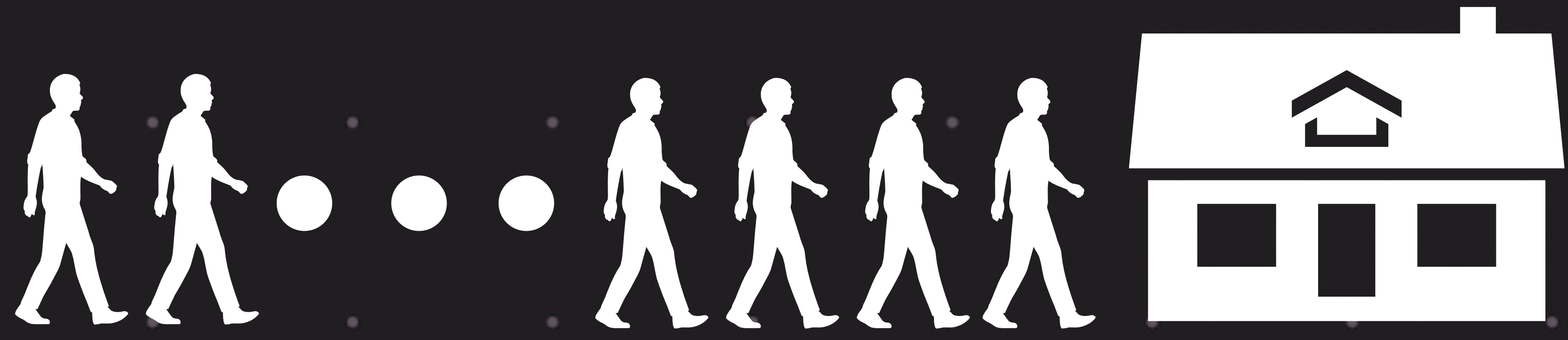
The average number of customers entering a store is 2000 per month

A marketing company is hired to improve this number

The next month, number of customers was seen to be 2128

With 95% confidence, is this improvement statistically significant?

2000 per month on average



What should the null and alternate hypothesis be?

$H_0 : \mu = 2000 \quad H_a : \mu > 2000$

What is the test statistic?

N: Number of people entering the store in a month

Distribution of the test statistic N? Poisson with rate 2000 per month

Right, left tailed, or two-tailed? Right tailed

What is the p-value?

$P[N \geq 2128 \,|\, H_0 \text{ is true}] = 1 - P[N \leq 2127 \,|\, H_0 \text{ is true}]$ = `1 - poisson.cdf(2127, mu=2000)` $= 0.002$

What is α? $\alpha = 0.05$

Is p-value < α? Yes

We reject the null hypothesis We say that the marketing worked

Recommender System

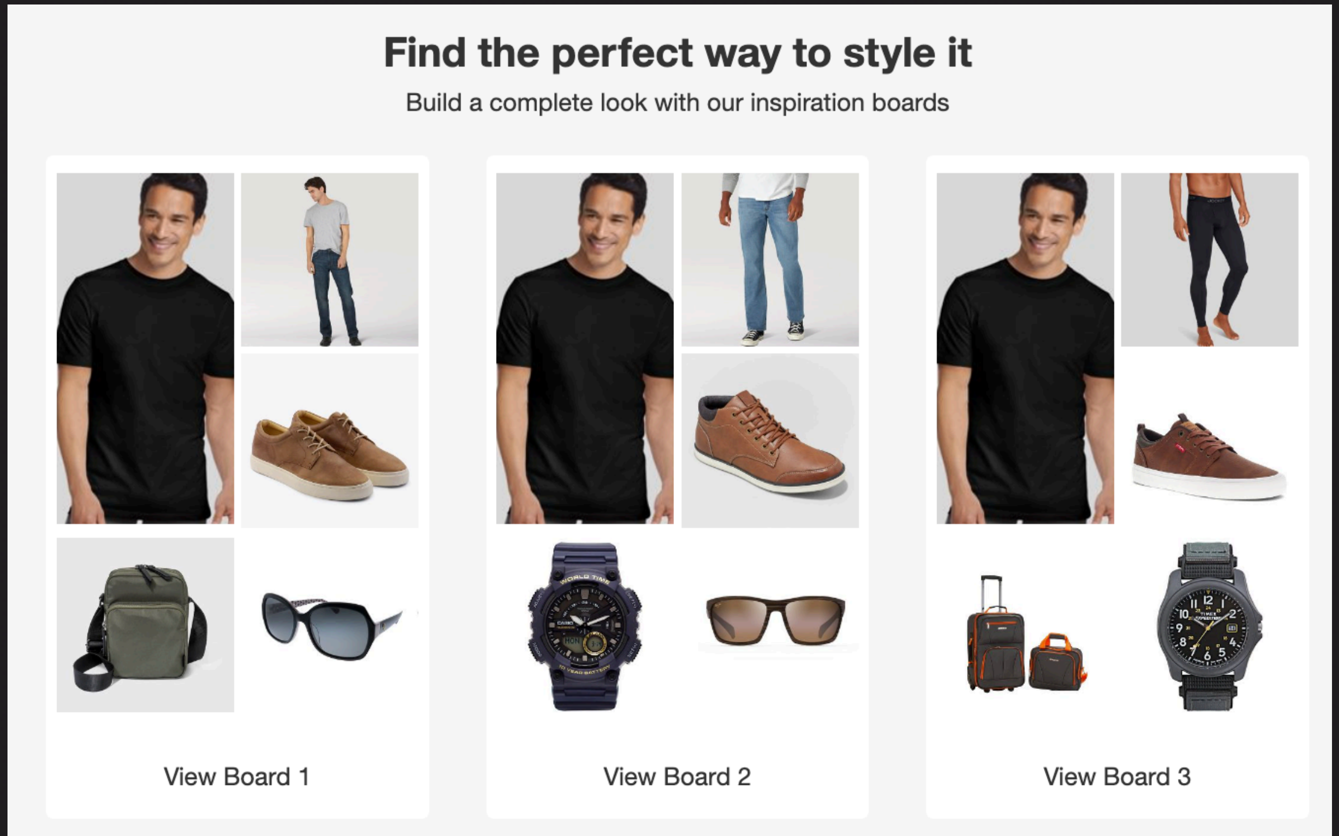
When a customer buys a T-shirt, a recommender algorithm also suggests a few related items

The recommender system in production (legacy) that has a success rate of 10%

You and your team have developed a new deep learning algorithm for recommendation

It is tested before deploying. Of the next 500 customers, 72 bought items recommended by the new model.

Is the improvement brought by the new model is statistically significant at 95% confidence?



Null and alternate hypothesis?

$H_0 : p = 0.1$ $\quad$ $H_a : p > 0.1$

H_0 assumes new model has same performance

This means that the $72/500 = 0.14$ of the new model is just fluke

What is the test statistic?

X : Number of people who bought the recommended items

Distribution of the test statistic X? Binom(n=500, p=0.1)

Right, left tailed, or two-tailed? Right tailed

What is the p-value?

$P[X \geq 72 \,|\, H_0 \text{ is true}] = 1 - P[X \leq 71 \,|\, H_0 \text{ is true}] =$ `1 - binom.cdf(71, n=500, p=0.1)` $= 0.001$

What is α? $\alpha = 0.05$

Is p-value $< \alpha$? Yes We reject the null hypothesis We say that the new model is better