

Applied Statistical Methods - Solution 6

AUTHOR
Peter von Rohr

PUBLISHED
April 15, 2024

WEBR STATUS
● Ready!

Problem 1: Regression On Dummy Variables

Use the dataset with the breeds assigned to every animal and find out the influence of the breed on the response variable `body weight`. The data is available from

https://charlotte-ngs.github.io/asmasss2024/data/asm_bw_flem.csv

Start by fitting a linear model with `Breed` as the only factor in the model, hence ignore the independent variables such as `Breast Circumference`, `BCS` and `HEI`.

Tasks

- Read the data

▶ Run Code

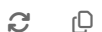


```
1 s_ex06p01_data_path <- "https://charlotte-ngs.github.io/asmasss2024/"
2 tbl_ex06p01_data <- read.table(s_ex06p01_data_path,
3                               header = T, sep = ",")
4 tbl_ex06p01_data
```

	Animal	Breast.Circumference	Body.Weight	BCS	HEI	Breed
1	1	176	471	5.0	161	Angus
2	2	177	463	4.2	121	Angus
3	3	178	481	4.9	157	Simmental
4	4	179	470	3.0	165	Angus
5	5	179	496	6.8	136	Simmental
6	6	180	491	4.9	123	Simmental
7	7	181	518	4.4	163	Limousin
8	8	182	511	4.4	149	Limousin
9	9	183	510	3.5	143	Limousin
10	10	184	541	4.7	130	Limousin

- Fit a linear model including breed as a factor

▶ Run Code



```
1 lm_reg_dummy_bw_breed <- lm(Body.Weight ~ Breed,
2                             data = tbl_ex06p01_data)
3 summary(lm_reg_dummy_bw_breed)
```

Call:

```
lm(formula = Body.Weight ~ Breed, data = tbl_ex06p01_data)
```

Residuals:

Min	1Q	Median	3Q	Max
-10.0000	-7.5000	-0.1667	2.7500	21.0000

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	468.000	6.097	76.758	1.68e-11 ***
BreedLimousin	52.000	8.066	6.447	0.000351 ***

BreedSimmental 21.333 8.623 2.474 0.042575 *

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

Residual standard error: 10.56 on 7 degrees of freedom

Multiple R-squared: 0.8597, Adjusted R-squared: 0.8196

F-statistic: 21.44 on 2 and 7 DF, p-value: 0.001035