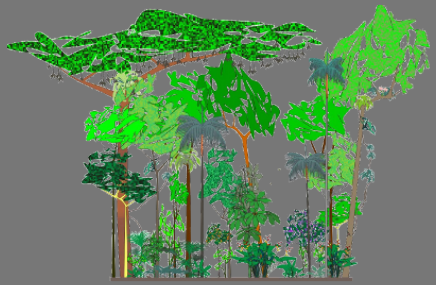
Modelos Lineares Misto

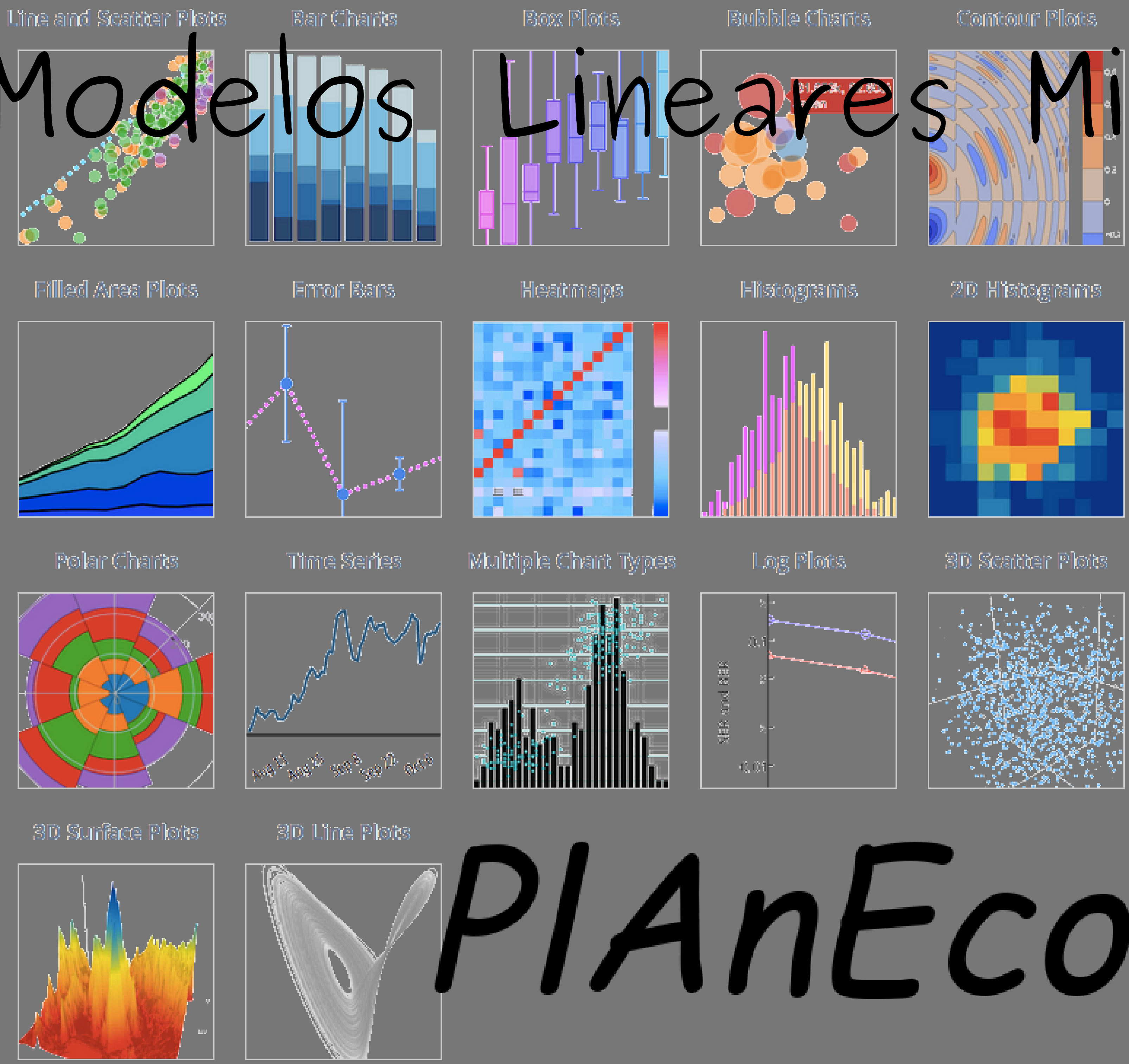
unificação metodológica

Alexandre Adalardo de Oliveira

PlanECO 2019



Modelos Lineares Misto



PlAnEco

Testes Clássicos

Resposta	Preditoras	Teste	Hipótese
Categórica	Categórica	Qui-quadrado	independência
Contínua	Categórica(2)	Teste-t	$\mu_1 = \mu_2$
Contínua	Categórica (>2)	Anova	$\mu_1 = \mu_2 = \mu_3$
Contínua	1 Contínua	Regressão	$\beta_1 = 0$
Contínua	>1 Contínua	Reg. múltipla	$\beta_1 = 0; \beta_n = 0$
Contínua	Cont + Cat	Ancova	$\beta_1 = \beta_2; \alpha_1 = \alpha_2$
Proporção	Contínua	Reg. Logística	$logit(\beta_1) = 1$

- independência entre observações
- distribuição dos resíduos: normal

Modelos Lineares (LM)

Características:

- resposta: variável contínua
- preditora: múltiplas (contínuas, discretas, fator)
  - fator: variável indicadora (nível basal no intercepto)
- simplificação do modelo: F (razão da variância)

Premissas:

- relação linear: $y = \alpha + \beta x$
- estrutura dos resíduos: $N(0, \sigma)$
- independência entre observações

Modelos Lineares Generalizados (GLM)

Características:

- resposta: contagem, proporção ou binária
- função de ligação: preditor linear
- simplificação do modelo: X^2 ou F

Premissas:

- relação linearizável
- estrutura dos resíduos: binomial, poisson, gama…
- independência entre observações

Modelos Lineares Mistos:

Variáveis preditoras: fixas e aleatórias

- incorporam dependência entre observações
  - espaço
  - tempo
- desenho em bloco
- aninhado

Efeito (Fator) Aleatório: agrupamento de observações

- preditora aleatória: $N(\mu, \sigma^2_{entre})$
- não há interesse em interpretá-la (pode ter!)
- o nível se repete para observações (correlação)
- dependência entre observações
- depende da sua pergunta

Efeito Aleatório

Box 13.1 WHEN TO TREAT A PREDICTOR VARIABLE AS A RANDOM EFFECT

You may want to treat a predictor variable as a random effect if you:

- don't want to test hypotheses about differences between responses at particular levels of the grouping variable;
- do want to quantify the variability among levels of the grouping variable;
- do want to make predictions about unobserved levels of the grouping variable;
- do want to combine information across levels of the grouping variable;
- have variation in information per level (number of samples or noisiness);
- have levels that are randomly sampled from/representative of a larger population;
- have a categorical predictor that is a nuisance variable (i.e., it is not of direct interest, but should be controlled for).

Cf. Crawley (2002); Gelman (2005)

If you have sampled fewer than five levels of the grouping variable, you should strongly consider treating it as a fixed effect even if one or more of the criteria above apply.

- Bolker (2015)

Definições: efeito aleatório

Frequentista

Modelos Mistos

- variáveis categóricas em que os níveis são amostras aleatórias da população

Bayesiana

Modelos hierarquicos

- variáveis em que os parâmetros são realizações de uma distribuição

Modelo Misto

Efeito aleatório no intercepto

$$y_{ij} = \bar{\alpha} + \beta x + \epsilon_j + \epsilon_{ij}$$

Efeito aleatório

$$\epsilon_j = N(0, \sigma^2_{entre})$$

Resíduo

$$\epsilon_{ij} = N(0, \sigma^2_{intra})$$

Variância total

$$\sigma^2_{total} = \sigma^2_{entre} + \sigma^2_{intra}$$

Modelo hierárquicos

$$y_{ij} = \alpha_j + \beta_j x + \epsilon ij$$

Efeito Randômico

$$\alpha_j = N(\bar{\alpha}, \sigma^2_{alpha})$$

e/ou

$$\beta_j = N(\bar{\beta}, \sigma^2_{beta})$$

Resíduo

$$\epsilon_{ij} = N(0, \sigma^2_{res})$$

Variância do LMM

Dependência entre observações:

- parte da variação total vai para entre grupos

$$\sigma^2_{total} = \sigma^2_{entre} + \sigma^2_{intra}$$

- correlação entre 2 observação de um mesmo grupo

$$\rho = \sqrt{\frac{\sigma^2_{entre}}{\sigma^2_{total}}}$$

- observações de grupos diferentes não são correlacionadas (independentes!)

Dependência entre observações

Verificar se espécies decíduas x perenes tem diferentes taxas de crescimento

- 5 espécies em cada grupo (decídua, perene)
- 10 indivíduos em cada espécie
- crescimento controlado pelo tamanho

Problema:

- os indivíduos de uma mesma espécie não são independentes
- as espécies tem relações de parentescos (**ISSO FICA PARA DEPOIS!**)

Soluções:

1. Calcular a média para cada espécie e comparar as médias

$$n : 100 \rightarrow 20$$

2. Anova com efeito fixo crescimento e aleatório de espécies

No segundo caso incorporamos:

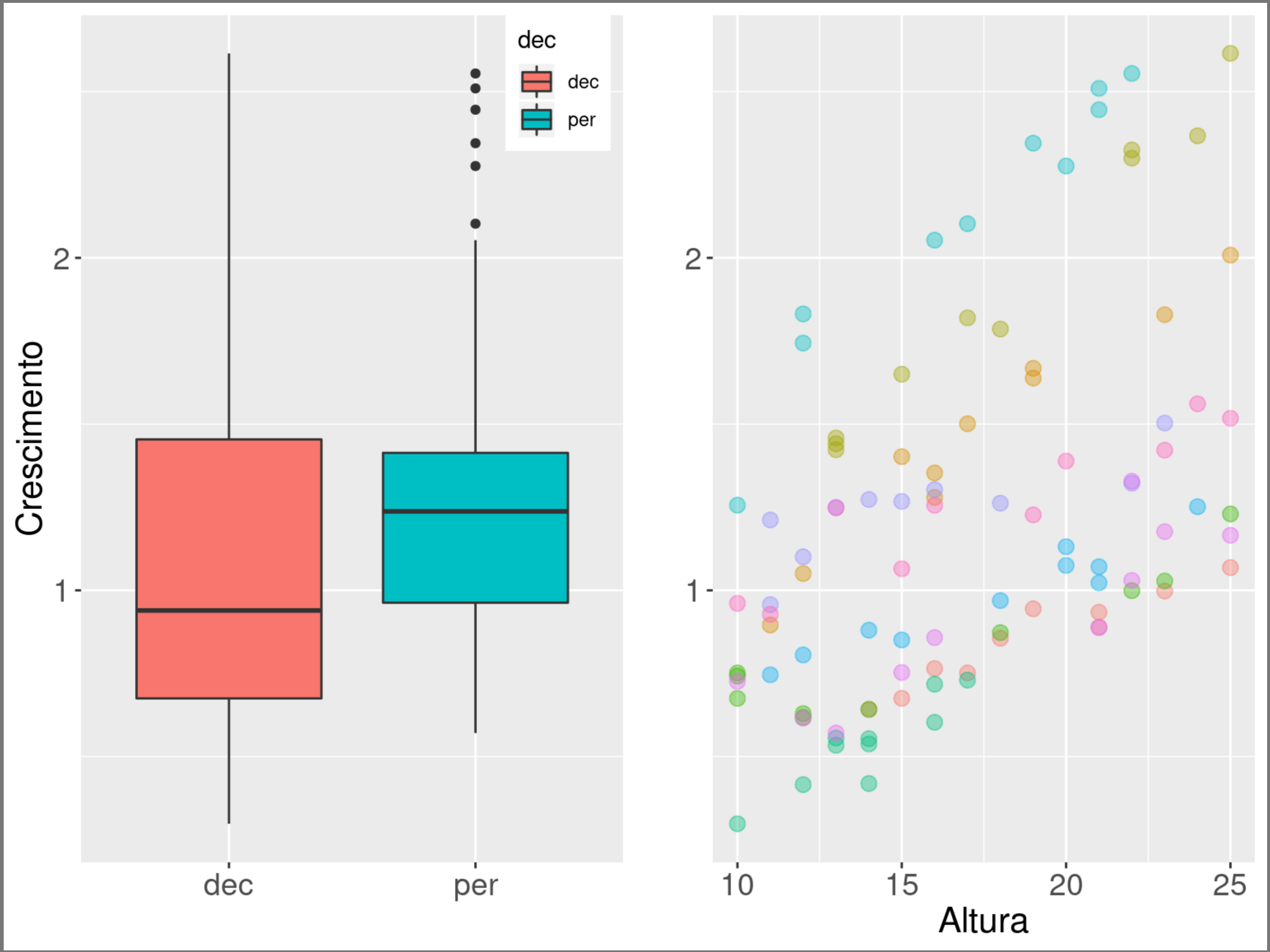
- diminuição da incerteza dentro do grupo (espécies)
- menos dados independentes do que observações

Exemplo: crescimento de árvores

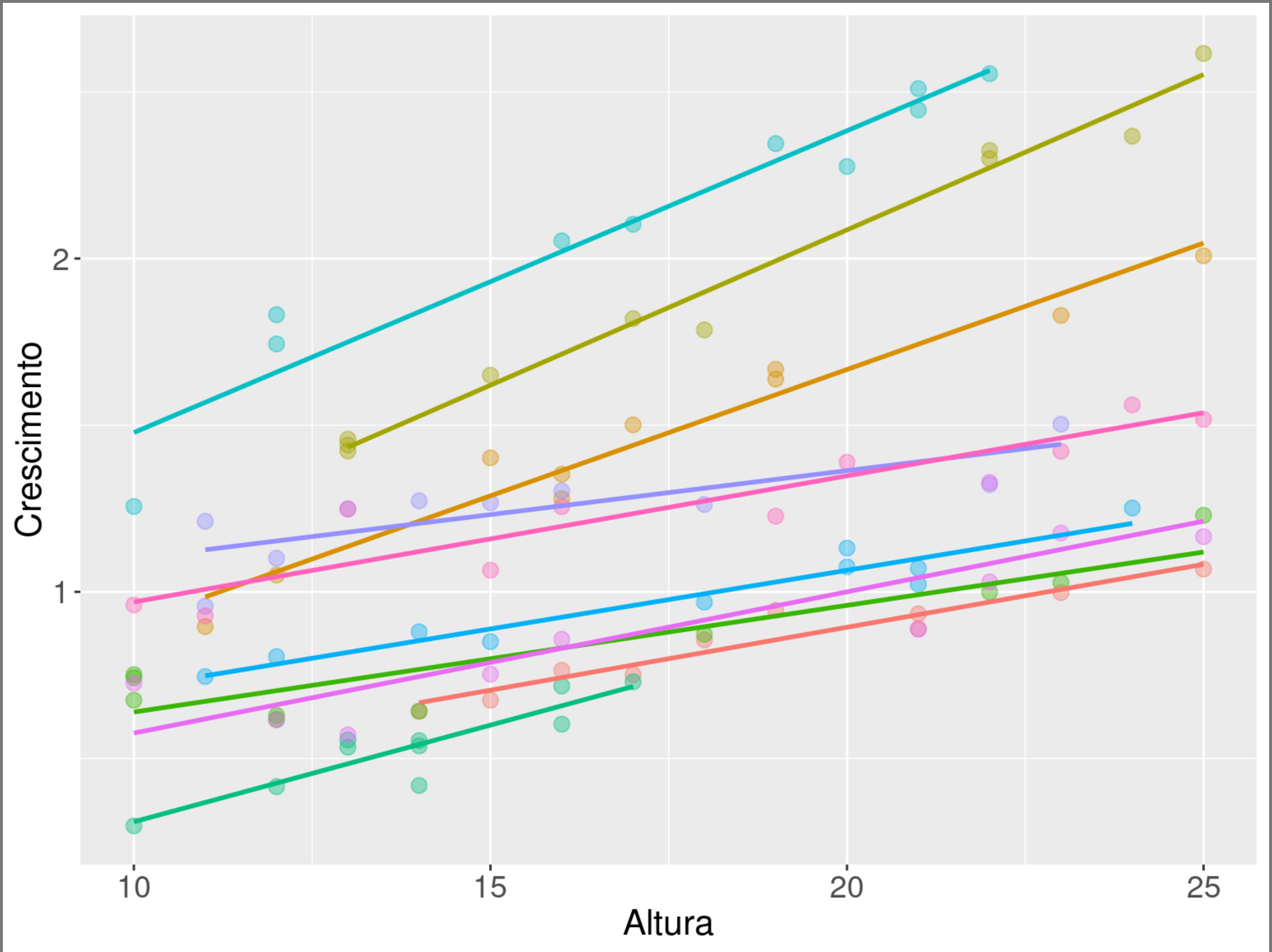
cresc	sp	dec	alt
0.64	sp01	dec	14
0.75	sp01	dec	17
2.01	sp02	dec	25
1.44	sp03	dec	13
0.63	sp04	dec	12
0.55	sp05	dec	14
2.45	sp06	per	21
1.07	sp07	per	21
1.25	sp08	per	13

cresc	sp	dec	alt
0.86	sp09	per	16

Crescimento de árvores



Diferentes Espécies



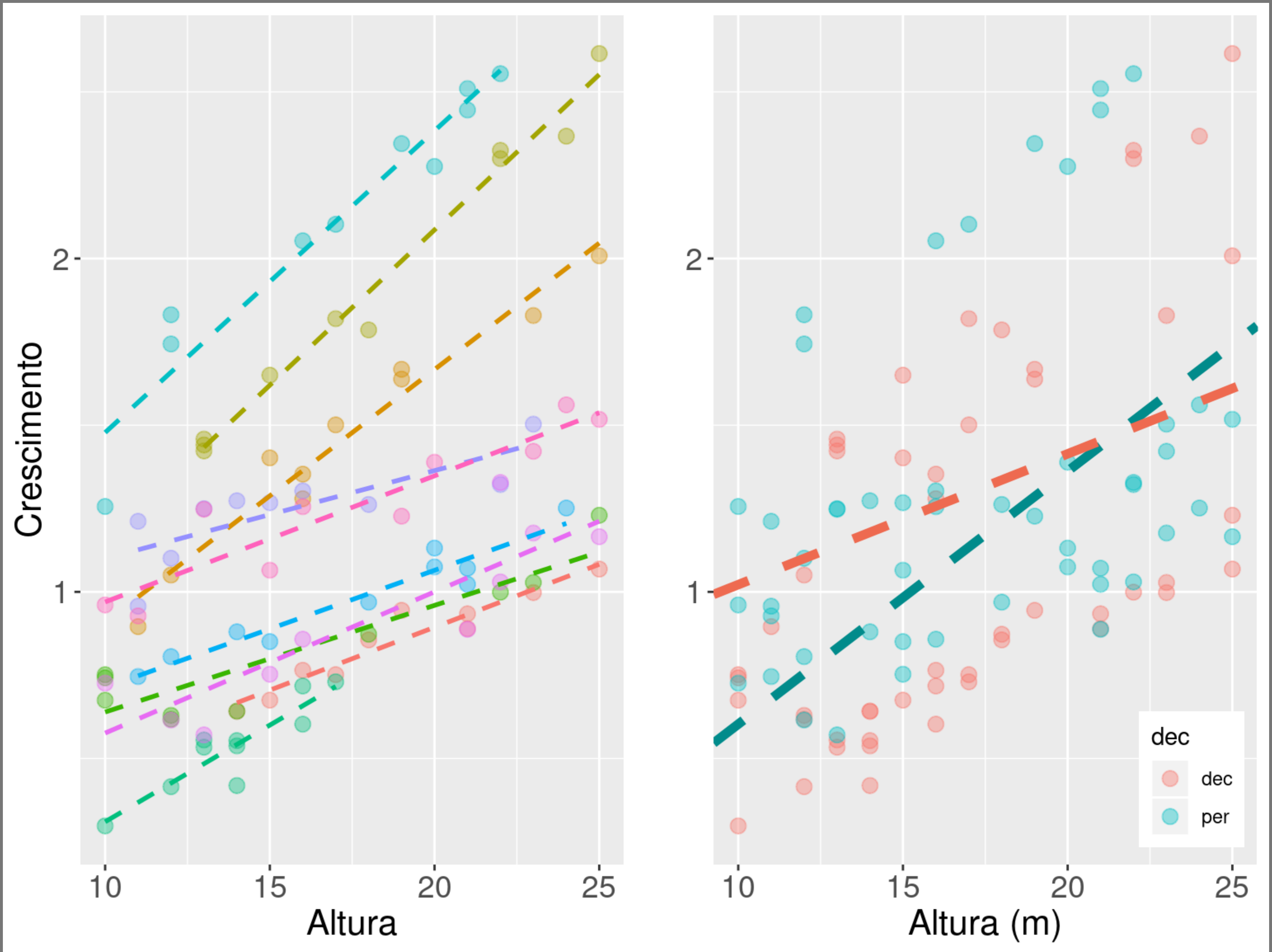
Modelo Linear Simples

```
##
## Call:
## lm(formula = cresc ~ alt * sp * dec, data = dados)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.221921 -0.042993 -0.006632  0.049369  0.243368
##
## Coefficients: (20 not defined because of singularities)
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.387e-01  1.511e-01   0.918 0.361317
## alt                3.776e-02  7.870e-03   4.798 7.32e-06 ***
## spsp02             1.202e-02  1.892e-01   0.064 0.949489
## spsp03             8.339e-02  1.877e-01   0.444 0.658085
## spsp04             1.808e-01  1.706e-01   1.060 0.292331
## spsp05            -4.084e-01  2.427e-01  -1.683 0.096314 .
## spsp06             4.340e-01  1.888e-01   2.299 0.024127 *
## spsp07             2.224e-01  1.910e-01   1.165 0.247645
## spsp08             6.973e-01  1.839e-01   3.792 0.000289 ***
## spsp09             1.398e-02  1.804e-01   0.077 0.938448
## spsp10             4.526e-01  1.786e-01   2.535 0.013211 *
## decper                    NA         NA      NA       NA
## alt:spsp02         3.807e-02  1.014e-02   3.754 0.000328 ***
## alt:spsp03         5.544e-02  9.861e-03   5.623 2.66e-07 ***
## alt:spsp04        -5.758e-03  9.204e-03  -0.626 0.533386
## alt:spsp05         2.023e-02  1.565e-02   1.293 0.199706
## alt:spsp06         5.278e-02  1.019e-02   5.180 1.62e-06 ***
## alt:spsp07        -2.579e-03  1.018e-02  -0.253 0.800683
## alt:spsp08        -1.137e-02  1.023e-02  -1.111 0.269797
## alt:spsp09         4.612e-03  9.491e-03   0.486 0.628301
## alt:spsp10         8.565e-05  9.426e-03   0.009 0.992773
## alt:decper                NA         NA      NA       NA
## spsp02:decper             NA         NA      NA       NA
## spsp03:decper             NA         NA      NA       NA
## spsp04:decper             NA         NA      NA       NA
## spsp05:decper             NA         NA      NA       NA
## spsp06:decper             NA         NA      NA       NA
## spsp07:decper             NA         NA      NA       NA
## spsp08:decper             NA         NA      NA       NA
## spsp09:decper             NA         NA      NA       NA
## spsp10:decper             NA         NA      NA       NA
```

Modelo Linear Simples

```
##
## Call:
## lm(formula = cresc ~ alt + sp + dec, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.49901 -0.09029 -0.00374  0.08075  0.35015
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.110632   0.073375  -1.508  0.13516
## alt          0.050953   0.003146  16.194  < 2e-16 ***
## spsp02       0.691668   0.060994  11.340  < 2e-16 ***
## spsp03       1.101669   0.060826  18.112  < 2e-16 ***
## spsp04       0.134515   0.061666   2.181  0.03179 *
## spsp05      -0.061143   0.062789  -0.974  0.33280
## spsp06       1.356294   0.061079  22.206  < 2e-16 ***
## spsp07       0.194196   0.060923   3.188  0.00198 **
## spsp08       0.565906   0.061720   9.169 1.68e-14 ***
## spsp09       0.109734   0.060867   1.803  0.07480 .
## spsp10       0.471219   0.060923   7.735 1.52e-11 ***
## decper             NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1359 on 89 degrees of freedom
## Multiple R-squared:  0.9435, Adjusted R-squared:  0.9371
## F-statistic: 148.5 on 10 and 89 DF,  p-value: < 2.2e-16
```

LM: desconsidera espécie



LM: desconsidera espécie

```
##
## Call:
## lm(formula = cresc ~ alt + dec + alt:dec, data = dados)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6731 -0.3533 -0.1294  0.2737  1.0620
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.15569    0.25488  -0.611   0.5427
## alt          0.07589    0.01467   5.172 1.27e-06 ***
## decper       0.78504    0.35938   2.184   0.0314 *
## alt:decper  -0.03665    0.02048  -1.789   0.0767 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4656 on 96 degrees of freedom
## Multiple R-squared:  0.2845, Adjusted R-squared:  0.2621
## F-statistic: 12.72 on 3 and 96 DF,  p-value: 4.532e-07
```

LM: desconsidera espécie

Simplificando o modelo

```
## Analysis of Variance Table
##
## Model 1: cresc ~ alt + dec + alt:dec
## Model 2: cresc ~ alt + dec
##   Res.Df    RSS Df Sum of Sq     F  Pr(>F)
## 1     96 20.807
## 2     97 21.501 -1  -0.69379 3.201 0.07675 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

LM: desconsidera espécie

```
##
## Call:
## lm(formula = cresc ~ alt + dec, data = dados)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6349 -0.3800 -0.1749  0.3577  1.0281
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.15991    0.18605   0.859   0.3922
## alt          0.05708    0.01035   5.513 2.91e-07 ***
## decper       0.16404    0.09423   1.741   0.0849 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4708 on 97 degrees of freedom
## Multiple R-squared:  0.2606, Adjusted R-squared:  0.2454
## F-statistic: 17.09 on 2 and 97 DF,  p-value: 4.368e-07
```

LM: desconsidera espécie

Simplifica o modelo

```
## Analysis of Variance Table
##
## Model 1: cresc ~ alt
## Model 2: cresc ~ alt + dec
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1     98 22.173
## 2     97 21.501  1   0.67183 3.0309 0.08486 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Modelo Linear Simples

Simplifica o modelo

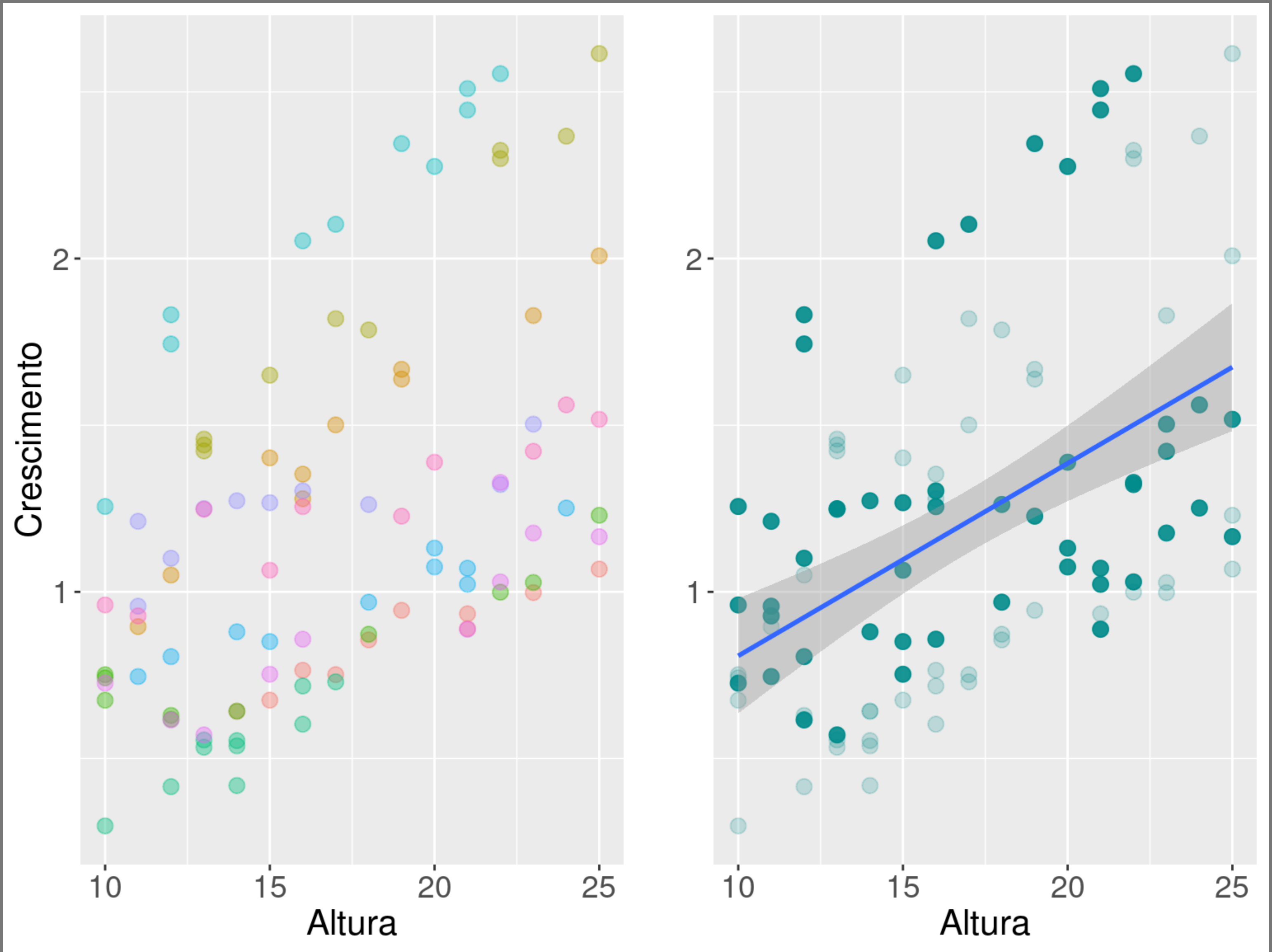
```
## Analysis of Variance Table
##
## Model 1: cresc ~ 1
## Model 2: cresc ~ alt
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     99 29.079
## 2     98 22.173  1    6.9067 30.526 2.716e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Modelo Linear Simples

Modelo Mínimo Adequado

```
##
## Call:
## lm(formula = cresc ~ alt, data = dados)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6201 -0.4005 -0.1092  0.2928  1.0665
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.23052    0.18345   1.257    0.212
## alt          0.05775    0.01045   5.525 2.72e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4757 on 98 degrees of freedom
## Multiple R-squared:  0.2375, Adjusted R-squared:  0.2297
## F-statistic: 30.53 on 1 and 98 DF,  p-value: 2.716e-07
```

Modelo Linear Simples



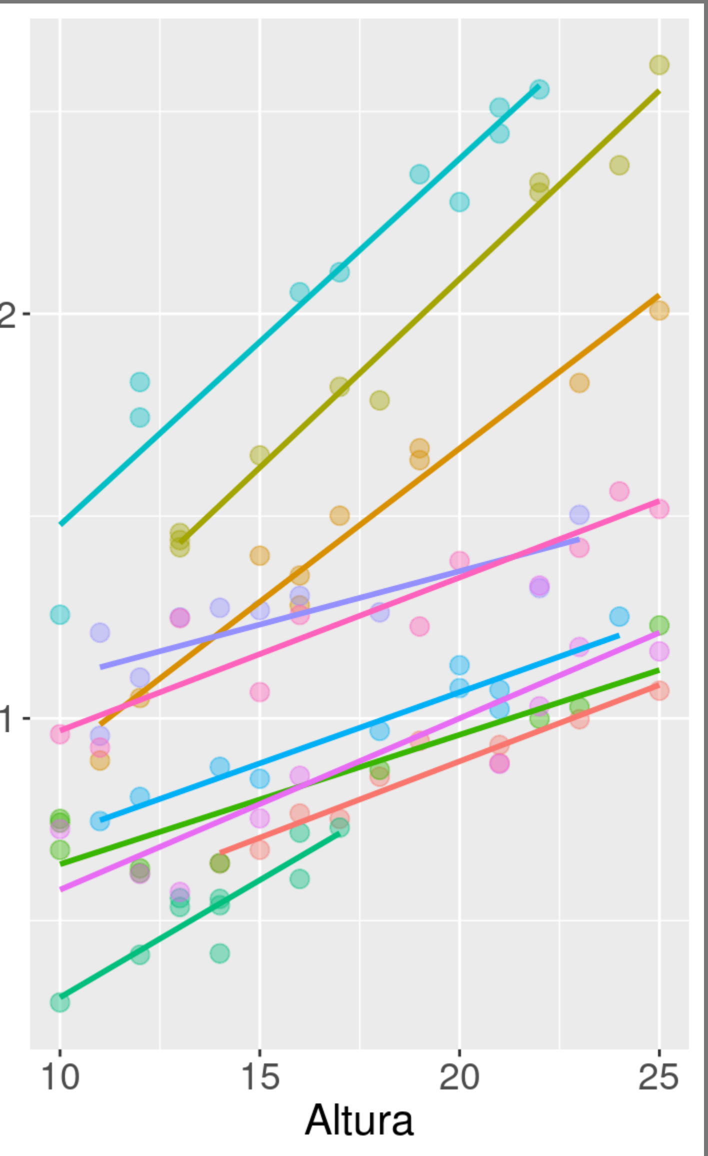
Modelo Linear Simples

```
##
## Call:
## lm(formula = cresc ~ alt + dec, data = dados)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6349 -0.3800 -0.1749  0.3577  1.0281
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.15991    0.18605   0.859   0.3922
## alt          0.05708    0.01035   5.513 2.91e-07 ***
## decper       0.16404    0.09423   1.741   0.0849 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4708 on 97 degrees of freedom
## Multiple R-squared:  0.2606, Adjusted R-squared:  0.2454
## F-statistic: 17.09 on 2 and 97 DF,  p-value: 4.368e-07
```

Não há diferença entre árvores decíduas e perenes

Procedimento em duas estapas

- Modelo Linear para cada espécie




```
lmsp1 <- lm(cres~alt,
            data= arv,
            subset =
  arv$sp==sp1)
lmsp2 <- lm(cres~alt,
            data= arv,
            subset =
  arv$sp==sp2)
...
```

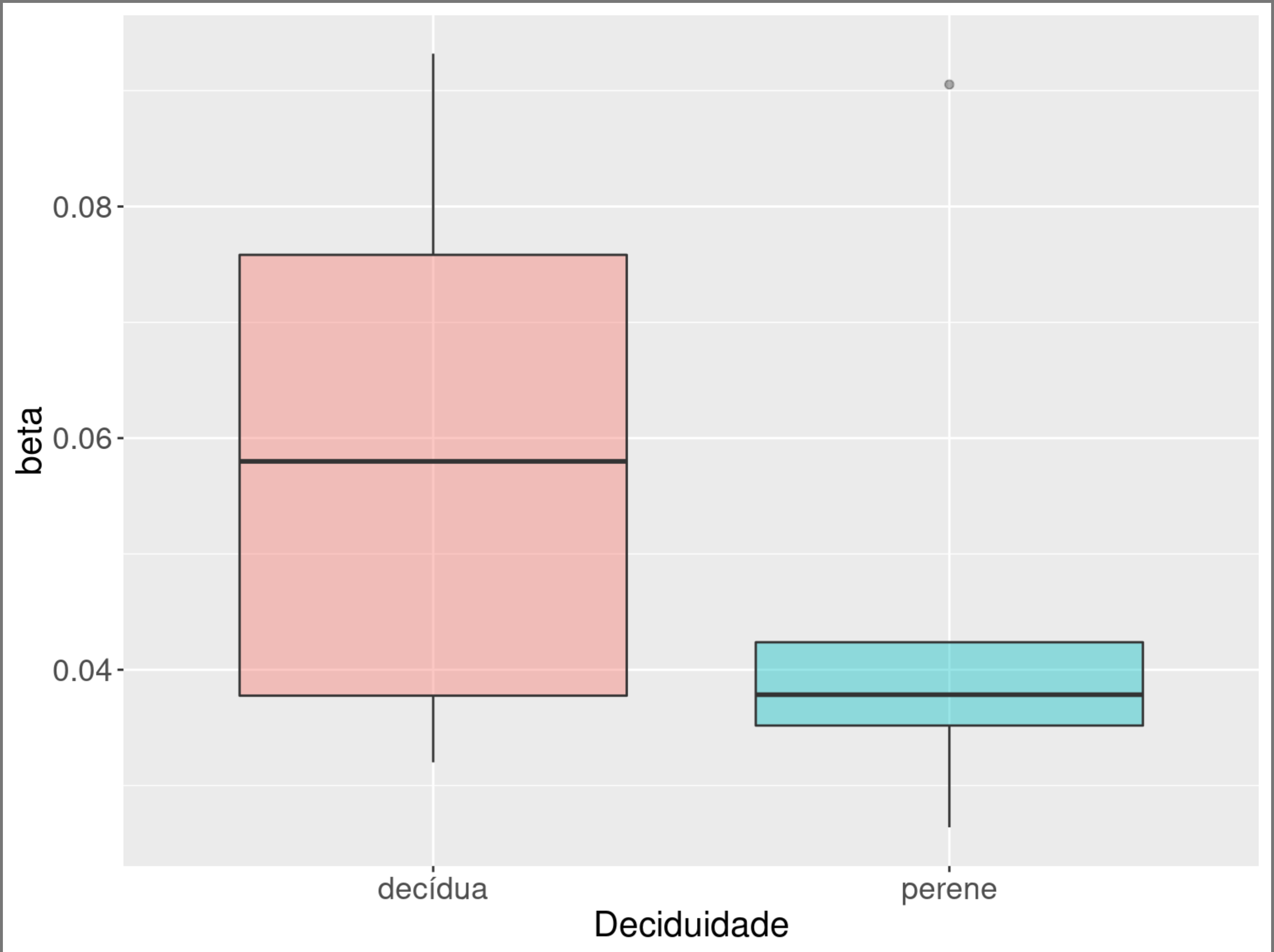
Coeficientes do LM

sp	decid	alpha	beta
sp01	decidua	0.14	0.04

sp	decid	alpha	beta
sp02	decidua	0.15	0.08
sp03	decidua	0.22	0.09
sp04	decidua	0.32	0.03
sp05	decidua	-0.27	0.06
sp06	perene	0.57	0.09
sp07	perene	0.36	0.04
sp08	perene	0.84	0.03
sp09	perene	0.15	0.04
sp10	perene	0.59	0.04

LM dos coeficientes

lm(beta ~ decid)



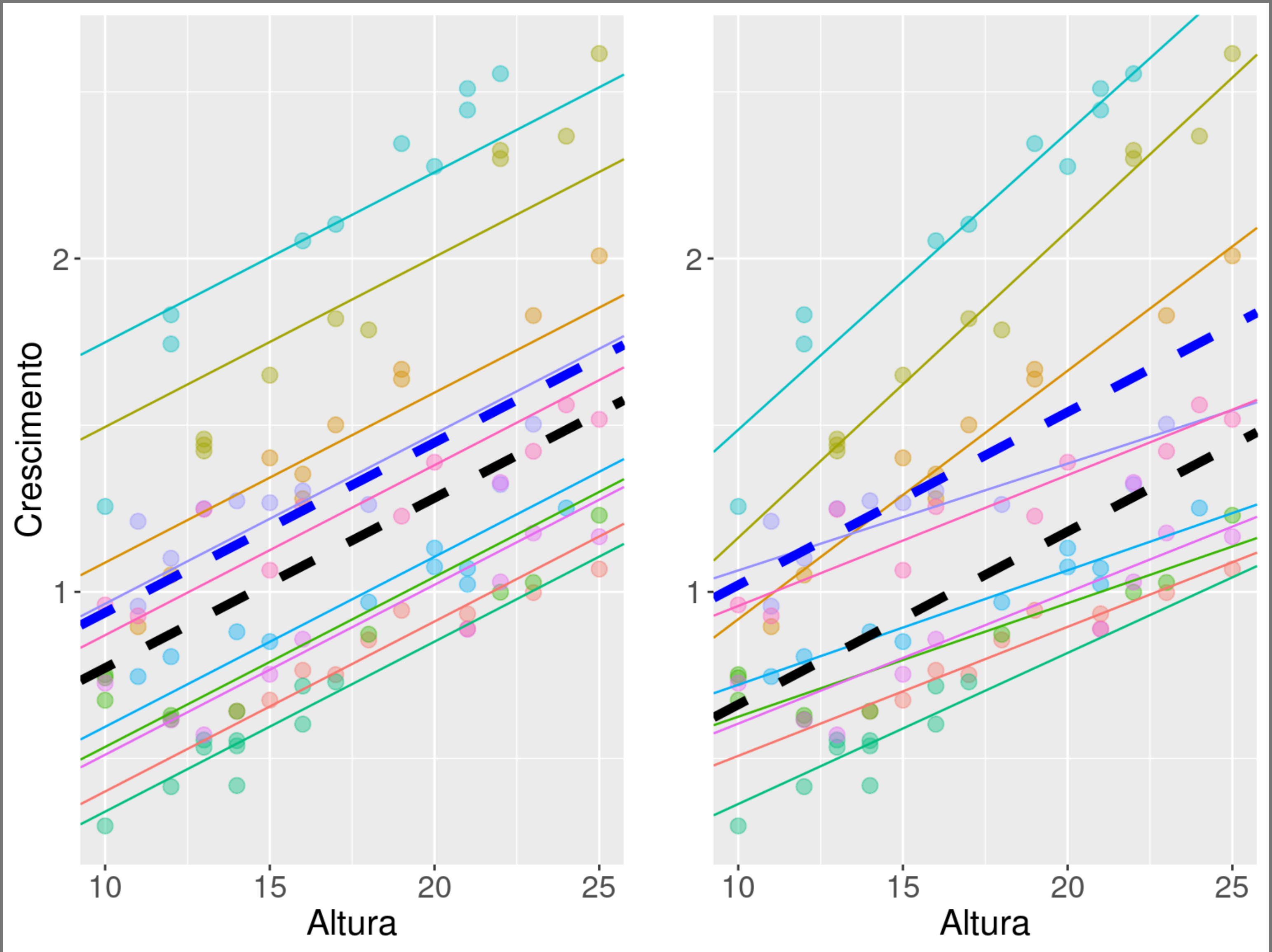
Comparando coeficientes

```
##
## Call:
## lm(formula = beta ~ decid, data = coefdados)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.027354 -0.017877 -0.006357  0.012011  0.044072
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.05936    0.01140   5.207 0.000815 ***
## decidperene -0.01289    0.01612  -0.800 0.446968
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02549 on 8 degrees of freedom
## Multiple R-squared:  0.07403,    Adjusted R-squared:  -0.04172
## F-statistic: 0.6395 on 1 and 8 DF,  p-value: 0.447
```

Não há diferença entre decíduas e perenes

Apenas tamanho afeta o crescimento

Modelo Misto: opções



Modelos Mistos: dados

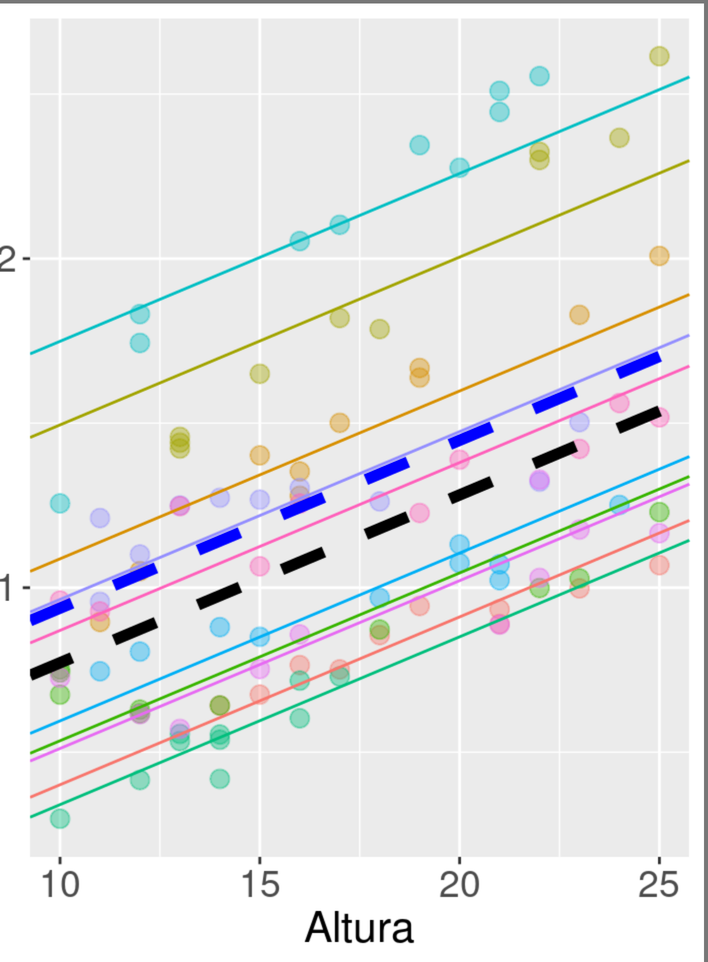
cresc	sp	dec	alt
0.63	sp04	dec	12
0.42	sp05	dec	12
1.06	sp10	per	15
2.45	sp06	per	21
0.86	sp09	per	16
0.81	sp07	per	12
2.10	sp06	per	17
1.39	sp10	per	20
0.74	sp04	dec	10
0.62	sp04	dec	12
0.75	sp09	per	15
1.74	sp06	per	12
0.72	sp05	dec	16
1.50	sp08	per	23
0.67	sp04	dec	10

- 10 espécies
  - 10 indivíduos por sp
    - crescimento
    - altura

100 observações

Modelo Misto: intercepto

```
lmer(cresc ~ alt + dec + (1|sp), data = arv)
```



Modelo Méd

$y = \bar{\alpha} + \beta_1 x_1$

$\epsilon_{sp} = \mathcal{I}$

$\alpha_{sp} = \mathcal{I}$

Modelo Co

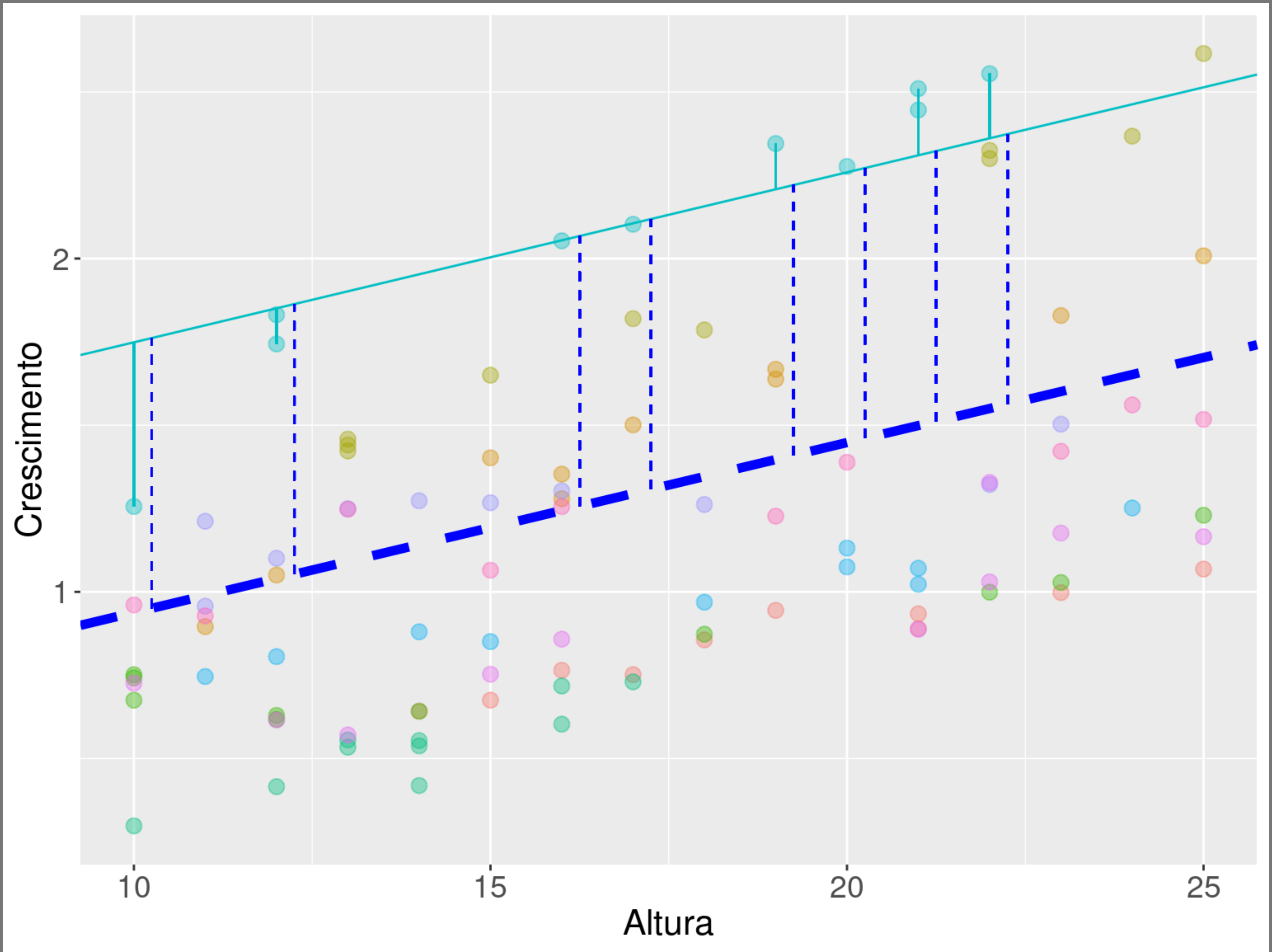
$cr_{ij} = \alpha_{sp_j} + \beta_1 a$

Modelo Misto: intercepto

$$y = \bar{\alpha} + \beta_1 x_1 + \beta_2 x_2 + \epsilon_{sp} + \epsilon_{res}$$

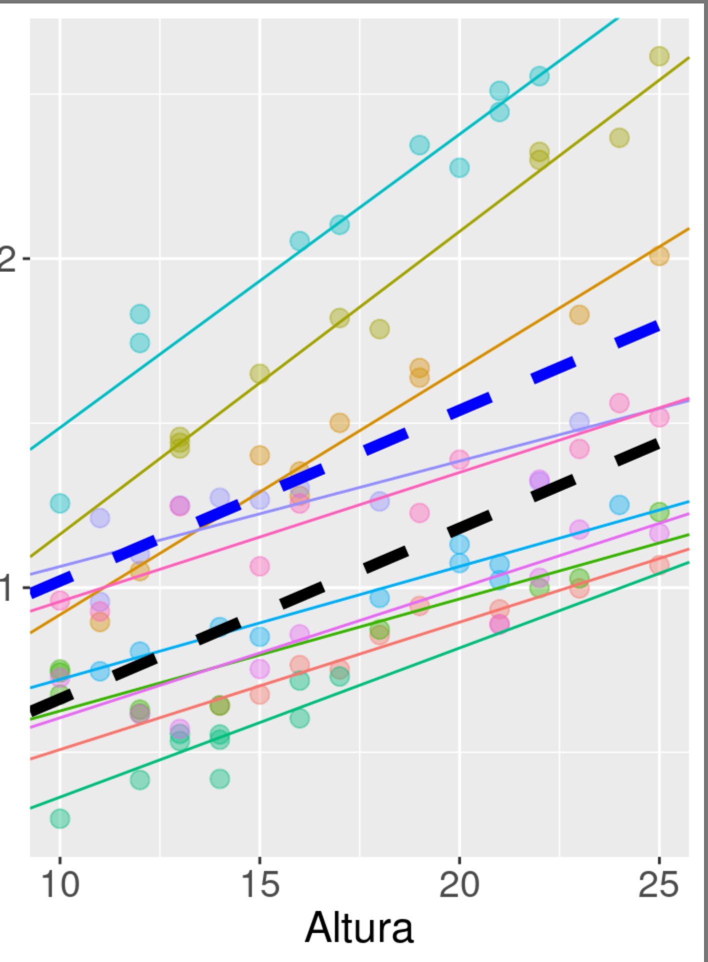
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: cresc ~ alt + dec + (1 | sp)
##    Data: dados
##
## REML criterion at convergence: -57.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6242 -0.6882 -0.0167  0.5874  2.6134
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  sp       (Intercept) 0.24734  0.4973
##  Residual             0.01847  0.1359
## Number of obs: 100, groups:  sp, 10
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 0.261866   0.229397   1.142
## alt         0.051003   0.003145  16.216
## decper      0.166111   0.315716   0.526
##
## Correlation of Fixed Effects:
##        (Intr) alt
## alt    -0.230
## decper -0.687 -0.003
```

Resíduos do modelo: sp06



Modelo Misto: inclinação

```r
lmer(cresc ~ alt + dec + (alt|sp), data = arv)
```



Modelo Mé

$$y = \bar{\alpha} +$$

$$\beta_{sp} =$$

$$\alpha_{sp} =$$

$$\epsilon_{mm}$$

Modelo Co

$$cr_{ij} = \alpha_{sp_j} + \beta_{sp_j}$$

Modelo Misto: inclinação

$$y = (\bar{\alpha} + \epsilon_{sp}) + (\bar{\beta_1} + \epsilon_{sp:alt})x_1 + \beta_2 x_2 + \epsilon_{res_{ij}}$$

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: cresc ~ alt + dec + (alt | sp)
##    Data: dados
##
## REML criterion at convergence: -130
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.73002 -0.53247 -0.05236  0.58671  2.97258
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  sp       (Intercept) 0.0416332 0.20404
##           alt         0.0005977 0.02445  0.07
##  Residual             0.0071124 0.08433
## Number of obs: 100, groups:  sp, 10
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 0.143124   0.107106   1.336
## alt         0.051887   0.008027   6.464
## decper      0.358508   0.148248   2.418
##
## Correlation of Fixed Effects:
##        (Intr) alt
## alt    -0.065
## decper -0.721  0.026
```

Modelo Misto: comparando efeitos aleatórios

efeito aleatório: (alt|sp) ou (1|sp)

```r
anova(lmincl, lmint, refit=FALSE)
```

```
## Data: dados
## Models:
## lmint: cresc ~ alt + dec + alt:dec + (1 | sp)
## lmincl: cresc ~ alt + dec + alt:dec + (alt | sp)
##        Df      AIC     BIC logLik deviance  Chisq Chi Df Pr
## lmint   6  -40.971 -25.340 26.486  -52.971
## lmincl  8 -108.018 -87.176 62.009 -124.018 71.046      2  3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

retemos o modelo mais complexo (alt|sp)

Modelo Misto: simplificando o efeito fixo

Retirando a interação alt:dec

```r
lmmsp02 <- lmer(cresc ~ alt + dec + (alt|sp), data=dados)
anova(lmmsp02, lmincl)
```

```
## Data: dados
## Models:
## lmmsp02: cresc ~ alt + dec + (alt | sp)
## lmincl: cresc ~ alt + dec + alt:dec + (alt | sp)
##         Df     AIC     BIC logLik deviance  Chisq Chi Df Pr
## lmmsp02  7 -129.42 -111.18 71.710  -143.42
## lmincl   8 -127.90 -107.06 71.952  -143.90 0.4841      1
```

O modelo com interação não fornece informação adicional relevante

LMM: simplificando o efeito fixo

Retirando deciduidade

```r
lmmsp01 <- lmer(cresc ~ alt  + (alt|sp), data=dados)

anova(lmmsp01, lmmsp02)
```

```
## Data: dados
## Models:
## lmmsp01: cresc ~ alt + (alt | sp)
## lmmsp02: cresc ~ alt + dec + (alt | sp)
##         Df     AIC     BIC logLik deviance  Chisq Chi Df Pr
## lmmsp01  6 -125.98 -110.35 68.993  -137.99
## lmmsp02  7 -129.42 -111.18 71.710  -143.42 5.4349      1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Deciduidade é informativa e deve ser retida no modelo

Modelo Misto: mínimo adequado

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: cresc ~ alt + dec + (alt | sp)
##    Data: dados
## 
## REML criterion at convergence: -130
## 
## Scaled residuals: 
##      Min       1Q   Median       3Q      Max 
## -2.73002 -0.53247 -0.05236  0.58671  2.97258 
## 
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  sp       (Intercept) 0.0416332 0.20404      
##           alt         0.0005977 0.02445  0.07
##  Residual             0.0071124 0.08433      
## Number of obs: 100, groups:  sp, 10
## 
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 0.143124   0.107106   1.336
## alt         0.051887   0.008027   6.464
## decper      0.358508   0.148248   2.418
## 
## Correlation of Fixed Effects:
##        (Intr) alt   
## alt    -0.065       
## decper -0.721  0.026
```

Interpretação Biológica

```r
fixef(lmmsp02)
```

```
## (Intercept)         alt      decper
##  0.14312365  0.05188655  0.35850769
```

- crescimento de espécies perenes é ~ 3.5x maior

$$sd_{sp} = 0.20$$
$$sd_{res} = 0.08$$

- variação interespecífica é muito maior que a intra

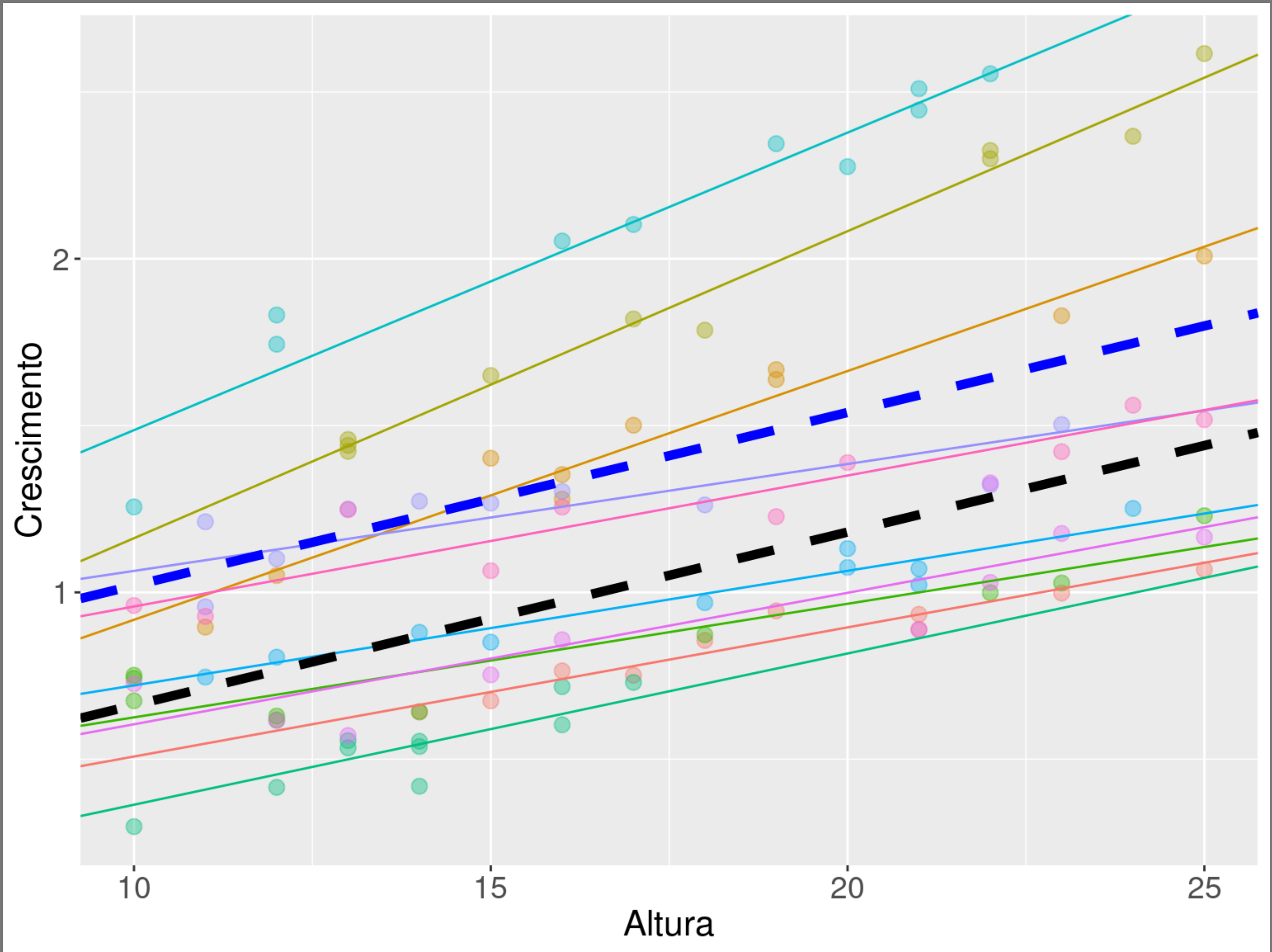
$$y = \bar{\alpha} + \bar{\beta}x + \epsilon_{res}$$

$$\alpha_{sp} = N(\bar{\alpha}, \sigma_{\alpha})$$

$$\beta_{sp} = N(\bar{\beta}, \sigma_{\beta})$$

$$\epsilon_{res} = N(0, \sigma_{res})$$

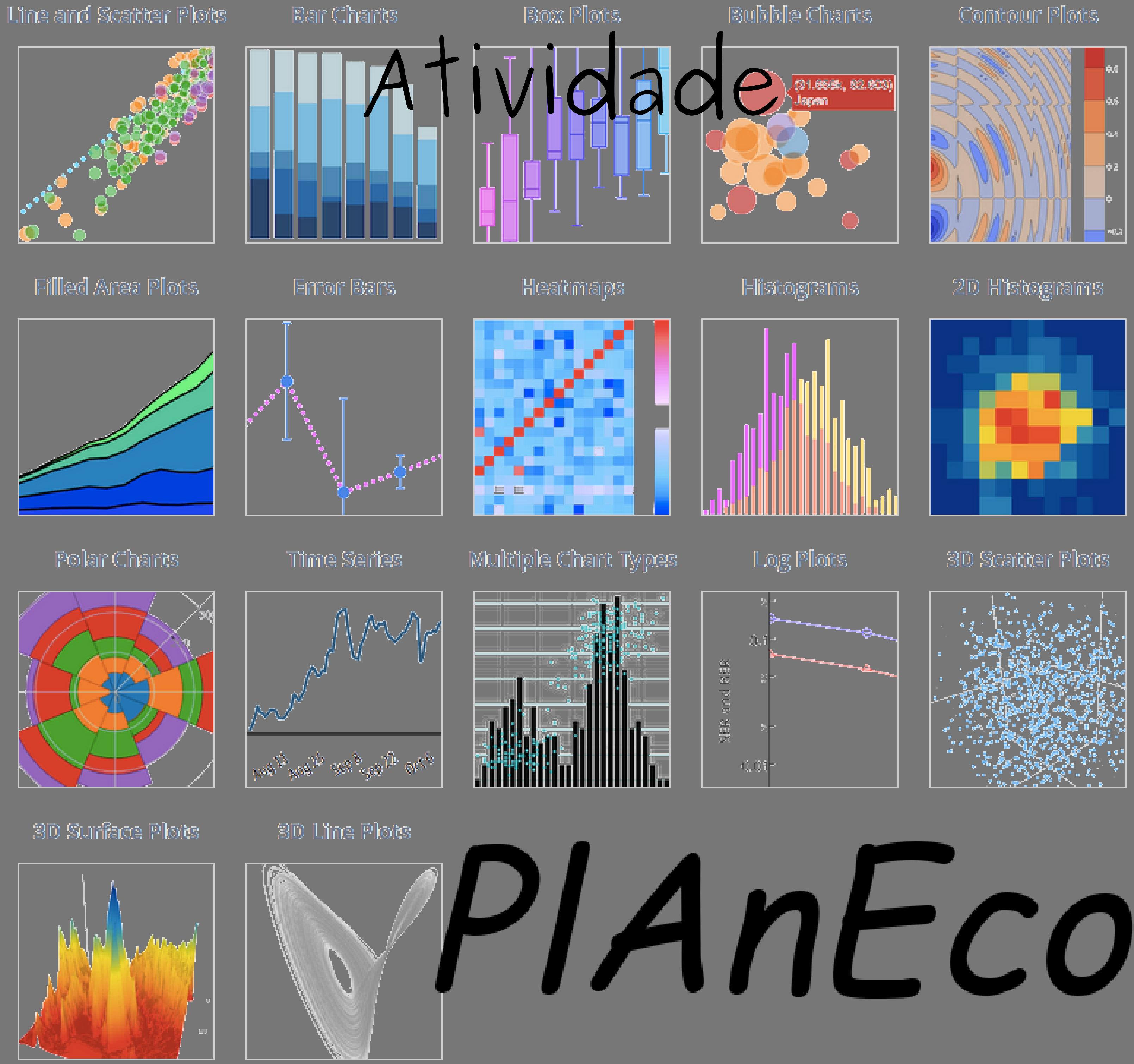
LMM: árvores



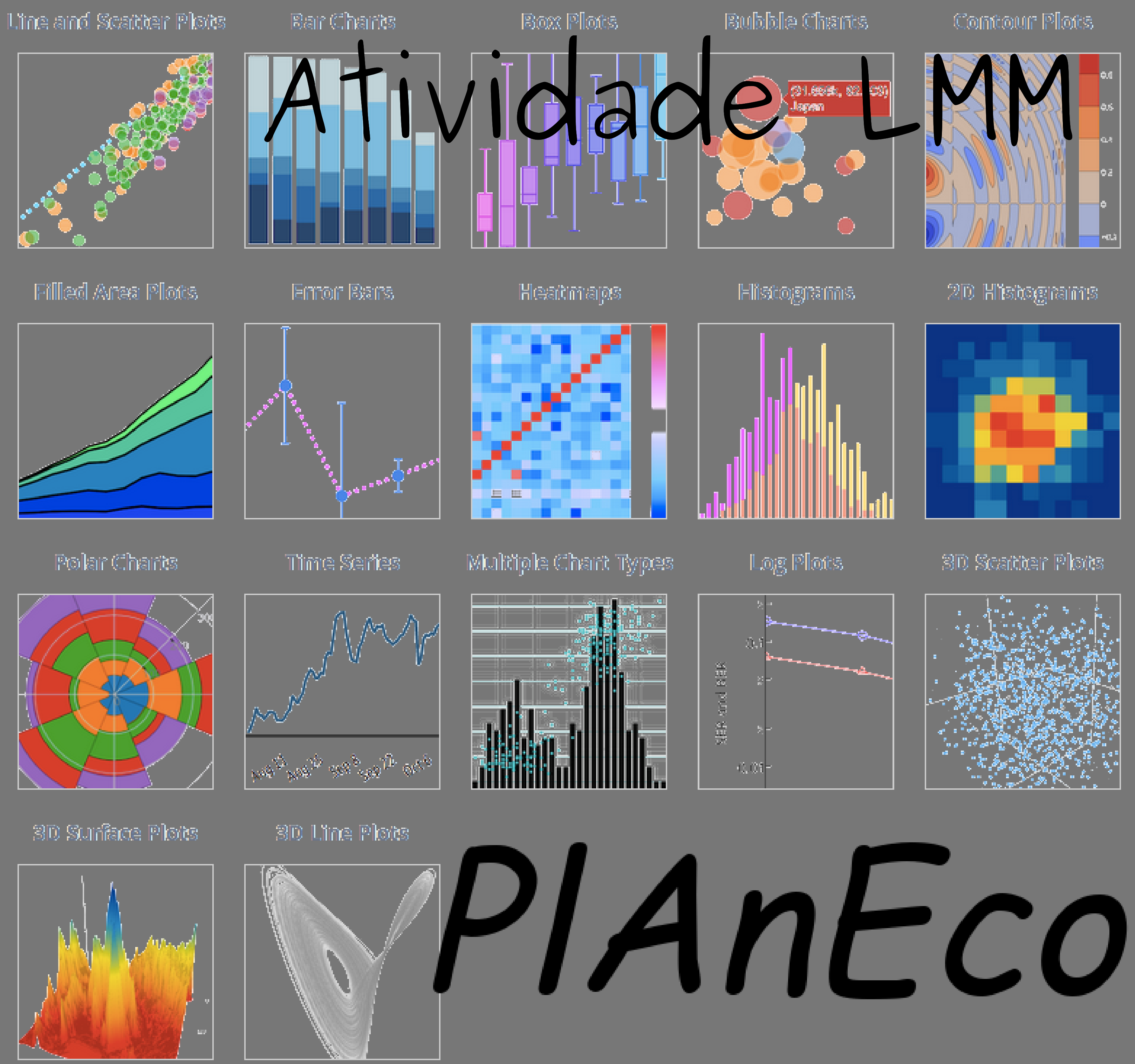
Modelo Mistos: uma abordagem

1. Parta do modelo com efeitos fixos e aleatórios cheio
2. Simplifique a estrutura aleatória:
   - retenha a estrutura aleatória minima adequada
   - mantenha termos associada ao desenho experimental (correlação entre amostras)
   - comparação com anova usando REML: refit= FALSE
3. Definido a estrutura do efeito aleatório:
   - simplifique a estrutra fixa do modelo
   - retenha as variáveis e interações minimas adequadas
   - use comparação por anova com ML (padrão)
4. Diagnóstico do Modelo
   - resíduos: homogeneidade, normalidade

Atividade



PlAnEco


Atividade LMM
Line and Scatter Plots
Bar Charts
Box Plots
Bubble Charts
Contour Plots
Filled Area Plots
Error Bars
Heatmaps
Histograms
2D Histograms
Polar Charts
Time Series
Multiple Chart Types
Log Plots
3D Scatter Plots
3D Surface Plots
3D Line Plots
PlAnEco

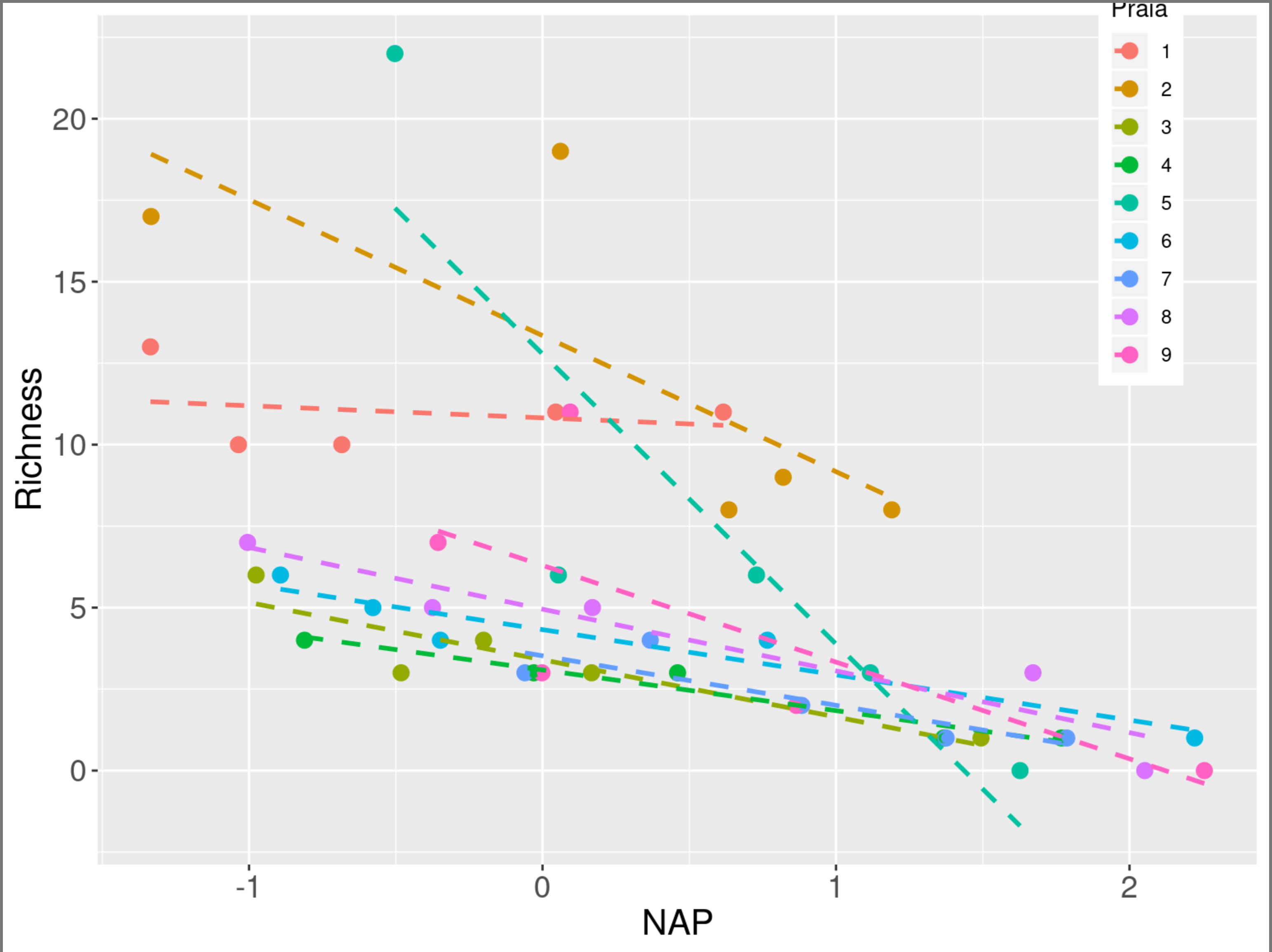
Atividade: riqueza na praia

A riqueza da macrofauna varia em função da altura e exposição da praia

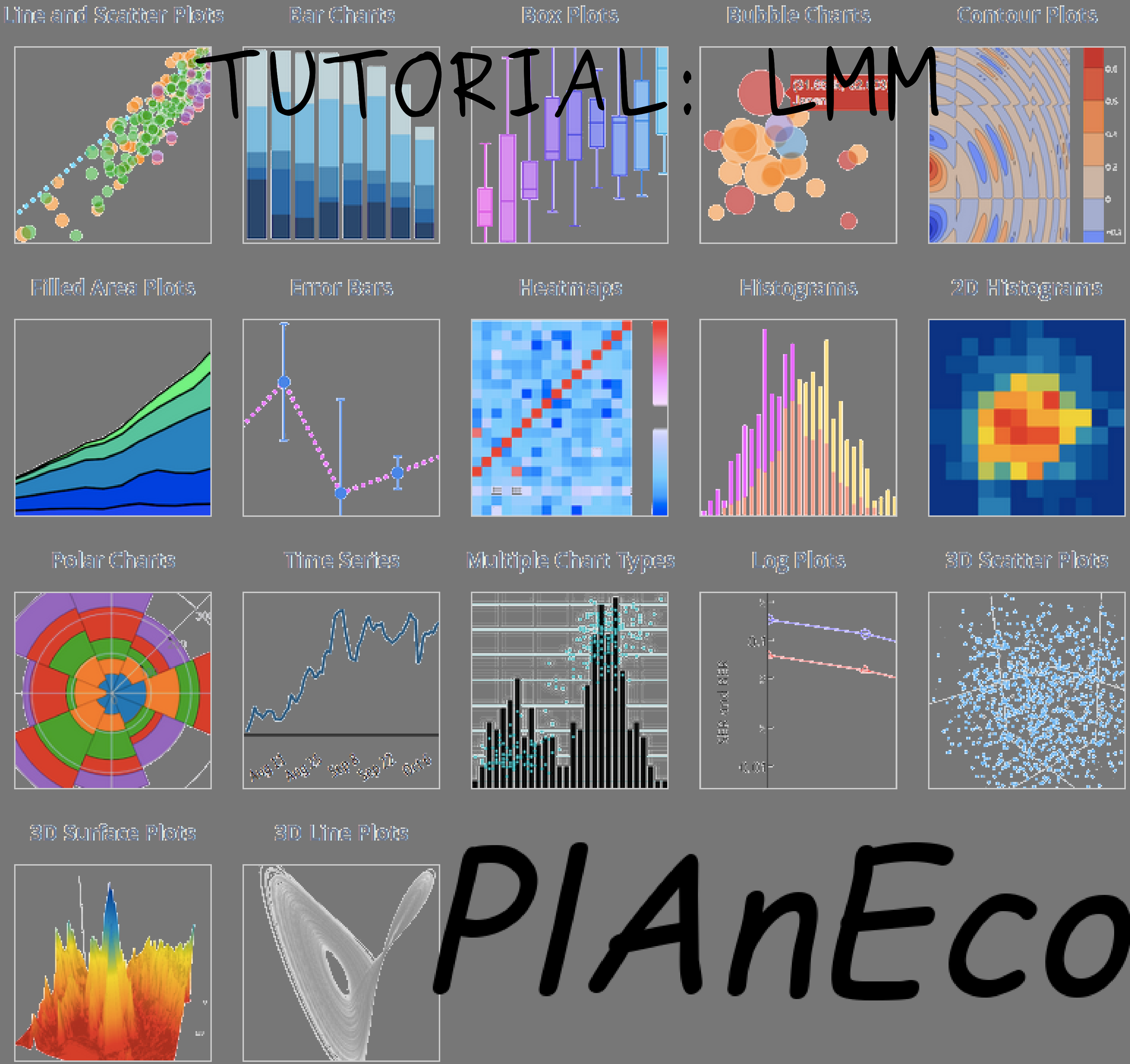
- 45 observações
- 9 praias
  - NAP: contínuo, altura em relação ao NMM
- fExp: 2 níveis de exposição da praia (10 , 11)

Sample	Beach	Richness	NAP	Praia	fExp
1	1	11	0.045	1	10
6	2	8	1.190	2	10
11	3	6	−0.976	3	11
16	4	1	1.768	4	11
21	5	3	1.117	5	10
26	6	5	−0.578	6	11
31	7	2	0.883	7	11
36	8	3	1.671	8	10
41	9	7	−0.356	9	10

LMM: riqueza na praia



TUTORIAL: LMM



PlAnEco

LMM: riqueza da praia

Comparando efeitos aleatórios

```r
lmm00s <- lmer(Richness ~ NAP*fExp + (NAP|Beach),
               data=praia, REML =  FALSE)
lmm00i <- lmer(Richness ~ NAP*fExp + (1|Beach),
               data=praia, REML =  FALSE)
anova(lmm00i, lmm00s, refit= FALSE)
```

```
## Data: praia
## Models:
## lmm00i: Richness ~ NAP * fExp + (1 | Beach)
## lmm00s: Richness ~ NAP * fExp + (NAP | Beach)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chis
## lmm00i  6 242.11 252.95 -115.06   230.11
## lmm00s  8 243.22 257.67 -113.61   227.22 2.8925      2     0.23
```

LMM: modelo mínimo adequado

Comparando LMM: estrutura fixa

- os modelos devem ser comparados por ML
- devem ser apresentados com REML

```
lmm01p <- lmer(Richness ~ NAP + fExp + NAP:fExp + (1|Beach), data=praia)
lmm02p <- lmer(Richness ~ NAP + fExp +  (1|Beach), data=praia)
anova(lmm01p, lmm02p)
```

```
## Data: praia
## Models:
## lmm02p: Richness ~ NAP + fExp + (1 | Beach)
## lmm01p: Richness ~ NAP + fExp + NAP:fExp + (1 | Beach)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmm02p  5 244.76 253.79 -117.38   234.76
## lmm01p  6 242.11 252.95 -115.06   230.11 4.6454      1    0.03114 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

A interação NAP:fExp fornece informação relevante

LMM: resultado do modelo

```
summary(lmm01p)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Richness ~ NAP + fExp + NAP:fExp + (1 | Beach)
##    Data: praia
##
## REML criterion at convergence: 224.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.4849 -0.4161 -0.0770  0.1521  3.7313
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Beach    (Intercept) 3.307    1.819
##  Residual             8.660    2.943
## Number of obs: 45, groups:  Beach, 9
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   8.8611     1.0208   8.680
## NAP          -3.4637     0.6279  -5.517
## fExp11       -5.2556     1.5452  -3.401
## NAP:fExp11    2.0005     0.9461   2.114
##
```

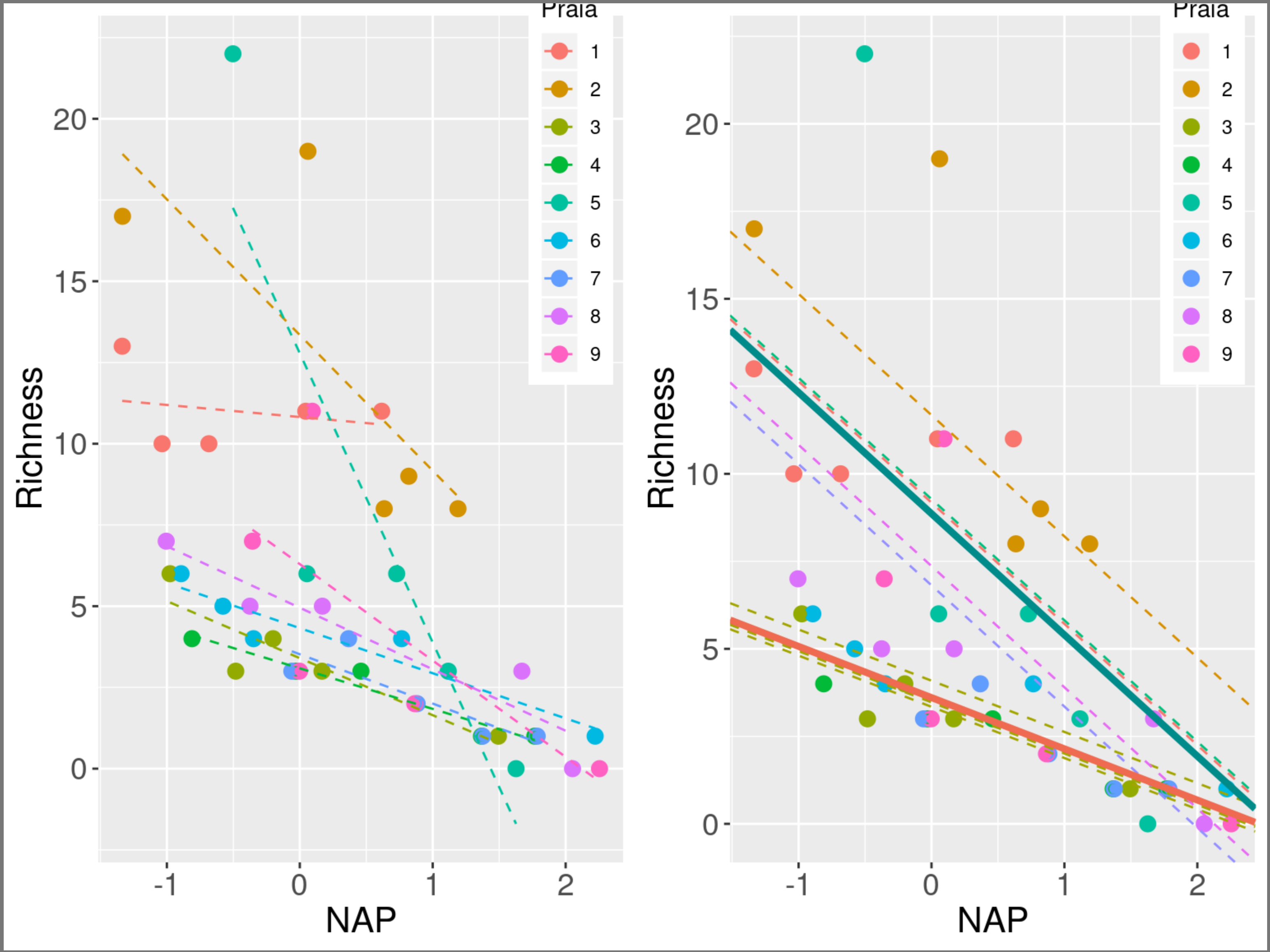
```
## Correlation of Fixed Effects:
##              (Intr) NAP    fExp11
```

LMM: resultado do modelo

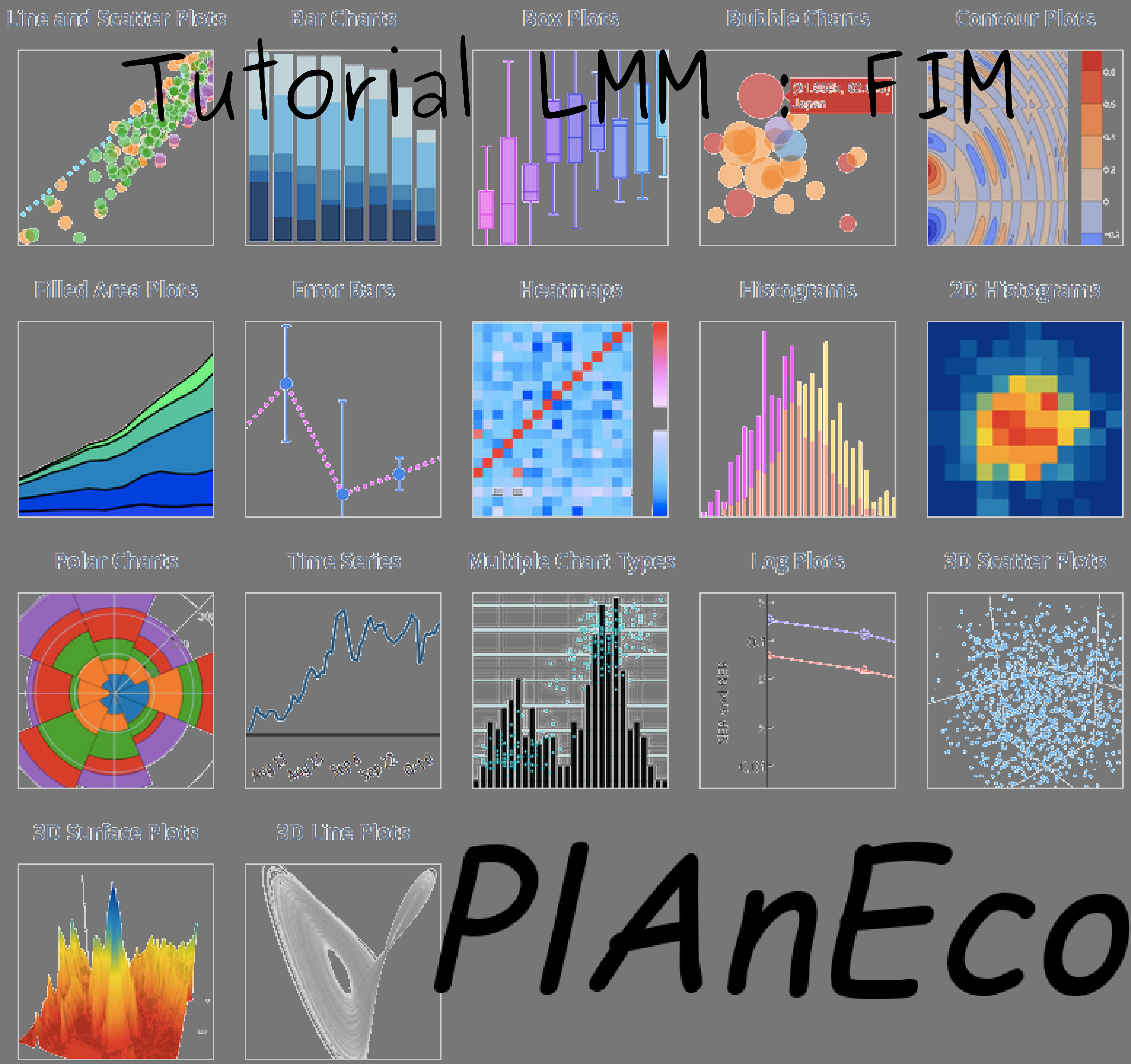
```
confint(lmm01p)

## Computing profile confidence intervals ...
##                  2.5 %    97.5 %
## .sig01       0.0000000  3.145294
## .sigma       2.3114668  3.681773
## (Intercept)  6.9045813 10.804288
## NAP         -4.7299177 -2.275315
## fExp11      -8.1969707 -2.303772
## NAP:fExp11   0.1919491  3.877650
```

LMM: resultado do modelo

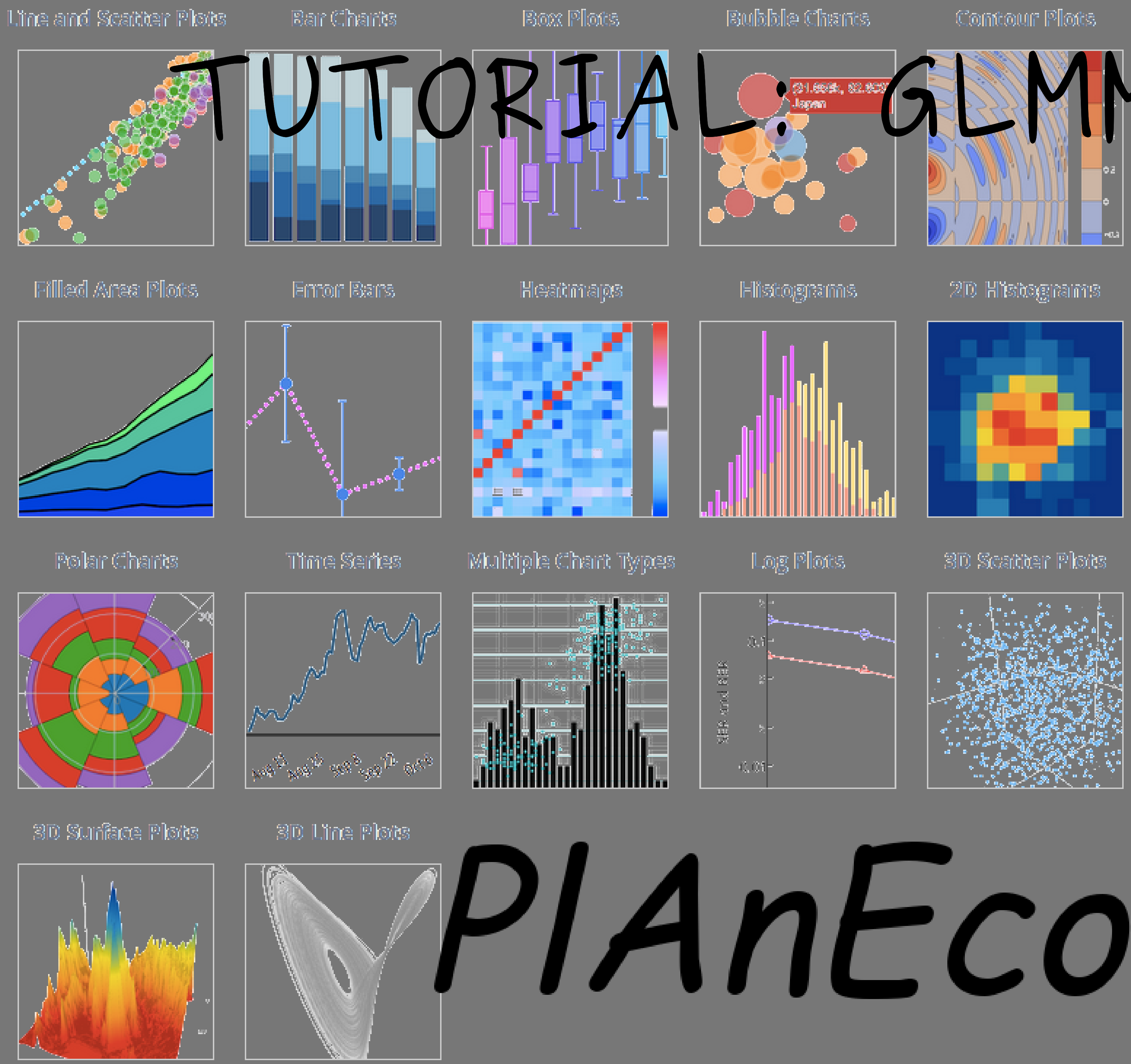


Tutorial LMM : FIM



PlAnEco

TUTORIAL: GLMM



PlAnEco

GLMM: riqueza da praia

Entre as premissas dos modelo mistos lineares estão aquelas básicas dos modelos lineares. Apesar de poder tratar alguma corelação entre observações, estamos trando de dados com:

- relação linear,
- homogeneidade de variância e
- resíduos com estrutura Gaussiana

A estrutura do componente dos efeitos fixos do modelo são:

$$y = a + bx + \epsilon$$

$$\epsilon = N(0, \sigma)$$

GLMM: generalizado misto

Para tratar com variáveis resposta de outra natureza, como contagem e proporção, precisamos usar a estrutura dos modelos generalizados, ou seja:

- **função de ligação**
- **outra estrutura do resíduo**: poisson, binomial, gama..

Da mesma forma que os modelos generalizados, a resposta do modelo estará na escala do **preditor linear**, portanto muitas vezes é necessário usar a função inversa para interpretar o modelo.

Modelos Lineares Misto Generalizados

Para usarmos os modelos generalizados precisamos do pacote `lmer4`

Instalando o pacote lmer4

```
install.packages("lmer4")
```

Carregando o pacote

```
library("lmer4")
```

Exemplo: riqueza das praias

A riqueza da macrofauna varia em função da altura e exposição da praia

- 45 observações
- 9 praias
  - NAP: contínuo, altura em relação ao NMM
- fExp: 2 níveis de exposição da praia (10 , 11)

Sample	Beach	Richness	NAP	Praia	fExp
1	1	11	0.045	1	low
6	2	8	1.190	2	low
11	3	6	−0.976	3	high
16	4	1	1.768	4	high
21	5	3	1.117	5	low
26	6	5	−0.578	6	high
31	7	2	0.883	7	high
36	8	3	1.671	8	low
41	9	7	−0.356	9	low

GLMM: riqueza da praia

Comparando efeitos aleatórios

Efeito aleatório na inclinação (NAP|Beach)

```
glmm00s <- glmer(Richness ~ NAP*fExp + (1|Beach) +
  (NAP|Beach),
                 family=poisson, data=praia)
glmm00i <- glmer(Richness ~ NAP*fExp + (1|Beach),
                 family = poisson, data=praia)
anova(glmm00i, glmm00s, refit= FALSE)
```

```
## Data: praia
## Models:
## glmm00i: Richness ~ NAP * fExp + (1 | Beach)
## glmm00s: Richness ~ NAP * fExp + (1 | Beach) + (NAP | Beach
##         Df    AIC    BIC   logLik deviance  Chisq Chi Df Pr
## glmm00i  5 215.42 224.45 -102.709   205.42
## glmm00s  8 215.11 229.57  -99.557   199.11 6.3042      3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Modelar a inclinção não trás informação relevante ao modelo

GLMM: riqueza da praia

Comparando efeitos aleatórios

Efeito aletório no intercepto??

```
glm00 <- glm(Richness ~ NAP*fExp, data=praia, family =
  poisson)
anova(glmm00i, glm00, refit= FALSE)
```

```
## Data: praia
## Models:
## glm00: Richness ~ NAP * fExp
## glmm00i: Richness ~ NAP * fExp + (1 | Beach)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(
## glm00    4 221.69 228.92 -106.84   213.69
## glmm00i  5 215.42 224.45 -102.71   205.42 8.2723      1    0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

A estrutura aleatória do intercepto deve ser mantido, mesmo que não fosse significativo

Simplificando: fatores fixos

```
glmm00 <- glmer(Richness ~ NAP + fExp + NAP:fExp + (1|Beach),
                family="poisson", data=praia)
glmm01 <- glmer(Richness ~ NAP + fExp +  (1|Beach),
                family="poisson", data=praia)
anova(glmm00, glmm01)
```

```
## Data: praia
## Models:
## glmm01: Richness ~ NAP + fExp + (1 | Beach)
## glmm00: Richness ~ NAP + fExp + NAP:fExp + (1 | Beach)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>
## glmm01  4 213.56 220.78 -102.78   205.56
## glmm00  5 215.42 224.45 -102.71   205.42 0.1366      1
```

Interação NAP:fExp não é informativa

Removendo NAP

```
glmm02 <- glmer(Richness ~  fExp +  (1|Beach),
                family="poisson", data=praia)
anova(glmm01, glmm02)
```

```
## Data: praia
## Models:
## glmm02: Richness ~ fExp + (1 | Beach)
## glmm01: Richness ~ NAP + fExp + (1 | Beach)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>
## glmm02  3 260.22 265.64 -127.11   254.22
## glmm01  4 213.56 220.78 -102.78   205.56 48.663      1  3.0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

NAP é muito informativa

Removendo fExp

```r
glmm03 <- glmer(Richness ~ NAP +  (1|Beach),
                family="poisson", data=praia)
anova(glmm01, glmm03)
```

```
## Data: praia
## Models:
## glmm03: Richness ~ NAP + (1 | Beach)
## glmm01: Richness ~ NAP + fExp + (1 | Beach)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>
## glmm03  3 220.78 226.20 -107.39   214.78
## glmm01  4 213.56 220.78 -102.78   205.56 9.2293      1   0.
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Exposição também é informativa e deve ser retida

Compara com o modelo MÍNIMO (NULO)

```
glmmNULL <- glmer(Richness ~ 1 +  (1|Beach),
                  family="poisson", data=praia)
anova(glmm01, glmmNULL)
```

```
## Data: praia
## Models:
## glmmNULL: Richness ~ 1 + (1 | Beach)
## glmm01: Richness ~ NAP + fExp + (1 | Beach)
##          Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr
## glmmNULL  2 265.88 269.49 -130.94   261.88
## glmm01    4 213.56 220.78 -102.78   205.56 56.323      2  5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

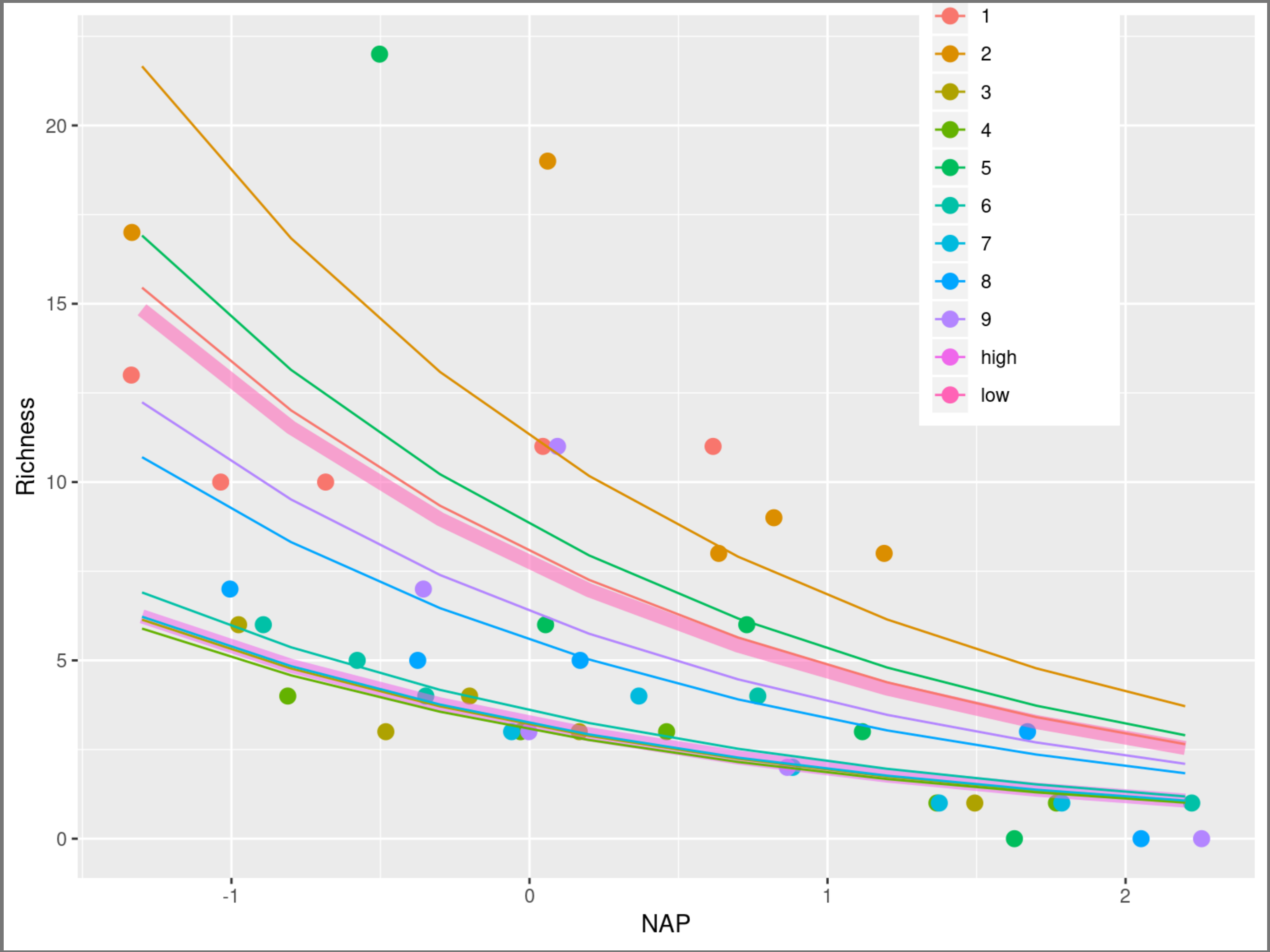
Resultado Final

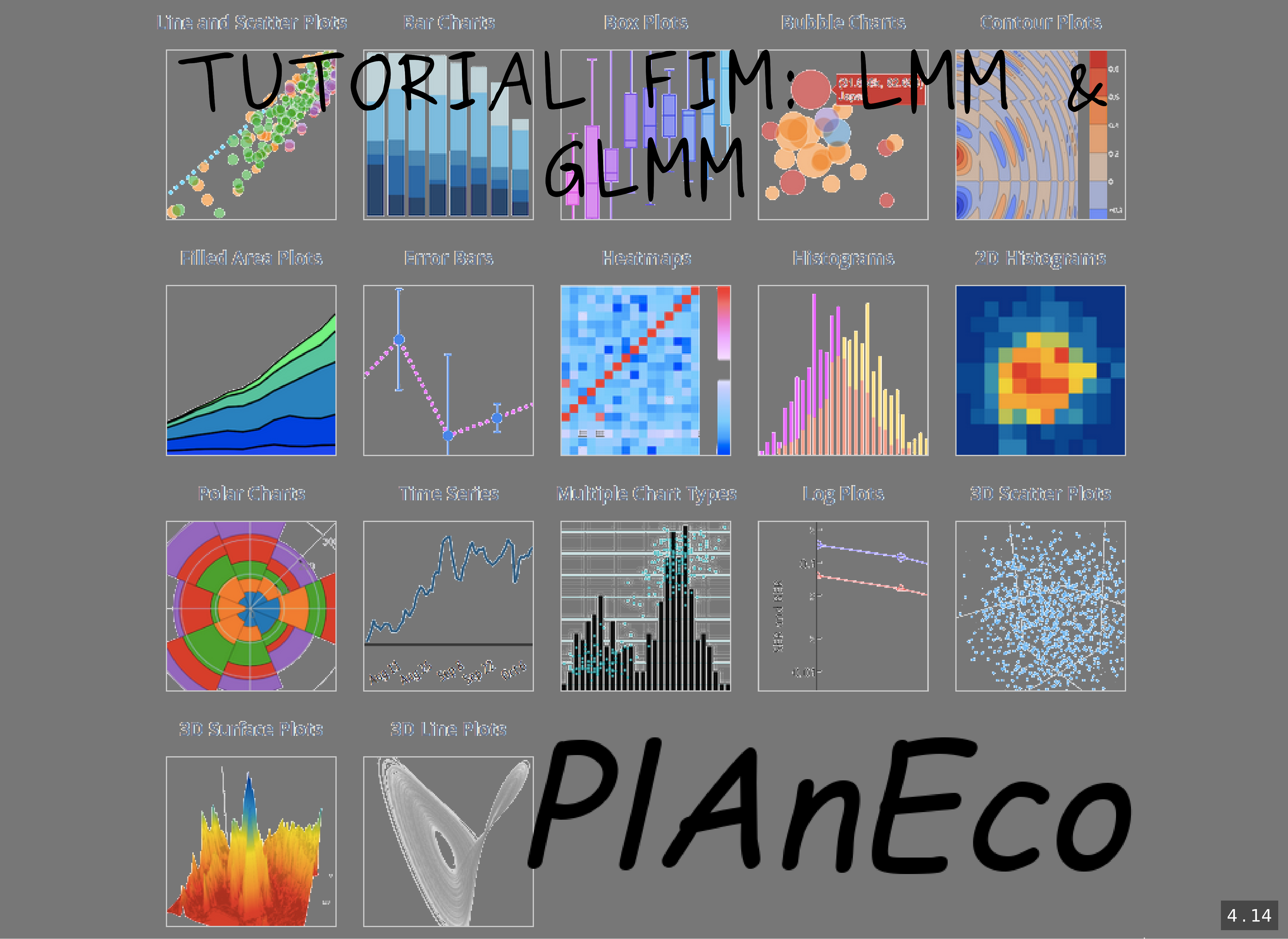
```r
summary(glmm01)
```

```
## Generalized linear mixed model fit by maximum likelihood (L
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: Richness ~ NAP + fExp + (1 | Beach)
##    Data: praia
## 
##      AIC      BIC   logLik deviance df.resid 
##    213.6    220.8   -102.8    205.6       41 
## 
## Scaled residuals: 
##     Min      1Q  Median      3Q     Max 
## -1.9673 -0.6611 -0.1579  0.3434  3.1746 
## 
## Random effects:
##  Groups Name        Variance Std.Dev.
##  Beach  (Intercept) 0.06044  0.2458  
## Number of obs: 45, groups:  Beach, 9
## 
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)  2.04183    0.13406  15.231  < 2e-16 ***
## NAP         -0.50383    0.07395  -6.814 9.52e-12 ***
## fExphigh    -0.86699    0.22332  -3.882 0.000103 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
## 
## Correlation of Fixed Effects:
##          (Intr) NAP
```

```
##             (Intr) NAP
## NAP          0.058
```

Gráfico Final




TUTORIAL FIM: LMM & GLMM
Line and Scatter Plots
Bar Charts
Box Plots
Bubble Charts
Contour Plots
Filled Area Plots
Error Bars
Heatmaps
Histograms
2D Histograms
Polar Charts
Time Series
Multiple Chart Types
Log Plots
3D Scatter Plots
3D Surface Plots
3D Line Plots
PlAnEco
4.14