

Novel insights in browse preferences in proboscis monkeys.

2022-12-01

Overview

Table 1: Overview Table x.xx

ID	Shannons	Simpsons	FNB	Ketapung	Total	Age	Sex
Anggun	1.079	0.586	0.180	43	363	3	F
Jeff	0.705	0.419	0.110	168	343	13	M
Kejora	0.719	0.401	0.105	47	330	3	F
Malaka	1.017	0.529	0.166	149	618	14	F
Sakti	0.923	0.518	0.163	135	397	23	F
Sayang	0.530	0.307	0.069	128	340	12	F
Wendy	1.165	0.608	0.188	161	656	12	F

Data collection:

- Consumption is determined by bite counts per individual browse species by individual monkey
- Order of observed monkey is random but split evenly across stipulated sampling duration ($n = 40$ days)
- Observations were only conducted during the ‘first’ feeding session in the morning (8.30 am to 9 am) and within the first 30 minutes.
- All browses provided were fresh in quality
- Browse provision (in terms of species offered) was random except for the 100% provision of ketapung and miracle (see Figure 1); species within browse bundles are randomly assigned.
- Missing data were imputed with the in-built function of random forest algorithms.
- Browse preference is “inferred” from importance (via random forest models) within diversity and consumption indexes.
- Diversity indexes ($n = 3$; Figures 3–5) used are: Shannons, Simpsons and Food niche breadth.
- Consumption indexes ($n = 2$; Figures 6 and 7) used are: total leaves consumed (regardless of species) and ketapung species only.

Spreads/Distribution

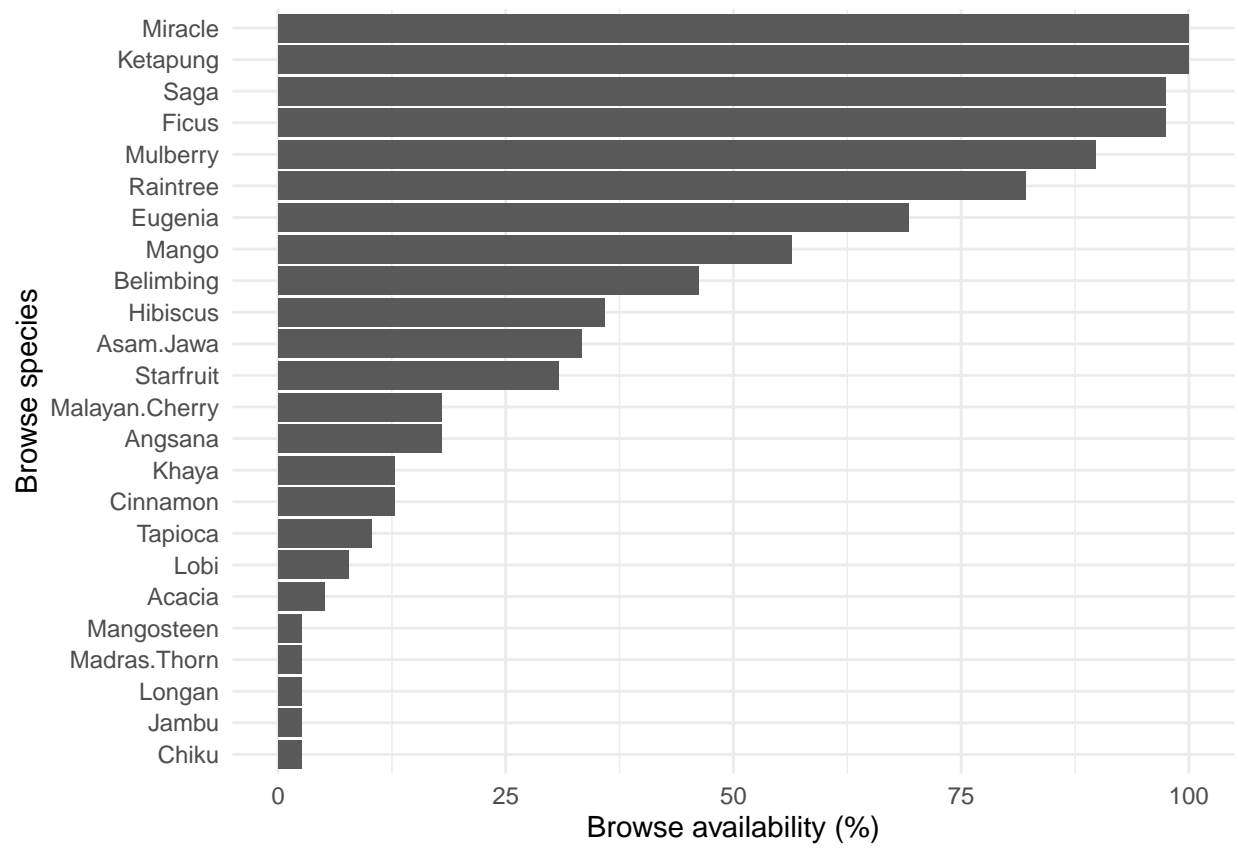


Figure 1: The spread (availability) of browse distribution across the sampling period (n = 39 days/sessions).

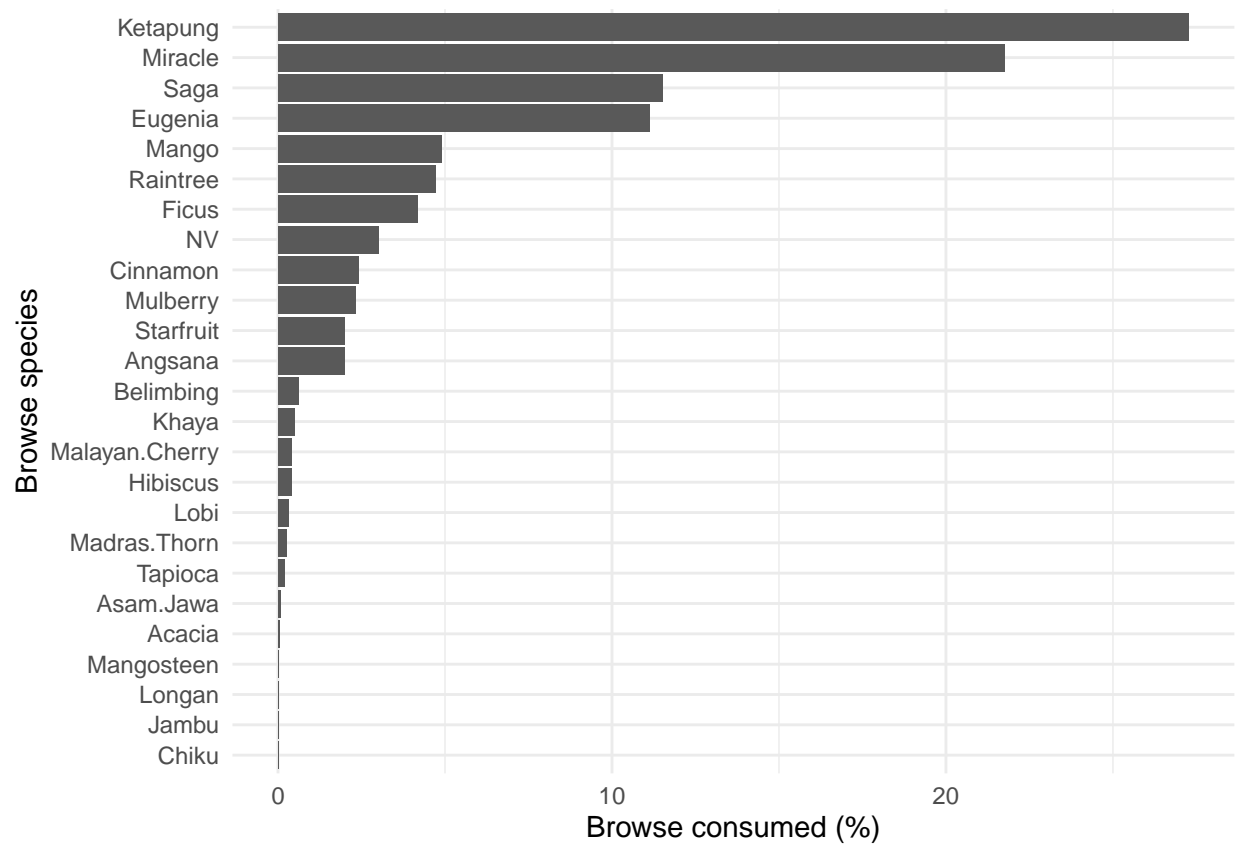


Figure 2: The spread of browse consumption across the sampling period (n = 39 days/sessions)

Variable importance of each browse by indexes

Some notes on RF model importance:

- All models are regression models.
- Negative value predictors have no influence on response variable (i.e., diversity and consumption index; see Figure 3–7)
- Residuals are an indicator of prediction error rates; lower residuals are better
- Importance on the x-axis determines %change in prediction; for e.g., removing miracle would result in a 10% decrease in accuracy in predicting H-index, see Figure 3
- Thus, greater importance means greater weight in improving the response variable (e.g., increase in H / D / FNC / Consumption index).
- Node purity decides variance in the removal of said predictor; at the moment, not really important.

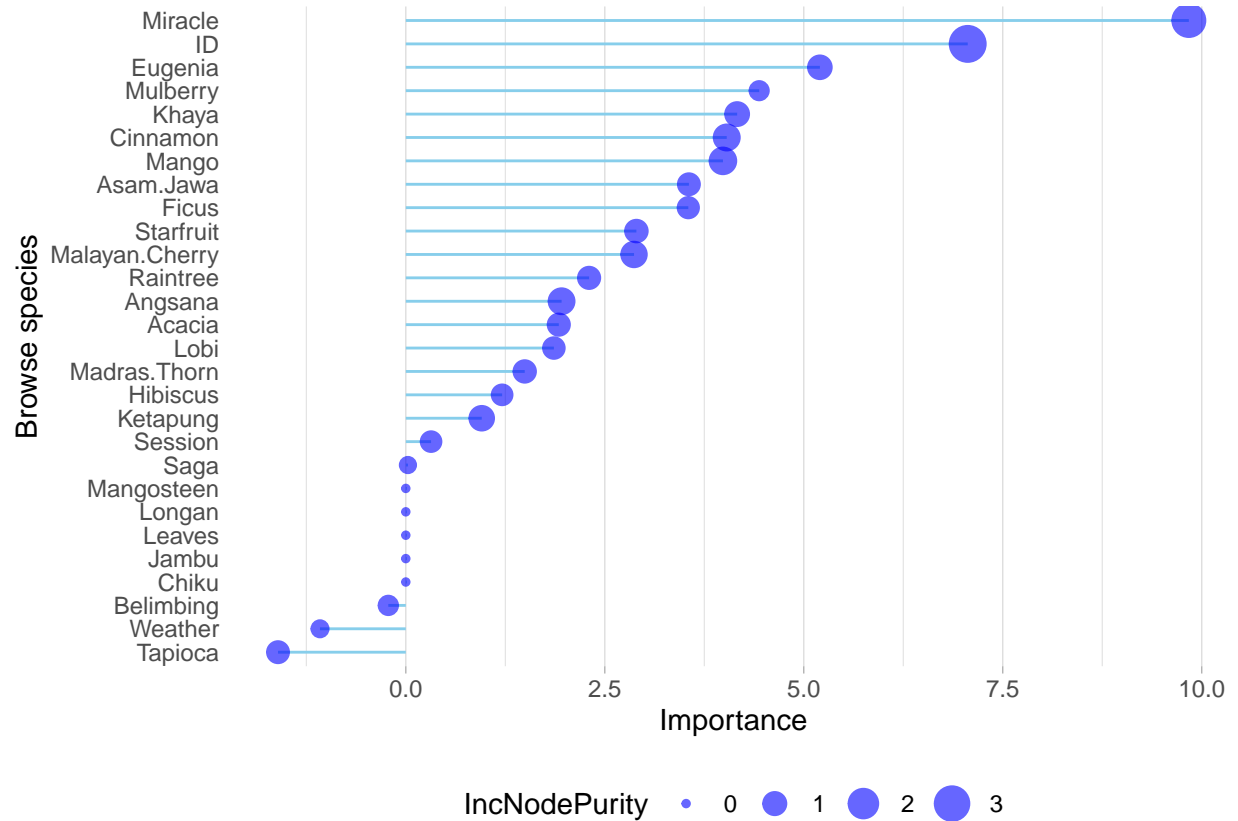


Figure 3: Variable importance plot by H-index / Shannons. Residuals: 0.18.

My interpretation / results:

- Across diversity and consumption indexes, ID (i.e., id of individual monkey) continually holds a high importance in each model; inter-individual variability plays a huge role.

- Availability has confounding influence on consumption; miracle is always provided and always consumed (hence high on the importance chart) but not the same for Saga (Figure 3–5).
- The consumption of ketapung is interesting (see Figure 7); the inclusion of high importance species (positive values: e.g., khaya, angšana, lobi; see Figure 7) results in a greater consumption of ketapung
- conversely, the inclusion of low importance species (negative values: miracle, ficus; see Figure 7) would then lead to a decrease in ketapung consumed.

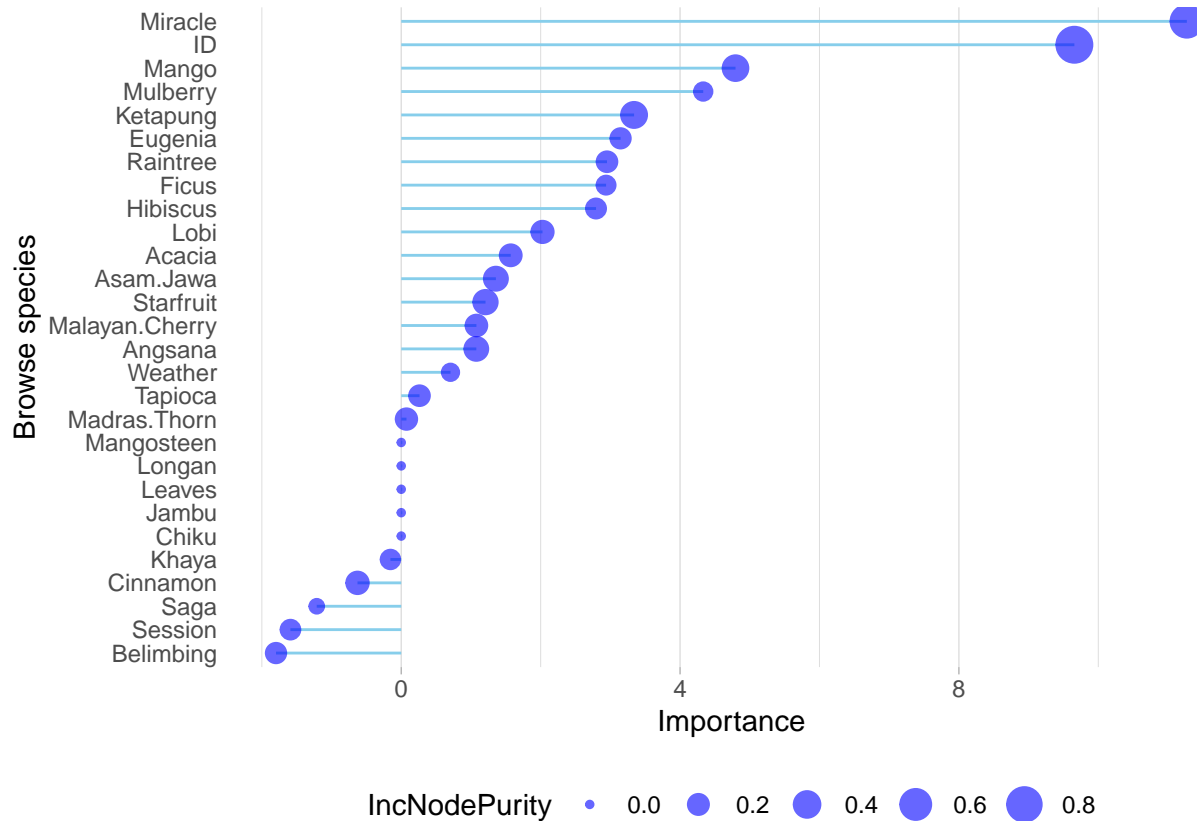


Figure 4: Variable importance plot by D-index / Simpsons. Residuals: 0.05.

Moving forward:

- I’ve discussed with Dr Shangz about the scope of the work and we feel that more can be done if nutritional composition is included. Here, I proposed PCA to examine how the ranks of certain browse species (when grouped) explains their importance in improving the diversity and/or consumption index.
- On a secondary note, the identification of individual variability in the current work is not surprising, and can be further explored / written as a manuscript? We both believe it has some relevance to other institutes/collection. For e.g, other zoo managers will need to do their own browse preference studies to properly identify what “works best” for their collection.
- That said, I am unsure if I can write this up on my own and I would gladly accept any form of help on this, if possible.
- I am also accepting any feedback on the research done today. There are two more days of data to analyse (i.e., 1st and 2nd Dec) and I’m waiting for Geraldene to get back to me on those observation.

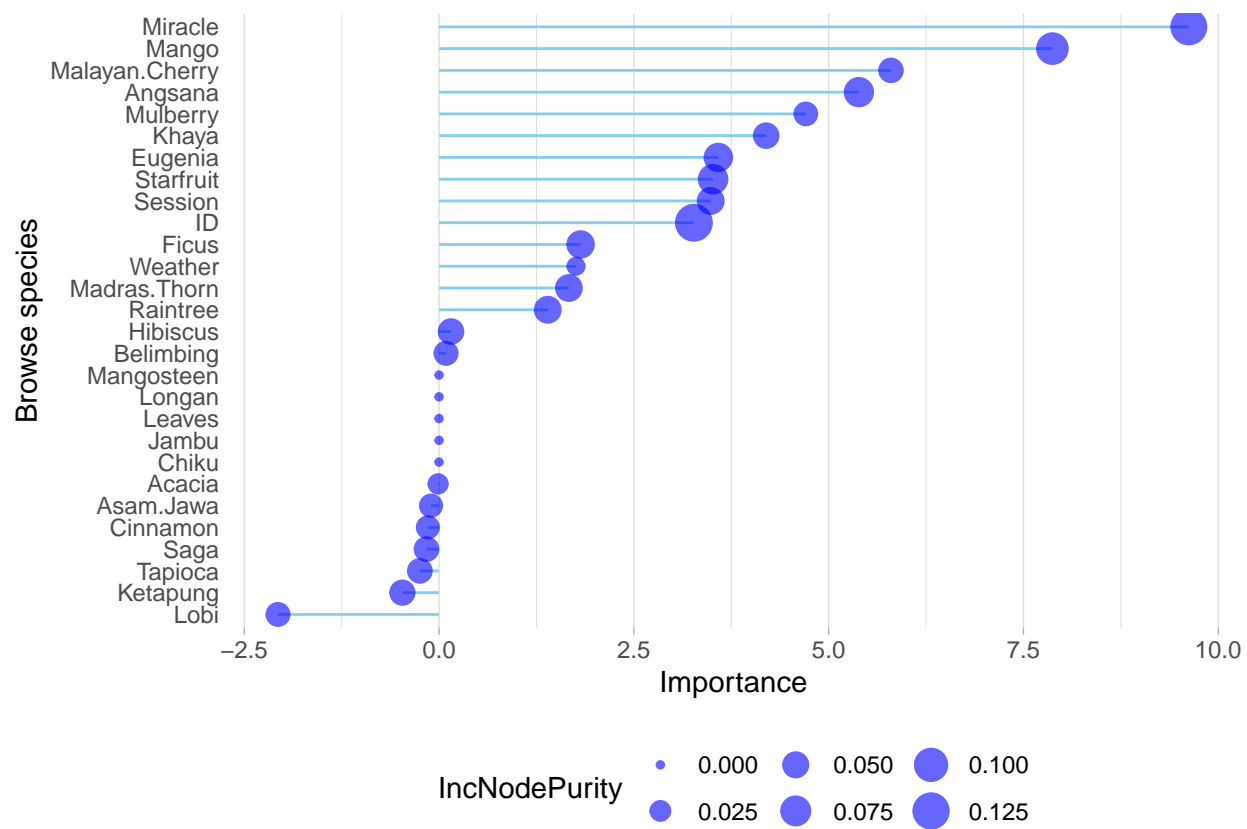


Figure 5: Variable importance plot by Food Niche Breadth. Residuals: 0.01.

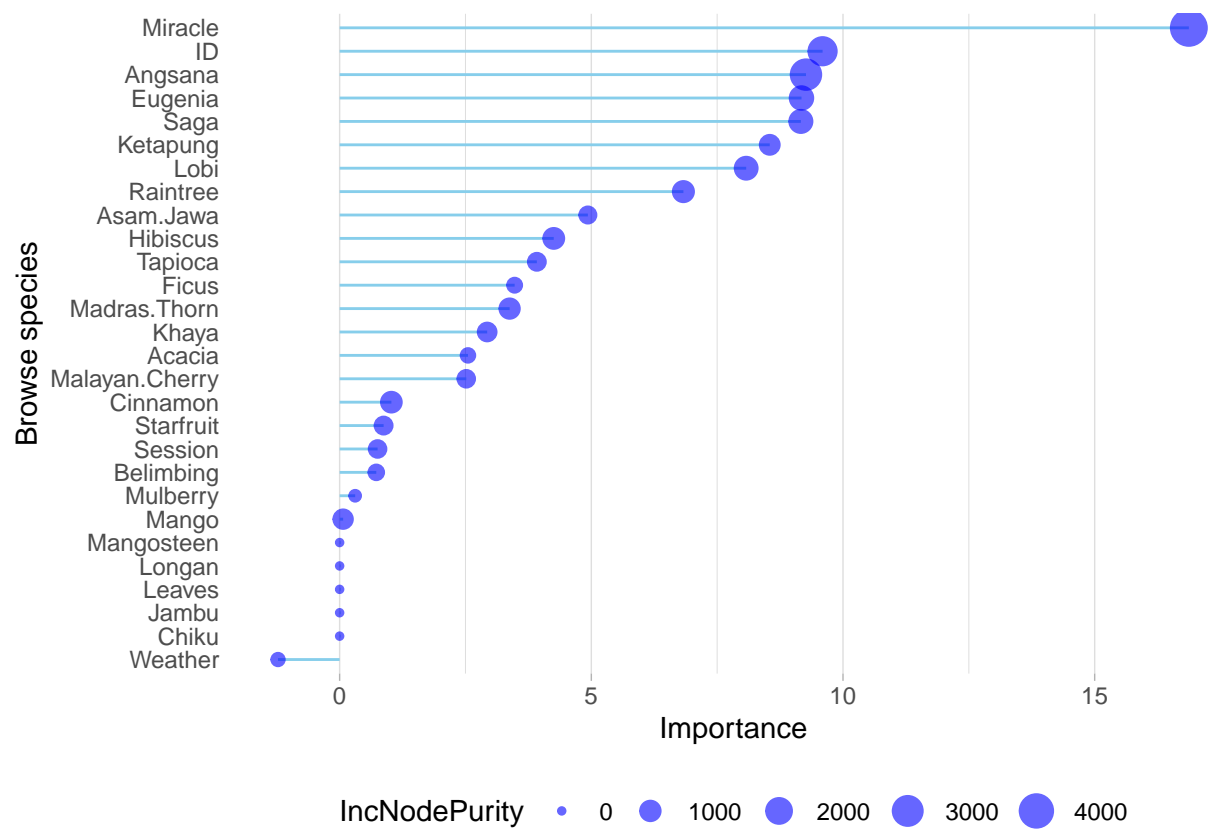


Figure 6: Variable importance plot by total consumption. Residuals: 119

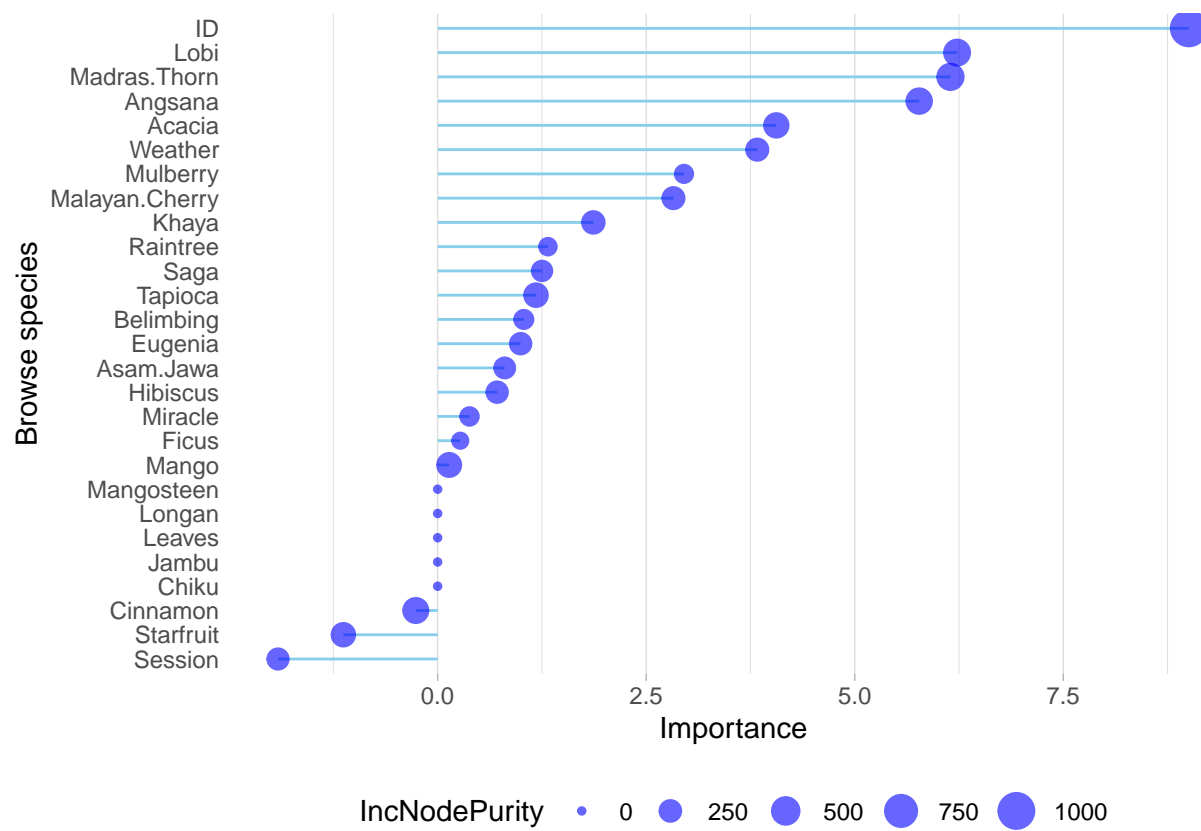


Figure 7: Variable importance plot by Ketapung consumption. Residuals: 57

Date: 2nd December 2022; without the mass data imputation

```
## boundary (singular) fit: see help('isSingular')
## boundary (singular) fit: see help('isSingular')
## boundary (singular) fit: see help('isSingular')
## boundary (singular) fit: see help('isSingular')

## Model selection table
##              (Int)      H      D  fnb brw.siz      Ktp      Sss df   logLik  AICc
## bundle.FNB      26.84      H      D  50.65 -0.9048      Ktp      Sss df   logLik  AICc
## FNB.only        19.49      D      49.20      Sss df   logLik  AICc
## bundle.H        30.80 12.32      -1.8200      Sss df   logLik  AICc
## bundle.D        28.69      21.38      -1.5050      Sss df   logLik  AICc
## H.only          17.62 10.02      Sss df   logLik  AICc
## D.only          17.82      17.85      Sss df   logLik  AICc
## session.only    30.67      -1.166      Sss df   logLik  AICc
## ktp.only        23.82      0.3523      Sss df   logLik  AICc
## null.model      26.39      Sss df   logLik  AICc
## bundle.ktp      27.00      -0.3715 0.3421      Sss df   logLik  AICc
## bundlesize.only 31.02      -0.5548      Sss df   logLik  AICc
##              delta weight
## bundle.FNB      0.00 0.474
## FNB.only        0.47 0.374
## bundle.H        3.07 0.102
## bundle.D        5.61 0.029
## H.only          7.27 0.013
## D.only          8.33 0.007
## session.only    19.36 0.000
## ktp.only        20.43 0.000
## null.model      20.78 0.000
## bundle.ktp      21.16 0.000
## bundlesize.only 21.16 0.000
## Models ranked by AICc(x)
## Random terms (all models):
## 1 | ID, 1 | Session
```

- Based on model selection analysis, the bundle.fnb model is the most parsimonious model that explains total consumption in the proboscis monkeys. Here, individual IDs ($n = 7$) are used as error terms, which accounts for individual slopes.
- Based on the summary output, a linear increase FNB results in an increase in total leaves consumed (by 49 counts).
 - In contrast, the negative coefficients for browse size simply suggest that every increase in bundle size will lead to a small decrease in total leaves consumed (by -0.9 counts).
 - Thankfully, we are also seeing low collinearity between bundle size and FNB ($VIF > 3$) as well as low correlation (0.13).
 - Therefore, higher FNB explains higher consumption rate than the other diversity indexes.
 - FNB does a good job in addressing present-but-not-consumed browses unlike the other diversity indexes

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: total ~ browse.size + fnb + (1 | ID) + (1 | Session)
```

```
## Data: rf.data
## REML criterion at convergence: 913.1172
## Random effects:
## Groups Name Std.Dev.
## Session (Intercept) 3.630
## ID (Intercept) 7.158
## Residual 12.004
## Number of obs: 116, groups: Session, 39; ID, 7
## Fixed Effects:
## (Intercept) browse.size fnb
## 26.8364 -0.9048 50.6451

## # Check for Multicollinearity
##
## Low Correlation
##
## Term VIF VIF 95% CI Increased SE Tolerance Tolerance 95% CI
## browse.size 1.01 [1.00, 29136.32] 1.01 0.99 [0.00, 1.00]
## fnb 1.01 [1.00, 29136.32] 1.01 0.99 [0.00, 1.00]
```

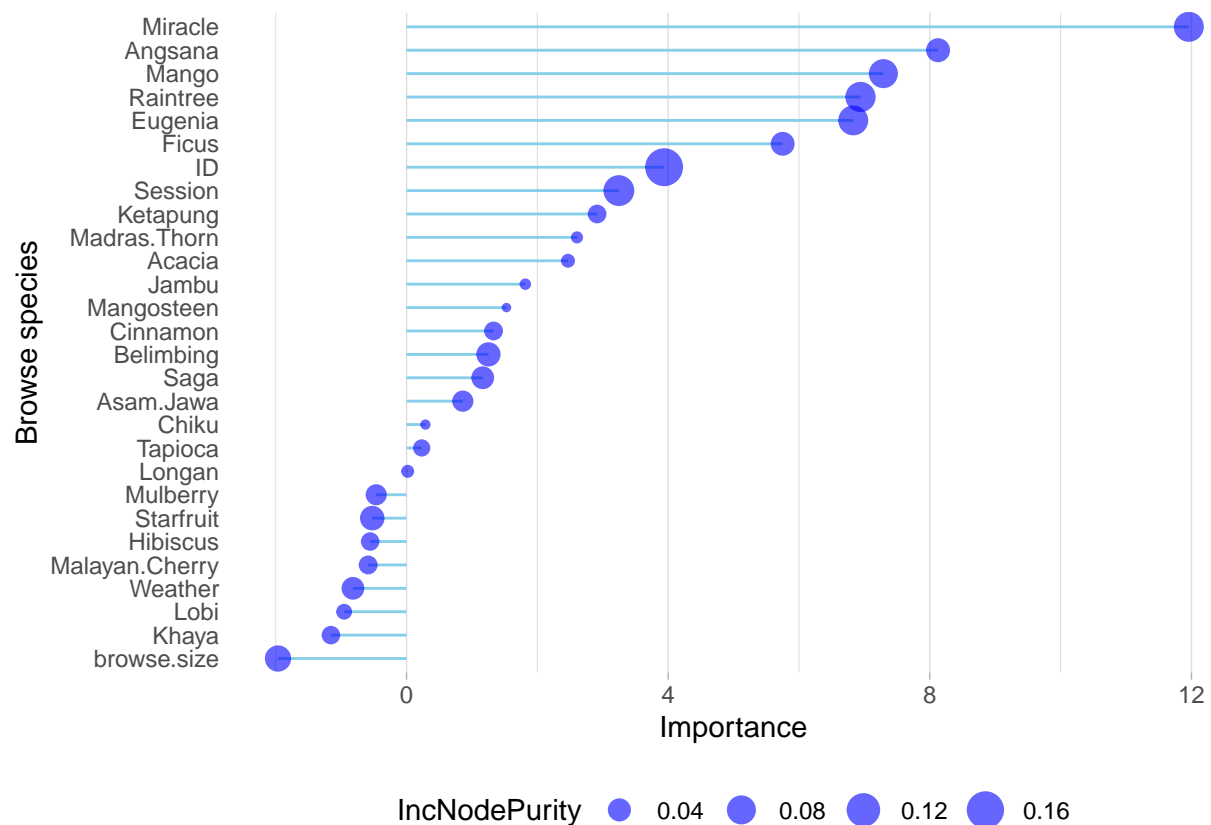


Figure 8: Variable importance plot by FNB. Residuals: 0.001

