Data Assignment 3

### Create Variables

data\_tw <- data\_tw %>% mutate(DV1 = ifelse(year >= 1959,1,0), DV2 = ifelse(year >= 2000,1,0))  
  
data\_tw <- data\_tw %>%  
 mutate(SplineBreak1 = (Trend - 58)\*DV1)  
  
data\_tw <- data\_tw %>%  
 mutate(SplineBreak2 = (Trend - 99)\*DV2)

### Implement First Regression

model1 <- lm(log(gdppc)~Trend+DV1+Trend\*DV1+DV2+Trend\*DV2,data=data\_tw)  
  
pander(summary(model1))

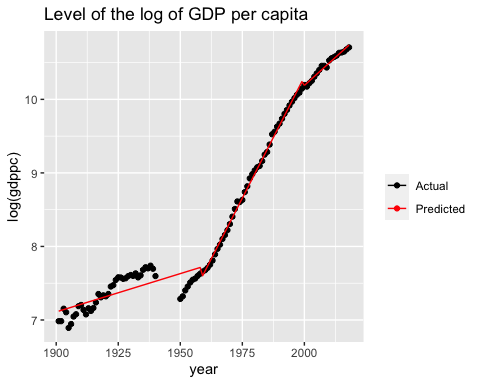
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| **(Intercept)** | 7.12 | 0.03025 | 235.4 | 1.662e-142 |
| **Trend** | 0.01042 | 0.0009885 | 10.54 | 4.729e-18 |
| **DV1** | -3.346 | 0.1239 | -27 | 1.584e-48 |
| **DV2** | 3.397 | 0.5365 | 6.333 | 6.39e-09 |
| **Trend:DV1** | 0.05558 | 0.001816 | 30.6 | 1.627e-53 |
| **Trend:DV2** | -0.0355 | 0.005069 | -7.004 | 2.632e-10 |

Fitting linear model: log(gdppc) ~ Trend + DV1 + Trend \* DV1 + DV2 + Trend \* DV2

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | Residual Std. Error |  | Adjusted |
| 109 | 0.1154 | 0.9919 | 0.9915 |

#### Graph

predicted\_df1 <- data.frame(gdppc\_prd = predict(model1, data\_tw), year=data\_tw$year)  
  
ggplot(data\_tw, aes(x=year,y=log(gdppc))) +  
 geom\_point(aes(colour= "Actual")) +  
 geom\_line(aes(x=year, y=gdppc\_prd, colour = "Predicted"),data = predicted\_df1) +  
 scale\_colour\_manual("",  
 values = c("Actual"="black", "Predicted"="red")) +  
 labs(title = "Level of the log of GDP per capita")



#ggplot(data\_tw, aes(x=year,y=log(gdppc))) +  
# geom\_point() +  
# geom\_line(color='red',data = predicted\_df1, aes(x=year, y=gdppc\_prd))

## Spline Method

spline\_model <- lm(log(gdppc)~Trend+SplineBreak1+SplineBreak2,data=data\_tw)  
  
pander(summary(spline\_model))

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| **(Intercept)** | 7.143 | 0.02899 | 246.4 | 7.012e-147 |
| **Trend** | 0.009077 | 0.0007743 | 11.72 | 8.607e-21 |
| **SplineBreak1** | 0.05384 | 0.001681 | 32.02 | 5.362e-56 |
| **SplineBreak2** | -0.03722 | 0.004222 | -8.815 | 2.762e-14 |

Fitting linear model: log(gdppc) ~ Trend + SplineBreak1 + SplineBreak2

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | Residual Std. Error |  | Adjusted |
| 109 | 0.1186 | 0.9913 | 0.991 |

### Graph

predicted\_df2 <- data.frame(gdppc\_prd2 = predict(spline\_model, data\_tw), year=data\_tw$year)  
  
ggplot(data\_tw, aes(x=year,y=log(gdppc))) +  
 geom\_point(aes(colour= "Actual")) +  
 geom\_line(aes(x=year, y=gdppc\_prd2, colour = "Predicted"),data = predicted\_df2) +  
 scale\_colour\_manual("",  
 values = c("Actual"="black", "Predicted"="red")) +  
 labs(title = "Level of the log of GDP per capita (Spline)")

