



# **WILL NEW AIRBNB USERS BOOK THEIR FIRST TRIP?**

Brought to you by  
Anne T Griffin

# WHY DOES AIRBNB WANT TO KNOW THIS?

G R O W T H





# WHY DOES AIRBNB WANT TO KNOW THIS?

- Convert new users to active users
- Personalize content to where they are likely to book
  - Reduce time between sign up and first booking

# THE DATA

- AirBnB provided some real data and some test data
  - Destination country dataset
  - Age, gender, and country destination summary set
  - Test user data with user attributes
  - Web sessions data including device type and actions

# NULL HYPOTHESIS

There is **no correlation** between a user's age, gender, and other user info and the destination country they will first book.

# ALTERNATIVE HYPOTHESIS

Using basic info about the user such as age, gender, and other attributes such as device and when and how they signed up, we can predict which country where they are likely to book their first trip.



**THANK YOU**

# APPENDIX



# WHAT'S IN THE DATA?

- test\_users.csv - the test set of users
  - id: user id
  - date\_account\_created: the date of account creation
  - timestamp\_first\_active: timestamp of the first activity, note that it can be earlier than date\_account\_created or date\_first\_booking because a user can search before signing up
  - date\_first\_booking: date of first booking
  - gender
  - age
  - signup\_method
  - signup\_flow: the page a user came to signup up from
  - language: international language preference
  - affiliate\_channel: what kind of paid marketing
  - affiliate\_provider: where the marketing is e.g. google, craigslist, other
  - first\_affiliate\_tracked: whats the first marketing the user interacted with before the signing up
  - signup\_app
  - first\_device\_type
  - first\_browser
  - country\_destination: this is the target variable you are to predict

# WHAT'S IN THE DATA?

- sessions.csv - web sessions log for users
  - user\_id: to be joined with the column 'id' in users table
  - action
  - action\_type
  - action\_detail
  - device\_type
  - secs\_elapsed
- countries.csv - summary statistics of destination countries in this dataset and their locations
- age\_gender\_bkts.csv - summary statistics of users' age group, gender, country of destination