
DRAMA

About Helge

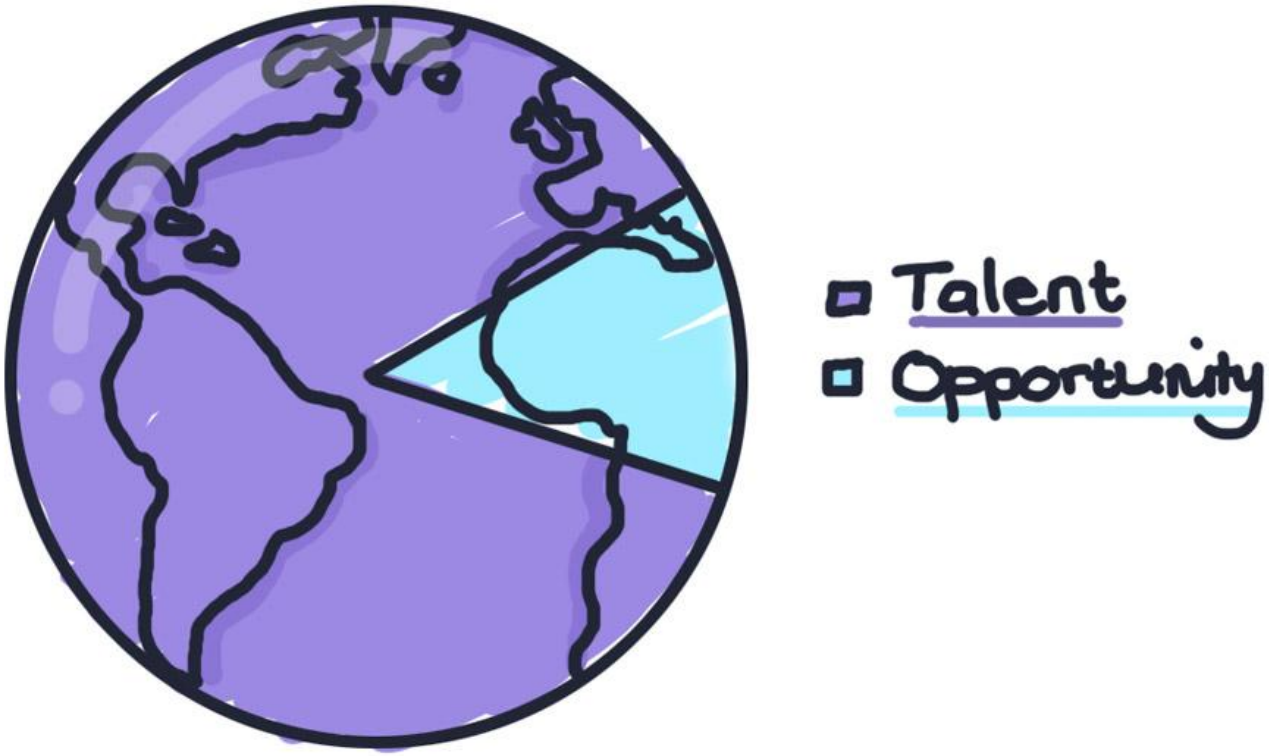
Data Scientist

OFFER ZEN

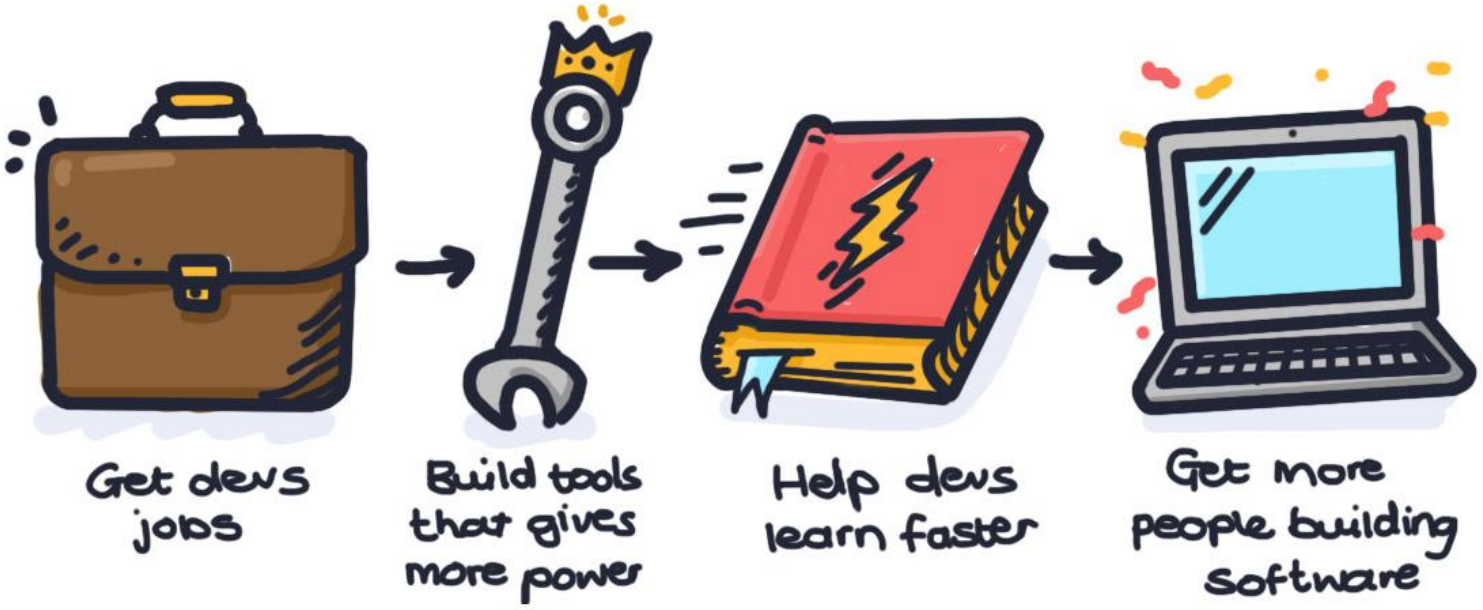
UNIVERSITEIT STELLENBOSCH UNIVERSITY

Applied Mathematics

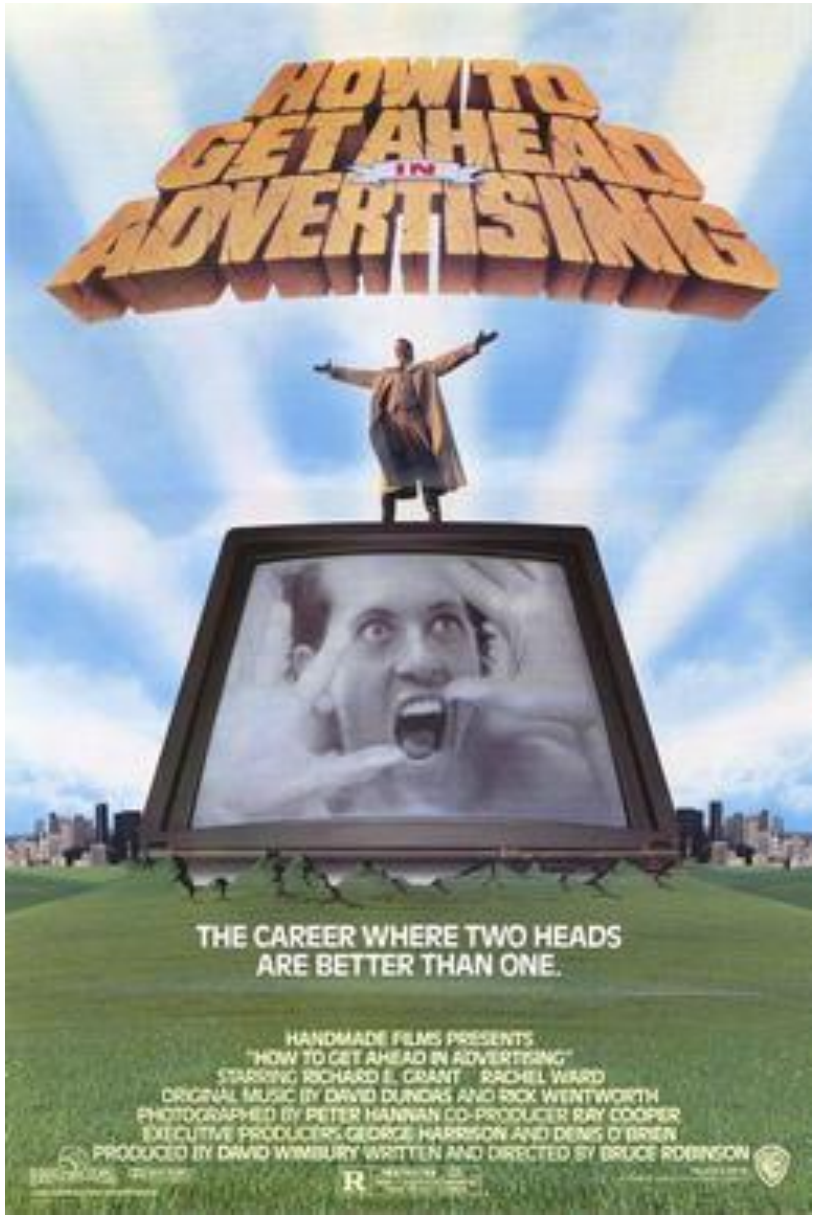
OfferZen's mission



The master plan



OfferZen **Root** **Source** **Make**


HOW TO GET AHEAD IN ADVERTISING
THE CAREER WHERE TWO HEADS ARE BETTER THAN ONE.
HANDMADE FILMS PRESENTS
"HOW TO GET AHEAD IN ADVERTISING"
STARRING RICHARD E. GRANT RACHEL WARD
ORIGINAL MUSIC BY DAVID DUNDAS AND RICK WENTWORTH
PHOTOGRAPHED BY PETER HANNAN CO-PRODUCER RAY COOPER
EXECUTIVE PRODUCERS GEORGE HARRISON AND DENIS O'BRIEN
PRODUCED BY DAVID WIMBURY WRITTEN AND DIRECTED BY BRUCE ROBINSON
R

In this talk

- **What is the marketing attribution problem?**

In this talk

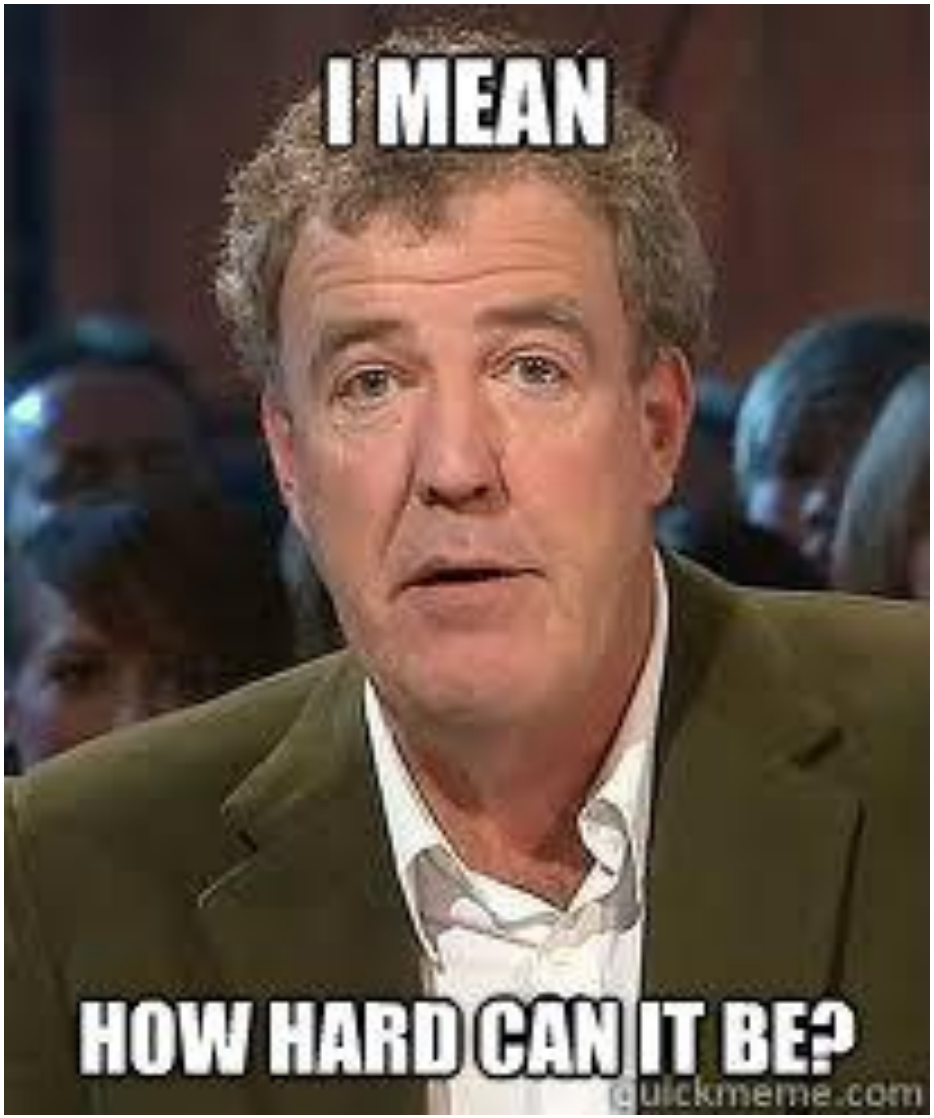
- What is the marketing attribution problem?
- **Models and methods**

In this talk

- What is the marketing attribution problem?
- Models and methods
- **Practical lessons learned**

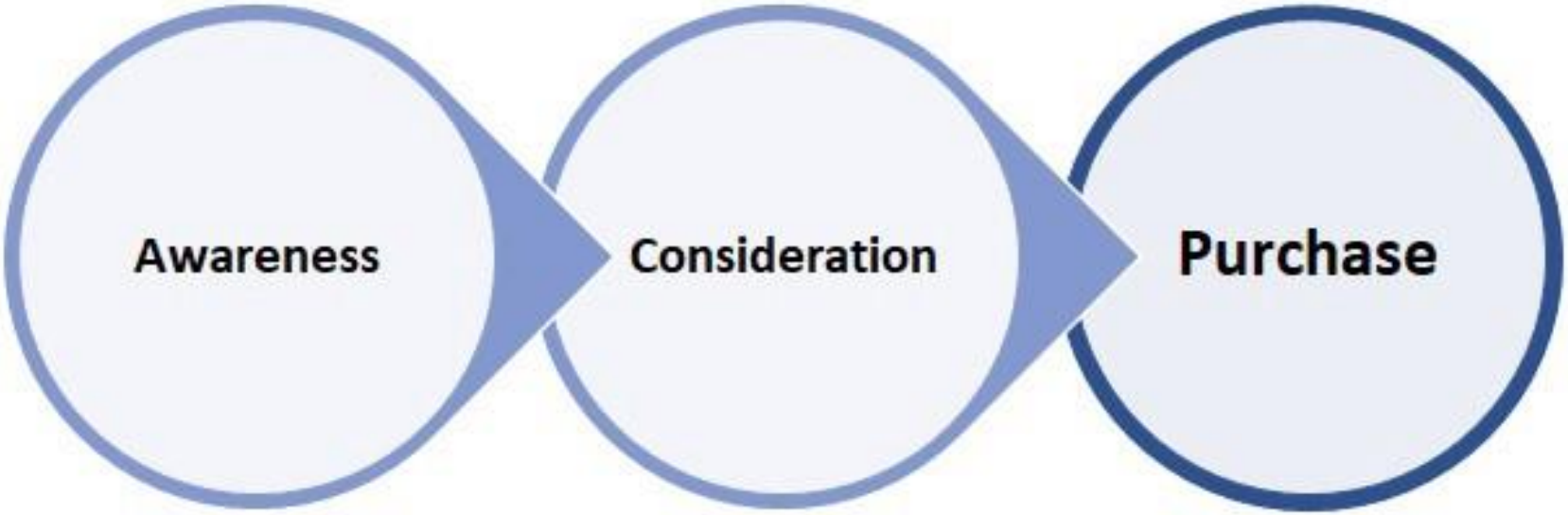
What is the marketing channel attribution problem?

Marketing channel attribution

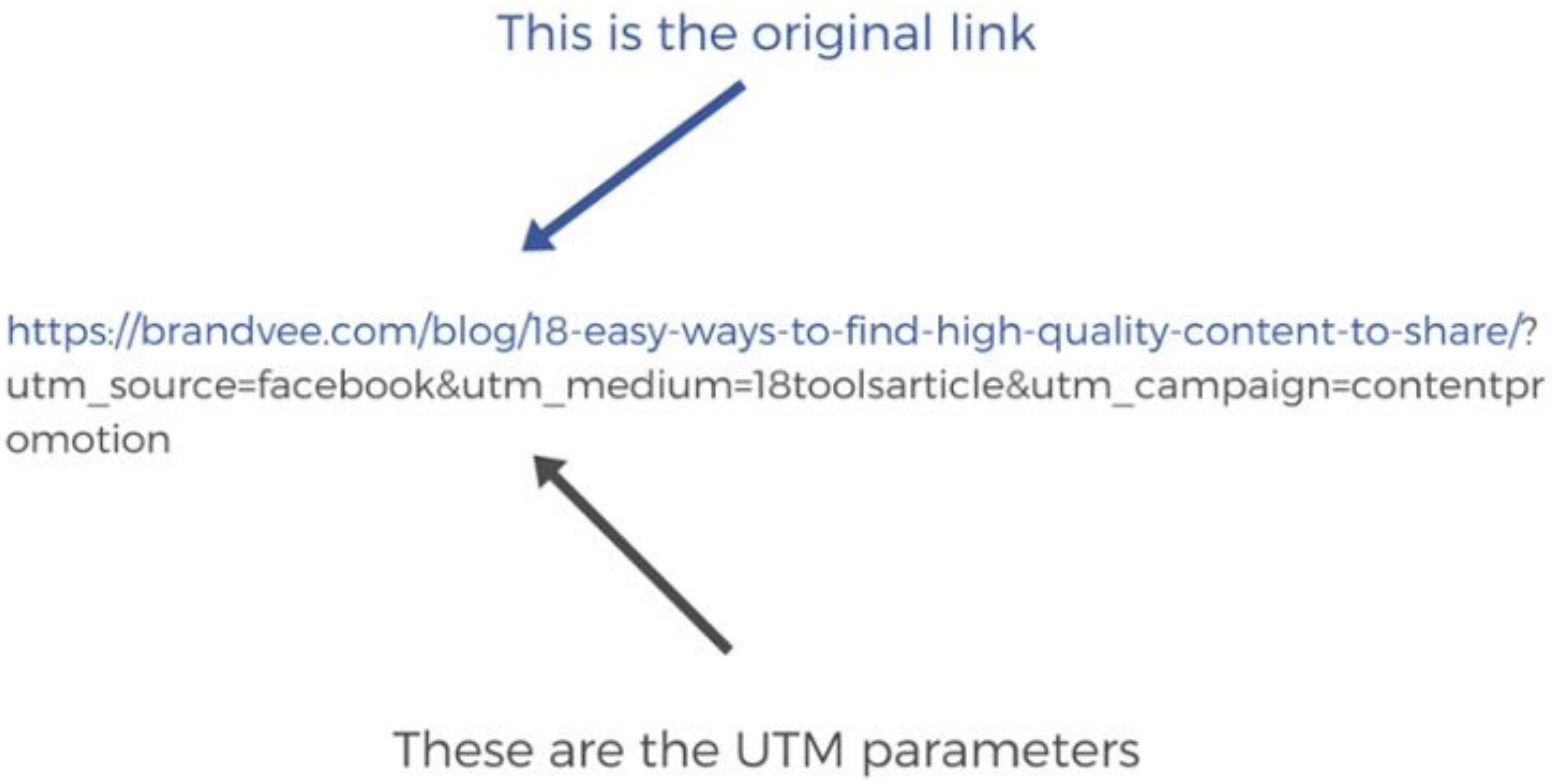


Marketing channel attribution

Marketing channel attribution

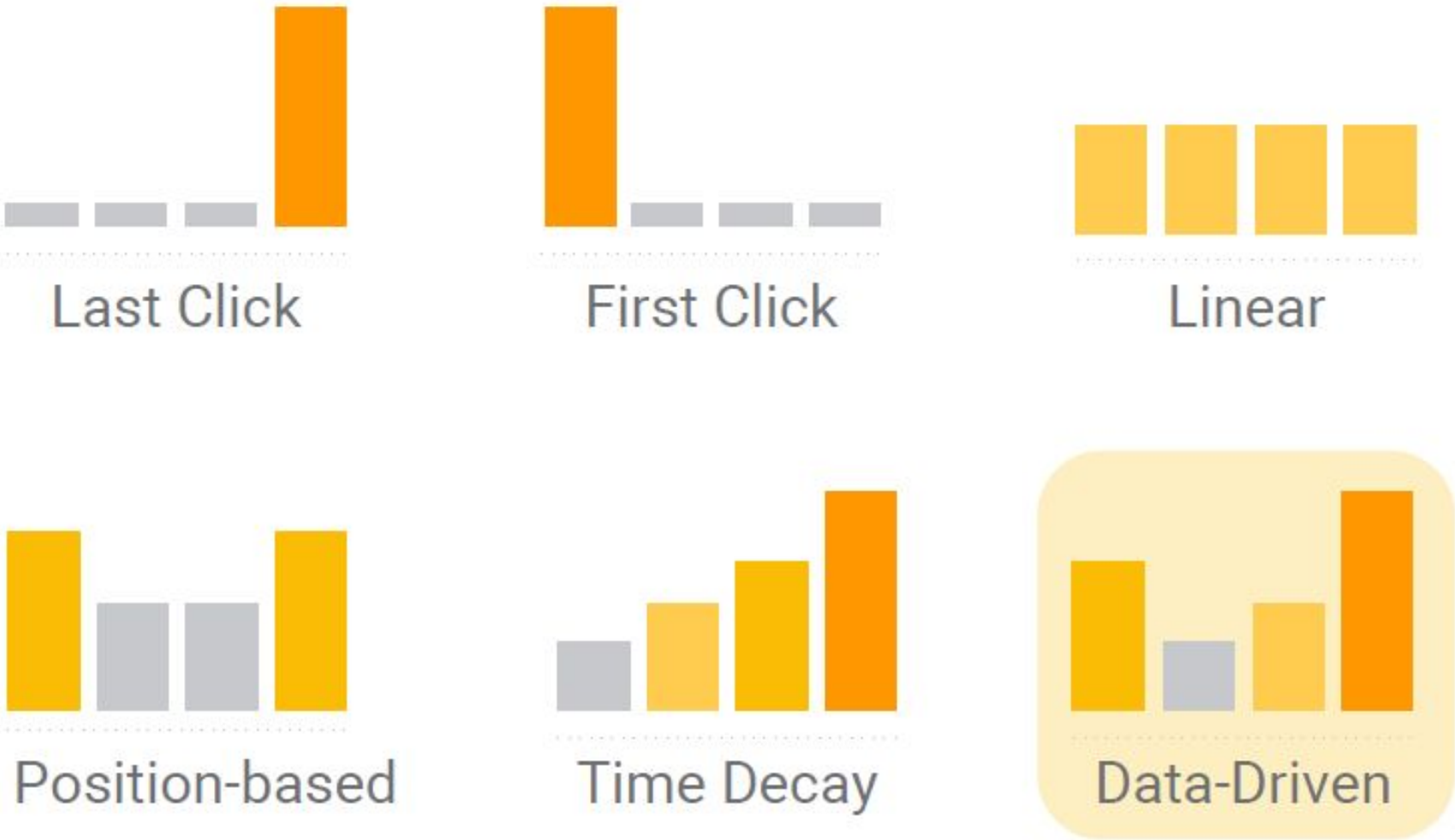


Marketing channel attribution



Models and methods

Models and methods



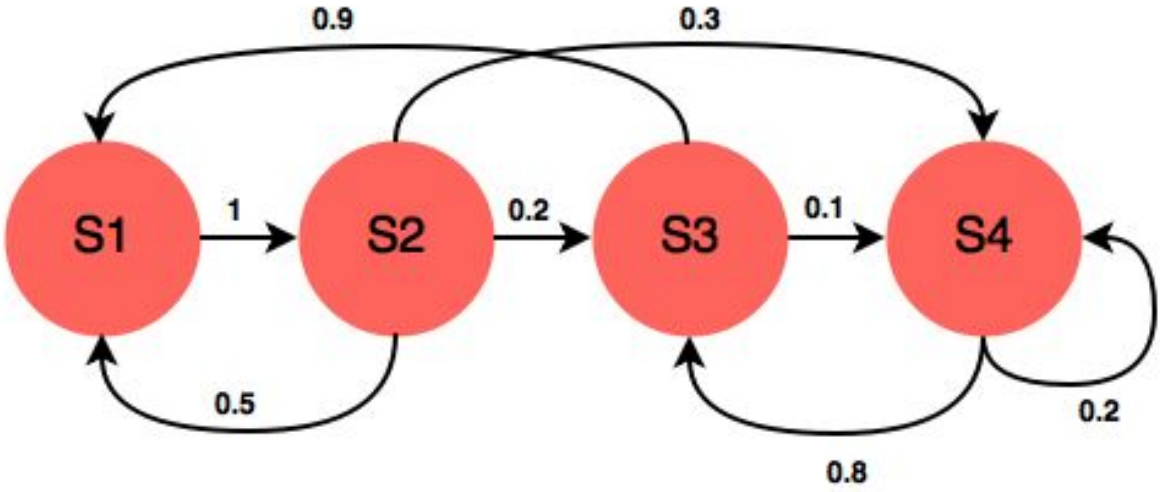
Models and methods

"all models are wrong,
but some are useful"

George Box
1919 –2013

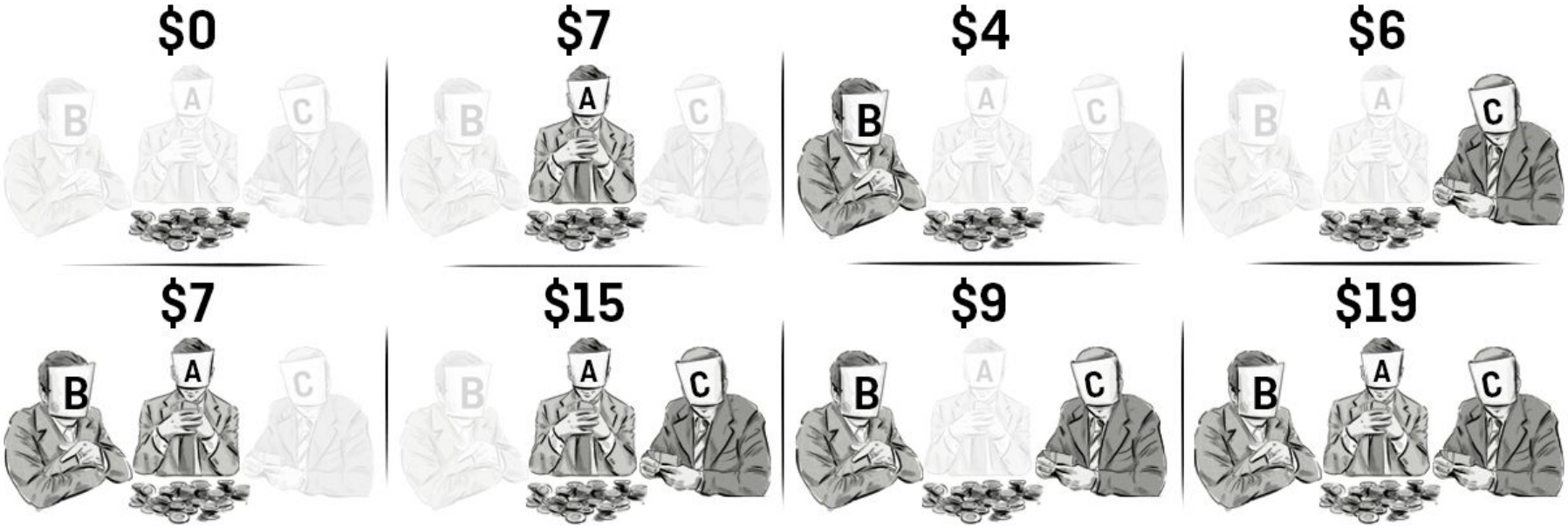
Markov Chain Attribution

Markov Chain Attribution



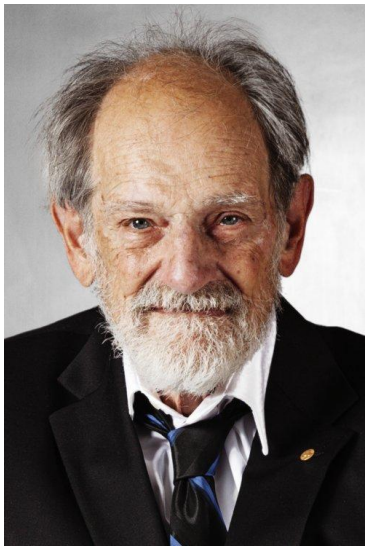
Game Theory Attribution

Game Theory Attribution



Game Theory Attribution

Shapley Values



$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(n - |S| - 1)!}{n!} (v(S \cup \{i\}) - v(S))$$

Markov Chain Attribution

In [1]:

```
library(ChannelAttribution)
library(reshape)
library(ggplot2)
```

Toy data

In [2]:

```
df1 <- data.frame(sequence = c('c1 > c2 > c3', 'c1', 'c2 > c3'), conv = c(1, 0, 0), no_conv =
c(0, 1, 1))
df1
```

sequence	conv	no_conv
c1 > c2 > c3	1	0
c1	0	1
c2 > c3	0	1

Define the model

In [3]:

```
model1 <- markov_model(df1,
                       var_path = 'sequence',
                       var_conv = 'conv',
                       var_null = 'no_conv',
                       out_more = TRUE)
```

Get the results

In [4]:

```
df_result_1 <- model1$result
df_result_1
```

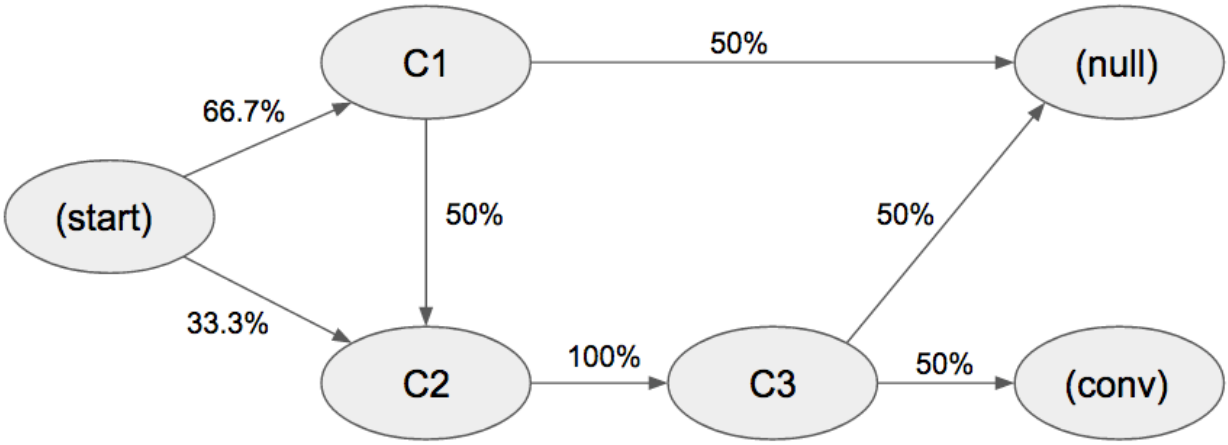
channel_name	total_conversions
c1	0.2002886
c2	0.3998557
c3	0.3998557

Removel effects

In [5]:

```
removel_effects = model1$removal_effects
removel_effects
```

channel_name	removal_effects
c1	0.5009023
c2	1.0000000
c3	1.0000000

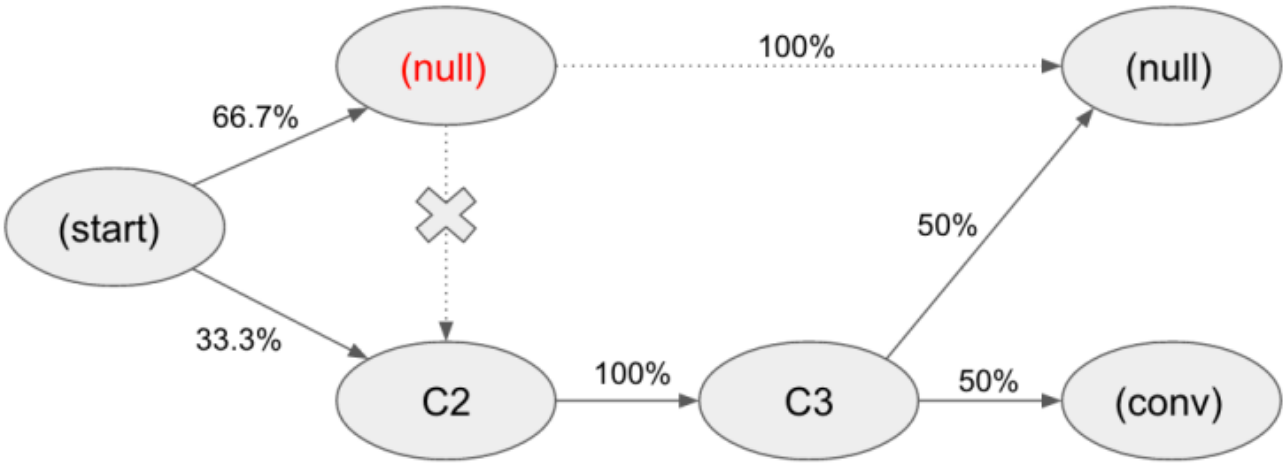


In [6]:

```
# 1. Probability of conversion

0.667 * 0.5 * 1 * 0.5 + 0.333 * 1 * 0.5
```

0.33325



In [7]:

```
# 2. Probability of conversion

0.333 * 1 * 0.5
```

0.1665

In [8]:

```
# Removal effect

1 - 0.167/0.333
```

0.498498498498498

In [9]:

```
removel_effects = removel_effects$removal_effects
data.frame(removel_effects/sum(removel_effects))
```

removel_effects.sum.removel_effects.
0.2002886
0.3998557
0.3998557

Game Theory Attribution

In [10]:

```
library('GameTheoryAllocation')
```

```
Loading required package: e1071
Loading required package: lpSolveAPI
```

Coalitions

In [11]:

```
df_B1 = data.frame(coalitions(3)$Binary)
names(df_B1) <- c('c1', 'c2', 'c3')
df_B1
```

c1	c2	c3
0	0	0
1	0	0
0	1	0
0	0	1
1	1	0
1	0	1
0	1	1
1	1	1

In [12]:

```
2**3
```

8

Characteristic function

In [13]:

```
characteristic_function <- c(0,7,4,6,7,15,9,19)
```

In [14]:

```
df_B1$conversions <- characteristic_function
df_B1
```

c1	c2	c3	conversions
0	0	0	0
1	0	0	7
0	1	0	4
0	0	1	6
1	1	0	7
1	0	1	15
0	1	1	9
1	1	1	19

Shapely values

In [15]:

```
shapley_values <- Shapley_value(characteristic_function, game="profit")
```

```
[1] "Shapley Value"
```

In [16]:

```
#shapley_values = data.frame(shapley_values)

data.frame("channel_name" = c('c1', 'c2', 'c3'), "shapley_values" = c(shapley_values))
```

channel_name	shapley_values
c1	7.666667
c2	3.166667
c3	8.166667

Marginal values

c1-c3-c2 -> 7 + 8 + 4

c1-c2-c3 -> 7 + 0 + 12

c2-c3-c1 -> 4 + 5 + 10

c2-c1-c3 -> 4 + 3 + 12

c3-c1-c2 -> 6 + 9 + 4

c3-c2-c1 -> 6 + 3 + 10

In [17]:

```
factorial(3)
```

6

Shapley values

In [18]:

```
# c1
(7 + 7 + 10 + 3 + 9 + 10)/6
```

7.66666666666667

In [19]:

```
# c2
(4 + 0 + 4 + 4 + 4 + 3)/6
```

3.16666666666667

In [20]:

```
# c3
(8 + 12 + 5 + 12 + 6 + 6)/6
```

8.16666666666667

In [21]:

```
shapley_values
```

	1	2	3
	7.666667	3.166667	8.166667

In [22]:

```
shapley_values/sum(shapley_values)
```

	1	2	3
	0.4035088	0.1666667	0.4298246

Practical lessons learned

- **Marketing attribution is an important and hard problem**

Practical lessons learned

- Marketing attribution is an important and challenging problem.
- **No best model or method (might not matter too much)**

Practical lessons learned

- Marketing attribution is an important and challenging problem.
- No best model or method (might not matter too much)
- **Good data quality + simple model -> often sufficient**

Conclusion

“The career where two heads are better than one.”

Questions?