

Plan:

1. Work through a conceptual example of a geospatial analysis

# Spatial Statistics: Example

Shannon E. Ellis, Ph.D  
UC San Diego



Department of Cognitive Science  
[sellis@ucsd.edu](mailto:sellis@ucsd.edu)

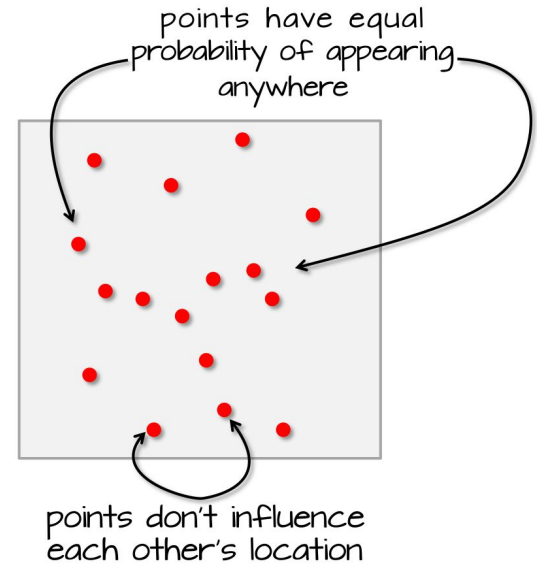
# Hypothesis Testing: CSR/IPR

(Distance-based Methods - how the points are distributed relative to one another)

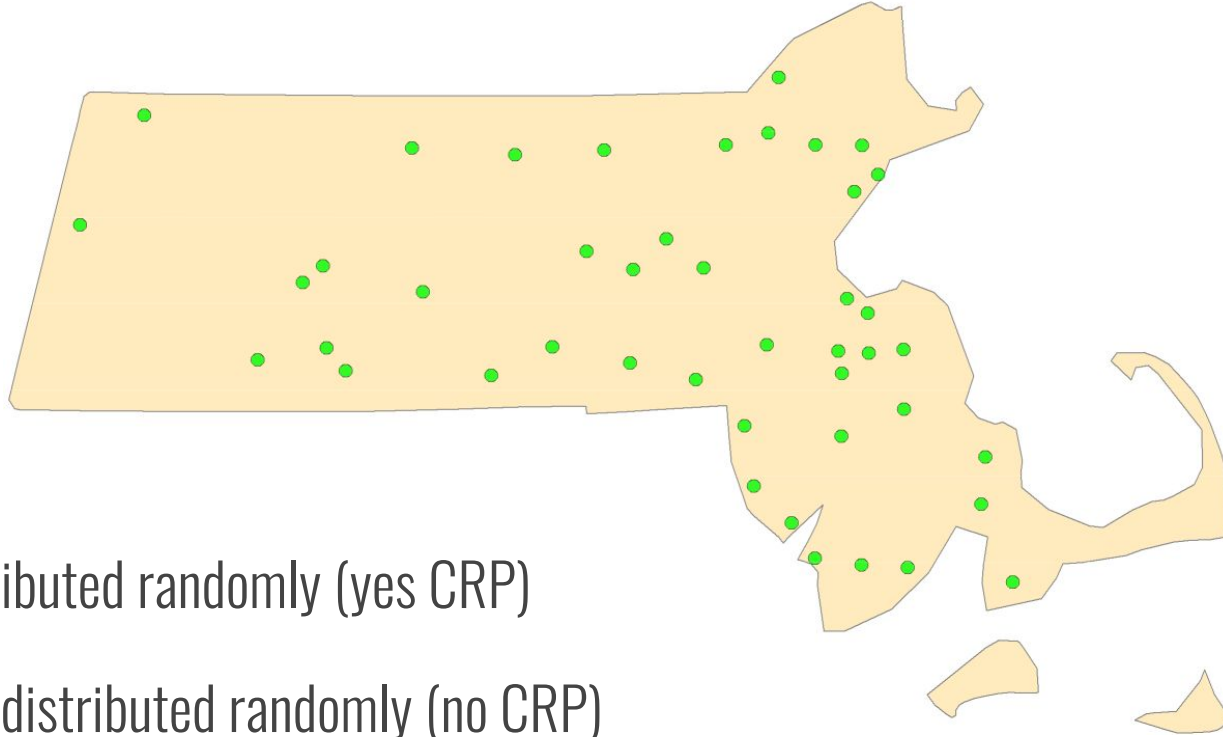
Compare observed point patterns to ones generated by an **independent random process (IRP)**, aka **complete spatial randomness (CSR)**.

CSR/IPR satisfy two conditions:

1. Any event has equal probability of being in any location, a 1st order effect.
2. The location of one event is independent of the location of another event, a 2nd order effect



# Is this distribution of Walmarts in MA the result of a CRP?



$H_0$ : Distributed randomly (yes CRP)

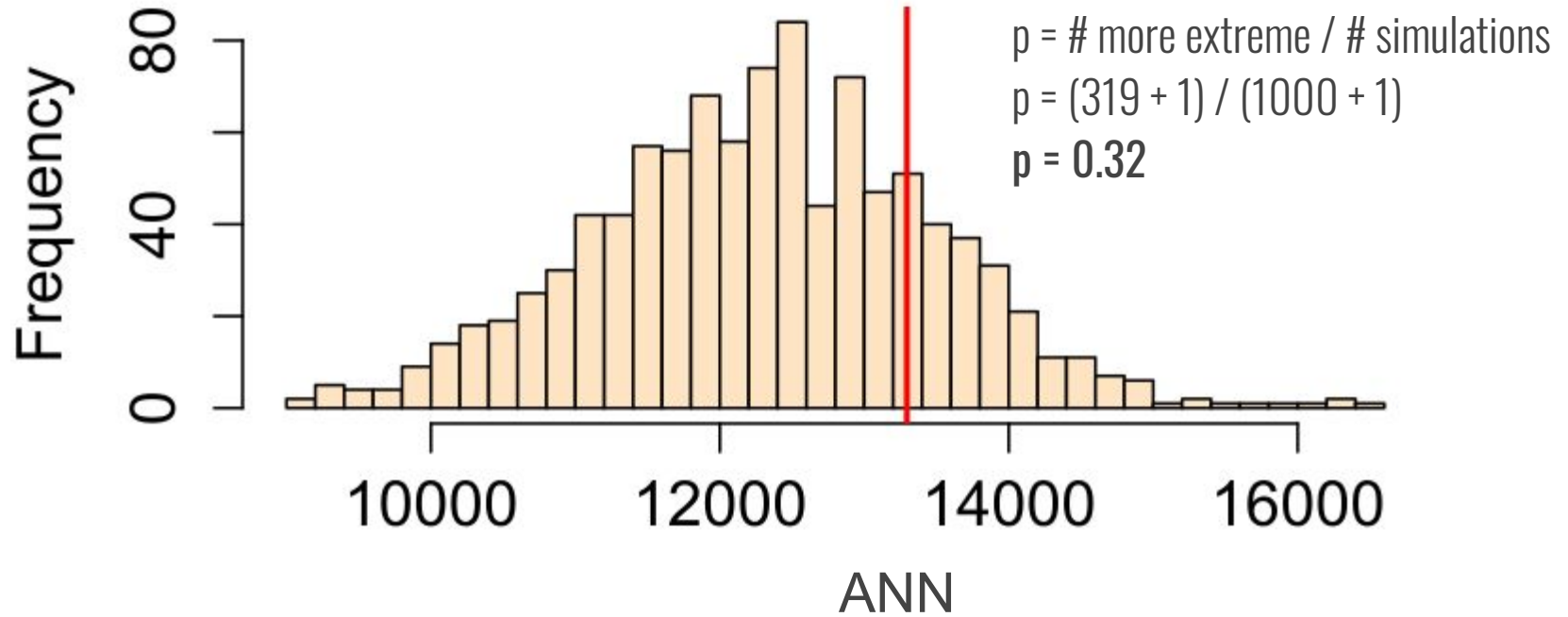
$H_a$ : NOT distributed randomly (no CRP)

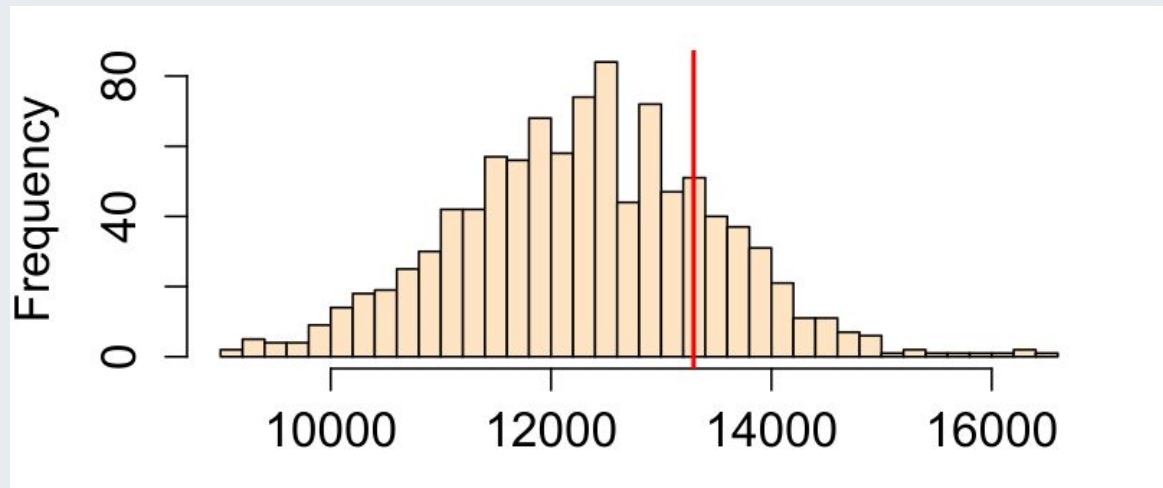
# Hypothesis Testing: A Monte Carlo Test

1. First, we postulate a process—**our null hypothesis,  $H_0$** . For example, we hypothesize that the distribution of Walmart stores is consistent with a completely random process (CRP).
2. Next, we **simulate** many realizations of our postulated process and compute a statistic (e.g. ANN) for each realization.
3. Finally, we **compare our observed data to the patterns generated by our simulated processes** and assess (via a measure of probability) if our pattern is a likely realization of the hypothesized process.



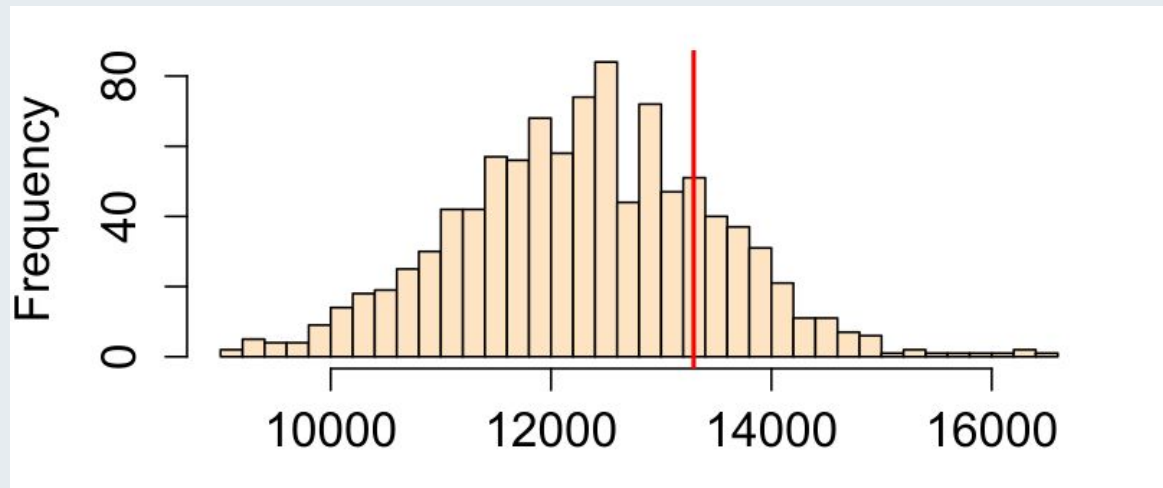
This is an example of bootstrapping!



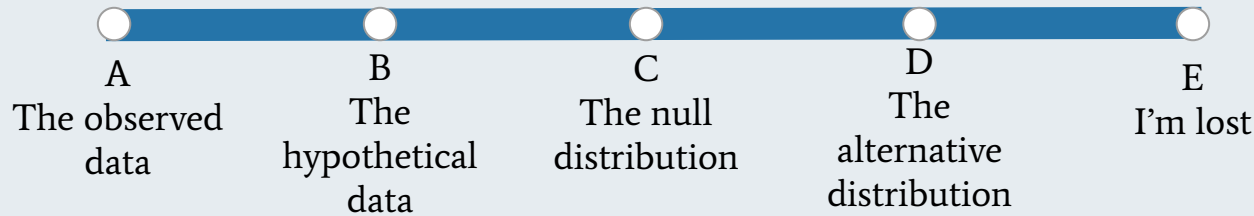


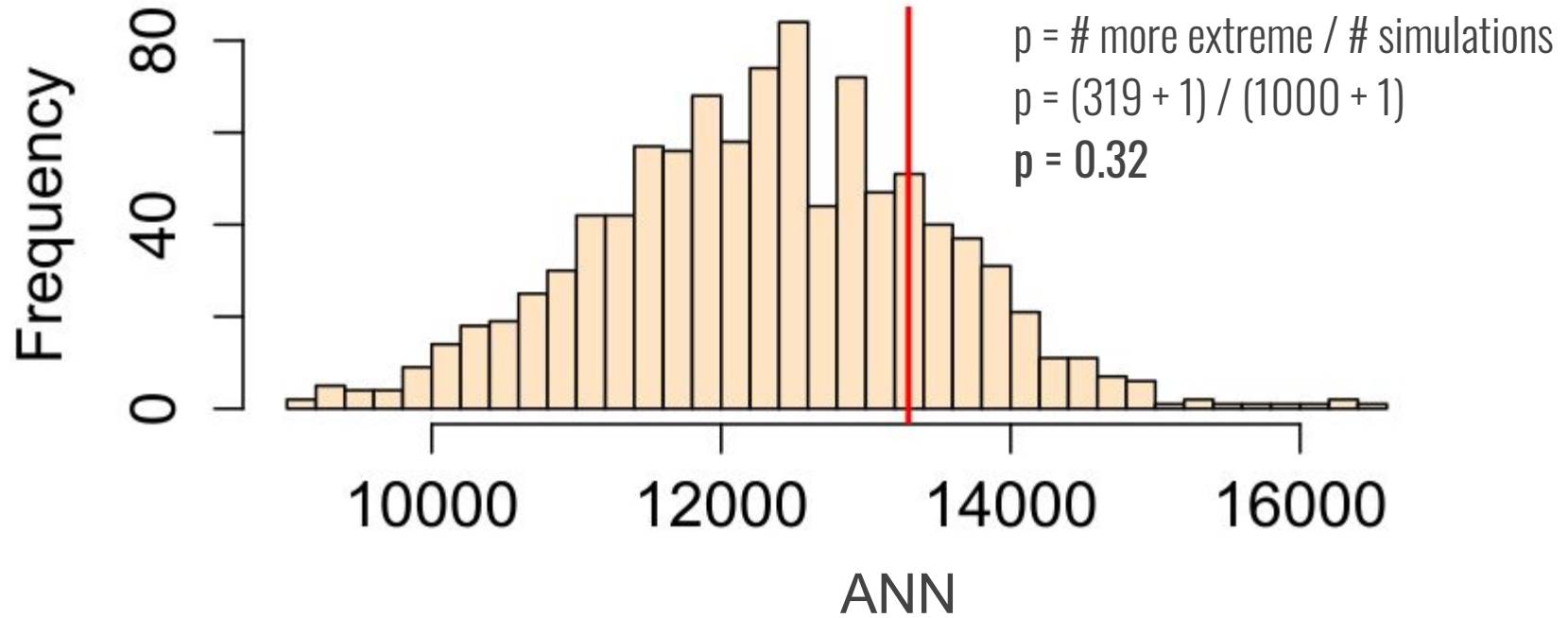
What does the **histogram** represent in this image?

- A B C D E
- Alternative Hypothesis Null Hypothesis Distribution of ANN generated under the null Distribution of ANN generated under the alternative I'm lost



What does the **red line** represent?



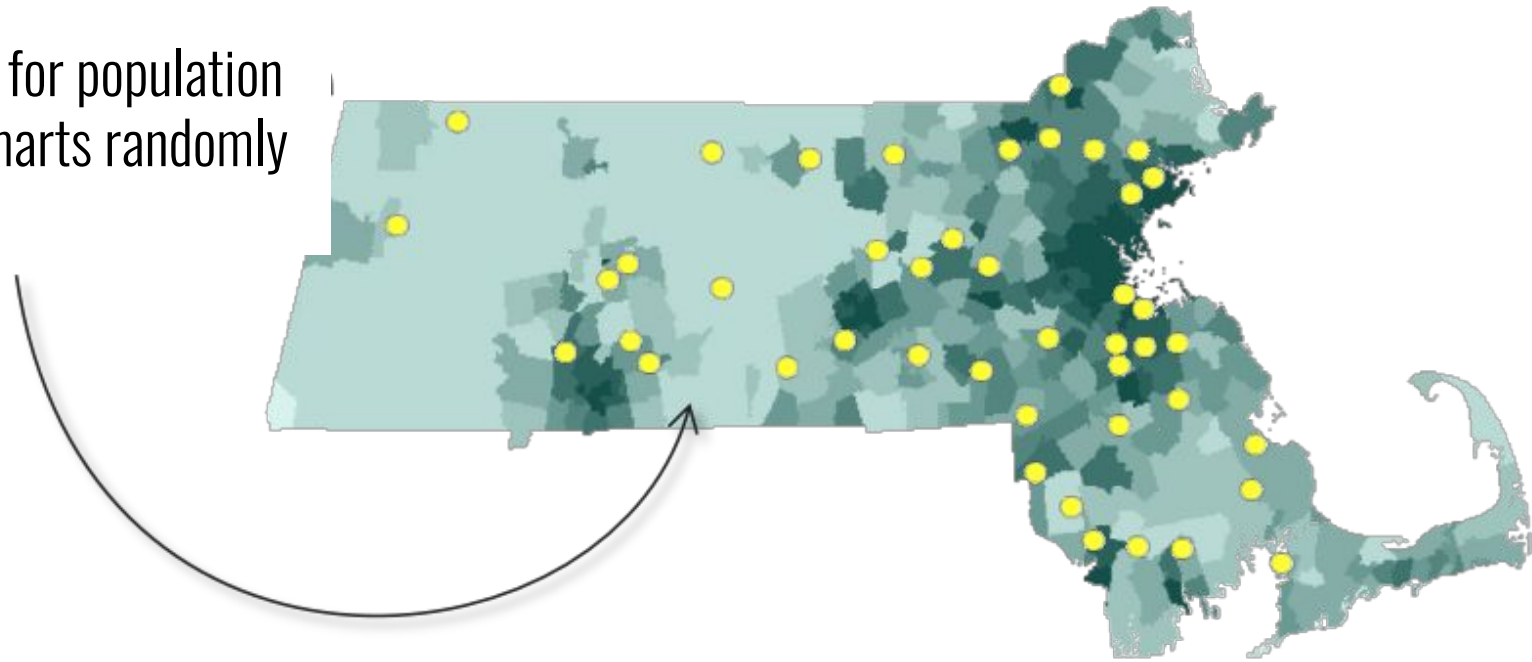


Fail to reject the null

Suggests that our results come from a CRP



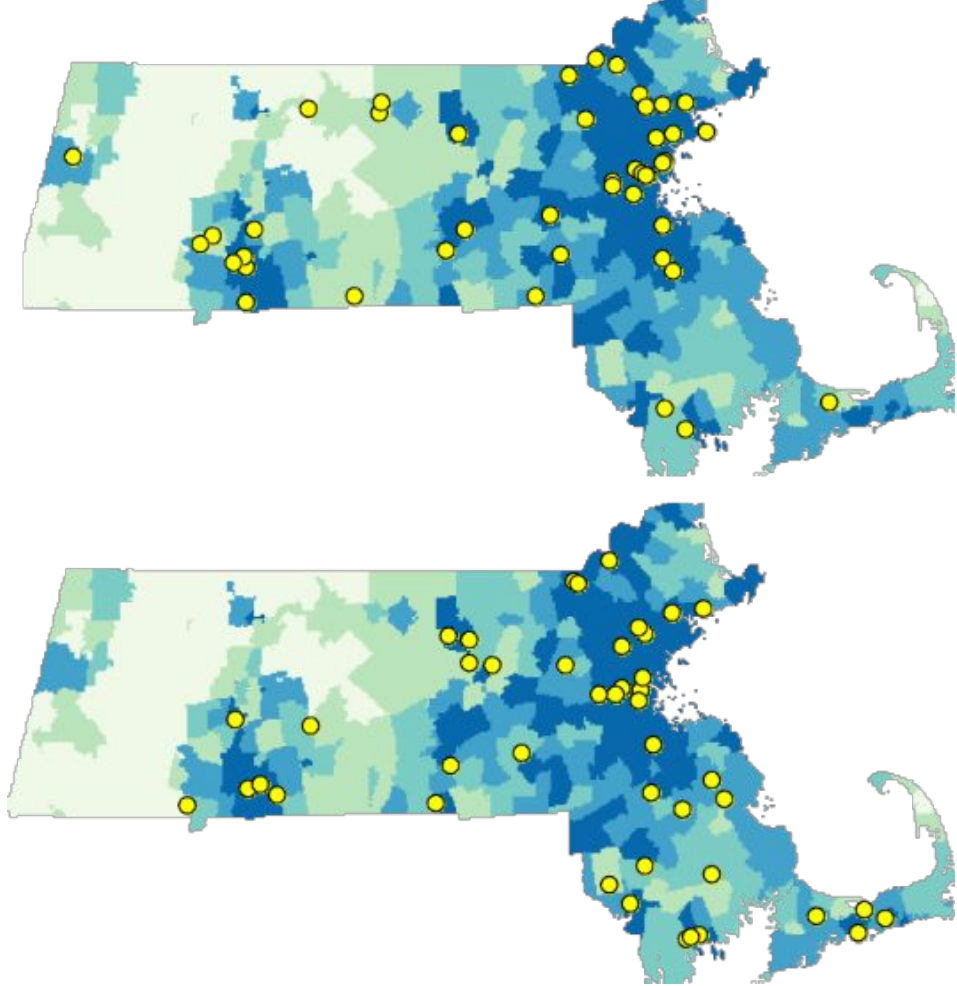
When controlling for population density, are Walmarts randomly distributed?

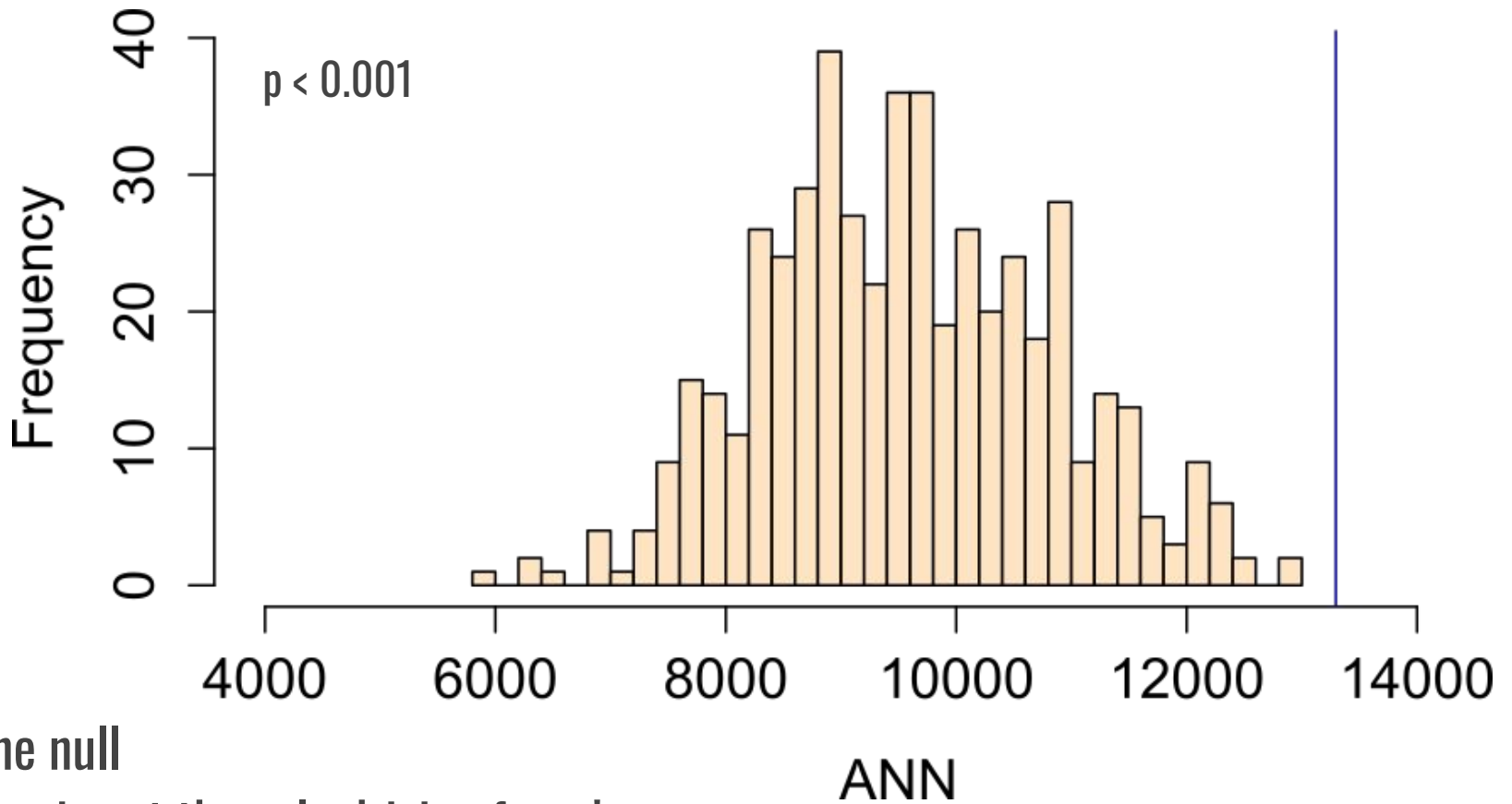


$H_0$ : Walmarts are distributed according to population density alone

$H_a$ : Walmarts are *not* distributed based on population density alone

two randomly generated  
point patterns using  
population density as the  
underlying process

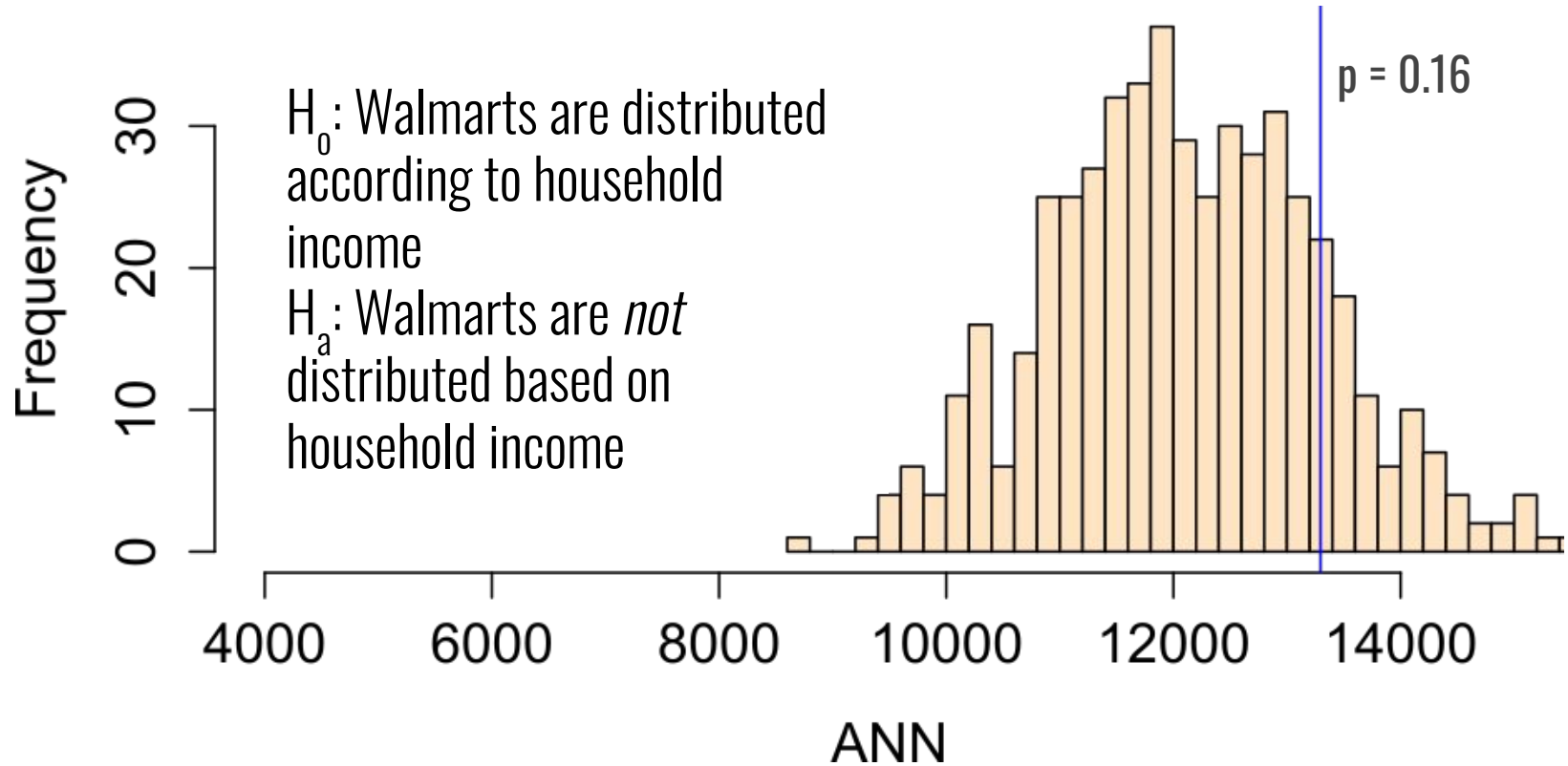




Reject the null

Population is not the sole driving force!

# Maybe median household income is the driving force...?



**...Is it a CRP or median household income?**

hints at plausible scenarios, but doesn't tell us which one it is definitively.

# Basic Geospatial Analysis: Summary

1. Considerations when visualizing spatial data important to conclusions drawn
  - a. values to plot?
  - b. map type?
  - c. color scale?
2. Traditional statistics fail with geospatial data:
  - a. Spatial autocorrelation
  - b. MAUP
  - c. Edge effects
  - d. Ecological fallacy
  - e. Nonuniformity of space
3. Analysis still possible
  - a. Global Point Density, Quadrat Density, Kernel Density
  - b. Poisson Point Process
  - c. K-Nearest Neighbor (KNN)
  - d. Comparison to a CRP (using simulation)