Statistical Analysis of the 'Luxuriant' Treatment Effect on Hair Growth

Anastasia Akchurina

2023-11-17

Abstract

This report provides a comprehensive analysis of a clinical trial evaluating the effectiveness of a new drug "Luxuriant" for treating male pattern baldness. The study compares Luxuriant with a placebo and two existing anti-baldness treatments "BaldBeGone" and "HairyGoodness" considering hair growth and patients' ages. The analysis involves data manipulation, linear regression modeling and thorough model checking, and data visualisation.

Key findings include evidence of non-normality in the residuals distribution and violations of the assumption of constant variance. On the other hand, the results of the analysis offer valuable insights into the efficiency of Luxuriant compared to the placebo and other treatments. Luxuriant appears to be more effective than the placebo, though it falls behind HairyGoodness and BaldBeGone.

Introduction

This report summarises the statistical analysis of a clinical trial for a new drug "Luxuriant" to treat male pattern baldness. The main objective of the report is to answer the following three questions and to present the results to the company:

1) Is there an effect of Luxuriant above and beyond the placebo?
2) Is Luxuriant more effective than the existing treatments on the market?
3) Is age relevant to any effect?

The data available for the analysis comprised hair growth in inches and the patients' age (from 30 to 73). The methods are commonly used for this sort of analysis included data manipulation, fitting a linear regression model, assessing the model assumptions, data visualisation and interpretation of results.

Analysis

The new drug "Luxuriant" was considered in the same trial with both a placebo and two existing anti-baldness treatments "BaldBeGone" and "HairyGoodness". The pateients were randomly allocated to different groups, shaved and had any hair growth measured after one month of treatment. The first four columns of the dataframe represent the hair growth results for "Luxuriant", the placebo, "BaldBeGone" and "HairyGoodness" respectively.

	Luxuriant	Placebo	BaldBeGone	HairyGoodness
0	0.792234	0.080613	0.280537	0.520870
1	0.605398	0.037706	2.251034	0.421710
2	0.694443	0.097158	1.421599	1.830612
3	0.054321	0.217363	0.694945	0.833303
4	1.478682	0.196568	0.751972	0.300818

The next four columns of the dataframe represent age of the patients allocated to the groups by treatment.

	AgeLuxuriant	AgePlacebo	AgeBaldBeGone	AgeHairyGoodness
0	58.7	44.2	50.6	51.0
1	60.7	52.7	52.2	52.3
2	53.7	48.4	66.4	52.5
3	62.7	46.1	46.0	54.4
4	49.5	56.7	60.9	40.0

Data Manupilation

The first step in the analysis was to create separate groups based on different product types: Luxuriant, Placebo, BaldBeGone, and HairyGoodness, rename columns for consistency across different product groups and concatenate the groups into one dataframe which was necessary for further analysis.

`Combined DataFrame:`

	Hair Growth	Age	Product
0	0.792234	58.7	Luxuriant
1	0.605398	60.7	Luxuriant
2	0.694443	53.7	Luxuriant
3	0.054321	62.7	Luxuriant
4	1.478682	49.5	Luxuriant

Next step in data manipulation was to change units from inches to mm, round age to whole years and change indexes so the dataframe looked more natural and appealing.

`Updated DataFrame:`

	Hair Growth	Age	Product
1	20.122736	58.0	Luxuriant
2	15.377107	60.0	Luxuriant
3	17.638841	53.0	Luxuriant
4	1.379751	62.0	Luxuriant
5	37.558535	49.0	Luxuriant

Furthermore, since we want to know how treatments and age of patients influence hair growth, the linear regression model was designed and fitted: Hair_Growth ~ Age + Product.

Model checking

A host of testing and plotting were performed to check the model assumptions.

For assessing normality:

1) Histogram of residuals
2) Quantile-Quantile plot
3) Shapiro-Wilk's test, and other residual plots

For testing independence:

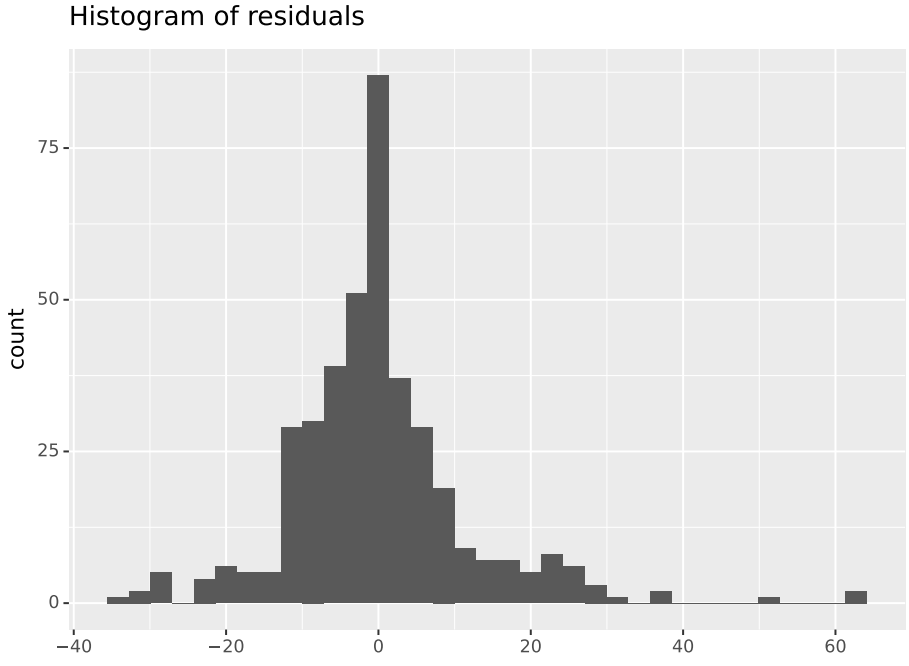
1) Residuals plot
2) Durbin-Watson test

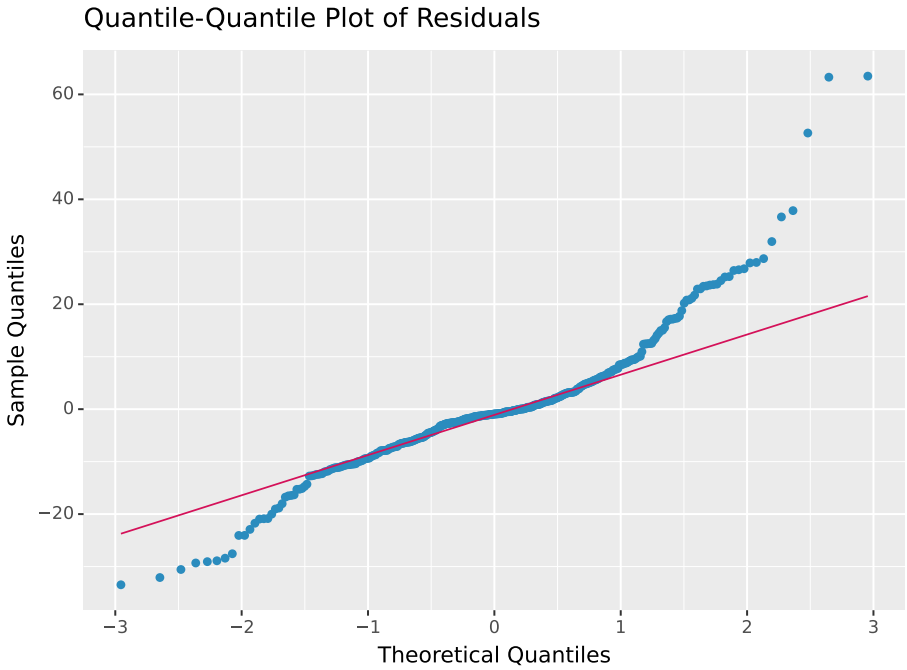
For variance testing:

1) Variance plot
2) Breusch-Pagan test

And creating a plot to check predicted vs observed values.

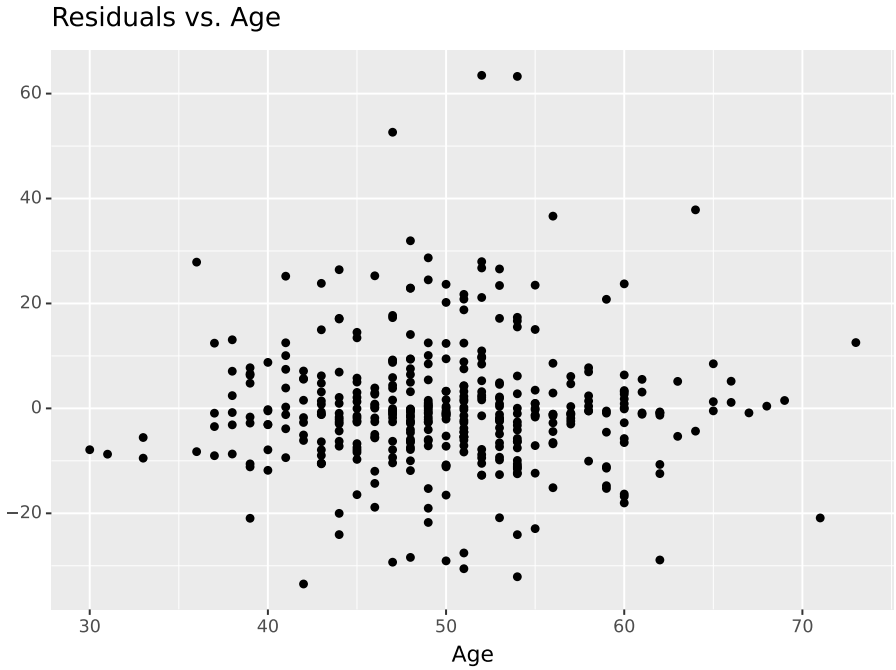
Both the "Histogram of residuals" and the Quantile-Quantile plot show evidence of non normality as does the Shapiro-Wilk's test. Therefore, we might want to consider transforming the response data to make it more normal or designing an alternative model.

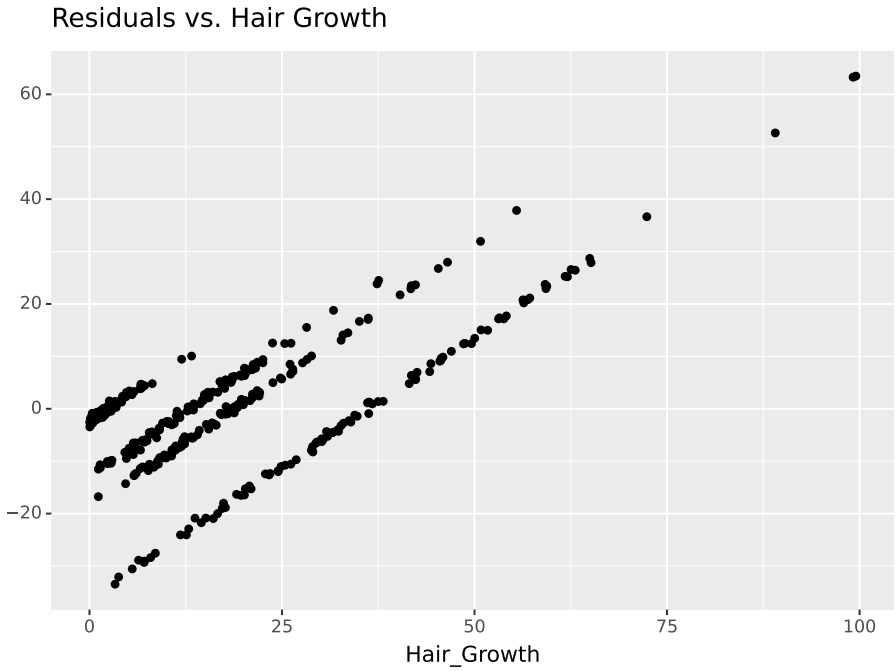




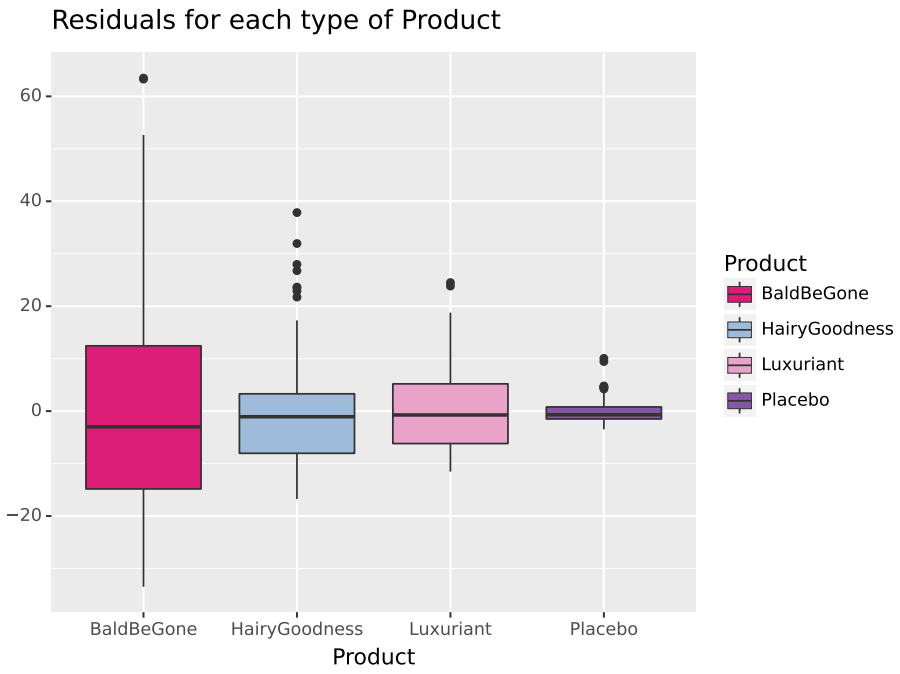

```
Normality test
ShapiroResult(statistic=0.9094886183738708, pvalue=9.946553739610849e-15)
```

The “Residuals vs. Age” and “Residuals vs. Hair Growth” scatter plots confirm the non-normality of the residuals distribution as we expect to see a roughly horizontal band around zero.

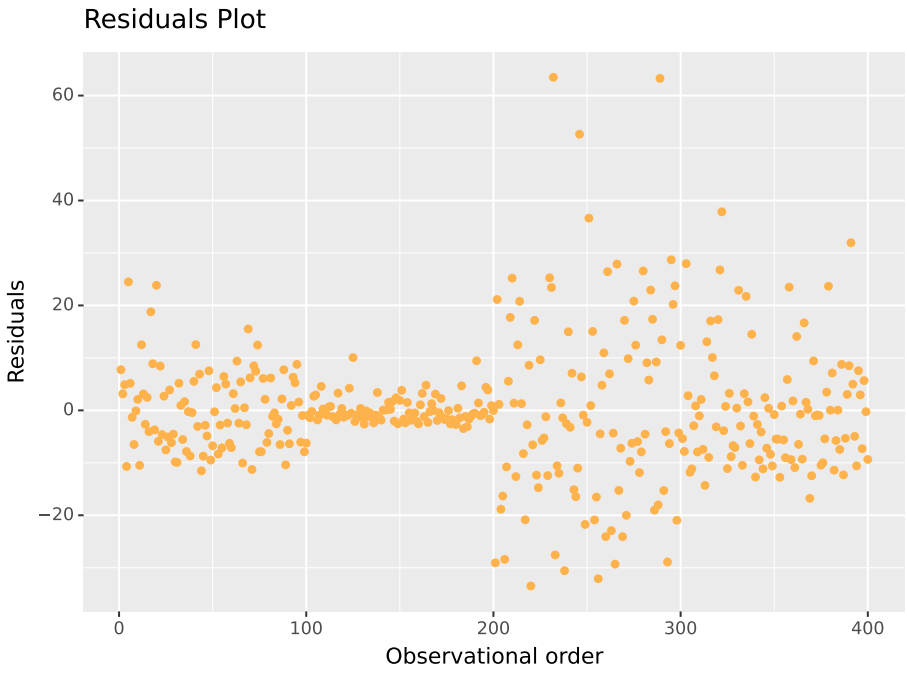



Residuals vs. Hair Growth
60
40
20
0
−20
0
25
50
75
100
Hair_Growth

The aim of creating the following boxplots is to identify if there are systematic differences in the residuals distribution among the drugs and placebo. We can see that the "BaldBeGone" drug has the highest level of noise around it and the placebo has the lowest.

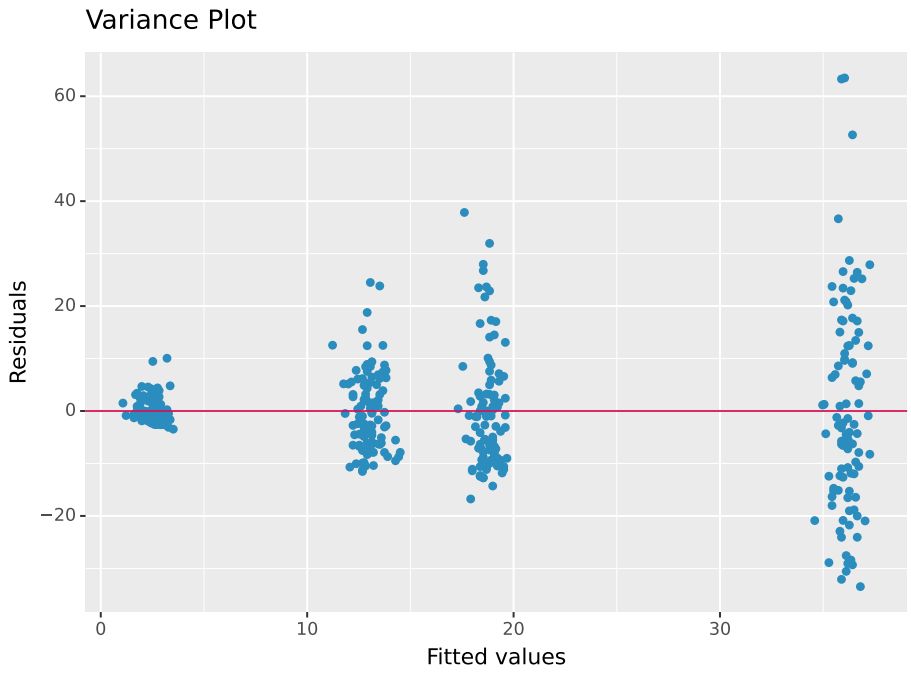


According to the "Residuals Plot" and the Durbin-Watson test, the residuals are independent.




```
Independnece Test
1.9768599638153963
```

The "Variance Plot" and the Breusch-Pagan test suggest that the assumption of constant variance of residuals is violated. The values in the output indicate that the test statistic (Lagrange multiplier) is relatively large and the associated p-values are extremely small. This provides a strong evidence against the assumption of constant variance.



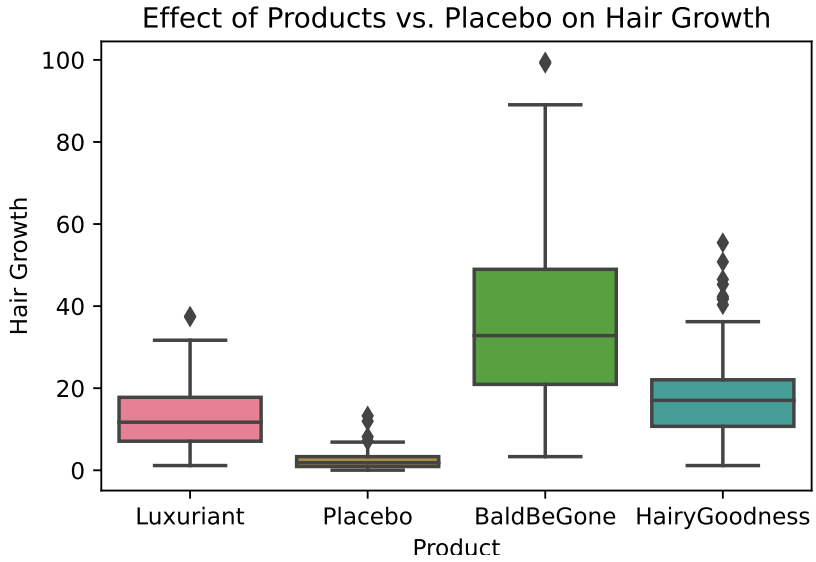

```
Constant Variance Test
[('Lagrange multiplier statistic', 61.063008737020574), ('p-value', 1.7341266798624277e-12),
```

The “Observed Values vs. Fitted Values” plot conveys there might be an evidence that the model predictions are lower than the actual values as we expect the fitted vs observed values to be on a 45 line.

Results

Is there an effect of Luxuriant above and beyond the placebo?

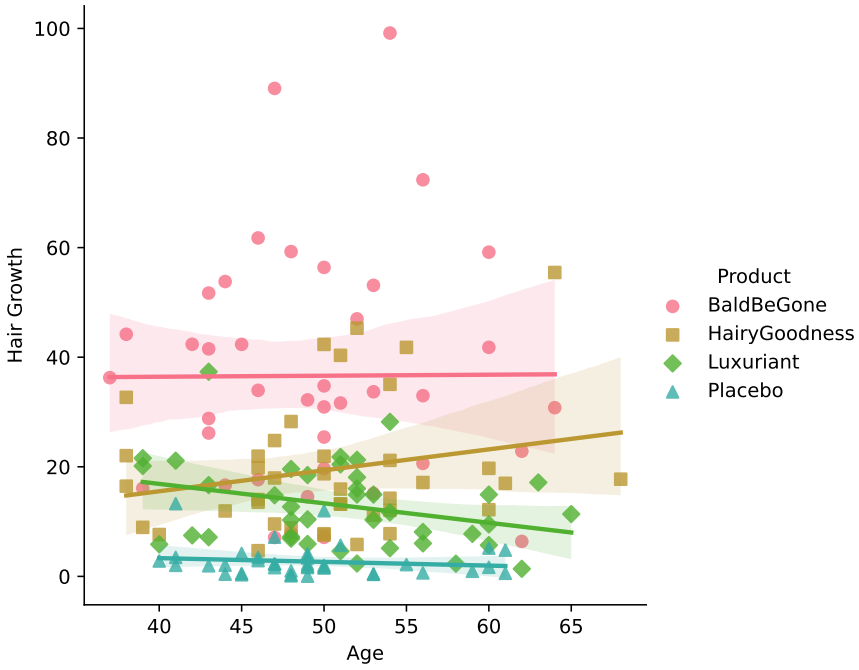
To answer this question we created a side-by-side boxplots to compare the effect of Luxuriant vs. Placebo on hair growth and conducted an independent two-sample t-test. The test statistic is 13.307 and the p-value is 2.604826e-25 which indicates a significant difference between Luxuriant and placebo. In addition, the boxplots "Effect of Products vs. Placebo on Hair Growth" conveyed that Luxuriant is more effective for hair growth than Placebo which is a good news for Luxuriant developers.

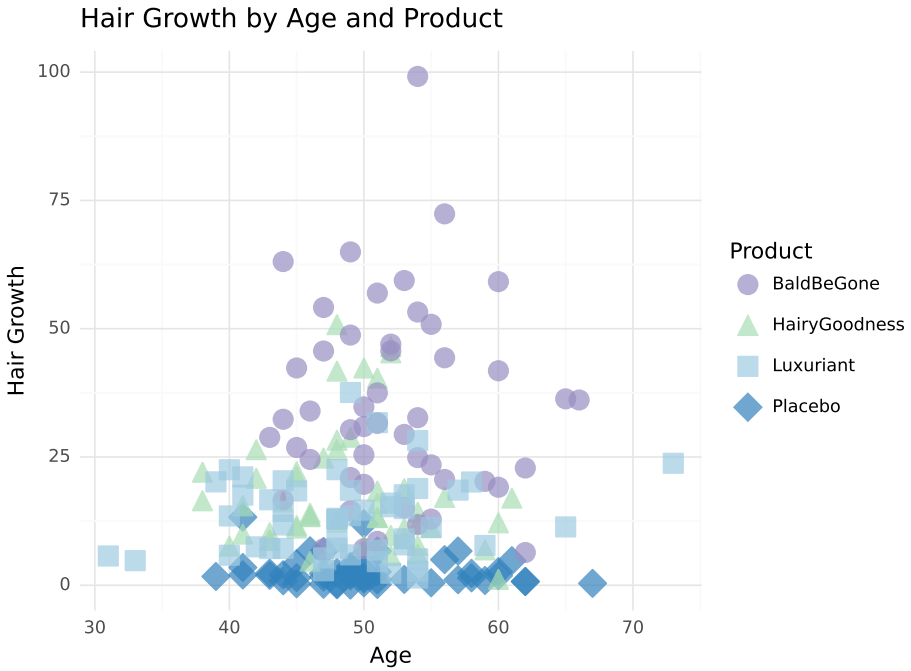



```
T-statistic: 13.307183044116524
P-value: 2.6048259801032925e-25
Luxuriant is more effective than Placebo.
```

2) Is Luxuriant more effective than the existing treatments on the market?

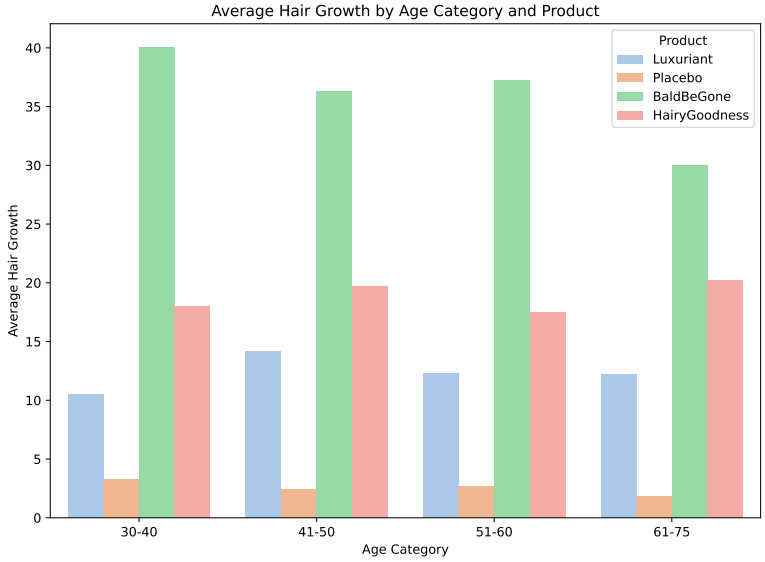
To analyse the effectiveness of Luxuriant for hair growth among other drugs, two scatter plots were created: a scatter plot with a regression line for each product and "Hair Growth by Age and Product" scatter plot. Both plots show that Luxuriant is not as effective as HairyGoodness and BaldBeGone treatments since it is scattered lower in the plot than other drugs except for placebo.


100
80
60
40
20
0
Hair Growth
40
45
50
55
60
65
Age
Product
BaldBeGone
HairyGoodness
Luxuriant
Placebo


Hair Growth by Age and Product
100
75
50
25
0
Hair Growth
30
40
50
60
70
Age
Product
BaldBeGone
HairyGoodness
Luxuriant
Placebo

3) Is age relevant to any effect?

Finally, creating a grouped bar plot was considered a good option to analyse the relationship between age, product, and average hair growth. The "Average Hair Growth by Age Category and Product" plot shows that there is no significant variability in age regarding other factors meaning age is not quite relevant to any of the effects.



Discussion / Conclusion

To sum it all up, the clinical trial suggests that Luxuriant has a positive effect on hair growth surpassing the placebo but falling short of the effectiveness demonstrated by HairyGoodness and BaldBeGone. Model checking procedures revealed some problems in meeting the assumptions of normality and constant variance in the residuals.

The assessment of age as a relevant factor in the treatment effects suggests that age does not significantly contribute to the variability observed in the responses to Luxuriant and other treatments.

The findings from this analysis are a good soil for the developers of Luxuriant drug, providing insights about its performance in comparison to existing treatments. Further investigations and adjustments in the modelling approach may be considered to refine the analysis and address the challenges discovered in the data. Overall, this report contributes to the understanding of the clinical trial

results and provides much food for thought in terms of research and development in the field of hair growth treatments.