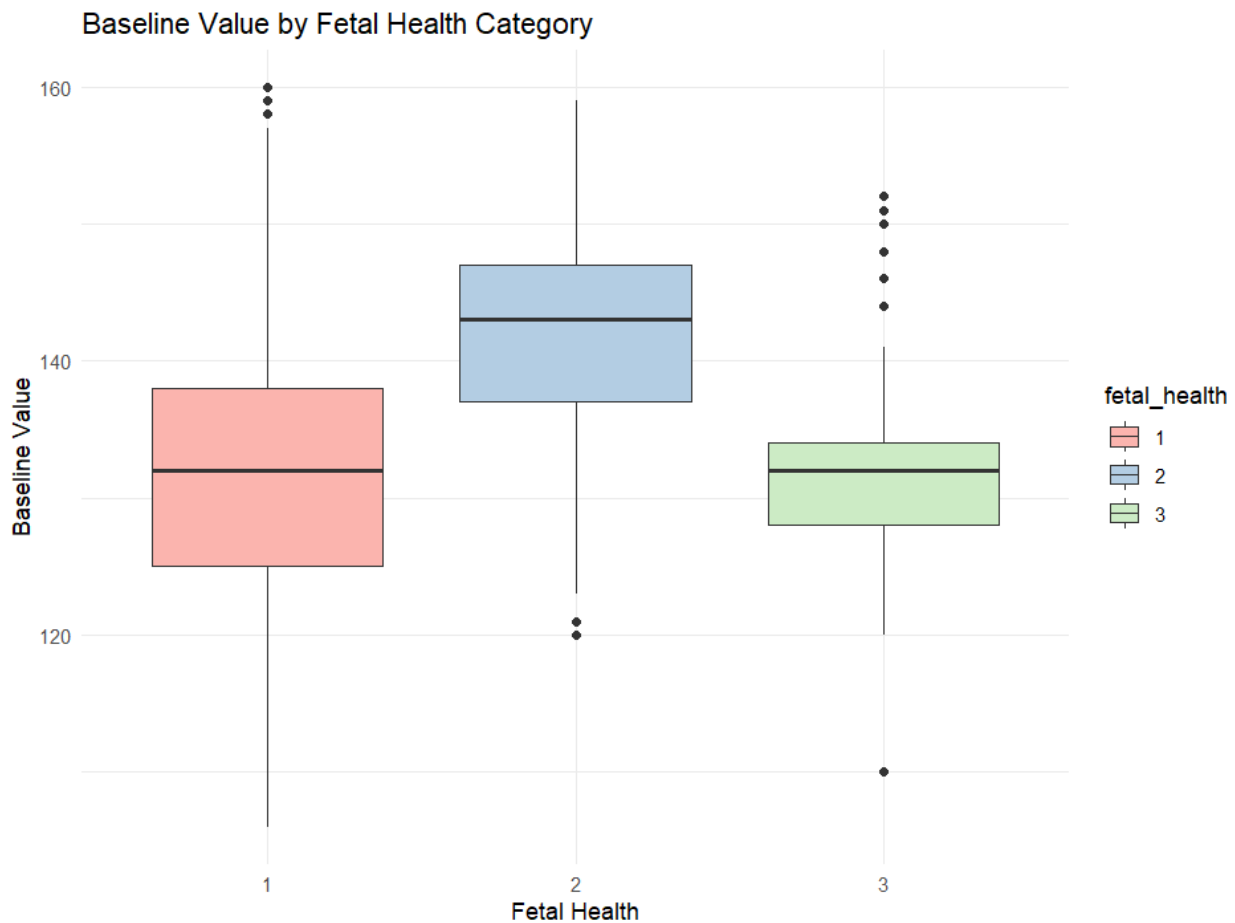
- **ANOVA and TukeyHSD output screenshots**

```
>
> # 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

Interpretation from Boxplot:
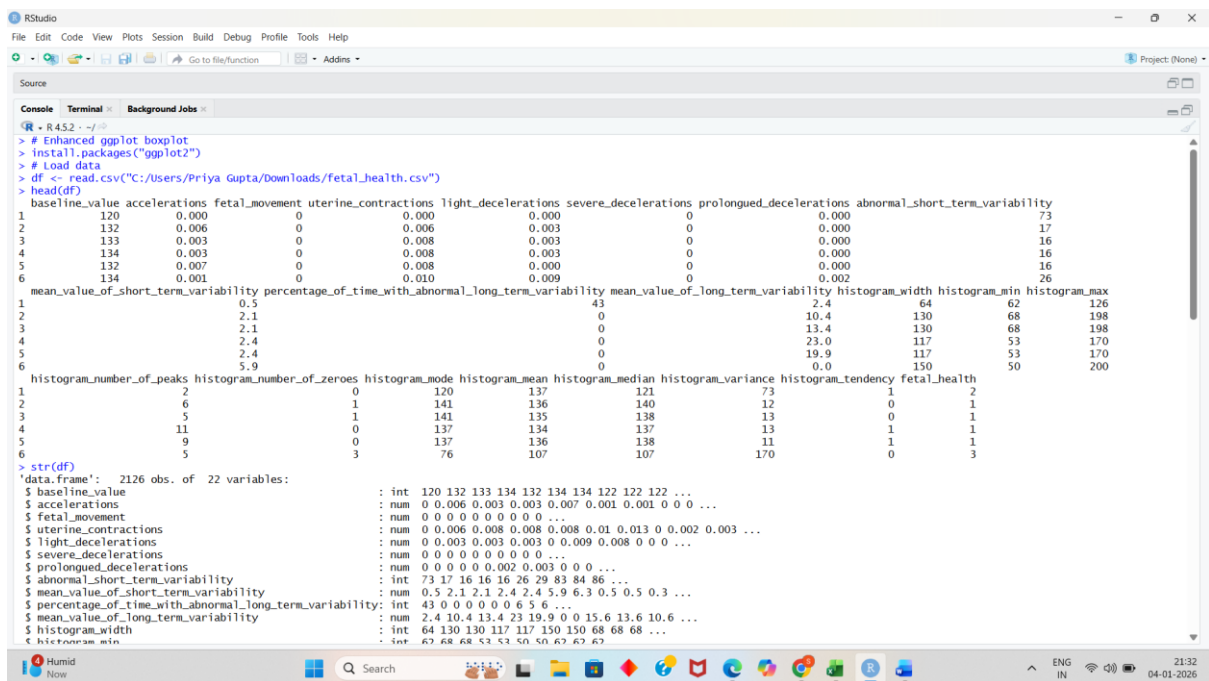
- Group 2 (Suspect) shows a higher median baseline value.

- 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

Name :- Priya Gupta
Roll no. :- S081

```
$ 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 1 0 0 3 0 0 0 ...
$ histogram_mode                    : int  120 141 141 137 137 76 71 122 122 122 ...
$ histogram_mean                    : int  137 136 135 134 136 107 107 122 122 122 ...
$ histogram_median                  : int  121 140 138 137 138 107 106 123 123 123 ...
$ histogram_variance                : int  73 12 13 13 11 170 215 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
>
```



```
        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)
              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

>
>
>
> # 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