

ETC5512: Wild Caught Data

Week 9

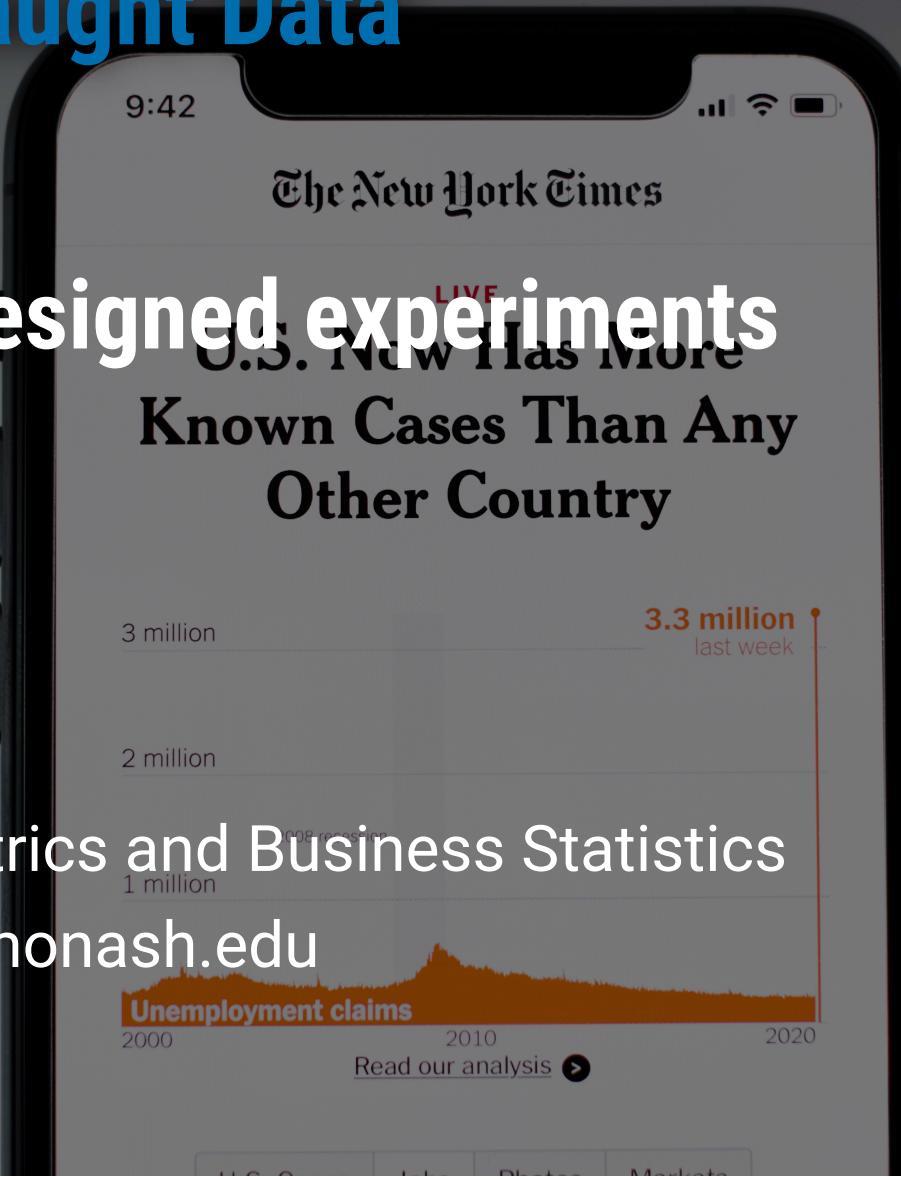
The language of designed experiments

Lecturer: *Emi Tanaka*

Department of Econometrics and Business Statistics

✉ ETC5512.Clayton-x@monash.edu

20th May 2020



 Experiments

- # Experiment is a procedure that is carried out to test a hypothesis or understand a phenomenon.
- # One of the most common experiment is a **comparative experiment** which compares different sets of conditions referred to as **treatments**.
- # These treatments are applied to **experimental units** - the smallest division of the experimental material such that any two units may receive different treatments in the actual experiment.
- # The smallest unit which the response is measured on is referred to as the **observational unit**.
- # Note that observational unit is *not* the observation nor the response!

Classical Design of Experiments

 Wheat Yield Trial

- ✿ A selective breeding experiment with **107 wheat varieties** (or *genotypes*) were conducted in South Australia in a field with plots laid out in a rectangular array with 22 rows and 15 columns.
- ✿ The breeders want to find a variety with *high yield*.
- ✿ The **treatments** are the *107 wheat varieties*.
- ✿ The **experimental units** are the *330 plots*.
- ✿ The **observational units** are also the *330 plots*.

Source: Gilmour et al. (1997) Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agric Biol Env Statistics*, 2, 269-293.

Wheat Yield Trial: Linear Model

1A

```
libraryagridat # data package  
head(gilmour.serpentine)  
  
##   col row rep          gen yield  
## 1   1   1  R1      ANGAS    483  
## 2   1   2  R1 BT_SCHOMBURG  526  
## 3   1   3  R1 DGR/MNX-9-9e  557  
## 4   1   4  R1 EXCALIBUR    564  
## 5   1   5  R1      JANZ    498  
## 6   1   6  R1     MACHETE    510  
  
ggplot(gilmour.serpentine) +  
  geom_histogram(aes(x = yield))
```

✿ Assuming the experiment is unstructured, we may propose a linear model: $\text{yield} = \text{mean} + \text{gen} + \text{error}$, where $\text{error} \sim N(0, \sigma^2)$.

✿ This model can be fitted in R as below.

```
fit1 <- lm(yield ~ 1 + gen,  
            data = gilmour.serpentine)
```

✿ Note that 1 (for the mean) is included by default and can be omitted.

Wheat Yield Trial: ANOVA

1B

- Analysis of variance (ANOVA) is historically used in the analysis of experimental data to test if any treatment is significantly different from others:

$\text{H}_0: \text{gen}_1 = \dots = \text{gen}_{107} = 0.$

- Although ANOVA is still used today, it is widely recognized as a special case of linear models.

- ANOVA table shows the decomposition of the total variation by source - we won't go into depth about ANOVA in this course.

```
anova(fit1)

## Analysis of Variance Table
##
## Response: yield
##              Df  Sum Sq Mean Sq F value Pr(>F)
## gen          106 2041055   19255  0.7428 0.9579
## Residuals  223  5781054   25924

fit2 <- lm(yield ~ 1, data = gilmour.serpentine)
anova(fit2, fit1)

## Analysis of Variance Table
##
## Model 1: yield ~ 1
## Model 2: yield ~ 1 + gen
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     329 7822108
## 2     223 5781054 106 2041055 0.7428 0.9579
```

Treatment Replications

CORRI	SUNLA	RAC80	VF300	TINCU	RAC80	VF655	VF519	RAC79	VG701	DSPRE	PELSA	VF300	MEERI	EXCAL
CADOL	SUNFI	RAC80	VF299	TINCU	RAC75	SWIFT	TRIDE	MOLIN	HOUTM	VF299	BT_SC	M5075	KATUN	WWH
BLADE	SUNBR	RAC80	VF508	CONDO	VF508	AMERY	WI221	RAC78	VG508	BD231	RAC81	SUNFI	BATAV	
BEULA	SHRIK	RAC80	VF655	RAC65	TINCU	SUNBR	RAC79	RAC81	(WWH)	RAC82	VG503	CADOL	VF519	WW183
BATAV	ROSEL	RAC80	(WWH)	M5097	DGR/M	PEROL	BD231	VF655	RAC77	COND0	VG878	RAC80	RAC82	RAC65
AROON	PEROL	RAC80	(WqKP)	K2011	VG878	JANZ	KITE	WI232	RAC80	GOROK	HALBE	VG701	RAC79	RAC75
AMERY	PELSA	RAC79	WI216	WW183	VG714	MACHE	BT_SC	RAC75	RAC71	VF302	RAC80	STILE	SUNLA	DUYEN
YARRA	OXLEY	RAC79	WW147	WW147	KULIN	VF664	TINCU	VF300	RAC81	CUNNI	DGR/M	WW140	SUNBR	RAC71
WYUNA	DUYEN	RAC79	WW140	M4997	SUNLA	LOWAN	WARBL	AROON	M5097	K2011	WILGO	JANZ	M5097	HOUTM
TRIDE	DSPRE	RAC78	VF519	WI232	TASMA	WW147	STILE	BEULA	VG506	DOLLA	RAC80	VG714	RAC81	RAC81
TATIA	LOWAN	RAC77	RAC82	WI231	EXCAL	(WqKP)	M4997	RAC81	RAC77	ROSEL	RAC79	KIATA	WW147	MOLIN
STILE	LARK	RAC77	RAC82	WI221	RAC79	WW183	BATAV	RAC81	RAC81	RAC79	RAC75	SPEAR	WYUNA	CORRI
SPEAR	KULIN	RAC77	RAC81	BD231	RAC65	RAC77	M5075	CORRI	RAC80	BEULA	SHRIK	RAC78	VF664	RAC80
SCHOM	KITE	RAC75	RAC81	VG878	VG302	MEERI	SHRIK	SCHOM	KIATA	RAC81	TASMA	RAC81	RAC81	RAC65
MOLIN	KIATA	RAC75	RAC81	VG714	RAC80	WILGO	CADOU	SUNFI	BLADE	WI216	SCHOM	RAC81	TRIDE	MACHE
MEERI	KATUN	RAC71	RAC81	VG701	OXLEY	KATUN	YARRA	CONDO	PELSA	RAC77	KULIN	WARBL	WI232	PEROL
MACHE	HOUTM	RAC69	RAC81	VG506	RAC65	RAC81	CUNNI	WI216	RAC81	AMERY	KITE	WI221	RAC81	YARRA
JANZ	HALBE	RAC65	RAC81	VG503	RAC80	DSPRE	RAC69	RAC82	RAC81	RAC80	OXLEY	TINCU	VF655	
EXCAL	GOROK	WILGO	RAC81	VG127	K2011	DUYEN	WW140	LARK	RAC81	RAC77	VG127	RAC80	TINCU	VG506
DGR/M	DOLLA	WARBL	RAC81	VF302	RAC82	VG127	HALBE	WI231	WYUNA	WW147	TATIA	LOWAN	RAC81	WI231
BT_SC	M5075	TASMA	RAC81	VF664	GOROK	VG503	RAC81	ROSEL	TATIA	ANGAS	RAC89	VF655	BLADE	M4997
ANGAS	CUNNI	SWIFT	RAC81	VF655	WW147	RAC81	ANGAS	SPEAR	DOLLA	AROON	(WqKP)	LARK	SWIFT	RAC77

- ✿ The varieties VF655, TINCURIN and WW1477 have a **replication** of 6, the remaining 104 varieties each have a replication of 3.
- ✿ Treatment **replications are essential** in an experiment; without any replication, no treatment variation can be measured nor distinguished from unit variation.
- ✿ More replications are desirable for accuracy, however, there is always a tension to balance the cost of the experiment.

Systematic Design of Experiments

BEULA	DGR/M	K2011	M5075	OXLEY	RAC75	RAC79	RAC81	RAC81	SPEAR	TATIA	VF508	VG701	WI232	YARRA
BD231	DGR/M	K2011	M4997	OXLEY	RAC75	RAC79	RAC81	RAC81	SHRIK	TATIA	VF508	VG508	WI232	YARRA
BD231	CUNNI	K2011	M4997	DUYEN	RAC75	RAC79	RAC80	RAC81	SHRIK	TASMA	VF508	VG508	WI231	YARRA
BD231	CUNNI	JANZ	M4997	DUYEN	RAC71	RAC79	RAC80	RAC81	SHRIK	TASMA	VF302	VG508	WI231	WYUNA
BATAV	CUNNI	JANZ	LOWAN	DUYEN	RAC71	RAC79	RAC80	RAC81	SCHON	TASMA	VF302	VG503	WI231	WYUNA
BATAV	CORRI	JANZ	LOWAN	SPRE	RAC71	RAC79	RAC80	RAC81	SCHON	SWIFT	VF302	VG503	WI221	WYUNA
BATAV	CORRI	HOUTM	LOWAN	SPRE	RAC69	RAC79	RAC80	RAC81	SCHON	SWIFT	VF300	VG503	WI221	WW183
AROON	CORRI	HOUTM	LARK	SPRE	RAC69	RAC78	RAC80	RAC81	ROSEL	SWIFT	VF300	VG127	WI221	WW183
AROON	COND	HOUTM	LARK	MOLIN	RAC69	RAC78	RAC80	RAC81	ROSEL	SUNLA	VF300	VG127	WI216	WW183
AROON	COND	HALBE	LARK	MOLIN	RAC65	RAC78	RAC80	RAC81	ROSEL	SUNLA	VF299	VG127	WI216	WW147
ANGAS	COND	HALBE	KULIN	MOLIN	RAC65	RAC77	RAC80	RAC81	RAC82	SUNLA	VF299	VF664	WI216	WW147
ANGAS	CADOU	HALBE	KULIN	MEERI	RAC65	RAC77	RAC80	RAC81	RAC82	SUNFI	VF299	VF664	WARBL	WW147
ANGAS	CADOU	GOROK	KULIN	MEERI	RAC65	RAC77	RAC80	RAC81	RAC82	SUNFI	TRIDE	VF664	WARBL	WW147
AMERY	CADOU	GOROK	KITE	MEERI	RAC65	RAC77	RAC80	RAC81	RAC82	SUNFI	TRIDE	VF664	WARBL	WW147
AMERY	BT_SC	GOROK	KITE	MACHE	RAC65	RAC77	RAC80	RAC81	RAC82	SUNBR	TRIDE	VF664	VG878	WW147
AMERY	BT_SC	EXCAL	KITE	MACHE	PEROU	RAC77	RAC80	RAC81	RAC82	SUNBR	TINCU	VF664	VG878	WW140
(WqKP)	BT_SC	EXCAL	KIATA	MACHE	PEROU	RAC77	RAC80	RAC81	RAC81	SUNBR	TINCU	VF664	VG878	WW140
(WqKP)	BLADE	EXCAL	KIATA	M5097	PEROU	RAC77	RAC80	RAC81	RAC81	STILE	TINCU	VF664	VG714	WW140
(WqKP)	BLADE	DOLLA	KIATA	M5097	PELSA	RAC77	RAC80	RAC81	RAC81	STILE	TINCU	VF664	VG714	WILGO
(WWH)	BLADE	DOLLA	KATUN	M5097	PELSA	RAC75	RAC80	RAC81	RAC81	STILE	TINCU	VF519	VG714	WILGO
(WWH)	BEULA	DOLLA	KATUN	M5075	PELSA	RAC75	RAC79	RAC81	RAC81	SPEAR	TINCU	VF519	VG701	WILGO
(WWH)	BEULA	DGR/M	KATUN	M5075	OXLEY	RAC75	RAC79	RAC81	RAC81	SPEAR	TATIA	VF519	VG701	WI232

- ✿ The treatments appear to be randomly ordered before.
- ✿ Why don't we order the treatments in a **systematic order** like on the left?
- ✿ Isn't this easier to manage the experiment?
- ✿ Systematic designs are prone to **bias** and **confounding**.

Randomisation

- ✿ Treatment must be allocated *randomly* to experimental units.
- ✿ This avoids:
 - systematic bias - e.g. all flu vaccine A tested in January (summer) and all flu vaccine B tested in July (winter).
 - selection bias - e.g. giving the treatment that you are testing to the sick patients and placebo to those that are healthy.
 - other bias - e.g. the lab technician giving the treatment to the first rat that is taken out of the cage.
- ✿ So how do we randomise?
- ✿ We can make a reproducible design using R.
- ✿ Be sure to use `set.seed` in the beginning of your script.

Completely randomised design using R

```
set.seed(2020) # for reproducibility
# first create the field array
expand_grid(col = 1:15, row = 1:22) %>%
  # create plot id (optional)
  mutate(plot = 1:n()) %>%
  # genotype 1-104 has 3 reps
  # genotype 105-107 has 6 reps
  mutate(gen = c(rep(1:104, each = 3),
                rep(105:107, each = 6))) %>%
  # now randomly permute the genotypes
  mutate(gen = sample(gen))

## # A tibble: 330 x 4
##       col     row   plot  gen
##   <int> <int> <int> <int>
## 1     1      1     1    79
## 2     1      2     2    29
## 3     1      3     3    8
## 4     1      4     4    72
## 5     1      5     5   106
## 6     1      6     6    91
## 7     1      7     7    55
## 8     1      8     8    57
## 9     1      9     9    66
## 10    1     10    10    37
## # ... with 320 more rows
```

- ✿ **Blocks** are used to group the experimental units into alike units.
- ✿ If well done, blocking can lower the variance of treatment contrasts which increase power.
- ✿ Blocking reduces the residual degrees of freedom which can decrease power if the sample size is small.
- ✿ A non-homogeneous block (i.e. units within block are *not* alike) can decrease the power of the experiment.

You can form blocks from:

- ✿ **Natural discrete divisions** between experimental units.
E.g. in experiments with people, gender make an obvious block.
- ✿ Grouping experimental units with similar **continuous gradients**.
E.g., if the experiment is spread out in time or space and there exists no obvious natural boundaries, then an arbitrary boundary may be chosen to group experimental units that are contiguous in time or space.



Blocking in field trial

- ✿ In agricultural field trials, it is common to have some underlying soil fertility trend.
- ✿ So contiguous plots may be grouped to form a block.
- ✿ The wheat yield trial actually employed 3 blocks (as colored on left) as recorded in the variable rep.
- ✿ The treatment is best to be balanced across the blocks.
- ✿ If possible, block sizes should have the same size.

How to randomise design if there are blocks?

```
set.seed(20052020)
expand_grid(col = 1:15, row = 1:22) %>%
  mutate(plot = 1:n()) %>%
  # 3 blocks ->
  # block 1 is col 1-5,
  # block 2 is col 6-10,
  # block 3 is col 11-15
  mutate(rep = case_when(
    col %in% 1:5 ~ "block1",
    col %in% 6:10 ~ "block2",
    col %in% 11:15 ~ "block3"
  )) %>%
  # every block contains:
  # - 1 replicate of gen 1-104
  # - 2 replicates of gen 105-107
  group_by(rep) %>%
  mutate(gen = c(1:107, 105:107)) %>%
  # randomise within `rep`
  mutate(gen = sample(gen))
```

```
## # A tibble: 330 x 5
## # Groups:   rep [3]
##       col     row   plot rep      gen
##       <int> <int> <int> <chr> <int>
## 1       1       1     1  1 block1    28
## 2       2       1     2  2 block1   106
## 3       3       1     3  3 block1    93
## 4       4       1     4  4 block1    41
## 5       5       1     5  5 block1    92
## 6       6       1     6  6 block1    71
## 7       7       1     7  7 block1    78
## 8       8       1     8  8 block1    38
## 9       9       1     9  9 block1    13
## 10    10       1    10 10 block1    89
## # ... with 320 more rows
```

Wheat Yield Trial: Analysis

2A

- ✿ We take the block effect into account in our linear model:

```
fitb <- lm(yield ~ 1 + rep + gen,  
            data = gilmour.serpentine)
```

- ✿ The ANOVA table takes into account block source of variation now:

```
anova(fitb)
```

```
## Analysis of Variance Table  
##  
## Response: yield  
##           Df  Sum Sq Mean Sq F value    Pr(>F)  
## rep        2 2828701 1414351 105.8720 < 2e-16 ***  
## gen       106 2041055   19255   1.4414  0.01235 *  
## Residuals 221 2952352   13359  
## ---
```

- ✿ Variation due to block is large!
- ✿ Take that into account, now the \ (p\)-value for gen is small.

- ✿ This indicates that at least one variety has significantly different mean than others provided model assumptions are satisfied.

(The assumption is violated in this case, but we won't go into this.)

Wheat Yield Trial: Analysis

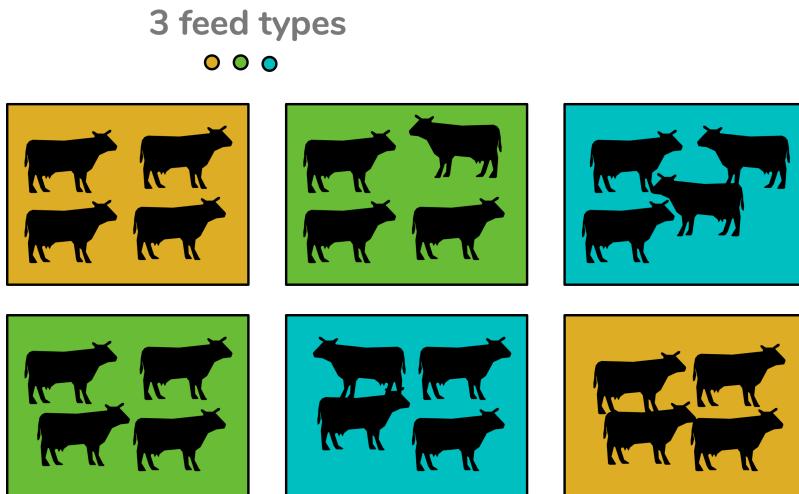
2B

```
broom::tidy(fitb) %>%
  select(term, estimate) %>%
  filter(str_detect(term, "gen")) %>%
  arrange(-estimate)

## # A tibble: 106 x 2
##   term      estimate
##   <chr>     <dbl>
## 1 genVG878    52.3
## 2 genRAC811    42.3
## 3 gen(WqKPWmH*3Ag) 24.3
## 4 genVF508     11.7
## 5 genRAC772     5.00
## 6 genWI216     4.00
## 7 genRAC779     3.67
## 8 genRAC820     -1.
## 9 genVF519     -1.
## 10 genRAC798    -1.67
## # ... with 96 more rows
```

✿✿✿ The variety VG878 is performing the best according to the analysis.

Replication vs Repetition

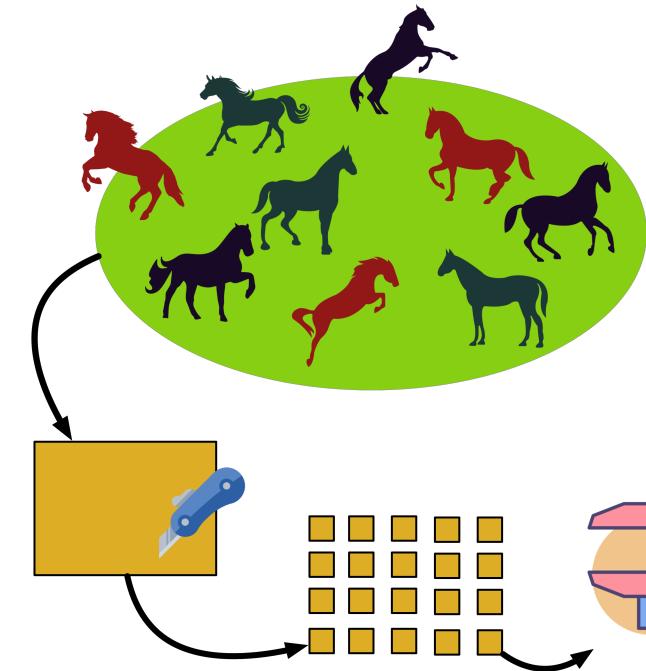


- ✿ Three feed treatments are compared on 24 calves
- ✿ The calves are kept in 6 pens with 4 calves per pen
- ✿ Each feed is applied to two whole pens
- ✿ Every calf is weighed individually

- ✿ What are the experimental units? Observational units?
 - ✿ The pens are the experimental units.
 - ✿ The calves are the observational units.
- ✿ In this experiment,
 - the **replication** of each treatment is 2, and
 - the **repetition** of each treatment is 8.
- ✿ Why do we need to distinguish this?

Example: Grafting on horses

3 grafting methods



- A surgeon is going to use 9 horses in an experiment
- He wants to compare 3 methods of grafting skin
- He intended to use 3 animals for each method
- After the graft was complete he would take a sample of new skin from each horse
- He would then cut each sample into 20 (tiny) pieces and use a precision instrument to measure the thickness of each piece

- ✿ **Treatments** are the 3 grafting methods.
- ✿ **Experimental units** are the 9 horses
- ✿ **Observational units** are the 20×9 skin pieces
- ✿ If we assume that the grafting results in uniform thickness, then any variation in thickness of the 20 pieces from the same skin is a result of measurement error.
- ✿ The variation of thickness between horse skins is variation due to grafting + residual variation.

Simulation: Grafting on horses

```
set.seed(1)
# no difference between trts
trt <- c(0, 0, 0)
# random deviation for horse
hordev <- rnorm(9, 0, 20)
# there are 9 horse
sim_df <- tibble(horse = 1:9) %>%
  # 3 grafting with 3 reps
  mutate(grft = rep(1:3, 3)) %>%
  # cut each grafted skin to 20 pieces
  mutate(piece = list(1:20)) %>%
  # let each piece be one row
  unnest(piece) %>%
  # now simulate the response
  mutate(y = 300 + # mean
         trt[grft] + # trt effect
         hordev[horse] + # horse dev
         rnorm(n(), 0, 5)) %>% # OU dev
  mutate_if(is.integer, as.factor)
```

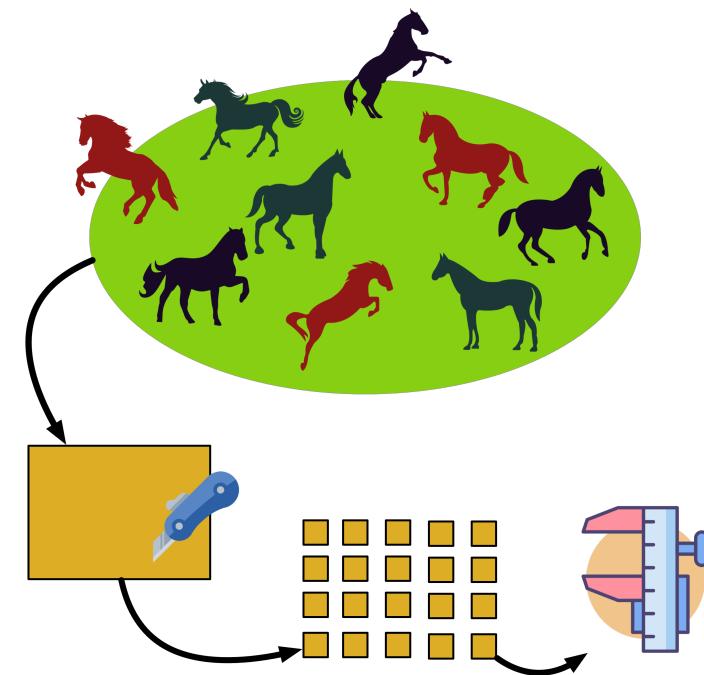
- ✿ Note we don't need to randomise here as we are doing a simulation and not a design.

```
anova(lm(y ~ graft + horse, data = sim_df))
## Analysis of Variance Table
##
## Response: y
##             Df Sum Sq Mean Sq F value    Pr(>F)
## graft       2   9035  4517.4  205.86 < 2.2e-16 ***
## horse        6  32225  5370.8  244.75 < 2.2e-16 ***
## Residuals 171   3752     21.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
```

- ✿ The p -value for graft is small indicating there is at least one grafting method is significantly different!

Pseudo-replication

3 grafting methods



- From the simulation there should be no difference between grafting methods.

- The previous analysis treats skin pieces as replications of treatment.
- The treatment that the skin pieces received are however not independent!
- The treatment of repetition as replication in the analysis is referred to as **pseudo-replication**.

```
summary(aov(y ~ graft + Error(horse/piece), data = sim_df))

##
## Error: horse
##           Df Sum Sq Mean Sq F value Pr(>F)
## graft      2   9035   4517    0.841  0.476
## Residuals  6  32225    5371

##
## Error: horse:piece
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals 171   3752    21.94
```

Case Study Vaccine Field Trials & *infectious disease*

Experimental vs. Observational Studies

i

In a **controlled experiment**, the investigators allocate the treatments to the units (that may be people, mice, plants, etc).

i

In an **observational study**, the investigators observe units without manipulation or intervention.

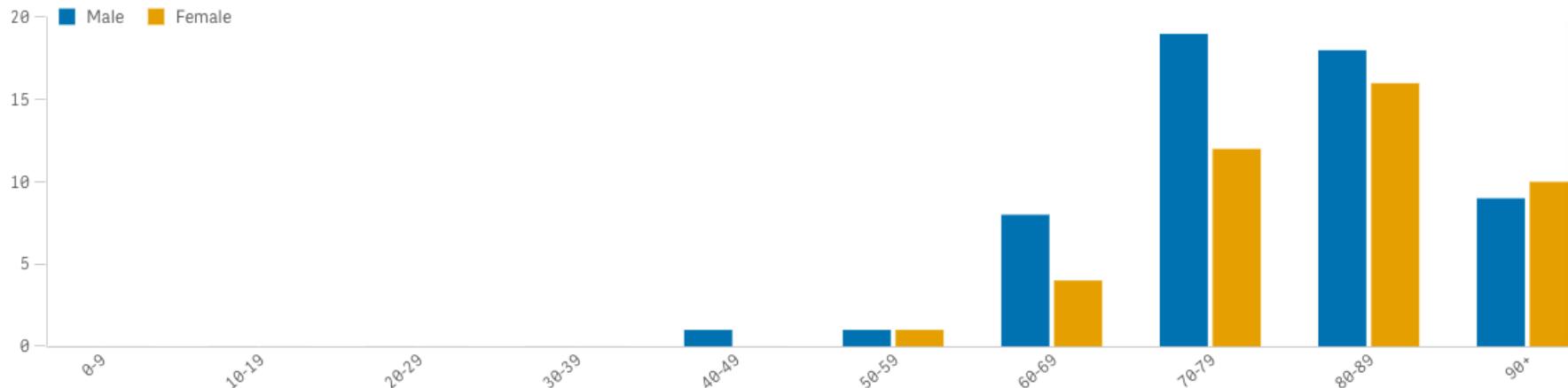
i

In a **well-controlled experiment**, the difference in response between treatment groups should be only due to the treatment.

COVID-19 deaths by age group and sex

This graph shows the number of COVID-19 deaths for males and females by age group since 22 January 2020.

Source: NNDSS data 19/5/2020



Note that the chart does not include cases where age or sex are unknown

Is this an experimental or observational study?

Claims from observational study

Based on previous graphs, which statements are true?

1. Age determines the risks of death from COVID-19.
2. Community transmission of coronavirus is rare. Most infected cases are from overseas or close contact with confirmed case.
3. Men are at a higher risk of death from COVID-19.
4. Children have a much higher immunity against coronavirus.
5. NSW has the highest number of coronavirus infected cases out of all the Australian states and territories.
6. Australia started to go into lock-down from Sun 22/03/2020. The shutdown measures were effective.

Correlation vs Causation

i

Correlation does not imply causation.

- ✿ Age may not be a defining factor that determines the risk of death from COVID-19.
- ✿ There is increasing observations that those with *underlying health conditions* are at a higher risk of death from COVID-19.
- ✿ Many underlying health conditions, such as hypertension, is prevalent in elderly.
- ✿ It may be the combination of COVID-19 and other health conditions that is the causal factor of death.
- ✿ You can read more about this in this Conversation article: [Coronavirus: the puzzle of why the risk of death is greater for men and for the elderly.](#)

What was the data collection procedure?

Chief Medical Officer Brendan Murphy said there was no point in testing Australians simply because they had respiratory or cold and flu symptoms.

Other than a "small and controlled" cluster of community transmission in Sydney, cases were largely confined to returned travelers.

"If you're a returned traveler or you've been in contact with someone who has been a confirmed case, then you should be tested. But other Australians do not need testing and all they're doing is putting an unnecessary burden on the testing," he said.

Read the article [here](#).

Statisticians urge random testing

- ✿ Nicholas Fisher (former chief scientist in statistics in CSIRO) and Dennis Teewin (Australian Statistician) urge random testing in Australia [\[link\]](#)
- ✿ Without an experimental study, it is hard to estimate the true level of community transmissions.
- ✿ In the beginning, the criteria for testing was for those who returned from overseas and those that were in close contact with a confirmed case.
- ✿ It is not surprising then that the number of cases almost all belonged to those two categories in the beginning.

Control

i

A **control** is an experimental unit that did not receive any treatment.

- ✿ In order to know the effect of treatment, e.g. vaccine, we must compare with something, e.g. the control.
- ✿ Confusingly, in experimental descriptions, some regard control as one of "treatments"; some when referring to treatments, exclude control; and then some use both with context needed to infer whether control is included or not.
- ✿ Note: you do not always need a control!
- ✿ If there is already effective treatment that is applied as a standard, then testing should be compared with this standard treatment (as was the case for breeding trial).
- ✿ Is "do-nothing" treatment wise comparison though?

Placebo

- ✿ When people are enrolled in a trial to test a potential treatment, the control group may be aware that they are not receiving the treatment; likewise the treatment group are aware they are receiving treatment.
- ✿ This may result in unconscious or conscious **bias** where the control group expects they will not get better and the treatment group expects that they will get better; thus the difference in the result may not be due to treatment but due to this bias.

i

A **placebo** is a medical treatment or procedure designed to have no therapeutic value.

- ✿ All participants enrolled in a study then will be assigned to a treatment or placebo group but will not be told which group

Double-Blind Study

- ✿ In a randomised controlled study, the participants are **blind** to whether they are in the treatment or placebo group.
- ✿ The experimenters, however, can still bias the results if they know which group the participant belongs to.

i

A **double-blind study** is an experimental study that neither the participants nor the experimenters know who is receiving which treatment.

- ✿ This again helps to reduce any potential bias in the study.

Confounding variable

i

A **confounding variable** is a variable that is associated with the variable of interest (usually the treatment) and the response.

- ✿ E.g., consider the lab technician giving the diet treatment to the first rat that is taken out of the case and leaving the other rats as control.
- ✿ The first rat taken out of cage may be slower or lazier than other rats (hence easier to catch to take out of the cage).
- ✿ In that case the genetics or character of the rat may be confounded with treatment.

The Salk Vaccine Field Trial

- ✿ The first **polio epidemic** hit the United States in 1916 claiming hundreds of thousands of victims, especially children.
- ✿ National Foundation for Infantile Paralysis (NFIP) was ready to test the vaccine developed by Jonas Salk in the real world.
- ✿ A controlled experiment was proposed to test the effectiveness of the vaccine on grade 1, 2 and 3 children at selected school districts though the country where the risk of polio was high.

- ✿ In total two million children were involved although not all parents consented to their children to be vaccinated.



Photo: Historical Society of Pennsylvania

Design for the NFIP Study

Vaccinate all grade 2 children whose parents would consent, leaving children in grades 1 and 3 as controls.

- ✿✿ Can grade 2 children whose parents did not consent be included as control?
- ✿✿ What are the potential issues with such a design?
- ✿✿ Polio is a contact disease. Would incidences of disease be higher in grade 2?

Results from Salk vaccine trial of 1954

The rate is the number of polio cases per 100,000 in each group.

```
## Warning in kableExtra::kable_styling(.) : Please specify format
## kableExtra can customize either HTML or LaTeX outputs. See http://
## haozhu233.github.io/kableExtra/ for details.
```

Table: Randomised controlled experiment

Group	Participants	Rate
Vaccinated	200,745	28
Placebo	201,229	71
(no consent)	338,778	46
Not Vaccination		
Incomplete Vaccination	8,484	24

Group	Participants	Rate
Vaccinated	200,745	28
Placebo	201,229	71
(no consent)	338,778	46
Not Vaccination		
Incomplete Vaccination	8,484	24

What does the result say?

Group	RCT Rate	NFIP Rate
Vaccinated	28	25
Placebo/Control	71	54
Not Vaccination (no consent)	46	44
Incomplete Vaccination	24	40

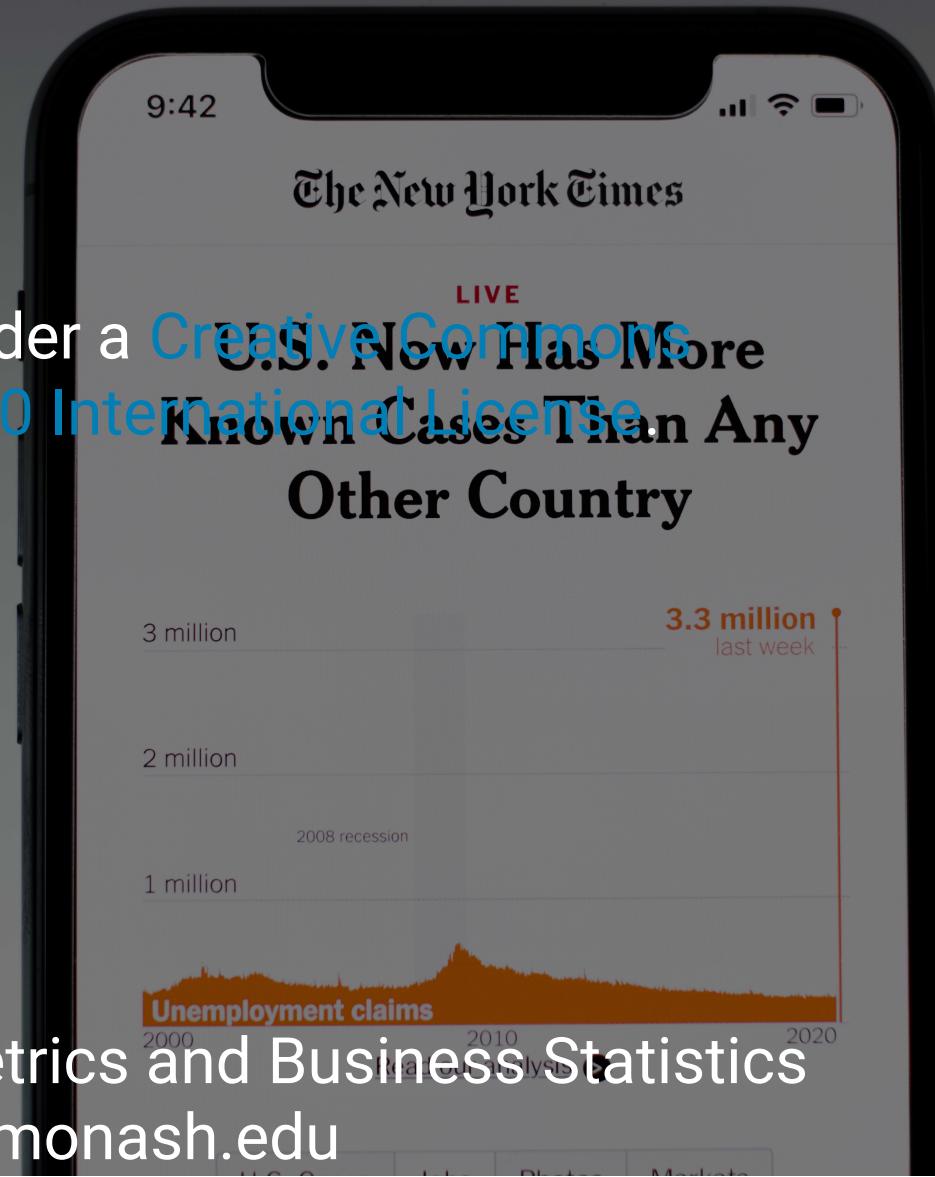
✿ RCT and NFIP trial sampled from school districts with similar exposures to the polio virus.

- ✿ Both the not vaccinated (no consent) and placebo/control group did not receive the treatment but why is the rate of polio cases less in the not vaccinated (no consent) group?
- ✿ Higher income parents would more likely consent to treatment than lower-income parents.
- ✿ Children of higher income parents are more vulnerable to polio.
- ✿ Many forms of polio are hard to diagnose and in borderline cases.

That's it!



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).



Lecturer: Emi Tanaka

Department of Econometrics and Business Statistics

ETC5512.Clayton-x@monash.edu