

Linear models

Example dataset: forest trees

- Go to <https://tinyurl.com/treesdata>

```
trees <- read.csv("data-raw/trees.csv")  
head(trees)
```

	site	dbh	height	sex	dead
1	4	29.68	36.1	male	0
2	5	33.29	42.3	male	0
3	2	28.03	41.9	female	0
4	5	39.86	46.5	female	0
5	1	47.94	43.9	female	0
6	1	10.82	26.2	male	0

Example dataset: forest trees

- ▶ Go to <https://tinyurl.com/treesdata>
- ▶ Download zip file and uncompress (within your project folder!)

```
trees <- read.csv("data-raw/trees.csv")  
head(trees)
```

	site	dbh	height	sex	dead
1	4	29.68	36.1	male	0
2	5	33.29	42.3	male	0
3	2	28.03	41.9	female	0
4	5	39.86	46.5	female	0
5	1	47.94	43.9	female	0
6	1	10.82	26.2	male	0

Questions

- ▶ What is the relationship between DBH and height?

Questions

- ▶ What is the relationship between DBH and height?
- ▶ Do taller trees have bigger trunks?

Questions

- ▶ What is the relationship between DBH and height?
- ▶ Do taller trees have bigger trunks?
- ▶ Can we predict height from DBH? How well?

Always plot your data first!

Always plot your data first!

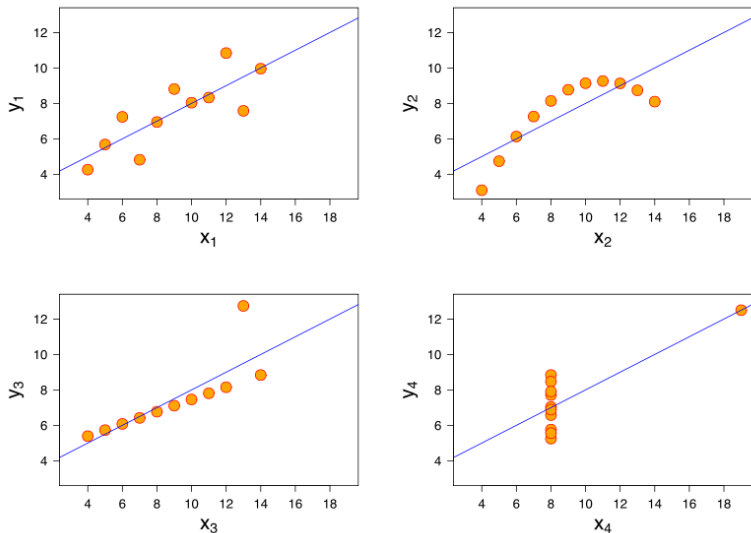
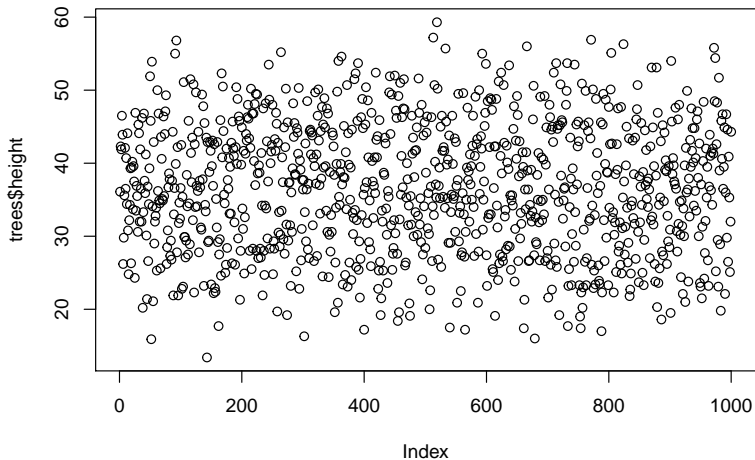


Figure 1

Exploratory Data Analysis (EDA)

Outliers

```
plot(trees$height)
```



Outliers impact on regression

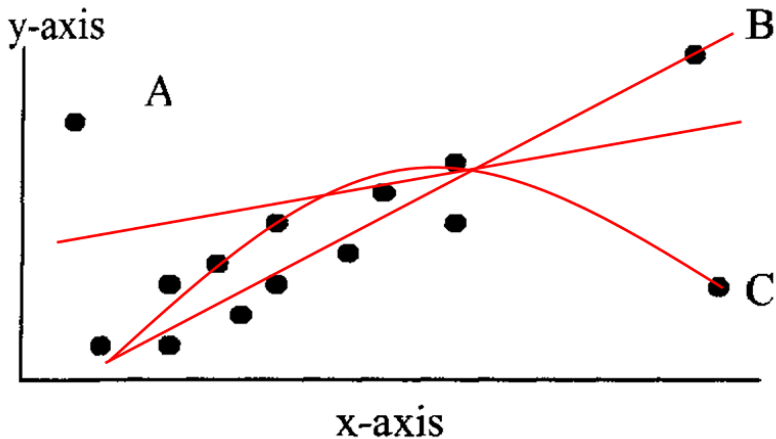
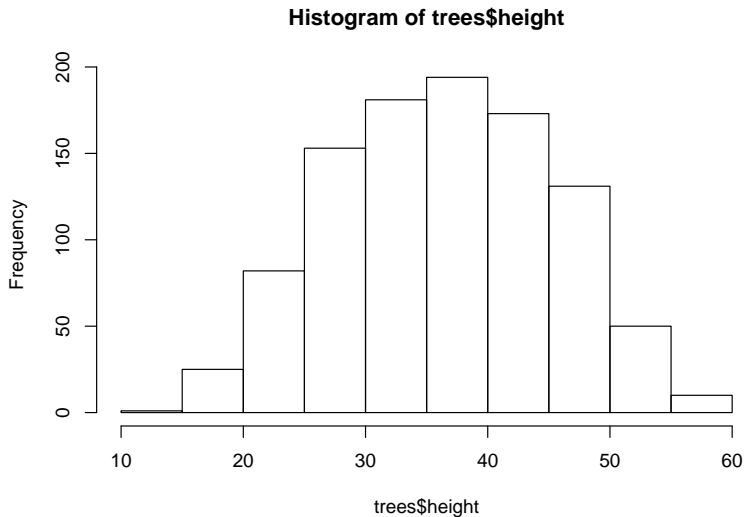


Figure 2

See <http://rpsychologist.com/d3/correlation/>

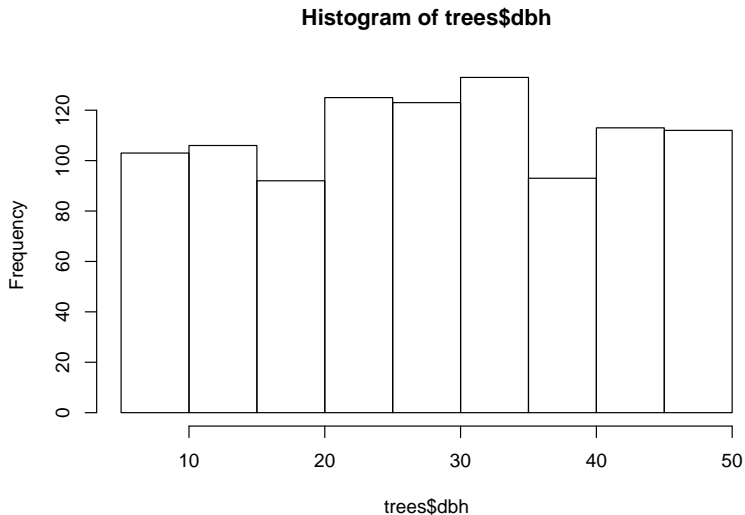
Histogram of response variable

```
hist(trees$height)
```



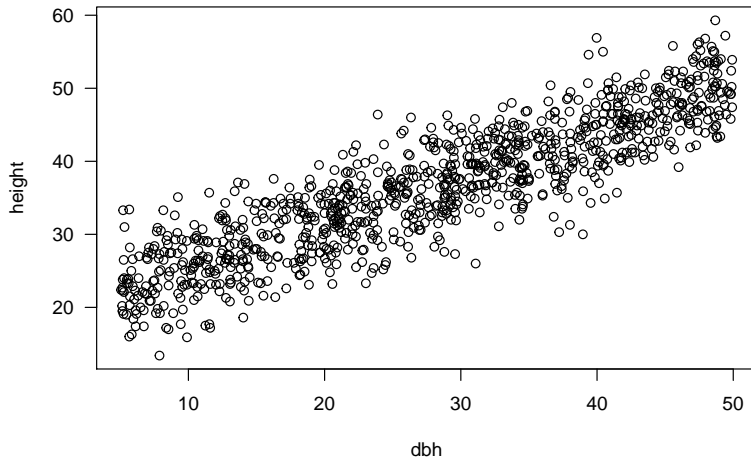
Histogram of predictor variable

```
hist(trees$dbh)
```



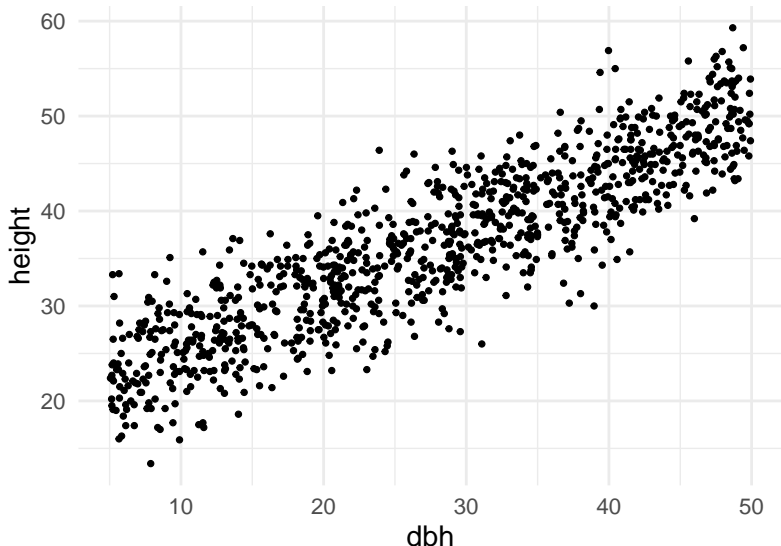
Scatterplot

```
plot(height ~ dbh, data = trees, las = 1)
```



Scatterplot

```
ggplot(trees) + aes(dbh, height) +  
  geom_point()
```



Model fitting

Now fit model

Hint: `lm`

Now fit model

Hint: `lm`

```
m1 <- lm(height ~ dbh, data = trees)
```

which corresponds to

$$\begin{aligned} \text{Height}_i &= a + b \cdot \text{DBH}_i + \varepsilon_i \\ \varepsilon_i &\sim N(0, \sigma^2) \end{aligned}$$

Model interpretation

What does this mean?

Call:

```
lm(formula = height ~ dbh, data = trees)
```

Residuals:

Min	1Q	Median	3Q	Max
-13.3270	-2.8978	0.1057	2.7924	12.9511

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	19.33920	0.31064	62.26	<2e-16 ***
dbh	0.61570	0.01013	60.79	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.093 on 998 degrees of freedom

Multiple R-squared: 0.7874, Adjusted R-squared: 0.7871

F-statistic: 3695 on 1 and 998 DF, p-value: < 2.2e-16

Avoid dichotomania of statistical significance



EDITORIAL • 20 MARCH 2019

It's time to talk about ditching statistical significance

- ▶ 'Never conclude there is 'no difference' or 'no association' just because $p > 0.05$ or CI includes zero'

Avoid dichotomania of statistical significance



EDITORIAL • 20 MARCH 2019

It's time to talk about ditching statistical significance

- ▶ 'Never conclude there is 'no difference' or 'no association' just because $p > 0.05$ or CI includes zero'
- ▶ Estimate and communicate effect sizes and their uncertainty

Avoid dichotomania of statistical significance



EDITORIAL • 20 MARCH 2019

It's time to talk about ditching statistical significance

- ▶ 'Never conclude there is 'no difference' or 'no association' just because $p > 0.05$ or CI includes zero'
- ▶ Estimate and communicate effect sizes and their uncertainty
- ▶ <https://doi.org/10.1038/d41586-019-00857-9>

Communicating results

We found a significant positive relationship between DBH and Height
($p < 0.05$) ($b = 0.61$, $SE = 0.01$).

Presenting model results

```
kable(xtable::xtable(m1), digits = 2)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	19.34	0.31	62.26	0
dbh	0.62	0.01	60.79	0

Presenting model results

```
texreg::texreg(m1, single.row = TRUE)
```

	Model 1
(Intercept)	19.34 (0.31)***
dbh	0.62 (0.01)***
R ²	0.79
Adj. R ²	0.79
Num. obs.	1000
RMSE	4.09

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 2: Statistical models

Retrieving model coefficients

```
coef(m1)
```

(Intercept)	dbh
19.3391968	0.6157036

Tidy up model coefficients with broom

```
library(broom)
tidy(m1)
```

```
# A tibble: 2 x 5
```

	term <chr>	estimate <dbl>	std.error <dbl>	statistic <dbl>	p.value <dbl>
1	(Intercept)	19.3	0.311	62.3	0
2	dbh	0.616	0.0101	60.8	0

```
glance(m1)
```

```
# A tibble: 1 x 11
```

	r.squared <dbl>	adj.r.squared <dbl>	sigma <dbl>	statistic <dbl>	p.value <dbl>	df <int>	logLik <dbl>	AIC <dbl>	BIC <dbl>
1	0.787	0.787	4.09	3695.	0	2	-2827.	5660.	5675.

```
# ... with 2 more variables: deviance <dbl>, df.residual <int>
```

Confidence intervals

```
confint(m1)
```

	2.5 %	97.5 %
(Intercept)	18.7296053	19.948788
dbh	0.5958282	0.635579

Using effects package

```
library(effects)  
summary(allEffects(m1))
```

model: height ~ dbh

dbh effect

dbh

	5	20	30	40	50
	22.41771	31.65327	37.81030	43.96734	50.12438

Lower 95 Percent Confidence Limits

dbh

	5	20	30	40	50
	21.89682	31.35487	37.55287	43.61733	49.61669

Upper 95 Percent Confidence Limits

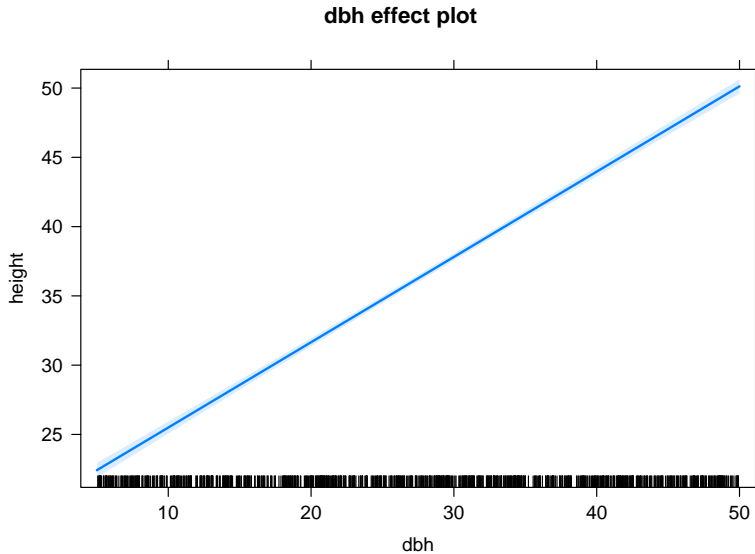
dbh

	5	20	30	40	50
	22.93861	31.95167	38.06774	44.31735	50.63207

Visualising fitted model

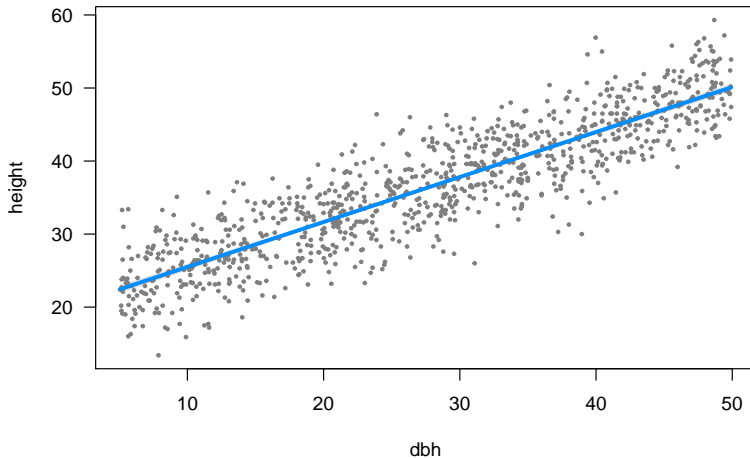
Plot effects

```
plot(allEffects(m1))
```



Plot model (visreg)

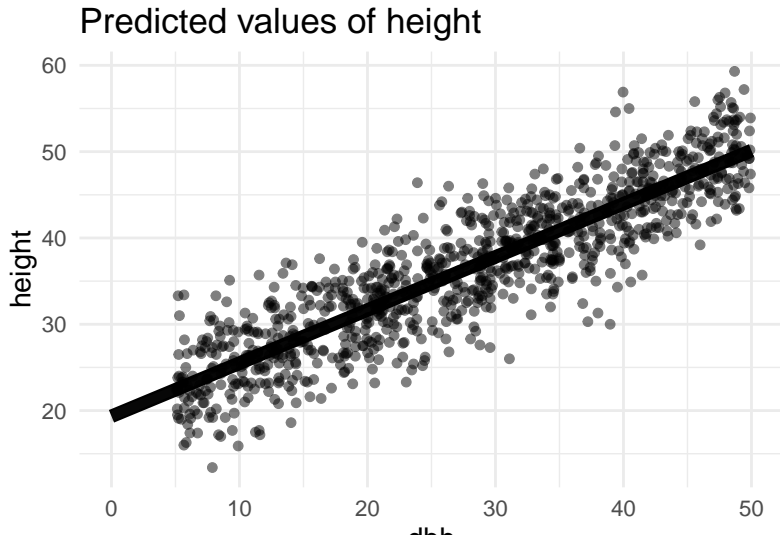
```
library(visreg)  
visreg(m1)
```



Plot model (sjPlot - ggplot2)

```
sjPlot::plot_model(m1, type = "eff", show.data = TRUE, line.size
```

\$dbh



Model checking

Linear model assumptions

- ▶ Linearity (transformations, GAM...)

Linear model assumptions

- ▶ Linearity (transformations, GAM...)
- ▶ Residuals:

Linear model assumptions

- ▶ Linearity (transformations, GAM...)
- ▶ Residuals:
 - ▶ Independent

Linear model assumptions

- ▶ Linearity (transformations, GAM...)
- ▶ Residuals:
 - ▶ Independent
 - ▶ Equal variance

Linear model assumptions

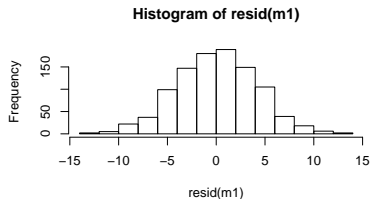
- ▶ Linearity (transformations, GAM...)
- ▶ Residuals:
 - ▶ Independent
 - ▶ Equal variance
 - ▶ Normal

Linear model assumptions

- ▶ Linearity (transformations, GAM...)
- ▶ Residuals:
 - ▶ Independent
 - ▶ Equal variance
 - ▶ Normal
- ▶ No measurement error in predictors

Are residuals normal?

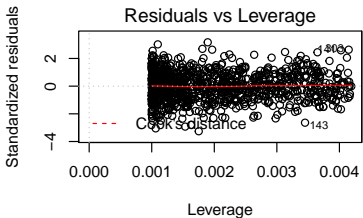
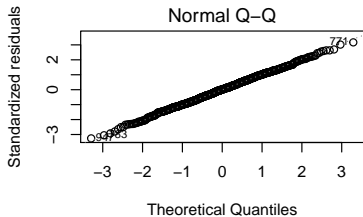
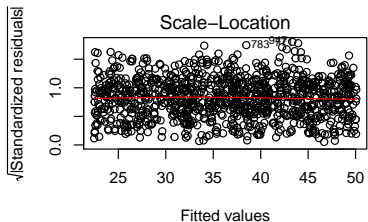
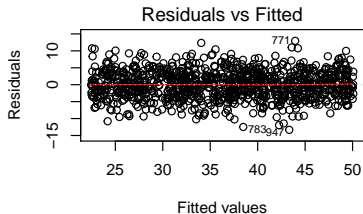
```
hist(resid(m1))
```



```
lm(formula = height ~ dbh, data = trees)
      coef.est coef.se
(Intercept) 19.34    0.31
      dbh      0.62    0.01
---
n = 1000, k = 2
residual sd = 4.09, R-Squared = 0.79
```

SD of residuals = 4.09 coincides with estimate of σ .

Model checking: residuals

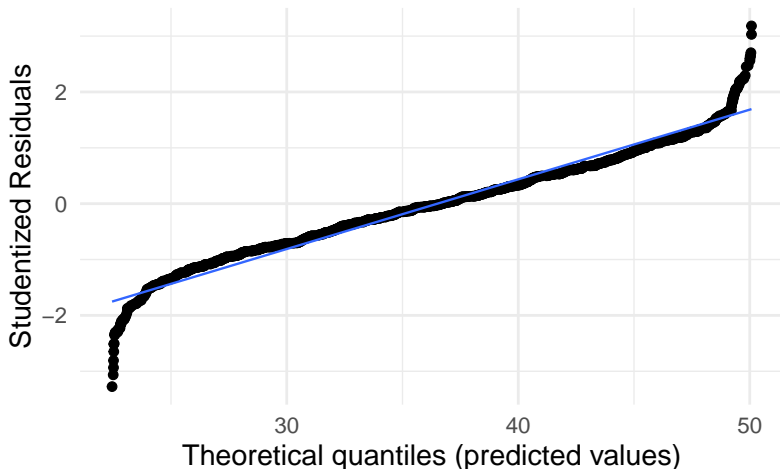


Model checking (sjPlot)

```
plot_model(m1, type = "diag")[[1]]
```

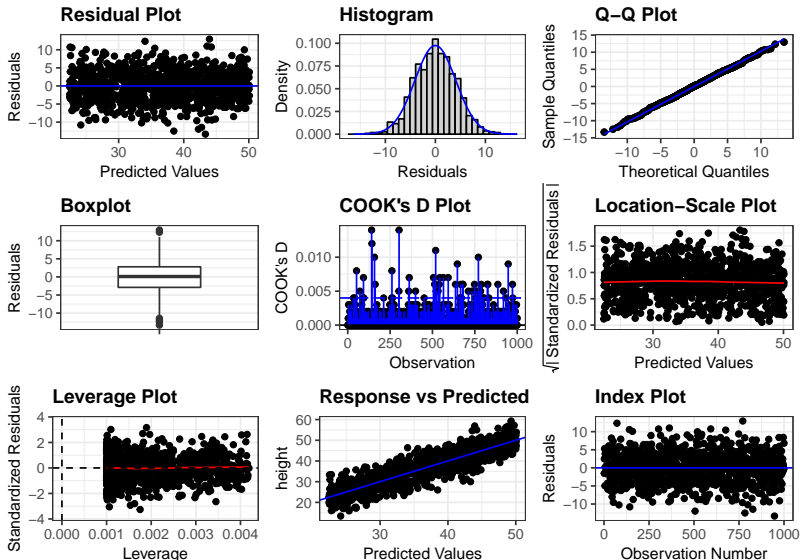
Non-normality of residuals and outliers

Dots should be plotted along the line



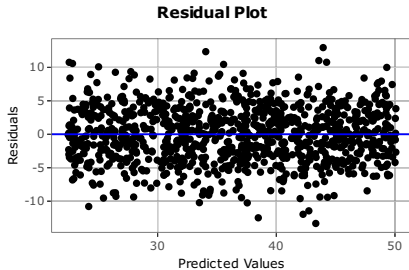
Model checking (ggResidpanel)

```
ggResidpanel::resid_panel(m1, plots = "all")
```



Interactive model checking (ggResidpanel)

```
ggResidpanel::resid_interact(m1)
```



Using model for prediction

How good is the model in predicting tree height?

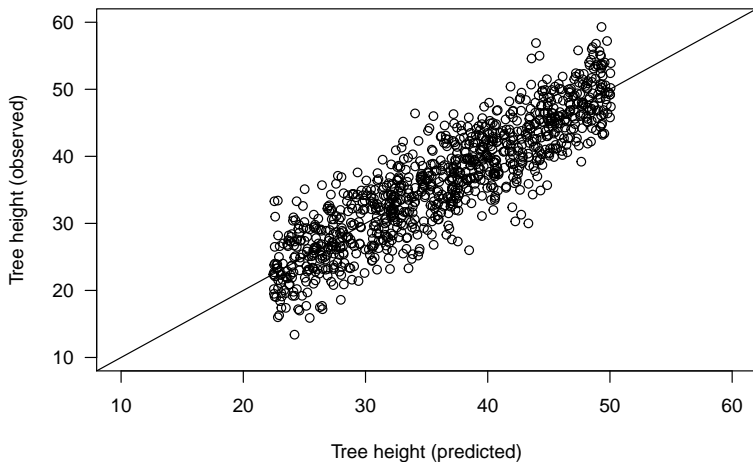
fitted gives predictions for each observation

```
trees$height.pred <- fitted(m1)
head(trees)
```

	site	dbh	height	sex	dead	height.pred
1	4	29.68	36.1	male	0	37.61328
2	5	33.29	42.3	male	0	39.83597
3	2	28.03	41.9	female	0	36.59737
4	5	39.86	46.5	female	0	43.88114
5	1	47.94	43.9	female	0	48.85603
6	1	10.82	26.2	male	0	26.00111

Calibration plot: Observed vs Predicted values

```
plot(trees$height.pred, trees$height,  
     xlab = "Tree height (predicted)", ylab = "Tree height (observed)")
```



Using fitted model for prediction

Q: Expected tree height if $\text{DBH} = 39 \text{ cm}$?

```
new.dbh <- data.frame(dbh = c(39))  
predict(m1, new.dbh, se.fit = TRUE)
```

```
$fit
```

```
1
```

```
43.35164
```

```
$se.fit
```

```
[1] 0.1715514
```

```
$df
```

```
[1] 998
```

```
$residual.scale
```

```
[1] 4.092629
```

Using fitted model for prediction

Q: Expected tree height if DBH = 39 cm?

```
predict(m1, new.dbh, interval = "confidence")
```

	fit	lwr	upr
1	43.35164	43.01499	43.68828

```
predict(m1, new.dbh, interval = "prediction")
```

	fit	lwr	upr
1	43.35164	35.31344	51.38983

Workflow

- ▶ **Visualise data**

Workflow

- ▶ **Visualise data**
- ▶ **Understand fitted model** (`summary`, `allEffects...`)

Workflow

- ▶ **Visualise data**
- ▶ **Understand fitted model** (`summary`, `allEffects...`)
- ▶ **Visualise model** (`plot(allEffects)`, `visreg`, `plot_model...`)

Workflow

- ▶ **Visualise data**
- ▶ **Understand fitted model** (`summary`, `allEffects`...)
- ▶ **Visualise model** (`plot(allEffects)`, `visreg`, `plot_model`...)
- ▶ **Check model** (`plot`, `resid_panel`, `calibration plot`...)

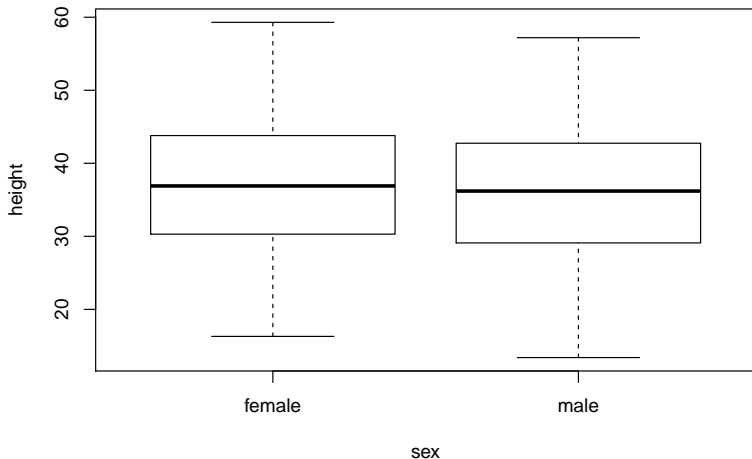
Workflow

- ▶ **Visualise data**
- ▶ **Understand fitted model** (`summary`, `allEffects`...)
- ▶ **Visualise model** (`plot(allEffects)`, `visreg`, `plot_model`...)
- ▶ **Check model** (`plot`, `resid_panel`, `calibration plot`...)
- ▶ **Predict** (`fitted`, `predict`)

Categorical predictors (factors)

Q: Does tree height vary with sex?

```
plot(height ~ sex, data = trees)
```



Model height ~ sex

```
m2 <- lm(height ~ sex, data = trees)
```

Call:

```
lm(formula = height ~ sex, data = trees)
```

Residuals:

Min	1Q	Median	3Q	Max
-22.6881	-6.7881	-0.0097	6.7261	22.3687

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36.9312	0.3981	92.778	<2e-16 ***
sexmale	-0.8432	0.5607	-1.504	0.133

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.865 on 998 degrees of freedom

Multiple R-squared: 0.002261, Adjusted R-squared: 0.001261

F-statistic: 2.261 on 1 and 998 DF, p-value: 0.133

Linear model with categorical predictors

```
m2 <- lm(height ~ sex, data = trees)
```

corresponds to

$$\begin{aligned} \text{Height}_i &= a + b_{\text{male}} + \varepsilon_i \\ \varepsilon_i &\sim N(0, \sigma^2) \end{aligned}$$

Model height ~ sex

```
m2 <- lm(height ~ sex, data = trees)
```

Call:

```
lm(formula = height ~ sex, data = trees)
```

Residuals:

Min	1Q	Median	3Q	Max
-22.6881	-6.7881	-0.0097	6.7261	22.3687

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36.9312	0.3981	92.778	<2e-16 ***
sexmale	-0.8432	0.5607	-1.504	0.133

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.865 on 998 degrees of freedom

Multiple R-squared: 0.002261, Adjusted R-squared: 0.001261

F-statistic: 2.261 on 1 and 998 DF, p-value: 0.133

Presenting model results

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36.93	0.40	92.78	0.00
sexmale	-0.84	0.56	-1.50	0.13

Effects: Height ~ sex

Compare CIs

```
summary(allEffects(m2))
```

```
model: height ~ sex
```

```
sex effect
```

```
sex
```

```
female      male
```

```
36.93125 36.08810
```

```
Lower 95 Percent Confidence Limits
```

```
sex
```

```
female      male
```

```
36.15012 35.31319
```

```
Upper 95 Percent Confidence Limits
```

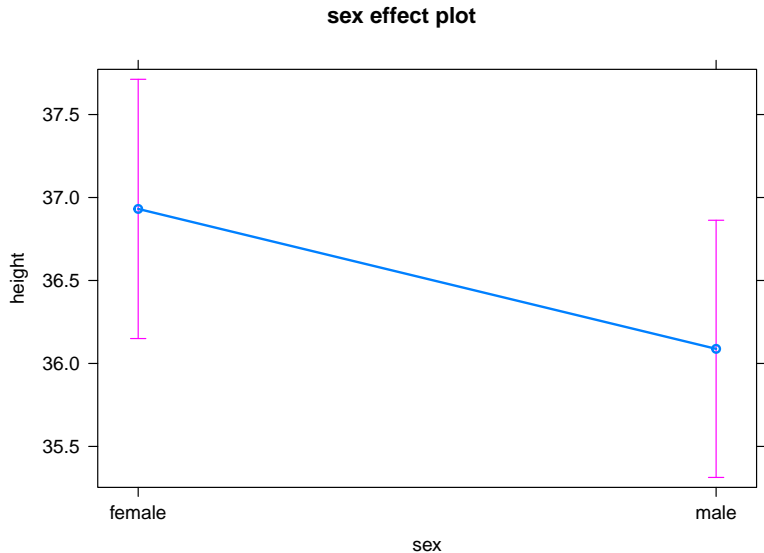
```
sex
```

```
female      male
```

```
37.71238 36.86300
```

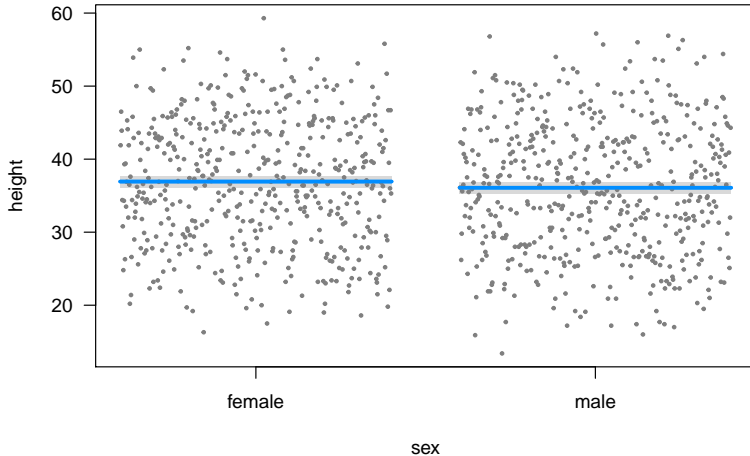
Plot

```
plot(allEffects(m2))
```



Plot (visreg)

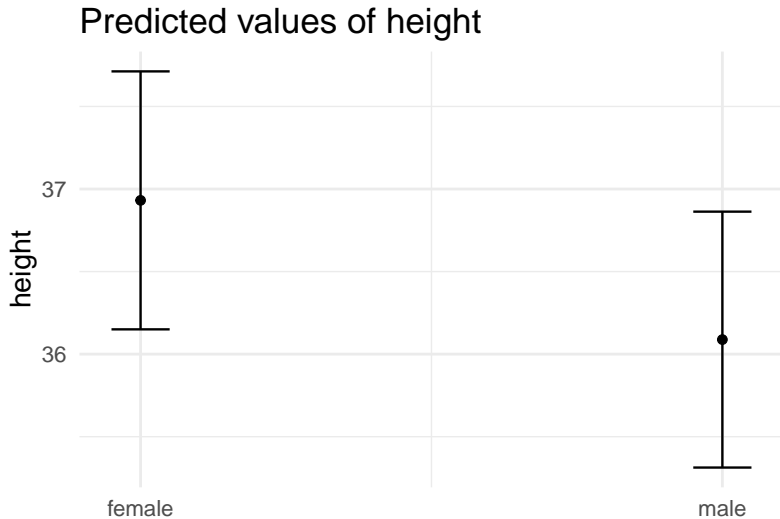
```
visreg(m2)
```



Plot model (sjPlot)

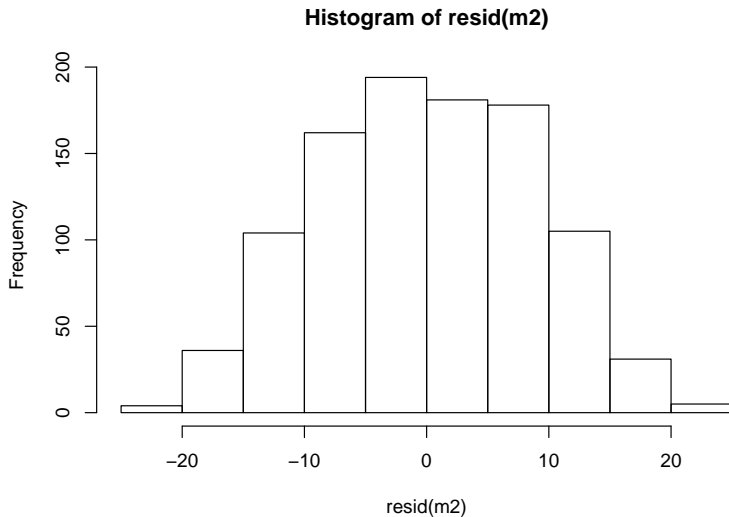
```
plot_model(m2, type = "eff")
```

\$sex

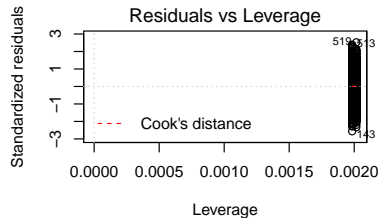
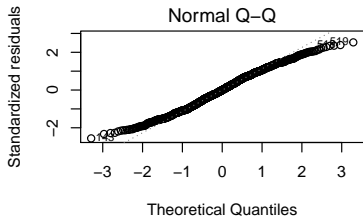
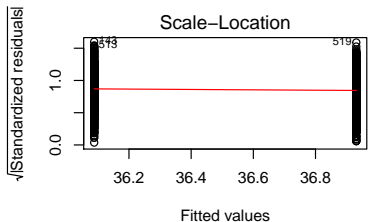
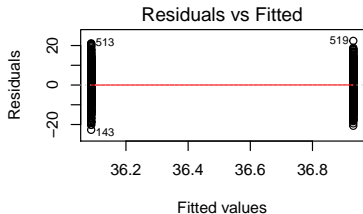


Model checking: residuals

```
hist(resid(m2))
```

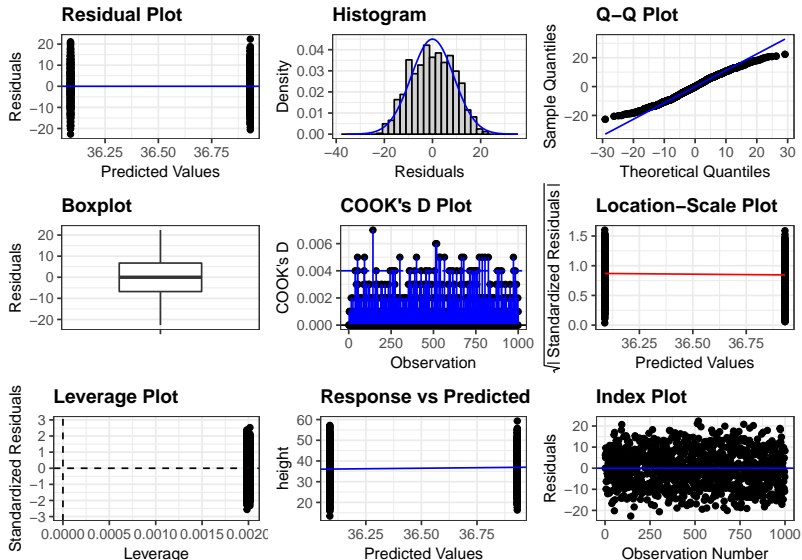


Model checking: residuals



Model checking (ggResidpanel)

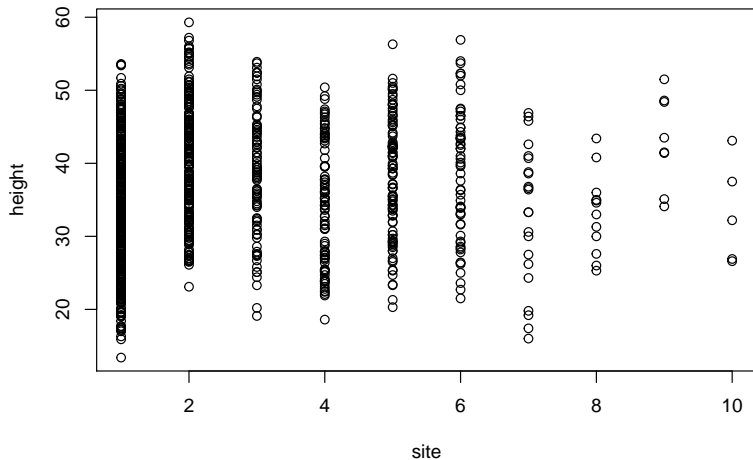
```
ggResidpanel::resid_panel(m2, plots = "all")
```



Q: Does height differ among field sites?

Plot data first

```
plot(height ~ site, data = trees)
```



Linear model with categorical predictors

```
m3 <- lm(height ~ site, data = trees)
```

$$y_i = a + b_{site2} + c_{site3} + d_{site4} + e_{site5} + \dots + \varepsilon_i$$
$$\varepsilon_i \sim N(0, \sigma^2)$$

Model Height ~ site

All right here?

```
m3 <- lm(height ~ site, data = trees)
```

Call:

```
lm(formula = height ~ site, data = trees)
```

Residuals:

Min	1Q	Median	3Q	Max
-22.4498	-6.7049	0.0709	6.7537	23.0640

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	35.4636	0.4730	74.975	< 2e-16 ***
site	0.3862	0.1413	2.733	0.00639 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.842 on 998 degrees of freedom

Multiple R-squared: 0.007429, Adjusted R-squared: 0.006435

site is a factor!

```
trees$site <- as.factor(trees$site)
```

Model Height ~ site

Call:

```
lm(formula = height ~ site, data = trees)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-20.4416	-6.9004	0.0379	6.3051	19.7584

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	33.8416	0.4266	79.329	< 2e-16	***
site2	6.3411	0.7126	8.899	< 2e-16	***
site3	4.9991	0.9828	5.086	4.36e-07	***
site4	0.5329	0.9872	0.540	0.58949	
site5	4.3723	0.9425	4.639	3.97e-06	***
site6	4.7601	1.1709	4.065	5.18e-05	***
site7	-0.7416	1.8506	-0.401	0.68871	
site8	-0.6832	2.4753	-0.276	0.78258	
site9	9.1709	3.0165	3.040	0.00243	**
site10	-0.5816	3.8013	-0.153	0.87843	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.446 on 990 degrees of freedom

Multiple R-squared: 0.1016, Adjusted R-squared: 0.09344

F-statistic: 12.44 on 9 and 990 DF, p-value: < 2.2e-16

Presenting model results

```
kable(xtable::xtable(m3), digits = 2)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	33.84	0.43	79.33	0.00
site2	6.34	0.71	8.90	0.00
site3	5.00	0.98	5.09	0.00
site4	0.53	0.99	0.54	0.59
site5	4.37	0.94	4.64	0.00
site6	4.76	1.17	4.07	0.00
site7	-0.74	1.85	-0.40	0.69
site8	-0.68	2.48	-0.28	0.78
site9	9.17	3.02	3.04	0.00
site10	-0.58	3.80	-0.15	0.88

Estimated tree heights for each site

```
summary(allEffects(m3))
```

```
model: height ~ site
```

```
site effect
```

```
site
```

1	2	3	4	5	6	7	8
33.84158	40.18265	38.84066	34.37444	38.21386	38.60167	33.10000	33.15833
9	10						
43.01250	33.26000						

```
Lower 95 Percent Confidence Limits
```

```
site
```

1	2	3	4	5	6	7	8
33.00444	39.06264	37.10317	32.62733	36.56463	36.46190	29.56629	28.37367
9	10						
37.15251	25.84764						

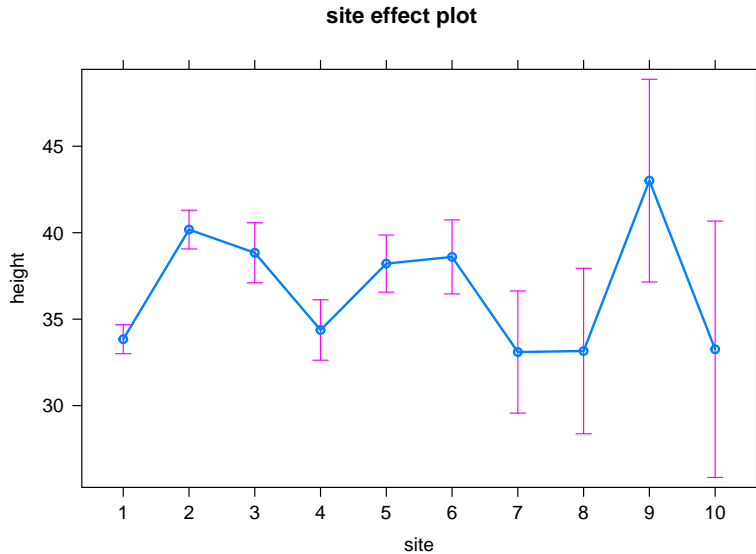
```
Upper 95 Percent Confidence Limits
```

```
site
```

1	2	3	4	5	6	7	8
34.67872	41.30265	40.57814	36.12156	39.86309	40.74143	36.63371	37.94299
9	10						
48.87249	40.67236						

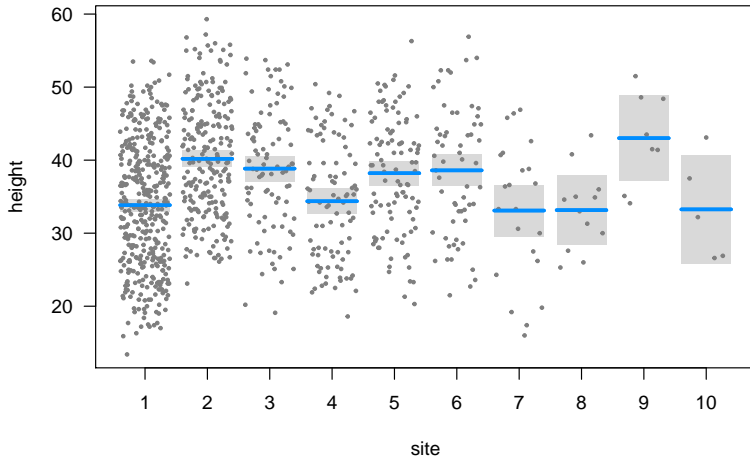
Plot

```
plot(allEffects(m3))
```



Plot (visreg)

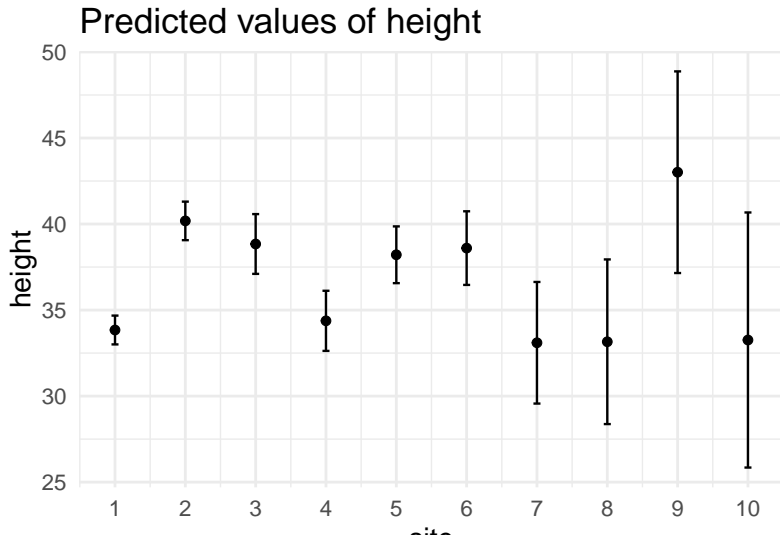
```
visreg(m3)
```



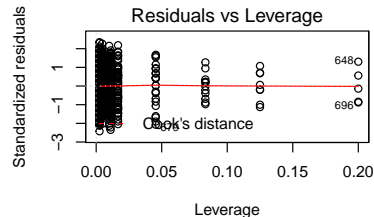
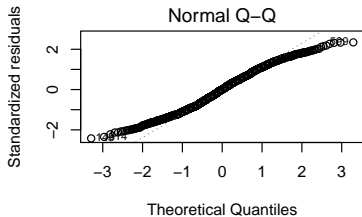
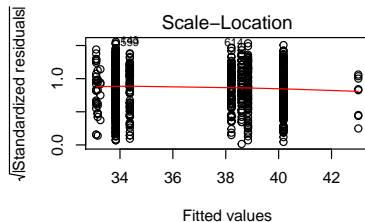
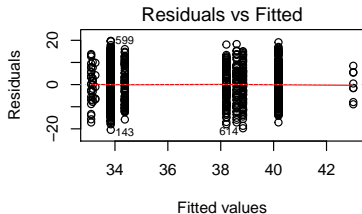
Plot model (sjPlot)

```
plot_model(m3, type = "eff")
```

\$site



Model checking: residuals



Combining continuous and categorical predictors

Predicting tree height based on dbh and site

```
lm(height ~ site + dbh, data = trees)
```

corresponds to

$$y_i = a + b_{site2} + c_{site3} + d_{site4} + e_{site5} + \dots + k \cdot DBH_i + \varepsilon_i$$
$$\varepsilon_i \sim N(0, \sigma^2)$$

Predicting tree height based on dbh and site

Call:

```
lm(formula = height ~ site + dbh, data = trees)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-10.1130	-1.9885	0.0582	2.0314	11.3320

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	16.699037	0.260565	64.088	< 2e-16	***
site2	6.504303	0.256730	25.335	< 2e-16	***
site3	4.357457	0.354181	12.303	< 2e-16	***
site4	1.934650	0.356102	5.433	6.98e-08	***
site5	3.637432	0.339688	10.708	< 2e-16	***
site6	4.204511	0.421906	9.966	< 2e-16	***
site7	-0.176193	0.666772	-0.264	0.7916	
site8	-5.312648	0.893603	-5.945	3.82e-09	***
site9	5.437049	1.087766	4.998	6.84e-07	***
site10	2.263338	1.369986	1.652	0.0988	.
dbh	0.617075	0.007574	81.473	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.043 on 989 degrees of freedom

Multiple R-squared: 0.8835, Adjusted R-squared: 0.8823

F-statistic: 750 on 10 and 989 DF, p-value: < 2.2e-16

Presenting model results

```
kable(xtable::xtable(m4), digits = 2)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	16.70	0.26	64.09	0.00
site2	6.50	0.26	25.34	0.00
site3	4.36	0.35	12.30	0.00
site4	1.93	0.36	5.43	0.00
site5	3.64	0.34	10.71	0.00
site6	4.20	0.42	9.97	0.00
site7	-0.18	0.67	-0.26	0.79
site8	-5.31	0.89	-5.95	0.00
site9	5.44	1.09	5.00	0.00
site10	2.26	1.37	1.65	0.10
dbh	0.62	0.01	81.47	0.00

Estimated tree heights for each site

```
summary(allEffects(m4))
```

```
model: height ~ site + dbh
```

```
site effect
```

```
site
```

1	2	3	4	5	6	7	8
33.90437	40.40868	38.26183	35.83902	37.54181	38.10889	33.72818	28.59173
9	10						
39.34142	36.16771						

```
Lower 95 Percent Confidence Limits
```

```
site
```

1	2	3	4	5	6	7	8
33.60276	40.00512	37.63569	35.20858	36.94739	37.33787	32.45495	26.86438
9	10						
37.22831	33.49623						

```
Upper 95 Percent Confidence Limits
```

```
site
```

1	2	3	4	5	6	7	8
34.20599	40.81223	38.88798	36.46947	38.13622	38.87990	35.00141	30.31907
9	10						
41.45454	38.83919						

```
dbh effect
```

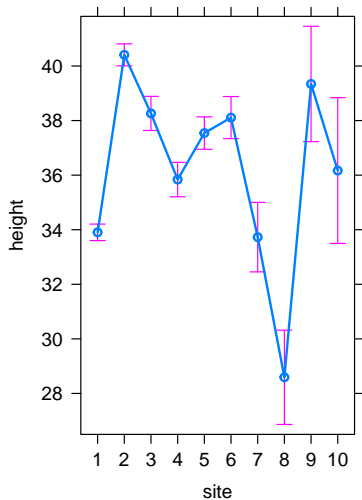
```
dbh
```

5	20	30	40	50
22.38634	31.64246	37.81321	43.98396	50.15471

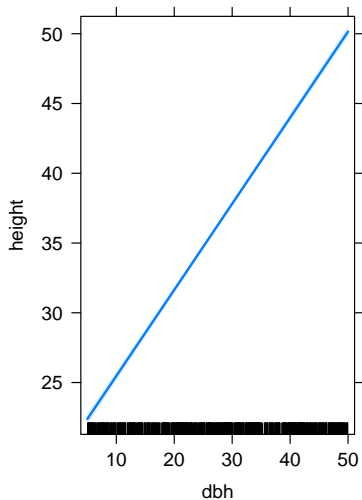
Plot

```
plot(allEffects(m4))
```

site effect plot

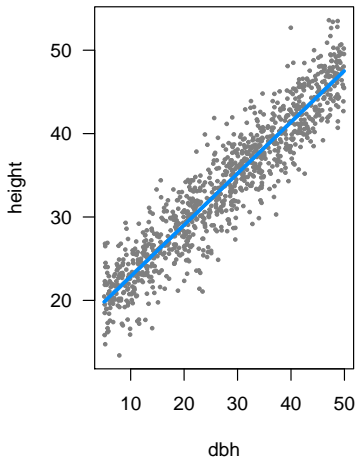
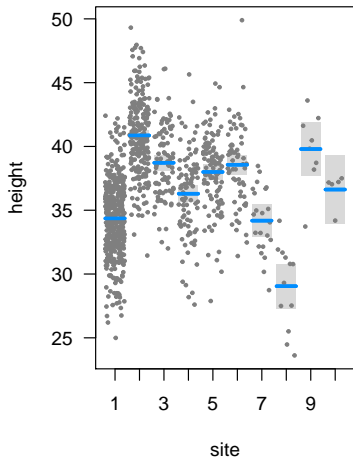


dbh effect plot



Plot (visreg)

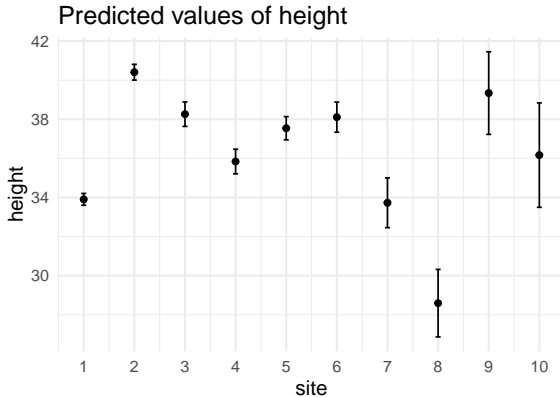
```
visreg(m4)
```



Plot model (sjPlot)

```
plot_model(m4, type = "eff")
```

\$site

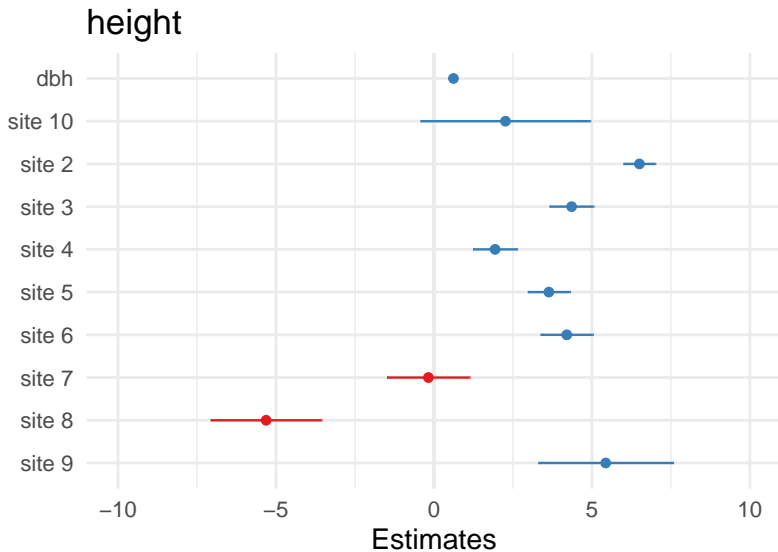


\$dbh

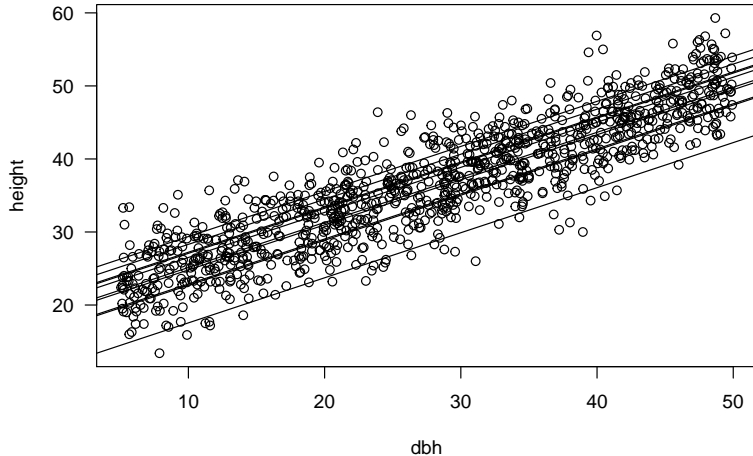
Predicted values of height

Plot model (sjPlot)

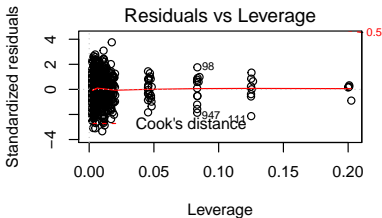
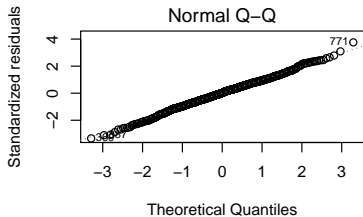
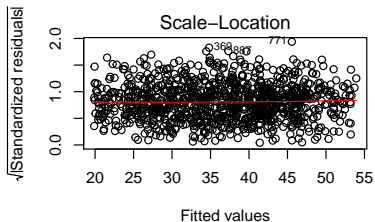
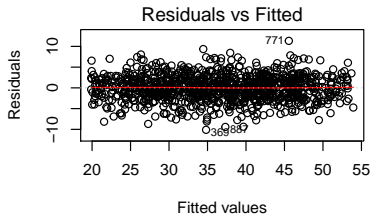
```
plot_model(m4, type = "est")
```



We have fitted model w/ many intercepts and single slope

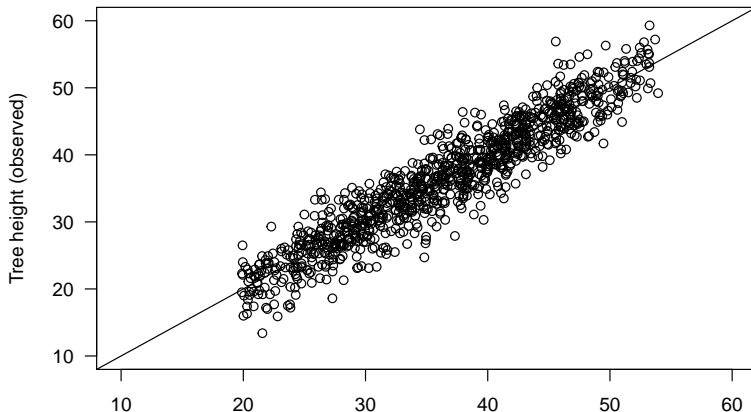


Model checking: residuals



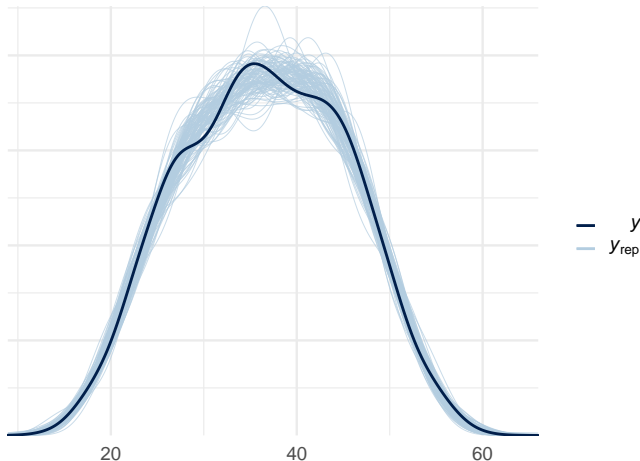
How good is this model? Calibration plot

```
trees$height.pred <- fitted(m4)
plot(trees$height.pred, trees$height, xlab = "Tree height (predi
abline(a = 0, b = 1)
```



Model checking with simulated data

```
library(bayesplot)
sims <- simulate(m4, nsim = 100)
ppc_dens_overlay(trees$height, yrep = t(as.matrix(sims)))
```



Q: Does allometric relationship between DBH
and Height vary among sites?

Model with interactions

Call:

```
lm(formula = height ~ site * dbh, data = trees)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-10.1017	-1.9839	0.0645	2.0486	11.1789

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	16.359437	0.360054	45.436	< 2e-16 ***
site2	7.684781	0.609657	12.605	< 2e-16 ***
site3	4.518568	0.867008	5.212	2.28e-07 ***
site4	2.769336	0.813259	3.405	0.000688 ***
site5	3.917607	0.870983	4.498	7.68e-06 ***
site6	4.155161	1.009379	4.117	4.17e-05 ***
site7	-2.306799	1.551303	-1.487	0.137334
site8	-2.616095	4.090671	-0.640	0.522630
site9	2.621560	5.073794	0.517	0.605492
site10	4.662340	2.991072	1.559	0.119378
dbh	0.629299	0.011722	53.685	< 2e-16 ***
site2:dbh	-0.042784	0.020033	-2.136	0.032950 *
site3:dbh	-0.006031	0.027640	-0.218	0.827312
site4:dbh	-0.031633	0.028225	-1.121	0.262677
site5:dbh	-0.010173	0.027887	-0.365	0.715334
site6:dbh	0.001337	0.032109	0.042	0.966797
site7:dbh	0.079728	0.052056	1.532	0.125951
site8:dbh	-0.079027	0.113386	-0.697	0.485984
site9:dbh	0.081035	0.146649	0.553	0.580679
site10:dbh	-0.101107	0.114520	-0.883	0.377522

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

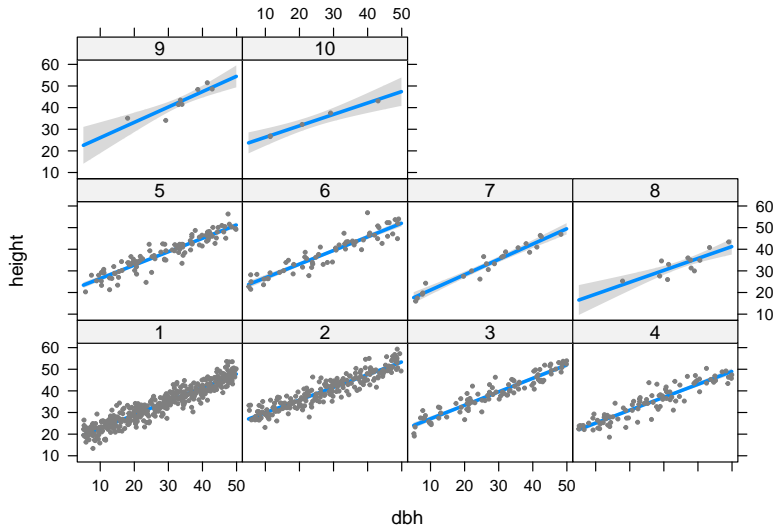
Residual standard error: 3.041 on 980 degrees of freedom

Multiple R-squared: 0.8847, Adjusted R-squared: 0.8825

F-statistic: 395.7 on 19 and 980 DF, p-value: < 2.2e-16

Does slope vary among forests?

```
visreg(m5, xvar = "dbh", by = "site")
```



Extra exercises

- ▶ paperplanes: How does flight distance differ with age, gender or paper type?

Extra exercises

- ▶ paperplanes: How does flight distance differ with age, gender or paper type?
- ▶ mammal sleep: Are sleep patterns related to diet?

Extra exercises

- ▶ paperplanes: How does flight distance differ with age, gender or paper type?
- ▶ mammal sleep: Are sleep patterns related to diet?
- ▶ iris: Predict petal length \sim petal width and species

Extra exercises

- ▶ paperplanes: How does flight distance differ with age, gender or paper type?
- ▶ mammal sleep: Are sleep patterns related to diet?
- ▶ iris: Predict petal length \sim petal width and species
- ▶ racing pigeons: is speed related to sex?