
Alice has a Blue Car:
Beginning the Conversation Around Ethically Aware Decision Making
BY ThoughtWorks®
Simon Aubury- Principal Data Engineer

Product - we’re an insurance co

Alice wants some insurance

What should Alice’s premiums be?

A machine can do this

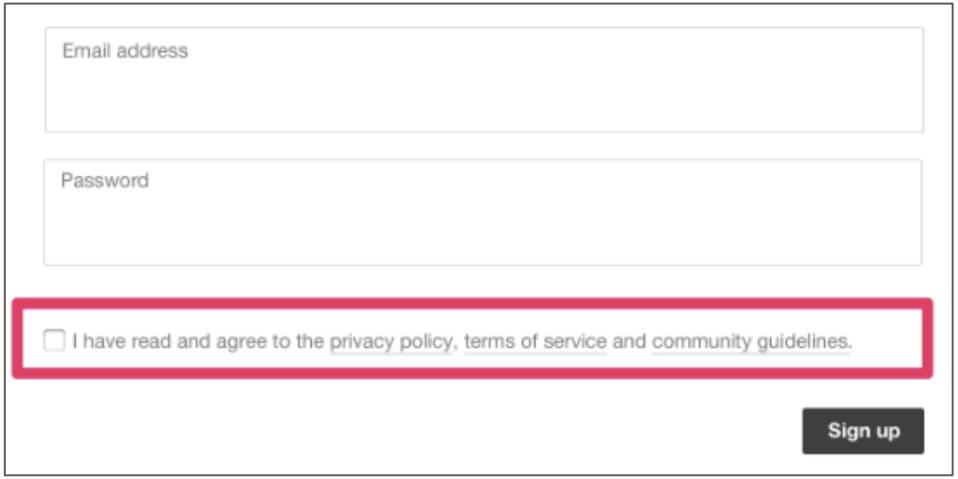
Premium ($) = Risk (%) X Value ($)

- Features
- Algorithms

UX

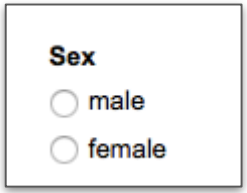
What are our assumptions?

- Consent - meaningful consent
  - Gathering
  - Storage
  - Aligned to 3rd parties
- TOS - did you read it?



UX - simple example

What are our assumptions?



UX - simple example

What are our assumptions?

UX - protect the individual

Randomized Response Technique

Table 1. Example of randomized response technique (RRT) application by using Warner's method.

Probability	Assignment by Rolling a Dice [1]	Question	Answer
2/3	1–4 → Sensitive Question:	I have an intention to dispose of waste illegally.	☐ Yes ☐ No
1/3	5–6 → Inverse Question:	I do not have any intention to dispose of waste illegally.	

[1] Respondents would answer either a sensitive question or the inverse question, according to the outcome of rolling a die.

Do you drive drunk?

Collect useful insights across a population

Short-list group

- Have you ever been abroad?
- Have you ever used a taxi?
- Have you been using a plane this week?
- Did you wash your car this week?

Long-list group

- Have you ever been abroad?
- Have you ever used a taxi?
- Have you been using a plane this week?
- Did you wash your car this week?
- Have you ever been driving a car although you had drunk too much alcohol?

What data to calculate risk?

What are our assumptions?

Who	Gender	Car	Model	Colour	Email	Postcode	DOB	Risk
Alice	a	Holden	Barina	Blue	alice@hotmail.com	2000	12/09/1990	?
Bob	b	BMW	5s	White	bob@gmail.com	2010	8/02/1972	?
Carole	c	Toyota	Carola	Red	carole@yahoo.com	2115	9/05/1998	?
Derek	d	Honda	CRv	Brown	derek@aol.com	2069	5/03/1985	?

Dev

What are our assumptions?

DOB	Star Sign
April	Aries
April	Aries
May	Taurus
June	Gemini
May	Taurus
May	Taurus
April	Aries
May	Taurus
July	Cancer
August	Leo

DOB	Star Sign
April	?
May	?
June	?
July	?
August	?

DOB	Star Sign
April	Aries
May	Taurus
June	Gemini
July	Cancer
August	Leo

Dev

What are our assumptions?

Infra & security & storage

What are our assumptions?

Legal - Aust

What are our assumptions?

Notifiable Data Breaches scheme

On 22 February 2018, new privacy laws come into effect to regulate the reporting and notification of eligible data breaches to the Office of the Australian Information Commissioner (OAIC) and impacted individuals.

Legal - UK

What are our assumptions?

Expectations

Is this fair?

example@**hotmail**.com

example@**gmail**.com

Expectations

Is this fair?

MONASH UNIVERSITY ACCIDENT RESEARCH CENTRE
REPORT DOCUMENTATION PAGE

Report No.	Date	ISBN	Pages
262	May 2007	0 7326 2222 2	20

Title An Investigation into the Relationship between Vehicle Colour and Crash Risk

Author(s): Stuart Newstead & Angelo D'Elia

Sponsoring Organisation(s): This project was funded as contract research by the following organisations: New South Wales Roads and Traffic Authority, NRMA Motoring and Services, Royal Automobile Club of Victoria, Transport Accident Commission, VicRoads.

Abstract:
This study has assessed the relationship between vehicle colour and crash risk through the analysis of real crash outcomes described in mass crash data reported to police in two Australian states. A stratified induced exposure study design was employed identifying vehicle to vehicle crashes and crashes involving unprotected road users as those having a risk dependent on vehicle colour whilst exposure was induced from single vehicle crash involvement. Analysis was stratified by vehicle type, light conditions and jurisdiction of crash.

Back to our risk algorithm

How’s this looking?

Premium ($) = Risk (%) X Value ($)

- Features
  - ~~Gender~~ if ! EU
  - Car Model
  - Colour (maybe)
  - Postcode
  - Customer Name
  - ~~Email~~

What data to calculate risk?

What are our assumptions?

Who	Gender	Car	Model	Colour	Email	Postcode	DOB	Risk
Alice	a	Holden	Barina	Blue	alice@hotmail.com	2000	12/09/1990	?
Bob	b	BMW	5s	White	bob@gmail.com	2010	8/02/1972	?
Carole	c	Toyota	Carola	Red	carole@yahoo.com	2115	9/05/1998	?
Derek	d	Honda	CRv	Brown	derek@aol.com	2069	5/03/1985	?

What data to calculate risk?

What are our assumptions?

Who	Gender	Car	Model	Colour	Email	Postcode	DOB	Risk
Alice	a	Holden	Barina	Blue	alice@hotmail.com	2000	12/09/1990	?
Bob	b	BMW	5s	White	bob@gmail.com	2010	8/02/1972	?
Carole	c	Toyota	Carola	Red	carole@yahoo.com	2115	9/05/1998	?
Derek	d	Honda	CRv	Brown	derek@aol.com	2069	5/03/1985	?

Alice has a blue car

Even if we don't explicitly set out to capture the data; there's implicit relationships in our feature set

Name	Is it Female?
Carol	Yes
Betsy	Yes
Hazel	Yes
Ike	No
Sandy	Yes
Agnes	Yes
Connie	Yes