

Untitled

L. Austin Hadamuscin

3/10/2022

Packages

```
library(dplyr)
```

Data

```
dat <- read.csv("Life Expectancy Data.csv") |>
  filter(Year==2014) |>
  select(-Year)

dat <- mutate_if(dat,
  is.numeric,
  function(x) ifelse(is.na(x), median(x, na.rm = T), x))
```

Partitioning The Data

```
train_ind <- sample(1:nrow(dat), size = 0.70*nrow(dat))

train_dat <- dat[train_ind,]
test_dat <- dat[-train_ind,]
```

Fitting a model on the training data

```
full_resid <- lm(Life.expectancy~., select(train_dat, -Country, -Status))$residuals
```

Plotting the Residuals (linearity check)











