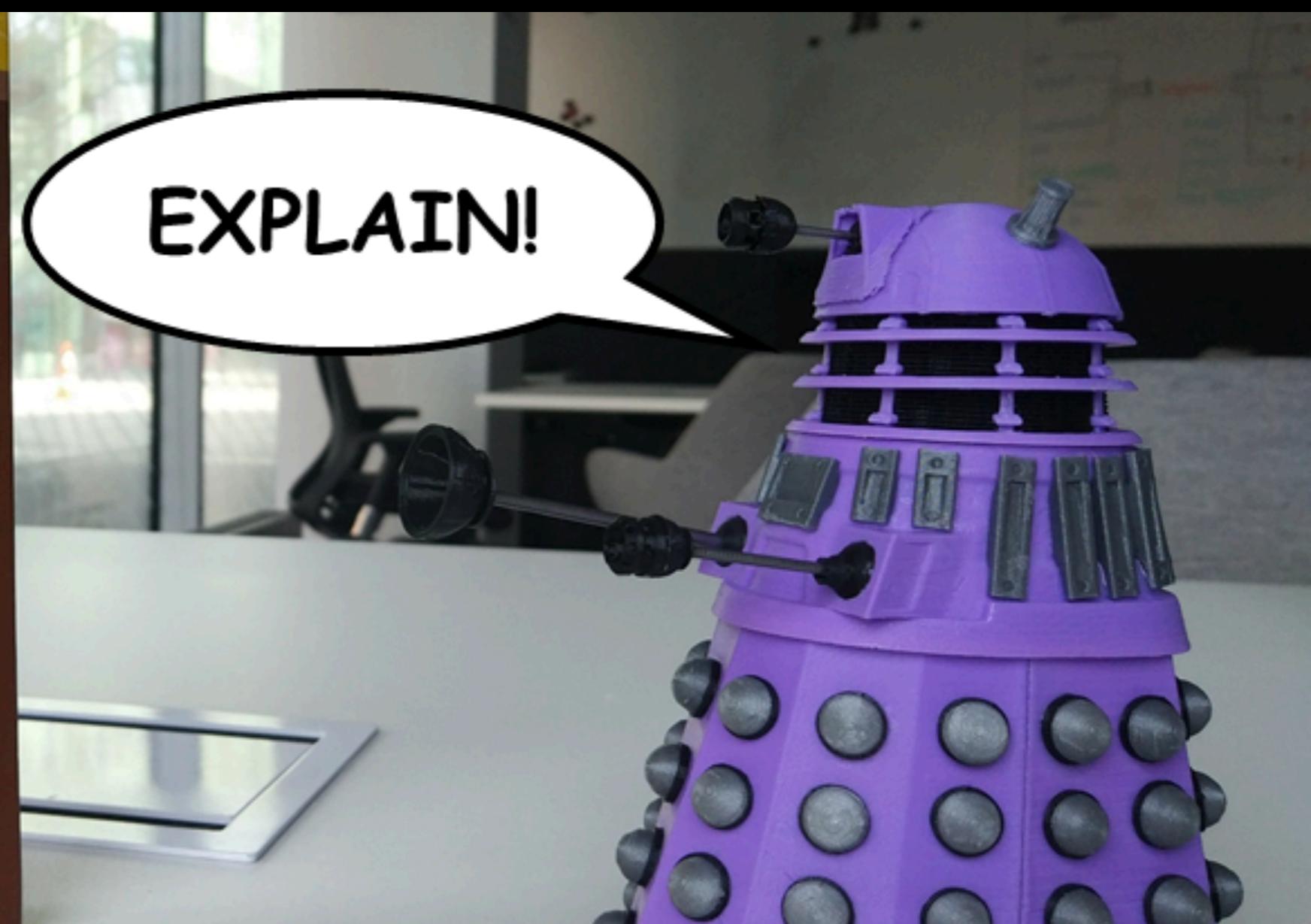
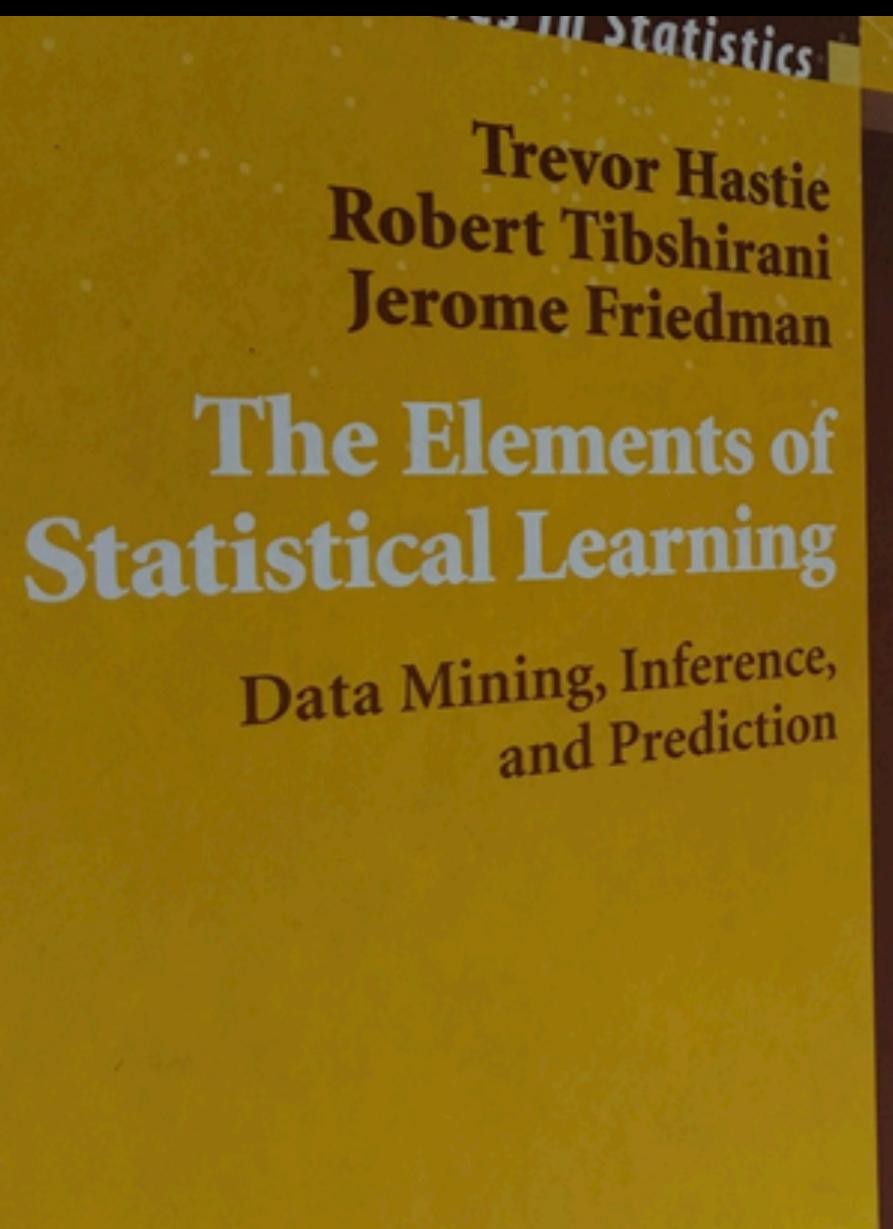


Show me your model 2.0!



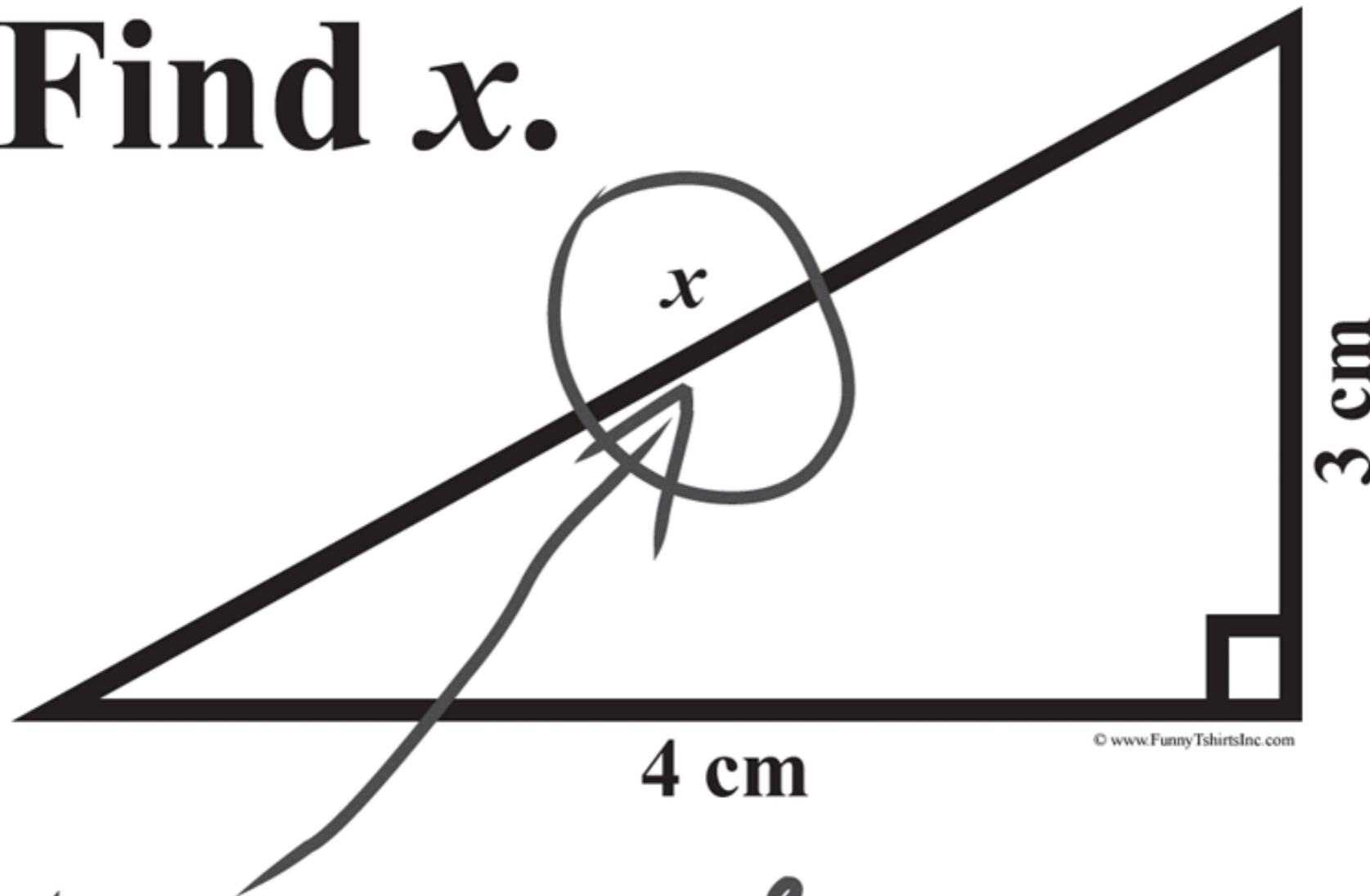
Przemysław Biecek

eRum 2018

How to visualise complex predictive models?

How to visualise complex predictive models?

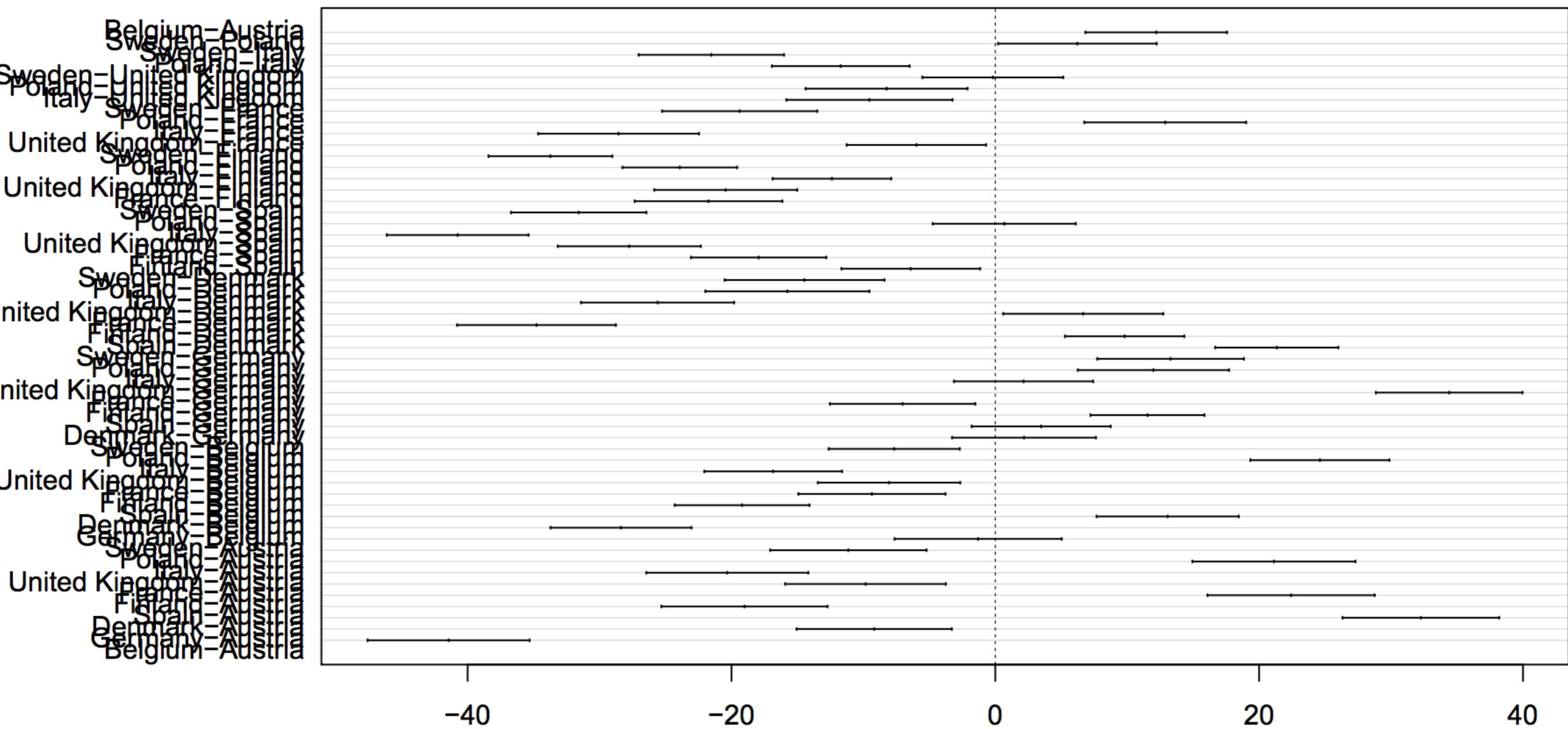
Find x .



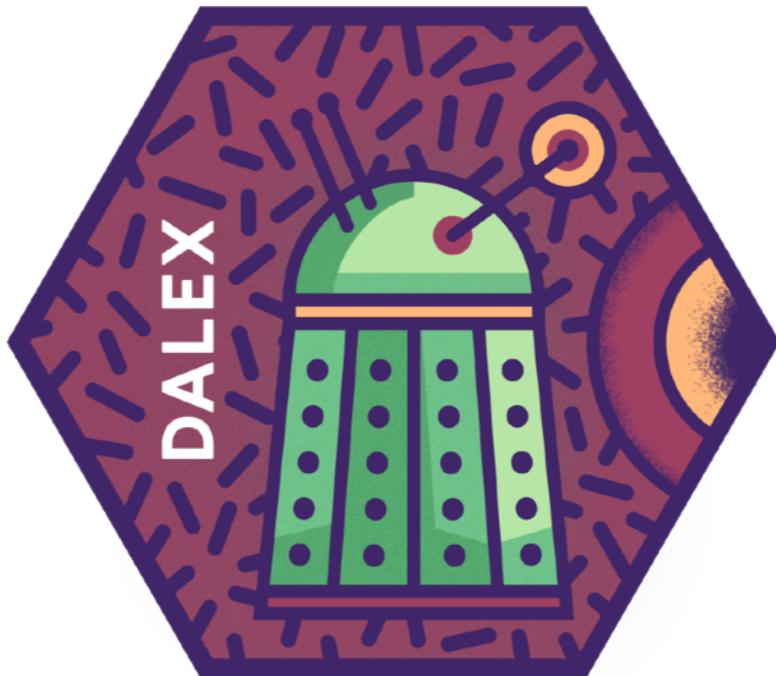
© www.FunnyTshirtsInc.com

Here it is

How to visualise complex predictive models?



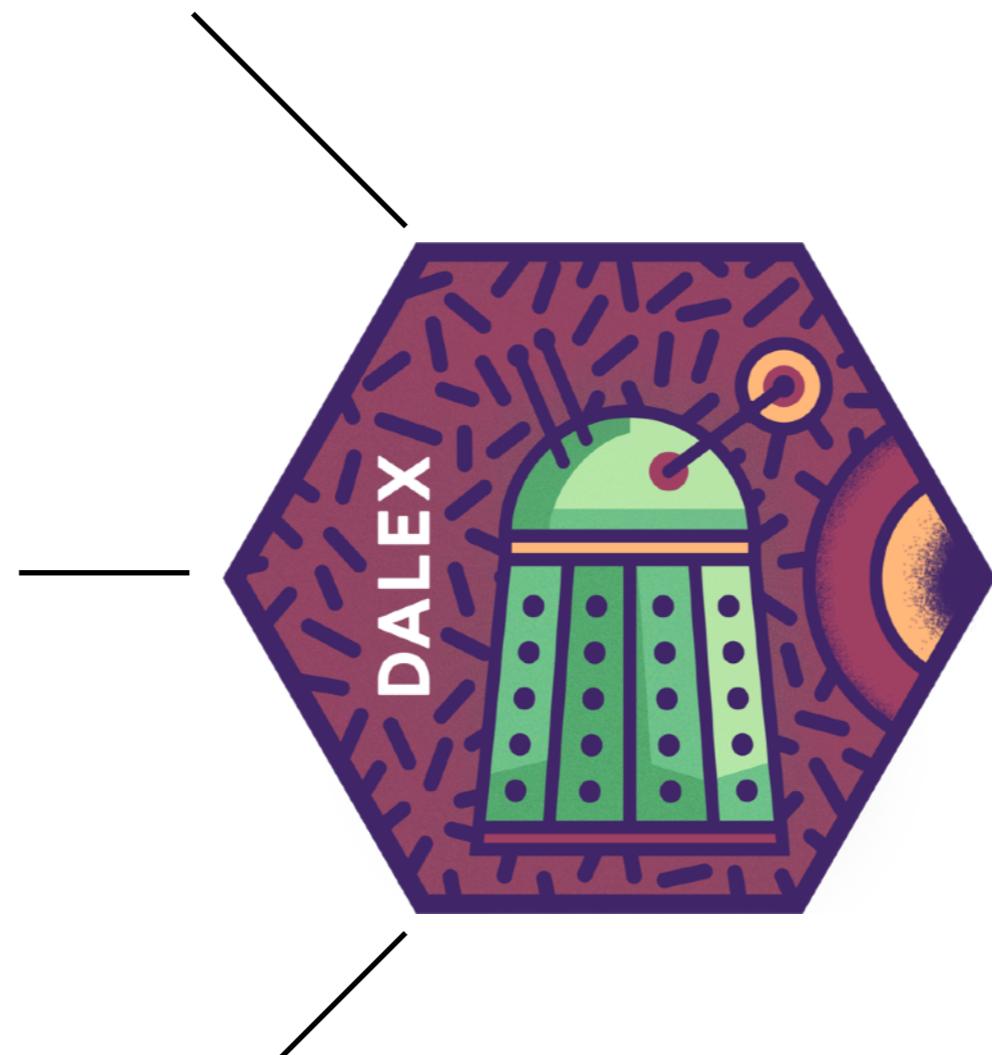
DALEX is the tool for model understanding



Model understanding

How good is the model?

Which variables
are important?



How a variable is
linked with model response?

Model understanding

How good is the model?

Which variables
are important?

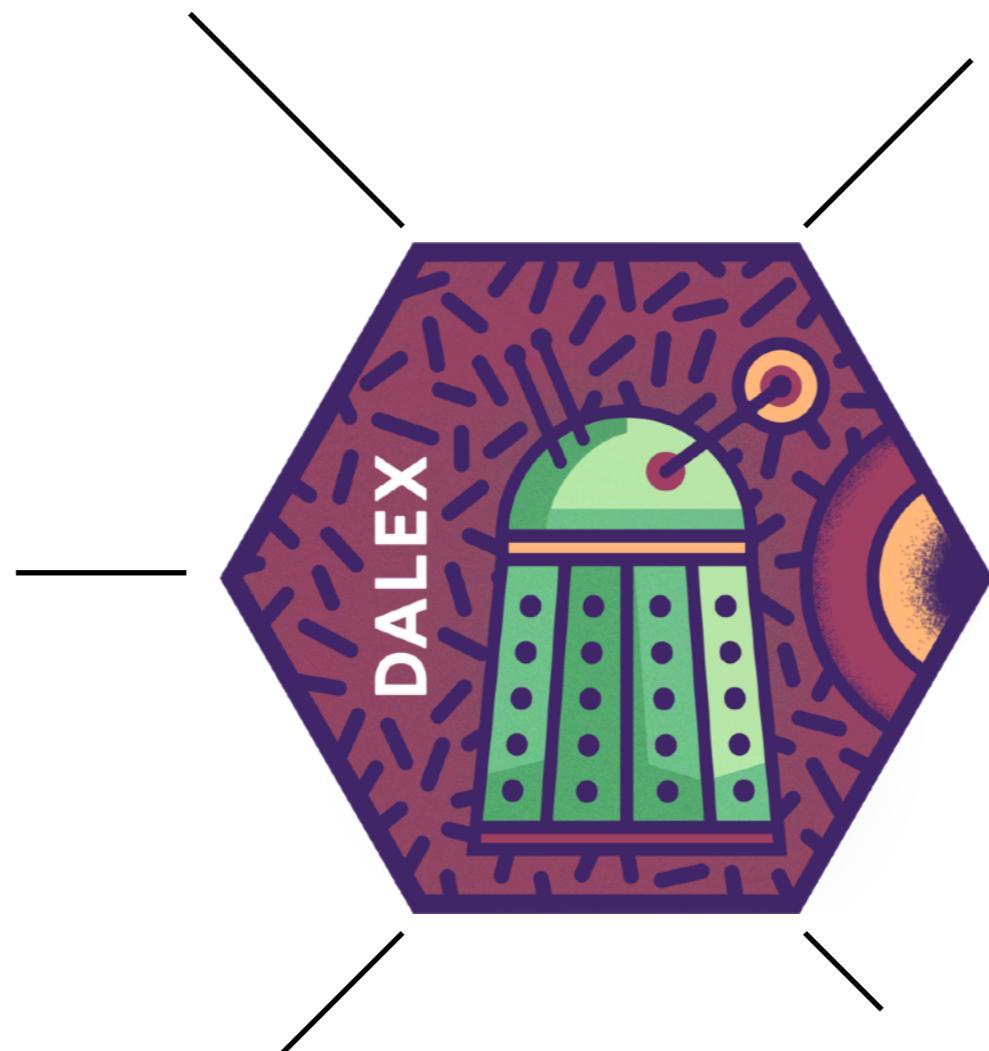
How a variable is
linked with model response?

Prediction understanding

How good is a specific
model prediction?

Which observations
have large residuals?

Which variables influence
a specific model prediction?





```
library("DALEX")
head(apartments)
```

m2.price	construction.year	surface	floor	no.rooms	district
5897	1953	25	3		1 Śródmieście
1818	1992	143	9		5 Bielany
3643	1937	56	1		2 Praga
3517	1995	93	7		3 Ochota
3013	1992 ⁸	144	6		5 Mokotów

A Tale of Two Models

```
apartments_lm_model <- lm(m2.price ~ construction.year + surface + floor  
                           no.rooms + district, data = apartments)  
  
library("randomForest")  
set.seed(59)  
apartments_rf_model <- randomForest(m2.price ~ construction.year + surface  
                           no.rooms + district, data = apartments)
```

A Tale of Two Models

Accuracy as a single number is not enough!

```
# root mean square
predicted_mi2_lm <- predict(apartments_lm_model, apartmentsTest)
sqrt(mean((predicted_mi2_lm - apartmentsTest$m2.price)^2))
## [1] 283.0865

# root mean square
predicted_mi2_rf <- predict(apartments_rf_model, apartmentsTest)
sqrt(mean((predicted_mi2_rf - apartmentsTest$m2.price)^2))
## [1] 283.3479
```

A Tale of Two Models

```
library("DALEX")
apartments_lm_model <- lm(m2.price ~ construction.year + surface + floor
                           + no.rooms + district, data = apartments)

library("randomForest")
set.seed(59)
apartments_rf_model <- randomForest(m2.price ~ construction.year + surface
                                       + no.rooms + district, data = apartments)

explainer_lm <- explain(apartments_lm_model,
                         data = apartmentsTest[,2:6], y = apartmentsTest$price)

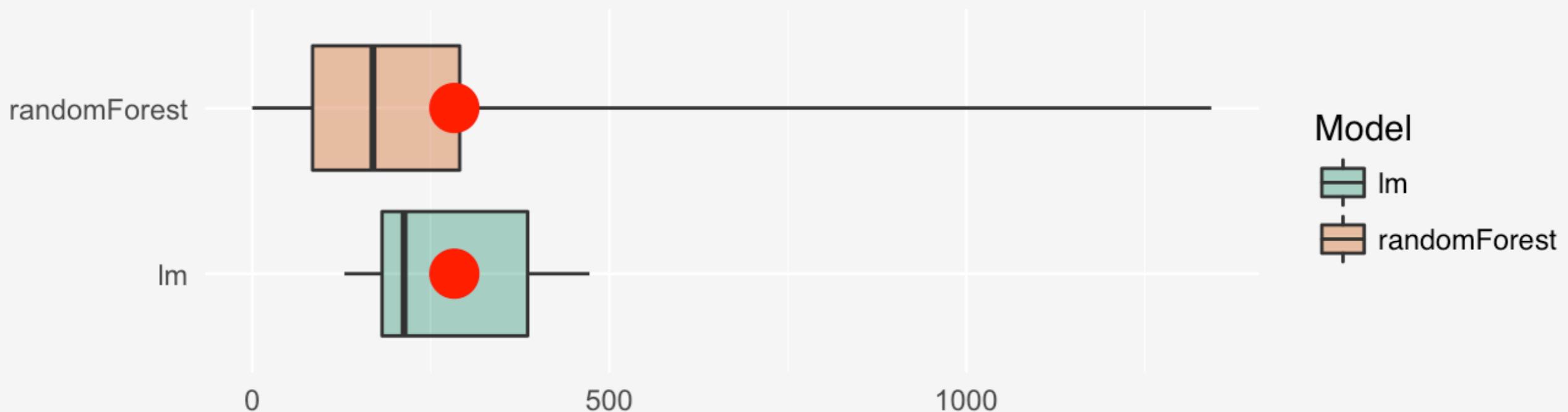
explainer_rf <- explain(apartments_rf_model,
                         data = apartmentsTest[,2:6], y = apartmentsTest$price)
```

How good are these models?

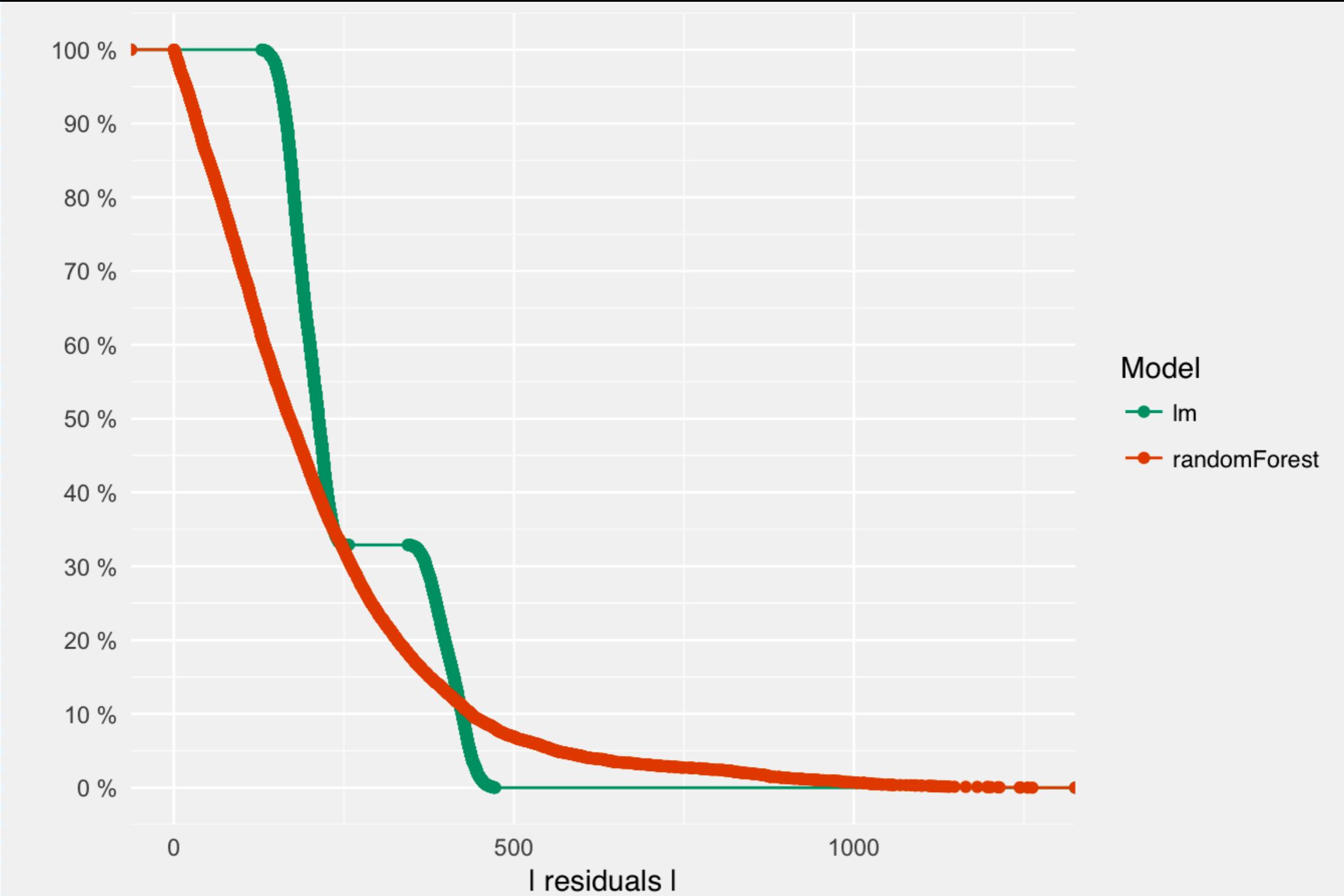
```
plot(mp_lm, mp_rf, geom = "boxplot")
```

Boxplots of I residuals I

Red dot stands for root mean square of residuals



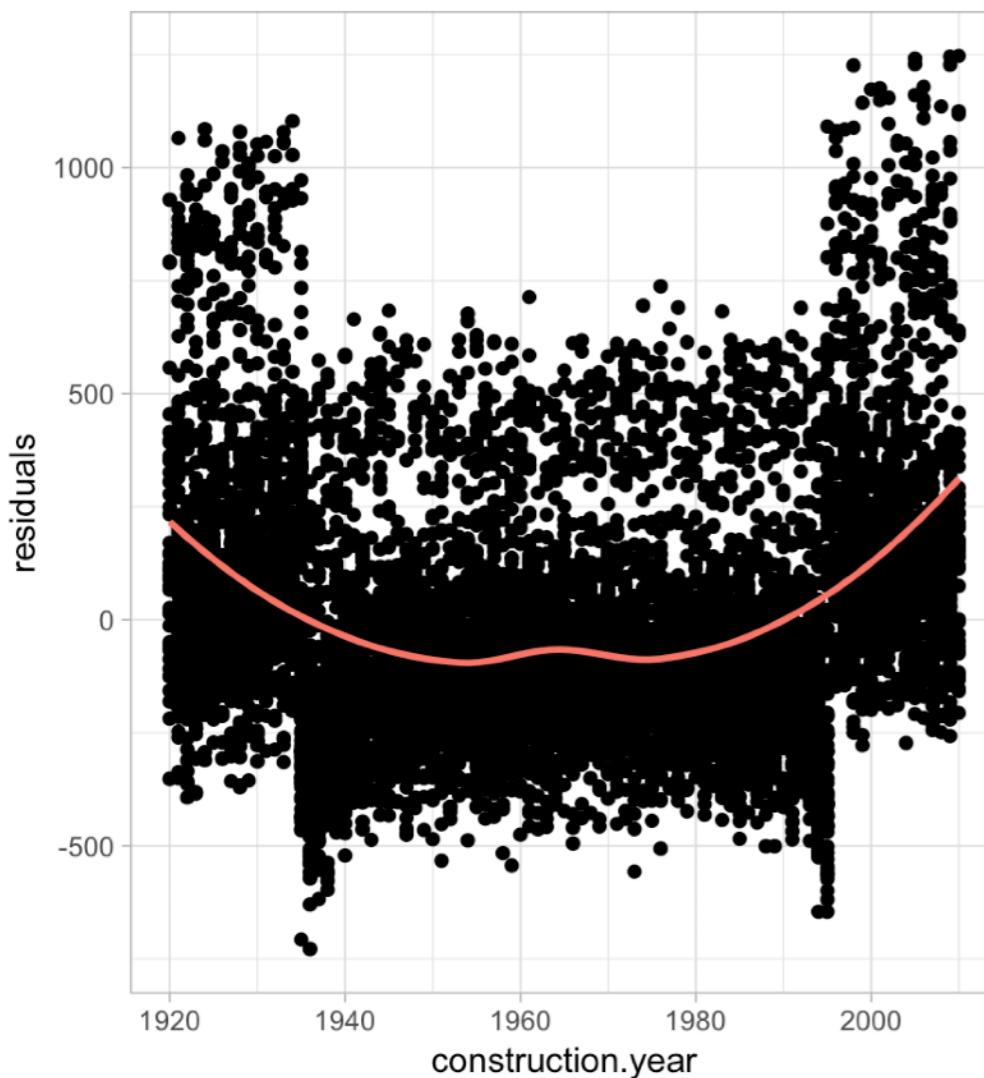
How good are these models?



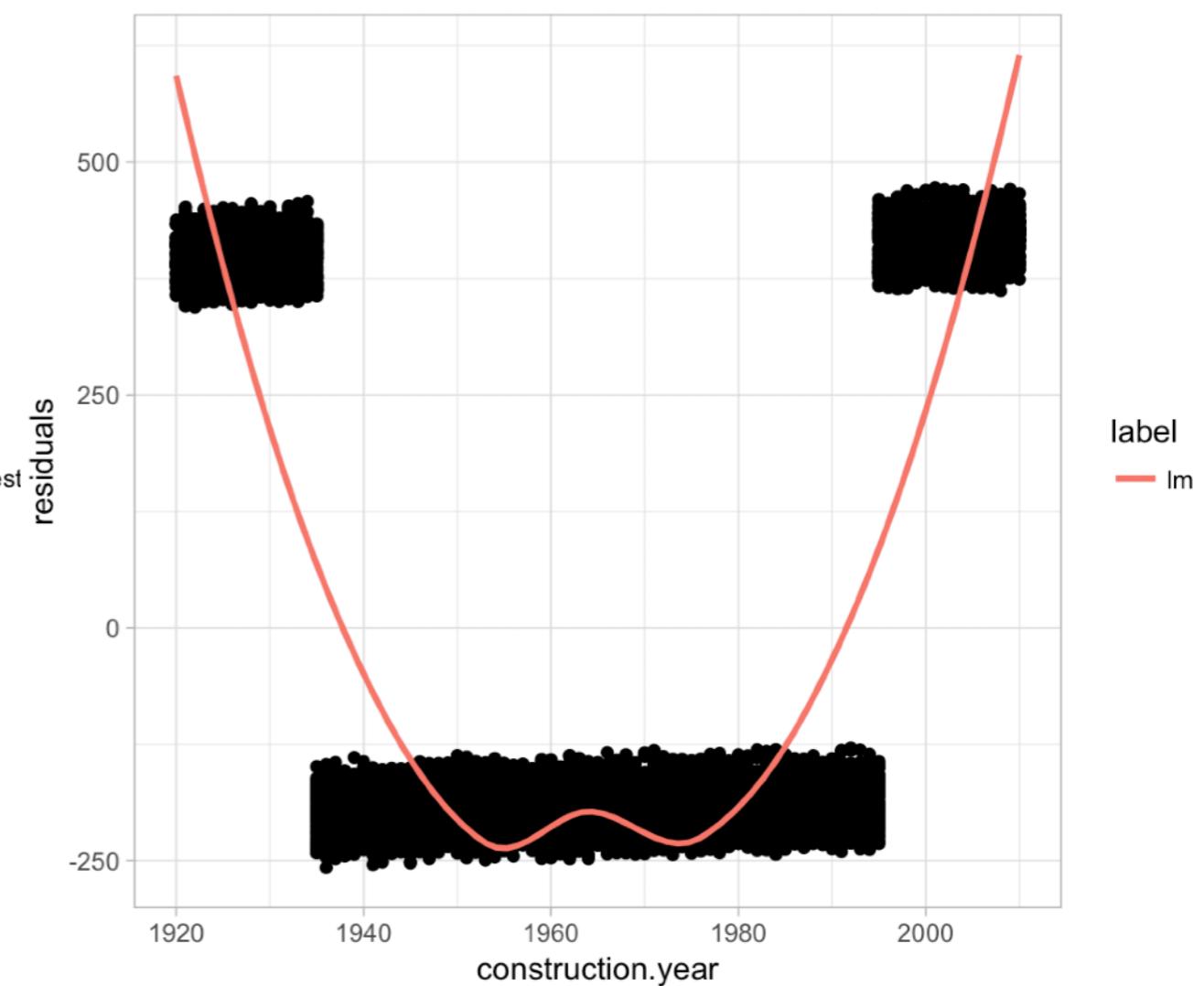
```
library(auditor)
audit_rf <- audit(explainer_rf)
plotResidual(audit_rf, variable = "construction.year")
```

```
audit_lm <- audit(explainer_lm)
plotResidual(audit_lm, variable = "construction.year")
```

Residuals vs construction.year



Residuals vs construction.year

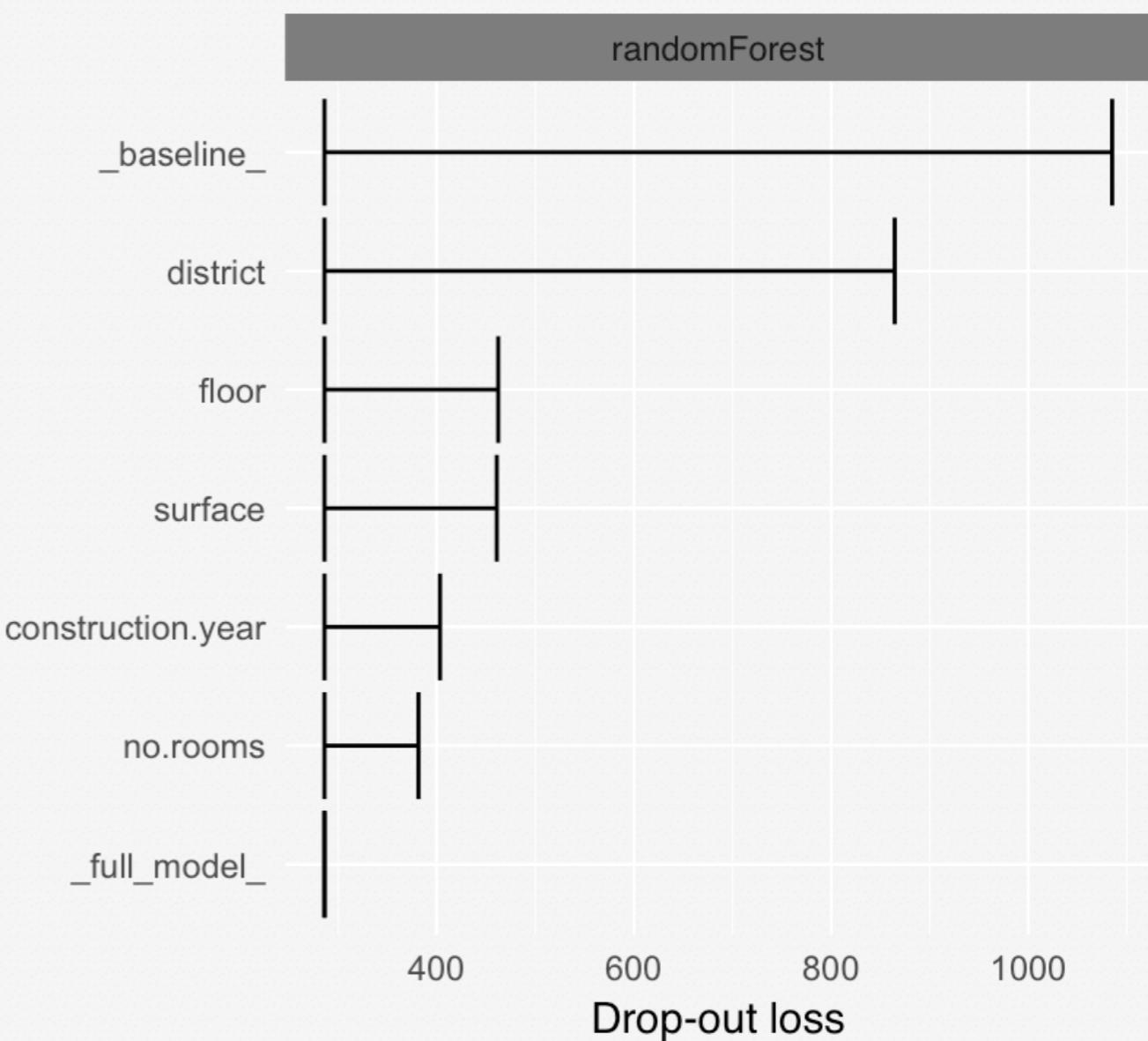


Which variables are important?

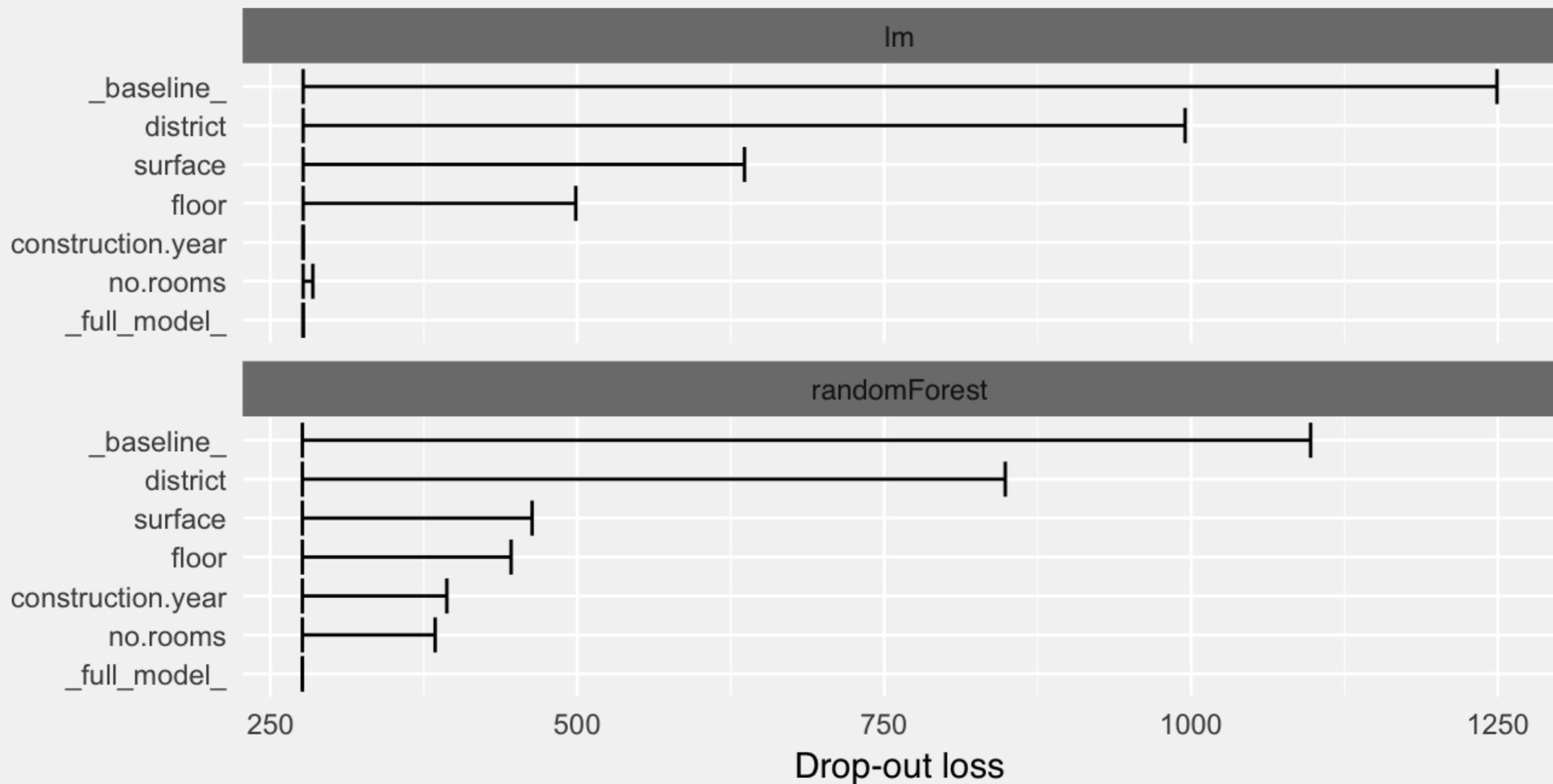
```
> vi_rf <- variable_importance(explainer_rf, loss_function = loss_root_mean_square)
> vi_rf
```

	variable	dropout_loss	label
1	_full_model_	286.2676	randomForest
2	no.rooms	381.5975	randomForest
3	construction.year	403.4376	randomForest
4	surface	461.0018	randomForest
5	floor	462.3999	randomForest
6	district	864.4315	randomForest
7	_baseline_	1084.9218	randomForest

```
>
> plot(vi_rf)
```

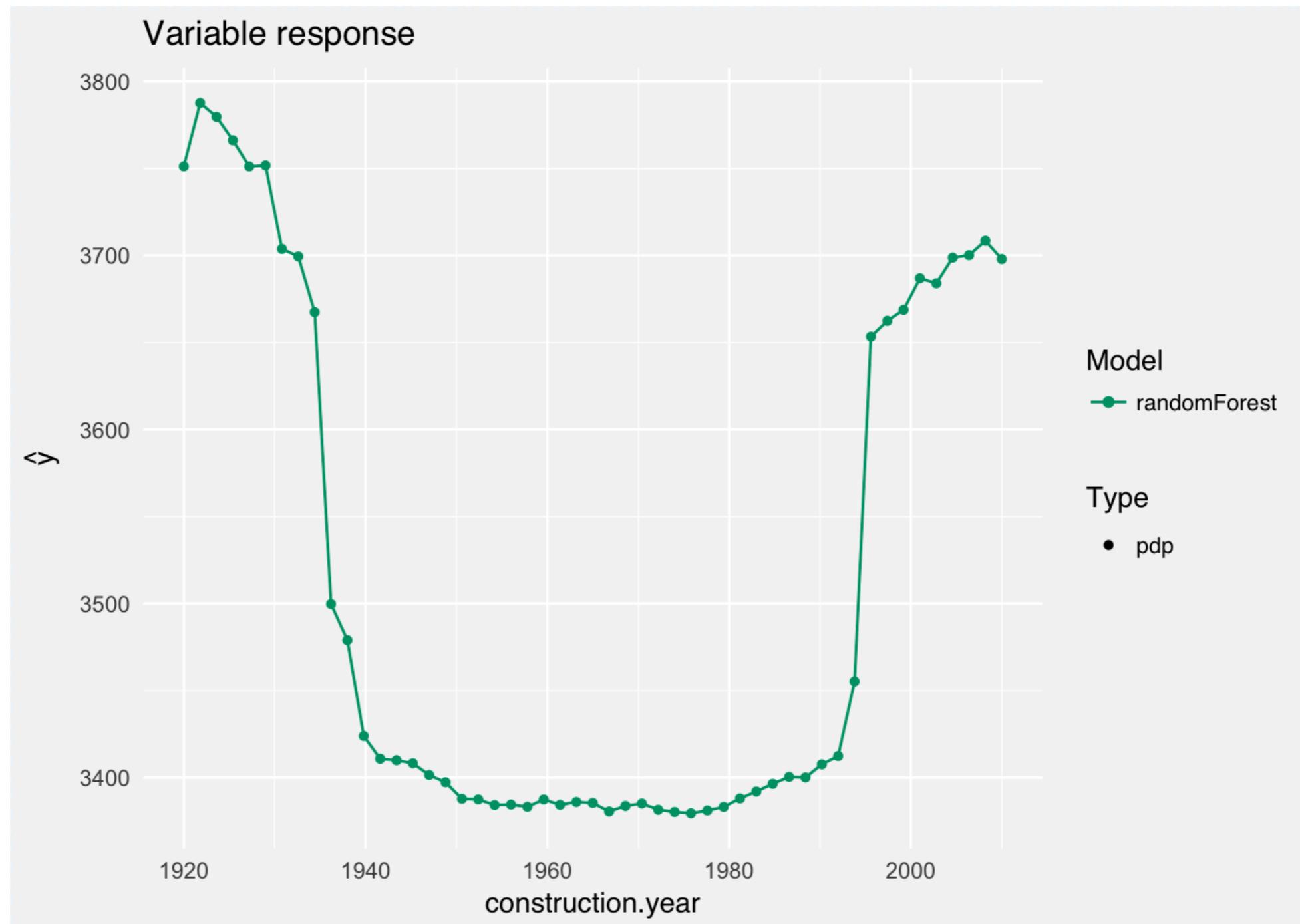


Which variables are important?



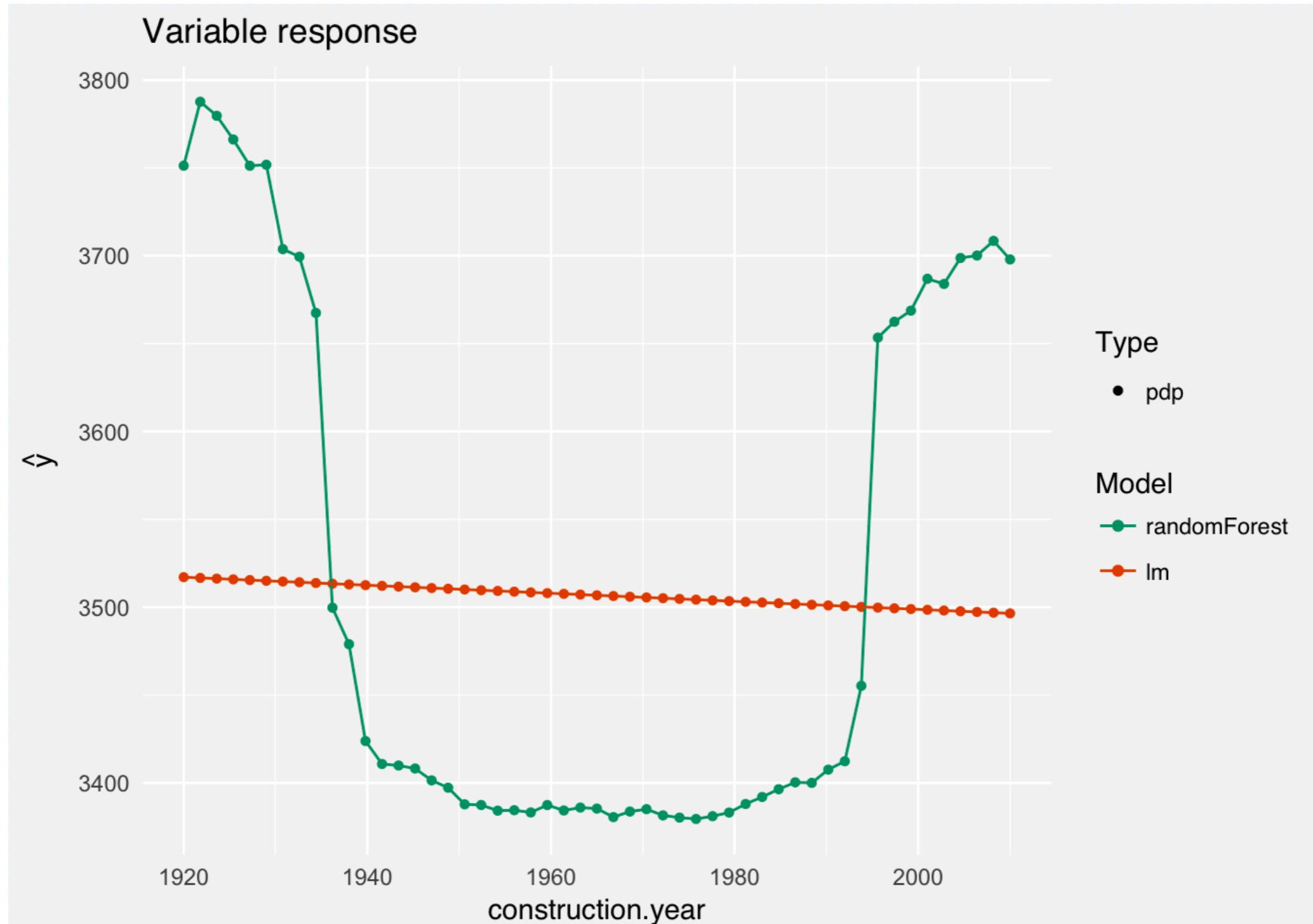
What is the conditional response?

```
sv_rf <- single_variable(explainer_rf, variable = "construction.year", type = "pdp")
plot(sv_rf)
```

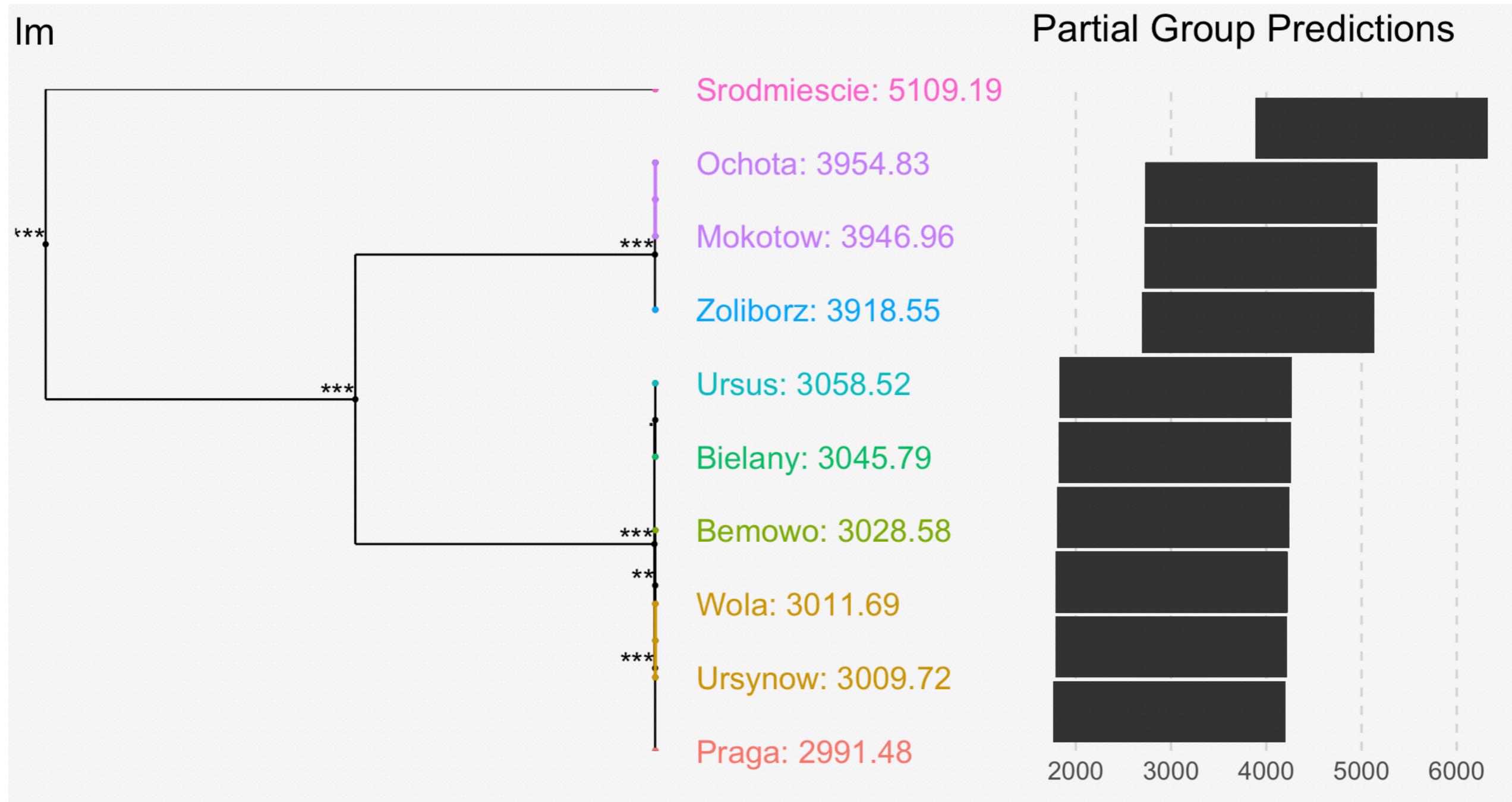


What is the conditional response?

```
plot(sv_rf, sv_lm)
```



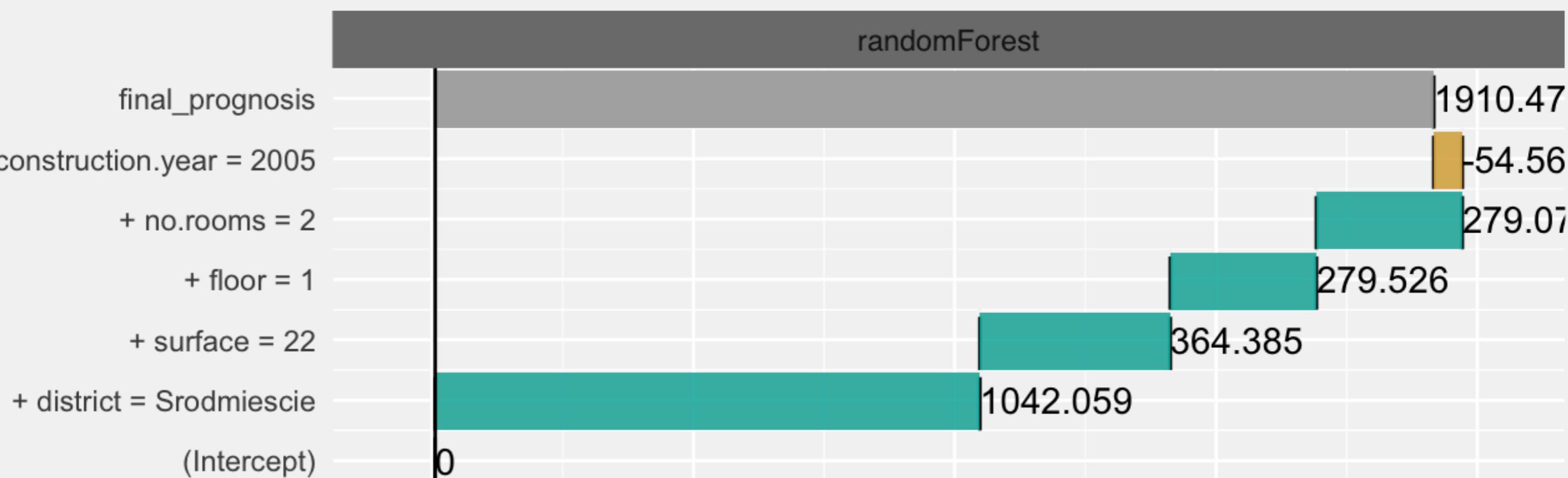
What is the conditional response?



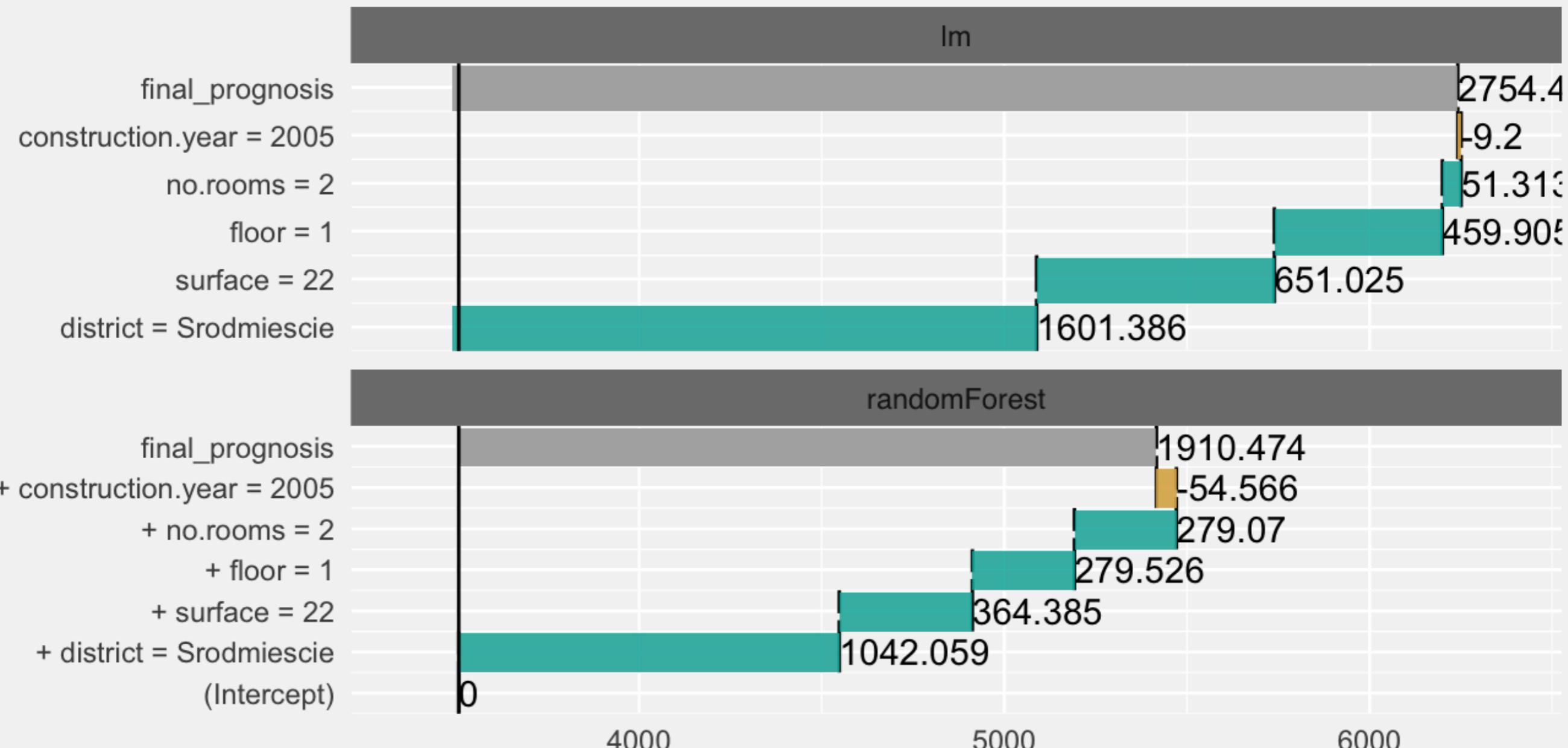
breakDown for predictions

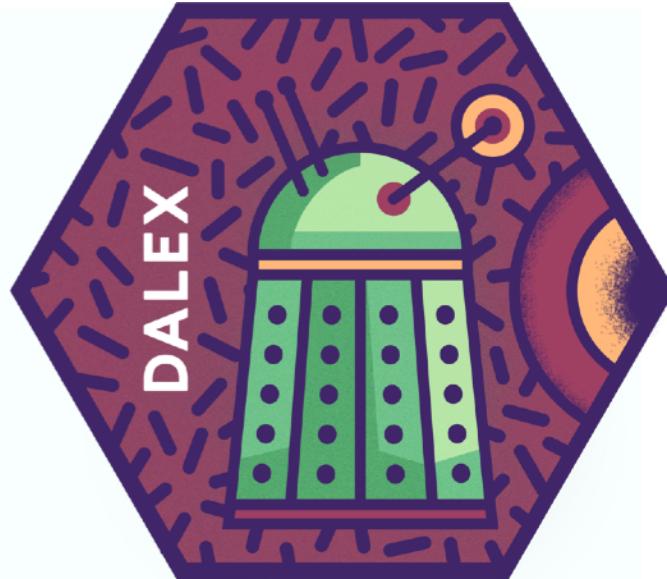
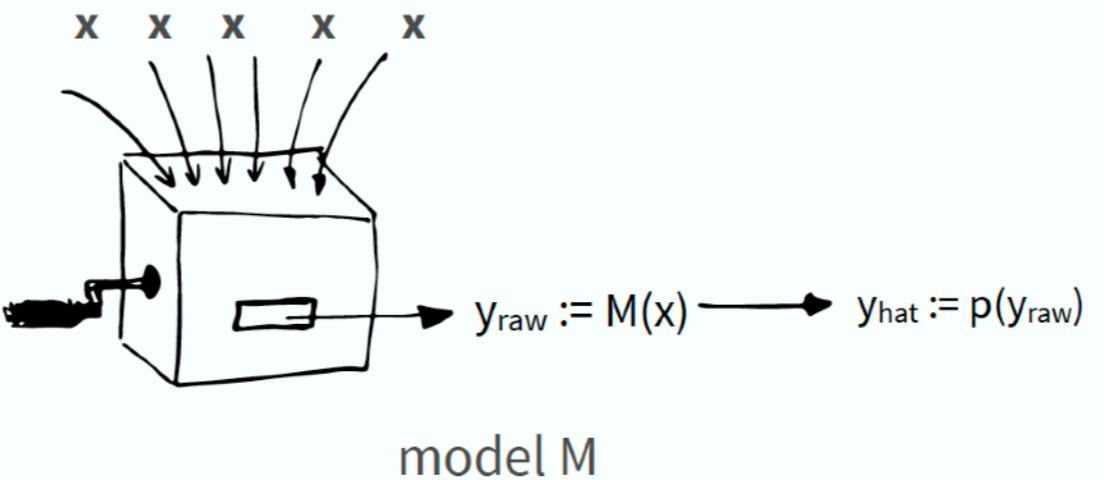
```
new_apartment_rf <- single_prediction(explainer_rf, observation = new_apartment)
breakDown:::print.broken(new_apartment_rf)
```

```
##                                     contribution
## (Intercept)                      0.000
## + district = Srodmiescie      1042.059
## + surface = 22                  364.385
## + floor = 1                     279.526
## + no.rooms = 2                  279.070
## + construction.year = 2005     -54.566
## final_prognosis                 1910.474
```



Asking for opinions

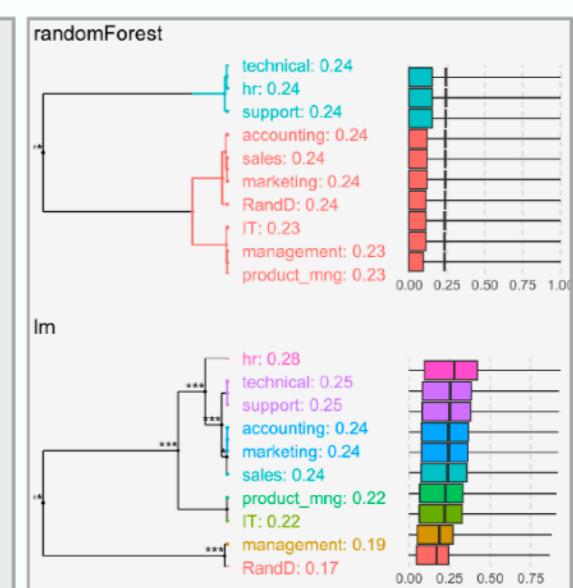
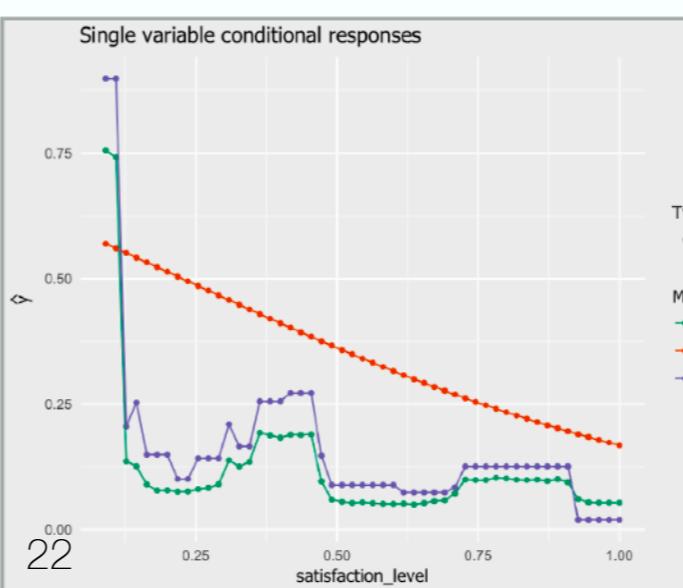
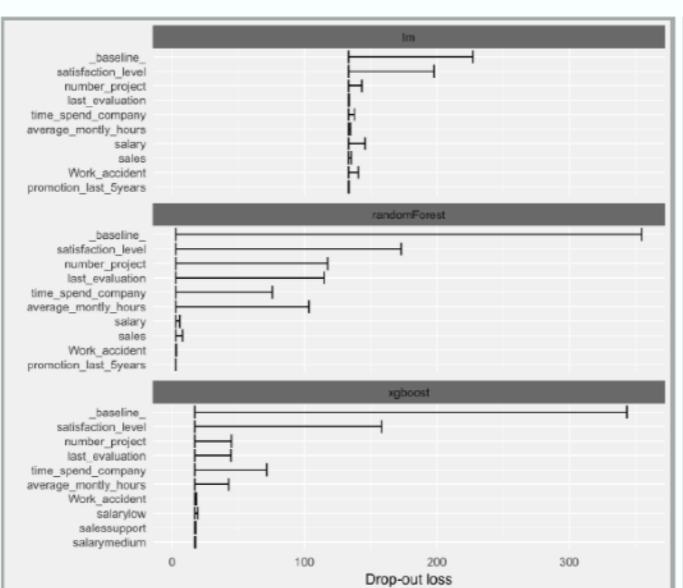
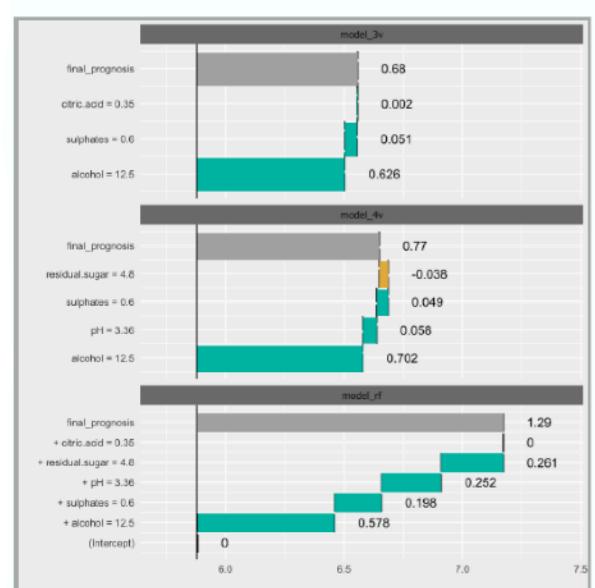


A)**B)**

`explain(model; data; y; predict_function; trans)`

C)

`single_prediction(explainer, x)` `variable_dropout(explainer)` `single_variable(explainer, variable)`



Language for Human-Model Interaction



Language for Human-Model Interaction

