25/02/2019 itslearning



Home

Courses

Groups

Calendar **BAU**



BDA500...

Overview

Plans

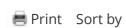
Resources

Status and follow-up



European Sales Data (with Derived Features)

- Cem Acun



Date



Cem Acun 1802131 Hello all,

TL;DR: I thought it would be interesting to derive some extra features from given ones (derived *GDP* and *Education Spending per Head*), and they perform pretty well by themselves with log-log transformation.

Results: Best ones are modelCapita4 and modelComputer4

- log(SalesPerCapita) ~ log(EdSpendHead) -> 0.7757 Adj. R²
- log(ComputerSales) ~ log(GDP)

-> 0.8807 Adj. R²

<CodeWithFigures>

#Get the dataset

europeData<-read.csv("EuropeanSales.csv",header=T)

#Take a look at it

attributes(europeData)

summary(europeData)

