Linear Modelling 2

Starting point:

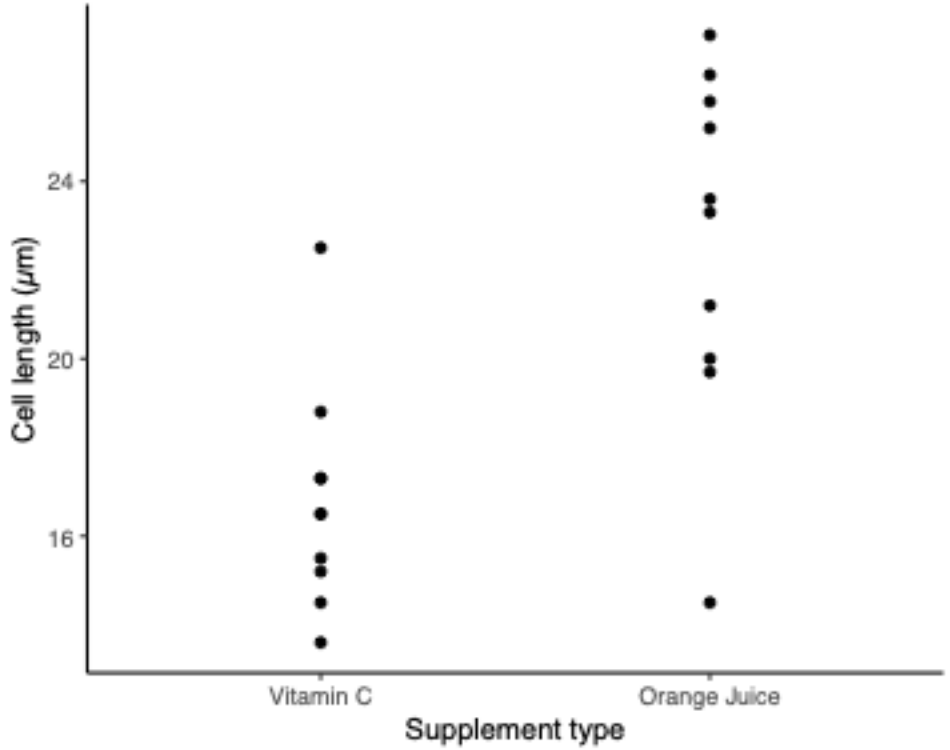
- Clarity about what is your response variable and what are explanatory variables
- Scientific questions about the relationship between response and explanatory variables
- Data in a long-form dataframe or tibble

Example Experiment – variables

- Classic experiment giving vitamin C to look at tooth growth in guinea pigs

- Two variables:
  - Cell length (continuous numeric)
  - Type of supplement: vitamin C versus orange juice (categorical)

- Roles in experiment
  - **Response variable**
    - What we care about
    - What we measured
    - Want to know if it responds to explanatory variables
    - Not set as part of experiment
    - Biological and technical '**noise**' in the measurement
  - **Explanatory variable**
    - Something we can manipulate
    - What we manipulated
    - Want to know if it affects the response variable
    - Set precisely as part of the experiment
    - Little or no '**noise**' in the value



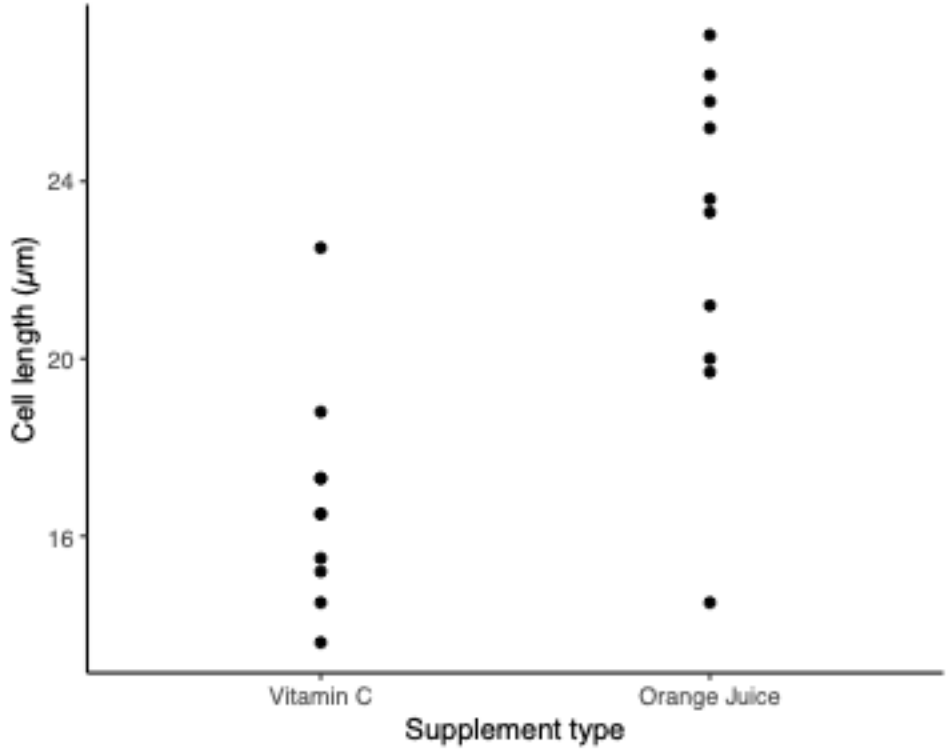
Crampton, E. W. (1947).

- Clarity about what is your response variable and what are explanatory variables

Example Experiment – questions

- Classic experiment giving vitamin C to look at tooth growth in guinea pigs

- Two variables:
  - Cell length
  - Type of supplement: vitamin C versus orange juice
- Questions
  - Does cell length depend on supplement type?
  - How much does cell length change with supplement type?
  - What's the expected cell length with Vitamin C?
  - How sure are we of the answers?



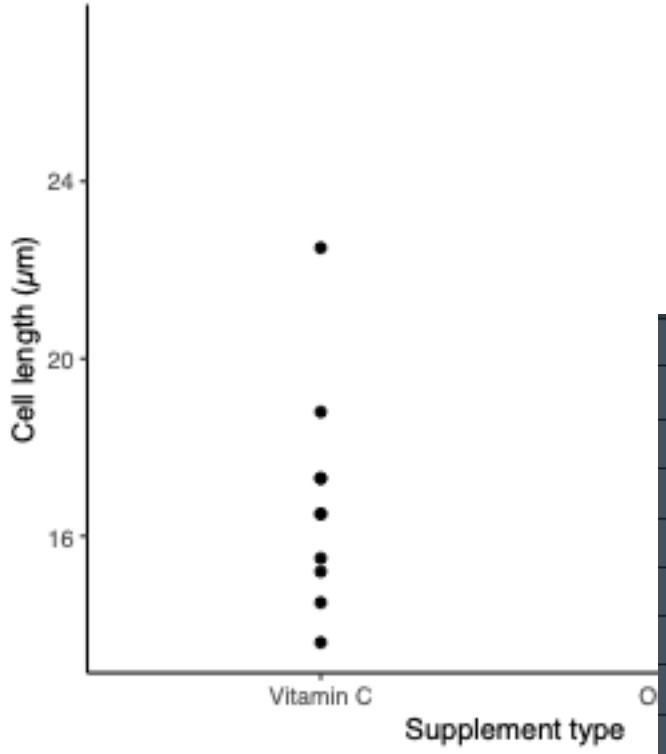
Crampton, E. W. (1947).

- Scientific questions about the relationship between response and explanatory variables
- Clarity about what is your response variable and what are explanatory variables

Example Experiment – data

- Classic experiment giving vitamin C to look at tooth growth in guinea pigs

- Two variables:
  - Cell length
  - Type of supplement: vitamin C versus orange juice
- Questions
  - Does cell length depend on supplement type?
  - How much does cell length change with supplement type?
  - What's the expected cell length with Vitamin C
  - How sure are we of the answers?



Crampton

```
> ToothGrowth2
```

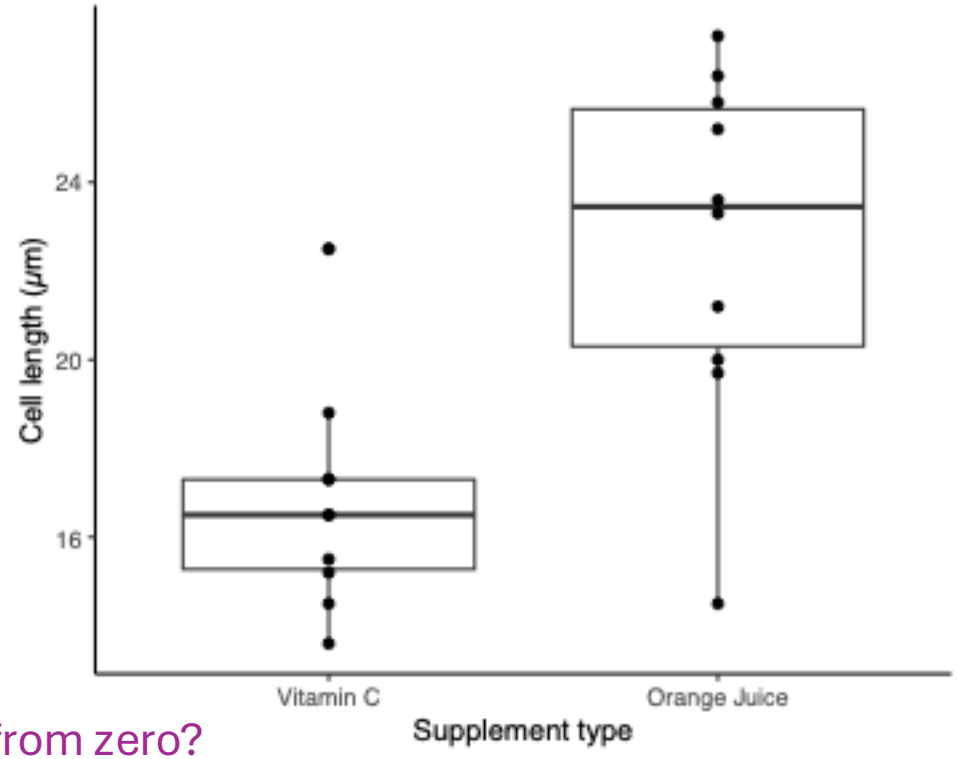
Filter

	len	supp
1	16.5	VC
2	16.5	VC
3	15.2	VC
4	17.3	VC
5	22.5	VC
6	17.3	VC
7	13.6	VC
8	14.5	VC
9	18.8	VC
10	15.5	VC
11	19.7	OJ
12	23.3	OJ
13	23.6	OJ
14	26.4	OJ
15	20.0	OJ
16	25.2	OJ

- Data in a long-form dataframe or tibble
- Scientific questions about the relationship between response and explanatory variables
- Clarity about what is your response variable and what are explanatory variables

Example Experiment – answering questions

- Classic experiment giving vitamin C to look at tooth growth in guinea pigs



Crampton, E. W. (1947).

- Two variables:
  - Cell length
  - Type of supplement: vitamin C versus orange juice
- Questions
  - Does cell length depend on supplement type?
  - How much does cell length change with supplement type?
  - What's the expected cell length with Vitamin C
  - How sure are we of the answers?
- Put lines through the data – can answer the questions:
  - Is the difference between the supplements **significantly** different from zero?
  - What is the difference between the supplements?
  - What's the estimate for Vitamin C?
  - What's the uncertainty estimate on any of the above?
- Data in a long-form dataframe or tibble
- Scientific questions about the relationship between response and explanatory variables
- Clarity about what is your response variable and what are explanatory variables

Example Experiment – Do the analysis

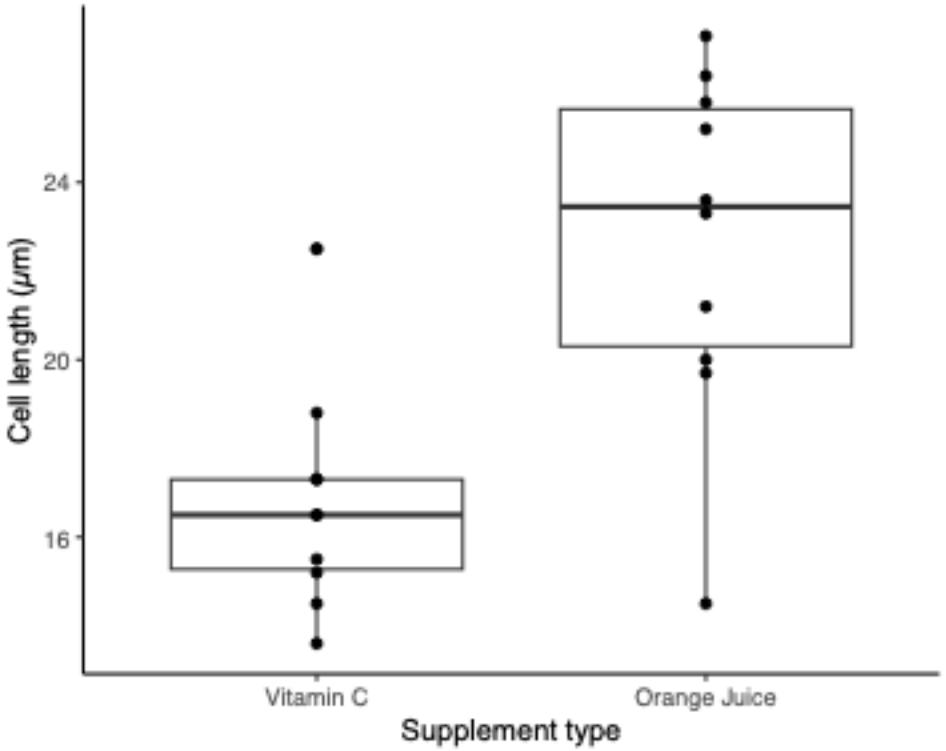
- Classic experiment giving vitamin C to look at tooth growth in guinea pigs

- Two variables:
  - Cell length
  - Type of supplement: vitamin C versus orange juice

```
> model <- lm(len ~ supp, data = ToothGrowth2)
```

- Minimises the squared distance from points to line
  - Only '**error**' in the response variable
  - Distance from the line the '**residual**'
  - Returns the value for one and the difference to the other

```
> summary(model)
```



Crampton, E. W. (1947).

- Data in a long-form dataframe or tibble
- Scientific questions about the relationship between response and explanatory variables
- Clarity about what is your response variable and what are explanatory variables

Example Experiment – Do the analysis

- Classic experiment giving vitamin C to look at tooth growth in guinea pigs

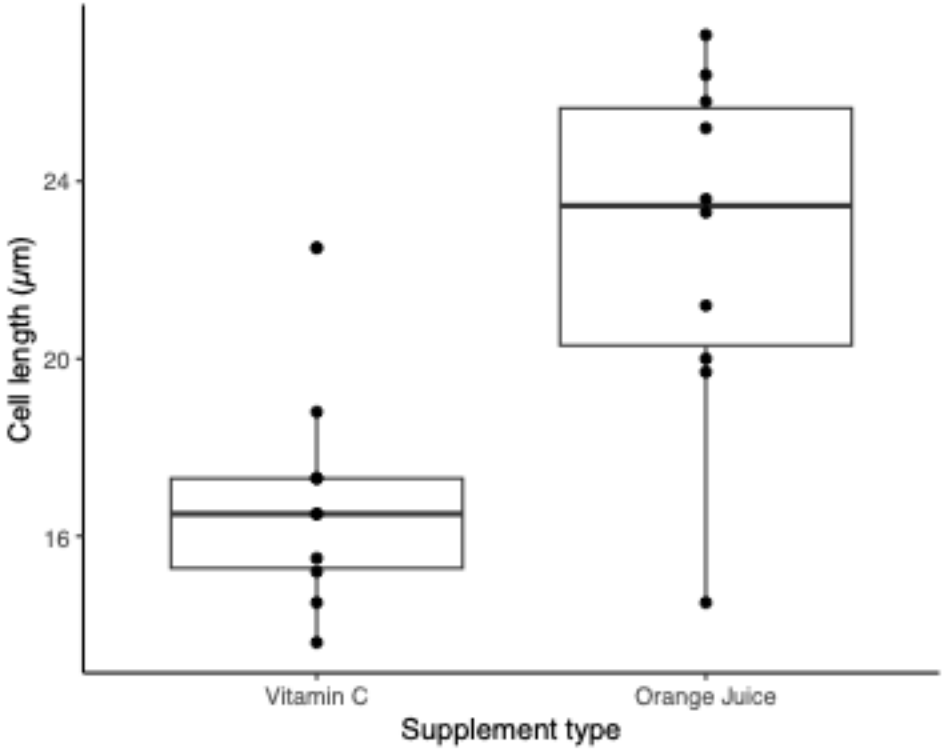
```
> model <- lm(len ~ supp, data = ToothGrowth2)
> summary(model)

Call:
lm(formula = len ~ supp, data = ToothGrowth2)

Residuals:
    Min      1Q  Median      3Q     Max
-8.200 -1.745  0.130  2.147  5.730

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    16.77       1.04  16.129 3.81e-12 ***
suppOJ          5.93       1.47   4.033 0.000781 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.288 on 18 degrees of freedom
Multiple R-squared:  0.4747,    Adjusted R-squared:  0.4455
F-statistic: 16.26 on 1 and 18 DF,  p-value: 0.0007807
```



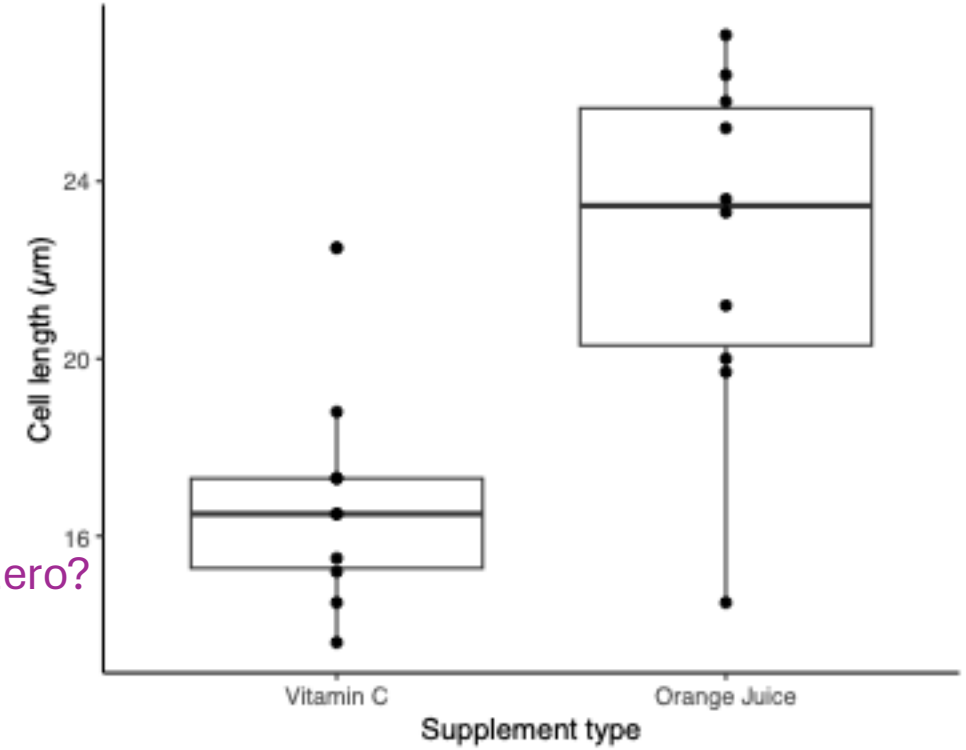
Crampton, E. W. (1947).

Example Experiment – Do the analysis

- Classic experiment giving vitamin C to look at tooth growth in guinea pigs

- Two variables:
  - Cell length
  - Type of supplement: vitamin C versus orange juice

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    16.77       1.04  16.129 3.81e-12 ***
suppOJ          5.93       1.47   4.033 0.000781 ***
```



Crampton, E. W. (1947).

- Is the difference between the supplements **significantly** different from zero?
- What is the difference between the supplements?
- What's the estimate for Vitamin C?
- What's the uncertainty estimate on any of the above?

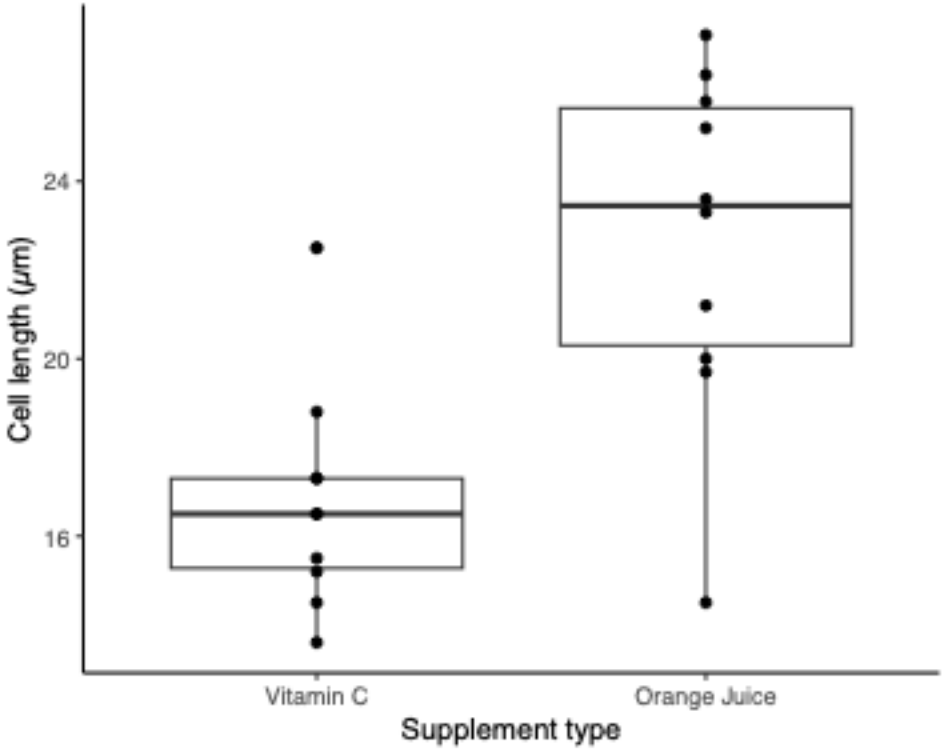
- Questions
  - Does cell length depend on supplement type? Yes, $P = 7.8 \times 10^{-4}$
  - How much does cell length change with supplement type? 5.9 µm
  - What's the expected cell length with Vitamin C 17 µm
  - How sure are we of the answers? 5.9 ±1.5 µm and 17 µm ±1.0 SE

Example Experiment – report it

- Classic experiment giving vitamin C to look at tooth growth in guinea pigs

- Two variables:
  - Cell length
  - Type of supplement: vitamin C versus orange juice

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    16.77       1.04  16.129 3.81e-12 ***
suppOJ          5.93       1.47   4.033 0.000781 ***
```



- Get the right units
- Give the uncertainty
- 2 or 3 significant figures sufficient – only report what's needed
- Describe it as an **Analysis of Variance**,
  - specifically **1-way analysis of variance** – only one explanatory variable (could add more!)
  - or just a **linear model** (many other sorts too)
- Questions
  - Does cell length depend on supplement type? Yes, $P = 7.8 \times 10^{-4}$
  - How much does cell length change with supplement type? 5.9 µm
  - What's the expected cell length with Vitamin C 17 µm
  - How sure are we of the answers? 5.9 ±1.5 µm and 17 µm ±1.0 SE

Crampton, E. W. (1947).

Crampton, E. W. **The growth of the odontoblasts of the incisor tooth as a criterion of the vitamin C intake of the guinea pig**. *J Nutr* 33, 491-504 (1947). https://doi.org/10.1093/jn/33.5.491