Bad Evidence

Paul Verbeek-Mast

All A/B tests and data shown in this presentation are not based on real experiments. They are made up just for this presentation.

Bad Evidence

Base

база

Base

база

Variant

вариант

Base

база

Variant

вариант

Base
база

Variant
вариант

Base
база

Variant
вариант

Base
база

Variant
вариант

Base
база

Variant
вариант

5234 searches

Base
база

Variant
вариант

5234 searches

Base

база

6252 searches

Variant

вариант

5234 searches

Base
база

6252 searches   +19.45%

Variant
вариант

5234 searches

Base
база

6252 searches   +19.45%

Variant
вариант

Найти предложения   Бад-Гаштайн   Брюссель   Париж   Лондон   Берлин   Амстердам   Антверпен   Барселона   Маастрихт   Рим   Мадрид

🔍 Найдите лучшие предложения
Выгодные цены на отели и апартаменты

Место, название или адрес варианта размещения

Можно искать по аэропорту, месту или виду отдыха

Приезжаю

📅 Дата заезда ⌄

Уезжаю

📅 Дата отъезда ⌄
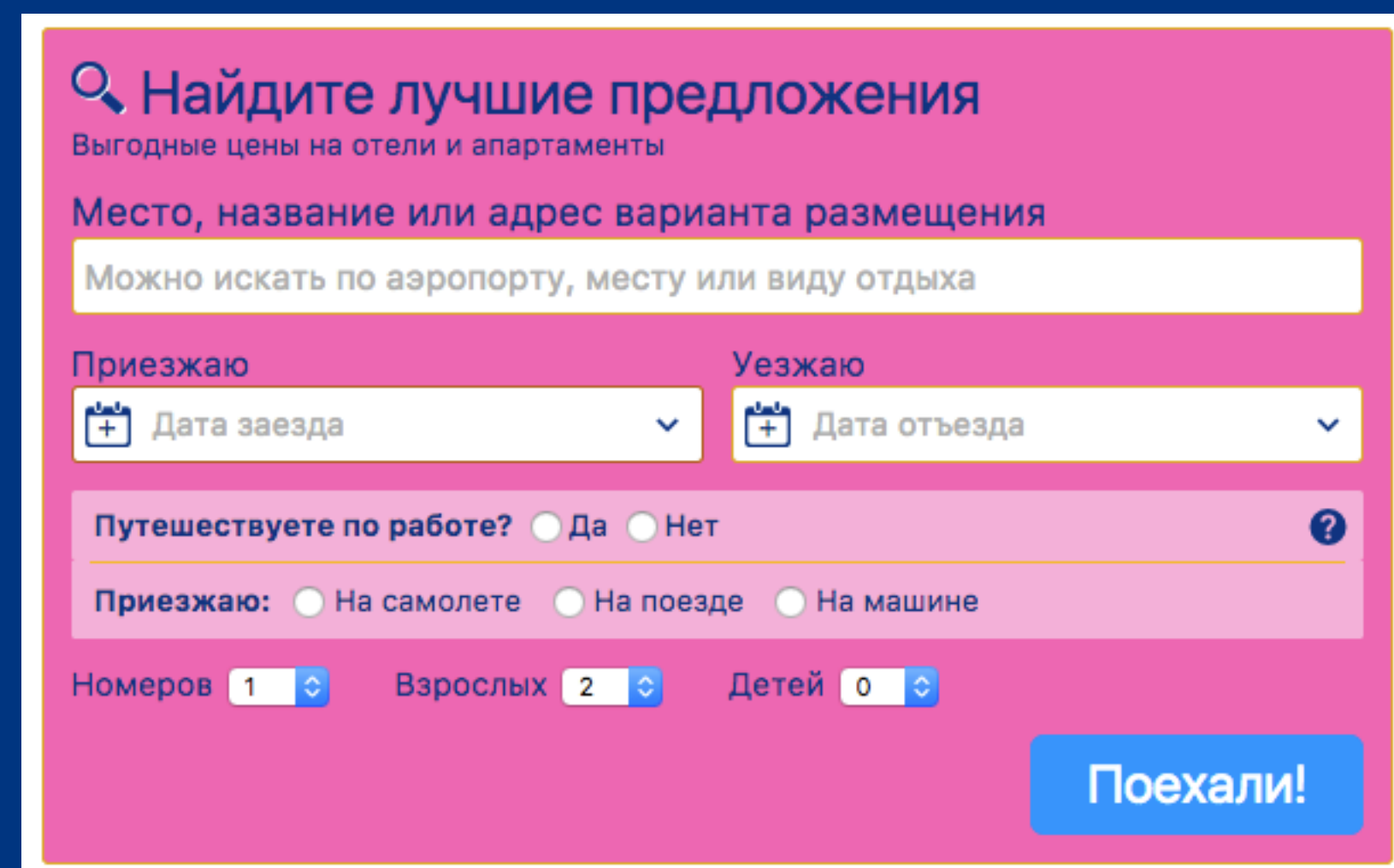
Путешествуете по работе? ○ Да ○ Нет   ❓

Приезжаю: ○ На самолете ○ На поезде ○ На машине

Номеров [1]   Взрослых [2]   Детей [0]

Поехали!

👤 Получите скидку до 50% на Секретные предложения! **Войти**
Цены будут ниже, как только вы войдете в аккаунт

Горящие предложения на любой бюджет!
Бронирование без комиссии • Регулярные специальные предложения • _Мы возвращаем разницу в цене_

1 168 501 вариант размещения во всем мире
Включая 599 698 вариантов аренды для отпуска по 107 180 направлениям в 227 странах

Вносите изменения в бронирование за пару шагов!
Нажмите здесь, чтобы изменить даты поездки, запросить парковку или отменить бронирование.

115 090 000 отзывов, которым можно доверять
Настоящие гости. Реальные бронирования. _Благодаря проверенному мнению вы сможете принять правильное решение._

Служба поддержки: круглосуточно и без выходных
Получайте быстрые ответы или _свяжитесь с нами здесь_

Все еще интересны эти места? Цены могут вырасти...

Бад-Гаштайн
17 марта - 18 марта
1 номер, 2 взрослых
📱 Последний поиск был на мобильном телефоне

Бад-Гаштайн
19 марта - 20 марта
1 номер, 2 взрослых
📱 Последний поиск был на мобильном телефоне

Чаттануга
18 апр. - 19 апр.
1 номер, 2 взрослых

Поделитесь этой ссылкой с друзьями и заработайте денег!
Пригласите друзей и заработайте!

💼 Управление деловыми поездками — бесплатно и без хлопот
Обновить мой аккаунт

Все еще хотите поехать сюда? Торопитесь, чтобы не упустить лучшую цену!

Бад-Гаштайн 🇦🇹
Главные причины для поездки: горные лыжи, горы, пейзажи

Найдите предложения на выходные в Бадгаштейне на двоих 👥

Эти выходные
17 марта – 19 марта

Следующие выходные
24 марта – 26 марта

Через две недели
31 марта – 2 апр.

Бад-Гаштайн: отели, апартаменты и многое другое »

Base
база

Variant
вариант

Making the search box hotpink will result in more searches

Base
база

Variant
вариант

Making the search box hotpink will result in more searches

Making the search box hotpink will result in more searches

6252 searches   +19.45%

Making the search box hotpink will result in more searches

6252 searches   +19.45%

242 bookings   -4.7%

Making the search box hotpink will result in more searches

Making the search box hotpink will result in more searches

SERIAL

SERIAL

SERIAL

How much do you want to create
"Bad Evidence"?

Насколько вы готовы получить
доказательство обратного ?
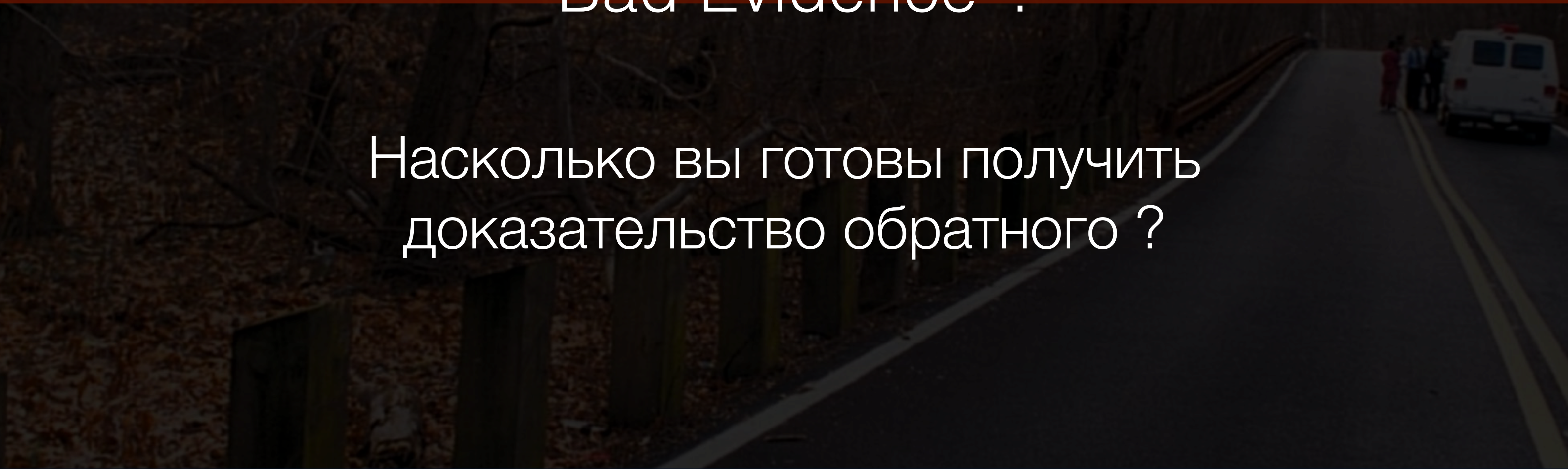
SERIAL

You don't want to do something if it is going to go against your theory of the case.

Вы не хотите делать что-то что повредит вашей теории
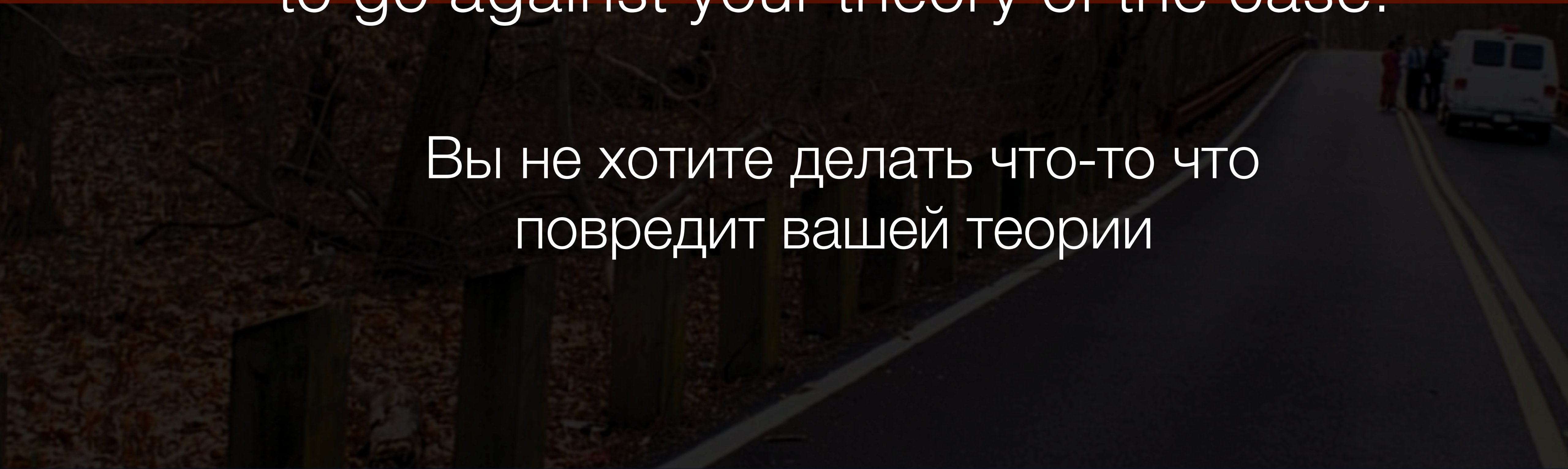
Rather than trying to get to the truth, what you're trying to do is build your case, and make it the strongest case possible.

Вместо того чтобы докопаться до истины вы пытаетесь защитить свою версию, сделав ее доказательство "пуленепробиваемым".
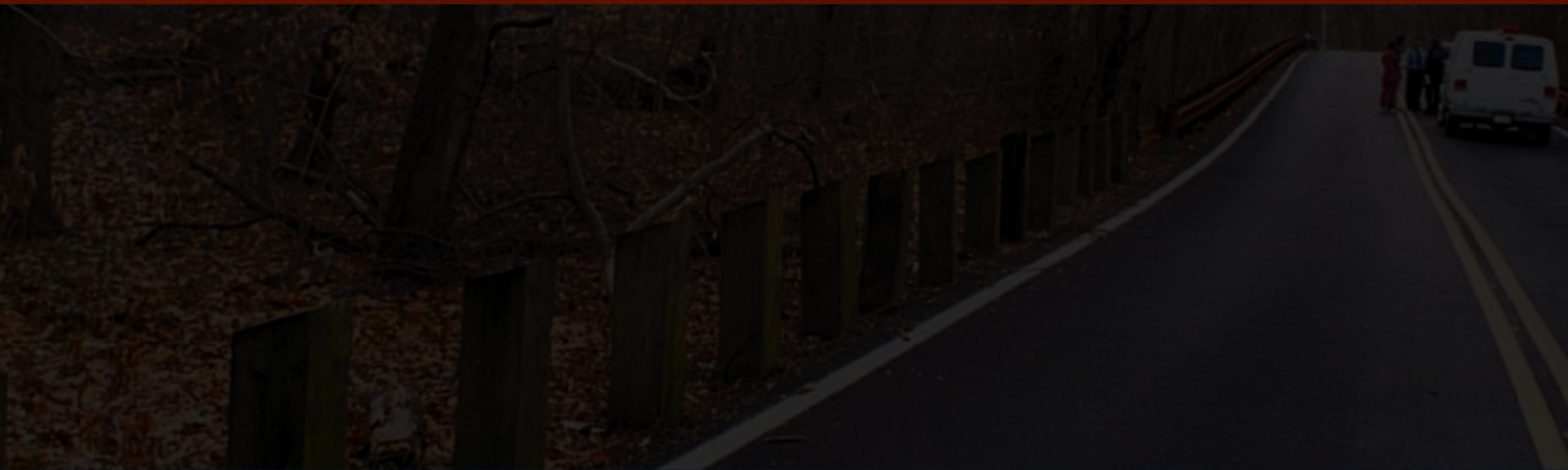
SERIAL

What does verification bias cause you to do? Ignore it and push it to the side.

Что вы будете делать со своей предвзятостью ? Просто игнорируйте ее.

SERIAL
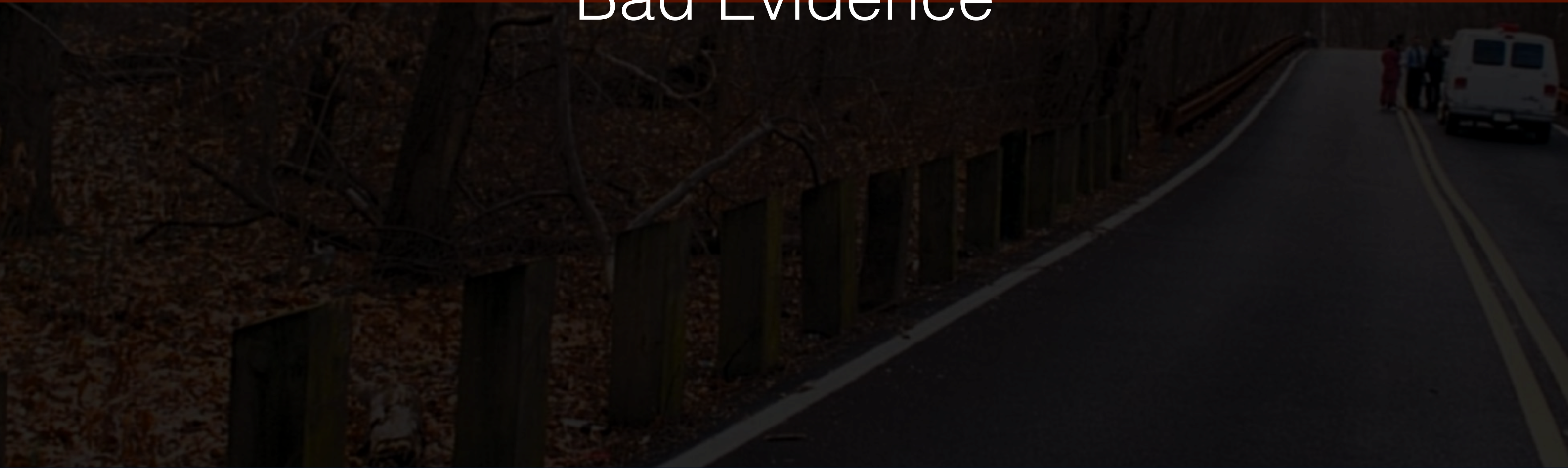
SERIAL

Verification bias
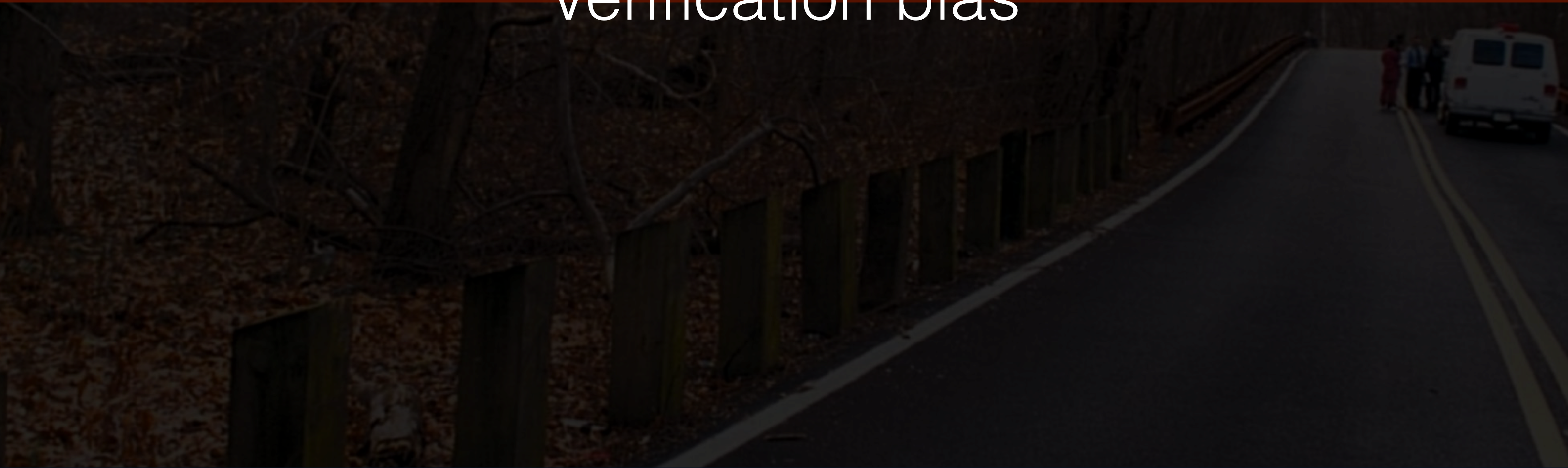
Because of **(why)** we believe that changing **(what)** for **(who)** will result into **(outcome)**

Why

Objective and based on data

Because of **(why)** we believe

Why
Objective and based on data

Bad examples

- Because of a gut feeling, we believe (…)

- Because I like it better, we believe (…)

- Because I saw it on another website, we believe (…)

Because of **(why)** we believe

Why
Objective and based on data

Good examples

- Because of research described in article (…), we believe (…)
- After done user research, we believe (…)
- Based on a previous experiment doing (…), we believe (…)

Because of **(why)** we believe

What

An accurate, short description of your change

Because of **(why)** we believe that changing **(what)**

What

An accurate, short description of your change

Bad examples

- changing it to pink
- changing the title that is on the top of the first block on the home page to 16px Arial #FF0000

Because of **(why)** we believe that changing **(what)**

What

An accurate, short description of your change

Good examples

- changing the background of search box to pink
- opening pictures in the search page in a lightbox when clicking on it

Because of **(why)** we believe that changing **(what)**

Who

A realistic, accurate description of your target group

Because of **(why)** we believe that changing **(what)** for **(who)**

Who

A realistic, accurate description of your target group

Bad examples

- everyone
- users booking a hotel in Novosibirsk, named Paul, from Amsterdam, with a big beard

Because of **(why)** we believe that changing **(what)** for **(who)**

Who

A realistic, accurate description of your target group

Good examples

- users visiting the home page
- users searching for a property in Novosibirsk
- users who are logged in

Because of **(why)** we believe that changing **(what)** for **(who)**

Outcome
measurable, expected changes

Because of **(why)** we believe that changing **(what)** for **(who)** will result into **(outcome)**

Outcome
measurable, expected changes

Bad examples

- users feeling better
- the site looking prettier
- an increase in loyalty

Because of **(why)** we believe that changing **(what)** for **(who)** will result into **(outcome)**

Outcome
measurable, expected changes

Good examples

- an increase in earnings

- a decrease in returned products

- an increase in sign-ups

Because of **(why)** we believe that changing **(what)** for **(who)** will result into **(outcome)**

Because of **(why)** we believe that changing **(what)** for **(who)** will result into **(outcome)**

Because of **user research** we believe that changing **(what)** for **(who)** will result into **(outcome)**

Because of **user research** we believe that changing **the background of the search box to pink** for **(who)** will result into **(outcome)**

Because of **user research** we believe that changing **the background of the search box to pink** for **users that visit the homepage** will result into **(outcome)**

Because of **user research** we believe that changing **the background of the search box to pink** for **users that visit the homepage** will result into **an increase in bookings**

Because of **user research** we believe that changing **the background of the search box to pink** for **users that visit the homepage** will result into **an increase in bookings**

Because of **user research** we believe that changing **the background of the search box to pink** for **users that visit the homepage** will result into **an increase in bookings**

You can **never** be 100% confident
that your test is correct

The more you measure, the higher the chance some things are incorrect

users going to search results

language changes

calls to customer service

logins

hover over button

bookings

returning visitors

bookings from IE8

clicks on button

bookings from Malaysia

buys with credit card

visits on page

scrolled to button

clicks on logo

number of rooms booked

price of booking

sign ups

time on page

users going to search results

language changes

calls to customer service

logins

hover over button

bookings

returning visitors

bookings from IE8

clicks on button

bookings from Malaysia

buys with credit card

visits on page

scrolled to button

clicks on logo

number of rooms booked

price of booking

sign ups

time on page

users going to search results

language changes

calls to customer service

logins

hover over button

bookings

returning visitors

bookings from IE8

clicks on button

bookings from Malaysia

buys with credit card

visits on page

scrolled to button

clicks on logo

number of rooms booked

price of booking

sign ups

time on page

users going to search results

language changes

calls to customer service

logins

hover over button

bookings

returning visitors

bookings from IE8

clicks on button

bookings from Malaysia

buys with credit card

visits on page

scrolled to button

clicks on logo

number of rooms booked

price of booking

sign ups

time on page

users going to search results	language changes
calls to customer service	logins
hover over button	bookings from IE8
returning visitors	bookings
clicks on button	buys with credit card
bookings from Malaysia	scrolled to button
visits on page	clicks on logo
number of rooms booked	price of booking
time on page	sign ups

users going to search results	+0.1%	language changes	-2.1%
calls to customer service	-0.2%	logins	+0.3%
hover over button	+2.3%	bookings from IE8	+2.1%
returning visitors	+0.3%	bookings	-1.8%
clicks on button	+4.7%	buys with credit card	-0.3%
bookings from Malaysia	-3.1%	scrolled to button	+0.0%
visits on page	+0.0%	clicks on logo	+0.5%
number of rooms booked	+3.5%	price of booking	+4.3%
time on page	-1.1%	sign ups	-0.2%

users going to search results	+0.1%	language changes	-2.1%
calls to customer service	-0.2%	logins	+0.3%
hover over button	+2.3%	bookings from IE8	+2.1%
returning visitors	+0.3%	bookings	-1.8%
clicks on button	+4.7%	buys with credit card	-0.3%
bookings from Malaysia	-3.1%	scrolled to button	+0.0%
visits on page	+0.0%	clicks on logo	+0.5%
number of rooms booked	+3.5%	price of booking	+4.3%
time on page	-1.1%	sign ups	-0.2%

users going to search results	+0.1%	language changes	-2.1%
calls to customer service	-0.2%	logins	+0.3%
hover over button	+2.3%	bookings from IE8	+2.1%
returning visitors	+0.3%	bookings	-1.8%
clicks on button	+4.7%	buys with credit card	-0.3%
bookings from Malaysia	-3.1%	scrolled to button	+0.0%
visits on page	+0.0%	clicks on logo	+0.5%
number of rooms booked	+3.5%	price of booking	+4.3%
time on page	-1.1%	sign ups	-0.2%

users going to search results	+0.1%	language changes	-2.1%
calls to customer service	-0.2%	logins	+0.3%
hover over button	+2.3%	bookings from IE8	+2.1%
returning visitors	+0.3%	bookings	-1.8%
clicks on button	+4.7%	buys with credit card	-0.3%
bookings from Malaysia	-3.1%	scrolled to button	+0.0%
visits on page	+0.0%	clicks on logo	+0.5%
number of rooms booked	+3.5%	price of booking	+4.3%
time on page	-1.1%	sign ups	-0.2%

Focus on your **defined metrics**, but also keep an eye on your **health metrics**

Be **honest** with yourself

Metrics that are not in hypothesis

Metrics that are not in hypothesis

"price is going up, so it must be doing well"

vs.

"price is going down, so it must be a false negative"

Newly implemented metrics

Newly implemented metrics

"this new metric is positive, it's working great!"

vs.

"this new metric is negative, must be having a bug"

Sample size

Sample size

"it's positive after 5 days, let's put it in production"

vs.

"it's negative after 5 days, let's run it for another few days"

How long should your run your A/B test?

- Number of **visitors**
- How **big of a change** you want to measure
- How **confident** you want to be, that your test is correct

Create a **solid hypothesis**, and stick to it

Make your decision based on **data**

There is no such thing as **bad evidence**, just a **bad hypothesis**

So stop building cases, and **find the truth**

Paul Verbeek-Mast

🐦 **@_paulverbeek**

✉️ **verbeek.p@gmail.com**

Questions?

Вопросы?