# Question 1 f)

Sheen Thusoo

10/11/2021

```r
data <- read.csv("Infectious.csv")
```

```r
powerfun <- function(x, alpha) {
  if(sum(x <= 0) > 1) stop("x must be positive")
  if (alpha == 0)
    log(x)
  else if (alpha > 0) {
    x^alpha
  } else -x^alpha
}
```

```r
power_transformation <- function(x, y) {
  function(alpha_x, alpha_y) {
    power_x <- powerfun(x + 1, alpha_x)
    power_y <- powerfun(y + 1, alpha_y)
    new_data <- list("x" = power_x, "y" = power_y)
  }
}

power_transformed = power_transformation(data$Infected, data$Deceased.Prop)
data_transformed <- power_transformed(0.3, 3.2)
```

```r
library("ggplot2")
library("gridExtra")

ggplot(data) +
  geom_point(
    aes(x = data_transformed$x,
        y = data_transformed$y,
        alpha = 0.5)
  ) +
  labs(
    title = "Power Transformation of Number Infected vs. Deceased Proportion",
    x = "Power-Transformed Number Infected "~alpha[x] == 0.3~"",
    y = "Power-Transformed Deceased Proportion "~alpha[y] == 3.2~""
  )
```

## Power Transformation of Number Infected vs. Deceased Proportion
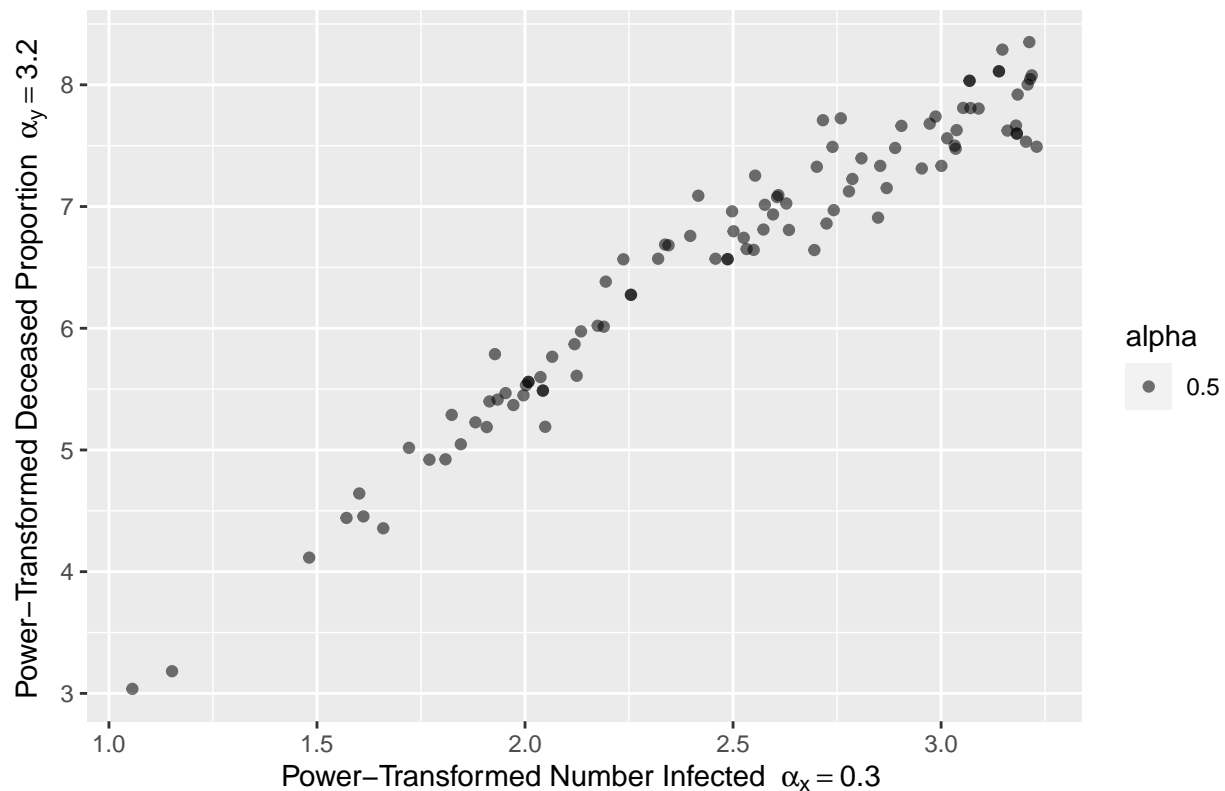


```
dataframe = data.frame(data_transformed$x, data_transformed$y)
fitted.model <- lm(data_transformed$y~data_transformed$x, dataframe)

summary(fitted.model)
```

```
##
## Call:
## lm(formula = data_transformed$y ~ data_transformed$x, data = dataframe)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.68140 -0.15446 -0.00016  0.16182  0.66243
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.23736    0.13470   9.186 7.07e-15 ***
## data_transformed$x  2.14742    0.05294  40.563  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.278 on 98 degrees of freedom
## Multiple R-squared:  0.9438, Adjusted R-squared:  0.9432
## F-statistic:  1645 on 1 and 98 DF,  p-value: < 2.2e-16
```

```
yhat <- fitted.model$coefficients[1] + fitted.model$coefficients[2] * data$Infected
yhat.realigned <- yhat ^ (1/3.2) - 1
```

```
ggplot(data) +
  geom_point(
    aes(x = yhat.realigned,
        y = Deceased.Prop,
        alpha = 0.5)
  ) +
  labs(
      y="Fitted Model",
      x="Deceased Proportion",
      title="Plot of Observed Proportions - Deceased Proportion vs. Fitted Model",
)
```



Plot of Observed Proportions – Deceased Proportion vs. Fitted Model
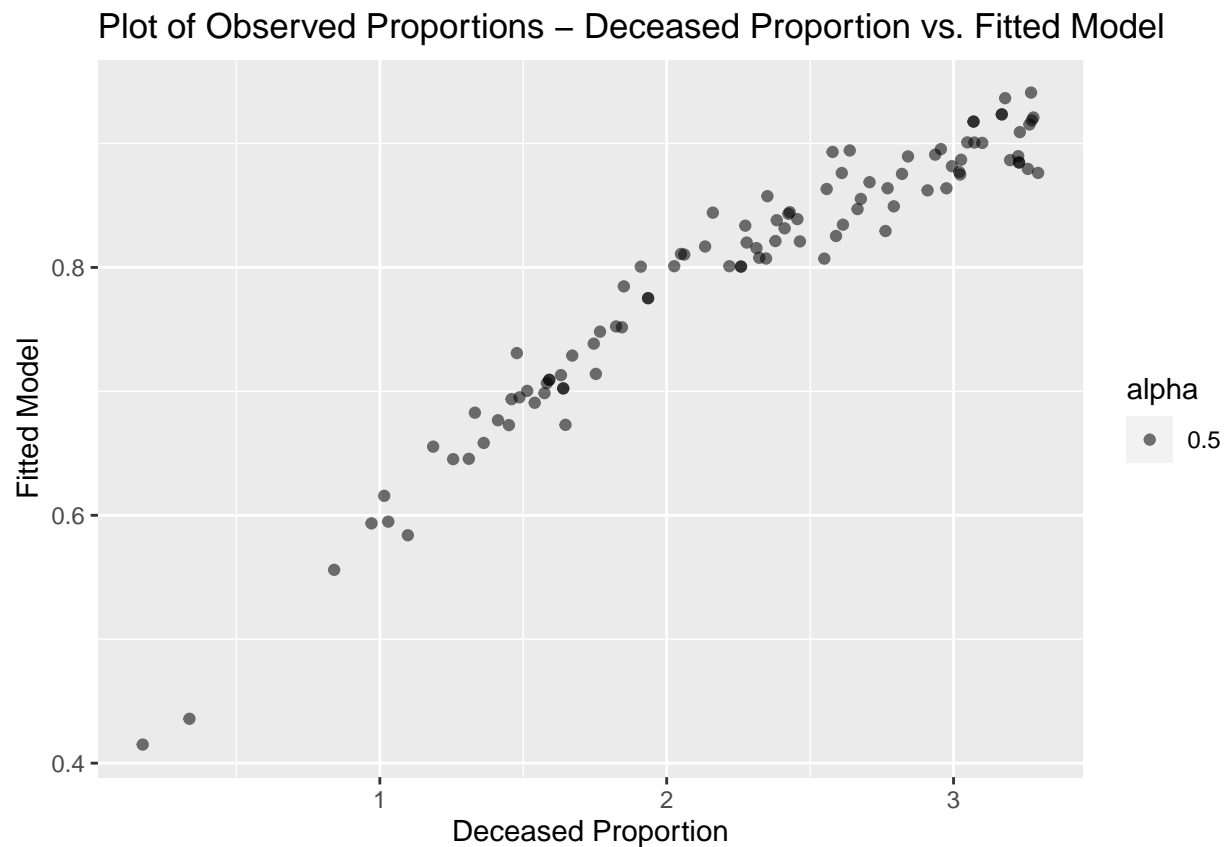
**Comment on the fit of the model**

The fit of the model is linear; almost resembling a $y = x$ line. However, it is slightly curved downwards towards the right of the plot. Thus, the fitted model works well to transform the data in part b) to a linear relationship.