ToothGrowth: Examining the Effect of Vitamin C Supplementation on Odontoblast Length

Executive Summary

Here, we examine the effect of vitamin supplementation on a proxy measure of tooth growth in guinea pigs. Both sources of vitamin C tested, orange juice and ascorbic acid, dose-dependently increased odontoblast length. Although at lower vitamin C-equivalent doses orange juice was more effective than ascorbic acid at increasing odontoblast length, both supplements were equally effective at the highest dose.

Explanatory Data Analysis and Basic Dataset Summary

After loading the ToothGrowth dataset and ggplot2 for visualization, let's examine the data using the str function.

```
str(ToothGrowth)
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

We see a simple data frame with 3 variables. According to the annotation of this dataset, 'len' is the length of odontoblasts cells of 60 guinea pigs at 3 dose levels (dose) and 2 delivery methods (supp), n = 10 each. We can further see the distribution of the variables using the summary function.

```
summary(ToothGrowth)
```

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

```
unique(ToothGrowth$dose)
```

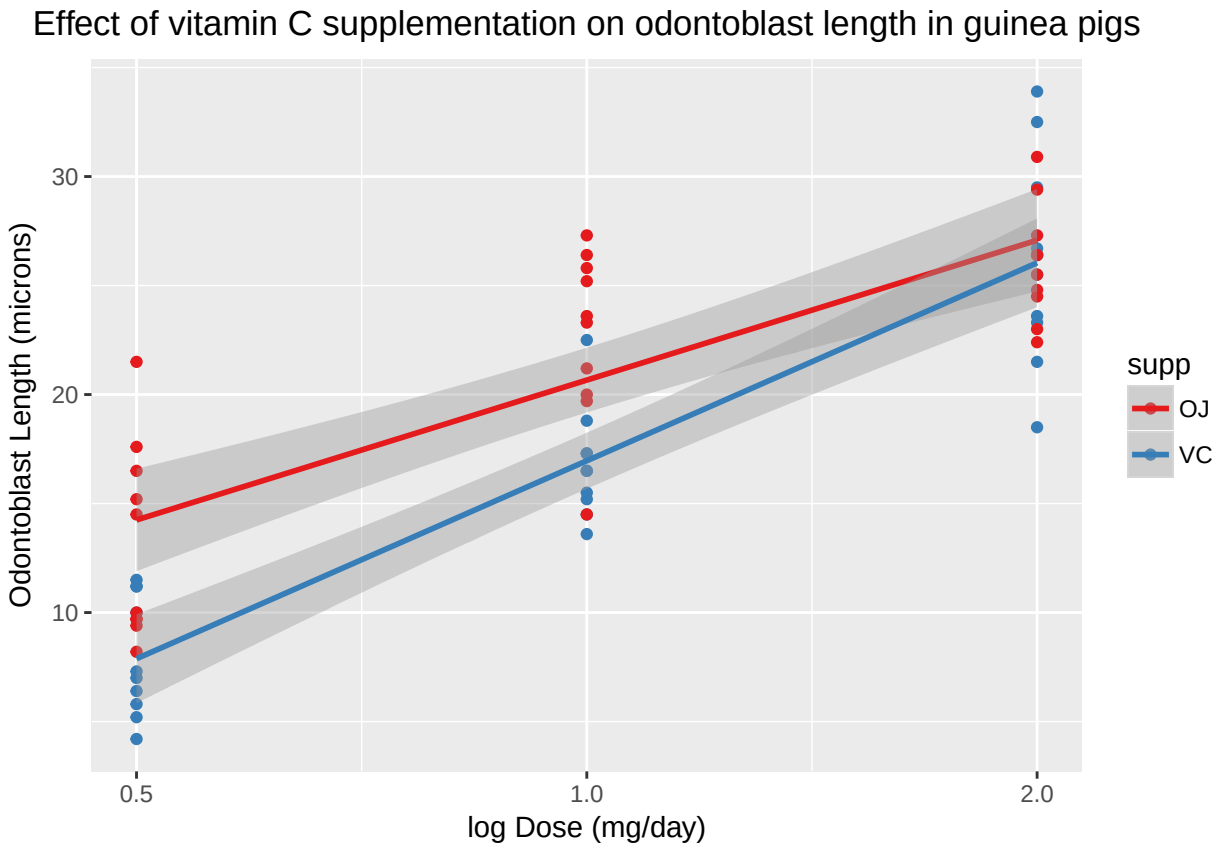
```
## [1] 0.5 1.0 2.0
```

As seen above, 'supp' is a factor variable taking one of 2 forms, OJ (orange juice) or VC (ascorbic acid, also known as vitamin C). Although specific doses were administered (0.5, 1, and 2 mg/day), 'dose' is a continuous numeric variable. 'Len' ranges from 4.2 microns up to 33.9 microns, with a mean of 18.8 and median of 19.25. Let's plot the data and see how the different experimental conditions affected length.

Importantly - this is dose-response data, which is traditionally visualized on a semi-log plot; we can tell ggplot2 to log10 transform our data.

```
ggplot(data = ToothGrowth, mapping = aes(x = dose, y = len, color = supp)) +
    scale_x_log10(breaks = c(0.5, 1, 2)) +
    scale_colour_brewer(palette = "Set1") +
    geom_point() +
    geom_smooth(method = "lm") +
```

```
    labs(title = "Effect of vitamin C supplementation on odontoblast length in guinea pigs", x = "log Do
    theme(plot.title = element_text(hjust = 0.5))
```



Statistical Inference

Generally, dose-response data is fit using a probit model or using nonlinear regression; however, with so few points and the assignment instructions, we will use a simple linear model. Assuming we are between 15% and 85% of maximal response, this linear fit is a fair representation as the sigmoidal curve is roughly linear in this range.

```
tooth_fit <- lm(len ~ log(dose)*supp, data = ToothGrowth)
tooth_fit
```

```
##
## Call:
## lm(formula = len ~ log(dose) * supp, data = ToothGrowth)
##
## Coefficients:
##      (Intercept)         log(dose)            suppVC  log(dose):suppVC
##           20.663             9.255            -3.700             3.845
```

```
anova(tooth_fit)
```

```
## Analysis of Variance Table
##
## Response: len
##                Df  Sum Sq Mean Sq  F value    Pr(>F)
```

```
## log(dose)       1 2400.95 2400.95 173.5134 < 2.2e-16 ***
## supp            1  205.35  205.35  14.8404 0.0003033 ***
## log(dose):supp  1   71.02   71.02   5.1327 0.0273658 *
## Residuals      56  774.89   13.84
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```