Example from Data Science Ready

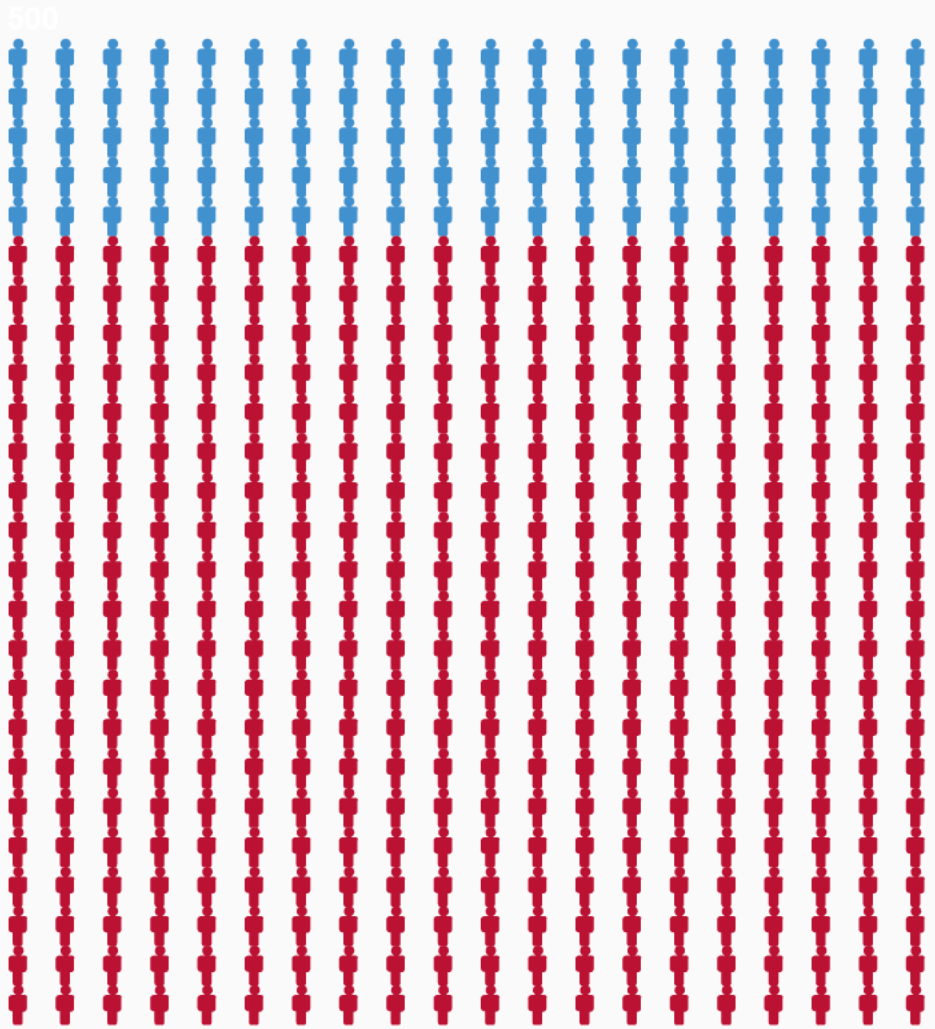
Decision thresholds: Simulation of sepsis diagnosis

1 Sample Size

In this simulation, 500 patients will have their sepsis status predicted by a computational model. Among the 500 patients, 100 of them (20%) actually have sepsis, and 400 (80%) do not. **Scroll down** to run the model prediction.

Sepsis:

Non-Sepsis:



2 Decision Threshold

Use this slider to update the

2 Decision Threshold

Use this slider to update the decision threshold. The higher the decision threshold, the more confident the model must be before you act on its prediction and treat a patient for sepsis.

Threshold: 50%

Patients are classified into four categories:

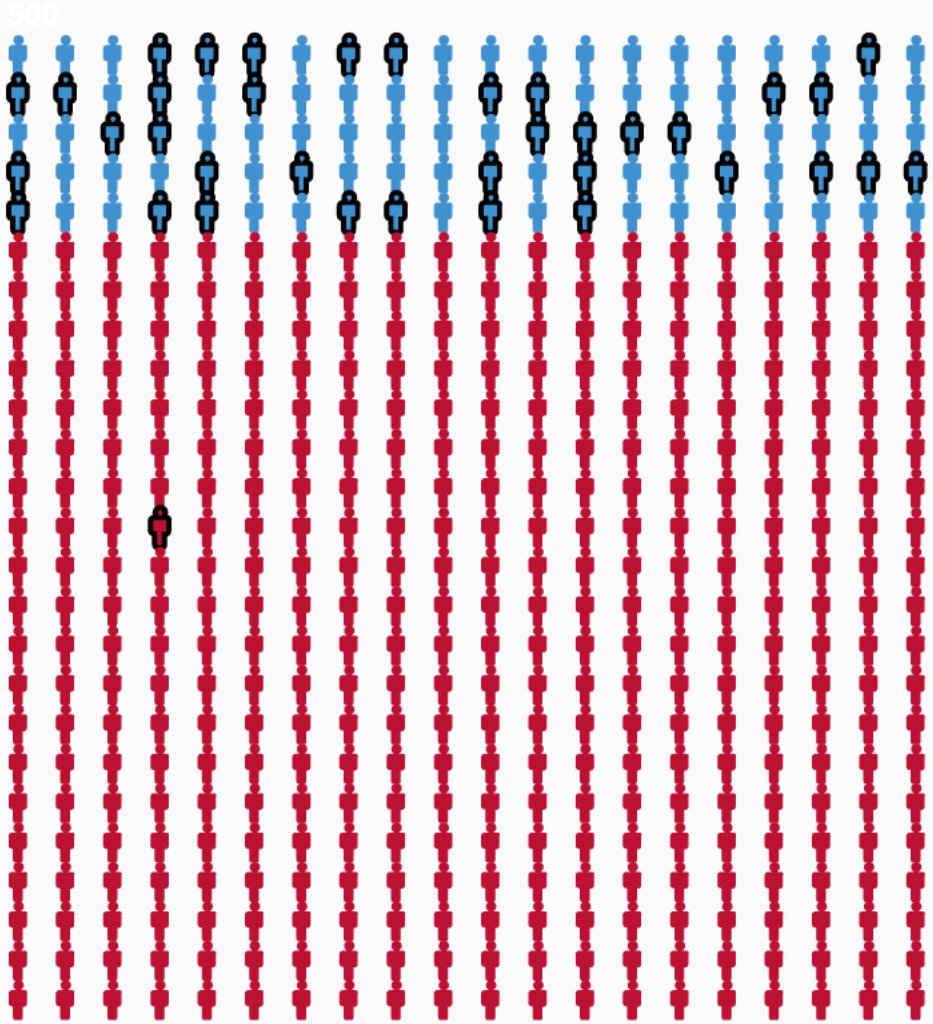
64 True Positives:

36 False Negatives:

399 True Negatives:

1 False Positives:

Scroll down to check the model accuracy.



3 Model Accuracy

Total Population: 500

65 people were diagnosed with sepsis, but 1 of them do not have it. 435 people were diagnosed as healthy, but 36 of them actually have sepsis.

1 false positives
36 false negatives

$$Accuracy = \frac{True\ Results}{Total\ Population}$$

93%

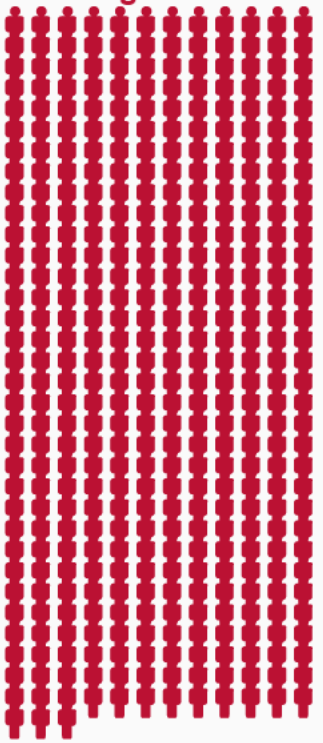
Scroll up if you want to try a different threshold.

True Positive: 64

False Negative: 36

False Positive: 1

True Negative: 399



Raj Chetty proposed interactive

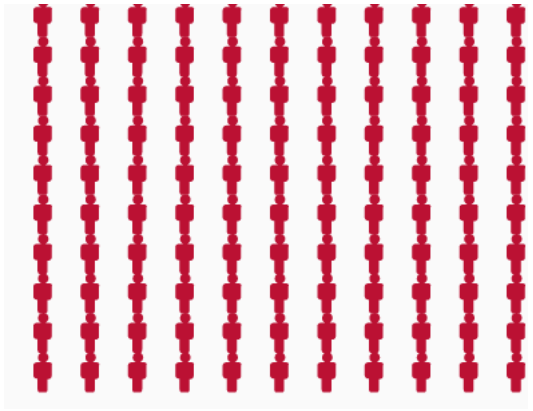
1 Treatment Effect

In this simulation, we will compare the earnings of 500 individuals, 250 growing up in the MLK Towers and 250 growing up in Wakefield in the Bronx.

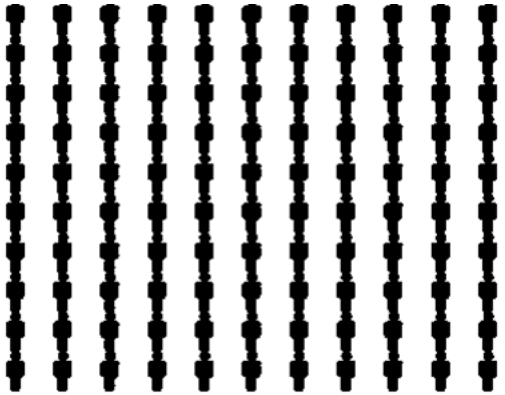
In this simulation, the only reason that different people earn different amounts is either because of their neighborhood or their race.

The causal effect of growing up in the Wakefield in the Bronx compared with growing up in Martin Luther King Jr. Towers in Harlem is **$2000.**

That is, if you grow up in Wakefield, you'd make $2000 more than you'd make if you grew up in the Martin Luther King Jr. Towers.



Martin Luther King Jr. Towers



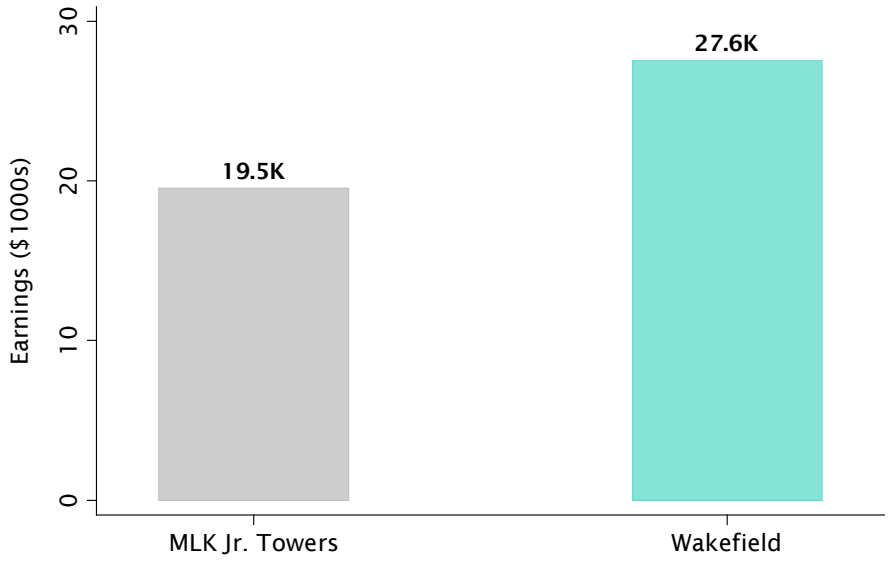
Wakefield

2 Differences in Earnings

The observed average earnings for the children growing up in Wakefield is $27.6K.

The observed average earnings for the children growing up in the MLK Towers is $19.5K

The difference is $8K, which is four times as large as the real causal effect!

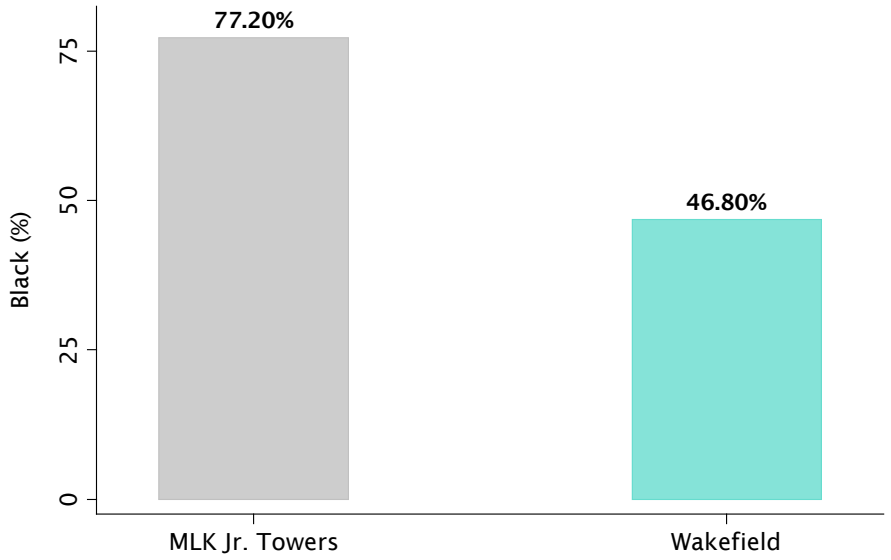


3 Differences in Racial Composition

Note also that 46.8% of children of the children growing up in Wakefield are Black.

In the MLK Towers, it is 77.2%.

The $8K difference in earnings is also picking up racial disparities in earnings, leading to an overestimate of the causal effect.

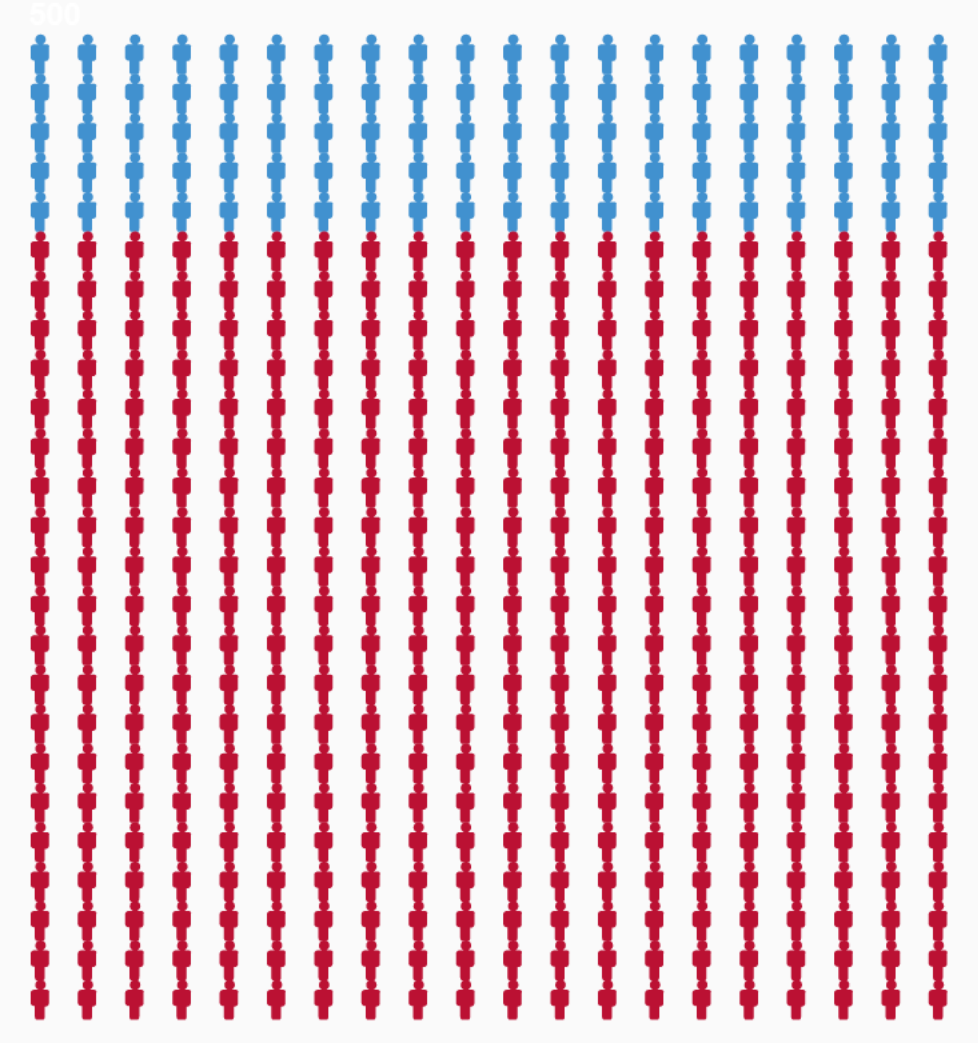


4 Random assignment

We will now simulate an experiment where we randomly assign half of the 500 children to grow up in the Martin Luther King Jr. Towers in Harlem and the other half to grow up in Wakefield in the Bronx.

Press the button below to randomly assign each observation to one of the two groups:

Randomize



5 Differences in Earnings and Racial Composition

The difference in earnings between the two groups is $2.21K

Why does this work better than the observational comparison? Randomization ensures that the treatment and control groups are comparable.

The difference in racial composition in the two groups is now only 1.4 percentage points.