DigiCAT Tutorials:

A Statistical App for Counterfactual Analysis

On behalf of Dr Aja Murray[1]

[1]Department of Psychology, University of Edinburgh, UK

aja.murray@ed.ac.uk

DIGICAT
DIGITAL
COUNTERFACTUAL
ANALYSIS TOOL

What is counterfactual analysis?

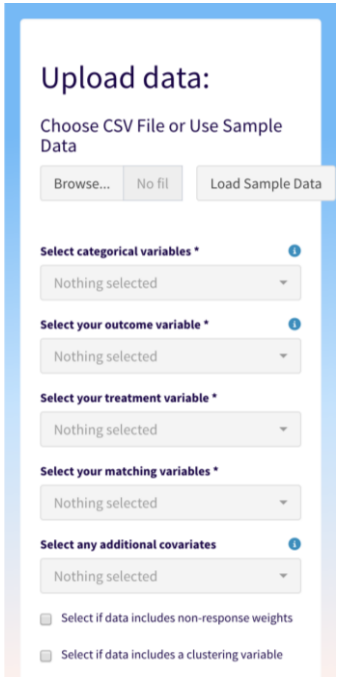
- understand the impact of a treatment
- comparing people who experienced that treatment to a hypothetical 'counterfactual' version of themselves who did not
- controls (who did not experience the treatment) who are very similar to the treated people in all important respects

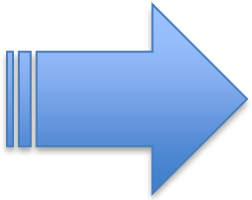
Three main variables:

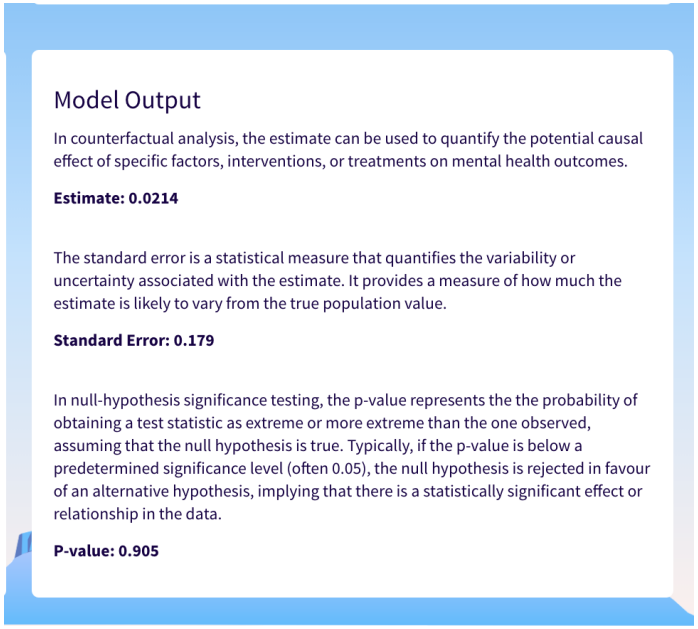
- A treatment variable (e.g., social media use) (e.g., VR treatment)
- An outcome variable (e.g., depression) (e.g., psychosis)
- Confounders or 'matching variables' (e.g., ADHD, prior depression, sex/gender, SES, physical health etc.) (e.g., VR experience, personality, SES, physical health etc.)

Why develop DigiCAT?

- RCTs are not always possible, they are expensive and challenging to do well. Their highly controlled nature also means that they may lack ecological validity so that results from RCTs might not apply very well in the real world.
- counterfactual analysis is not as widely used in mental health research as one might expect.
- it can take rather a long time to code counterfactual analyses using existing packages and things get complicated when we had to consider things like missingness, clustering, large numbers of confounders etc.







What counterfactual analysis approaches are offered in DigiCAT?

1. Propensity score matching for binary (traditional propensity score analysis)
2. Propensity score matching for ordinal treatment variables (non-bipartite optimal matching)
3. Propensity score weighting for binary variables (inverse propensity of treatment weighting; IPTW)

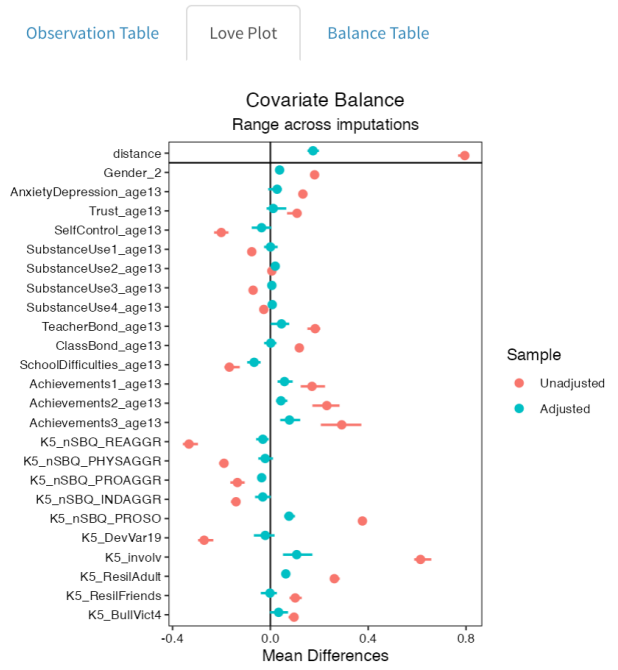
Sample size: > 25 after matching

Mean Difference <0.15

Propensity score matching?

Matching Ratio: 1:1

1:1 is the most commonly used matching ratio, although this ratio can be changed, for example, if there are many more members belonging to the control group than members belonging to the treatment group. In 1:1 matching, we match one unit belonging to the intervention group with one unit belonging to the control group that has the closest propensity score. The default matching strategy in DigiCAT is a nearest neighbour ('greedy') matching algorithm, with a 1:1 matching ratio. As our default is matching without replacement, the selected units will be taken out of the donor pool once matched. If you think this may not be best for you, see our tutorial pages for more detailed guidance on choosing an appropriate matching ratio.

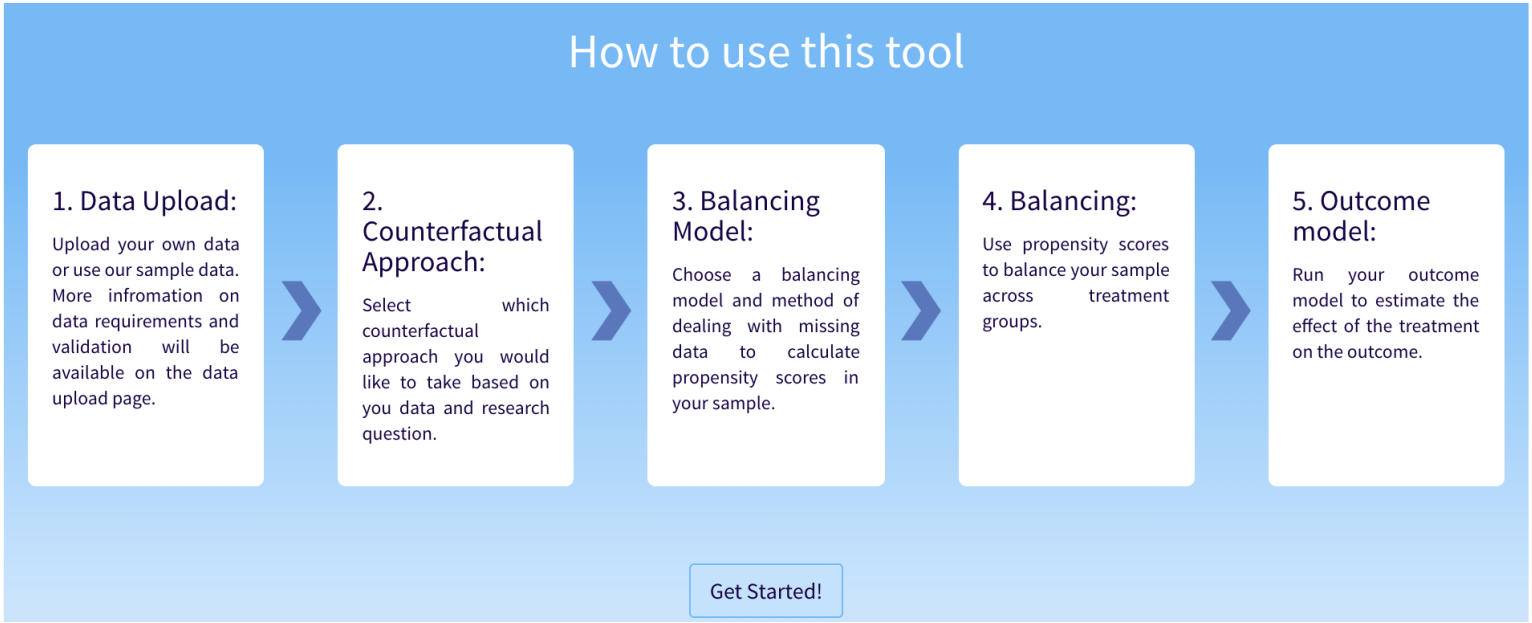


How to use DigiCAT?

1. R package for local computer

```
install.packages("remotes")
remotes::install_github("josiahpjking/DigiCAT@develop")
DigiCAT::run_DigiCAT(enableLocal = TRUE)
```

2. Web version: https://digicatapp.shinyapps.io/DigiCAT/
3. Tutorials: https://uoe-digicat.github.io/



How to use DigiCAT?

DIGICAT
DIGITAL COUNTERFACTUAL ANALYSIS TOOL

- Analysis
- Terms & Conditions
- About
- Dark Mode

uoe_digicat-group@uoe.onmicrosoft.com

GET STARTED — DATA UPLOAD — APPROACH — BALANCING — OUTCOME

Prev | Validate Data | Next

Upload data:

Choose CSV File or Use Example Data

Please install DigiCAT locally to enable file upload. See the DigiCAT github page.

Browse... | No fil | Load Example Data

Select categorical variables *
Gender, Reading_age15, Substancel

Select your outcome variable *
Anxiety_age17

Select your treatment variable *
Reading_age15

Select your matching variables *
Gender, AnxietyDepression_age13, T

Select any additional covariates
Nothing selected

Select if data includes survey weights †

Select if survey weight variable above

Requirements | Data | Validation

Data Dimensions:

Data contains 27 columns and 643 rows.

Your data has an appropriate number of columns.

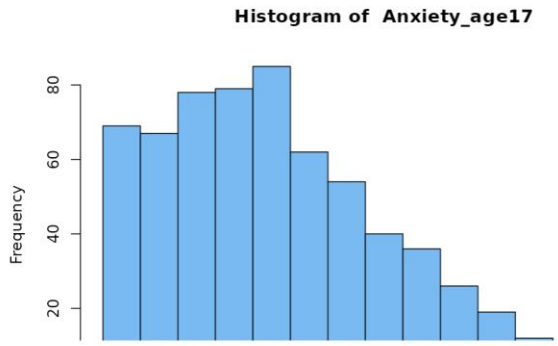
Your data has an appropriate number of rows.

Outcome Variable Type:

There are 16 unique observations in your chosen outcome variable.

Outcome Variable Skewness:

Your outcome variable does not appear highly skewed, skewness: 0.6



How to use DigiCAT?

DIGICAT
DIGITAL COUNTERFACTUAL ANALYSIS TOOL

- Analysis
- Terms & Conditions
- About
- Dark Mode

uoe_digicat-group@uoe.onmicrosoft.com

GET STARTED — DATA UPLOAD — APPROACH — BALANCING — OUTCOME

Outcome: Anxiety_age17
Treatment: Reading_age15

Prev Next

1. Choose a Counterfactual Approach:

- (●) Propensity Matching (PSM)
- ( ) Inverse probability of treatment weighting (IPTW)

Description:

The aim of counterfactual analysis is to estimate the causal effects of 'exposures' or 'treatments' by comparing what actually happened (observed outcomes) with what would have happened if a different action had been taken (counterfactual outcomes). Because we can never directly observe counterfactual outcomes, we compare groups who differ in their treatment. However, in observational settings where random allocation into different treatments is not possible, researchers may employ methods that such as 'matching' or 'weighting' of participants to ensure that the different treatment groups being compared are 'balanced' with respect to other characteristics. For further guidance on choosing an approach see the 'choosing a counterfactual analysis approach' section of the tutorials.

Note: If counterfactual analysis has already been run, changing this parameter will require the rerun of all subsequent steps.

Propensity Matching:

You've chosen propensity matching. This approach involves creating balanced comparison groups by matching treated individuals with similar untreated

2. Choose a Method of Dealing with Missingess:

- (●) Multiple Imputation
- ( ) Complete Case

Note: If counterfactual analysis has already been run, changing this parameter will require the rerun of all subsequent steps.

Description:

Missing data can reduce statistical power and introduce bias when it is related to the variables under study. In order to minimise the impact of missingness in our data, DigiCAT offers several approaches to handle missing values. For more information on missingness, see the 'missing data' section of the tutorials.

Multiple Imputation:

The fundamental idea behind multiple imputation is to create several (M) completed datasets by predicting what the missing values would have been if we could observe them. These datasets are then analysed separately, and the results of each of these datasets are combined, or 'pooled', together. The variability across datasets allows the standard errors to take account of the uncertainty due to the fact that some of the data are predicted rather than observed. DigiCAT imputes 5 complete datasets, conducts the analysis in each and then combines the results across them. However, don't forget that this statistical method relies on assumptions, which can be difficult to test. If

3. Choose a Balancing Model:

- (●) Logistic Regression

Note: If counterfactual analysis has already been run, changing this parameter will require the rerun of all subsequent steps.

Description:

In order to balance our dataset, we must first calculate the probability each individual has of being in the treatment group, based on their observed characteristics, i.e., their propensity score. On this page, you will select a balancing model to use with your dataset to predict the likelihood of an individual being treated. Different modelling approaches can be used for this step.

Logistic Regression:

Logistic regression is a method of modelling the relations between a set of predictors (e.g., matching variables) and a binary outcome (such as whether or not someone is in the treatment group). In propensity score analysis it can be used to estimate the propensity scores for each individual. See the logistic regression section of our tutorials for more information.

How to use DigiCAT?

DIGICAT
DIGITAL COUNTERFACTUAL ANALYSIS TOOL

- Analysis
- Terms & Conditions
- About
- Dark Mode

uoe_digicat-group@uoe.onmicrosoft.com

Prev | Run | Next

Choose a Matching Method:

- ○ Optimal
- ◉ Nearest Neighbour (NN)

Matching Method:

In order to balance covariates between treatment groups, propensity score matching involves matching individuals based on their propensity scores, which represent their likelihood of being treated based on observed characteristics. The goal is to create a pseudo-randomized comparison between the treatment and control groups by matching individuals who have similar or close propensity scores. See our tutorial section on propensity score matching for more detailed guidance on choosing an appropriate matching method.

Nearest Neighbour (NN):

Nearest neighbour matching is used in counterfactual analysis as a method for pairing treated and control cases with similar propensity scores. In DigiCat, nearest neighbour greedy matching is implemented whereby the most similar treated and control cases are matched first. Then, from those left the remaining most similar treated and control cases are paired and so on and so forth until all viable matches have been made. See the nearest neighbour matching tutorial for more details.

Choose a Matching Ratio:

- ◉ 1:1
- ○ 1:k

Matching Ratio:

Matching ratios in propensity score matching refer to the number of control/untreated individuals that are matched to each treated individual. In DigiCAT, matching ratios can be specified to control the trade-off between achieving better balance between treatment groups and maintaining a larger sample size. See

Output:

Once you have selected your matching method and ratio, press 'Run' to get output.

How to use DigiCAT?

DIGICAT
DIGITAL COUNTERFACTUAL ANALYSIS TOOL

- Analysis
- Terms & Conditions
- About
- Dark Mode

uoe_digicat-group@uoe.onmicrosoft.com

GET STARTED — DATA UPLOAD — APPROACH — BALANCING — OUTCOME

Outcome: Anxiety_age17
Treatment: Reading_age15

Approach: psm
Missingness: mi
Model: glm

Matching Method: nearest
Matching Ratio: 1:1

Prev | Run

Choose an Outcome Model:

- (●) Linear Regression (Outcome ~ Treatment * Matching Variables)
- ( ) Linear Regression (Outcome ~ Treatment + Matching Variables)
- ( ) Linear Regression (Outcome ~ Treatment)

Note: Changing this parameter will require outcome model to be rerun

Outcome Model:

Linear regression is a way of modelling the associations between exploratory variable(s) and a continuous outcome variable. The model takes the form y = bX + e, where y and x are our outcome and explanatory variables respectively, and e is the random error. Of interest here is b, which is the estimated effect of our treatment variable on our outcome. By fitting a model to a sample that is either matched or weighted according to the propensity of treatment, we can better estimate the **causal** effect of the treatment on our outcome variable. As you select an outcome model option more information will appear.

Linear regression adjusted for matching variables as covariates and their interactions with the treatment variable:

This option will fit a linear regression with your chosen outcome and the following predictors: i) the treatment variable ii) all matching variables iii) all interactions between treatment and matching variables. This method addresses the fact that dependencies (or 'clustering') are introduced into the data due to matching and can help to deal with the effects of any remaining imbalance that isn't completely addressed by the matching or weighting. It additionally deals with imbalance and bias that might be related to the fact that treatment status might interact with the matching variables (e.g., a treatment might affect people of different genders differently so if there isn't a perfect gender balance there might be some bias in estimating the treatment effect).

Model Output

In counterfactual analysis, the estimate can be used to quantify the potential causal effect of specific factors, interventions, or treatments on mental health outcomes.

Estimate: 0.057

The standard error is a statistical measure that quantifies the variability or uncertainty associated with the estimate. It provides a measure of how much the estimate is likely to vary from the true population value.

Standard Error: 0.2341

In null-hypothesis significance testing, the p-value represents the the probability of obtaining a test statistic as extreme or more extreme than the one observed, assuming that the null hypothesis is true. Typically, if the p-value is below a predetermined significance level (often 0.05), the null hypothesis is rejected in favour of an alternative hypothesis, implying that there is a statistically significant effect or relationship in the data.

P-value: 0.8099

95% Confidence Interval: -0.4268 to 0.5407