Analysis of ClimMob project Home tasting

Automated Report by ClimMob

08 June, 2020

You are reading a report generated by ClimMob. This is a software package to analyze data generated by citizen science or crowdsourcing.

# Introduction

In agriculture, the local environmental conditions determine to a large degree which technological solutions are the most suitable. In dry soils, for example, drought-resistant crop varieties will outperform other varieties, but in wet soils these same varieties may do worse than most. Not only drought, but an entire range of problems including excessive heat, floods, new pests and diseases tend to intensify under climate change. This multitude of limiting factors requires multiple technological solutions, tested in diverse environments.

Citizen science is based on the cooperation of citizen scientist or observers (paid or unpaid). Researchers assign microtasks (observations, experiments) that, once completed and gathered, contribute with a great amount of information to science. One of the advantages of citizen science is that agricultural researchers can get access to many environments by crowdsourcing their experiments. As farmers contribute with their time, skills and knowledge to the investigation, researchers are able to do more tests than in a traditional setup. Also citizen scientists acquire new knowledge, abilities and information useful for future challenges of their work.

## ClimMob

The primary goal of ClimMob is to help farmers adapt to variable and changing climates. ClimMob was created as part of Bioversity International’s research in the CGIAR Research Programme on Climate Change, Agriculture, and Food Security (CCAFS). It serves to prepare and analyze citizen science experiments in which a large number of farmers observe and compare different technological options under a wide range environmental conditions (*1*).

ClimMob software assigns a limited number of items (typically 3 crop varieties or agricultural practices) to each participant, who will compare their performance. Each participant gets a different combination of items drawn from a much larger set of items. Comparisons of this kind are thought to be a very reliable way to obtain data. Once the results of the microtasks have been collected, ClimMob builds an image of the whole set of assigned objects, combining all observations. ClimMob not only reconstructs the overall ordering of items, but also takes into account differences and similarities between participants and the conditions under which they observe (e.g. socio-economic and plot environmental characteristics). It assigns similar participants to groups that each corresponds among different group profiles. Groups are created on the basis of whichever items which have been collected, that are found to be significantly linked to the observed rankings.

ClimMob uses Plackett-Luce models to analyze ranking data with the R (*2*) package ‘PlackettLuce’ (*3*). It automatically generates analytical reports, as well as individualized information sheets for each participant using the R packages ‘knitr’ (*4*) and ‘rmarkdown’ (*5*). Organizing the data relies on packages ‘ClimMobTools’ (*6*), ‘gosset’ (*7*), ‘gtools’ (*8*), ‘jsonlite’ (*9*), ‘partykit’ (*10*), ‘psychotools’ (*11*) and ‘qvcalc’ (*12*). Summaries and data visualization are supported by packages ‘igraph’ (*13*), ‘ggparty’ (*14*), ‘ggplot2’ (*15*), ‘ggrepel’ (*16*), ‘mapview’ (*17*), ‘multcompView’ (*18*), ‘patchwork’ (*19*), and ‘pls’ (*20*).

## How to cite

If you publish any results generated with ClimMob, you should cite a number of articles as the package builds on various contributions. Van Etten et al. (2019) (*1*) introduced the crowdsourcing philosophy behind ClimMob. It is important to mention that ClimMob is implemented in R, a free, open-source analysis software (*2*). Methodologically, if you report on the Plackett-Luce tree results, you should mentioned that ClimMob applies the Plackett-Luce model published by Turner et al. 2020 (*3*). To cite ClimMob itself, mention van Etten et al. (2020) (*21*).

# Section 1: Headline results

Overall there were 383 participants contributing to this project. Each participant assessed 3 different varieties and ranked them in order of its overall performace. In addition they also provided rankings for 2 additional characteristics:

|  |  |  |
| --- | --- | --- |
| Characteristic | Short name | Number of valid answers |
| Overall Characteristic | Overall Characteristic | 276 |
| Color | Color | 276 |
| Taste | Taste | 276 |

Table 1 shows the varieties assessed within this project, with the frequency and percentage of participants who assessed each variety. If the varieties has large names (> 10 characters) an abbreviation was applied across the figures in this report.

Table 1. Frequency of varieties assessed.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variety | Abbreviation | Freq | Relative freq | Man (n=170) | Woman (n=210) |
| Ejumula | Ejumula | 216 | 56.4% | 100 | 116 |
| Kakamega | Kakamega | 195 | 50.9% | 86 | 109 |
| Naspot 10 | Ns10 | 186 | 48.6% | 77 | 106 |
| Naspot 12 | Ns12 | 161 | 42% | 69 | 89 |
| Naspot 13 | Ns13 | 189 | 49.3% | 81 | 105 |
| Naspot 8 | Naspot 8 | 202 | 52.7% | 97 | 105 |

This is how the different varieties were connected.

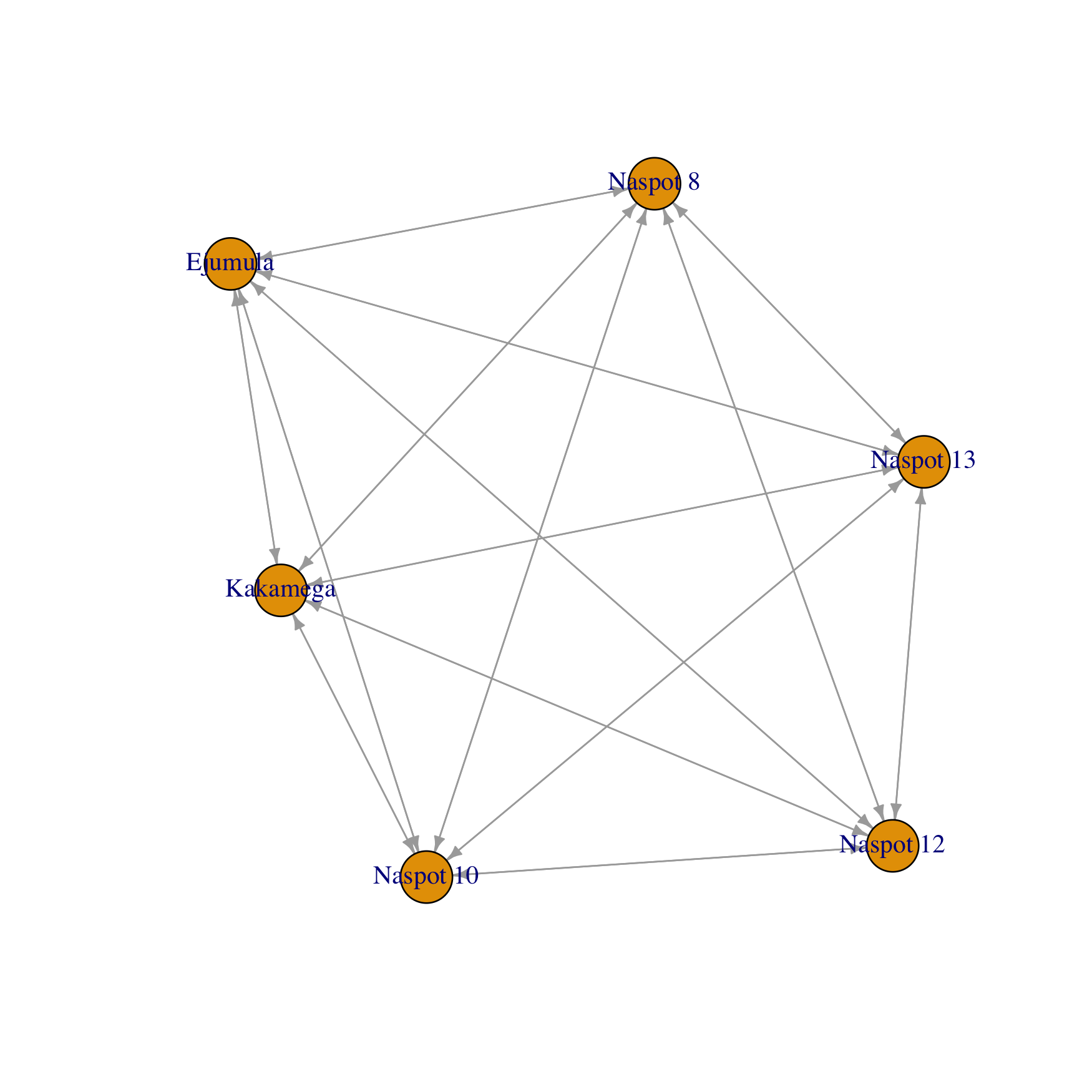


Figure 1. Network representation of varieties tested in this project.

## Overall differences in rankings

There were statistically significant differences found in the rankings of varieties in the overall performance (p=< 2.22e-16 \*\*\*). The best ranked varieties overall were Ejumula . Statistically significant differences were also found in the characteristic(s) Color

A summary of the p-values testing the hypothesis that there exist differences in the rankings within each of the Plackett-Luce models fitted for each of the assessed characteristics, and the list of varieties which were significantly highest and lowest ranked overall, are summarised in Table 1.1.

Table 1.1. Summary of differences found in varieties by characteristic.

|  |  |  |  |
| --- | --- | --- | --- |
| Ranking | Best Ranked | Worst Ranked | p.value |
| Overall | Ejumula | Naspot 12, Naspot 8, Naspot 10, Kakamega | < 2.22e-16 \*\*\* |
| Color | Naspot 8, Naspot 10, Naspot 13, Kakamega, Ejumula, Naspot 12 | Naspot 12, Ejumula, Kakamega, Naspot 13, Naspot 10, Naspot 8 | 0.39459 |
| Taste | No significant differences | No significant differences | 5.5639e-12 \*\*\* |

## Effect of covariates

Statistically significant relationship to the overall performance was found for 1 of the covariates tested. The significant relationship was found for: District (p=<2e-16)

A summary of the univariate signifance levels for all covariates that were able to be tested is shown in Table 1.2.1.

Table 1.2.1. Summary of univariate p-values for first split in Plackett-Luce tree model for the ‘overall performance’.

|  |  |
| --- | --- |
| Covariate | p.value |
| Gender | 0.924 |
| District | <2e-16 \*\*\* |

Table 1.2.2 shows which varieties where identified as the best and worst ranked in the subgroups identified by including explanatory variables into the ‘overall performance’.

Table 1.2.2. Summary of different subgroups identified by multivariate Plackett-Luce tree model.

|  |  |  |  |
| --- | --- | --- | --- |
| Split | Number of Respondents | Best Ranked | Worst Ranked |
| District %in% c(“Gulu”) | 64 | Naspot 12, Ejumula | Naspot 8, Kakamega, Naspot 13, Naspot 10 |
| District %in% c(“Soroti/Serere”, “Wakiso”) | 209 | Ejumula | Naspot 12 |

## Relationships between characteristics

Table 1.3 shows, for each characteristic in the project, the frequency with which the rankings matched with the overall performance.

The characteristic which had the strongest relationship with the overall performance was Taste. Overall the rankings for Taste matched the rankings for the overall performance 79% of the time.

The characteristic which had the weakest relationship with the overall performance was Color. Overall the rankings for Color matched the rankings for the overall performance only 0% of the time.

Table 1.3. Relationship between individual characteristics and ‘overall performance’.

|  |  |
| --- | --- |
| Variety | Complete Ranking Agreement |
| Color | -0.2% |
| Taste | 79.2% |

# Section 2: Data summary and exploratory analysis of characteristics

## Assessment of varieties

Exploratory analysis within the following section summarises results from the data directly. Given the structure of a ClimMob trail, where each participant only assessed 3 of the 6 possible varieties these results may be skewed if certain varieties were randomly assigned to face worse varieties than others. This is particularly a potential issue within a smaller trial, as due to the randomisation process the potential for an unbalanced assignment decreases as the sample size increases. Results from other sections, and in the overall summary use Plackett-Luce models (*3*), to adjust for any imbalance.

### Overall performance

Overall performance of each of the varieties is summarised in Table 2.1.

Table 2.1. Summary of overall performance.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variety | N | Top Ranked | Bottom Ranked | Net Favourability Score |
| Ejumula | 154 | 68.2% | 11.7% | 56.5 |
| Naspot 13 | 135 | 41.5% | 25.2% | 16.3 |
| Kakamega | 147 | 21.8% | 37.4% | -15.6 |
| Naspot 12 | 119 | 26.1% | 42.9% | -16.8 |
| Naspot 10 | 130 | 21.5% | 41.5% | -20.0 |
| Naspot 8 | 143 | 16.8% | 44.8% | -28.0 |

This shows the percentage of participants who assessed the varieties as the best among the 3 varieties they were provided, the percentage of participants who included the variety as their worst, the percentage of ‘head to head contests’ for which the variety won and the net favourability score. A score of +100 indicates the variety won all ‘contests’ it was involved in, a score of 0 indicates an equal number of wins and losses, a score of -100 indicates the variety lost all contests.

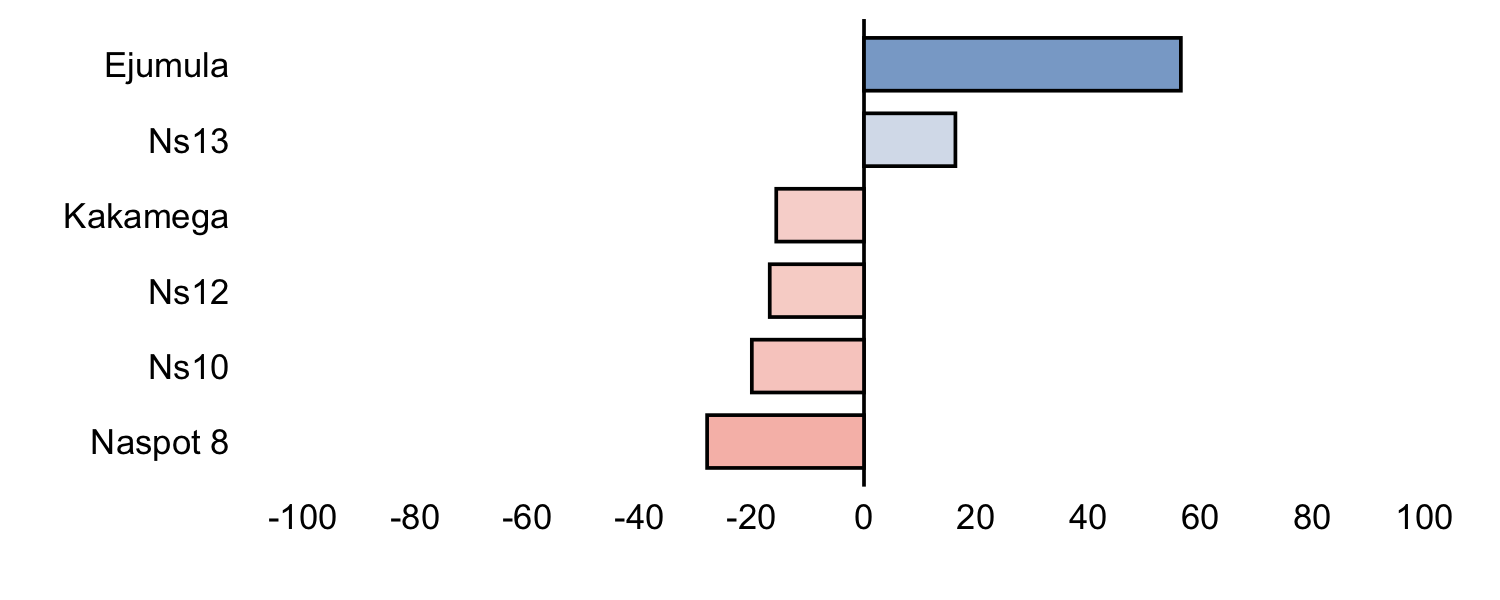


Figure 2.1. Net favourability score for ‘overall performance’.

The variety Ejumula was the ‘best’ variety overall being ranked highest by 68.2% of the 154 participants who assessed this variety.

### Other characteristics

Net favourability scores are shown below for the other characteristics assessed in this project.

**Color**

Table 2.2.1. Favourability statistics for ‘color’

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variety | N | Top Ranked | Bottom Ranked | Net Favourability Score |
| Naspot 12 | 119 | 23.5% | 37% | -13.4 |
| Kakamega | 147 | 27.2% | 32% | -4.8 |
| Ejumula | 154 | 37% | 39% | -1.9 |
| Naspot 13 | 135 | 40% | 37% | 3.0 |
| Naspot 10 | 130 | 32.3% | 26.2% | 6.2 |
| Naspot 8 | 143 | 38.5% | 28.7% | 9.8 |

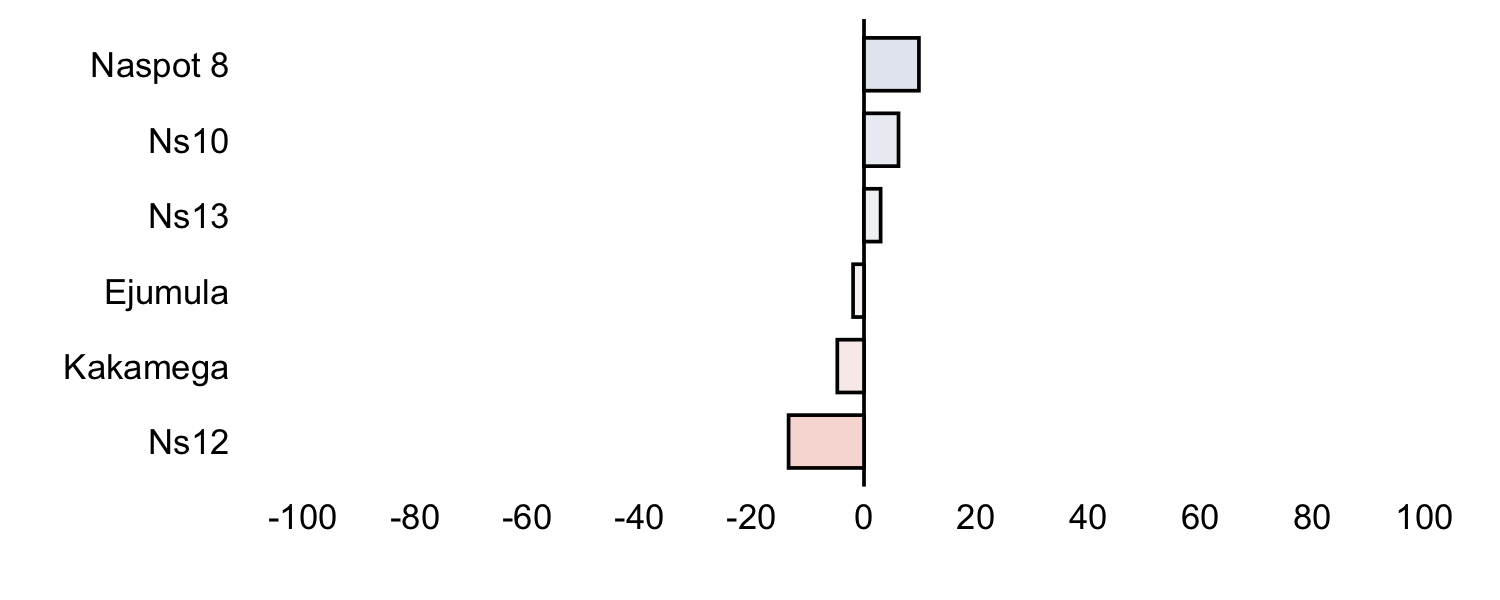


Figure 2.2.1. Net favourability score for ‘color’.

**Taste**

Table 2.2.2. Favourability statistics for ‘taste’

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variety | N | Top Ranked | Bottom Ranked | Net Favourability Score |
| Naspot 8 | 143 | 22.4% | 43.4% | -21.0 |
| Kakamega | 147 | 22.4% | 38.8% | -16.3 |
| Naspot 10 | 130 | 22.3% | 36.2% | -13.8 |
| Naspot 12 | 119 | 29.4% | 40.3% | -10.9 |
| Naspot 13 | 135 | 39.3% | 31.9% | 7.4 |
| Ejumula | 154 | 61% | 12.3% | 48.7 |

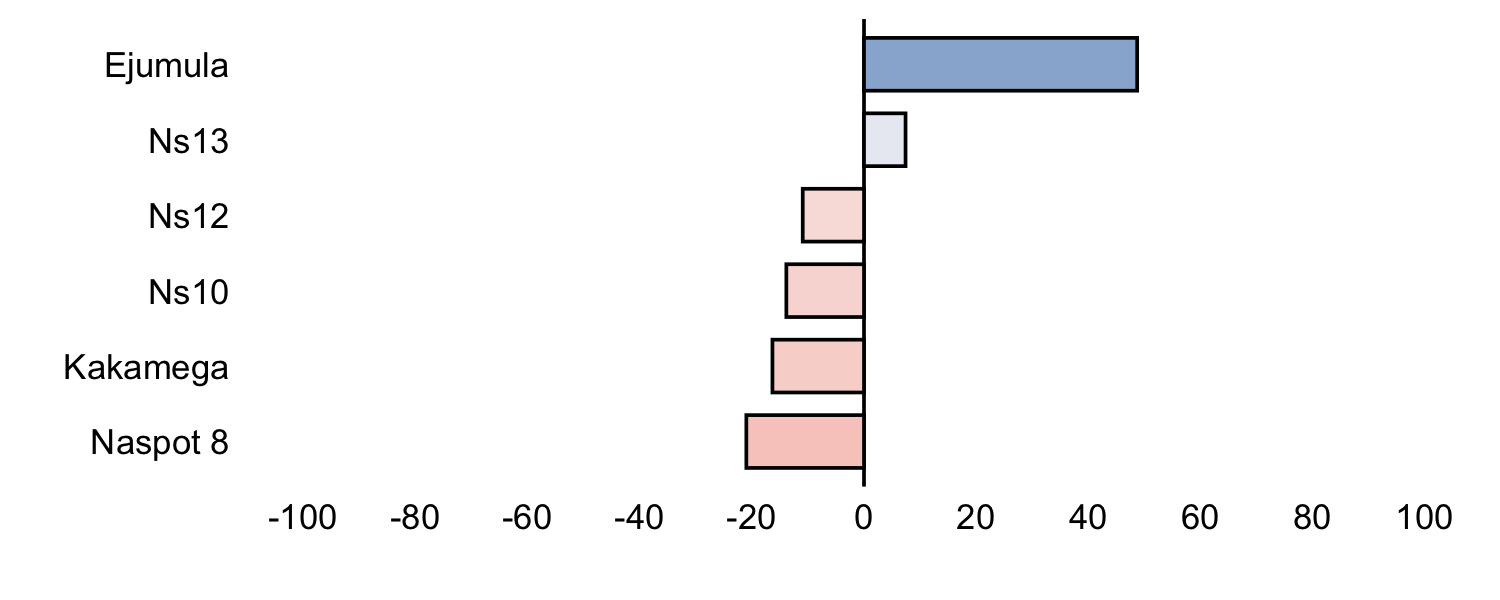


Figure 2.2.2. Net favourability score for ‘taste’.

## Pairwise contests

### Overall performance

Figure 2.3 shows the outcomes of all pairwise contests between the varieties included in the project. Each panel shows the performance of one variety against all the other varieties, and shows the percentage of the times in which the panelled variety was ranked above the other varieties shown as bars. The most winner variety is shown in the top left panel and the least winner is shown in the bottom right panel.

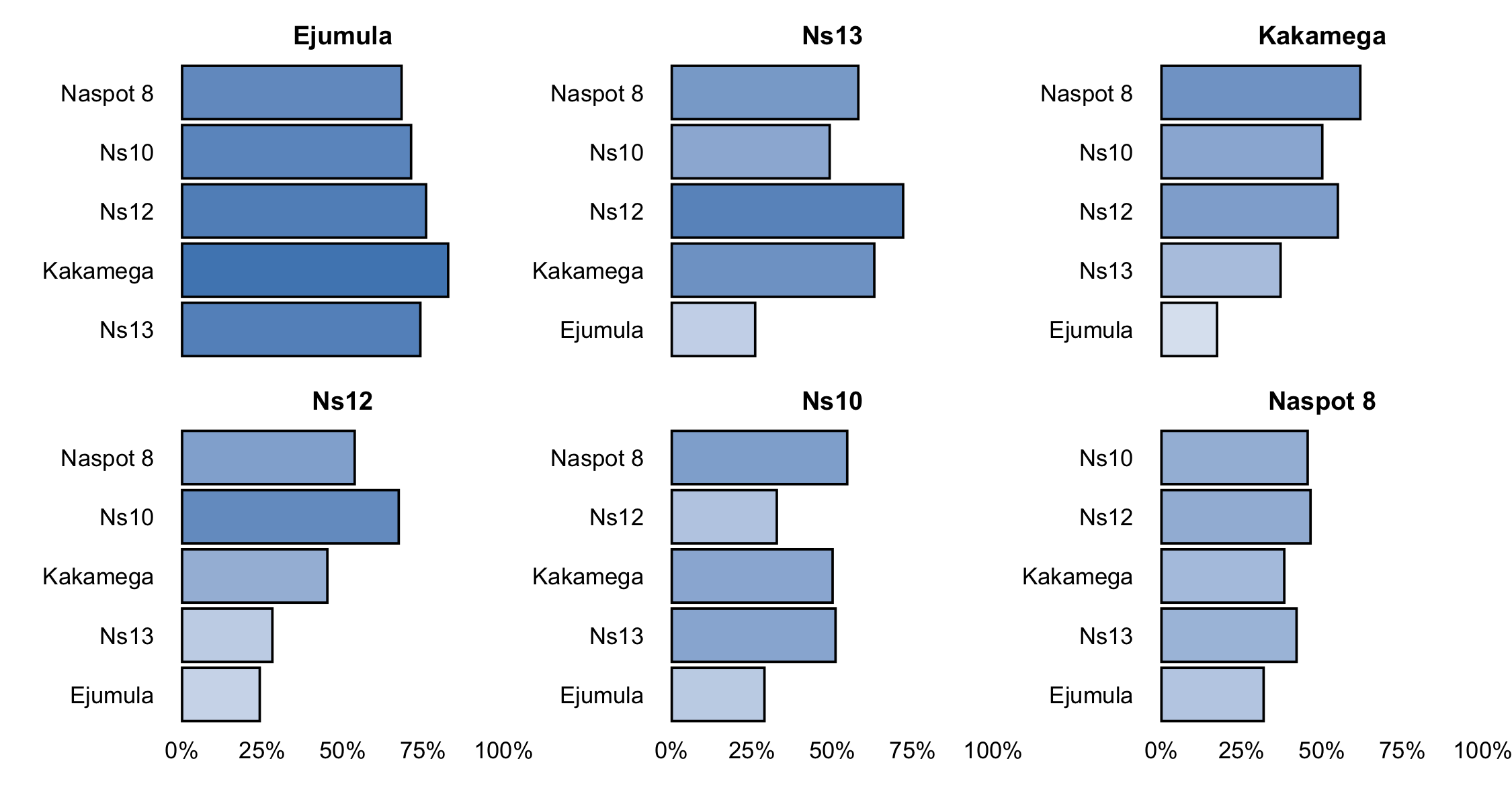
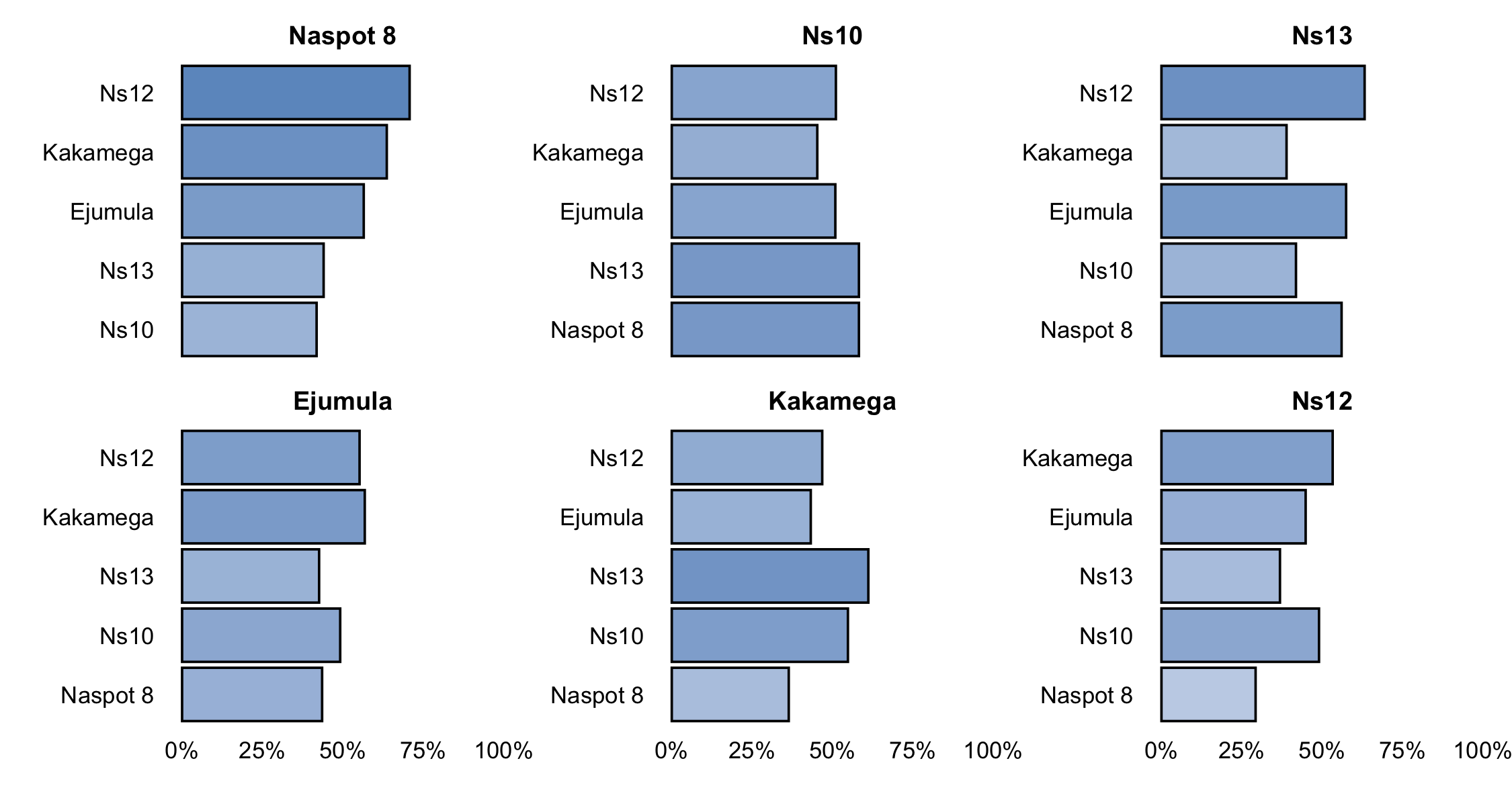


Figure 2.3. Head to head performance for ‘overall performance’.

### Other characteristics

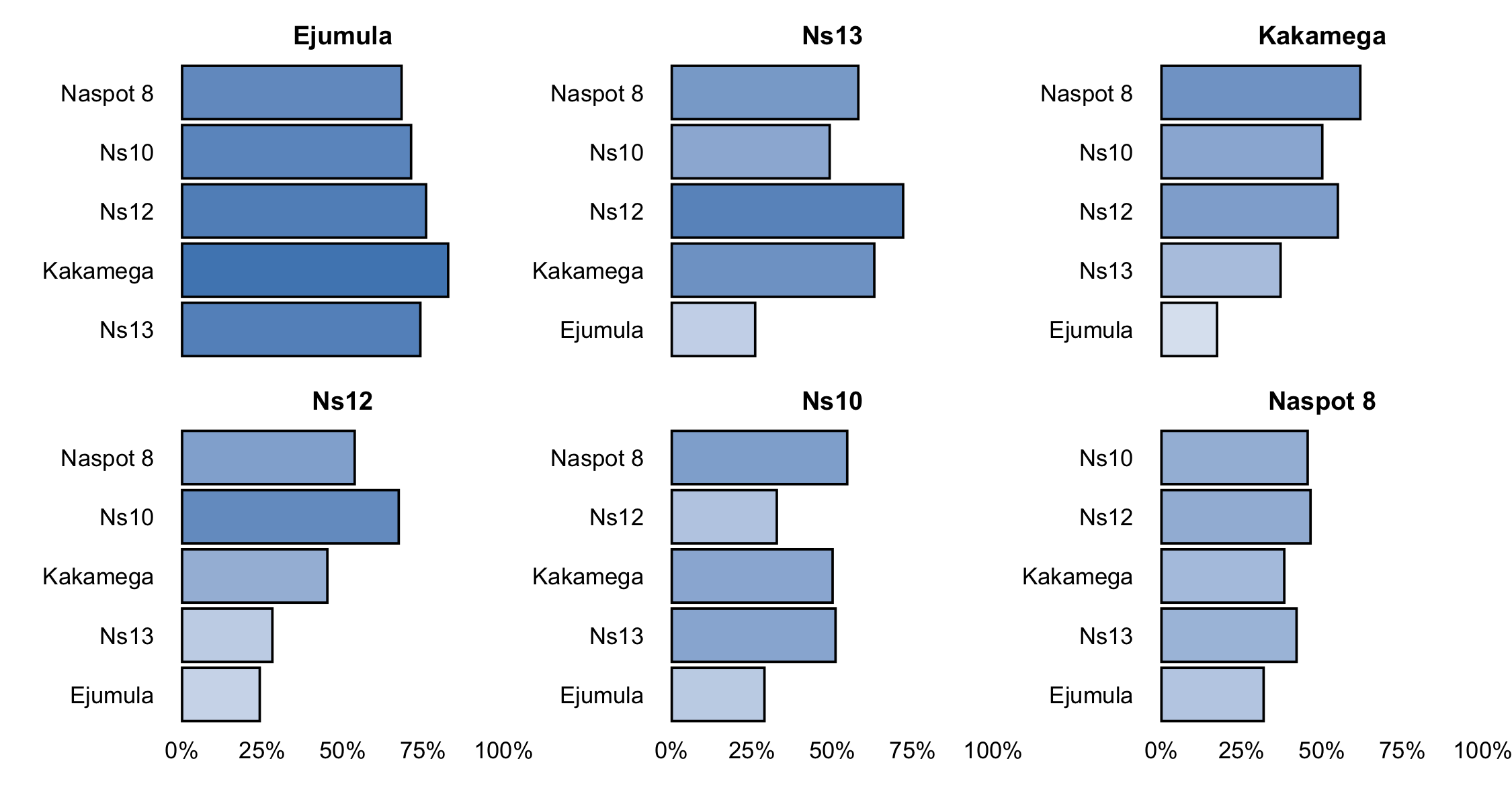
Results from the pairwise contests of the other individual characteristics assessed are shown below.

**Color**



2.4.1. Head to head performance for ‘color’.

**Taste**



2.4.2. Head to head performance for ‘taste’.

## Relationship between characteristics

Table 2.5 shows the relationship between the individual characteristic rankings and the overall performance. Complete agreement represents the percentage of participants for which the ranking of the 3 varieties in respect to the characteristic is an exact match to the overall performance. Best and worst agreement represents the percentage for which the best and worst variety for the characteristic matched the overall best and worst. Complete ranking agreement shows the proportion of correlation on the full ranking with the overall performance as baseline using the Kendall correlation coefficient (*22*).

Table 2.5. Relationship between individual characteristic and ‘overall performance’.

|  |  |  |  |
| --- | --- | --- | --- |
| Variety | Complete Ranking Agreement | Agreement with Overall Best | Agreement with Overall Worst |
| Color | -0.2% | 28.6% | 21% |
| Taste | 79.2% | 87.7% | 81.2% |

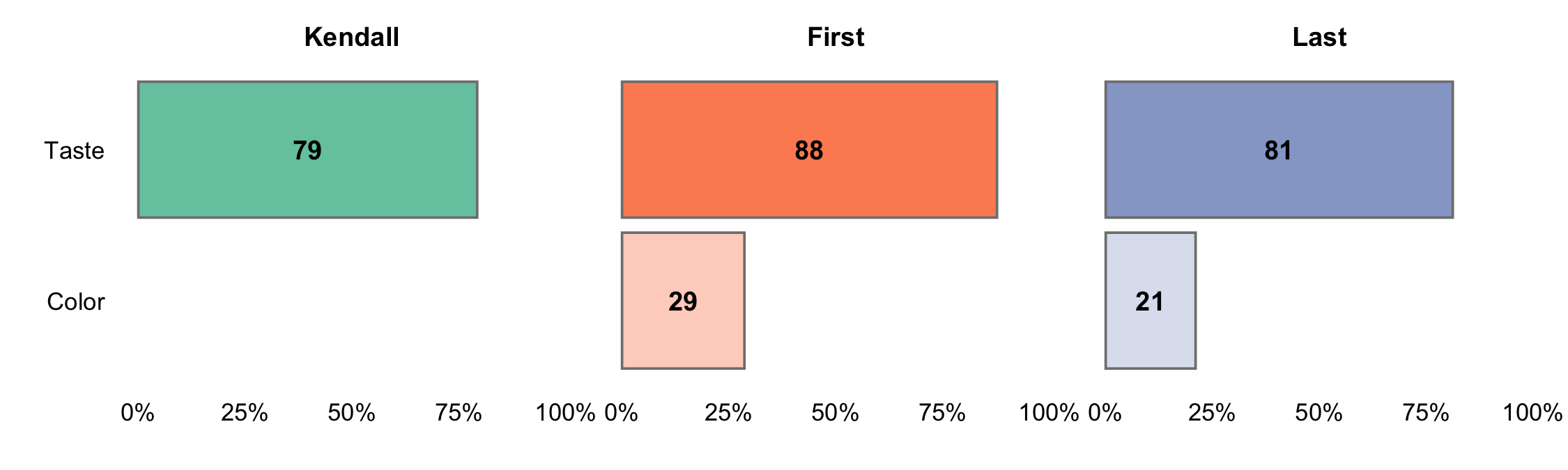


Figure 2.4. Agreement between individual characteristic and ‘overall performance’.

The characteristic which had the strongest relationship with the overall performance was Taste, with identical rankings being given as the overall performance 79.23% of the time.

Partial least squares regression was used to determine relationship between the specific characteristics and the overall performance.

The first two components recombining the specific characteristics are able to explain 97.79% of the variability in the overall performance. The dashed line represents the overall performance, with an increase in performance as the x and y increase. The variety positioned close to the dashed line will be performing equally across all characteristics; the variety positioned further away from the dashed line, on either side, will have varying performance in different characteristics. Better performance in characteristics will correspond with arrows pointing in the direction away from the dashed line and worse performance in characteristics directed on the opposite side.

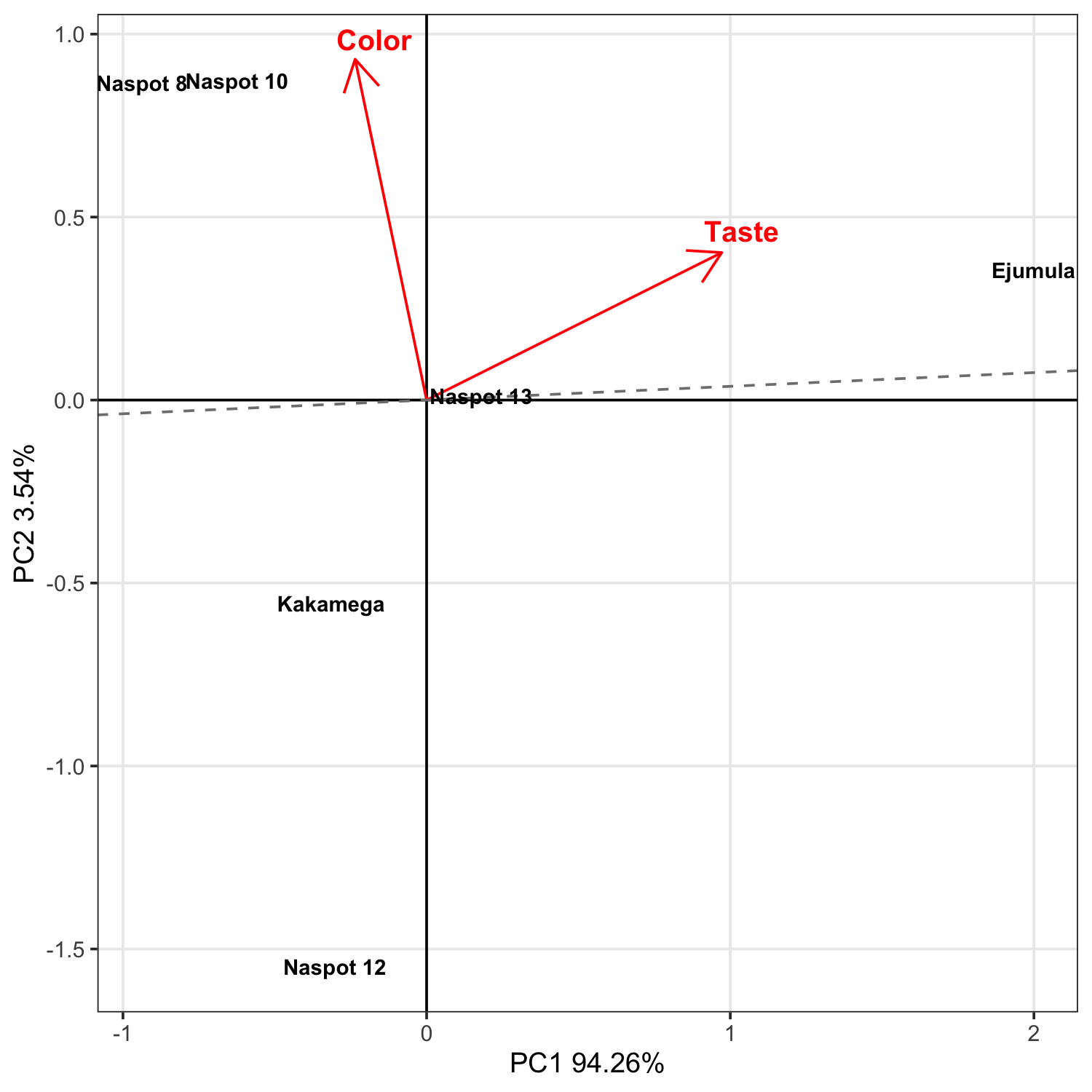


Figure 2.5. Partial least squares biplot of relationship between other characteristics and overall performance.

# Section 3: Data summary and exploratory analysis of covariates

Table 3.1. Covariates used in this project.

|  |  |
| --- | --- |
| Short name | Question |
| REG\_gender | Gender |
| district | District |

## Overall performance

Table 3.2 shows the results from the likelihood ratio test from the Plackett-Luce model for overall performance of the different varieties. The hypothesis being tested is that there is no difference in the assessments of any of the different varieties.

Table 3.2. Likelihood ratio test results from fitted Plackett-Luce model with rankings from ‘overall performance’.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| model | logLikelihood | DF | Statistic | Pr(>Chisq) |
| NULL | -494.5256 | 828 | NA | NA |
| Overall performance | -448.9137 | 823 | 91.2238 | < 2.22e-16 \*\*\* |

Figure 3.1 shows the estimates of the model coefficients with 84% confidence intervals. The purpose of this graph is to be able to best distinguish between the relative strength of each of the varieties assessed. As such the coefficient estimates themselves are not directly interpretable, but it can be concluded that a higher value for the coefficient indicates that a variety has been ranked as best more often. The 84% confidence width is chosen so that non-overlapping confidence intervals could be interpreted as indicating significant differences at the 10% significance level. This may not match exactly with the mean separation groupings, as these groupings also take into account multiple testing through the Benjamini and Hochberg adjustment.

Mean separation analysis was also conducted to indicate, using letters, which varieties are significantly more preferred than others; when varieties have at least one letter in common, there is not enough evidence from the experiment to be confident about their relative order at the 10% significance level.

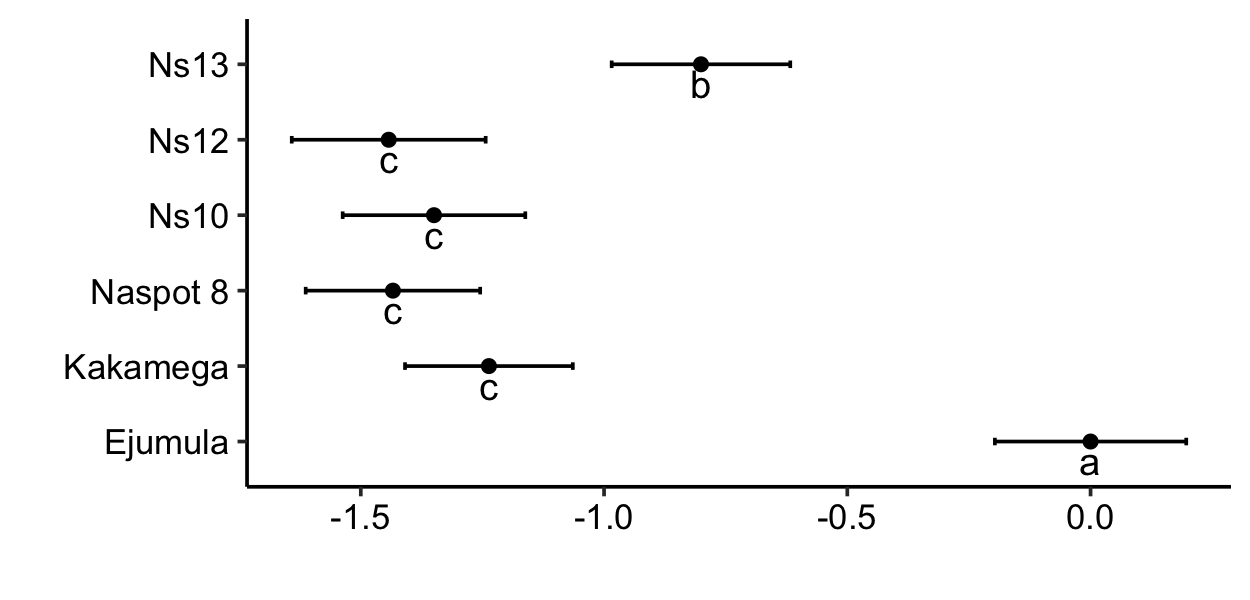


Figure 3.1. Model coefficients and mean separation of Plackett-Luce model for ‘overall performance’ with 84% confidence intervals.

The same information as Figure 3.1 is shown in Table 3.3 below.

Table 3.3. Model coefficients and mean separation of varieties with 10% level with BH adjustment.

|  |  |  |  |
| --- | --- | --- | --- |
| Variety | Estimate | quasiSE | Group |
| Ejumula | 0.0000 | 0.1400 | a |
| Naspot 13 | -0.8008 | 0.1306 | b |
| Kakamega | -1.2367 | 0.1227 | c |
| Naspot 10 | -1.3496 | 0.1336 | c |
| Naspot 8 | -1.4340 | 0.1276 | c |
| Naspot 12 | -1.4426 | 0.1418 | c |

Table 3.4 and Figure 3.2 use the coefficients from the Plackett-Luce model to estimate the probability of each variety being considered to be the top ranked variety in a direct comparison between all of the possible varieties.

Table 3.4. Percentage probability of being the best ranked for ‘overall performance’.

|  |  |
| --- | --- |
| Variety | Win probability |
| Ejumula | 40.4% |
| Naspot 13 | 18.2% |
| Kakamega | 11.7% |
| Naspot 10 | 10.5% |
| Naspot 8 | 9.6% |
| Naspot 12 | 9.6% |

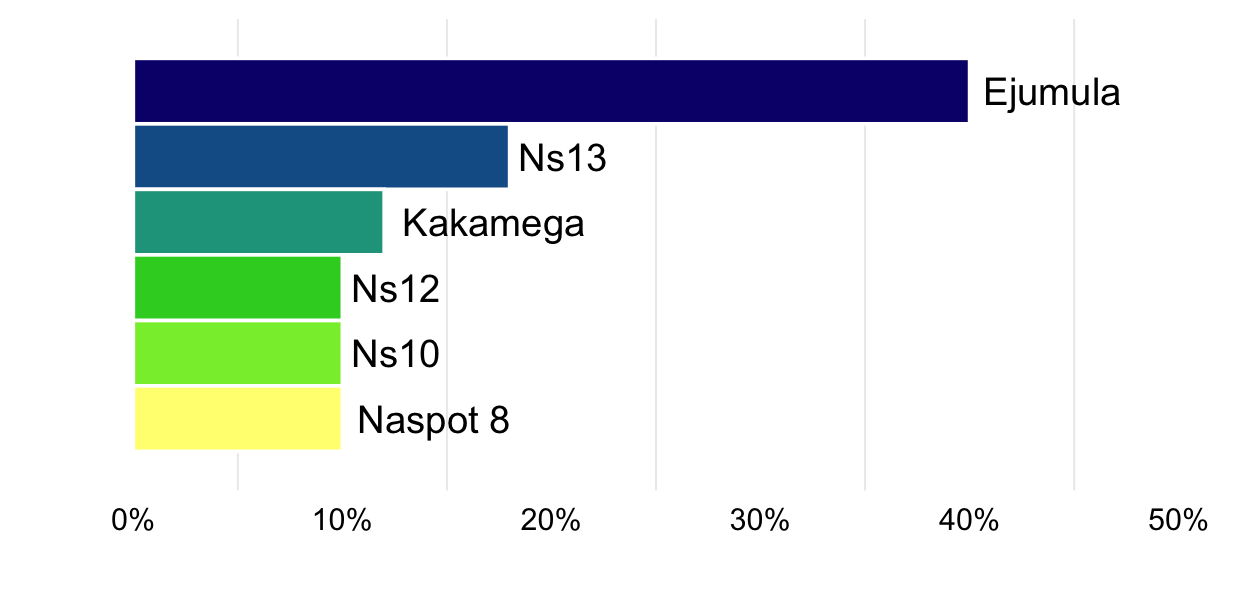


Figure 3.2. Probability of being the best ranked for ‘overall performance’.

# Section 4: Plackett-Luce models of other characteristics

**Color**

Table 4.1.1. Likelihood ratio test results from fitted Plackett-Luce model with rankings from ‘color’.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| model | logLikelihood | DF | Statistic | Pr(>Chisq) |
| NULL | -494.5256 | 828 | NA | NA |
| Color | -491.9368 | 823 | 5.177607 | 0.39459 |

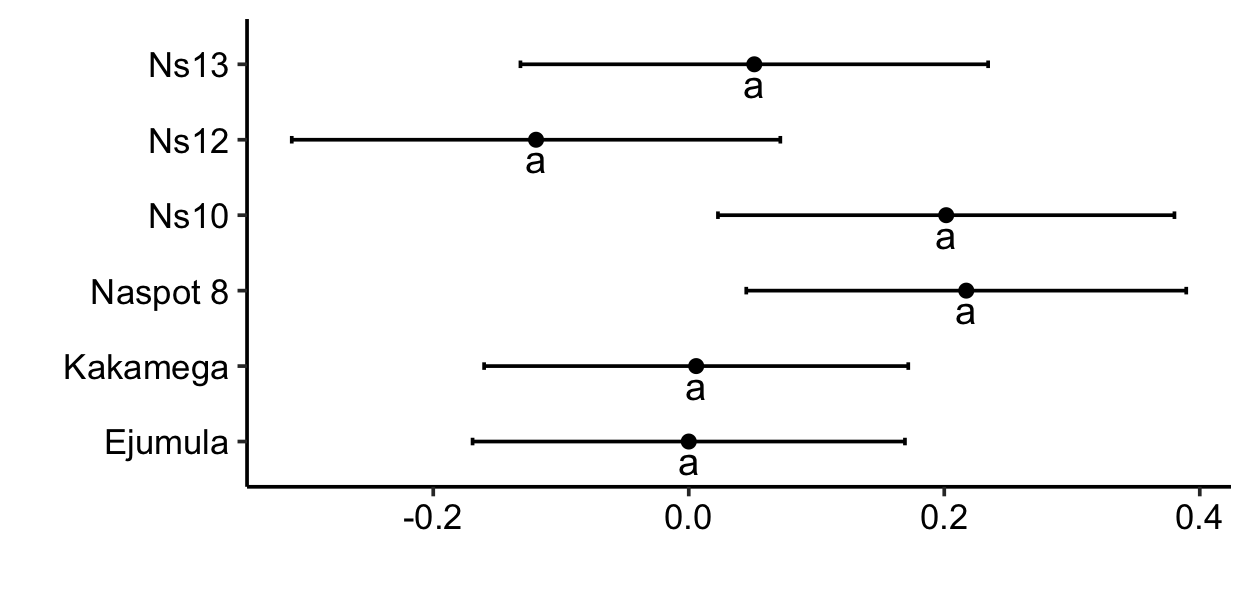


Figure 4.1.1. Model coefficients and mean separation of Plackett-Luce model for ‘color’ with 84% confidence intervals.

Table 4.1.2. Model coefficients and mean separation of Plackett-Luce model for ‘color’ with 84% confidence intervals.

|  |  |  |  |
| --- | --- | --- | --- |
| Variety | Estimate | quasiSE | Group |
| Naspot 8 | 0.2172 | 0.1226 | a |
| Naspot 10 | 0.2015 | 0.1272 | a |
| Naspot 13 | 0.0513 | 0.1303 | a |
| Kakamega | 0.0058 | 0.1181 | a |
| Ejumula | 0.0000 | 0.1204 | a |
| Naspot 12 | -0.1195 | 0.1361 | a |

Table 4.1.3 Percetage probability of being highest ranked for ‘color’.

|  |  |
| --- | --- |
| Variety | Win probability |
| Naspot 8 | 19.4% |
| Naspot 10 | 19.1% |
| Naspot 13 | 16.4% |
| Kakamega | 15.7% |
| Ejumula | 15.6% |
| Naspot 12 | 13.8% |

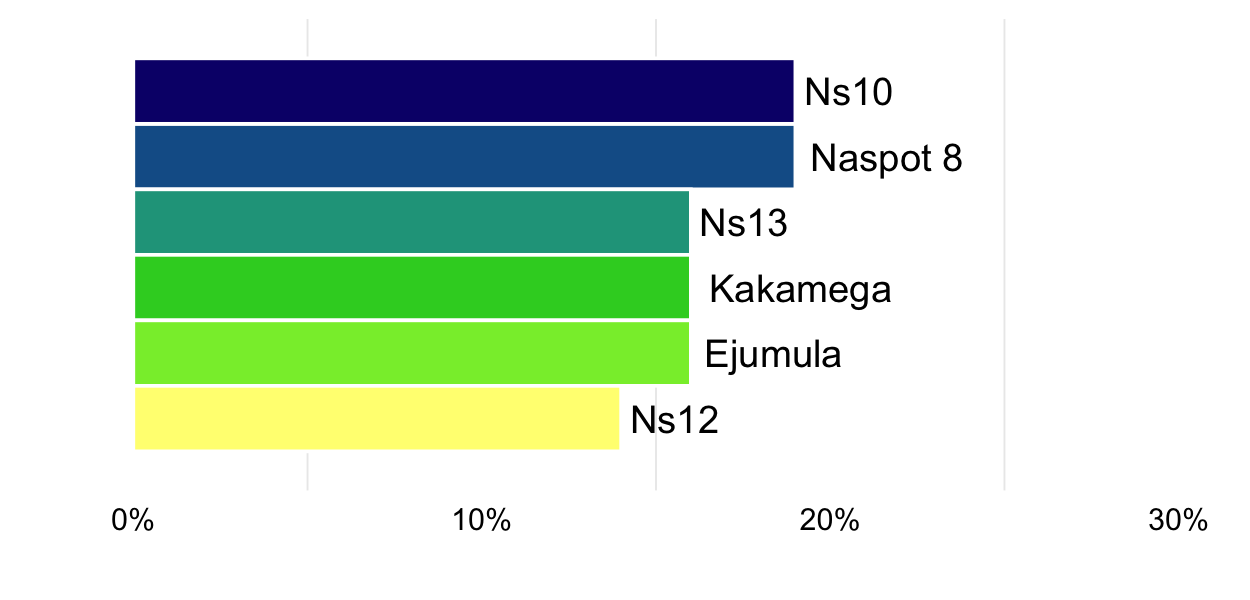


Figure 4.1.2. Probability of being the best ranked for ‘color’.

**Taste**

Table 4.2.1. Likelihood ratio test results from fitted Plackett-Luce model with rankings from ‘taste’.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| model | logLikelihood | DF | Statistic | Pr(>Chisq) |
| NULL | -494.5256 | 828 | NA | NA |
| Taste | -463.7050 | 823 | 61.64116 | 5.5639e-12 \*\*\* |

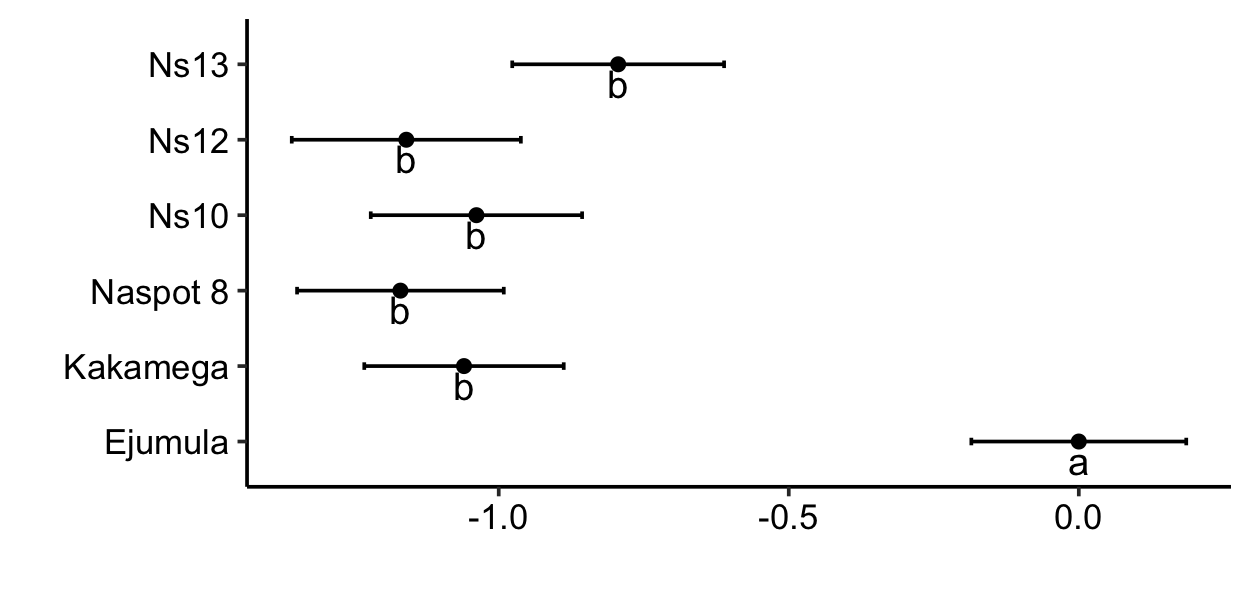


Figure 4.2.1. Model coefficients and mean separation of Plackett-Luce model for ‘taste’ with 84% confidence intervals.

Table 4.2.2. Model coefficients and mean separation of Plackett-Luce model for ‘taste’ with 84% confidence intervals.

|  |  |  |  |
| --- | --- | --- | --- |
| Variety | Estimate | quasiSE | Group |
| Ejumula | 0.0000 | 0.1318 | a |
| Naspot 13 | -0.7941 | 0.1300 | b |
| Naspot 10 | -1.0384 | 0.1297 | b |
| Kakamega | -1.0599 | 0.1224 | b |
| Naspot 12 | -1.1593 | 0.1405 | b |
| Naspot 8 | -1.1695 | 0.1268 | b |

Table 4.2.3 Percetage probability of being highest ranked for ‘taste’.

|  |  |
| --- | --- |
| Variety | Win probability |
| Ejumula | 36% |
| Naspot 13 | 16.3% |
| Naspot 10 | 12.7% |
| Kakamega | 12.5% |
| Naspot 12 | 11.3% |
| Naspot 8 | 11.2% |

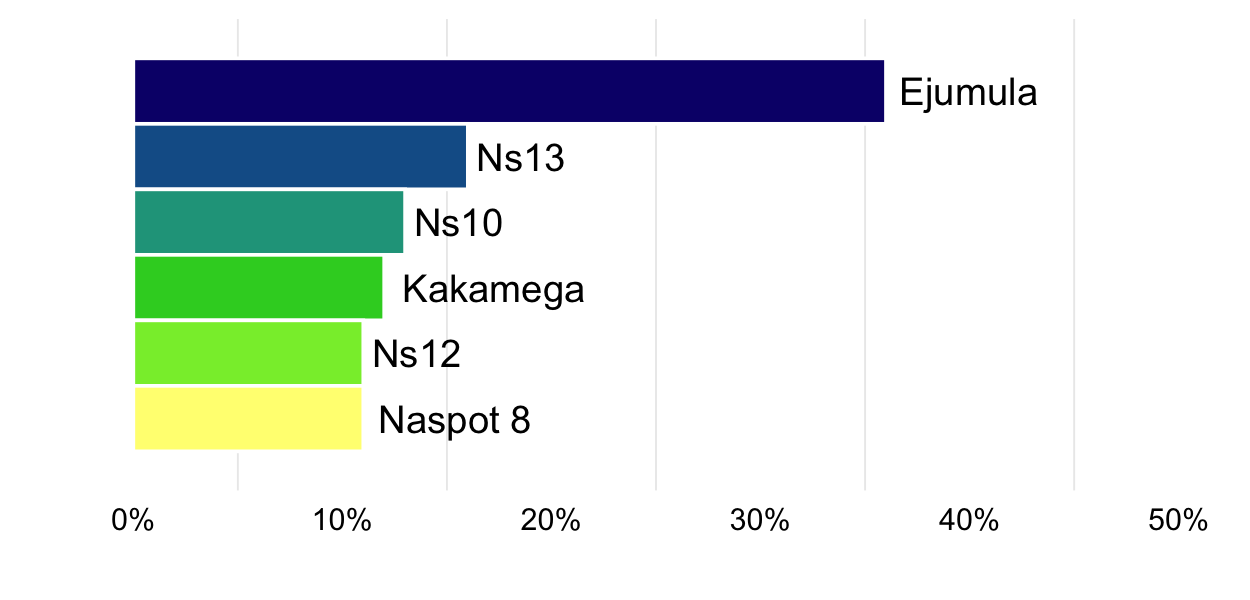


Figure 4.2.2. Probability of being the best ranked for ‘taste’.

# Section 5: Plackett-Luce models with covariates

## Overall performance

A model-based recursive partitioning method (*23*) was used to determine which of the explanatory variables, if any, had significant relationships with the rankings. This approach identifies sub-groups in the data for which the rankings of the different varieties are significantly different to each other. Table 5.1 shows the p-values for each of the covariates tested, one-by-one, showing whether or not the covariate could be used to define sub-groups with significantly different rankings.

Table 5.1. Univariate p-values for first split in Plackett-Luce tree model for the overall ranking.

|  |  |
| --- | --- |
| Covariate | p.value |
| Gender | 0.924 |
| District | <2e-16 \*\*\* |

Figure 5.1 shows the partitioning of the rankings based on the most significantly different sub-groups which could be identified from the data using a 10% significance level. At the top of the tree is the full dataset, then working down through the different levels of the tree shows the combinations of variables which define each subgroup. The model parameters are shown for the final subgroups (“terminal nodes”) in the plots at the bottoms of the tree. The model quasi-variance estimates, along with 84% confidence intervals are provided. This will help identification of which varieties were better suited to particular sub-groups identified by the analysis.

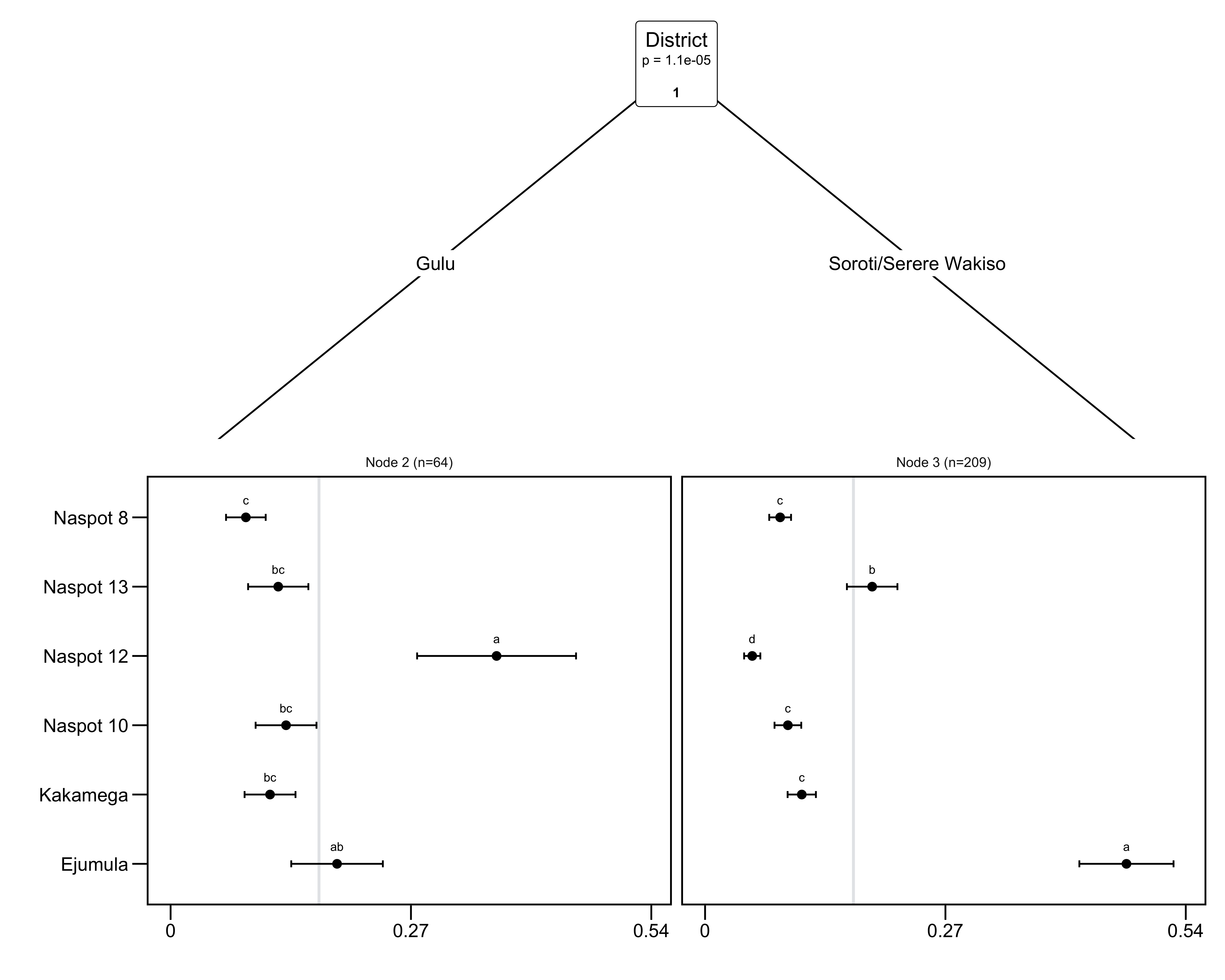


Figure 5.1. Plackett-Luce tree for the overall performance. The horizontal axis is the probability of winning. Error bars show quasi-SEs. The gray vertical line indicates the average probability of winning (1/number of items).

The highest and lowest performing variety within each sub-group is identified within Table 5.2.

Table 5.2. Summary of performance in each tree node.

|  |  |  |  |
| --- | --- | --- | --- |
| Split | Number of Respondents | Best Ranked | Worst Ranked |
| District %in% c(“Gulu”) | 64 | Naspot 12, Ejumula | Naspot 8, Kakamega, Naspot 13, Naspot 10 |
| District %in% c(“Soroti/Serere”, “Wakiso”) | 209 | Ejumula | Naspot 12 |

Table 5.3 outlines the p-values for each covariate at each of the nodes in the tree, outlining whether an additional significant split could be determined from within the existing sub-group at that node.

Table 5.3. Goodnesss-of-fit for the effects of covariate(s) at each node in the Plackett-Luce tree.

|  |  |  |
| --- | --- | --- |
| Covariate | Node | p.value |
| Gender | 1 | 0.924 |
| **District** | **1** | **<2e-16** \*\*\* |
| Gender | 3 | 0.712 |
| District | 3 | 0.702 |

# References

1. J. van Etten, E. Beza, L. Calderer, K. Van Duijvendijk, C. Fadda, B. Fantahun, Y. G. Kidane, J. van de Gevel, A. Gupta, D. K. Mengistu, D. Kiambi, P. N. Mathur, L. Mercado, S. Mittra, M. J. Mollel, J. C. Rosas, J. Steinke, J. G. Suchini, K. S. Zimmerer, First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Experimental Agriculture*. **55**, 275–296 (2019).

2. R Core Team, R: A language and environment for statistical computing. version 3.6.2. (2019), (available at <https://r-project.org/>).

3. H. L. Turner, J. van Etten, D. Firth, I. Kosmidis, Modelling rankings in R: the PlackettLuce package. *Computational Statistics* (2020), doi:[10.1007/s00180-020-00959-3](https://doi.org/10.1007/s00180-020-00959-3).

4. Y. Xie, *Dynamic documents with R and knitr* (Chapman; Hall/CRC, Boca Raton, Florida, ed. 2nd, 2015; <https://yihui.org/knitr/>).

5. Y. Xie, J. J. Allaire, G. Grolemund, *R markdown: The definitive guide* (Chapman; Hall/CRC, Boca Raton, Florida, 2018; <https://bookdown.org/yihui/rmarkdown>).

6. K. de Sousa, J. van Etten, ClimMobTools: API client for the ’ClimMob’ platform (2020), (available at <https://CRAN.R-project.org/package=ClimMobTools>).

7. K. de Sousa, J. van Etten, S. Dumble, N. Greliche, J. Steinke, gosset: Modelling metadata and crowdsourced citizen science (2020), (available at <https://agrobioinfoservices.github.io/gosset/>).

8. G. R. Warnes, B. Bolker, T. Lumley, gtools: Various r programming tools (2018), (available at <https://CRAN.R-project.org/package=gtools>).

9. J. Ooms, The jsonlite package: A practical and consistent mapping between json data and r objects. *arXiv:1403.2805 [stat.CO]* (2014) (available at <https://arxiv.org/abs/1403.2805>).

10. A. Zeileis, T. Hothorn, K. Hornik, Model-based recursive partitioning. *Journal of Computational and Graphical Statistics*. **17**, 492–514 (2008).

11. J. Kopf, C. Strobl, A. Zeileis, Anchor selection strategies for DIF analysis: Review, assessment, and new approaches. *Educational and Psychological Measurement*. **75**, 22–56 (2015).

12. D. Firth, qvcalc: Quasi variances for factor effects in statistical models (2020), (available at <https://CRAN.R-project.org/package=qvcalc>).

13. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal*. **Complex Systems**, 1695 (2006).

14. M. Borkovec, N. Madin, *ggparty: ’Ggplot’ visualizations for the ’partykit’ package* (2019; <https://CRAN.R-project.org/package=ggparty>).

15. H. Wickham, *ggplot2: Elegant graphics for data analysis* (Springer-Verlag New York, 2016; <https://ggplot2.tidyverse.org>).

16. K. Slowikowski, *ggrepel: Automatically position non-overlapping text labels with ’ggplot2’* (2020; <https://CRAN.R-project.org/package=ggrepel>).

17. T. Appelhans, F. Detsch, C. Reudenbach, S. Woellauer, *mapview: Interactive viewing of spatial data in r* (2020; <https://CRAN.R-project.org/package=mapview>).

18. S. Graves, H.-P. Piepho, L. S. with help from Sundar Dorai-Raj, multcompView: Visualizations of paired comparisons (2019), (available at <https://CRAN.R-project.org/package=multcompView>).

19. T. L. Pedersen, *patchwork: The composer of plots* (2019; <https://CRAN.R-project.org/package=patchwork>).

20. B.-H. Mevik, R. Wehrens, K. H. Liland, pls: Partial least squares and principal component regression (2019), (available at <https://CRAN.R-project.org/package=pls>).

21. J. van Etten, C. Quirós, K. de Sousa, J. Steinke, S. Dumble, B. Madriz, A. Coto, J. van de Gevel, N. Greliche, N. Murchan, J. Hook, ClimMob: Software to support experimental citizen science in agriculture (2020), (available at <https://climmob.net/climmob3>).

22. M. G. Kendall, A new measure of ranking correlation. *Biometrika*. **30**, 81–93 (1938).

23. C. Strobl, J. Malley, G. Tutz, An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological methods*. **14**, 323–48 (2009).