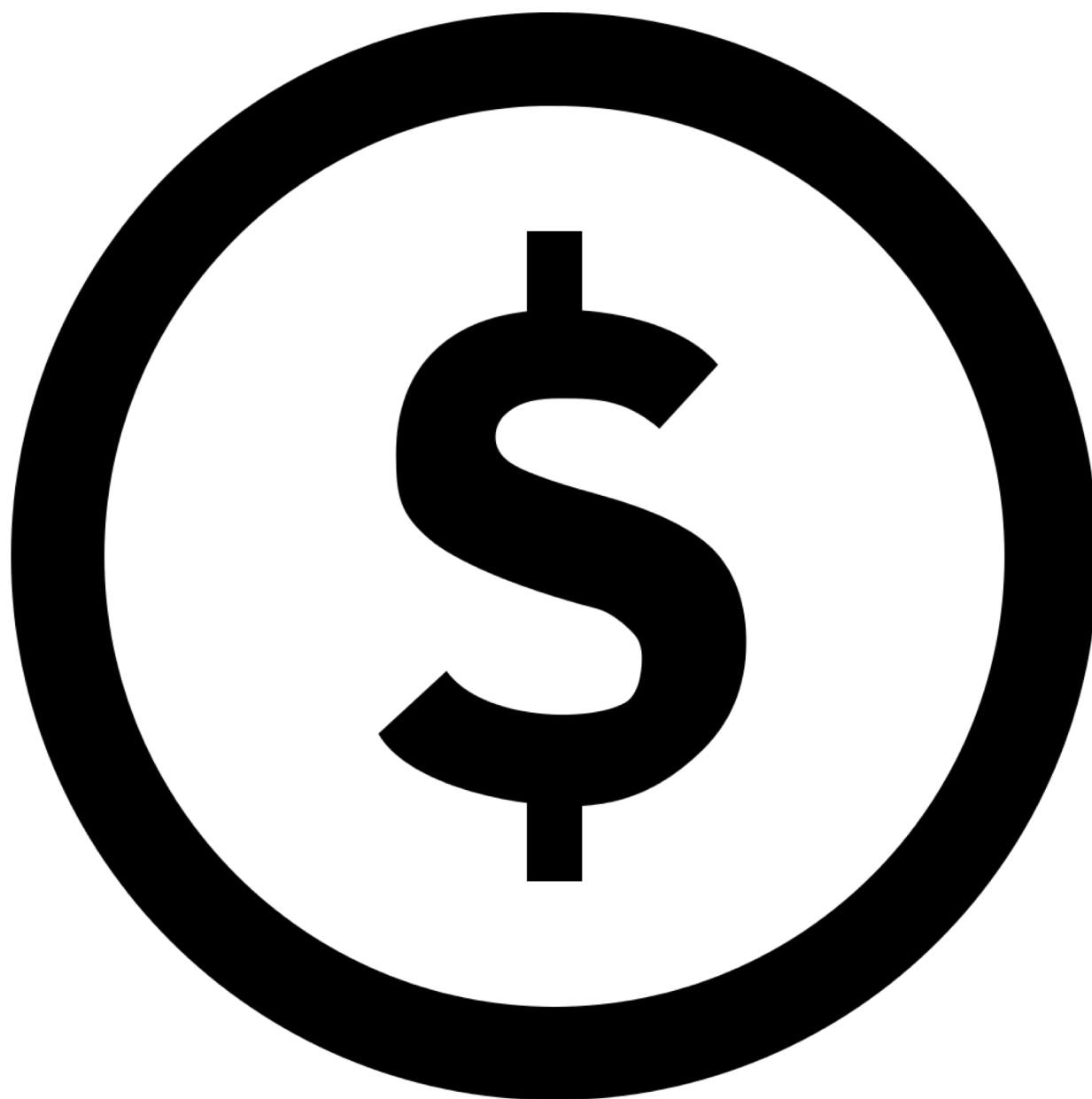


Leveraging publicly available data sources in evaluation.

Gerard Atkinson, ARTD Consultants

AES 2018 Conference, 19 September 2018

The logo for ARTD CONSULTANTS features the company name in a bold, sans-serif font. A thin, dark blue curved line arches over the text. The word 'ARTD' is in dark blue, and 'CONSULTANTS' is in a greenish-yellow color.




“An investment with a high return is likely to be high risk.” (True or false?)

- True:
- False:

“Buying shares in a single company usually provides a safer return than buying shares in a number of different companies.” (True or false?)

- True:
- False:



HILDA Survey reveals striking gender and age divide in financial literacy. Test yourself with this quiz

July 31, 2018 5:52am AEST

The latest HILDA data found women exhibiting much lower levels of financial literacy than men. Shutterstock

Email

Twitter

Facebook

LinkedIn

Print

25

596

The [Household, Income and Labour Dynamics in Australia \(HILDA\) Survey](#) tells the stories of the same group of Australians over the course of their lives. Starting in 2001, the survey now tracks more than 17,500 people in 9,500 households.

One of the most striking findings from [this year's HILDA report](#) is the large gender divide in financial literacy. Women exhibit much lower levels of financial literacy than men.

The [OECD International Network on Financial Education](#) defines financial literacy as:

Author



Roger Wilkins

Professorial Research Fellow and Deputy Director (Research), HILDA Survey, Melbourne Institute of Applied Economic and Social Research, University of Melbourne


Disclosure statement

Basic HILDA money test exposes dumb and dumber population



Shane Wright | The West Australian

Tuesday, 31 July 2018 4:00AM

 **Shane Wright**

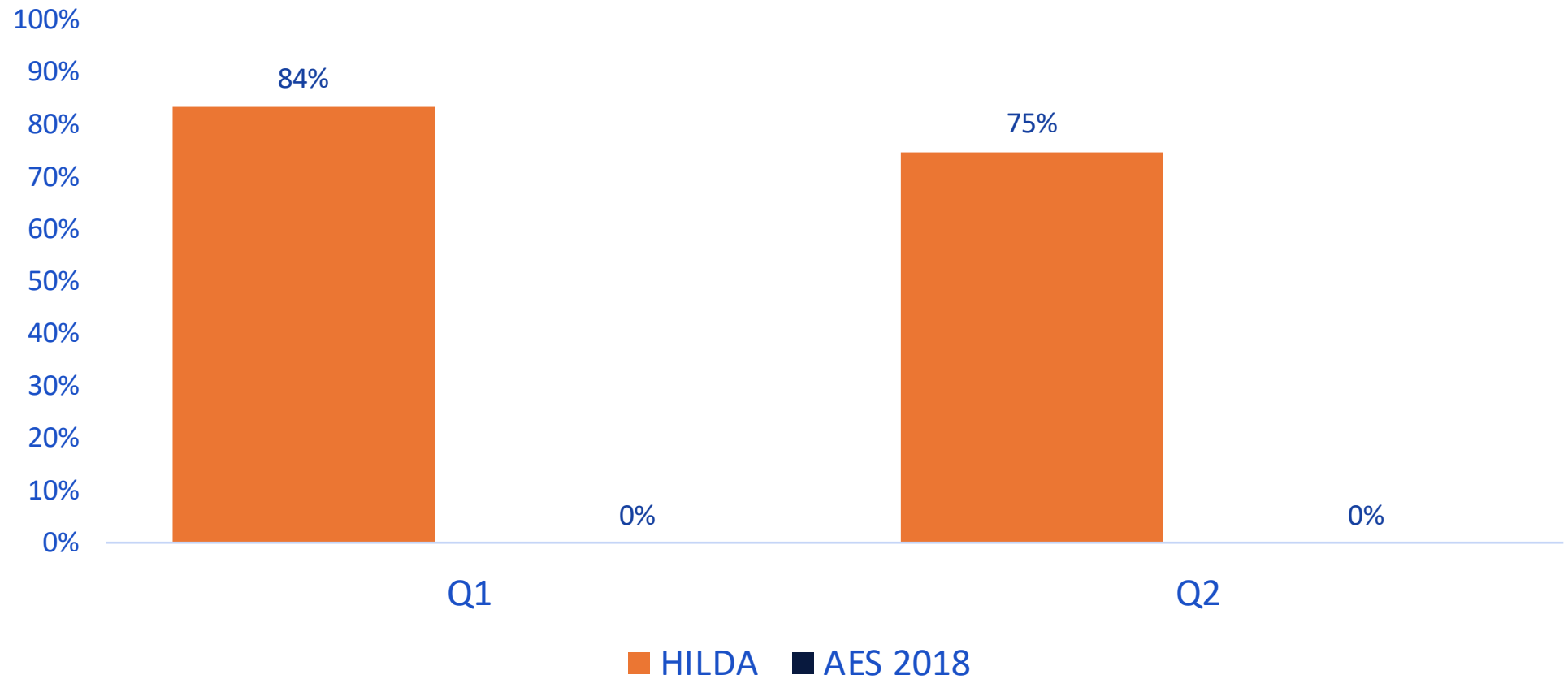


The answers

Q1 (Risk-return): True

Q2 (Diversification): False

Let's compare



What's the point of this?

Open
Data



Your
Data



Stronger
Evaluation

The rise of open data



```
graph TD; A[The rise of open data] --> B[Privacy and ethics]; B --> C[Applying open data in evaluation]; C --> D[An illustrative example]; D --> E[Summary and useful resources];
```

Privacy and ethics

Applying open data in evaluation

An illustrative example

Summary and useful resources

The rise of open data



```
graph TD; A[The rise of open data] --> B[Privacy and ethics]; B --> C[Applying open data in evaluation]; C --> D[An illustrative example]; D --> E[Summary and useful resources];
```

Privacy and ethics

Applying open data in evaluation

An illustrative example

Summary and useful resources



The rise of open data



```
graph TD; A[The rise of open data] --> B[Privacy and ethics]; B --> C[Applying open data in evaluation]; C --> D[An illustrative example]; D --> E[Summary and useful resources];
```

Privacy and ethics

Applying open data in evaluation

An illustrative example

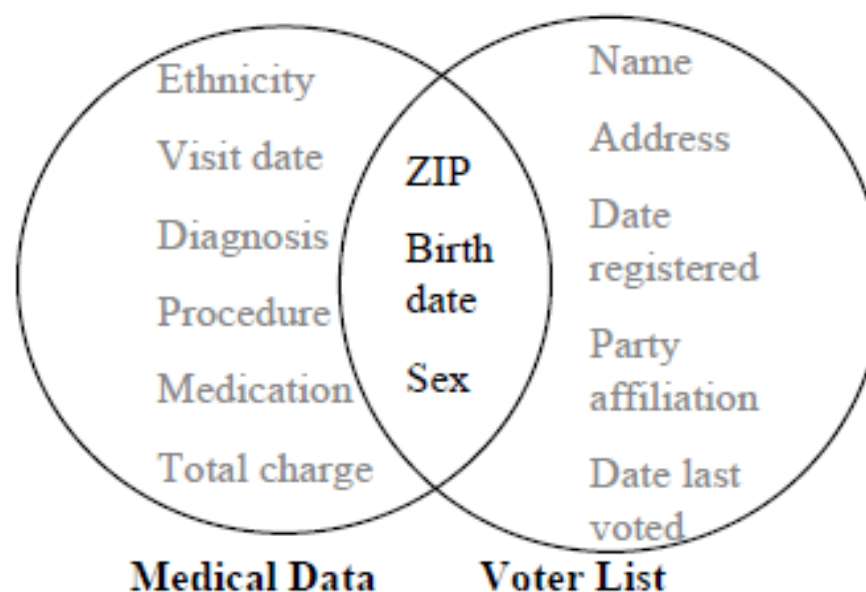
Summary and useful resources

Re-identification





For twenty dollars I purchased the voter registration list for Cambridge Massachusetts and received the information on **two diskettes** [2]. The rightmost circle in Figure 1 shows that these data included the name, address, ZIP code, birth date, and gender of each voter. This information can be linked using *ZIP*, *birth date* and *gender* to the medical information, thereby linking diagnosis, procedures, and medications to particularly named individuals. The question that remains of course is how unique would such linking be.



4-R Framework for De-Identification

Remove

Replace

Randomise

Roll up

Code of ethics

Confidentiality

5. Members should ensure responsible use of information obtained in the course of their evaluation practice and respect confidentiality undertakings. *(Refer to Guidelines 5, 9 and 15.)*

Australian Privacy Principles — a summary for APP entities

from 12 March 2014



Australian Government
Office of the
Australian Information Commissioner

APP 1 — Open and transparent management of personal information

Ensures that APP entities manage personal information in an open and transparent way. This includes having a clearly expressed and up to date APP privacy policy.

APP 2 — Anonymity and pseudonymity

Requires APP entities to give individuals the option of not identifying themselves, or of using a pseudonym. Limited exceptions apply.

APP 3 — Collection of solicited personal information

Outlines when an APP entity can collect personal information that is solicited. It applies higher standards to the collection of 'sensitive' information.

APP 4 — Dealing with unsolicited personal information

Outlines how APP entities must deal with unsolicited personal information.

APP 5 — Notification of the collection of personal information

Outlines when and in what circumstances an APP entity that collects personal information must notify an individual of certain matters.

APP 6 — Use or disclosure of personal information

Outlines the circumstances in which an APP entity may use or disclose personal information that it holds.

APP 7 — Direct marketing

An organisation may only use or disclose personal information for direct marketing purposes if certain conditions are met.

APP 8 — Cross-border disclosure of personal information

Outlines the steps an APP entity must take to protect personal information before it is disclosed overseas.

APP 9 — Adoption, use or disclosure of government related identifiers

Outlines the limited circumstances when an organisation may adopt a government related identifier of an individual as its own identifier, or use or disclose a government related identifier of an individual.

APP 10 — Quality of personal information

An APP entity must take reasonable steps to ensure the personal information it collects is accurate, up to date and complete. An entity must also take reasonable steps to ensure the personal information it uses or discloses is accurate, up to date, complete and relevant, having regard to the purpose of the use or disclosure.

APP 11 — Security of personal information

An APP entity must take reasonable steps to protect personal information it holds from misuse, interference and loss, and from unauthorised access, modification or disclosure. An entity has obligations to destroy or de-identify personal information in certain circumstances.

APP 12 — Access to personal information

Outlines an APP entity's obligations when an individual requests to be given access to personal information held about them by the entity. This includes a requirement to provide access unless a specific exception applies.

APP 13 — Correction of personal information

Outlines an APP entity's obligations in relation to correcting the personal information it holds about individuals.

www.oaic.gov.au

For private sector organisations,
Australian Government
and Norfolk Island agencies
covered by the *Privacy Act 1988*

GDPR

Assess risk

Put in place
governance
requirements

Ensure clear
information on
what data are
gathered and
stored

Ensure clear
explicit
permission is
provided

Have processes
in place for
extraction and
removal of
individual
records

The rise of open data



```
graph TD; A[The rise of open data] --> B[Privacy and ethics]; B --> C[Applying open data in evaluation]; C --> D[An illustrative example]; D --> E[Summary and useful resources];
```

Privacy and ethics

Applying open data in evaluation

An illustrative example

Summary and useful resources

Program
Design



Program
Delivery



Post-program
Evaluation

Quasi-experimental approaches

Pre-post

Regression
adjustment

Propensity
scoring

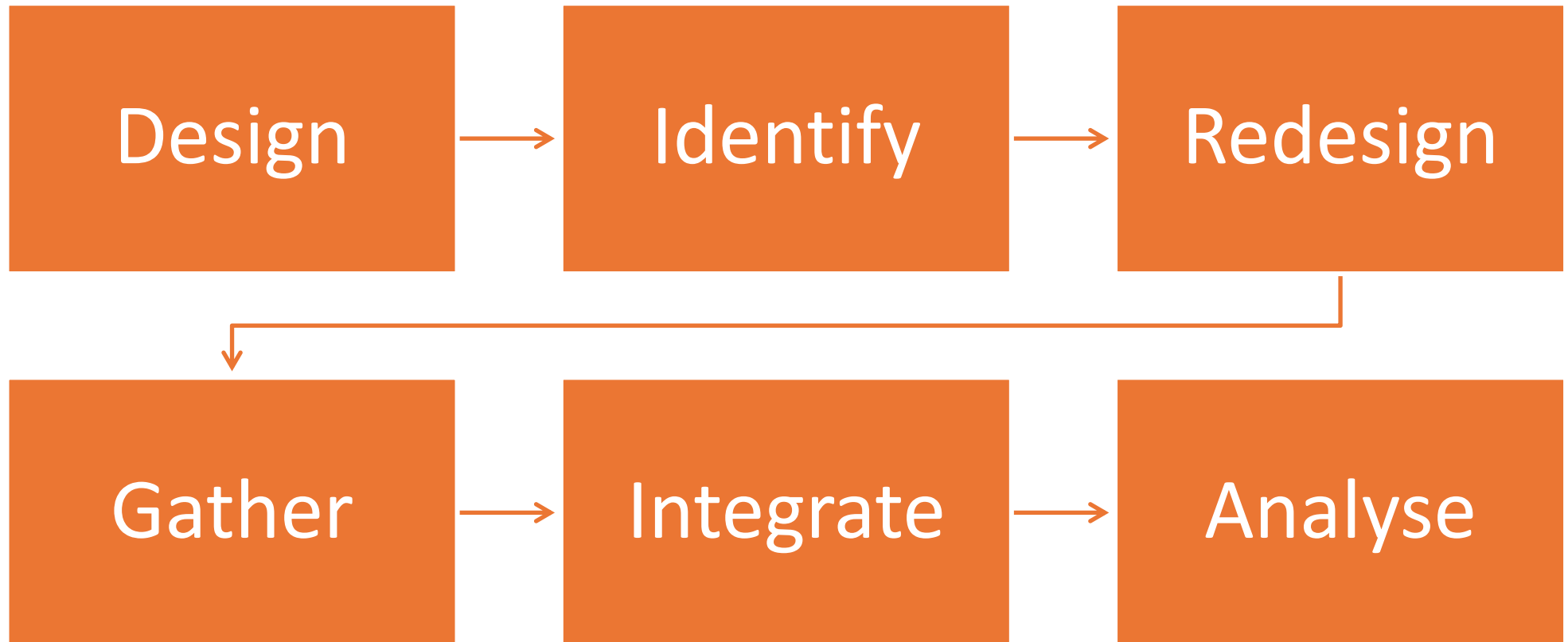
Difference-in-
difference

Interrupted
time series

Synthetic
controls

Regression
discontinuity

Instrumental
variables



“Give a small boy a hammer, and he will find that everything he encounters needs pounding.”

- Abraham Kaplan

Aligning your evaluation







The rise of open data



```
graph TD; A[The rise of open data] --> B[Privacy and ethics]; B --> C[Applying open data in evaluation]; C --> D[An illustrative example]; D --> E[Summary and useful resources];
```

Privacy and ethics

Applying open data in evaluation

An illustrative example

Summary and useful resources

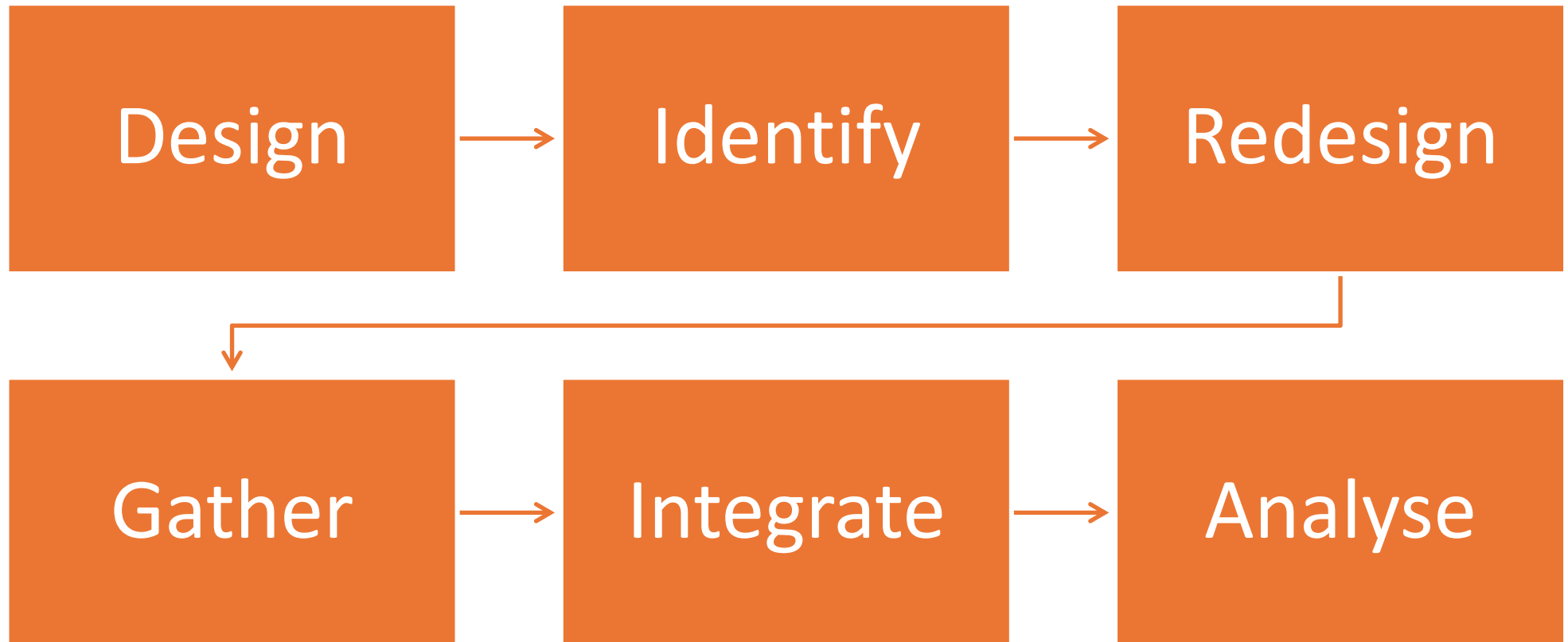
“Leveraging publicly available longitudinal and transactional data sources to create comparison groups in quasi-experimental and natural experimental evaluation scenarios.”

Metrics:

- Title Length (characters and words)
- Word length

Metric	My paper	AES 2018 (S.D.)
Title length (characters)	171	96.4 (39.8)
Title length (words)	19	13.2 (5.8)
Average word length (characters)	7.95	6.4 (1.1)

We can use open data to create a
comparative analysis!



How do AES conference titles compare
with other conferences?

Design

Metrics:

- Title Length (characters and words)
- Word length
- Question marks
- Double-barrelled titles
- Quotation marks

Identify

Sources:

- AES conferences (2014-2018)
- AEA conferences (2014-2018)
- IBM Centre for Advanced Studies conference “CASCON” (1991-2016)
- Neural Information Processing Systems “NIPS” (1987-2017)

Redesign

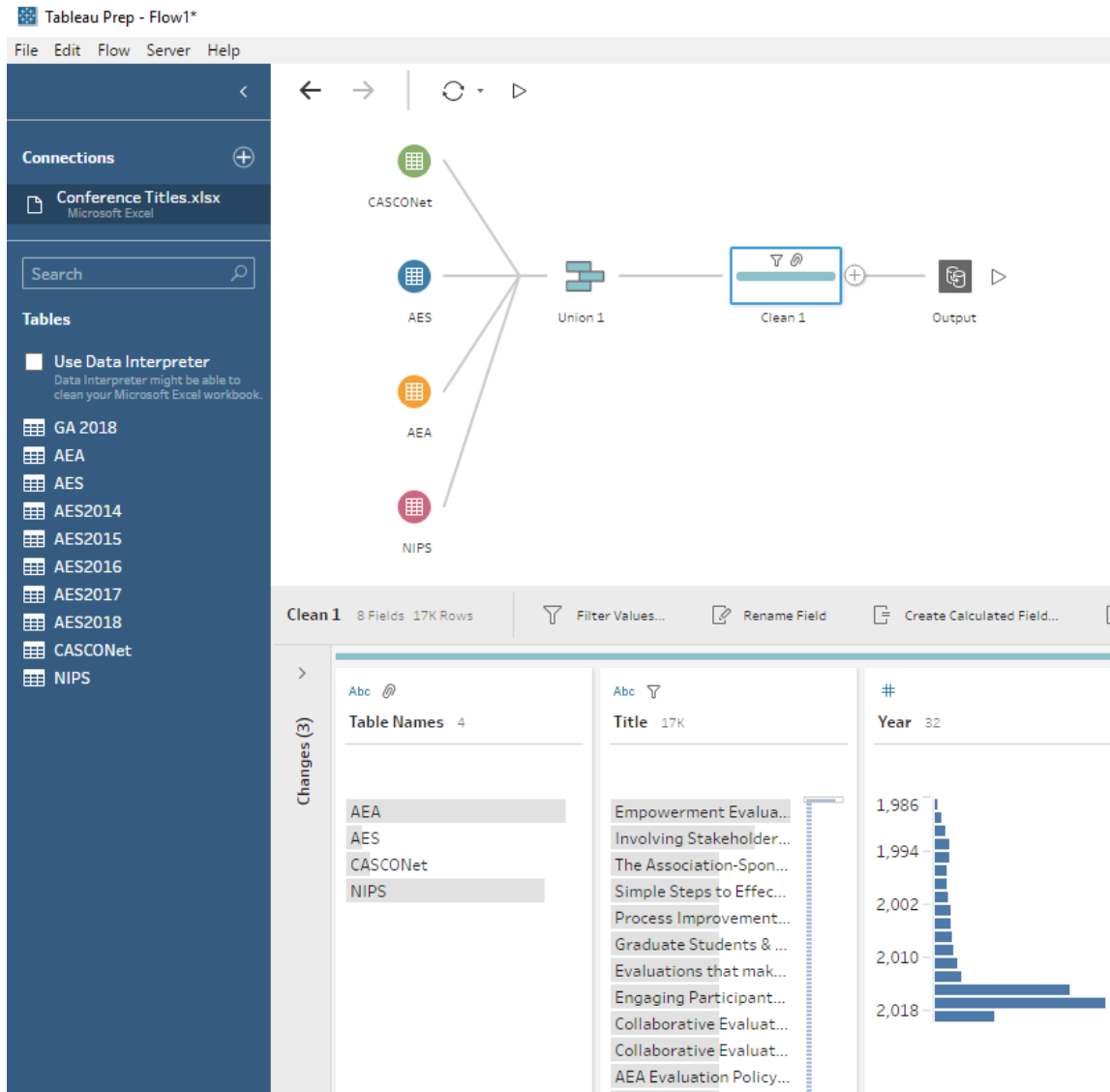
Metrics:

- Title Length (characters and words)
- Word length
- Question marks
- Double-barrelled titles
- ~~Quotation marks~~

Gather

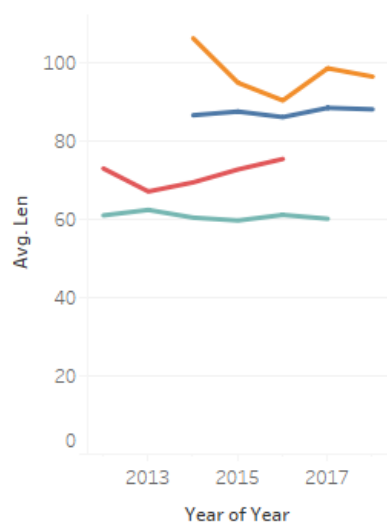


DOWNLOADING...

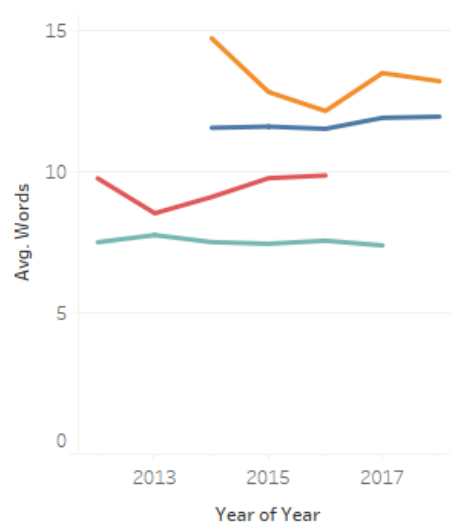


Metric	This paper	All my AES 2018 papers	AES (2014-2018)	AEA (2014-2018)	CASCON (1991-2016)	NIPS (1987-2017)
n	1	3	557	8016	845	7241
Title length (characters)	171	91.3 (57.7)	96.9 (39.6)	87.3 (34.6)	63.0 (21.0)	60.2 (20.8)
Title length (words)	19	11.3 (6.1)	13.2 (5.6)	11.7 (4.8)	8.2 (2.9)	7.6 (2.8)
Average word length (characters)	7.95	6.9 (1.3)	6.4 (1.1)	6.5 (1.2)	6.7 (1.3)	6.9 (1.2)
Questions	0%	0%	11%	6%	1%	2%
Double-barrelled	0%	33%	66%	48%	22%	12%

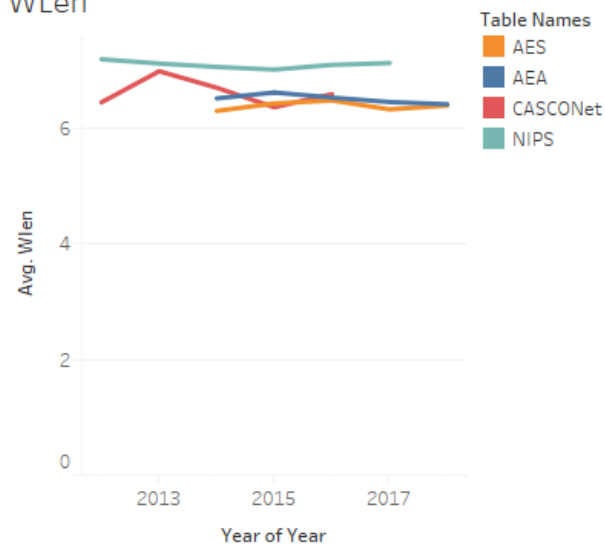
TLen Char



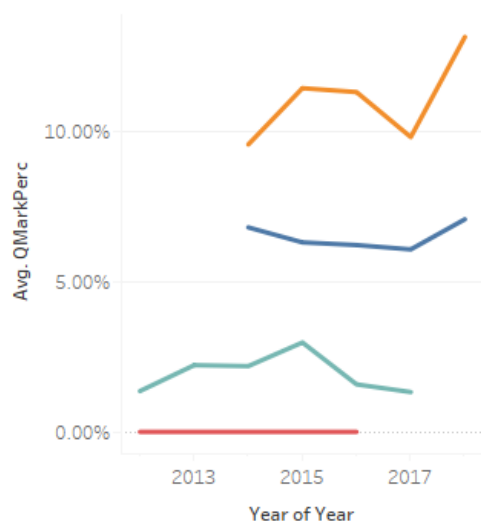
TLen Word



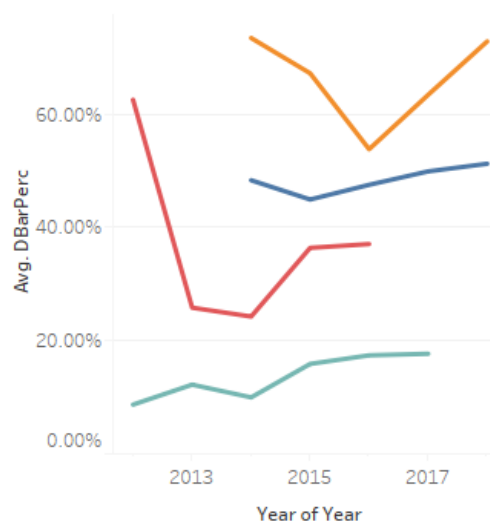
WLen



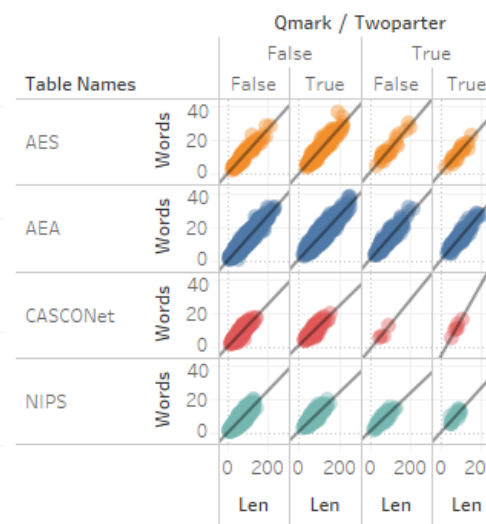
Questions

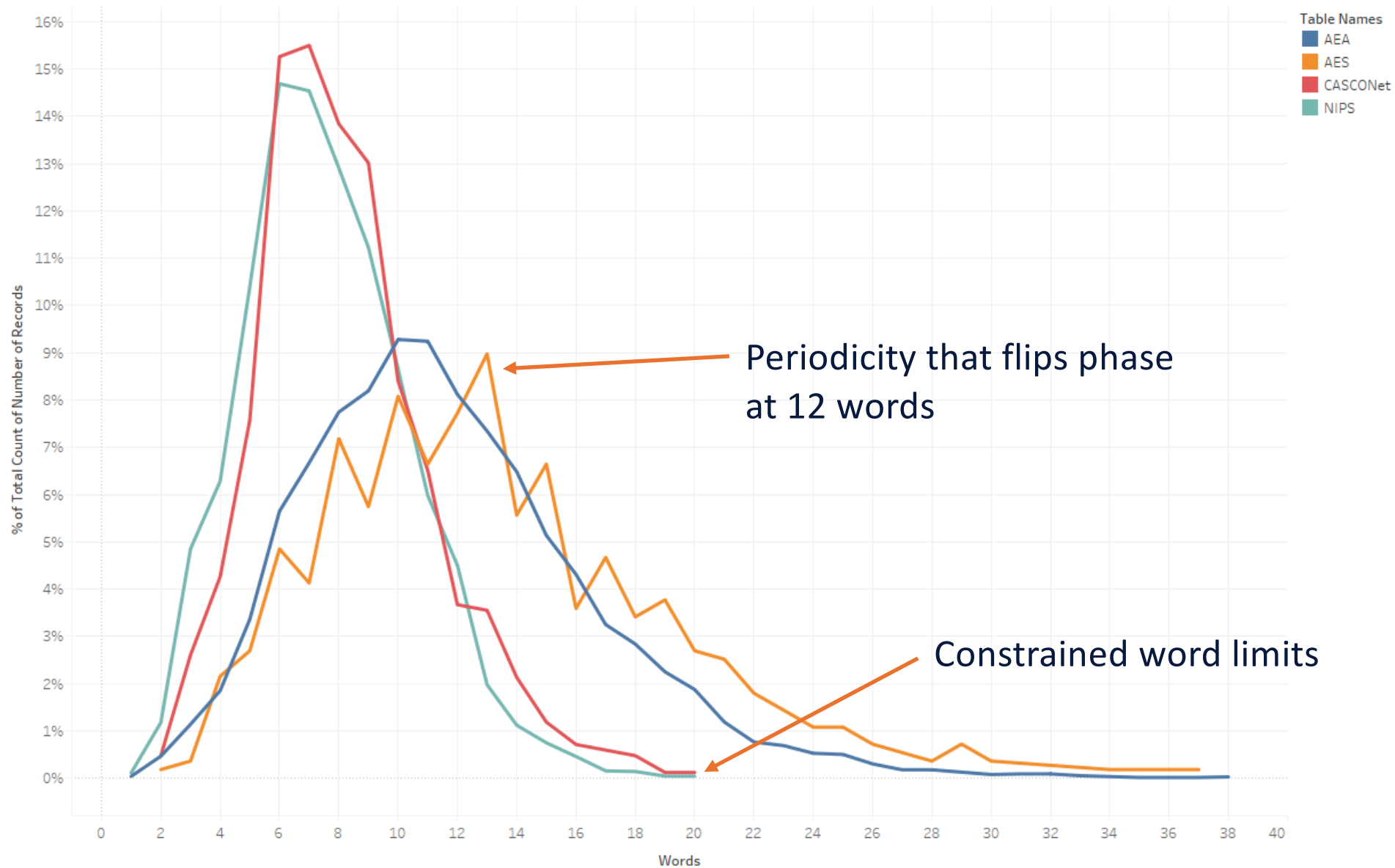


Twoparter

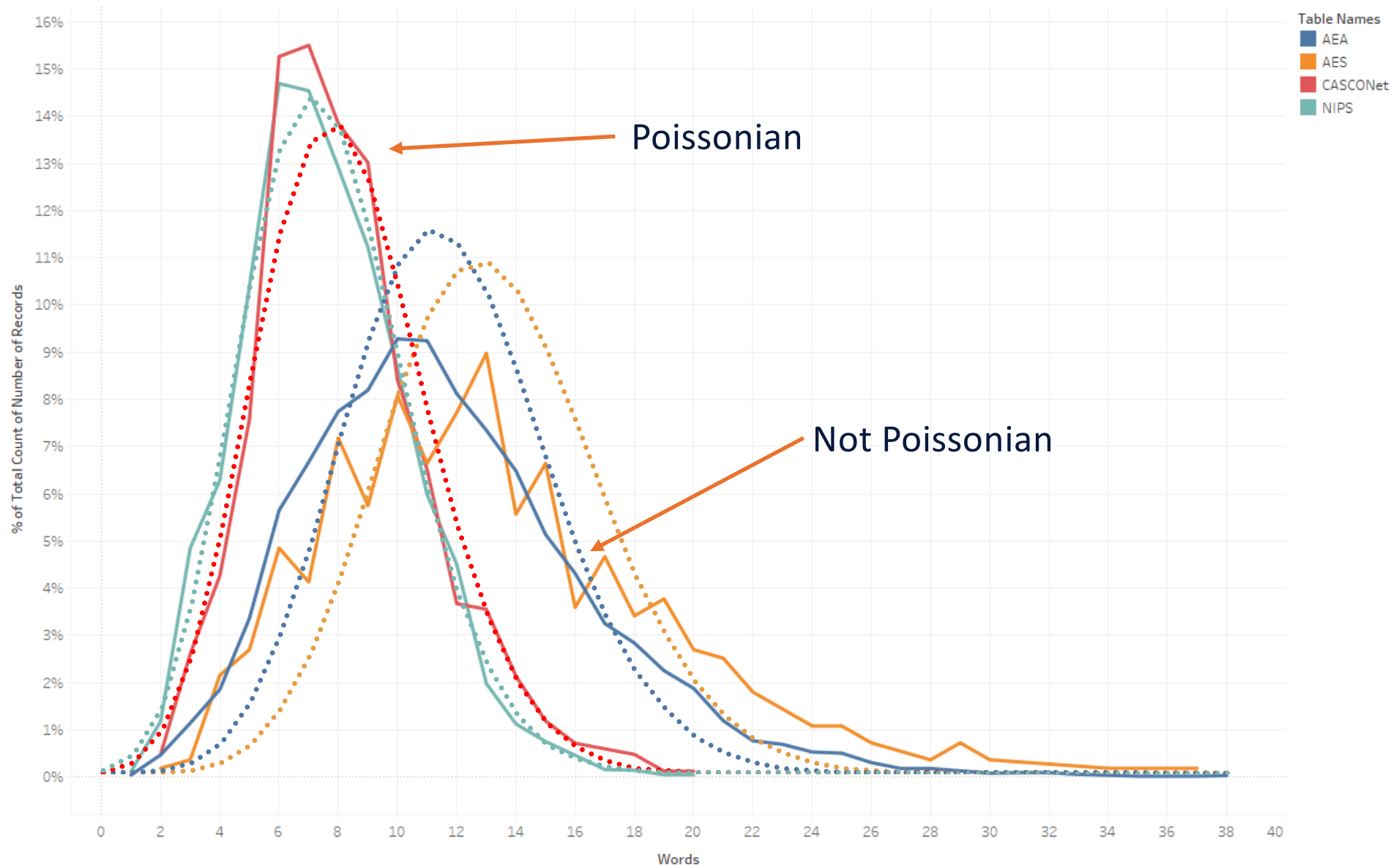


Words v Len





The trend of % of Total Count of Number of Records for Words. Color shows details about Table Names.



The trend of % of Total Count of Number of Records for Words. Color shows details about Table Names.

The rise of open data



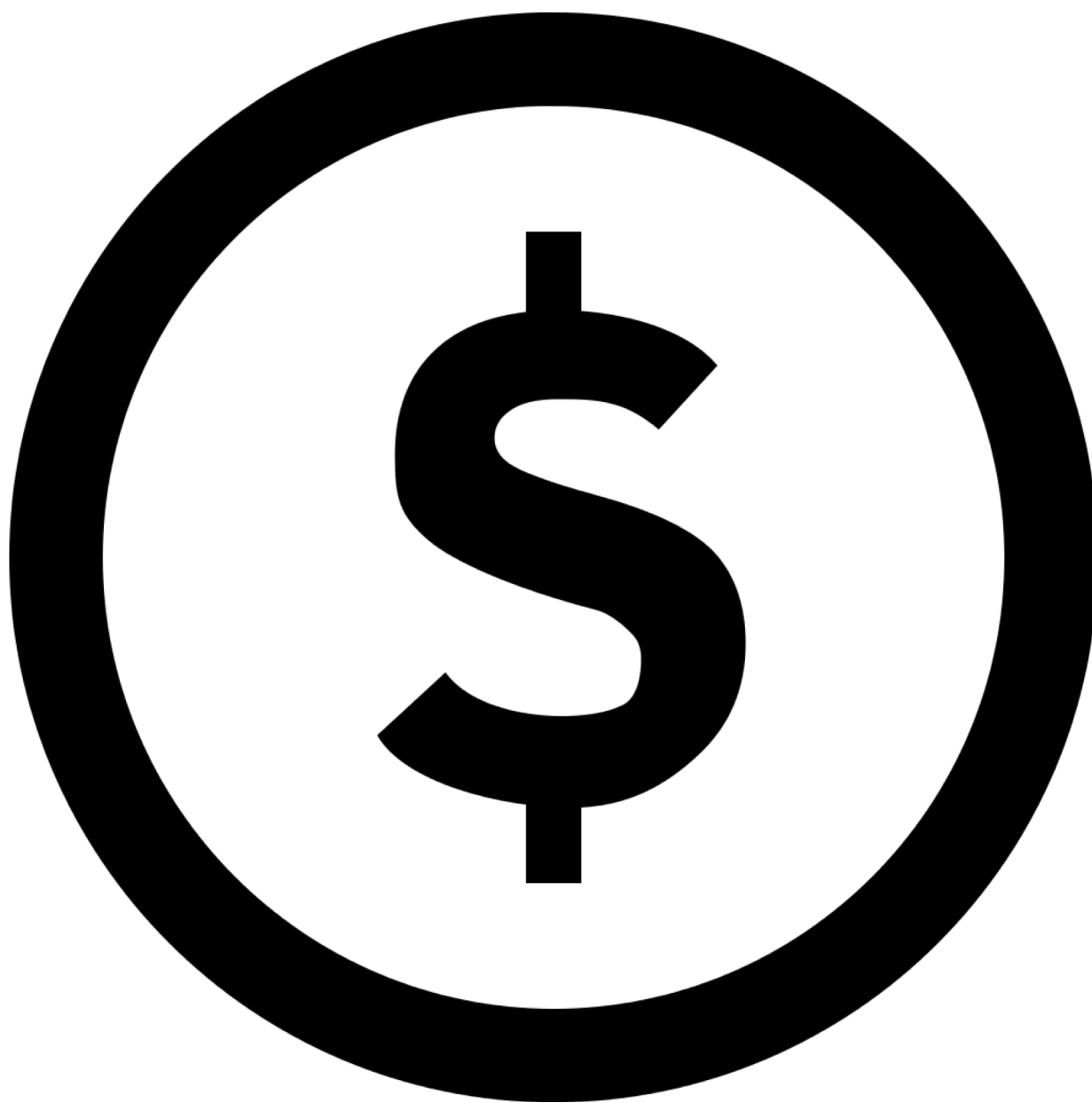
```
graph TD; A[The rise of open data] --> B[Privacy and ethics]; B --> C[Applying open data in evaluation]; C --> D[An illustrative example]; D --> E[Summary and useful resources];
```

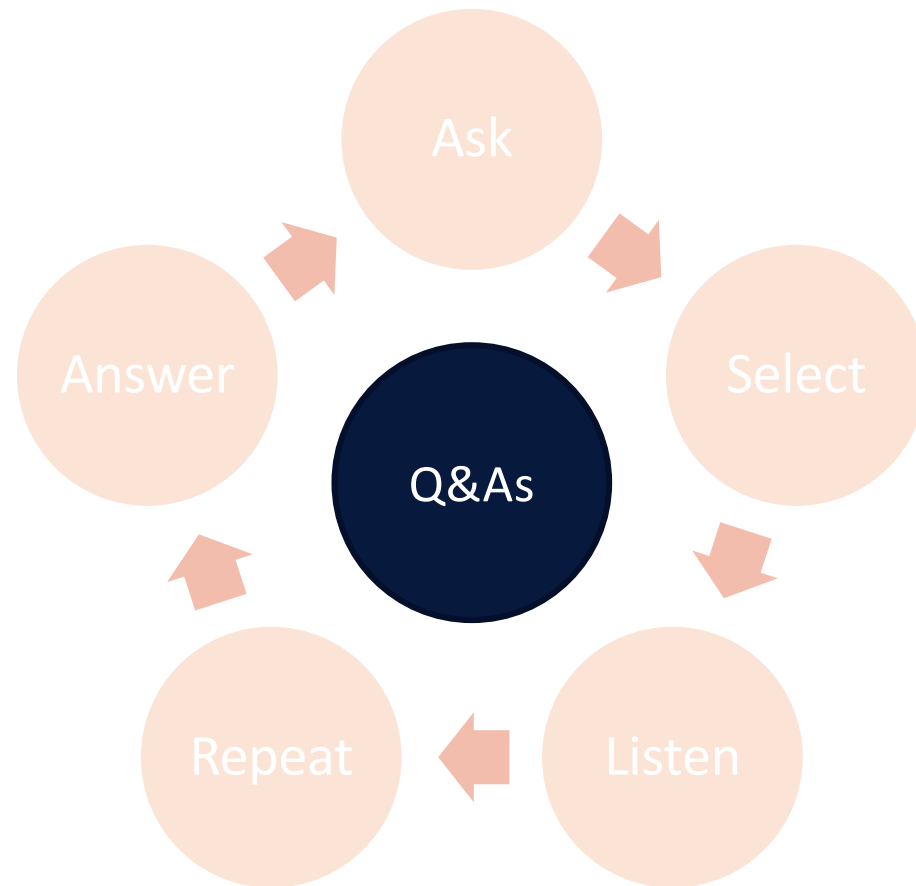
Privacy and ethics

Applying open data in evaluation

An illustrative example

Summary and useful resources





Open
Data



Your
Data



Stronger
Evaluation

Contact



Gerard Atkinson
Manager
gerard.atkinson@artd.com.au

Thanks to:

- ARTD Consultants for letting me do this analysis in the first place
- Milos Popovic at the AEA for preparing and providing their conference title dataset
- Ben Hamner for preparing the NIPS dataset
- Dixin Luo and Kelly Lyons for preparing the CASCONet dataset

