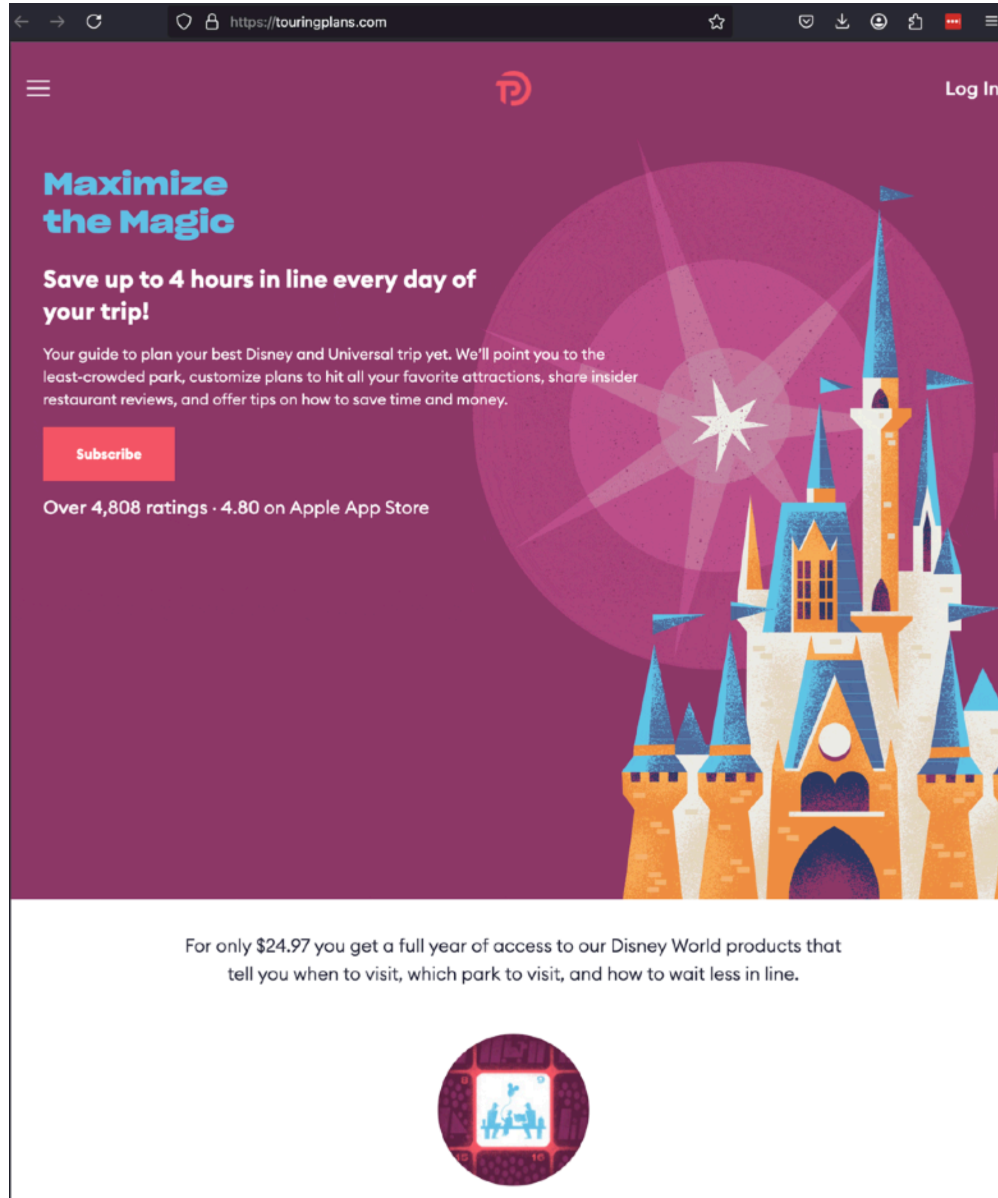
Causal inference with regression

Today's case study

EXTRA MAGIC HOURS

- Some days, **one** of the Disney World **theme parks** would open for an **extra morning hour**...
- ...but only for guests staying at the Disney Resort hotels
- What does this do to ride wait times?



*From Barrett et al., 2025, **Causal Inference in R**.*
Introduction to the datasets at https://www.r-causal.org/chapters/07-prep-data

Tour of the `touringplans` dataset

`alien_saucers`	Alien Swirling Saucers Wait Times
`attractions`	Attraction meta data
`attractions_metadata`	Metadata about attractions included in datasets
`dinosaur`	DINOSAUR Wait Times
`expedition_everest`	Expedition Everest Wait Times
`flight_of_passage`	Avatar Flight of Passage Wait Times
`kilimanjaro_safaris`	Kilimanjaro Safaris Wait Times
`navi_river`	Na'vi River Journey Wait Times
`parks_metadata_raw`	Meta data for Disney Touring Plan Park Data
`pirates_of_caribbean`	Pirates of the Caribbean Wait Times
`rock_n_rollercoaster`	Rock n Rollercoaster Wait Times
`seven_dwarfs_train`	Seven Dwarfs Mine Train Wait Times
`seven_dwarfs_train_2018`	Seven Dwarfs Mine Train wait times in 2018
`slinky_dog`	Slinky Dog Dash Wait Times
`soarin`	Soarin' Wait Times
`spaceship_earth`	Spaceship Earth Wait Times
`splash_mountain`	Splash Mountain Wait Times
`touringplans_2018`	Reduced data set for Disney World attraction wait times in 2018
`toy_story_mania`	Toy Story Mania! Wait Times

The theme parks

```
> xtabs(~park,attractions)
park
      California Adventure
                        14
    Disney's Animal Kingdom
                         8
Disney's Hollywood Studios
                        12
                Disneyland
                        25
                     Epcot
                        11
             Magic Kingdom
                        25
```

Embark on a Magical Adventure You'll Remember Forever

Visit all 4 theme parks at Walt Disney World Resort––
filled with classic attractions, exciting events and
enchanting entertainment!

Share a Family Vacation Filled with Can't-Miss Experiences at All 4 Parks!

Disney's Hollywood Studios
Venture to a place where legendary *Star Wars* stories come to life—and put you in the middle of the action!

Disney's Animal Kingdom Theme Park
Celebrate the magic of nature with some favorite Disney pals, animal adventures and lively entertainment.

Magic Kingdom Park
Race through the diamond mine from *Snow White and the Seven Dwarfs* on a swaying family coaster ride.

EPCOT
Discover worlds of wonder around the globe and set sail to Arendelle on *Frozen Ever After.*

Reminder: three criteria for identifiability

Reminder: three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$

Reminder: three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$
  - Consequence: different individuals' outcomes don't affect each other

Reminder: three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$
  - Consequence: different individuals' outcomes don't affect each other
  - Consequence: there can be no "multiple versions" of the same treatment A in terms of their influence on Y

Reminder: three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$
  - Consequence: different individuals' outcomes don't affect each other
  - Consequence: there can be no "multiple versions" of the same treatment A in terms of their influence on Y
- **Conditional Exchangeability**: for all i, $Y^{a=i} \perp A \mid Z$ for some set of observed variables Z

Reminder: three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$
  - Consequence: different individuals' outcomes don't affect each other
  - Consequence: there can be no "multiple versions" of the same treatment A in terms of their influence on Y
- **Conditional Exchangeability**: for all i, $Y^{a=i} \perp A \,|\, Z$ for some set of observed variables Z
  - Consequence: there can be no "hidden common causes" or "hidden mediators" of A and $Y^{a=i}$

Reminder: three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$
  - Consequence: different individuals' outcomes don't affect each other
  - Consequence: there can be no "multiple versions" of the same treatment A in terms of their influence on Y
- **Conditional Exchangeability**: for all i, $Y^{a=i} \perp A \,|\, Z$ for some set of observed variables Z
  - Consequence: there can be no "hidden common causes" or "hidden mediators" of A and $Y^{a=i}$
- **Positivity**: for all i and all values of Z, $P(A = i \,|\, Z) > 0$

Reminder: three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$
  - Consequence: different individuals' outcomes don't affect each other
  - Consequence: there can be no "multiple versions" of the same treatment A in terms of their influence on Y
- **Conditional Exchangeability**: for all i, $Y^{a=i} \perp A \,|\, Z$ for some set of observed variables Z
  - Consequence: there can be no "hidden common causes" or "hidden mediators" of A and $Y^{a=i}$
- **Positivity**: for all i and all values of Z, $P(A = i \,|\, Z) > 0$
  - e.g., in our example, it can't be the case that individuals with heart disease are *always* given transplants

Reminder: three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$
  - Consequence: different individuals' outcomes don't affect each other
  - Consequence: there can be no "multiple versions" of the same treatment A in terms of their influence on Y
- **Conditional Exchangeability**: for all i, $Y^{a=i} \perp A \,|\, Z$ for some set of observed variables Z
  - Consequence: there can be no "hidden common causes" or "hidden mediators" of A and $Y^{a=i}$
- **Positivity**: for all i and all values of Z, $P(A = i \,|\, Z) > 0$
  - e.g., in our example, it can't be the case that individuals with heart disease are *always* given transplants
- If all three criteria hold, we can estimate causal effects