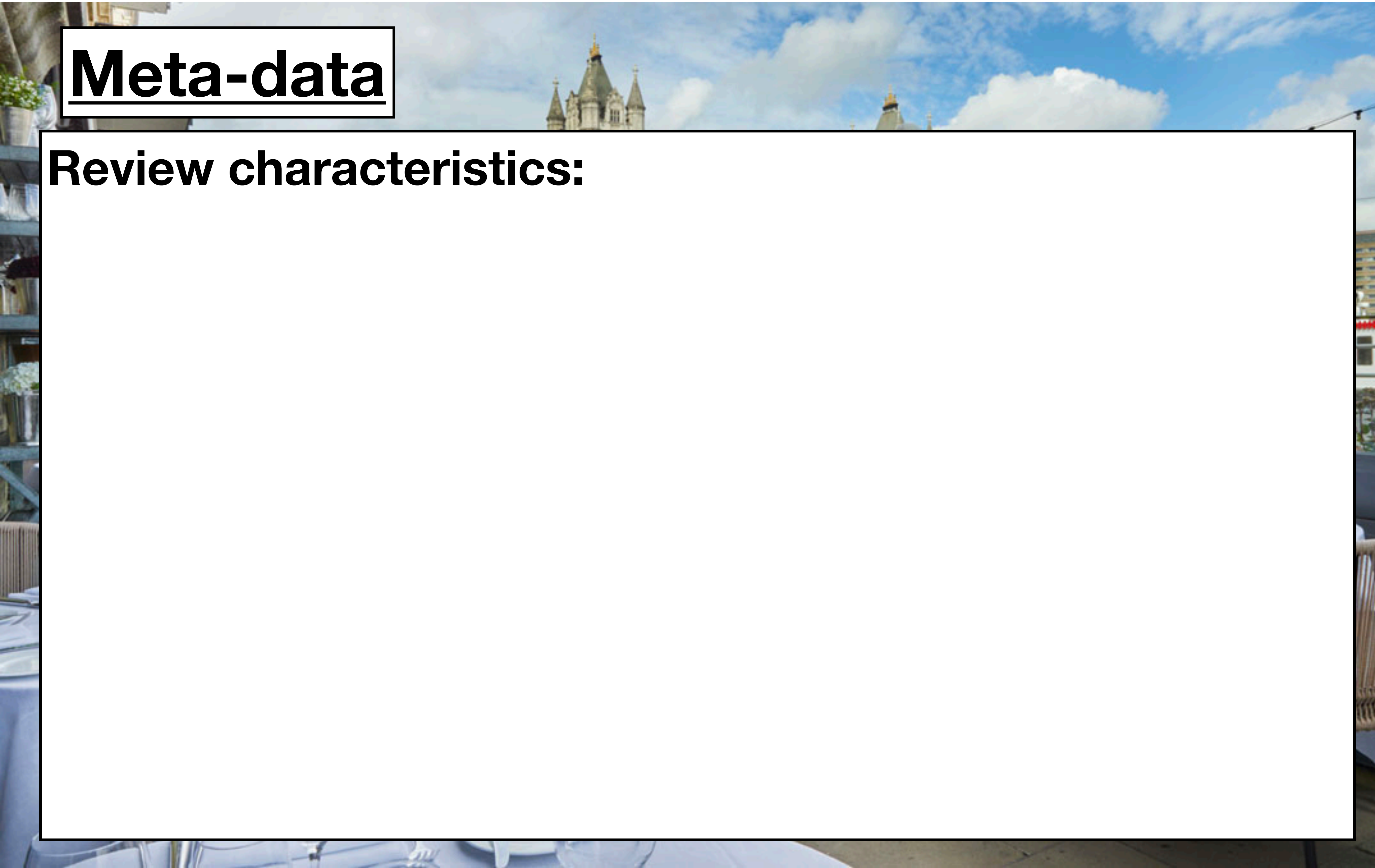
Text Analysis for Social Scientists and Leaders

Class 2: Interpretability & Clustering

Prof. Michael Yeomans

Application in Your Code

Application in Your Code



Meta-data

Review characteristics:

Application in Your Code

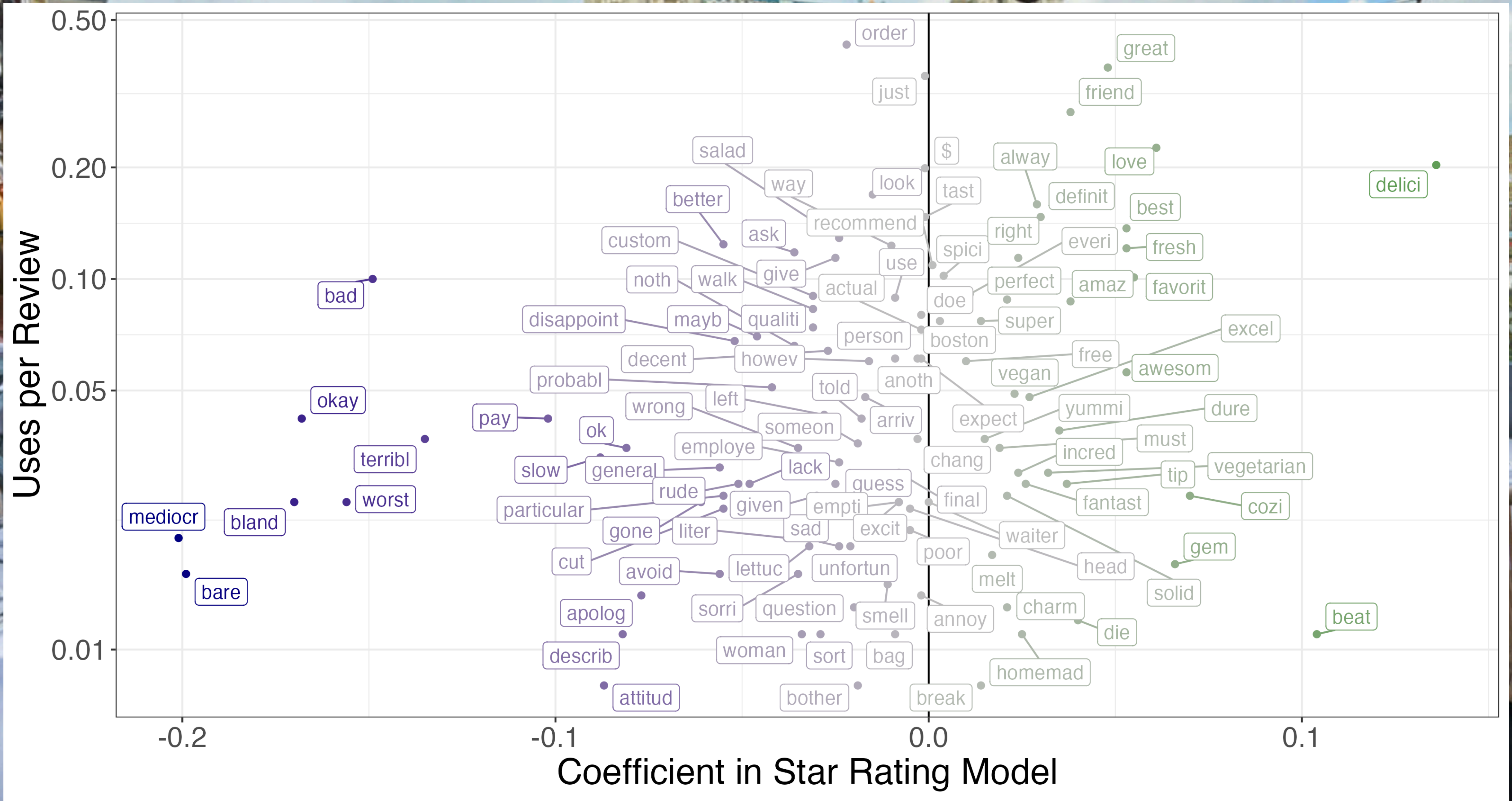
Meta-data

Review characteristics:

Star rating, time, likes from others, replies

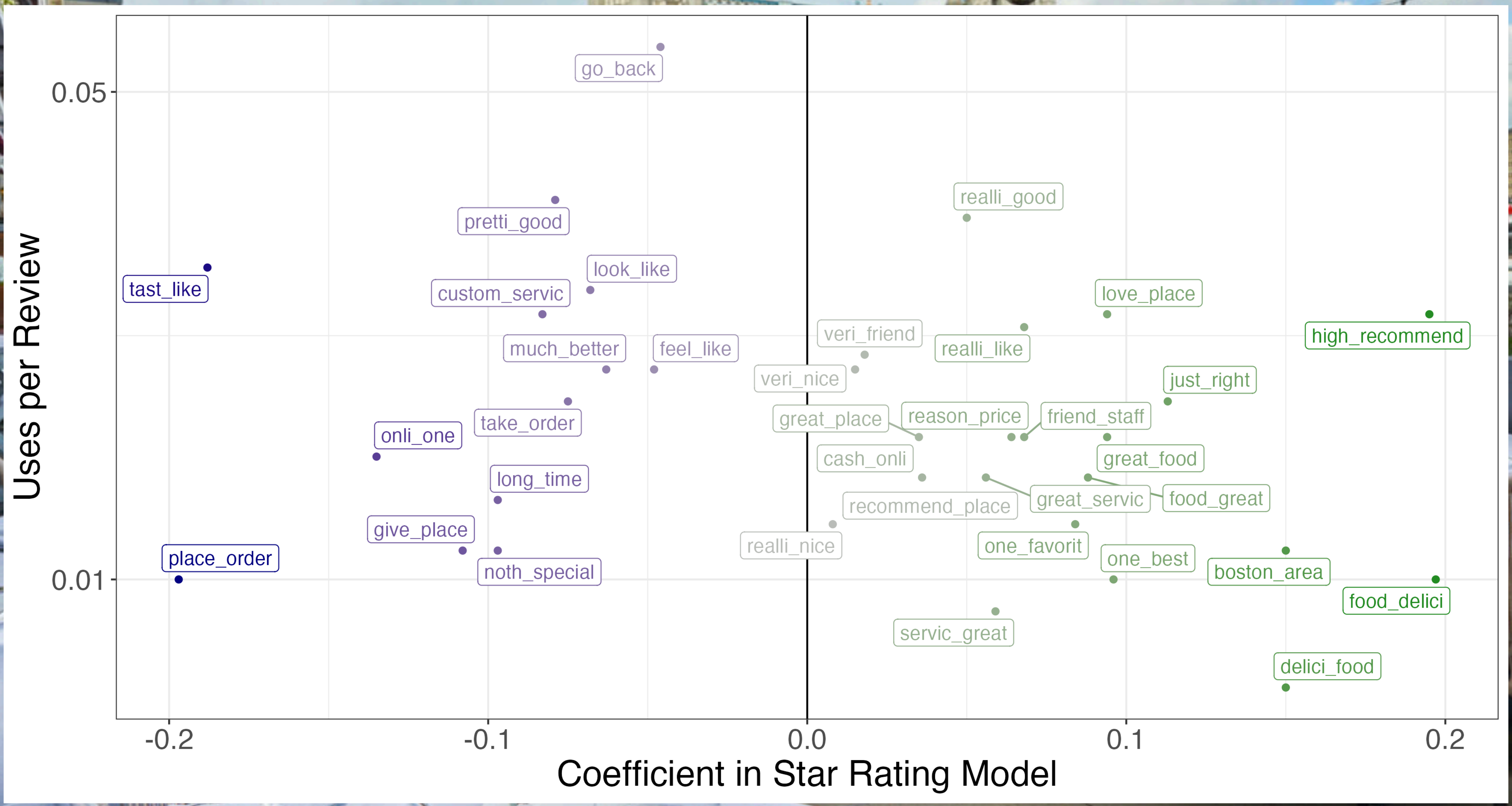
Application in Your Code

Meta-data



Application in Your Code

Meta-data



Application in Your Code

Meta-data

Review characteristics:

Star rating, time, likes from others, replies

Reviewer characteristics:

Application in Your Code

<u>Meta-data</u>

Review characteristics:

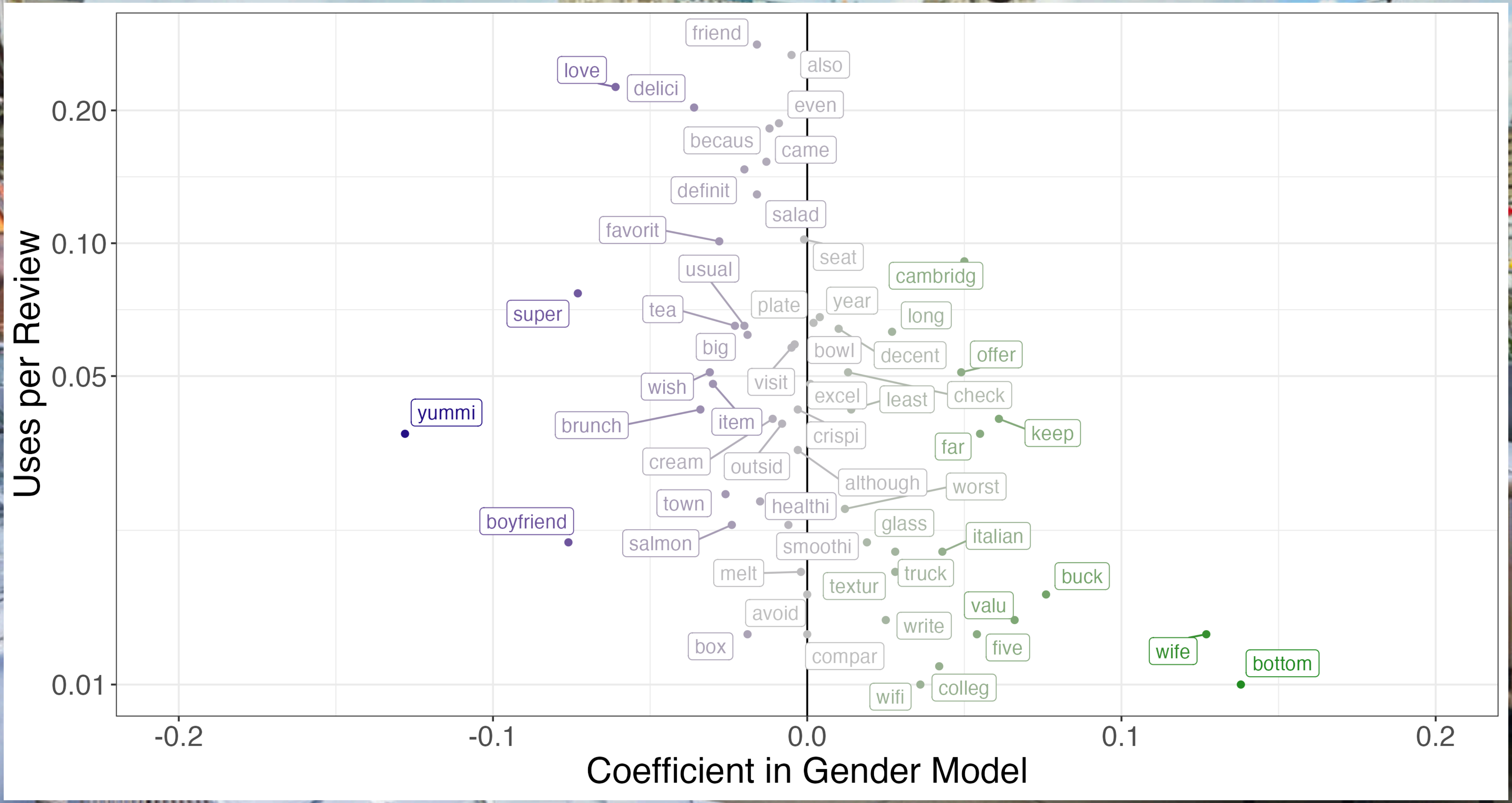
Star rating, time, likes from others, replies

Reviewer characteristics:

Demographics - gender, ethnicity, age, SES, education

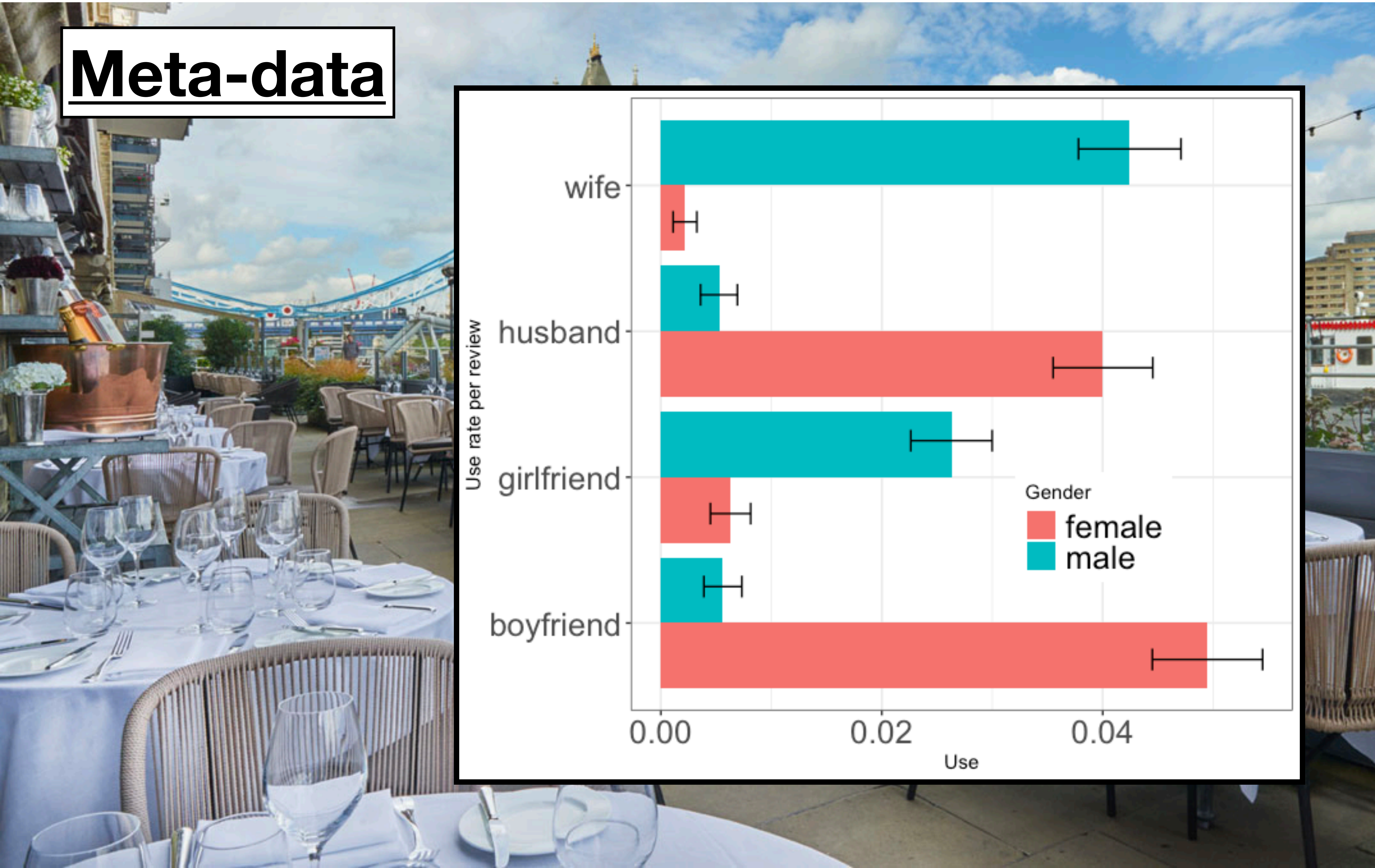
Application in Your Code

Meta-data



Application in Your Code

Meta-data



Application in Your Code

Meta-data

Review characteristics:

Star rating, time, likes from others, replies

Reviewer characteristics:

Demographics - gender, ethnicity, age, SES, education

Yelp data - Experience, other ratings

Application in Your Code

Meta-data

Review characteristics:
Star rating, time, likes from others, replies

Reviewer characteristics:
Demographics - gender, ethnicity, age, SES, education
Yelp data - Experience, other ratings

Restaurant characteristics:
Yelp metadata - price range, food type, good for kids, etc.

Application in Your Code

Meta-data

Review characteristics:

Star rating, time, likes from others, replies

Reviewer characteristics:

Demographics - gender, ethnicity, age, SES, education

Yelp data - Experience, other ratings

Restaurant characteristics:

Yelp metadata - price range, food type, good for kids, etc.

Location - city, high/low SES, partisanship, etc.

Application in Your Code

Meta-data

Review characteristics:
Star rating, time, likes from others, replies

Reviewer characteristics:
Demographics - gender, ethnicity, age, SES, education
Yelp data - Experience, other ratings

Restaurant characteristics:
Yelp metadata - price range, food type, good for kids, etc.
Location - city, high/low SES, partisanship, etc.
Business metadata - size, rent, revenues, rent, ownership

Application in Your Code

Meta-data

Review characteristics:

Star rating, time, likes from others, replies

Reviewer characteristics:

Demographics - gender, ethnicity, age, SES, education

Yelp data - Experience, other ratings

Restaurant characteristics:

Yelp metadata - price range, food type, good for kids, etc.

Location - city, high/low SES, partisanship, etc.

Business metadata - size, rent, revenues, rent, ownership

Government metadata - taxes, health inspections

Health Inspections

(Kang et al., 2013)

Health Inspections

(Kang et al., 2013)

Meta-data

Hygienic:	
Cooking Method & Garnish: brew, frosting, grill, crush, crust, taco, burrito, toast **Healthy or Fancier Ingredients**: celery, calamity, wine, broccoli, salad, flatbread, olive, pesto	
Cuisines : Breakfast, Fish & Chips, Fast Food, German, Diner, Belgian, European, Sandwiches, Vegetarian	
Whom & When: date, weekend, our, husband, evening, night	
Sentiment: lovely, yummy, generous, friendly, great, nice	
Service & Atmosphere: selection, attitude, atmosphere, ambiance, pretentious	

Table 3: Lexical Cues & Examples - Hygienic (clean)

Health Inspections

(Kang et al., 2013)

Meta-data

Hygienic:
Cooking Method & Garnish: brew, frosting, grill, crush, crust, taco, burrito, toast **Healthy or Fancier Ingredients**: celery, calamity, wine, broccoli, salad, flatbread, olive, pesto
Cuisines : Breakfast, Fish & Chips, Fast Food, German, Diner, Belgian, European, Sandwiches, Vegetarian
Whom & When: date, weekend, our, husband, evening, night
Sentiment: lovely, yummy, generous, friendly, great, nice
Service & Atmosphere: selection, attitude, atmosphere, ambiance, pretentious

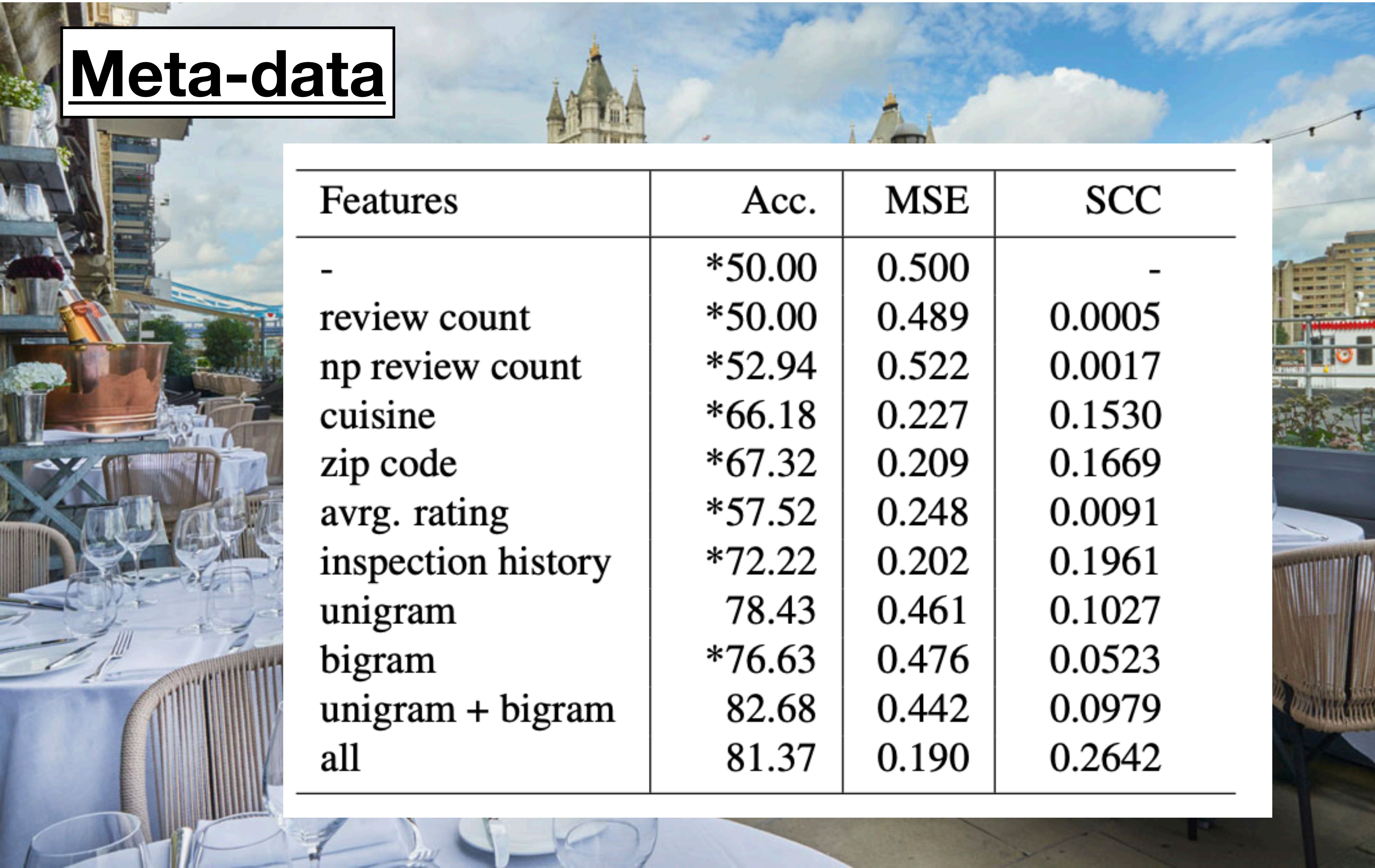
Table 3: Lexical Cues & Examples - Hygienic (clean)

Hygienic gross, mess, sticky, smell, restroom, dirty
Basic Ingredients: beef, pork, noodle, egg, soy, ramen, pho,
Cuisines Vietnamese, Dim Sum, Thai, Mexican, Japanese, Chinese, American, Pizza, Sushi, Indian, Italian, Asian
Sentiment: cheap, never,
Service & Atmosphere cash, worth, district, delivery, think, really, thing, parking, always, usually, definitely - door: "The wait is always out the *door* when I actually want to go there", - sticker: "I had *sticker* shock when I saw the prices.", - student: "heap, large portions and tasty = the perfect *student* food!", - the size: "i was pretty astonished at *the size* of all the plates for the money.", - was dry: "The beef *was dry*, the sweet soy and anise-like sauce was TOO salty (almost inedible).", - pool: "There are *pool* tables, TV airing soccer games from around the globe and of course - great drinks!"

Table 2: Lexical Cues & Examples - Unhygienic (dirty)

Health Inspections

(Kang et al., 2013)



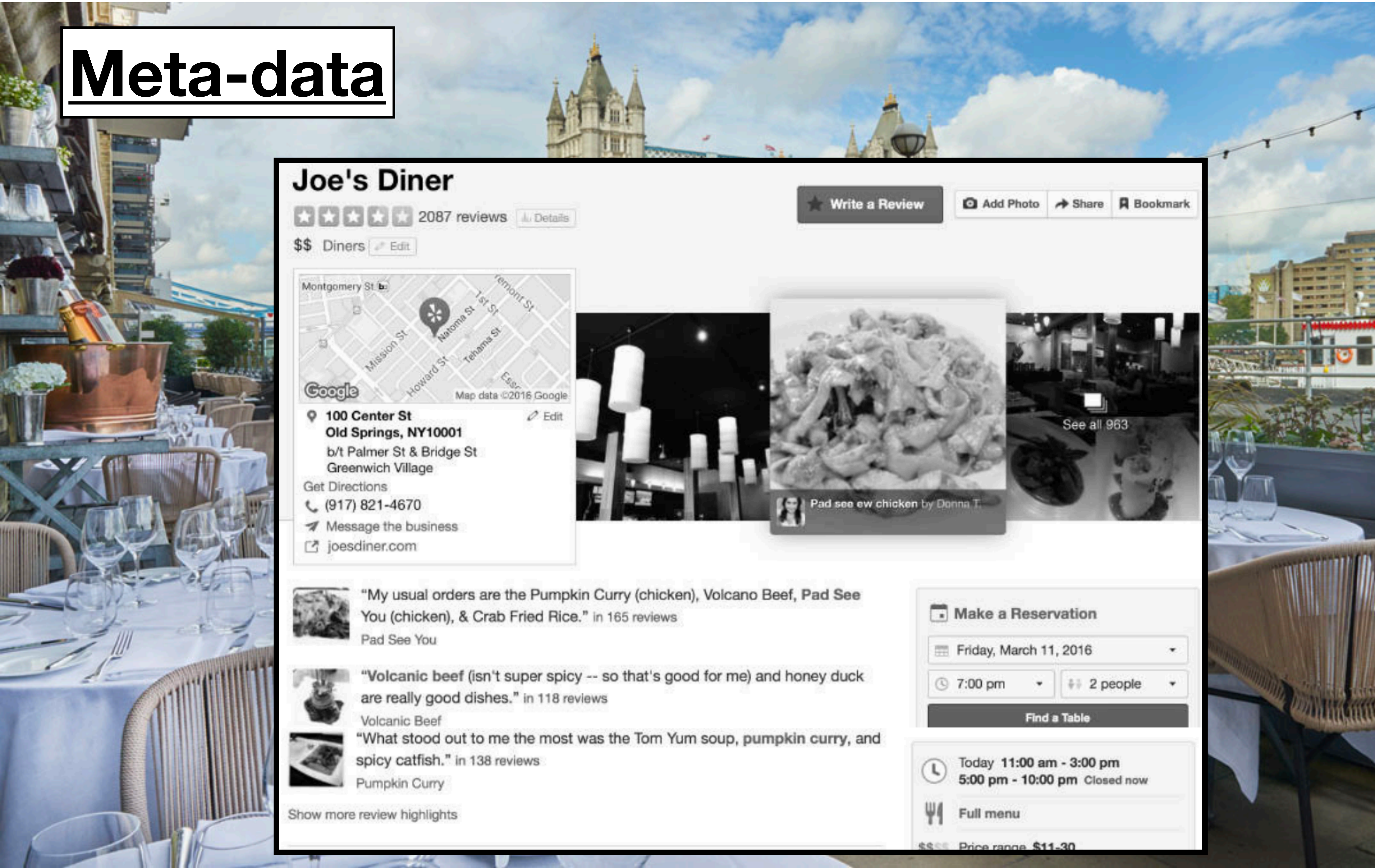
Meta-data

Features	Acc.	MSE	SCC
-	*50.00	0.500	-
review count	*50.00	0.489	0.0005
np review count	*52.94	0.522	0.0017
cuisine	*66.18	0.227	0.1530
zip code	*67.32	0.209	0.1669
avrg. rating	*57.52	0.248	0.0091
inspection history	*72.22	0.202	0.1961
unigram	78.43	0.461	0.1027
bigram	*76.63	0.476	0.0523
unigram + bigram	82.68	0.442	0.0979
all	81.37	0.190	0.2642

Health Inspections

(Dai & Luca, 2020)

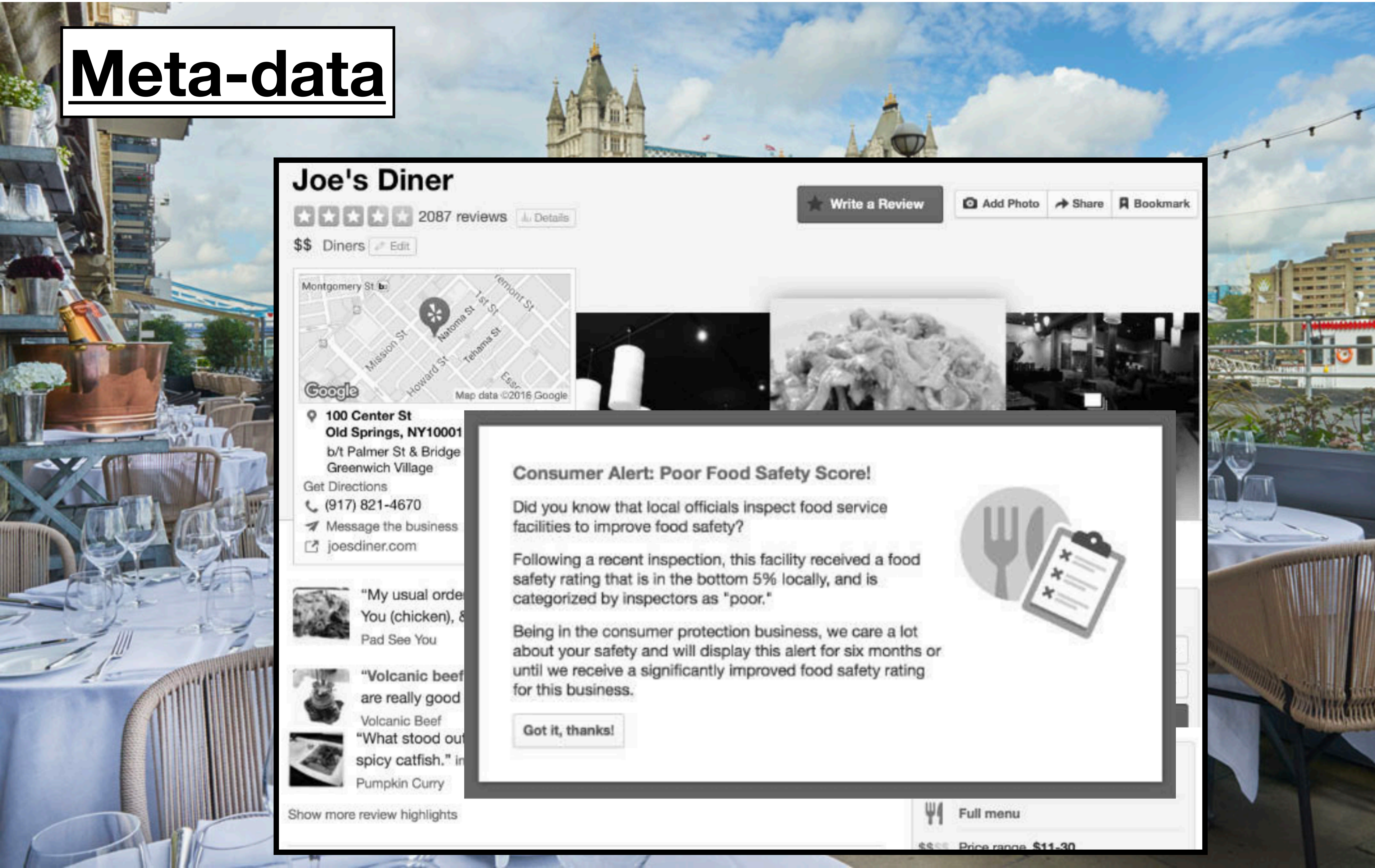
Meta-data



Health Inspections

(Dai & Luca, 2020)

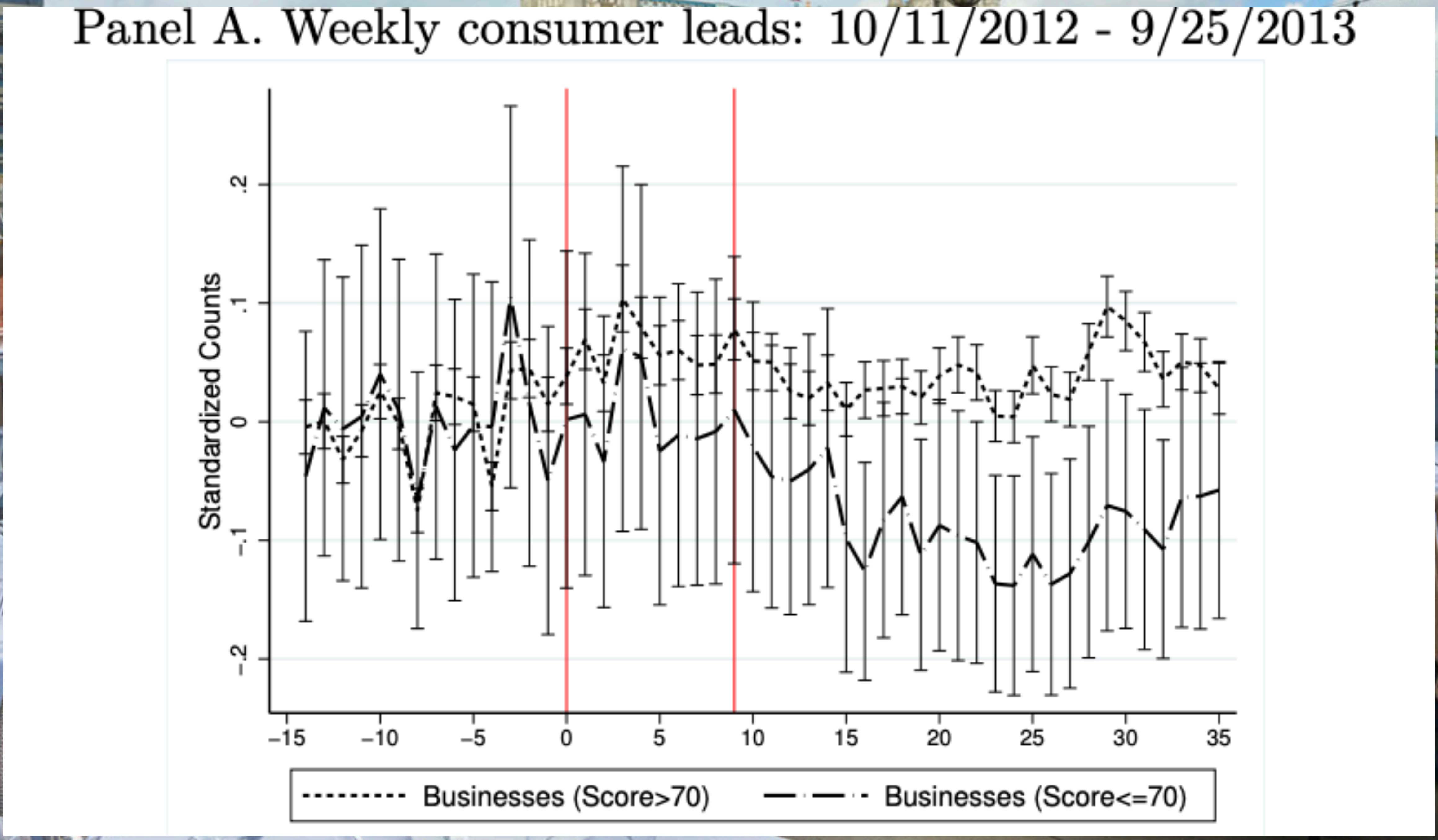
Meta-data



Health Inspections

(Dai & Luca, 2020)

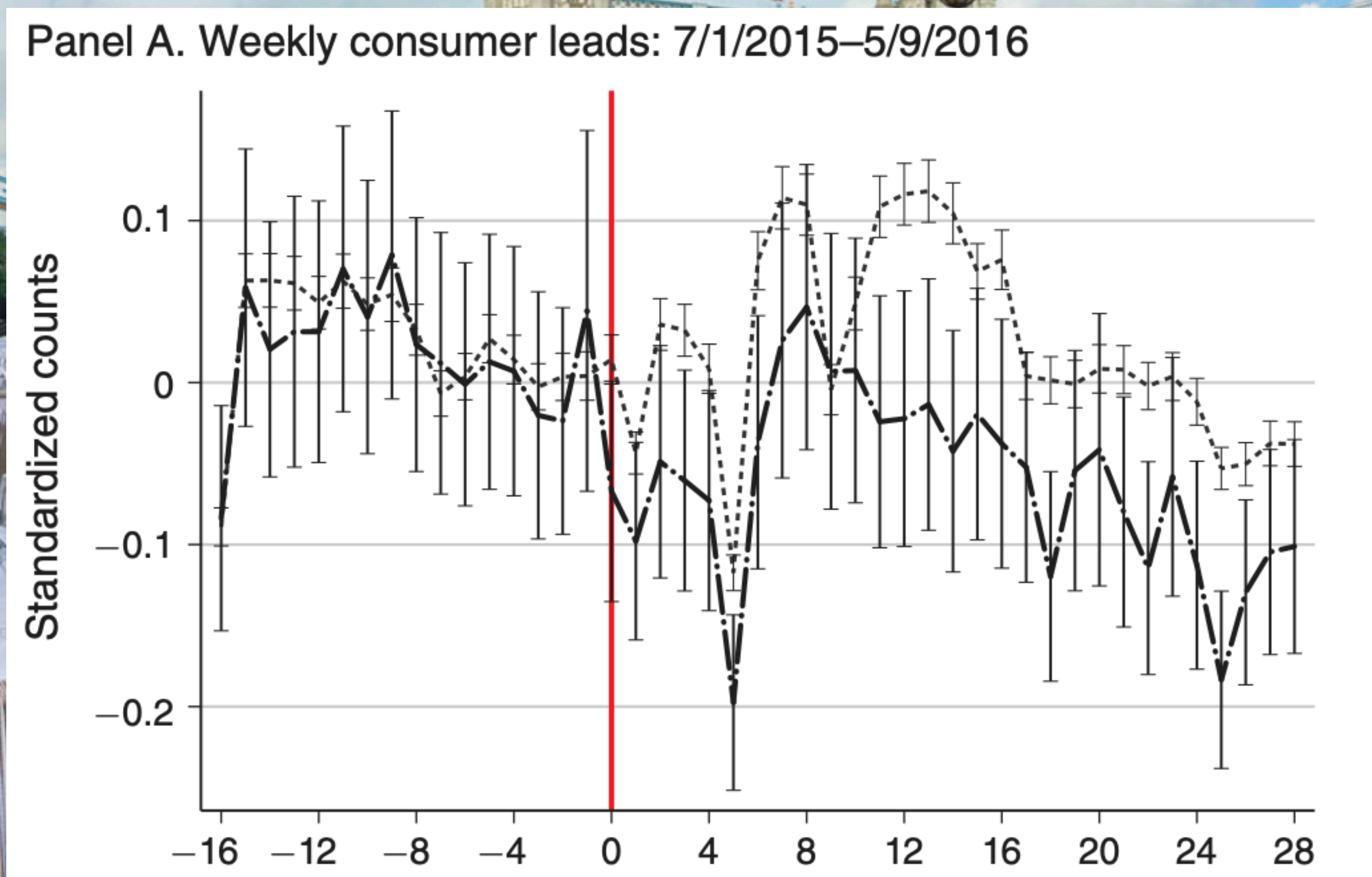
Meta-data



Health Inspections

(Dai & Luca, 2020)

Meta-data

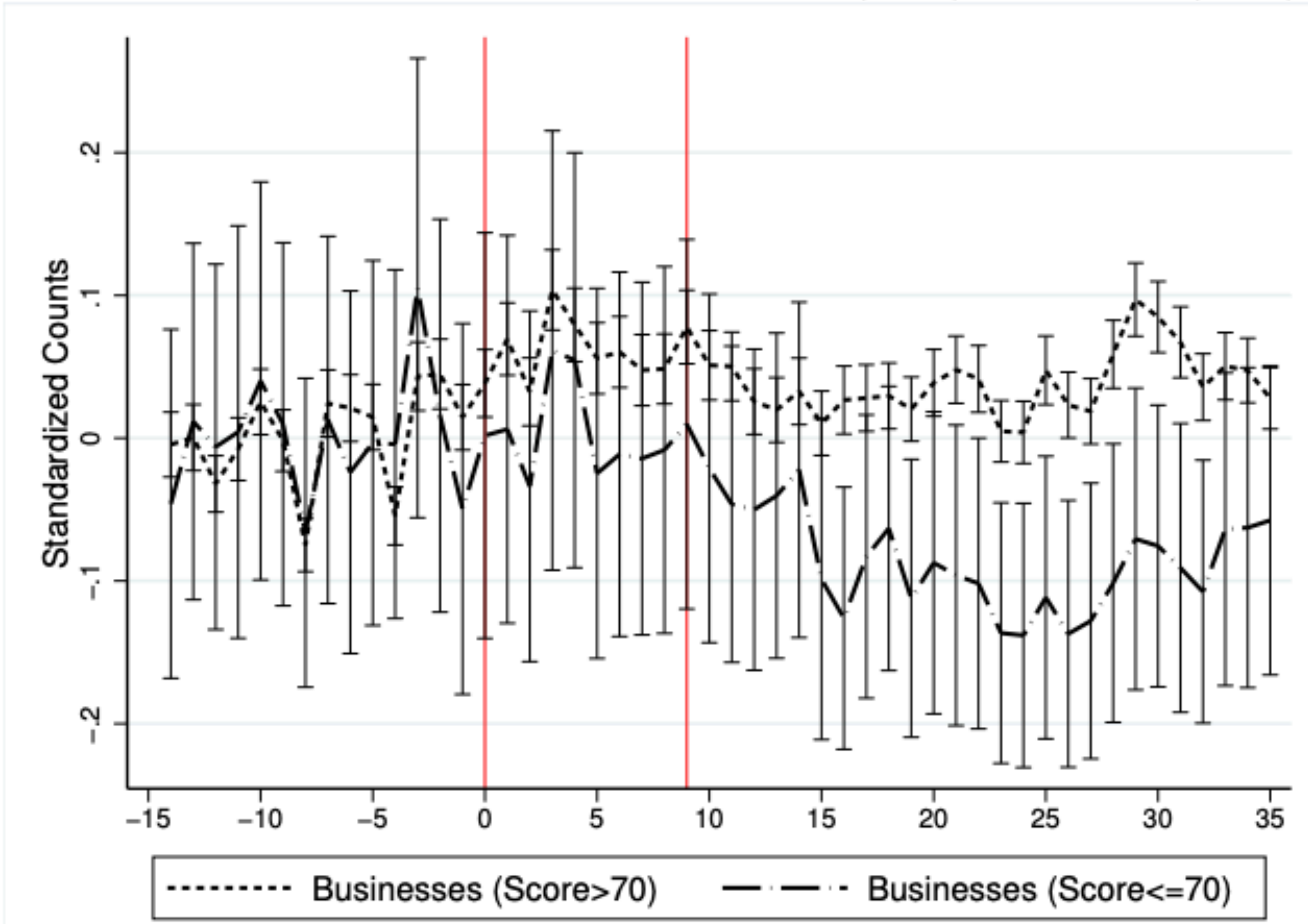


Health Inspections

(Dai & Luca, 2020)

Meta-data

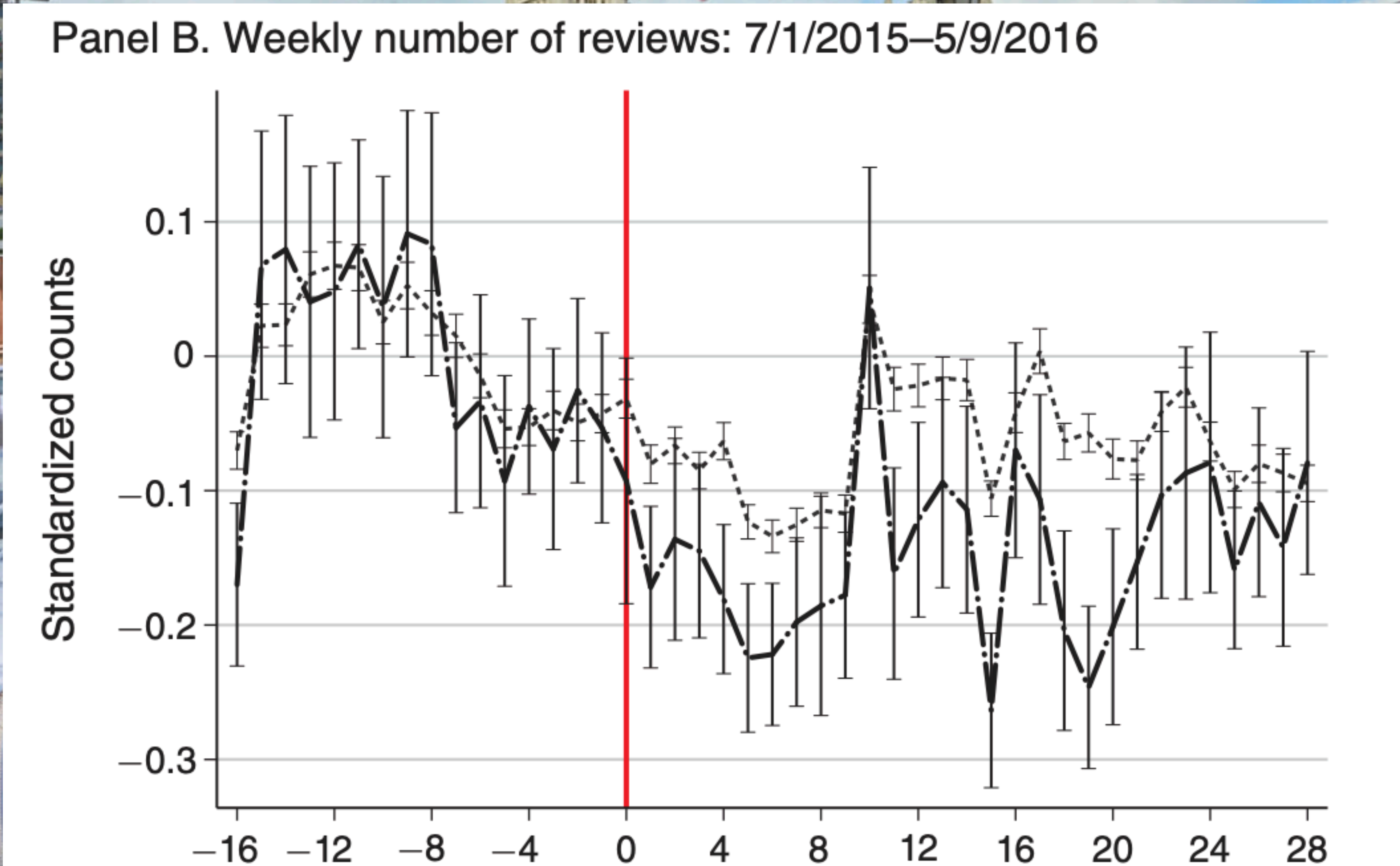
Panel A. Weekly consumer leads: 10/11/2012 - 9/25/2013



Health Inspections

(Kang et al., 2013)

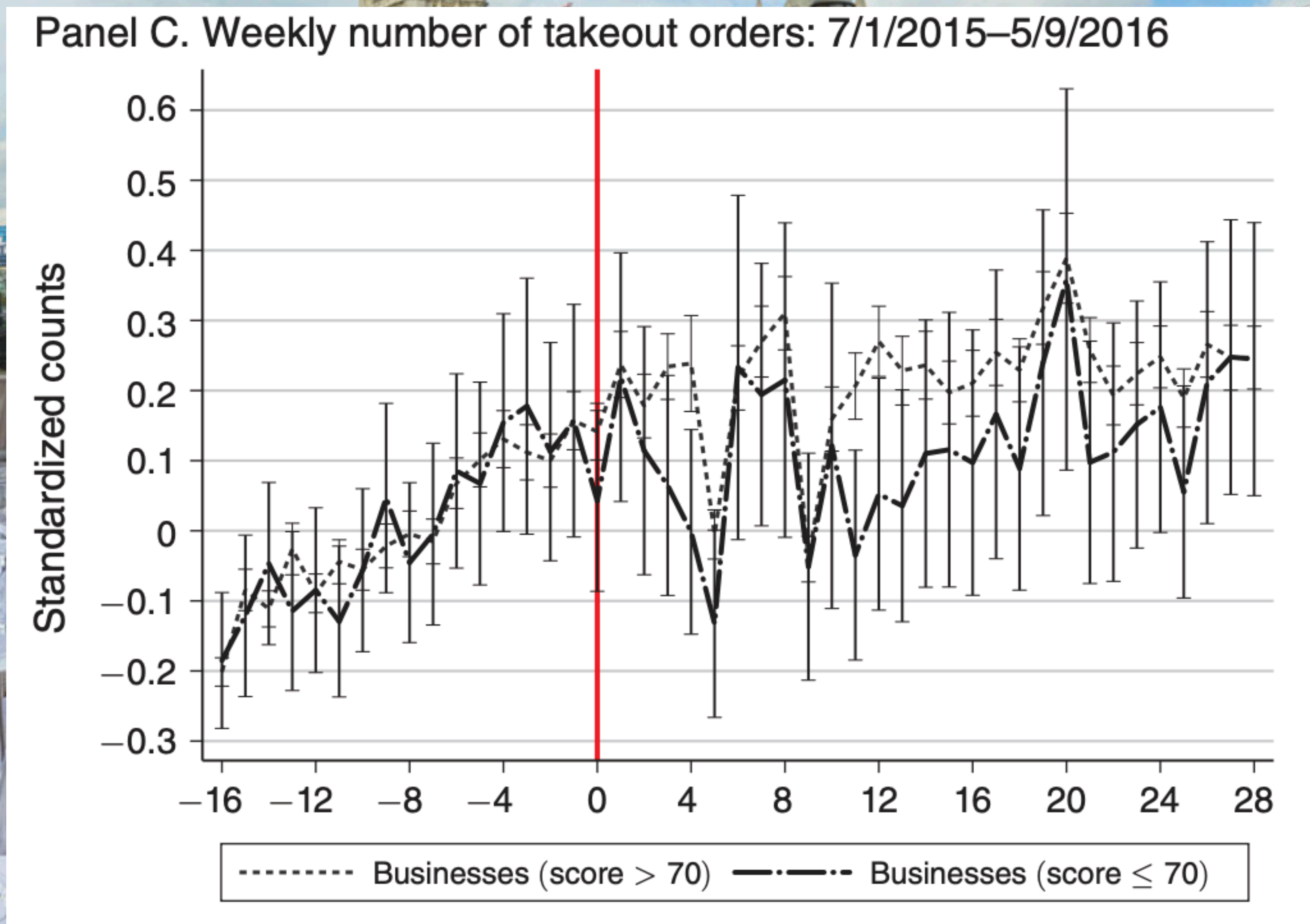
Meta-data



Health Inspections

(Dai & Luca, 2020)

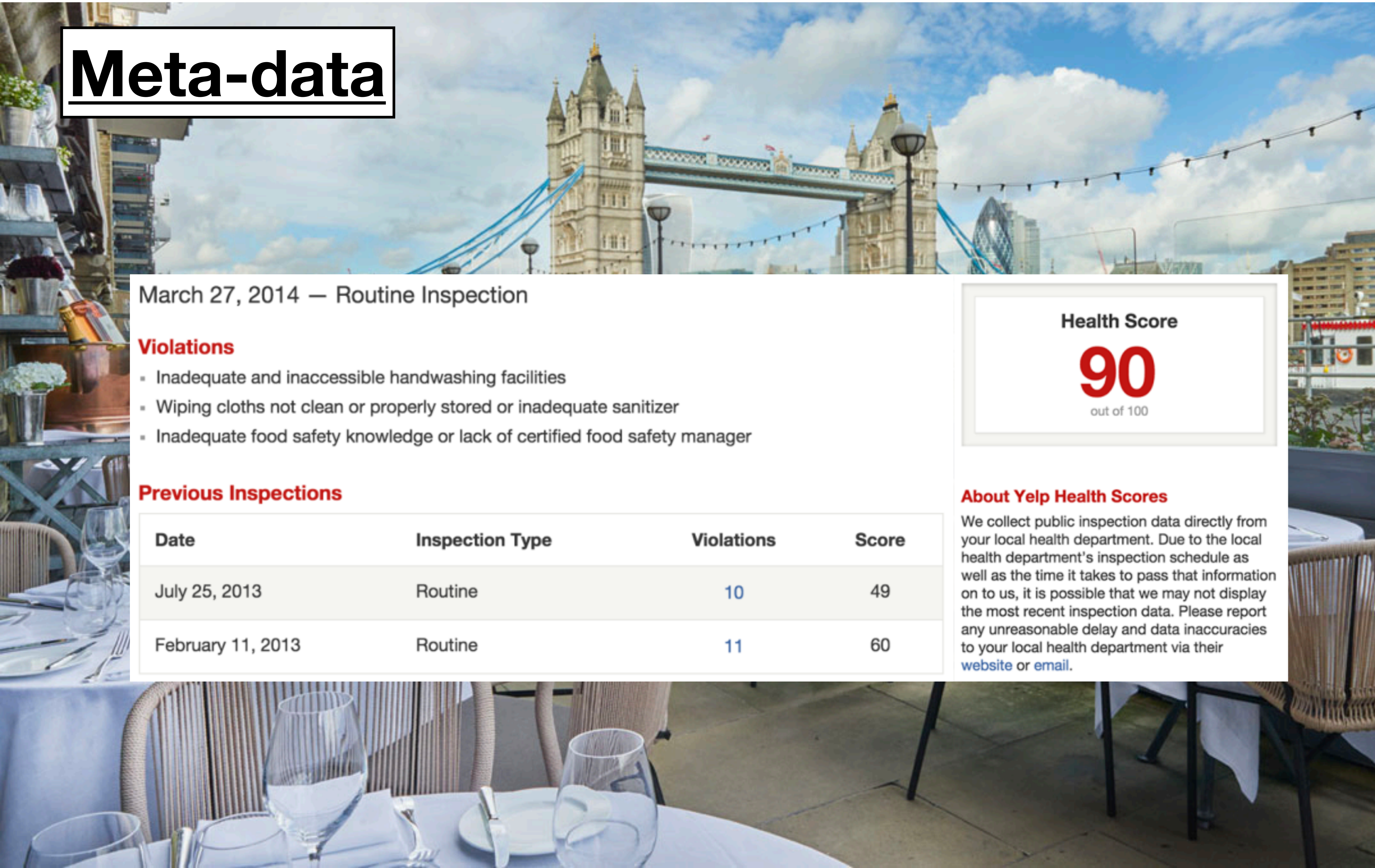
Meta-data



Health Inspections

(Dai & Luca, 2020)

Meta-data



March 27, 2014 — Routine Inspection

Violations

- Inadequate and inaccessible handwashing facilities
- Wiping cloths not clean or properly stored or inadequate sanitizer
- Inadequate food safety knowledge or lack of certified food safety manager

Previous Inspections

Date	Inspection Type	Violations	Score
July 25, 2013	Routine	10	49
February 11, 2013	Routine	11	60

Health Score

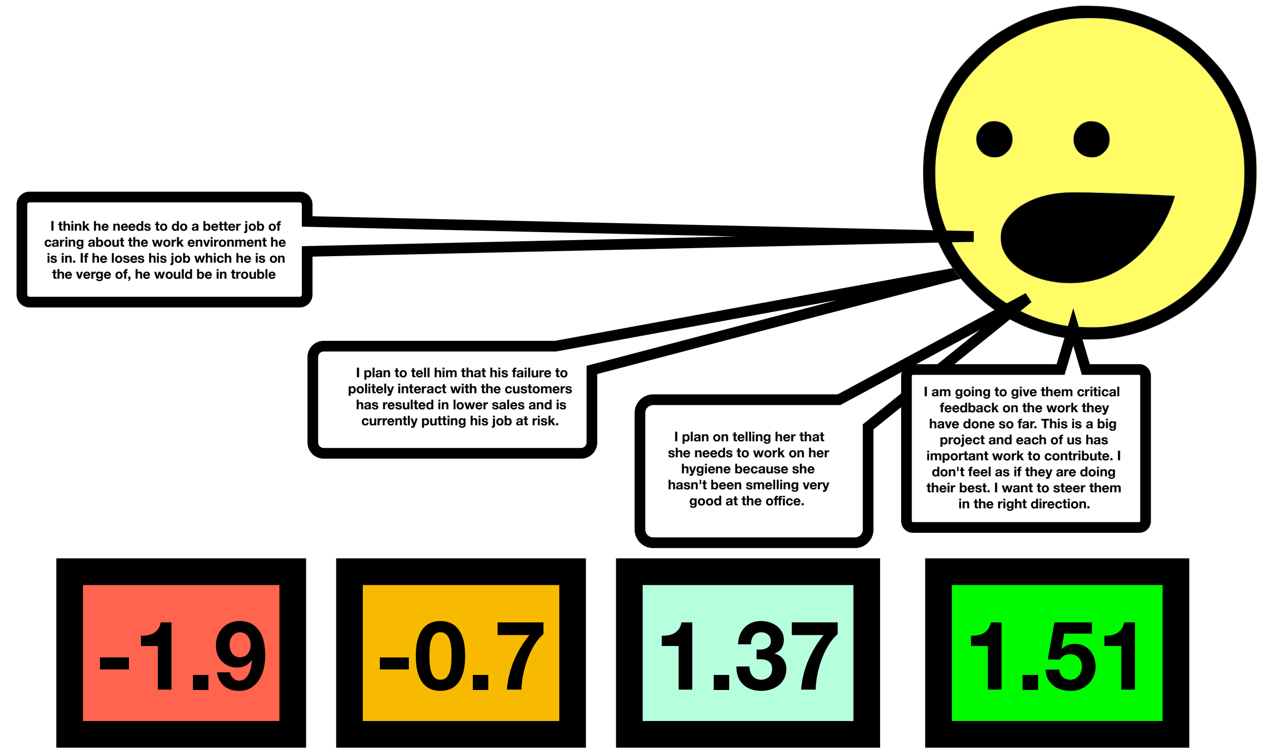
90

out of 100

About Yelp Health Scores

We collect public inspection data directly from your local health department. Due to the local health department's inspection schedule as well as the time it takes to pass that information on to us, it is possible that we may not display the most recent inspection data. Please report any unreasonable delay and data inaccuracies to your local health department via their website or email.

Measurement in Language



Humans: Plusses

What we've always been doing

More accurate than algorithms for complex tasks

Can understand context, nuance

Humans: Minuses

High marginal cost of labor

Not reliable

Not transparent

Goals for NLP in Social Science

Approximate good things about humans

Improve on bad things about humans

Goals for NLP in Social Science

Approximate good things about humans

- Validate measures in existing theoretical/empirical framework
- Measure with roughly the same accuracy as trained RAs
- Account for context-specificity

Goals for NLP in Social Science

Approximate good things about humans

Validate measures in existing theoretical/empirical framework

- define quantities/populations of interest **first**

Goals for NLP in Social Science

Approximate good things about humans

Validate measures in existing theoretical/empirical framework

- define quantities/populations of interest **first**

Measure with roughly the same accuracy as trained RAs

- depends on model… BUT
- use accuracy tools to simulate decisions

Goals for NLP in Social Science

Approximate good things about humans

Validate measures in existing theoretical/empirical framework

- define quantities/populations of interest **first**

Measure with roughly the same accuracy as trained RAs

- depends on model… BUT
- use accuracy tools to simulate decisions

Account for context-specificity

- non-random train/test splits
- “transfer learning” from one context to another

Goals for NLP in Social Science

Improve on bad things about humans

Goals for NLP in Social Science

Improve on bad things about humans

Decrease marginal cost of measurement in new texts

Increase replicability/reliability of measures

Increase interpretability of resulting models

Goals for NLP in Social Science

Improve on bad things about humans

Decrease marginal cost of measurement in new texts

Increase replicability/reliability of measures

- writing our own robust code <u>well</u>

Goals for NLP in Social Science

Improve on bad things about humans

Decrease marginal cost of measurement in new texts

Increase replicability/reliability of measures

- writing our own robust code <u>well</u>

Increase interpretability of resulting models

????

Illusion of Explanatory Depth

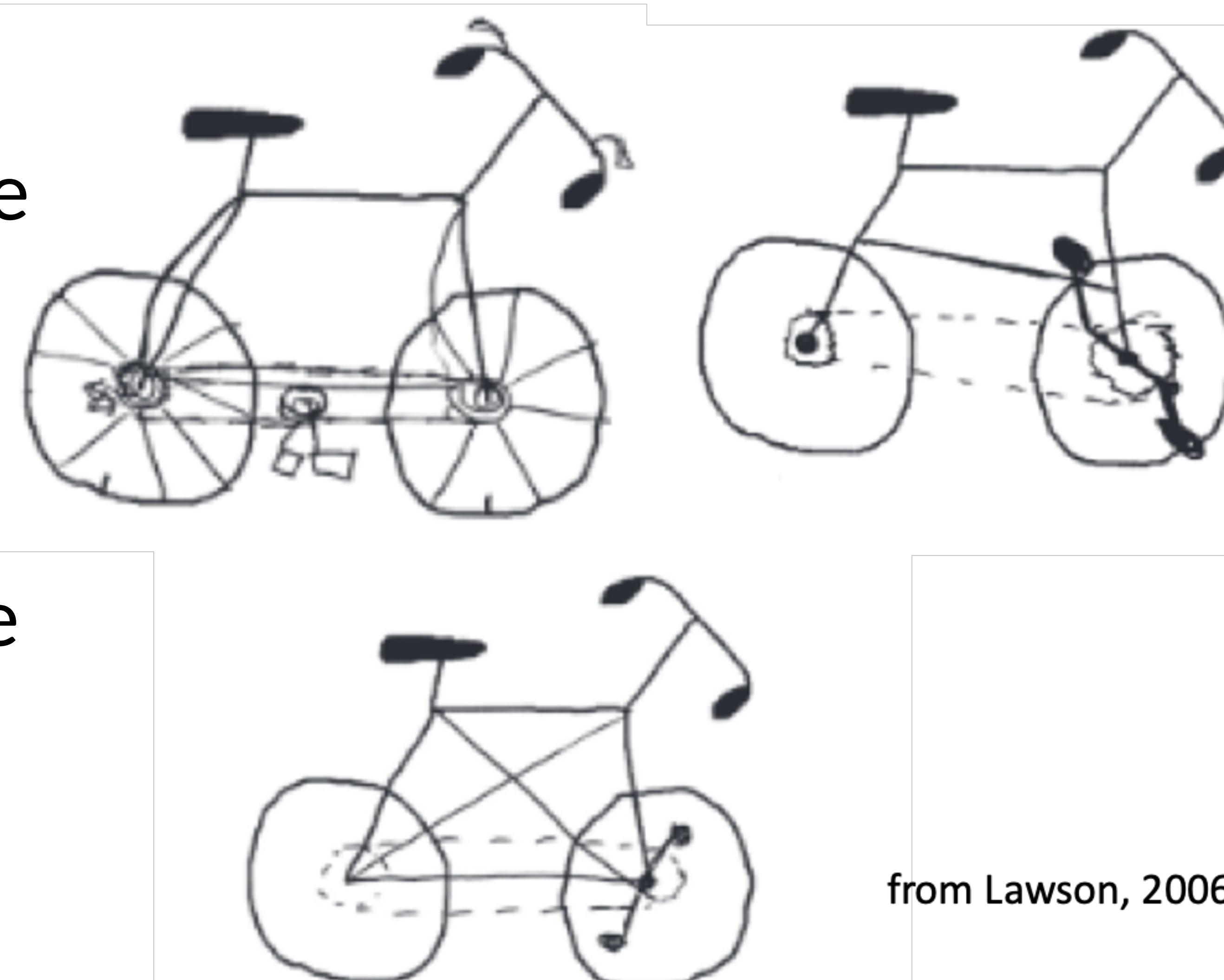
(Rozenblit & Keil, 2002)

“People feel they understand complex phenomena with far greater precision, coherence, and depth than they really do; they are subject to an illusion—an illusion of explanatory depth. The illusion is far stronger for explanatory knowledge than many other kinds of knowledge, […] and is most robust where the environment supports real-time explanations with visible mechanisms”

Illusion of Explanatory Depth

(Rozenblit & Keil, 2002)

“People feel they understand complex phenomena with far greater precision, coherence, and depth than they really do; they are subject to an illusion—an illusion of explanatory depth. The illusion is far stronger for explanatory knowledge than many other kinds of knowledge, […] and is most robust where the environment supports real-time explanations with visible mechanisms”



from Lawson, 2006

Interpretations of Interpretability

Interpretations of Interpretability

Many goals of interpretability

Interpretations of Interpretability

Many goals of interpretability

End users:

- increased trust in outputs

Interpretations of Interpretability

Many goals of interpretability

End users:

- increased trust in outputs
- selective trust in outputs

Interpretations of Interpretability

Many goals of interpretability

End users:

- increased trust in outputs
- selective trust in outputs
- support own learning/exploration

Interpretations of Interpretability

Many goals of interpretability

End users:

- increased trust in outputs
- selective trust in outputs
- support own learning/exploration
- fairness/accountability

Right to an Explanation

1The data subject should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application or e-recruiting practices without any human intervention.

Article 22, Recital 71, of GDPR

Right to an Explanation

1The data subject should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application or e-recruiting practices without any human intervention.

2 [Caveats - if required by contract or government (e.g. fraud detection), or if consent is explicitly given by user]

Article 22, Recital 71, of GDPR

Right to an Explanation

1The data subject should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application or e-recruiting practices without any human intervention.

2 [Caveats - if required by contract or government (e.g. fraud detection), or if consent is explicitly given by user]

4In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, **to obtain an explanation of the decision reached after such assessment** and to challenge the decision.

Article 22, Recital 71, of GDPR

Interpretations of Interpretability

Many goals of interpretability

End users:

- increased trust in outputs
- selective trust in outputs
- support own learning/exploration
- fairness/accountability

Model builders:

- identify room for improvement

Interpretations of Interpretability

Many goals of interpretability

End users:

- increased trust in outputs
- selective trust in outputs
- support own learning/exploration
- fairness/accountability

Model builders:

- identify room for improvement
- intuitions for bounds on generalisation

Interpretations of Interpretability

Many goals of interpretability

End users:

- increased trust in outputs
- selective trust in outputs
- support own learning/exploration
- fairness/accountability

Model builders:

- identify room for improvement
- intuitions for bounds on generalisation
- discovery of “true” world model

Interpretations of Interpretability

Many definitions, no clear answer

(Lipton, 2017)

Interpretations of Interpretability

Many definitions, no clear answer

(Lipton, 2017)

Two broad categories:

Interpretations of Interpretability

Many definitions, no clear answer

(Lipton, 2017)

Two broad categories:

Transparency

"What is the model doing?"

Interpretations of Interpretability

Many definitions, no clear answer

(Lipton, 2017)

Two broad categories:

Transparency

“What is the model doing?”

Post-hoc Explanations

“What else can the model tell me?”

Interpretability as Transparency

Simulatability

“can the model be contemplated all at once?”

Interpretability as Transparency

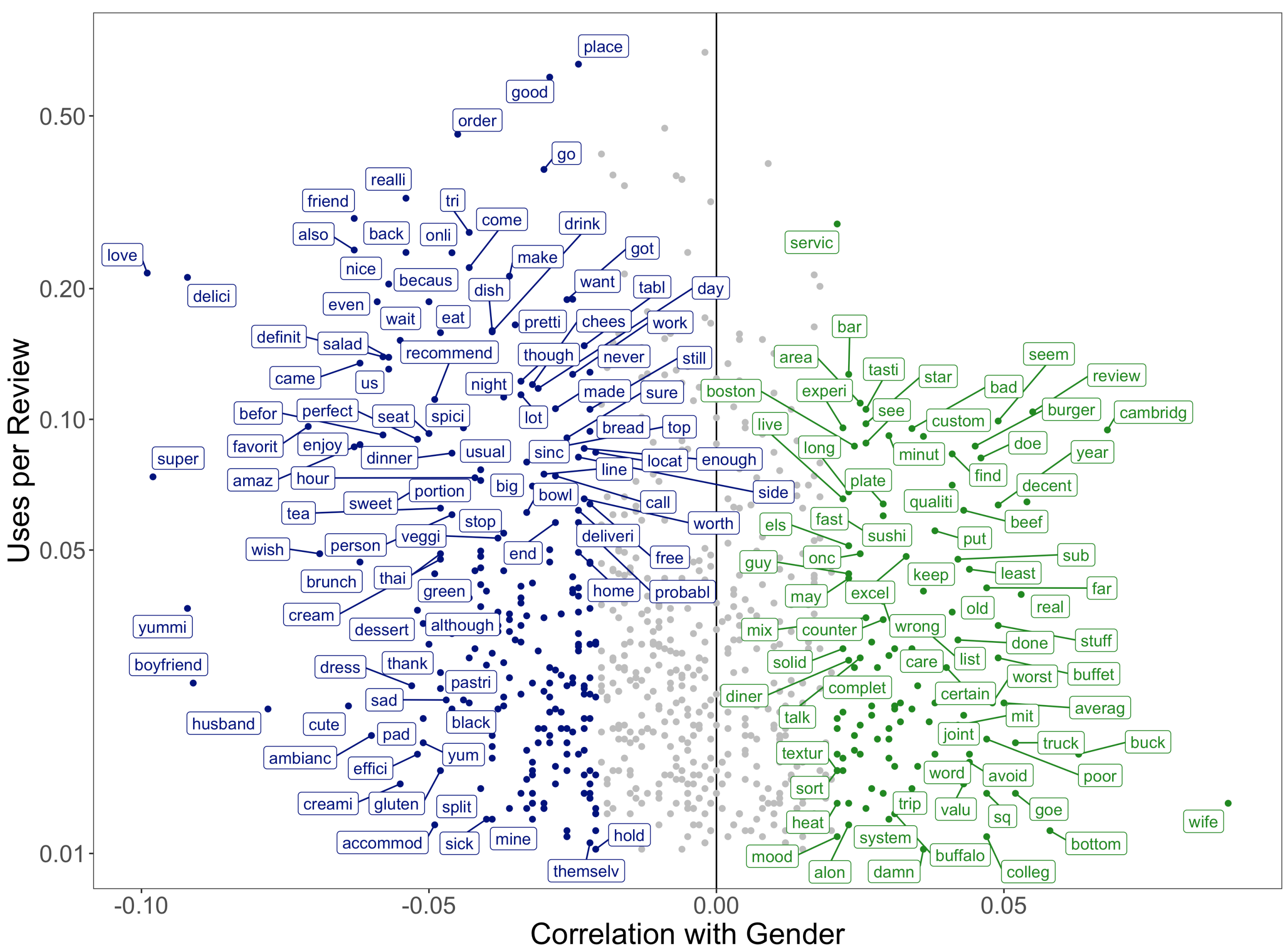
Simulatability

"can the model be contemplated all at once?"

- preference for low number of features
- LASSO: the "bet on sparsity principle"

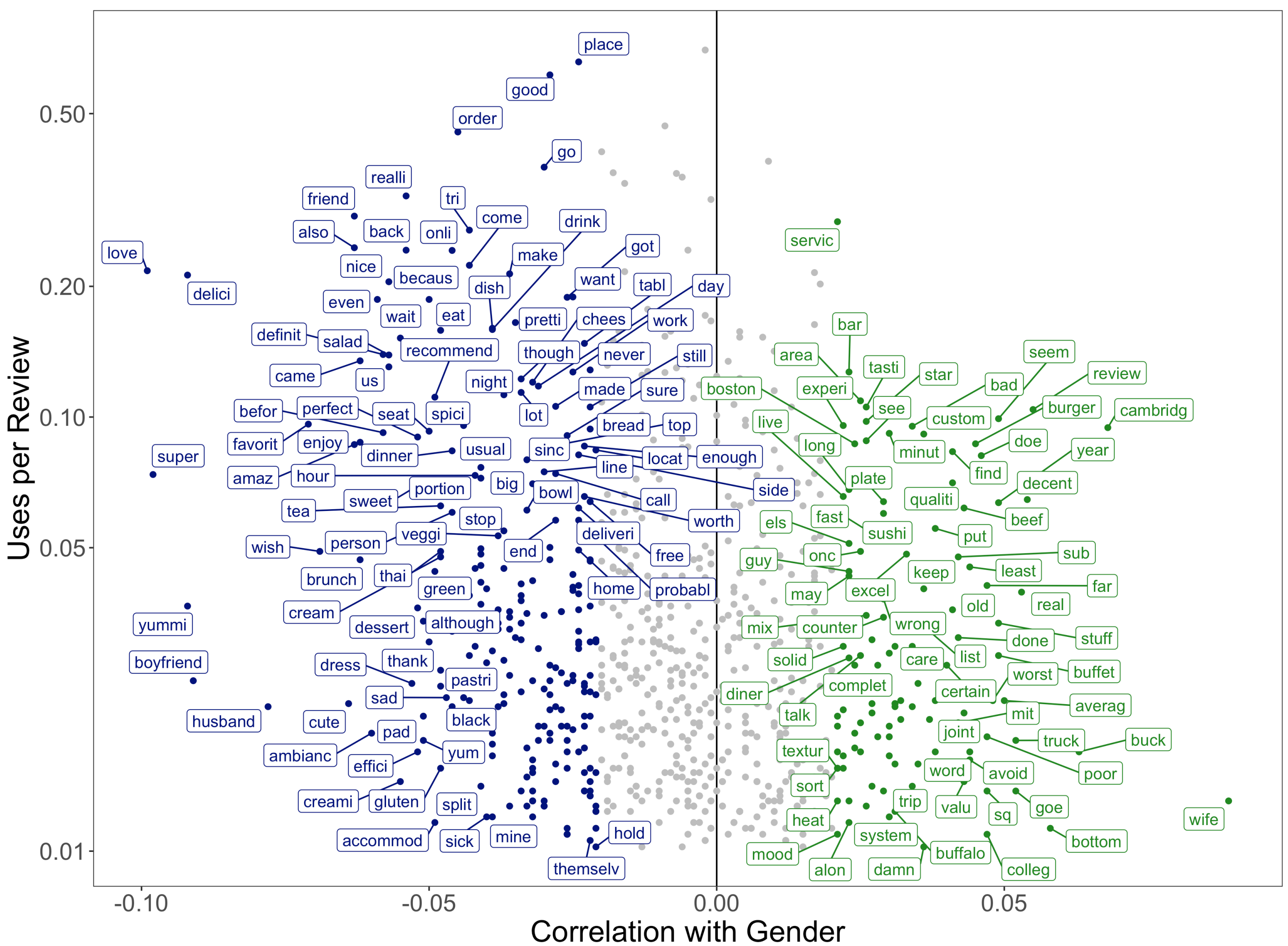
Interpretability as Transparency

Raw Correlation Plot

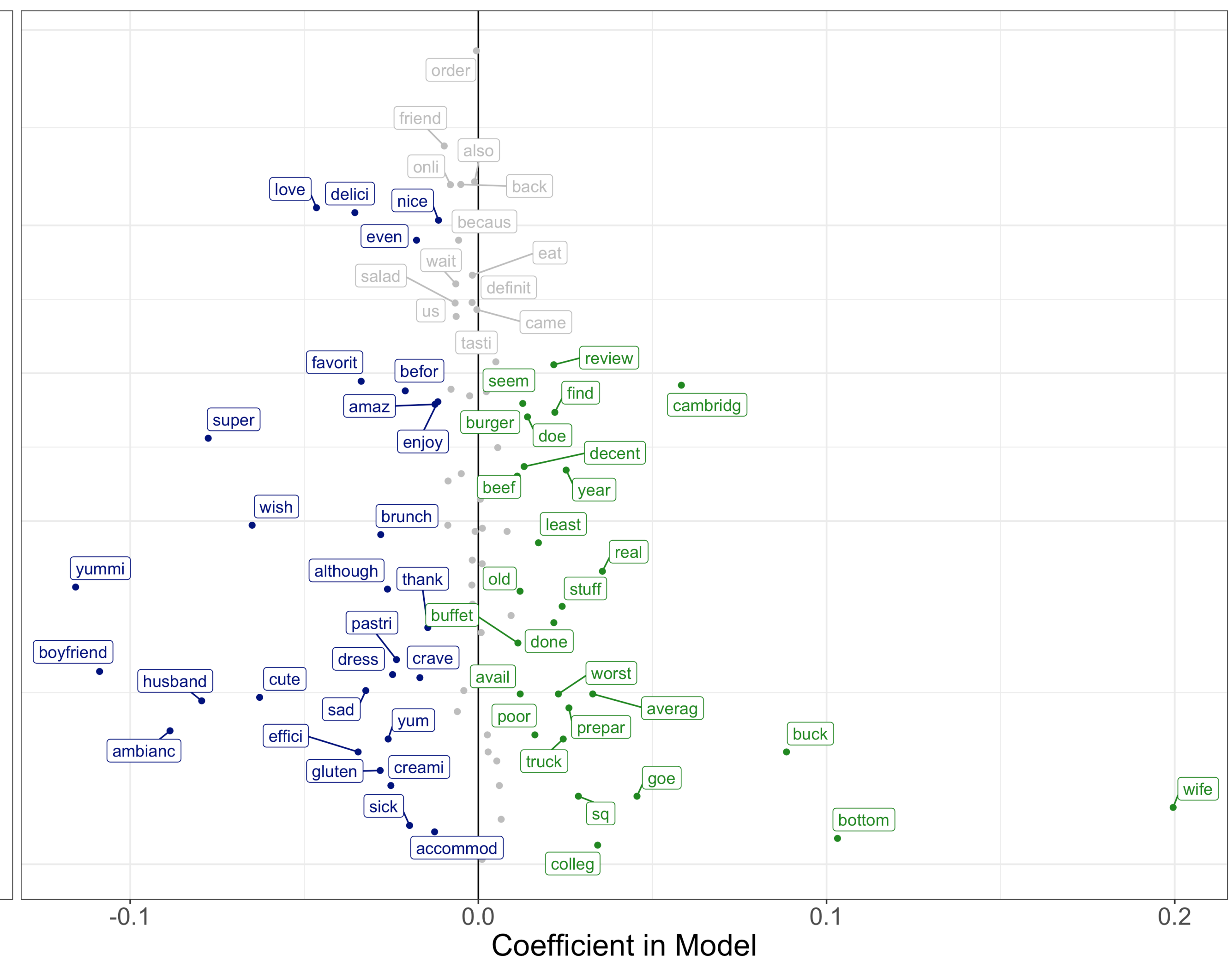


Interpretability as Transparency

Raw Correlation Plot



LASSO Coefficient Plot



Interpretability as Transparency

Simulatability

"can the model be contemplated all at once?"

- preference for low number of features
- LASSO: the "bet on sparsity principle"

Decomposability

"can each input and parameter be described?"

Interpretability as Transparency

Simulatability

"can the model be contemplated all at once?"

- preference for low number of features
- LASSO: the "bet on sparsity principle"

Decomposability

"can each input and parameter be described?"

- preference for clear feature sets
- preference for linear parameters

Interpretability as Transparency

Simulatability

“can the model be contemplated all at once?”

- preference for low number of features
- LASSO: the “bet on sparsity principle”

Decomposability

“can each input and parameter be described?”

- preference for clear feature sets
- preference for linear parameters

Algorithmic Transparency

“can the contents of the model be scrutinised?”

Interpretability as Transparency

Simulatability

"can the model be contemplated all at once?"

- preference for low number of features
- LASSO: the "bet on sparsity principle"

Decomposability

"can each input and parameter be described?"

- preference for clear feature sets
- preference for linear parameters

Algorithmic Transparency

"can the contents of the model be scrutinised?"

- preference for open source code
- preference for well-documented code
- training population is known/public

Model Cards for Model Reporting

Model Card

- **Model Details**. Basic information about the model.
  - Person or organization developing model
  - Model date
  - Model version
  - Model type
  - Information about training algorithms, parameters, fairness constraints or other applied approaches, and features
  - Paper or other resource for more information
  - Citation details
  - License
  - Where to send questions or comments about the model
- **Intended Use**. Use cases that were envisioned during development.
  - Primary intended uses
  - Primary intended users
  - Out-of-scope use cases
- **Factors**. Factors could include demographic or phenotypic groups, environmental conditions, technical attributes, or others listed in Section 4.3.
  - Relevant factors
  - Evaluation factors
- **Metrics**. Metrics should be chosen to reflect potential real-world impacts of the model.
  - Model performance measures
  - Decision thresholds
  - Variation approaches
- **Evaluation Data**. Details on the dataset(s) used for the quantitative analyses in the card.
  - Datasets
  - Motivation
  - Preprocessing
- **Training Data**. May not be possible to provide in practice. When possible, this section should mirror Evaluation Data. If such detail is not possible, minimal allowable information should be provided here, such as details of the distribution over various factors in the training datasets.
- **Quantitative Analyses**
  - Unitary results
  - Intersectional results
- **Ethical Considerations**
- **Caveats and Recommendations**

(Mitchell et al., 2018)

Interpretability as Explanation

Interpretability as Explanation

Text Explanation

- Describe what latent variables the selected features represent

Interpretability as Explanation

Text Explanation

- Describe what latent variables the selected features represent

Visualisation

- Use a plot to represent the weights on selected features

Interpretability as Explanation

Text Explanation

- Describe what latent variables the selected features represent

Visualisation

- Use a plot to represent the weights on selected features

Local Explanation

- Train a simple model to predict a more complex model

Interpretability as Explanation

Text Explanation

- Describe what latent variables the selected features represent

Visualisation

- Use a plot to represent the weights on selected features

Local Explanation

- Train a simple model to predict a more complex model

Explanation by Example

- Select example data that represents typical model output

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

Categorical data: confusion matrix

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

Categorical data: confusion matrix

Labels

Predictions

	Mexican	Chinese	Indian
Mexican			
Chinese			
Indian			

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

Categorical data: confusion matrix

Labels

Predictions		Mexican	Chinese	Indian
	Mexican	401	19	80
	Chinese	55	212	101
	Indian	41	207	367

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

Categorical data: confusion matrix

Labels

Predictions

	Mexican	Chinese	Indian
Mexican	401	19	80
Chinese	55	212	101
Indian	41	207	367

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

Categorical data: confusion matrix

Labels

Predictions

	Mexican	Chinese	Indian
Mexican	401	19	80
Chinese	55	212	101
Indian	41	207	367

Continuous data: quantile table

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

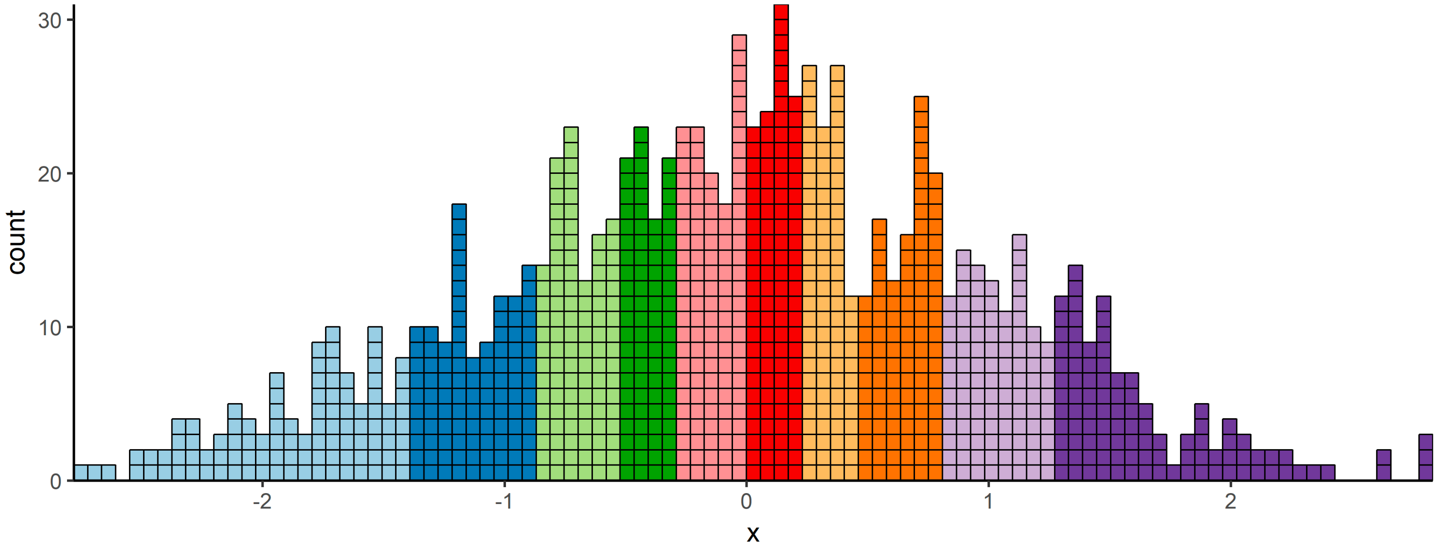
Categorical data: confusion matrix

Labels

Predictions

	Mexican	Chinese	Indian
Mexican	401	19	80
Chinese	55	212	101
Indian	41	207	367

Continuous data: quantile table



Related: Error Analysis

Search for examples of model errors -> fix, if possible!

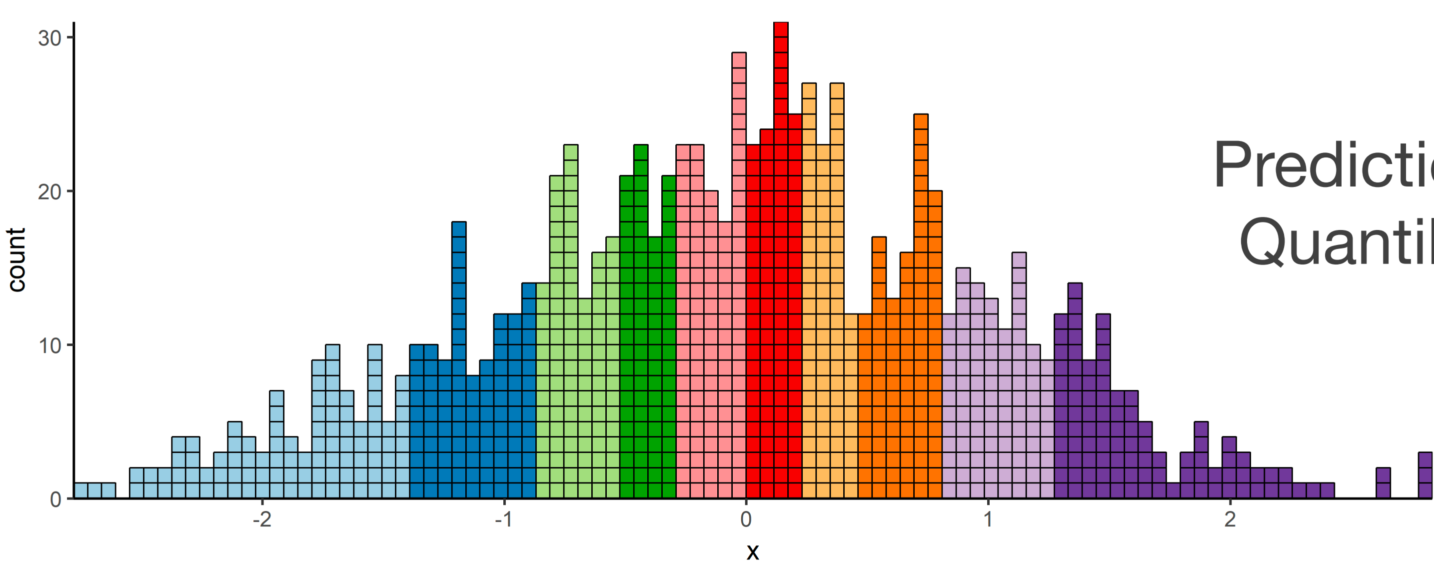
Categorical data: confusion matrix

Labels

Predictions

	Mexican	Chinese	Indian
Mexican	401	19	80
Chinese	55	212	101
Indian	41	207	367

Continuous data: quantile table



Outcome Quantile

Prediction Quantile

	0-20	20-40	40-60	60-80	80-100
0-20					
20-40					
40-60					
60-80					
80-10					

Related: Error Analysis

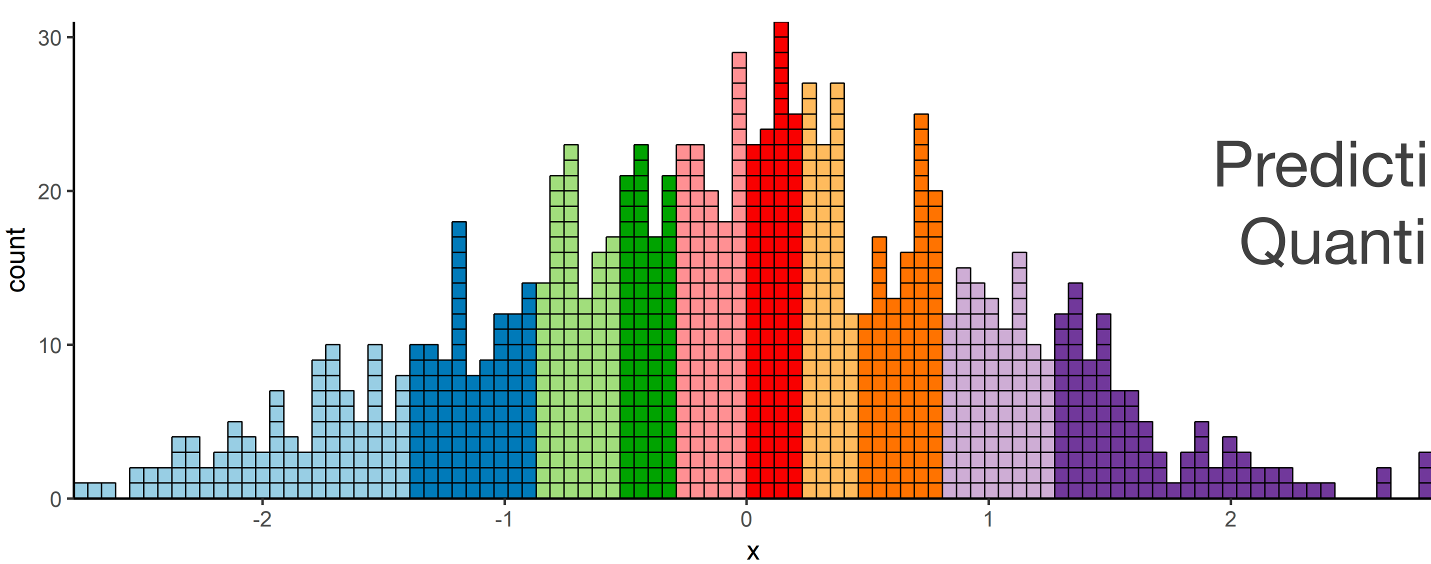
Search for examples of model errors -> fix, if possible!

Categorical data: confusion matrix

Labels

Predictions	Mexican	Chinese	Indian
Mexican	401	19	80
Chinese	55	212	101
Indian	41	207	367

Continuous data: quantile table



Outcome Quantile

Prediction Quantile	0-20	20-40	40-60	60-80	80-100
0-20	401	19	34	19	80
20-40	55	212	119	55	130
40-60	32	121	441	64	68
60-80	50	40	88	280	43
80-10	41	27	73	76	367

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

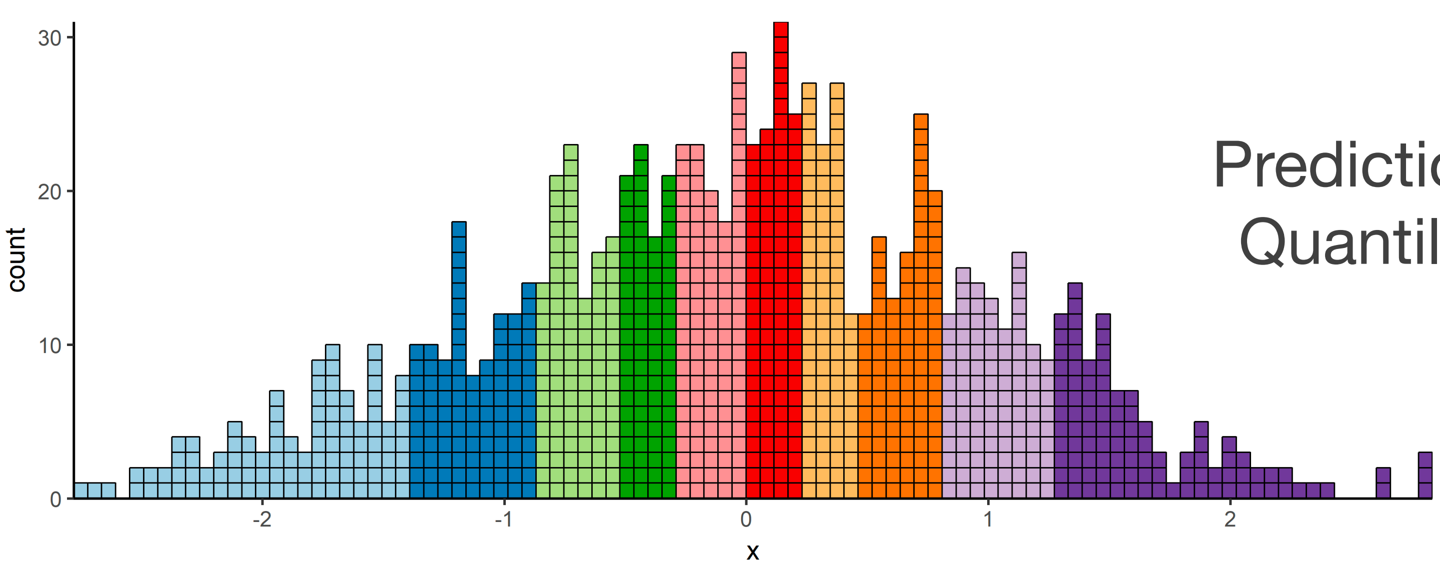
Categorical data: confusion matrix

Labels

Predictions

	Mexican	Chinese	Indian
Mexican	401	19	80
Chinese	55	212	101
Indian	41	207	367

Continuous data: quantile table



Outcome Quantile

Prediction Quantile

	0-20	20-40	40-60	60-80	80-100
0-20	401	19	34	19	80
20-40	55	212	119	55	130
40-60	32	121	441	64	68
60-80	50	40	88	280	43
80-10	41	27	73	76	367

Related: Error Analysis

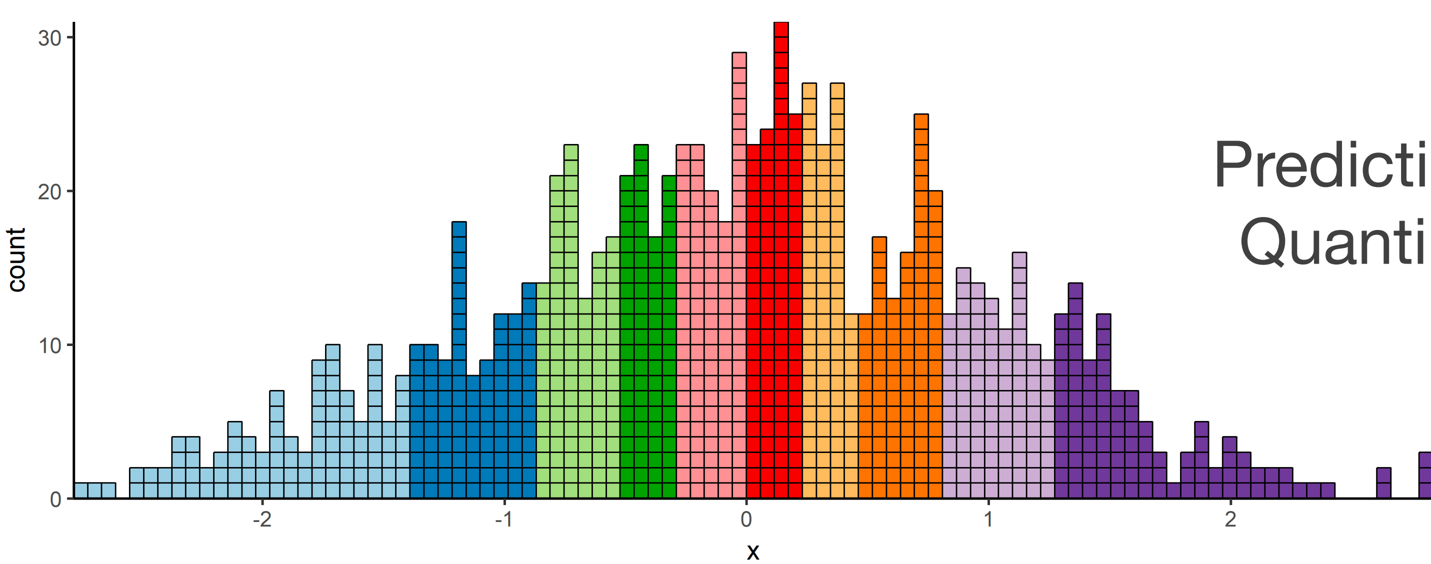
Search for examples of model errors -> fix, if possible!

Error Analysis

Labels

Predictions

	Mexican	Chinese	Indian
Mexican	401	19	80
Chinese	55	212	101
Indian	41	207	367



Outcome Quantile

Prediction Quantile

	0-20	20-40	40-60	60-80	80-100
0-20	401	19	34	19	80
20-40	55	212	119	55	130
40-60	32	121	441	64	68
60-80	50	40	88	280	43
80-10	41	27	73	76	367

Related: Error Analysis

Search for examples of model errors -> fix, if possible!

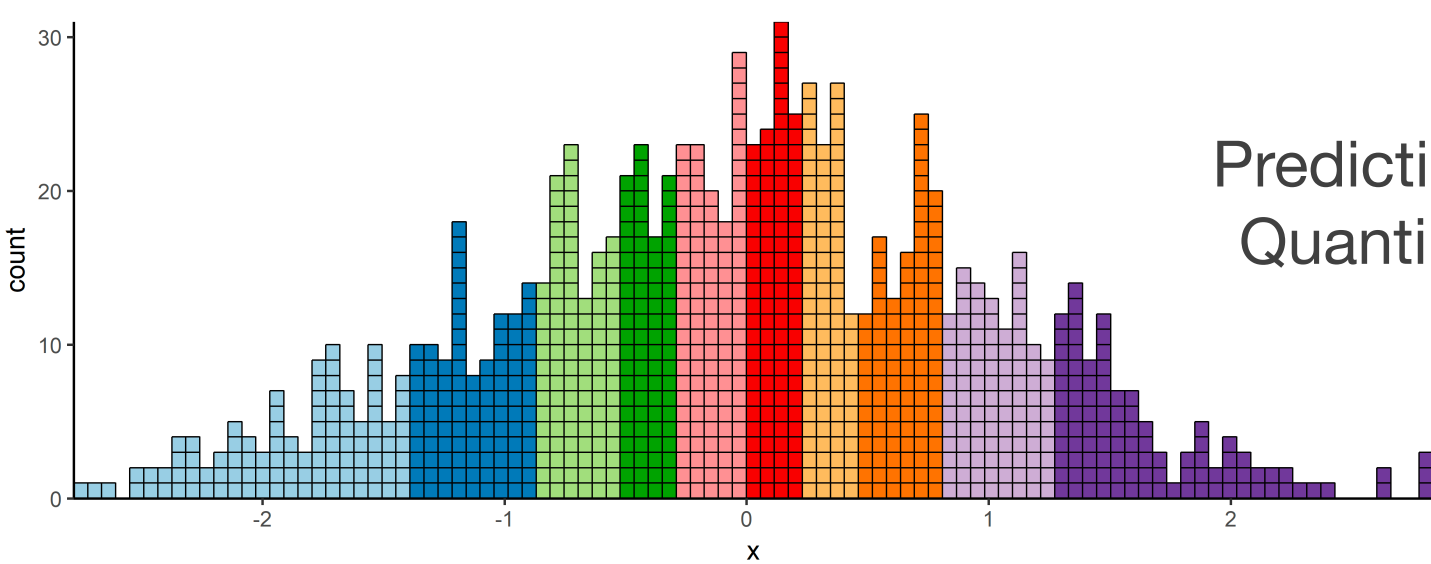
Error Analysis

Explanation By Example

Labels

Predictions

	Mexican	Chinese	Indian
Mexican	401	19	80
Chinese	55	212	101
Indian	41	207	367



Outcome Quantile

Prediction Quantile

	0-20	20-40	40-60	60-80	80-100
0-20	401	19	34	19	80
20-40	55	212	119	55	130
40-60	32	121	441	64	68
60-80	50	40	88	280	43
80-10	41	27	73	76	367

Interpretability as Explanation

Text Explanation
- Describe what latent variables the selected features represent

Visualisation
- Use a plot to represent the weights on selected features

Local Explanation
- Train a simple model to predict a more complex model

Explanation by Example
- Select example data that represents typical model output

Explanation by Comparison
- Compare performance to benchmark model

Benchmarking

Is 60% accurate?

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

Bottom-up: How well can a simple model perform?

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

Bottom-up: How well can a simple model perform?

- random predictions

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

<u>Bottom-up:</u> How well can a simple model perform?

- random predictions
- same prediction for everything

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

<u>Bottom-up:</u> How well can a simple model perform?

- random predictions
- same prediction for everything
- word count

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

<u>Bottom-up:</u> How well can a simple model perform?

- random predictions
- same prediction for everything
- word count
- complexity (e.g. syllables/word, words/sentence)

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

Bottom-up: How well can a simple model perform?
- random predictions
- same prediction for everything
- word count
- complexity (e.g. syllables/word, words/sentence)
- sentiment (# positive words - # negative words)

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

<u>Bottom-up:</u> How well can a simple model perform?

- random predictions
- same prediction for everything
- word count
- complexity (e.g. syllables/word, words/sentence)
- sentiment (# positive words - # negative words)
- other dictionaries (e.g. % of words about food)

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

Bottom-up: How well can a simple model perform?

- random predictions
- same prediction for everything
- word count
- complexity (e.g. syllables/word, words/sentence)
- sentiment (# positive words - # negative words)
- other dictionaries (e.g. % of words about food)

Top-down: How well can a similar model perform?

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

Bottom-up: How well can a simple model perform?

- random predictions
- same prediction for everything
- word count
- complexity (e.g. syllables/word, words/sentence)
- sentiment (# positive words - # negative words)
- other dictionaries (e.g. % of words about food)

Top-down: How well can a similar model perform?

- State of the Art leaderboards

Benchmarking

Is 60% accurate? FTSE returns vs. Siri commands

Two approaches: **Bottom-up** vs. **Top-down**

<u>Bottom-up:</u> How well can a simple model perform?
- random predictions
- same prediction for everything
- word count
- complexity (e.g. syllables/word, words/sentence)
- sentiment (# positive words - # negative words)
- other dictionaries (e.g. % of words about food)

<u>Top-down:</u> How well can a similar model perform?
- State of the Art leaderboards
- Ablation test: remove features and re-evaluate

Final Note on Interpretation

Three types of estimates

Detection

Prediction

Causation

Final Note on Interpretation

Know what it is you are trying to interpret

Quantity of interest & Population of interest

Three types of estimates

Detection

Prediction

Causation

Final Note on Interpretation

Know what it is you are trying to interpret

Quantity of interest & Population of interest

Three types of estimates

Detection

Prediction

Causation

Know what it is you are trying to interpret Quantity of interest & Population of interest Three types of estimates Detection Revealing latent variable e.g. messy kitchens Prediction Anticipating future state e.g. failed inspectio Causation Identifying text-induced future state e.g. reduced customer f

Final Note on Interpretation

Know what it is you are trying to interpret

Quantity of interest & Population of interest

Three types of estimates

Detection Revealing latent variable

e.g. messy kitchens

Prediction

Causation

Know what it is you are trying to interpret Quantity of interest & Population of interest Three types of estimates Detection Revealing latent variable e.g. messy kitchens Prediction Anticipating future state e.g. failed inspectio Causation Identifying text-induced future state e.g. reduced customer f

Final Note on Interpretation

Know what it is you are trying to interpret

Quantity of interest & Population of interest

Three types of estimates

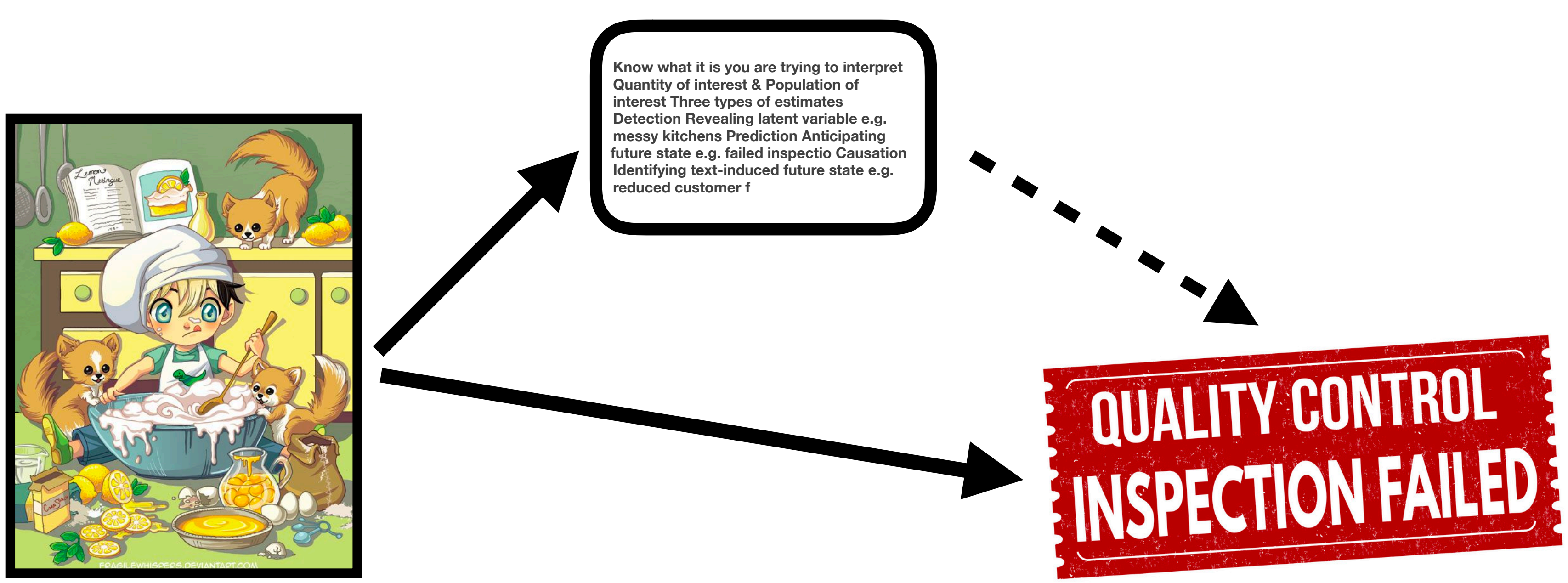
Detection Revealing latent variable

e.g. messy kitchens

Prediction Anticipating future state

e.g. failed inspection

Causation



Final Note on Interpretation

Know what it is you are trying to interpret

Quantity of interest & Population of interest

Three types of estimates

Detection Revealing latent variable

e.g. messy kitchens

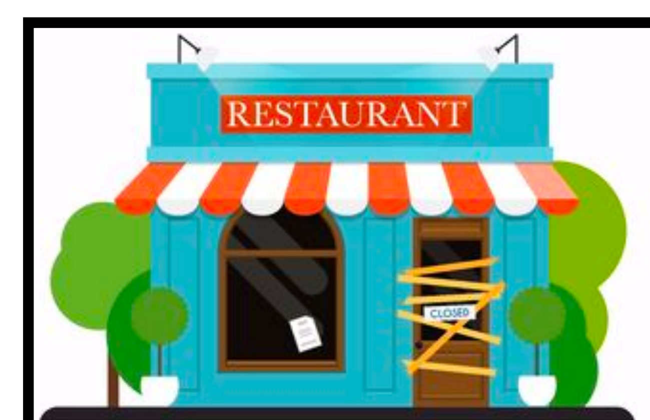
Prediction Anticipating future state

e.g. failed inspection

Causation Identifying text-induced future state

e.g. reduced customer flow

Know what it is you are trying to interpret Quantity of interest & Population of interest Three types of estimates Detection Revealing latent variable e.g. messy kitchens Prediction Anticipating future state e.g. failed inspectio Causation Identifying text-induced future state e.g. reduced customer f



Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Treat individual features as *dimensions* in a feature space

Clustering Documents

Single-Member Model

Each document belongs to a single cluster

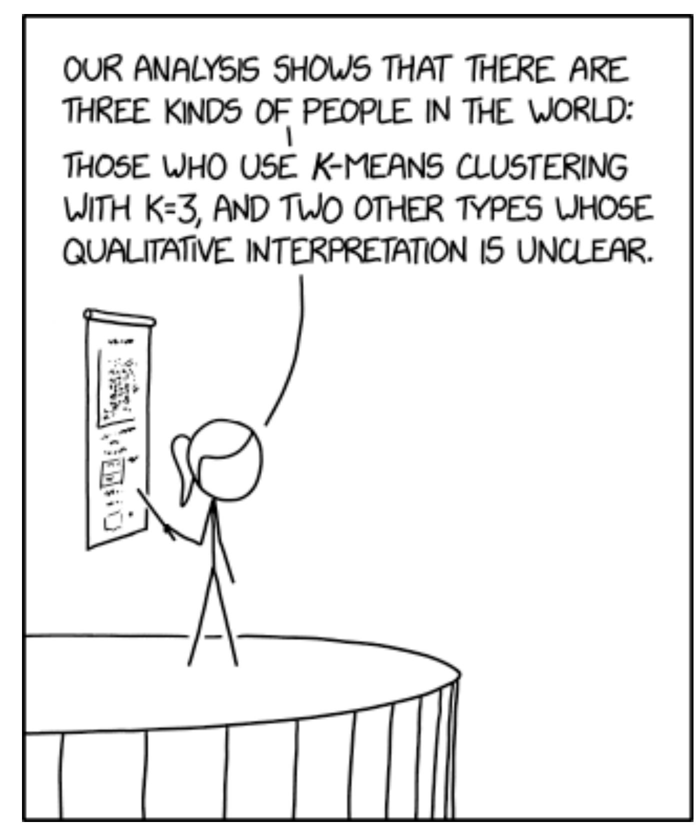
e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Treat individual features as *dimensions* in a feature space

k-means clustering algorithm

assign clusters to minimize intra-cluster distance



Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Treat individual features as *dimensions* in a feature space

k-means clustering algorithm

assign clusters to minimize intra-cluster distance

"This is bad. So bad."

"This is good"

"That is bad, very bad"

Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Treat individual features as *dimensions* in a feature space

k-means clustering algorithm

assign clusters to minimize intra-cluster distance

"This is bad. So bad."

"This is good"

"That is bad, very bad"

this	is	bad	good	that	very
1	1	2	0	0	0
1	1	0	1	0	0
0	1	2	0	1	1

Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Treat individual features as *dimensions* in a feature space

k-means clustering algorithm

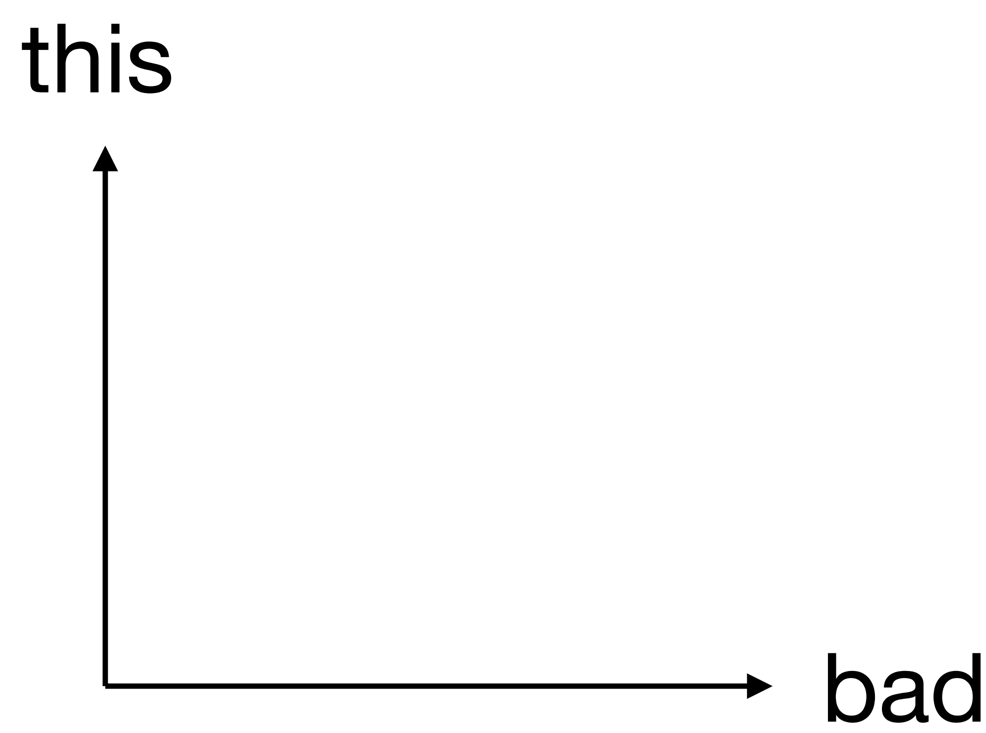
assign clusters to minimize intra-cluster distance

"This is bad. So bad."

"This is good"

"That is bad, very bad"

this	is	bad	good	that	very
1	1	2	0	0	0
1	1	0	1	0	0
0	1	2	0	1	1



Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Treat individual features as *dimensions* in a feature space

k-means clustering algorithm

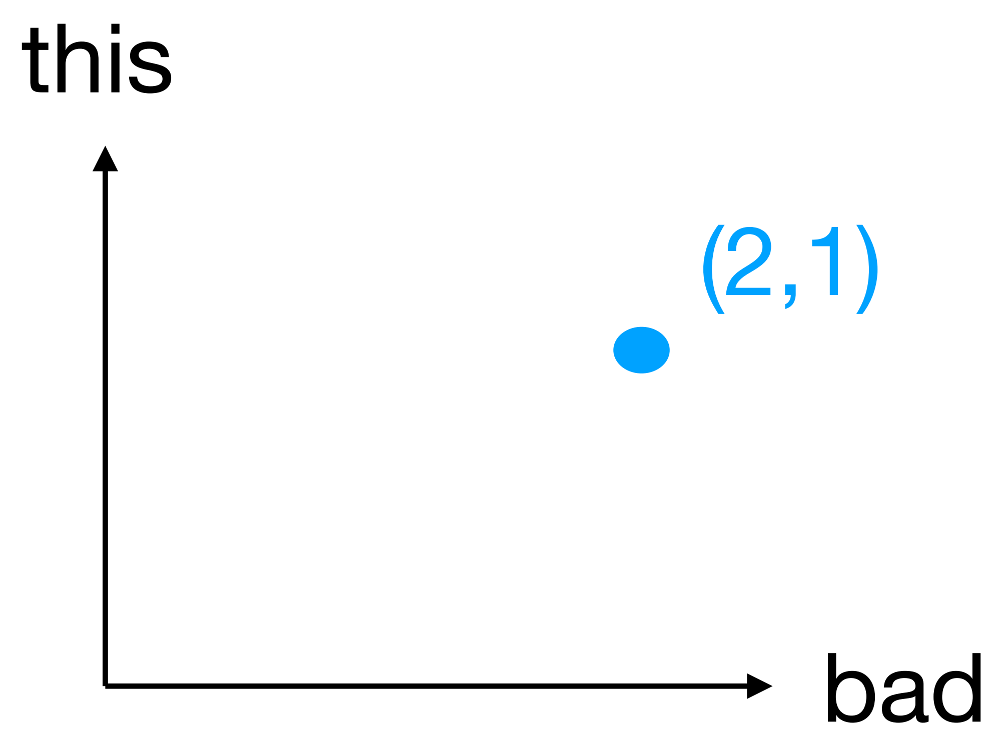
assign clusters to minimize intra-cluster distance

"This is bad. So bad."

"This is good"

"That is bad, very bad"

this	is	bad	good	that	very
1	1	2	0	0	0
1	1	0	1	0	0
0	1	2	0	1	1



Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Treat individual features as *dimensions* in a feature space

k-means clustering algorithm

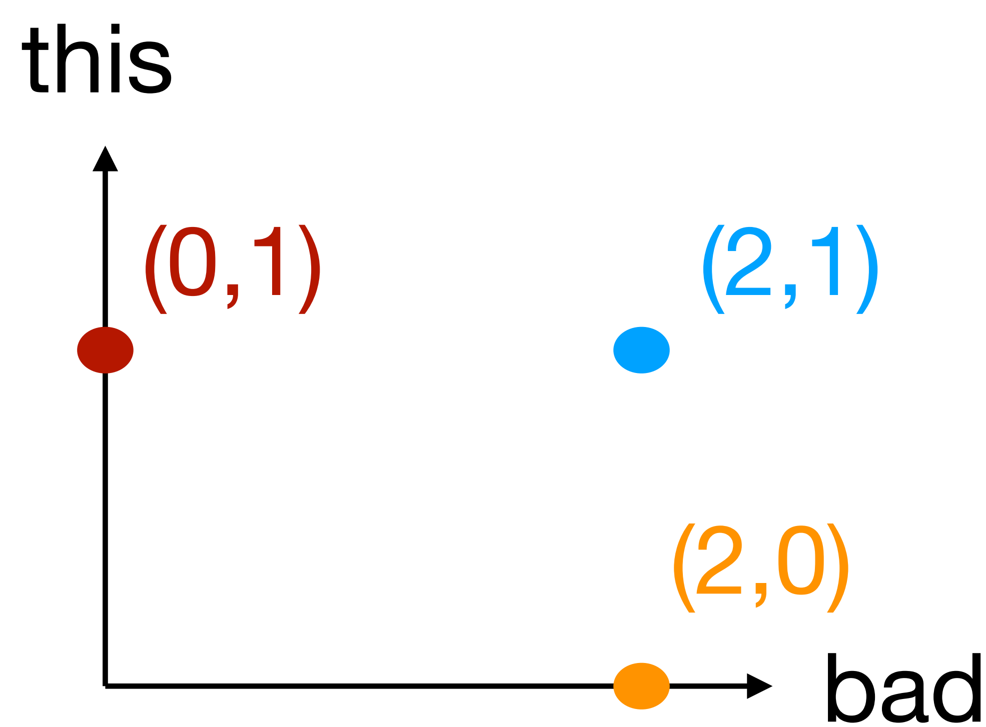
assign clusters to minimize intra-cluster distance

"This is bad. So bad."

"This is good"

"That is bad, very bad"

this	is	bad	good	that	very
1	1	2	0	0	0
1	1	0	1	0	0
0	1	2	0	1	1



Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents

Treat individual features as *dimensions* in a feature space

k-means clustering algorithm

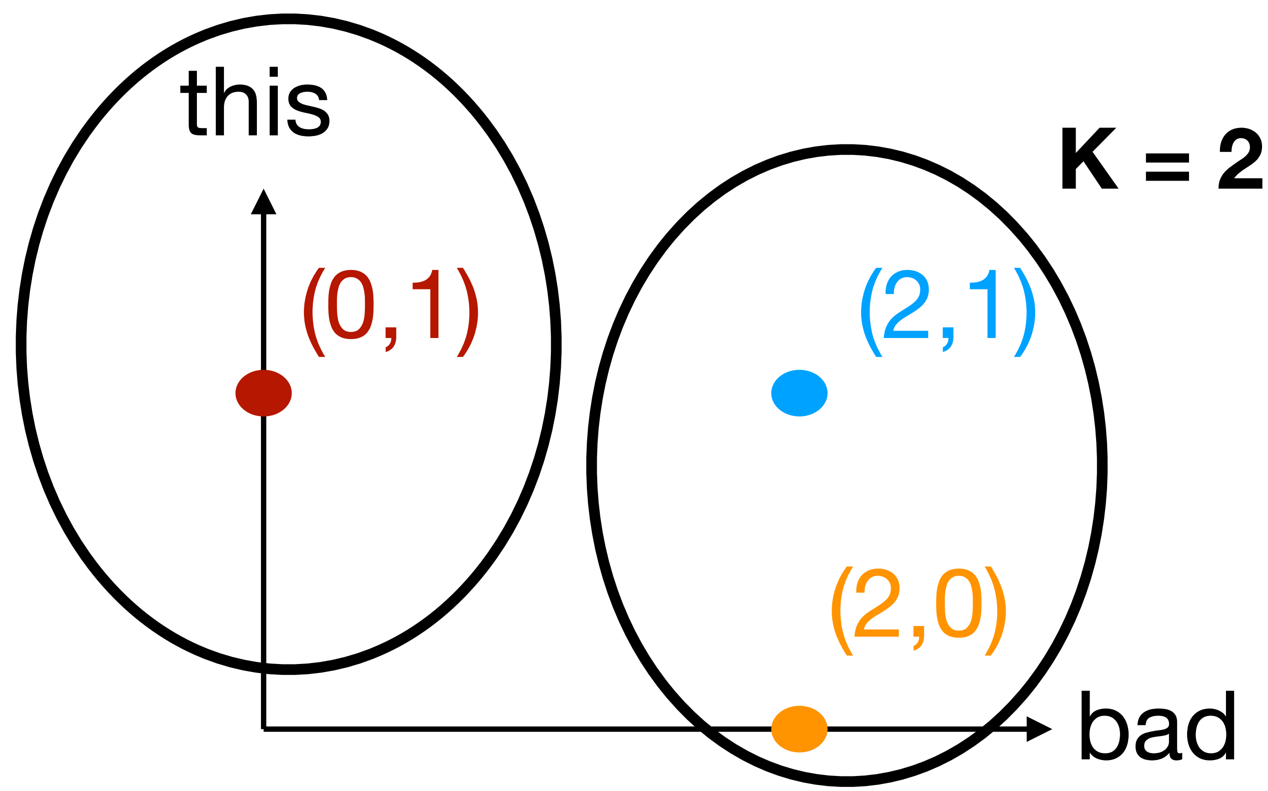
assign clusters to minimize intra-cluster distance

"This is bad. So bad."

"This is good"

"That is bad, very bad"

this	is	bad	good	that	very
1	1	2	0	0	0
1	1	0	1	0	0
0	1	2	0	1	1



Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Supervised approach: if you have labeled documents

Treat labels as outcome of regression problem

Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Supervised approach: if you have labeled documents

Treat labels as outcome of regression problem

cv.glmnet(x=X, y=Y, family="multinomial")

Clustering Documents

Single-Member Model

Each document belongs to a single cluster

e.g. job descriptions are "accounting", "teaching", "IT", etc.

Supervised approach: if you have labeled documents

Treat labels as outcome of regression problem

```
cv.glmnet(x=X, y=Y, family="multinomial")
```

Each document given a probability of having each label

Clustering Documents

Single-Member Model

Each document belongs to a single cluster

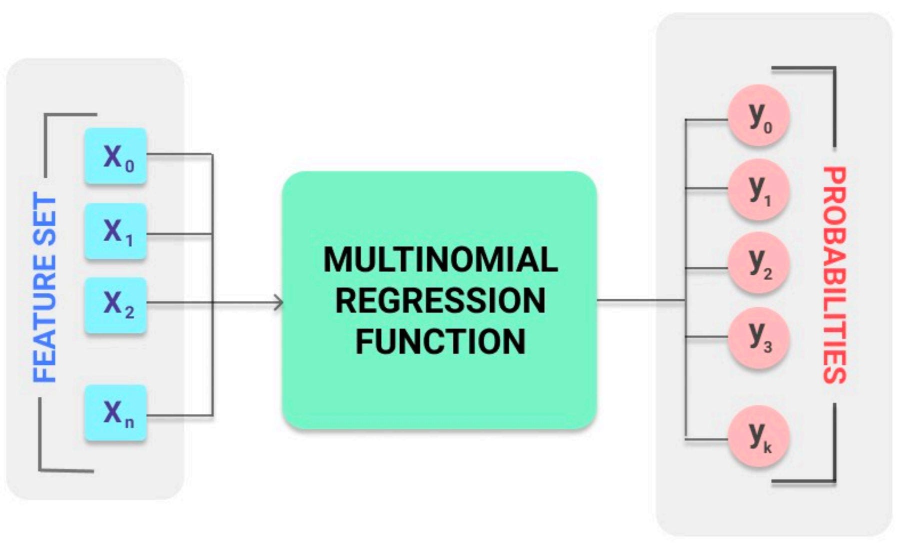
e.g. job descriptions are "accounting", "teaching", "IT", etc.

Supervised approach: if you have labeled documents

Treat labels as outcome of regression problem

```
cv.glmnet(x=X, y=Y, family="multinomial")
```

Each document given a probability of having each label



Does Question Asking Influence Interpersonal Liking?

(Huang, Yeomans, Brooks, Minson & Gino, 2017)

Does Question Asking Influence Interpersonal Liking?

(Huang, Yeomans, Brooks, Minson & Gino, 2017)

Studies 1 & 2

- 15 min open-ended conversations
- Strangers paired in dyads over ChatPlat
- Treatment Effect: question-asking instructions

Does Question Asking Influence Interpersonal Liking?

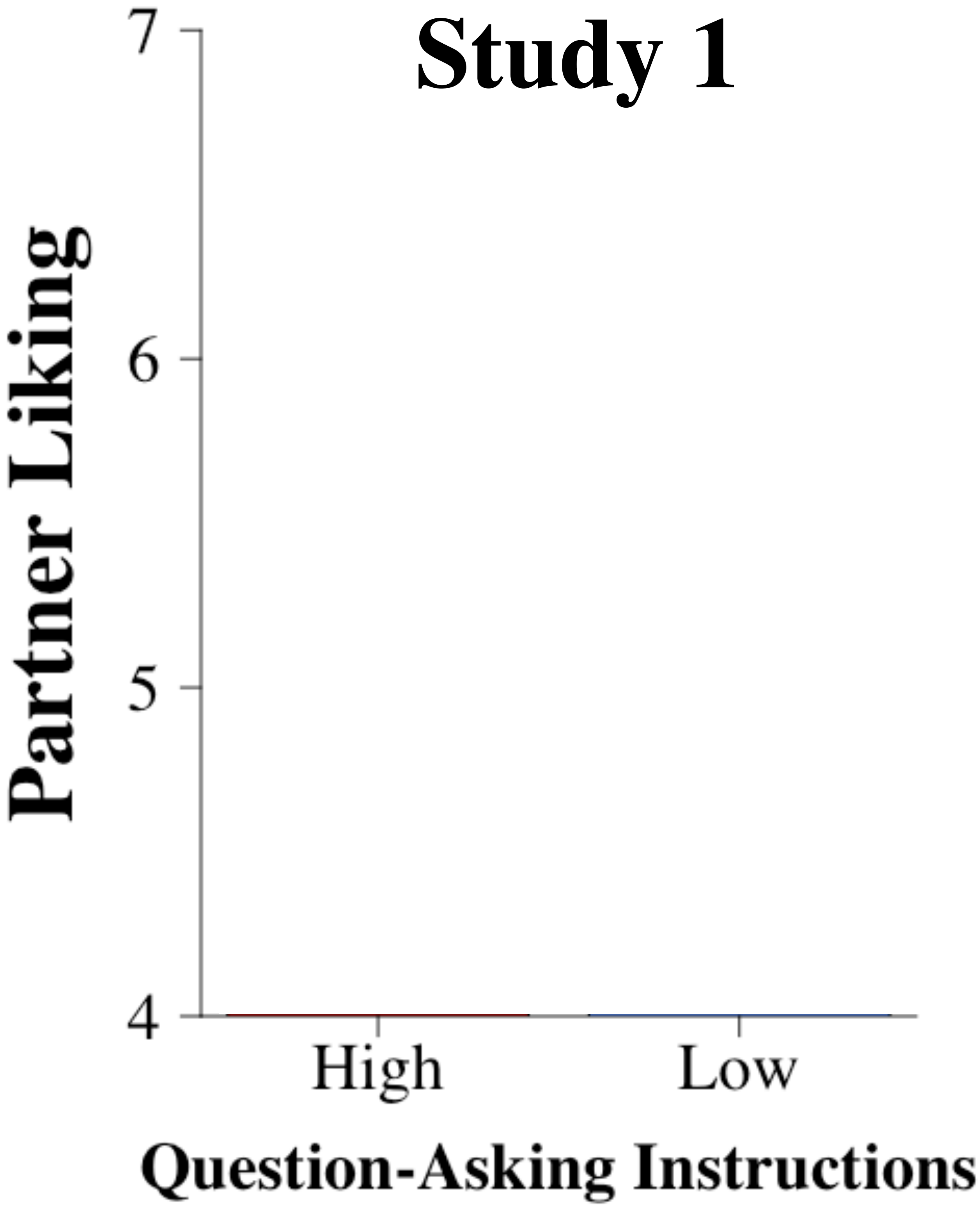
(Huang, Yeomans, Brooks, Minson & Gino, 2017)

Outcome: Interpersonal Liking

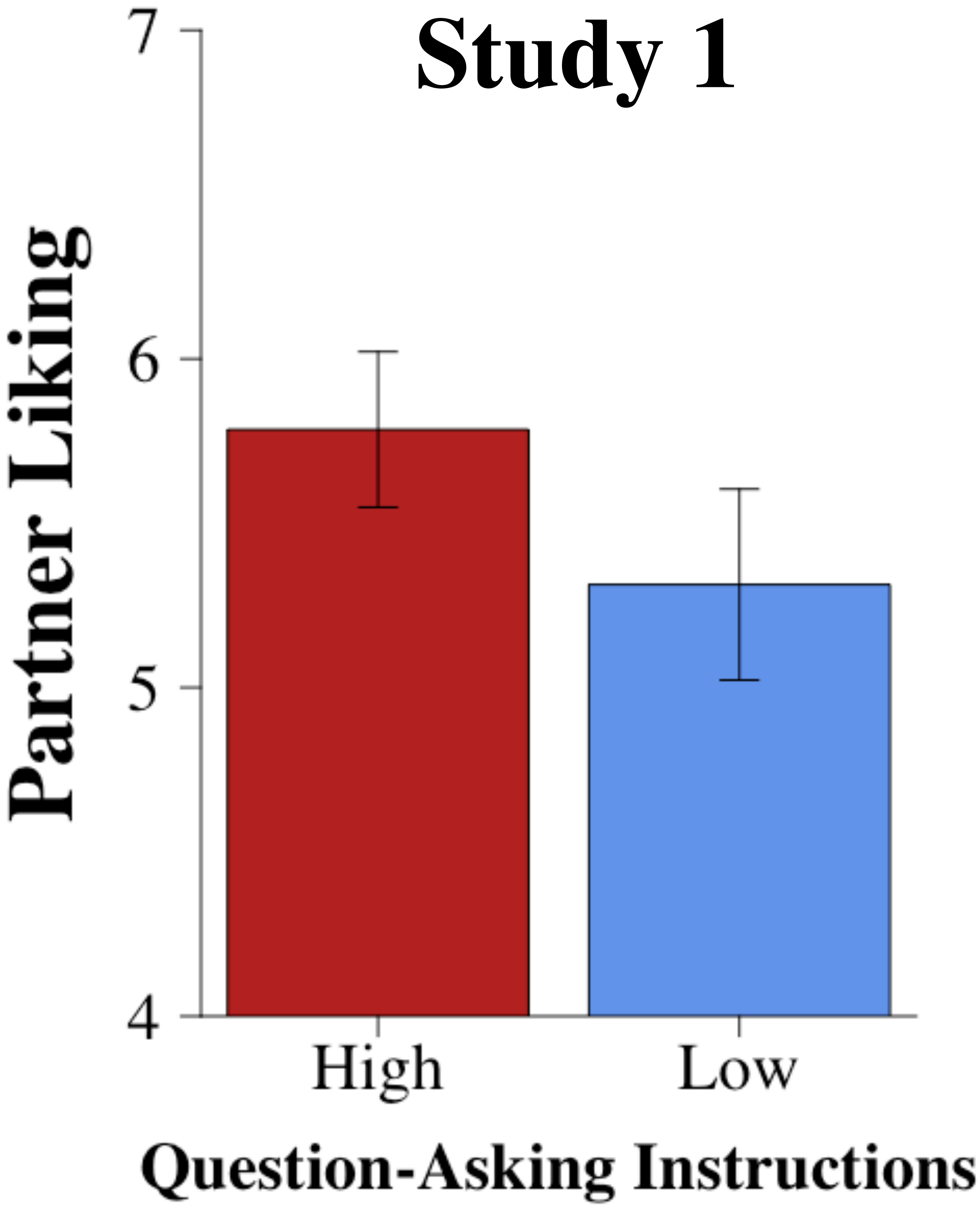
- My partner is likable.
- I liked my partner.
- I would enjoy spending time with my partner.
- I dislike my partner. *(reversed)*

(Cronbach's $\boldsymbol{\alpha}$ *= 0.96)*

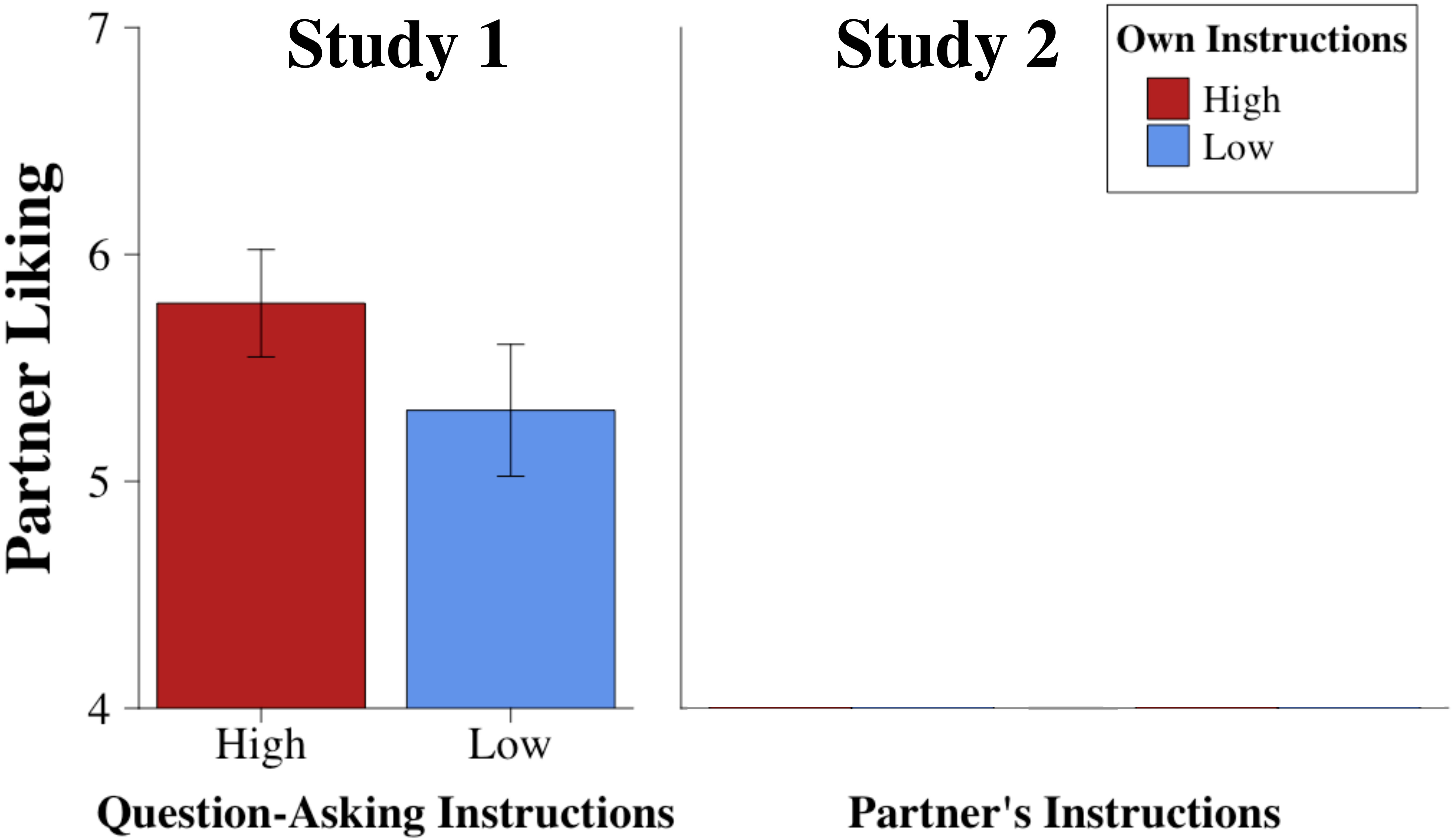
Main Effect of Question-Asking



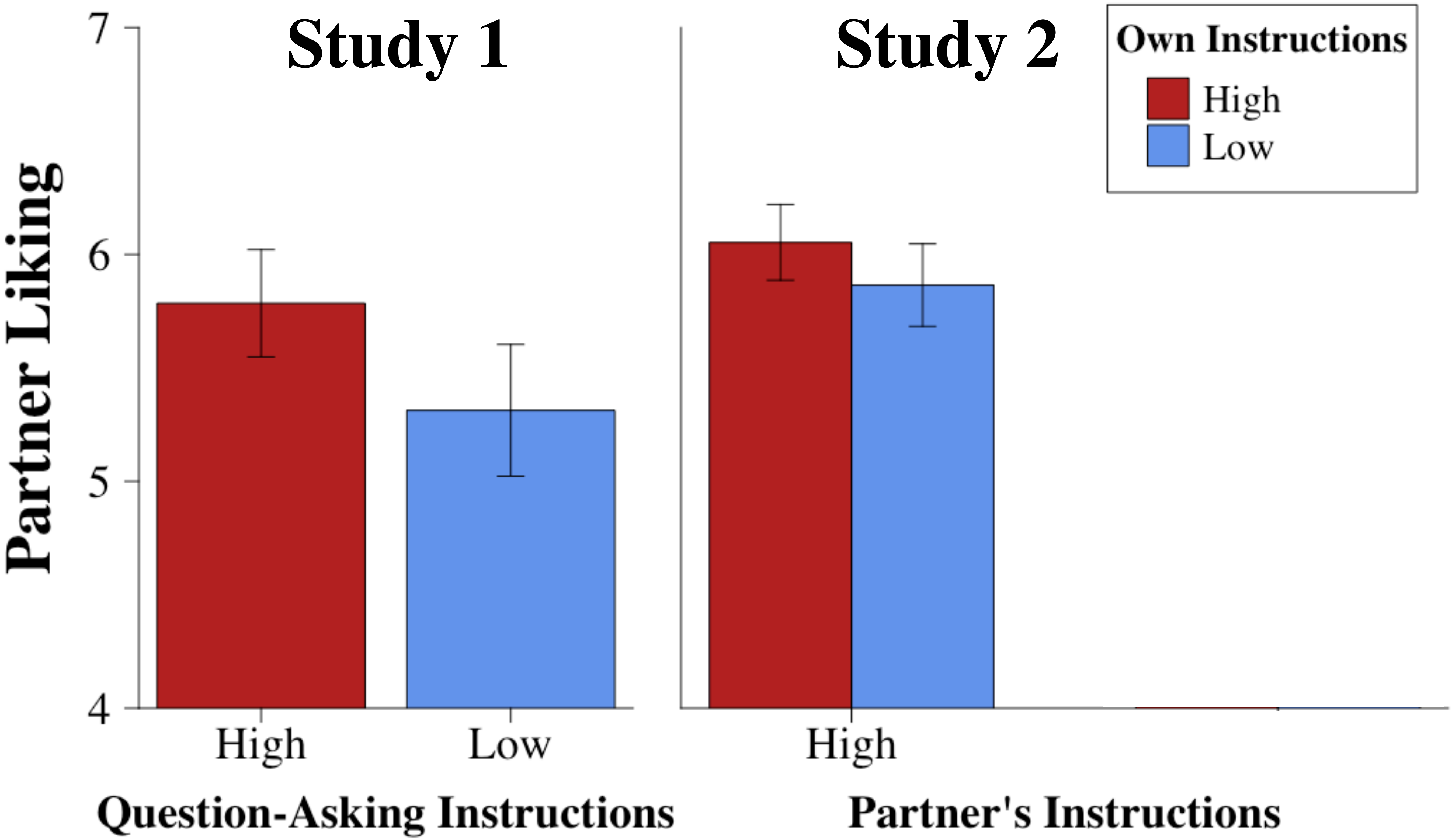
Main Effect of Question-Asking



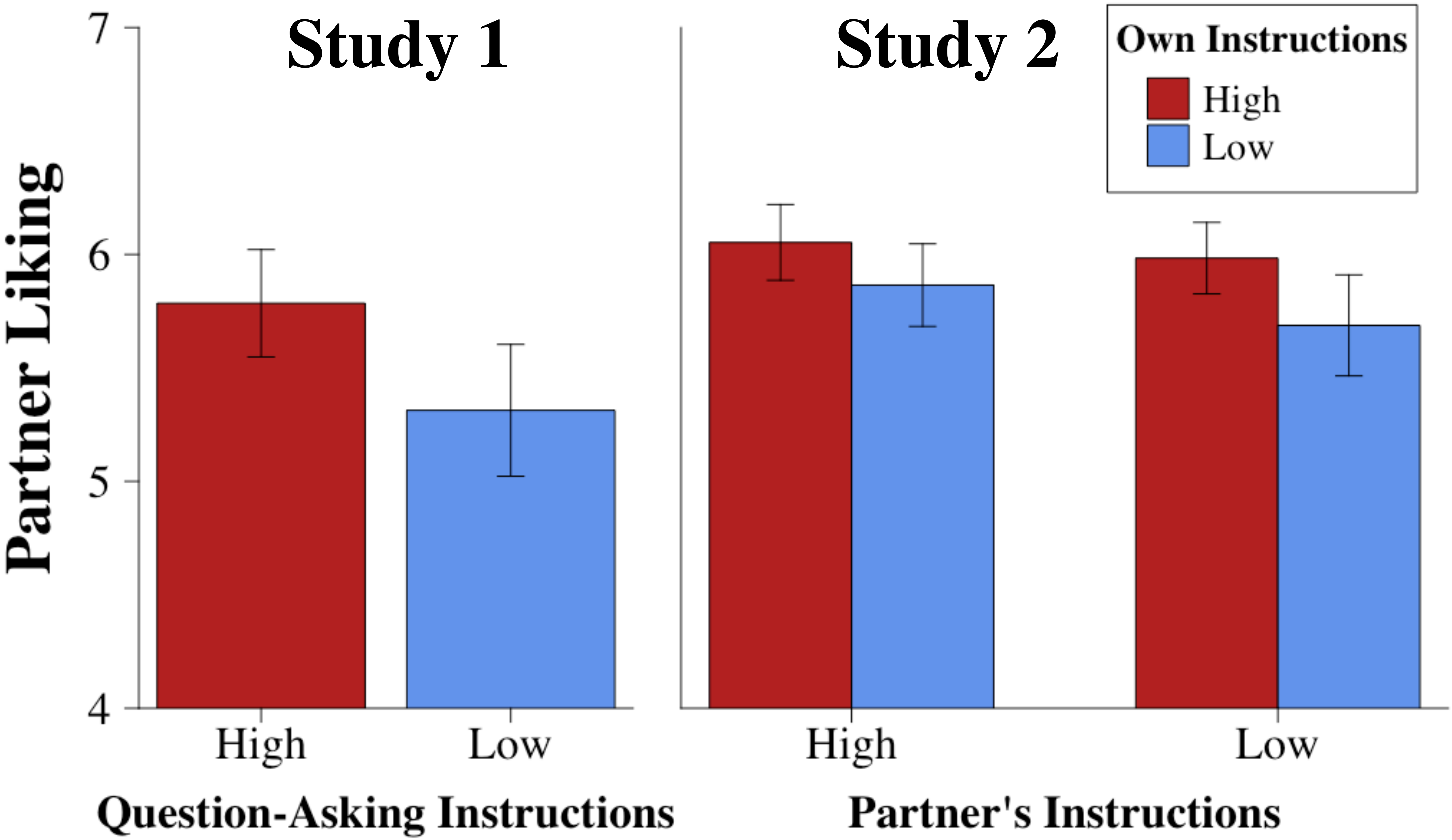
Main Effect of Question-Asking



Main Effect of Question-Asking



Main Effect of Question-Asking



Types of Annotation

Turn	ID	Text	Start	End
1	9152	Hi.	0	0
2	9150	Hi.	0	0
3	9152	Okay, [inaudible] nice to meet you.	4	4
4	9150	Janine nice to meet you.	5	6
5	9152	So, do you want me to start, I guess ?	8	10
6	9150	Sure.	11	11
7	9152	Okay, so, let's see the first one, if you were able to live to the age of 90 and retain either the mind or the body of a 30 year old the for the last 60 years of your life which would you want? We could skip	11	28
8	9150	I would say the body actually because that assuming that, like, I'm, like, you know, like dementia or Alzheimer's or anything like that it would be kind of cool to, like, sort of, like, stay young for a long	29	40
9	9152	Yeah, same with me only from experience my grandmother, who's gonna be 90 in July, actually, this year, she is still healthy, really, but like body wise but she has a little bit of dementia and a lot of memory loss. But my mother, on the other hand, is really like physically incapacitated and has lung problems and back problems so, but she's sharp as a tack, so you know, seeing how each	41	84
10	9150	If it works out and you're like sharp and you're --	87	89
11	9152	Well, if you're both that would be great but . Yeah.	91	95
12	9150	Okay, so, what's the strangest thing about where you grew up?	96	100
13	9152	There really isn't, I grew up in a pretty average town with, I went to public high school and it was pretty normal, so for me, I couldn't really answer anything about that. Was there anything for you?	100	119
14	9150	Well, I mean, I'm from, like, a suburb of, like, D.C. so, I guess, like, in high school or at least, or like, when I was kid, like, everyone's parents worked for, like the government. I guess that was	120	132
15	9152	Okay, let's see, when did you last sing to yourself or to someone	132	139
16	9150	Sing to myself? Well, probably, like, last night, actually, because I was cleaning my room and I had, like, my Spotify playlist, like,	141	152
17	9152	You were by yourself?	152	153
18	9150	I live by myself.	154	154

Types of Annotation

Holistic

Turn	ID	Text	Start	End
1	9152	Hi.	0	0
2	9150	Hi.	0	0
3	9152	Okay, [inaudible] nice to meet you.	4	4
4	9150	Janine nice to meet you.	5	6
5	9152	So, do you want me to start, I guess ?	8	10
6	9150	Sure.	11	11
7	9152	Okay, so, let's see the first one, if you were able to live to the age of 90 and retain either the mind or the body of a 30 year old the for the last 60 years of your life which would you want? We could skip	11	28
8	9150	I would say the body actually because that assuming that, like, I'm, like, you know, like dementia or Alzheimer's or anything like that it would be kind of cool to, like, sort of, like, stay young for a long	29	40
9	9152	Yeah, same with me only from experience my grandmother, who's gonna be 90 in July, actually, this year, she is still healthy, really, but like body wise but she has a little bit of dementia and a lot of memory loss. But my mother, on the other hand, is really like physically incapacitated and has lung problems and back problems so, but she's sharp as a tack, so you know, seeing how each	41	84
10	9150	If it works out and you're like sharp and you're --	87	89
11	9152	Well, if you're both that would be great but . Yeah.	91	95
12	9150	Okay, so, what's the strangest thing about where you grew up?	96	100
13	9152	There really isn't, I grew up in a pretty average town with, I went to public high school and it was pretty normal, so for me, I couldn't really answer anything about that. Was there anything for you?	100	119
14	9150	Well, I mean, I'm from, like, a suburb of, like, D.C. so, I guess, like, in high school or at least, or like, when I was kid, like, everyone's parents worked for, like the government. I guess that was	120	132
15	9152	Okay, let's see, when did you last sing to yourself or to someone	132	139
16	9150	Sing to myself? Well, probably, like, last night, actually, because I was cleaning my room and I had, like, my Spotify playlist, like,	141	152
17	9152	You were by yourself?	152	153
18	9150	I live by myself.	154	154

How responsive was this conversation?

Types of Annotation

Holistic

N=368 dyads

Turn-Level

N=4,545 questions

Turn	ID	Text	Start	End
1	9152	Hi.	0	0
2	9150	Hi.	0	0
3	9152	Okay, [inaudible] nice to meet you.	4	4
4	9150	Janine nice to meet you.	5	6
5	9152	So, do you want me to start, I guess ?	8	10
6	9150	Sure.	11	11
7	9152	Okay, so, let's see the first one, if you were able to live to the age of 90 and retain either the mind or the body of a 30 year old the for the last 60 years of your life which would you want? We could skip	11	28
8	9150	I would say the body actually because that assuming that, like, I'm, like, you know, like dementia or Alzheimer's or anything like that it would be kind of cool to, like, sort of, like, stay young for a long	29	40
9	9152	Yeah, same with me only from experience my grandmother, who's gonna be 90 in July, actually, this year, she is still healthy, really, but like body wise but she has a little bit of dementia and a lot of memory loss. But my mother, on the other hand, is really like physically incapacitated and has lung problems and back problems so, but she's sharp as a tack, so you know, seeing how each	41	84
10	9150	If it works out and you're like sharp and you're --	87	89
11	9152	Well, if you're both that would be great but . Yeah.	91	95
12	9150	Okay, so, what's the strangest thing about where you grew up?	96	100
13	9152	There really isn't, I grew up in a pretty average town with, I went to public high school and it was pretty normal, so for me, I couldn't really answer anything about that. Was there anything for you?	100	119
14	9150	Well, I mean, I'm from, like, a suburb of, like, D.C. so, I guess, like, in high school or at least, or like, when I was kid, like, everyone's parents worked for, like the government. I guess that was	120	132
15	9152	Okay, let's see, when did you last sing to yourself or to someone	132	139
16	9150	Sing to myself? Well, probably, like, last night, actually, because I was cleaning my room and I had, like, my Spotify playlist, like,	141	152
17	9152	You were by yourself?	152	153
18	9150	I live by myself.	154	154

What types of questions are they asking?

Typology of Question-Asking

(Huang et al., 2017)

Question	Example from Study 1		
Follow-Up			
Full Switch			
Partial Switch			
Mirror			
Introductory			
Rhetorical			

Typology of Question-Asking

(Huang et al., 2017)

Question	Example from Study 1		
Follow-Up	Person 1: I'm planning a trip to Canada. Person 2: Oh, cool. **Have you been there before?**		
Full Switch	Person 1: I am working at a dry cleaners. Person 2: **What do you like doing for fun?**		
Partial Switch	Person 1: Not super outdoorsy, but not opposed to a hike or something once in awhile. Person 2: **Have you been to the beach in Boston?**		
Mirror	Person 1: What did you have for breakfast? Person 2: I had eggs and fruit. **How about you?**		
Introductory	Person 1: hello! Person 2: **Hey, how's it going?**		
Rhetorical	Person 1: What's the craziest event you've been to? Person 2: Yesterday I followed a marching band around. **Where were they going?** It's a mystery.		

Typology of Question-Asking

(Huang et al., 2017)

Question	Example from Study 1	α	Q%
Follow-Up	Person 1: I'm planning a trip to Canada. Person 2: Oh, cool. **Have you been there before?**	**.87**	**40.5%**
Full Switch	Person 1: I am working at a dry cleaners. Person 2: **What do you like doing for fun?**	**.86**	**27.6%**
Partial Switch	Person 1: Not super outdoorsy, but not opposed to a hike or something once in awhile. Person 2: **Have you been to the beach in Boston?**	.47	5.5%
Mirror	Person 1: What did you have for breakfast? Person 2: I had eggs and fruit. **How about you?**	**.94**	**19.0%**
Introductory	Person 1: hello! Person 2: **Hey, how's it going?**	**.93**	**5.8%**
Rhetorical	Person 1: What's the craziest event you've been to? Person 2: Yesterday I followed a marching band around. **Where were they going?** It's a mystery.	.74	1.9%

Typology of Question-Asking

(Huang et al., 2017)

Question	Example from Study 1	α	Q%
Follow-Up	Person 1: I'm planning a trip to Canada. Person 2: Oh, cool. **Have you been there before?**	**.87**	**40.5%**
Full Switch	Person 1: I am working at a dry cleaners. Person 2: **What do you like doing for fun?**	.86	27.6%
Partial Switch	Person 1: Not super outdoorsy, but not opposed to a hike or something once in awhile. Person 2: **Have you been to the beach in Boston?**	.47	5.5%
Mirror	Person 1: What did you have for breakfast? Person 2: I had eggs and fruit. **How about you?**	.94	19.0%
Introductory	Person 1: hello! Person 2: **Hey, how's it going?**	.93	5.8%
Rhetorical	Person 1: What's the craziest event you've been to? Person 2: Yesterday I followed a marching band around. **Where were they going?** It's a mystery.	.74	1.9%

Question Type Detector

Bag of Words Features

- spell-check, expand contractions, strip punctuation
- stem words (e.g. “studied” → “study”)
- treat every 1-, 2- and 3-word sequence as a feature
- “only” 348 features used in >1% of question turns

Features for Question Type Detector

Common "stop" words (NLTK, quanteda, etc)

i	down	is	should've	she's	more	because	haven	yourself	why	a	didn't
me	in	are	now	her	most	as	haven't	yourselves	how	an	doesn
my	out	was	d	hers	other	until	isn	he	all	the	doesn't
myself	on	were	ll	herself	some	while	isn't	him	any	and	hadn
we	off	be	m	it	such	of	ma	his	both	but	hadn't
our	over	been	o	it's	no	at	mightn	himself	each	if	hasn
ours	under	being	re	its	nor	by	mightn't	she	few	or	hasn't
ourselves	again	have	ve	itself	not	for	mustn	who	s	before	wasn
you	further	has	y	they	only	with	mustn't	whom	t	after	wasn't
you're	then	had	ain	them	own	about	needn	this	can	above	weren
you've	once	having	aren	their	same	against	needn't	that	will	below	weren't
you'll	here	do	aren't	theirs	so	between	shan	that'll	just	to	won
you'd	there	does	couldn	themselves	than	into	shan't	these	don	from	won't
your	when	did	couldn't	what	too	through	shouldn	those	don't	up	wouldn
yours	where	doing	didn	which	very	during	shouldn't	am	should		wouldn't

Features for Question Type Detector

Common “stop” words (NLTK, quanteda, etc)

i	down	is	should've	she's	more	because	haven	yourself	why	a	didn't		
me	in	are	now	her	most	as	haven't	yourselves	how	an	doesn		
my	out	was	d	hers	other	until	isn	he	all	the	doesn't		
myself	on	were	ll	herself	some	while	isn't	him	any	and	hadn		
we	off	be	m	it	such	of	ma	his	both	but	hadn't		
our	over	been	o	it's	no	at	mightn	himself	each	if	hasn		
ours	under	being	re	its	nor	by	mightn't	she	few	or	hasn't		
ourselves	again	have	ve	itself	not	for	mustn	who	s	before	wasn		
you	further	has	y	they	only	with	mustn't	whom	t	after	wasn't		
you're	then	had	ain	them	own	about	needn	this	can	above	weren		
you've	once	having	aren	their	same	against	needn't	that	will	below	weren't		
you'll	here	do	aren't	theirs	so	between	shan	that'll	just	to	won		
you'd	there	does	couldn	themselves	than	into	shan't	these	don	from	won't		
your	when	did	couldn't	what	too	through	shouldn	those	don't	up	wouldn		
yours	where	doing	didn	which	very	during	shouldn't	am	should		wouldn't		

Distinctive Question Features

Follow-up	Switch	Introductory	Mirror
which	how old	how are you	how about
why	do you like	hello	what about
what kind	travel	your name	yourself?
cool	fun	how are	and
nice	do you live	hi how	i am
wow	interests	today?	and you
is it	hobbies?	what is	what about you
how do	you a student	go?	and
where do	weather	name?	no, i
want to	you from?	are you?	yes, i

Question Type Detector

Bag of Words Features

- spell-check, expand contractions, strip punctuation
- stem words (e.g. “studied” → “study”)
- treat every 1-, 2- and 3-word sequence as a feature
- “only” 348 features used in >1% of question turns

Contextual Features

- word count of question turn
- time into conversation
- questions in askee’s previous turn
- questions in asker’s previous turn
- multiple questions in the turn
- size of pre-question statement in the turn

The Detection Model

Q1
Q2
Q3
Q4
Q5

The Detection Model

	who	what	when	why	how	hobbies
Q1	0	1	0	0	0	0
Q2	0	0	1	0	0	0
Q3	0	1	0	1	1	1
Q4	1	0	0	1	0	1
Q5	0	0	0	0	0	0

Ngram Features

The Detection Model

	who	what	when	why	how	hobbies	Word count	prevQ asker	prevQ askee	turn Q count
Q1	0	1	0	0	0	0	11	0	0	1
Q2	0	0	1	0	0	0	26	0	1	2
Q3	0	1	0	1	1	1	8	1	0	1
Q4	1	0	0	1	0	1	15	1	0	1
Q5	0	0	0	0	0	0	19	0	0	2

Ngram Features

Context Features

The Detection Model

	who	what	when	why	how	hobbies	Word count	prevQ asker	prevQ askee	turn Q count
Q1	0	1	0	0	0	0	11	0	0	1
Q2	0	0	1	0	0	0	26	0	1	2
Q3	0	1	0	1	1	1	8	1	0	1
Q4	1	0	0	1	0	1	15	1	0	1
Q5	0	0	0	0	0	0	19	0	0	2

Ngram Features

Context Features

X

The Detection Model

	who	what	when	why	how	hobbies	Word count	prevQ asker	prevQ askee	turn Q count
Q1	0	1	0	0	0	0	11	0	0	1
Q2	0	0	1	0	0	0	26	0	1	2
Q3	0	1	0	1	1	1	8	1	0	1
Q4	1	0	0	1	0	1	15	1	0	1
Q5	0	0	0	0	0	0	19	0	0	2

Qtype
switch
follow
mirror
intro
follow

Ngram Features

Context Features

Annotated Labels

X

The Detection Model

	who	what	when	why	how	hobbies	Word count	prevQ asker	prevQ askee	turn Q count
Q1	0	1	0	0	0	0	11	0	0	1
Q2	0	0	1	0	0	0	26	0	1	2
Q3	0	1	0	1	1	1	8	1	0	1
Q4	1	0	0	1	0	1	15	1	0	1
Q5	0	0	0	0	0	0	19	0	0	2

Qtype
switch
follow
mirror
intro
follow

Ngram Features

Context Features

Annotated Labels

X

Y

The Detection Model

	who	what	when	why	how	hobbies	Word count	prevQ asker	prevQ askee	turn Q count
Q1	0	1	0	0	0	0	11	0	0	1
Q2	0	0	1	0	0	0	26	0	1	2
Q3	0	1	0	1	1	1	8	1	0	1
Q4	1	0	0	1	0	1	15	1	0	1
Q5	0	0	0	0	0	0	19	0	0	2

Qtype
switch
follow
mirror
intro
follow

Ngram Features

Context Features

Annotated Labels

X

Y

Output:

	Follow-up	Switch	Introductory	Mirror
Q1	0.3	0.6	0.05	0.05
Q2	0.06	0.02	0.12	0.8
Q3	0.4	0.4	0.14	0.06
Q4	0.1	0.1	0.7	0.1
Q5	0.9	0.03	0.03	0.04

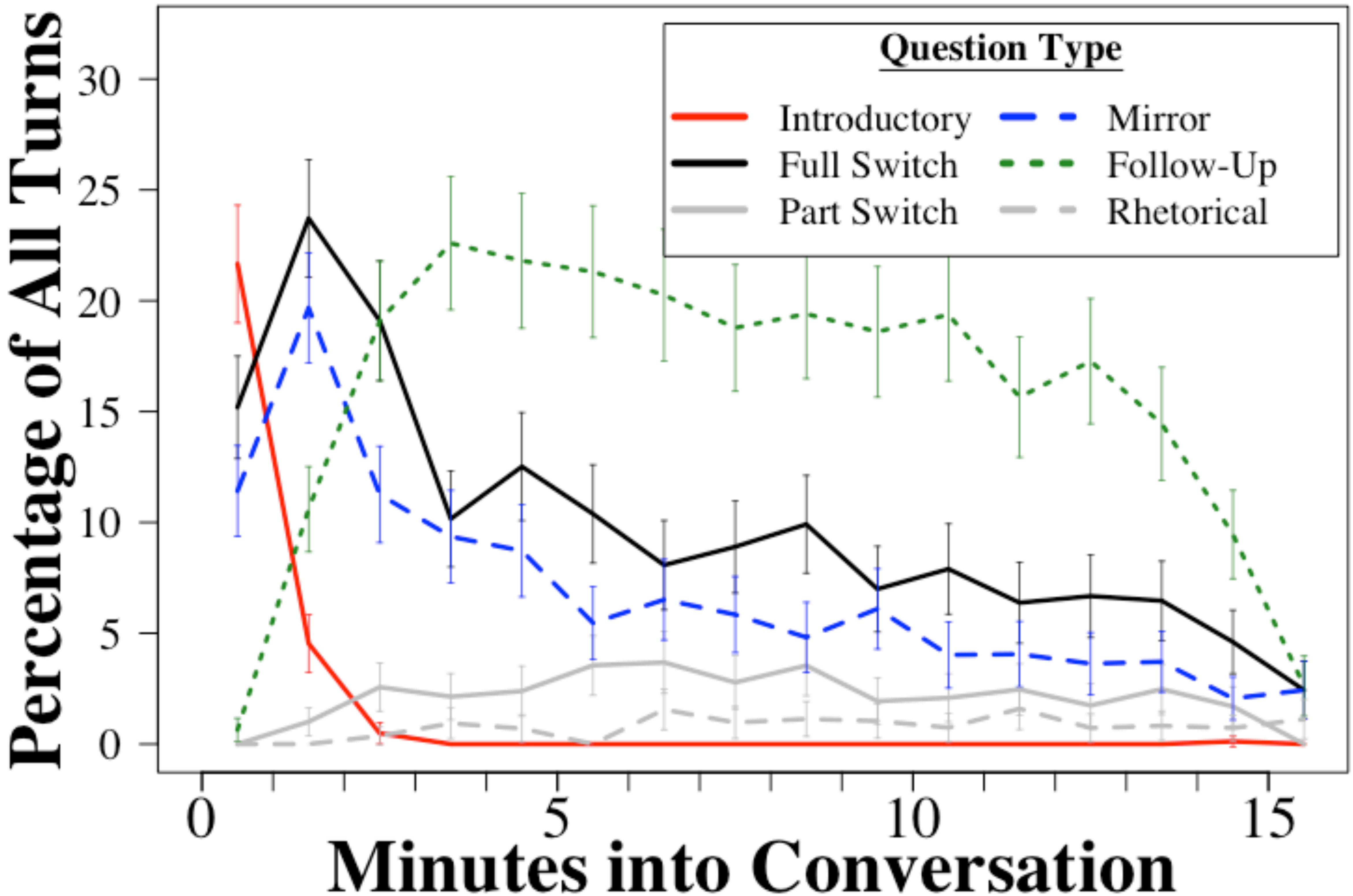
Distinctive Question Features

Contextual Feature	Follow-Up	Switch	Intro	Mirror
Word Count of turn	**.20**	—	—	—

Distinctive Question Features

Contextual Feature	Follow-Up	Switch	Intro	Mirror
Word Count of turn	**.20**	—	—	—
Time into Conversation	**.25**	—	**-1.34**	—

Time Course of Question-Asking



Distinctive Question Features

Contextual Feature	Follow-Up	Switch	Intro	Mirror
Word Count of turn	**.20**	—	—	—
Time into Conversation	**.25**	—	**-1.34**	—
Question in askee’s last turn	—	—	**-.43**	**.61**
Pre-question statement in turn	**-.10**	—	—	**.37**

Distinctive Question Features

hello, what things do you enjoy doing?
i love travelling
me too
are you a boston native ?
no, I am from Colombia. and you?

Distinctive Question Features

Contextual Feature	Follow-Up	Switch	Intro	Mirror
Word Count of turn	**.20**	—	—	—
Time into Conversation	**.25**	—	**-1.34**	—
Question in askee's last turn	—	—	**-.43**	**.61**
Pre-question statement in turn	**-.10**	—	—	**.37**
Question in asker's last turn	**.30**	**.12**	**-.33**	**-.12**
Multiple questions in the turn	**.08**	**-.04**	—	—

Distinctive Question Features

Do you play an instrument or enjoy listening to music?
I played the piano pretty seriously for about 13 years
whoa that's awesome. Did you like that?

Distinctive Question Features

Do you play an instrument or enjoy listening to music?
I played the piano pretty seriously for about 13 years
whoa that's awesome. Did you like that?
I don't play that much anymore, just occasionally when I go home. Yeah I loved it. Only classical music but I still enjoyed it so much. And now obviously I still really like listening to music, and producing it.
Classical? or other genres
you should try out Snarky Puppy. it's an instrumental based music. my favorite band
Cool! I'll definitely check it out. What kind of music do you mostly produce? And how'd you get started doing it?

Question Type Detector Accuracy

(Huang et al., 2017)

		Human Label of Question			
		Follow-up	**Switch**	**Introductory**	**Mirror**
Machine Label of Question	**Follow-up**	1377	358	7	99
	Switch	373	806	30	139
	Introductory	2	17	202	18
	Mirror	89	71	12	609

Question Type Detector Accuracy

(Huang et al., 2017)

Accuracy = 71.0%

95%CI: [69.6,72.4]

		Human Label of Question			
		Follow-up	**Switch**	**Introductory**	**Mirror**
Machine Label of Question	**Follow-up**	**1377**	358	7	99
	Switch	373	**806**	30	139
	Introductory	2	17	**202**	18
	Mirror	89	71	12	**609**

Question Type Detector Accuracy

(Huang et al., 2017)

Accuracy = 87.0%

95%CI: [86.0,88.0]

Human Label of Question

		Follow-up	Switch	Introductory	Mirror
Machine Label of Question	Follow-up	1377	358	7	99
	Switch	373	806	30	139
	Introductory	2	17	202	18
	Mirror	89	71	12	609

Question Types in New Data

[…]

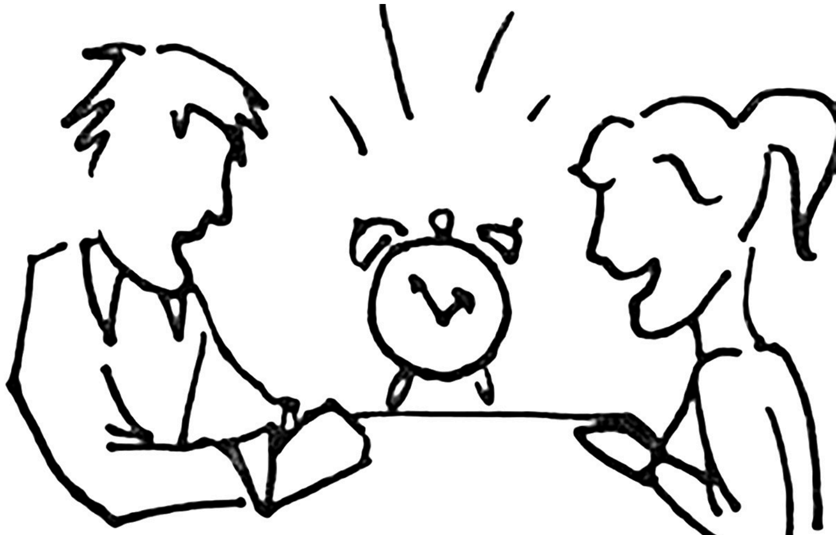
Male: So did you do anything fun over the weekend?

Female: Yes, I went to Yosemite with my friends.

Male: Nice! I love it there. Had you been there before?

[…]

(Ranganath, Jurafsky & McFarland, 2009)



Question Types in New Data

[…]

Male: So did you do anything fun over the weekend?

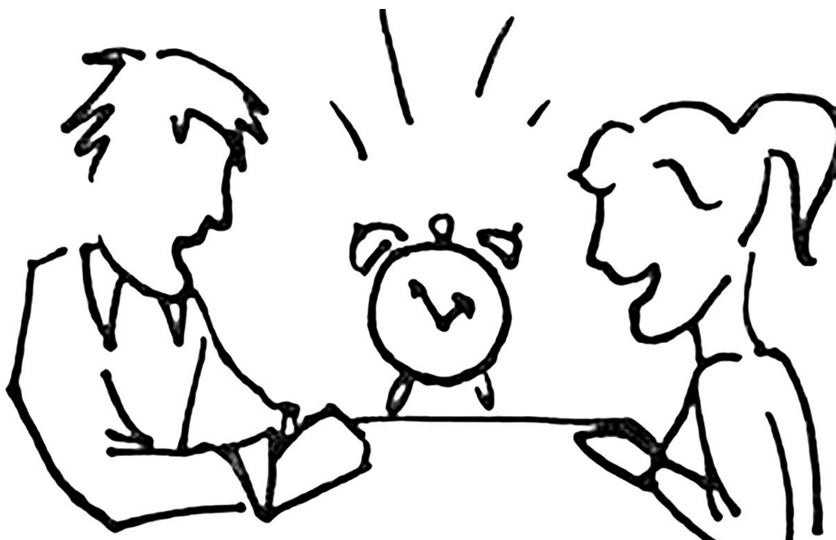
Female: Yes, I went to Yosemite with my friends.

Male: Nice! I love it there. Had you been there before?

[…]

Follow-Up	0.10
Full Switch	**0.63**
Mirror	0.19
Introductory	0.08

(Ranganath, Jurafsky & McFarland, 2009)



Question Types in New Data

[…]

Male: So did you do anything fun over the weekend?

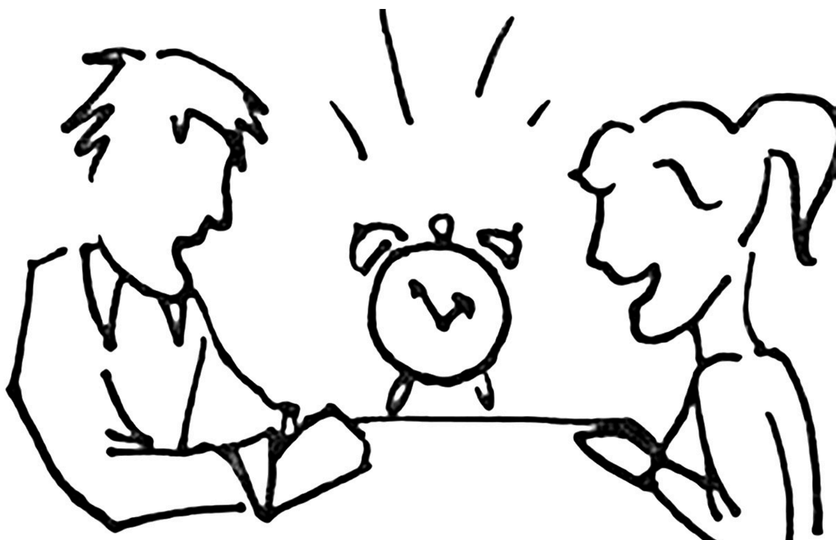
Female: Yes, I went to Yosemite with my friends.

Male: Nice! I love it there. Had you been there before?

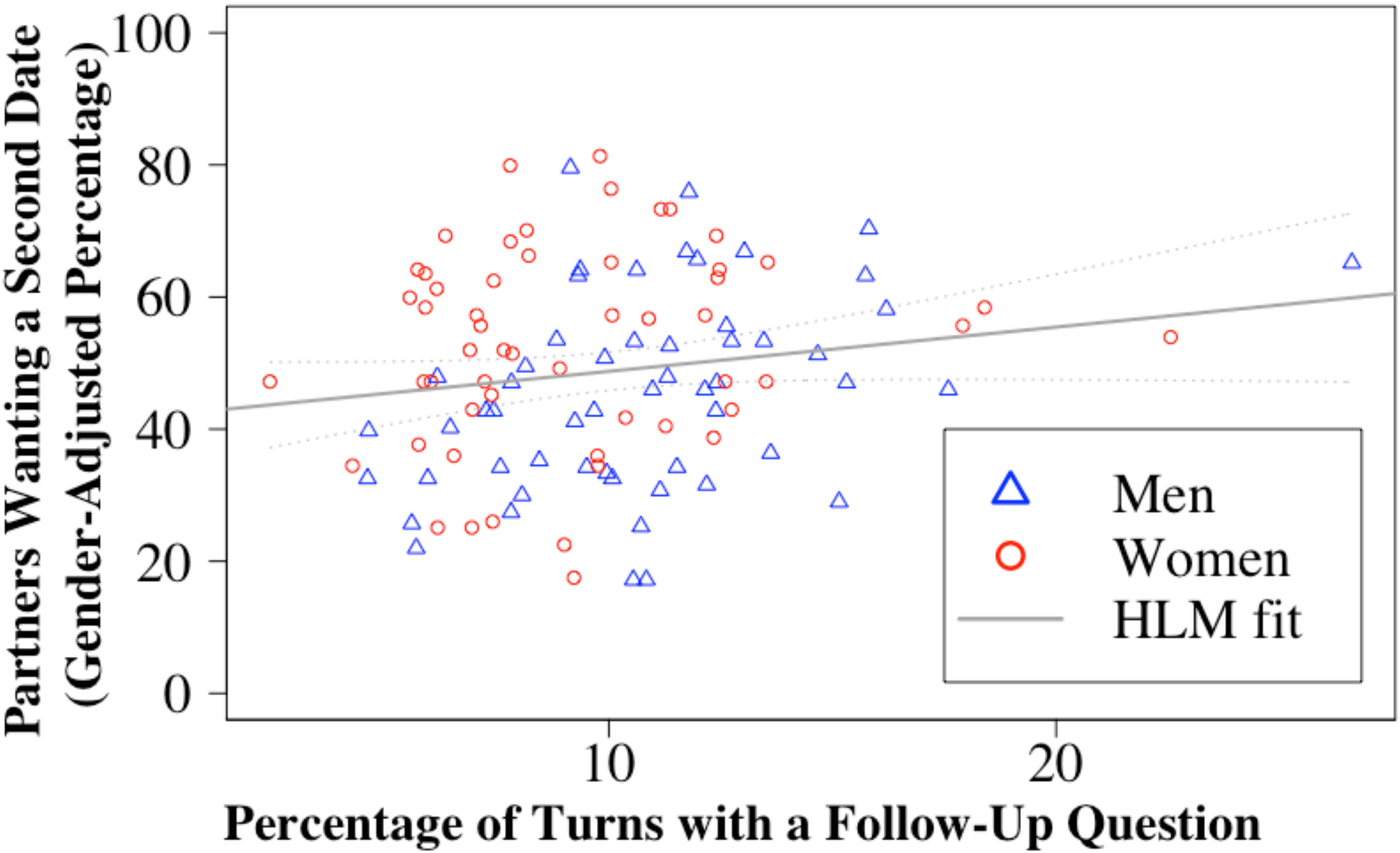
[…]

Follow-Up	**0.84**
Full Switch	0.11
Mirror	0.03
Introductory	0.11

(Ranganath, Jurafsky & McFarland, 2009)



Follow-Up Questions in Speed Dates



Clustering Documents

Single-Member Model

Each document belongs to a single cluster
  e.g. job descriptions are "accounting", "teaching", "IT", etc.

Unsupervised approach: if you don't have labeled documents
  Treat individual features as *dimensions* in a feature space
  k-means clustering algorithm
    assign clusters to minimize intra-cluster distance

Supervised approach: if you have labeled documents
  Treat labels as outcome of regression problem

Clustering Documents

Multinomial Mixture Model

Each document is a combination of underlying clusters
Each cluster is defined by a combination of words

Clustering Documents

Multinomial Mixture Model

Each document is a combination of underlying clusters
Each cluster is defined by a combination of words

A common approach: Latent Semantic Analysis

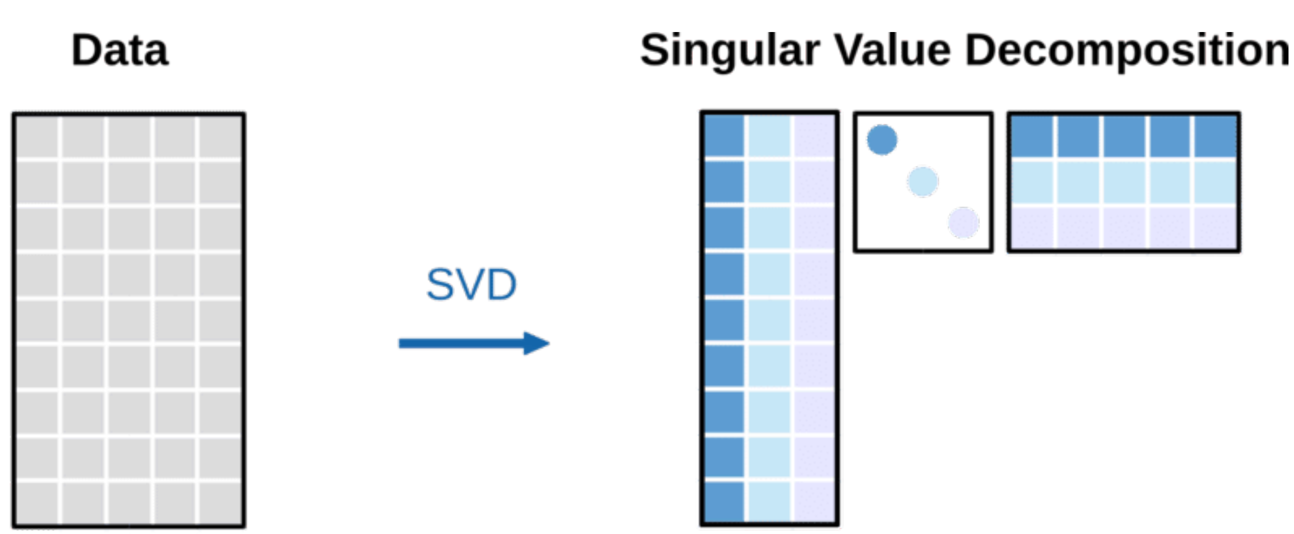
Clustering Documents

Multinomial Mixture Model

Each document is a combination of underlying clusters
Each cluster is defined by a combination of words

A common approach: Latent Semantic Analysis

principal components/singular value decomposition for text



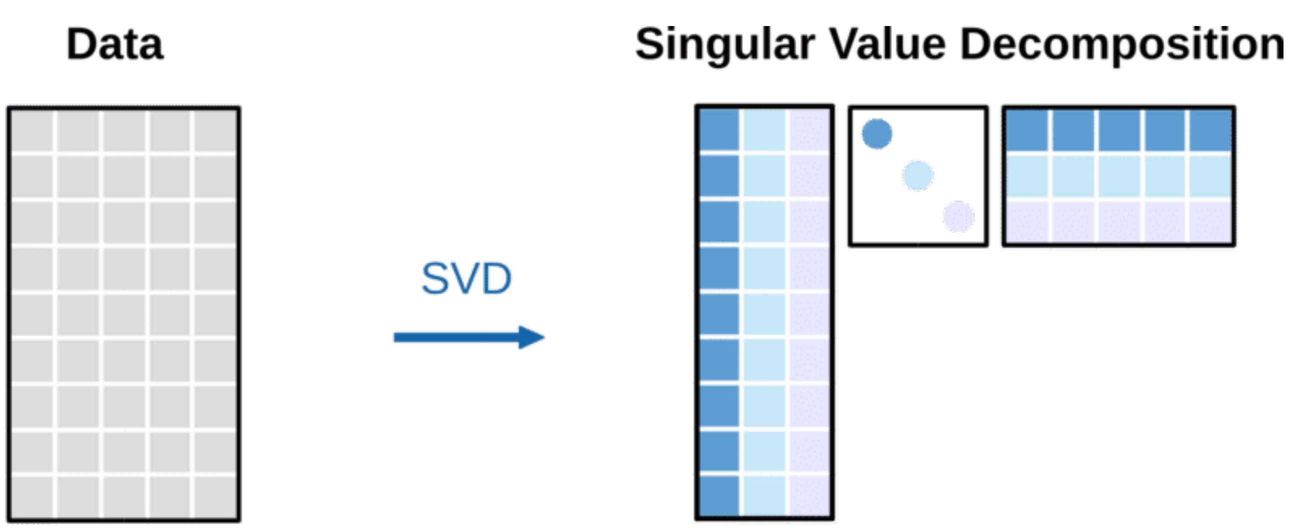
Clustering Documents

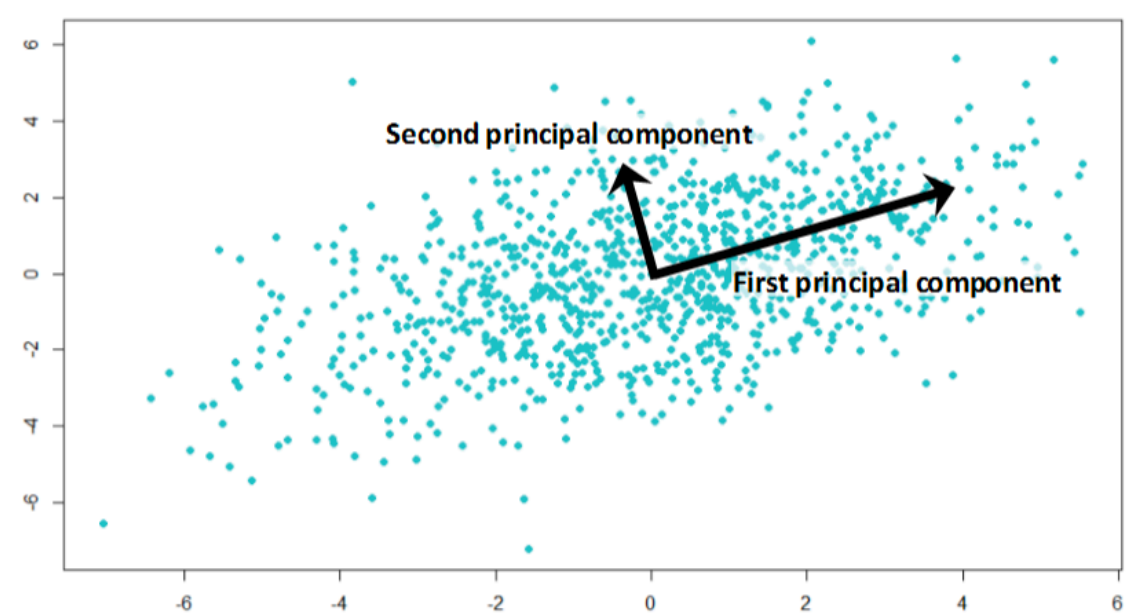
Multinomial Mixture Model

Each document is a combination of underlying clusters
Each cluster is defined by a combination of words

A common approach: Latent Semantic Analysis

principal components/singular value decomposition for text





Clustering Documents

Multinomial Mixture Model

Each document is a combination of underlying clusters
Each cluster is defined by a combination of words

A common approach: Latent Semantic Analysis

principal components/singular value decomposition for text

Clustering Documents

Multinomial Mixture Model

Each document is a combination of underlying clusters
Each cluster is defined by a combination of words

A common approach: Latent Semantic Analysis

principal components/singular value decomposition for text

The most common approach: Latent Dirichlet Allocation

“Topic models”

Clustering Documents

Multinomial Mixture Model

Each document is a combination of underlying clusters
Each cluster is defined by a combination of words

A common approach: Latent Semantic Analysis

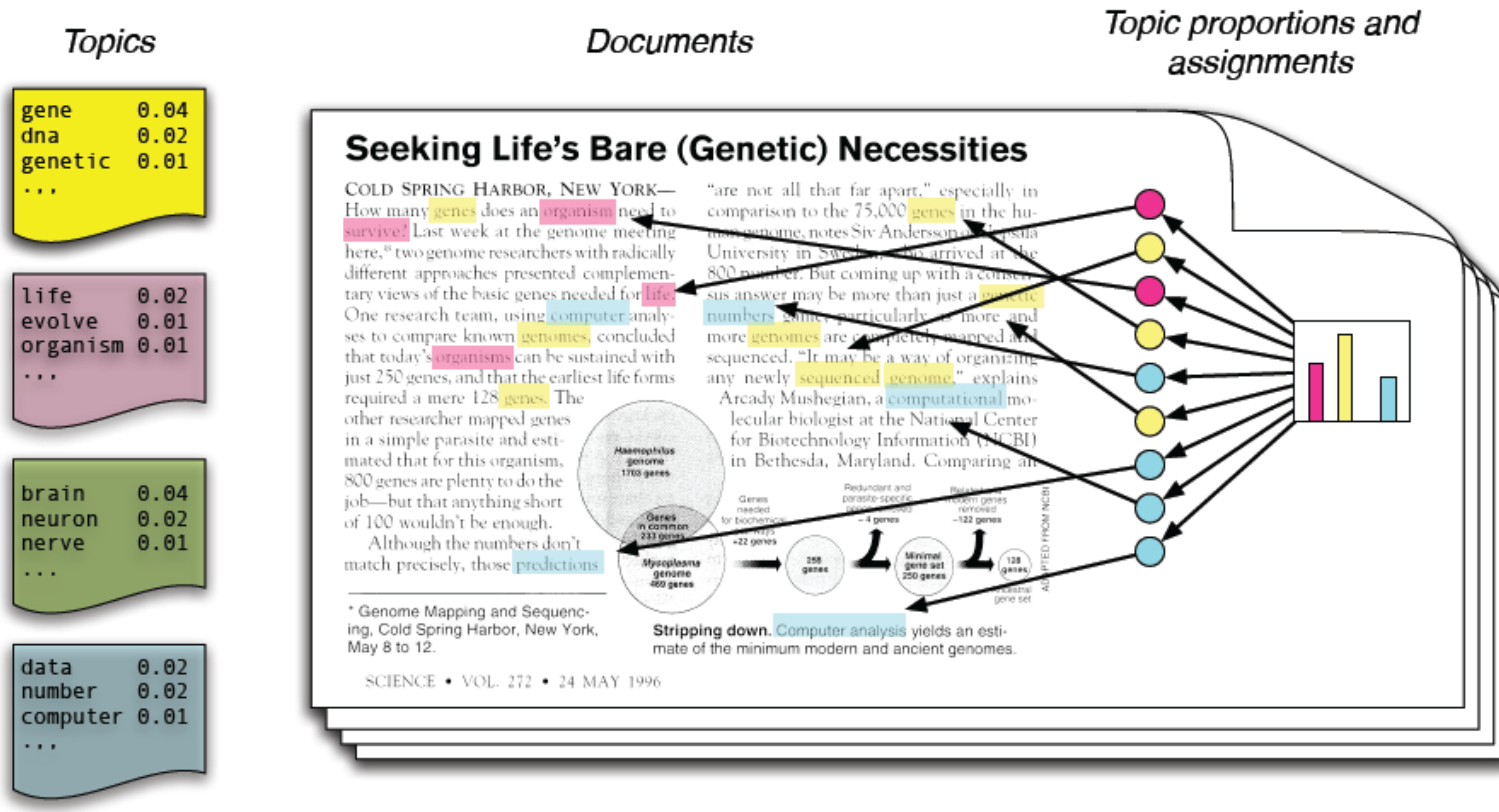
principal components/singular value decomposition for text

The most common approach: Latent Dirichlet Allocation

"Topic models"

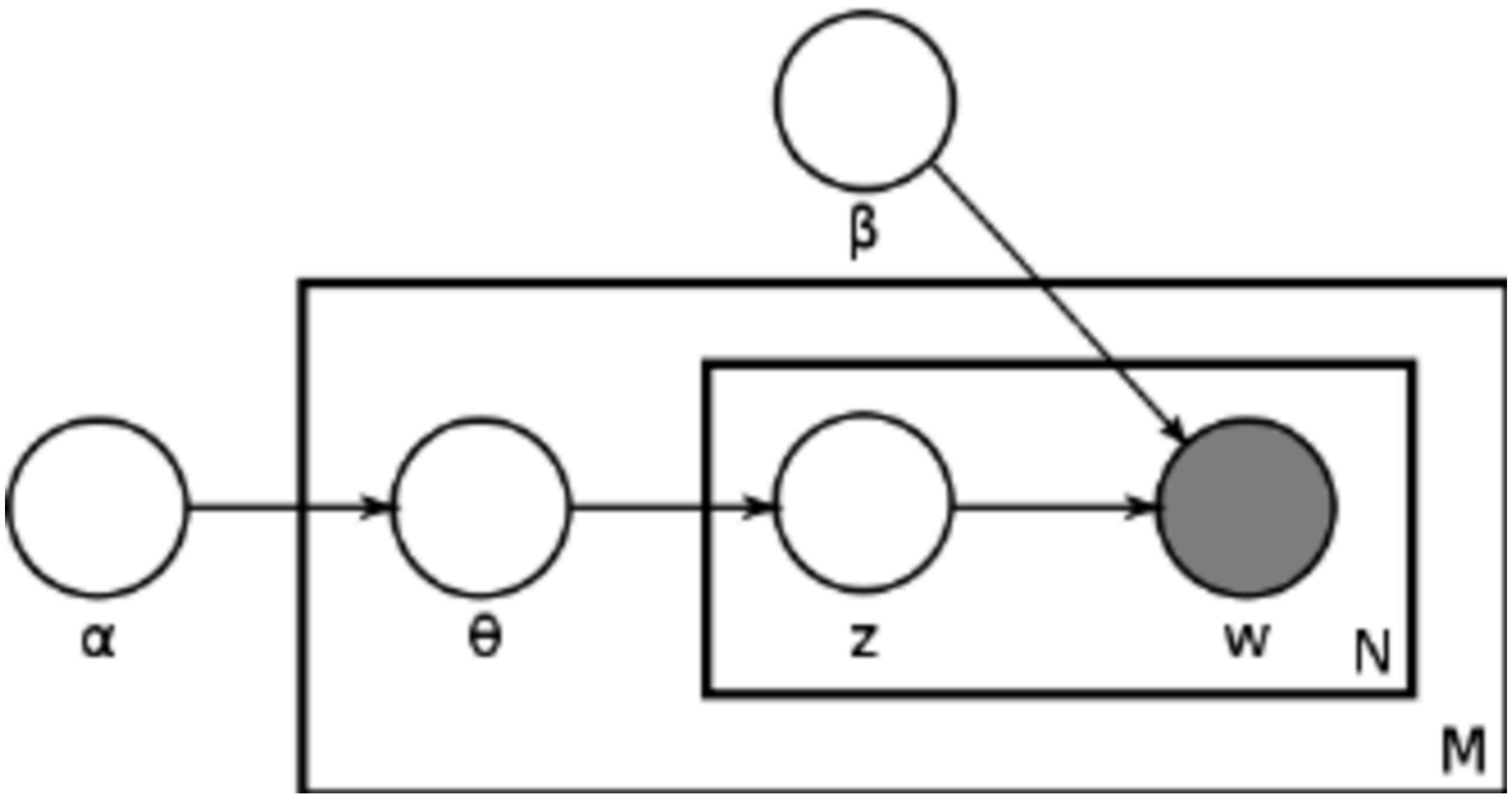
Both LSA and LDA rely on co-occurrence in documents
- if words appear together often, they are a similar topic

Topic Models



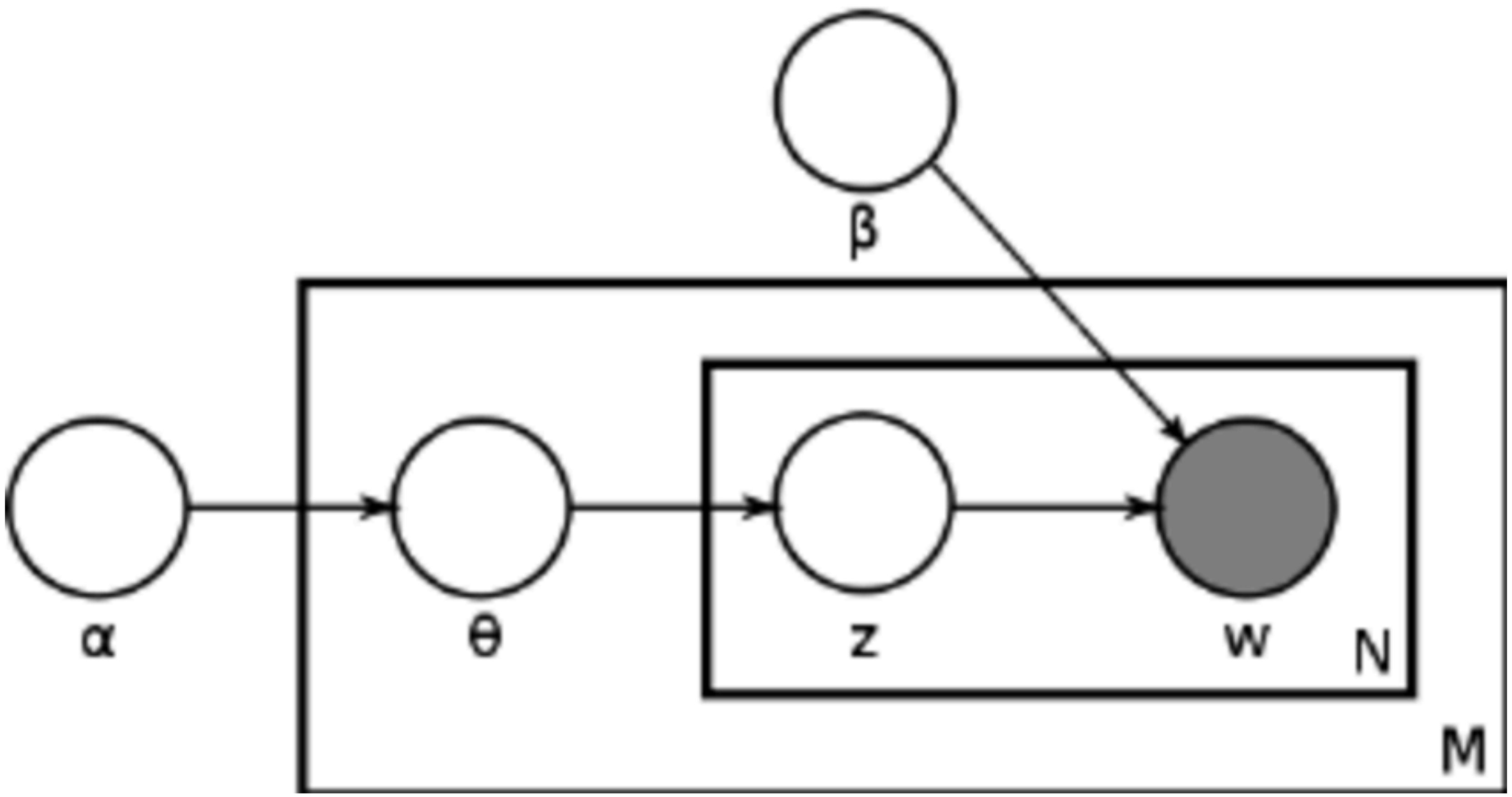
(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models



(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models



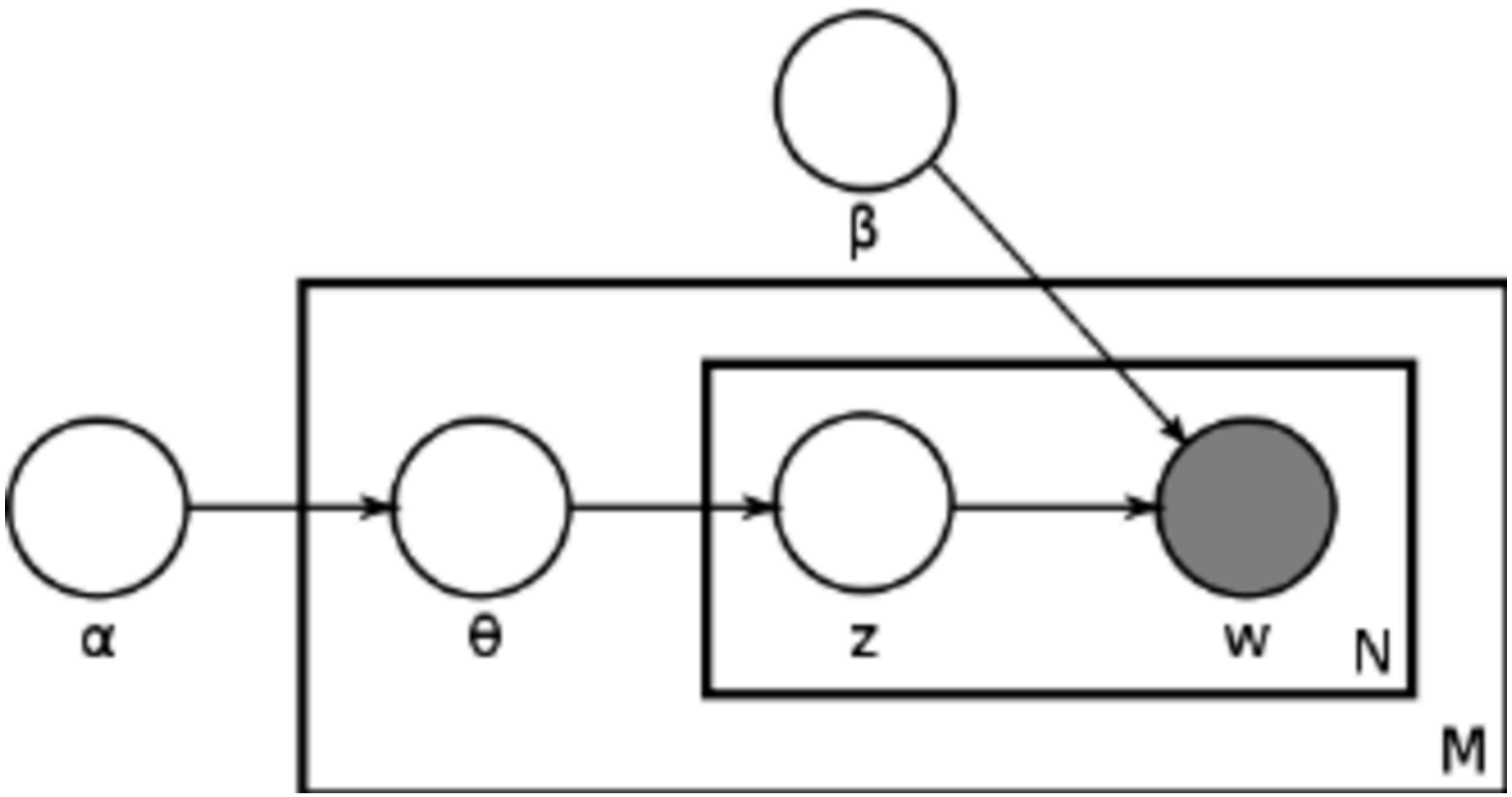
Beta matrix

One row per topic

One column per word

(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models



Beta matrix

One row per topic

One column per word

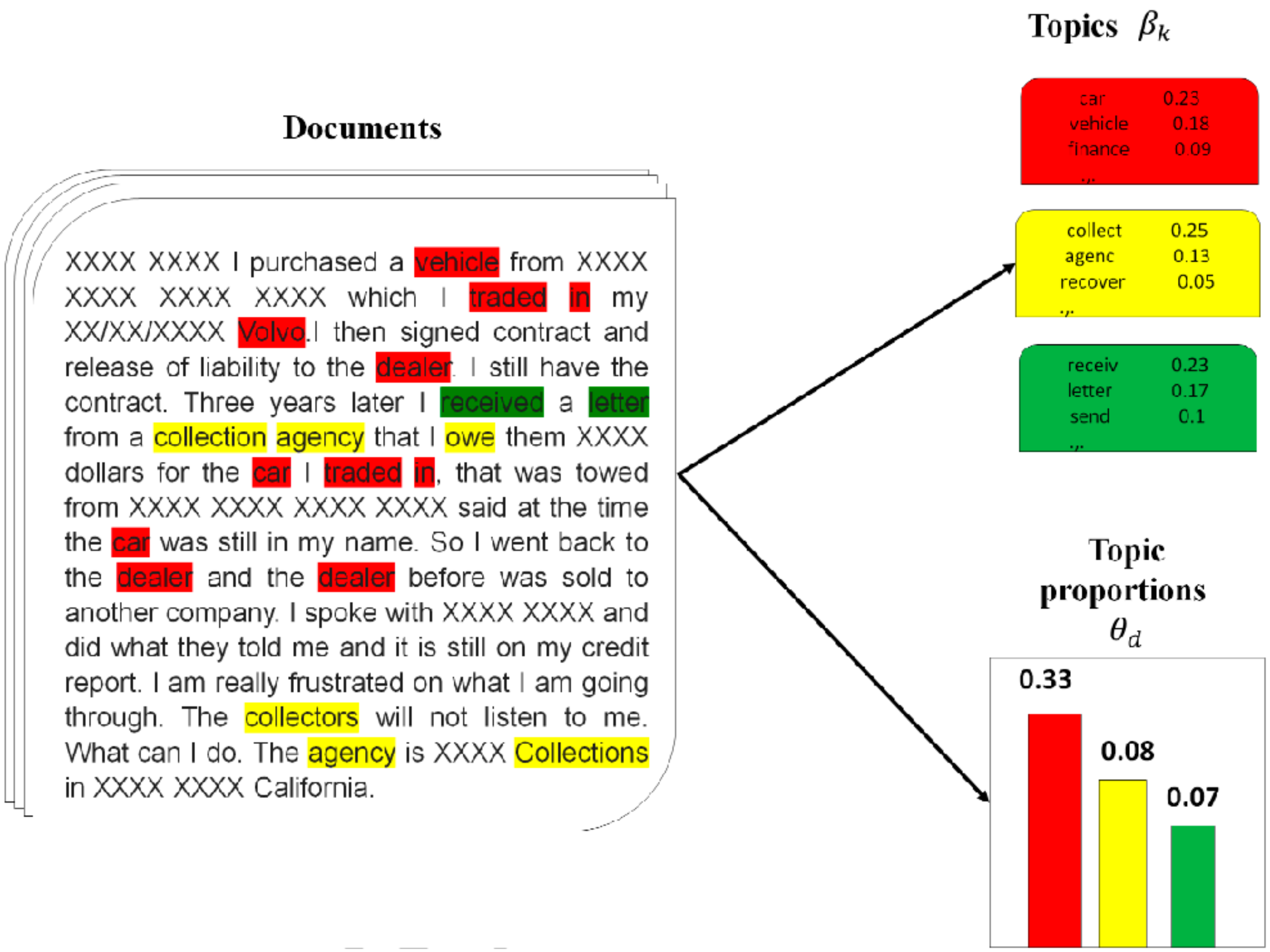
Theta matrix

One row per document

One column per topic

(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models



Beta matrix

One row per topic

One column per word

Theta matrix

One row per document

One column per topic

(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models

Relies on Dirichlet distributions

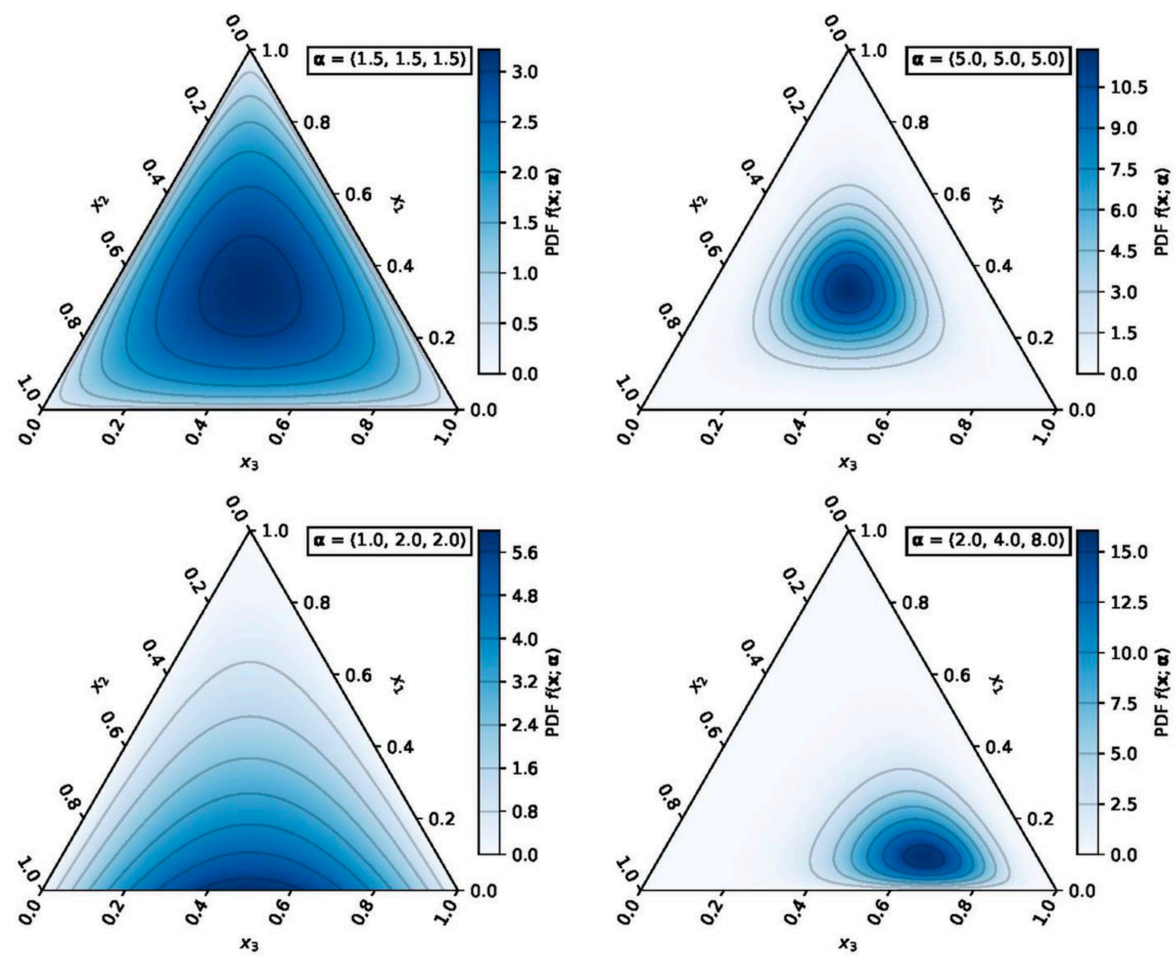
Dependent proportions, not orthogonal gaussians (as in LSA)

(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models

Relies on Dirichlet distributions

Dependent proportions, not orthogonal gaussians (as in LSA)



(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models

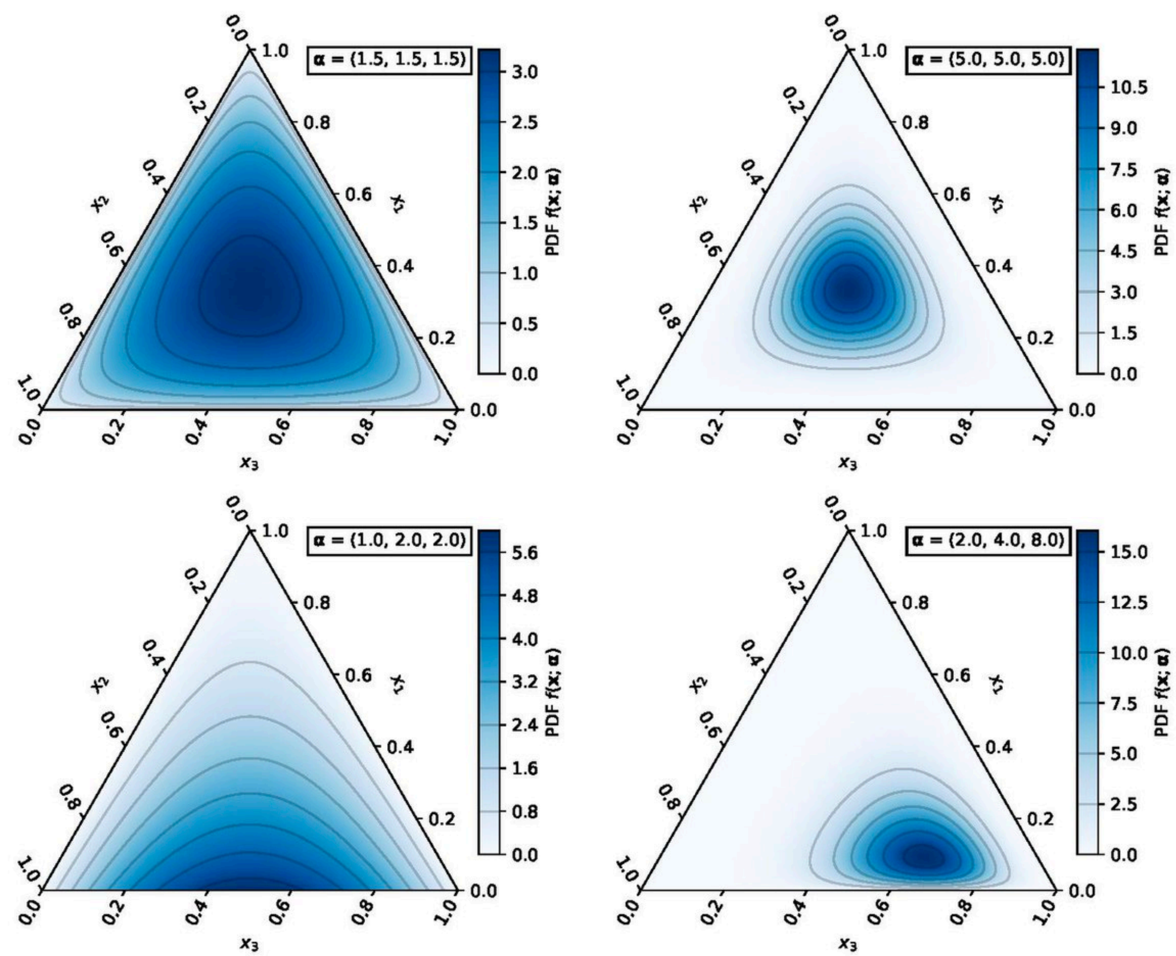
Relies on Dirichlet distributions

Dependent proportions, not orthogonal gaussians (as in LSA)

Many, many variations…

Common aspects:

- Bayesian estimation
- Unigram feature set
- Rarely stop words



(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models

Relies on Dirichlet distributions

Dependent proportions, not orthogonal gaussians (as in LSA)

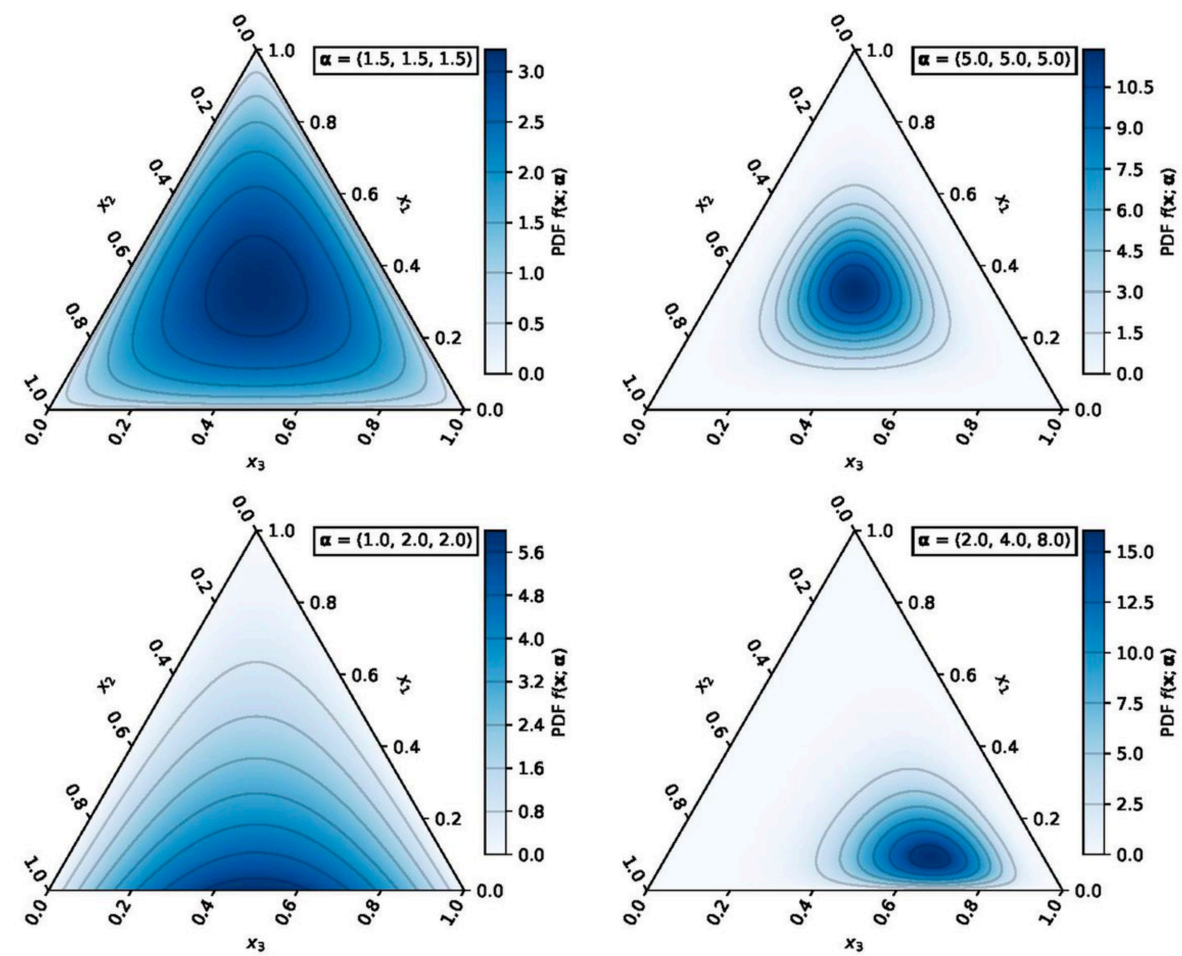
Many, many variations…

Common aspects:

- Bayesian estimation
- Unigram feature set
- Rarely stop words

No rule for # of topics

User tests many iterations



(Blei, Ng & Jordan, 2003; Blei & Lafferty, 2007)

Topic Models

How to choose the topic labels?

Topic Models

How to choose the topic labels? Read! And follow your heart

Topic Models

How to choose the topic labels? Read! And follow your heart

How to choose the example documents?

Topic Models

How to choose the topic labels? Read! And follow your heart

How to choose the example documents?

- Theta matrix tells which documents have most of each topic

Topic Models

How to choose the topic labels? Read! And follow your heart

How to choose the example documents?

- Theta matrix tells which documents have most of each topic

How to choose the example words? More complicated….

Topic Models

How to choose the topic labels? Read! And follow your heart

How to choose the example documents?
- Theta matrix tells which documents have most of each topic

How to choose the example words? More complicated….
Frequency: proportion of word within topic
- some words are very common, occur in multiple topics

Topic Models

How to choose the topic labels? Read! And follow your heart

How to choose the example documents?
- Theta matrix tells which documents have most of each topic

How to choose the example words? More complicated….

Frequency: proportion of word within topic
- some words are very common, occur in multiple topics

Exclusivity: ratio of word in topic vs other topics
- most diagnostic words are very rare

Topic Models

How to choose the topic labels? Read! And follow your heart

How to choose the example documents?
- Theta matrix tells which documents have most of each topic

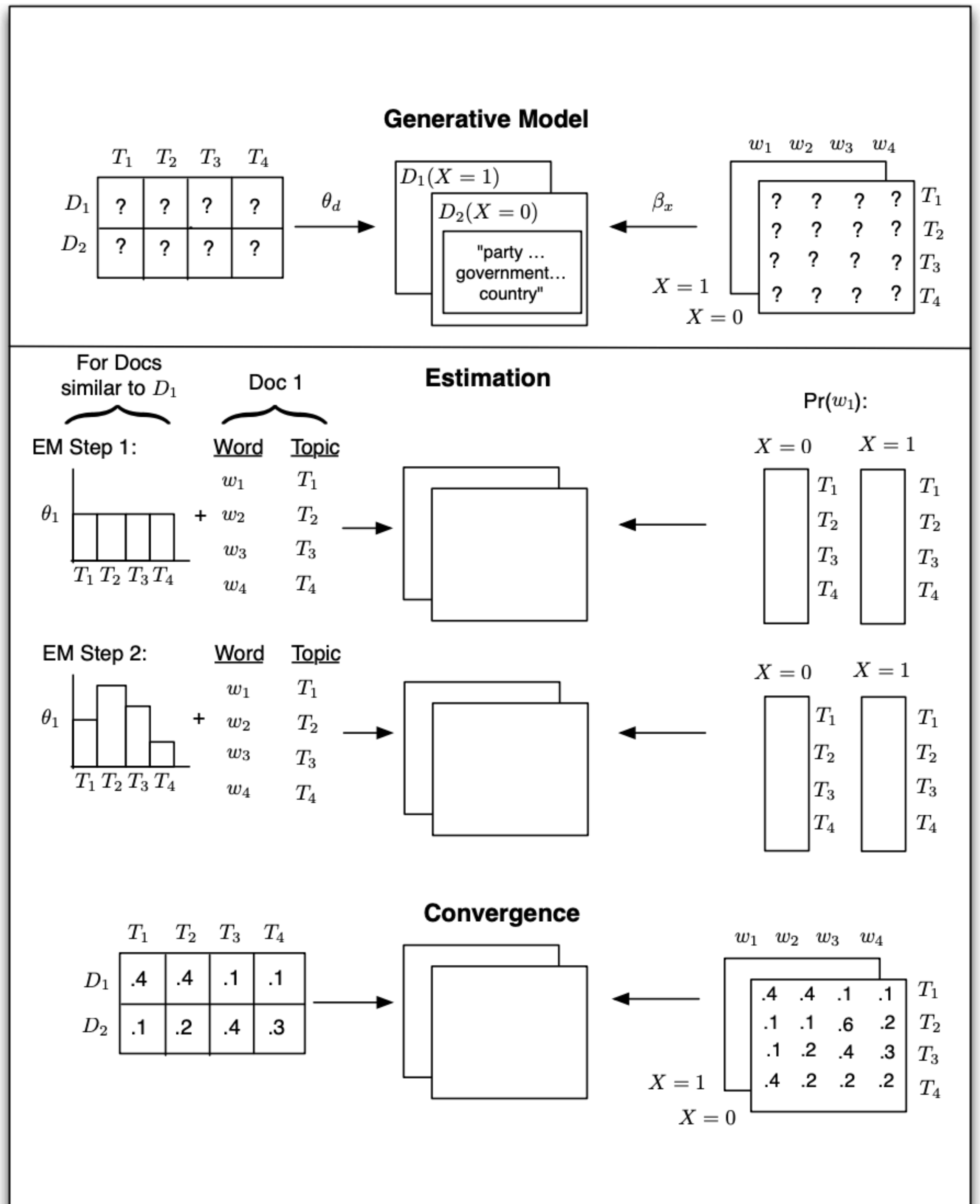
How to choose the example words? More complicated….

Frequency: proportion of word within topic
- some words are very common, occur in multiple topics

Exclusivity: ratio of word in topic vs other topics
- most diagnostic words are very rare

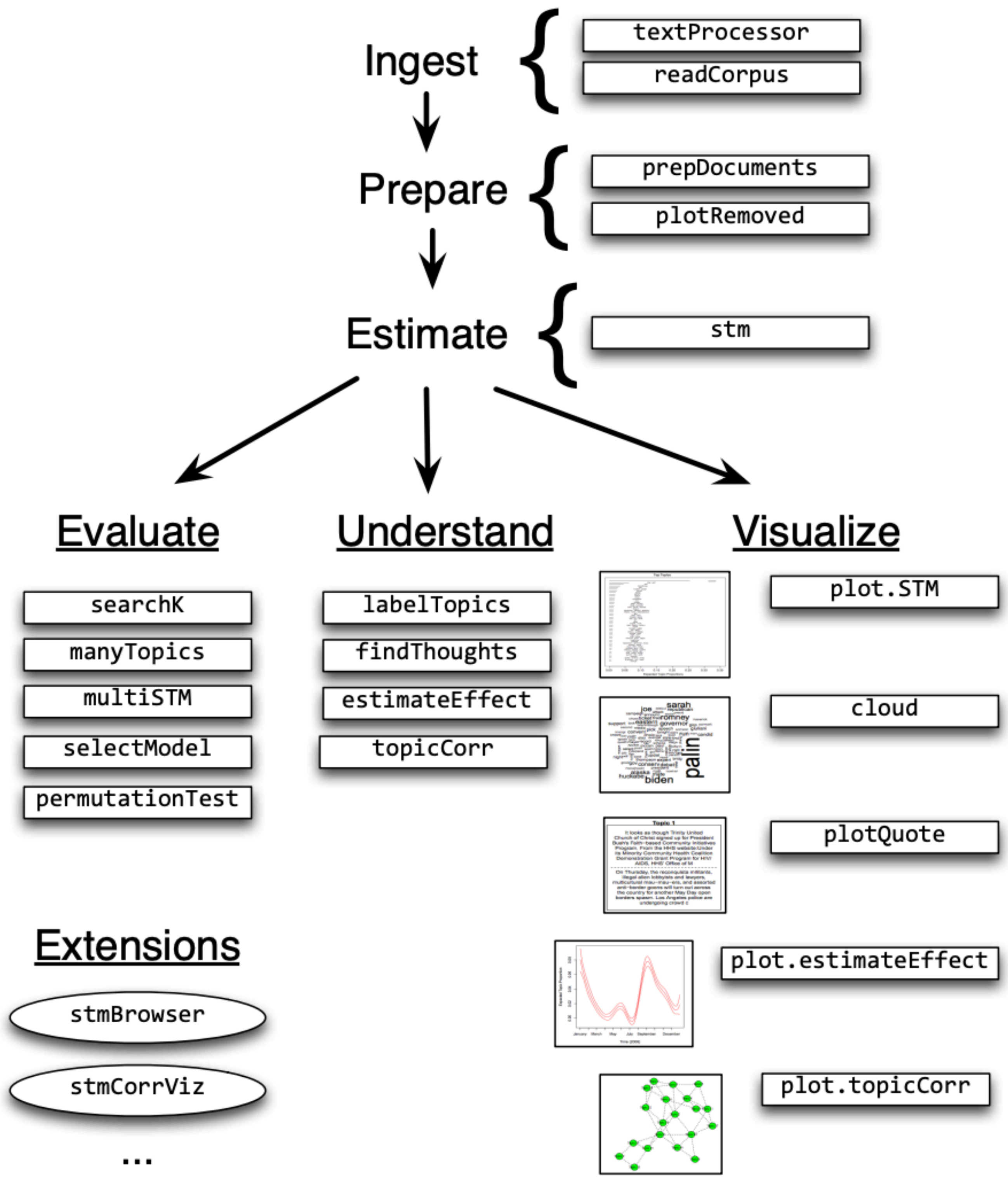
FREX - rank-order combination of Frequency & Exclusivity

The stm Package



(Roberts, Stewart & Tingley, 2017)

The stm Package



(Roberts, Stewart & Tingley, 2017)

Massive Open Online Courses

Massive Open Online Courses

EDUCATION LIFE

The New York Times

The Year of the MOOC

By LAURA PAPPANO NOV. 2, 2012

Massive Open Online Courses

EDUCATION LIFE

The New York Times

The Year of the MOOC

By LAURA PAPPANO NOV. 2, 2012

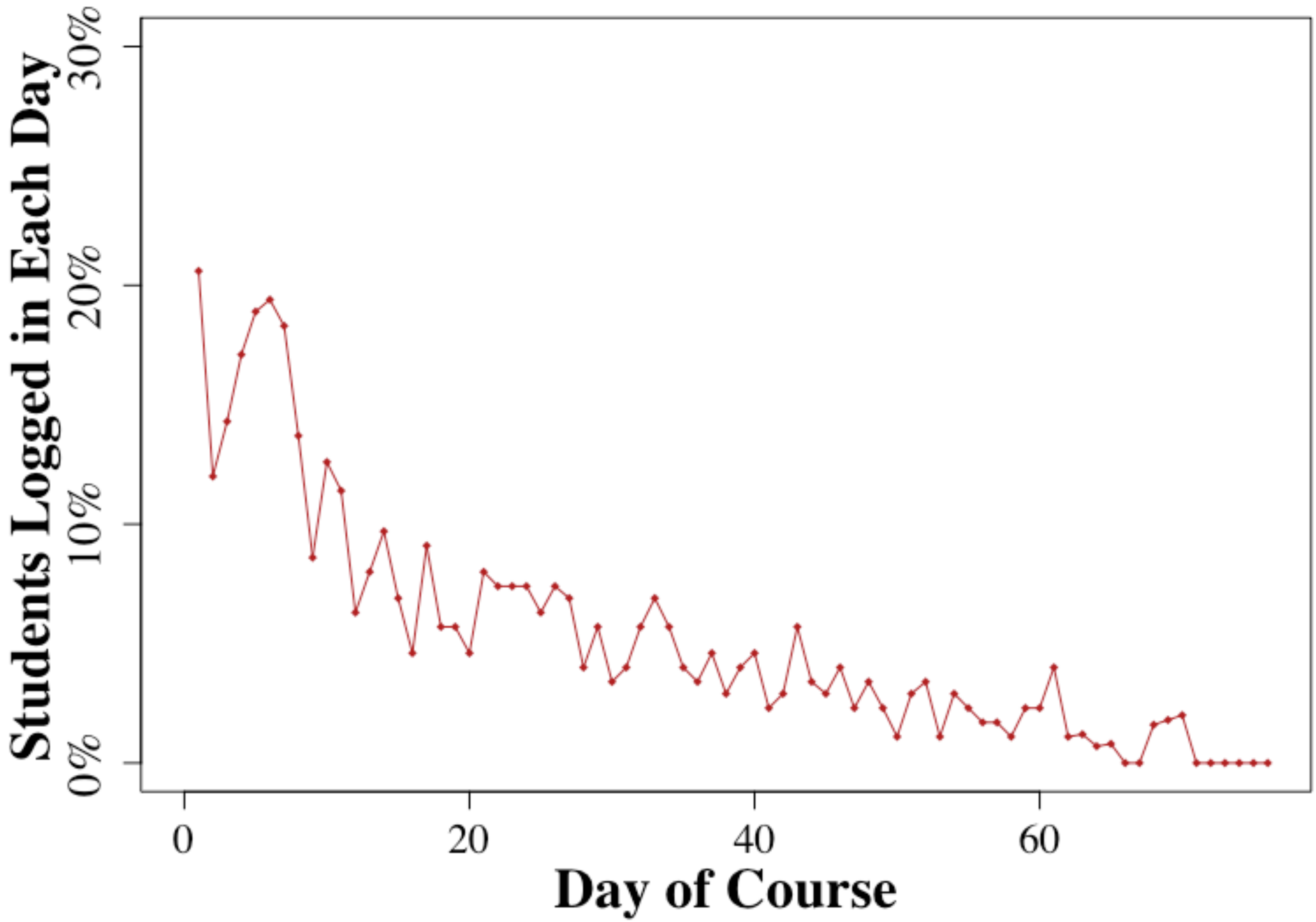
THE NEW YORKER

MARIA KONNIKOVA

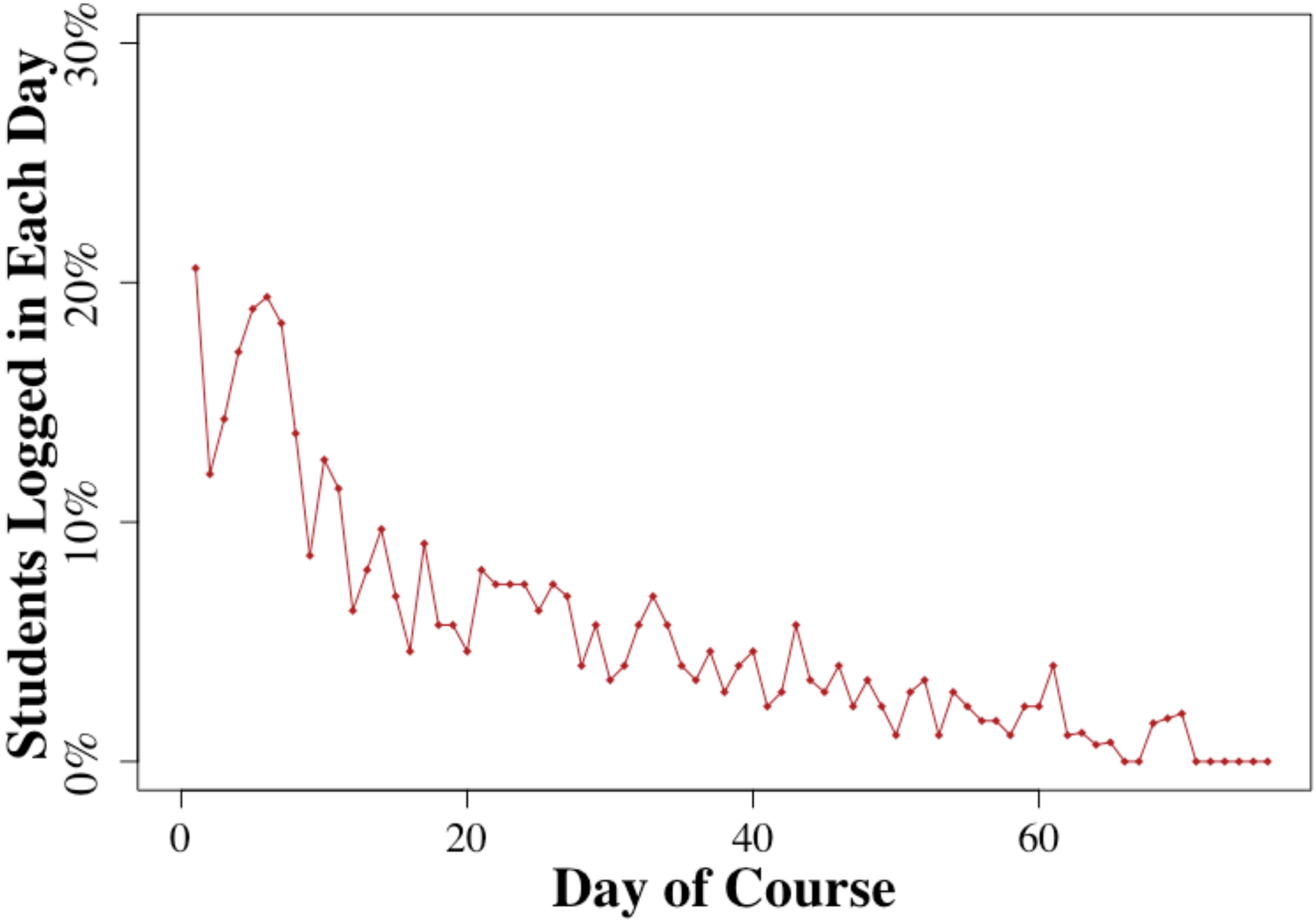
WILL MOOCS BE FLUKES?

By Maria Konnikova, NOVEMBER 7, 2014

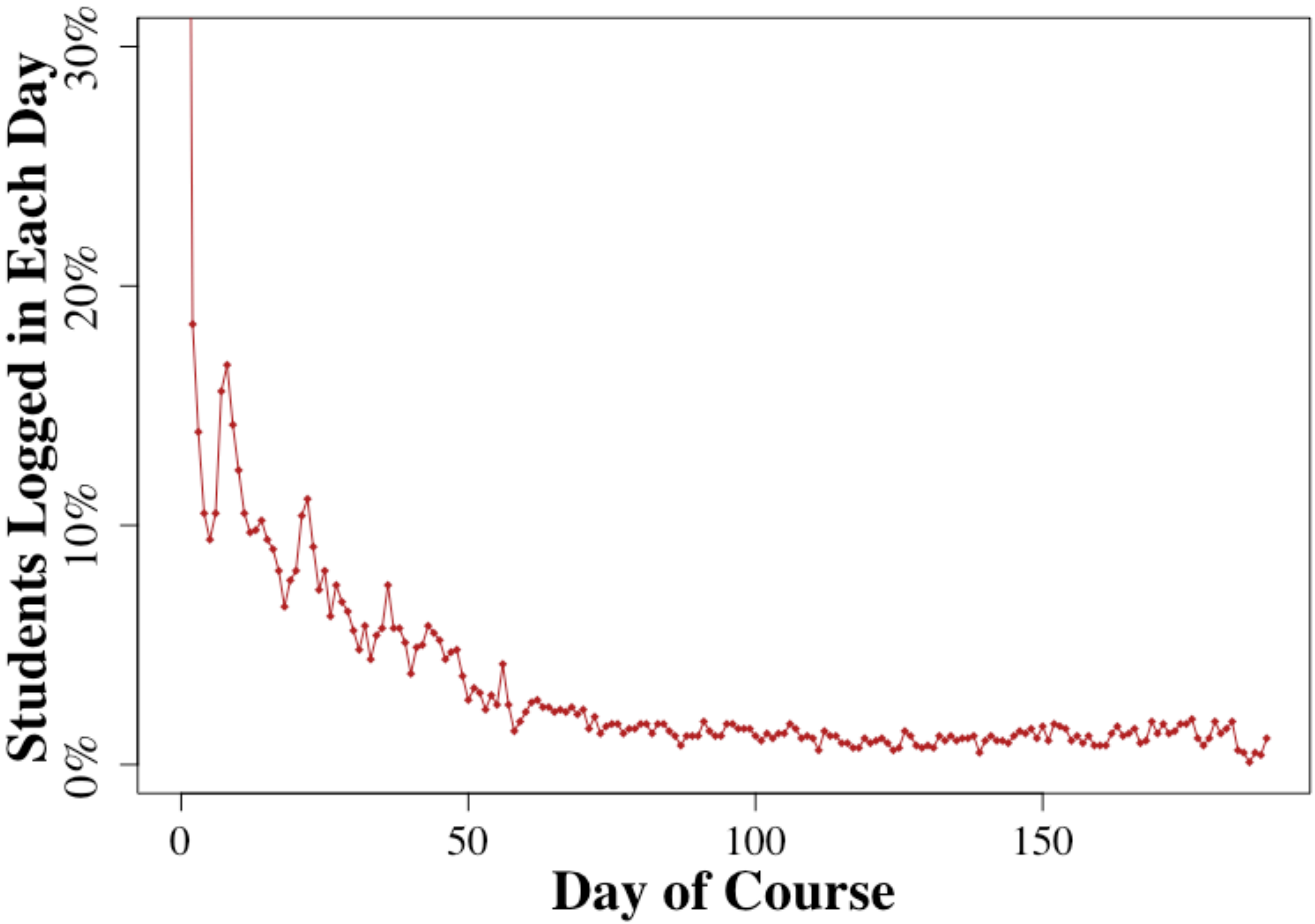
Massive Open Online Courses



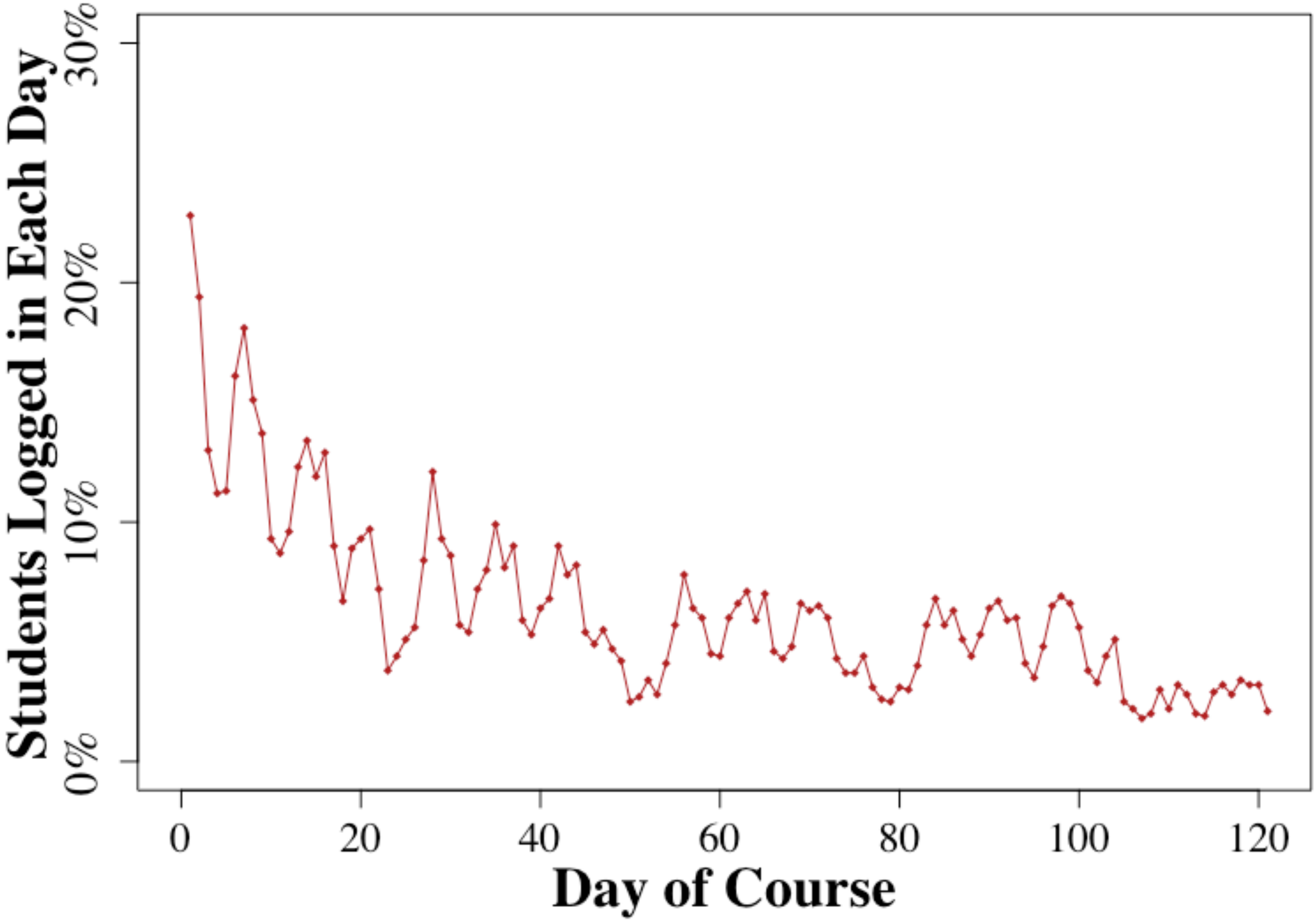
Business of Healthcare



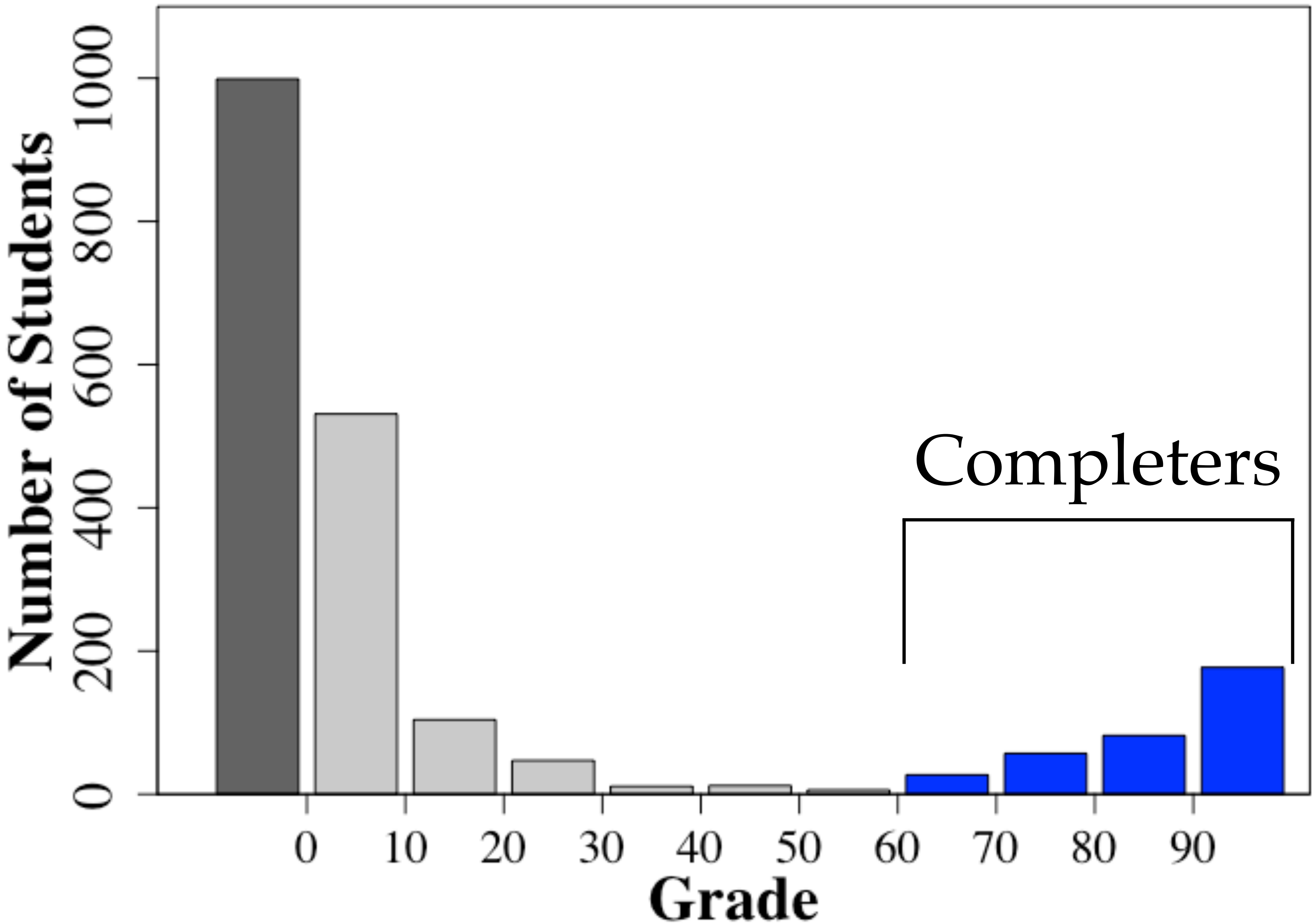
Introduction to Biochemistry



American Government

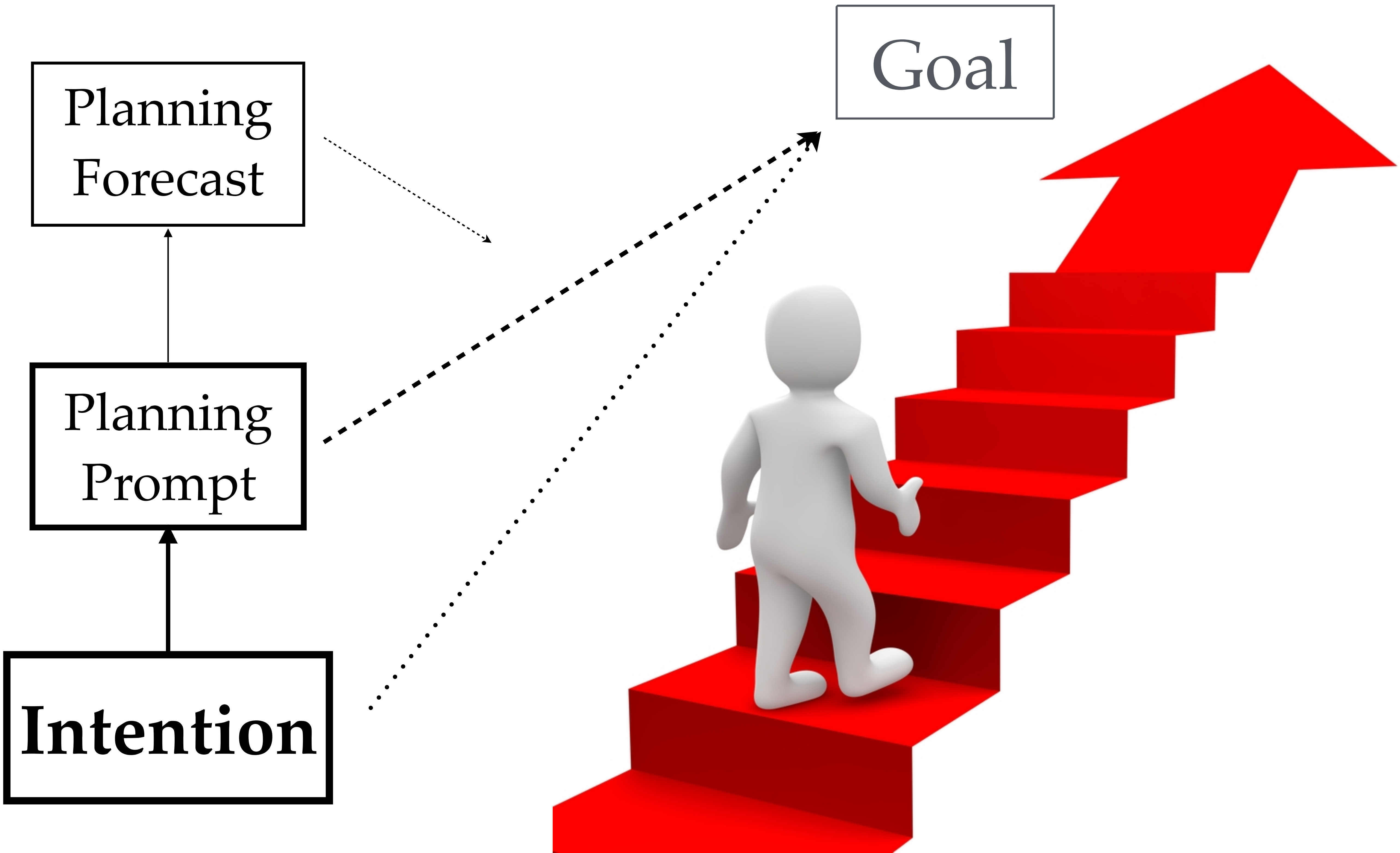


Outcomes



Number of Students
0
200
400
600
800
1000
0
10
20
30
40
50
60
70
80
90
Grade
Completers

Two Reasons to Prompt Plans



Planning Prompts

In the space below, write down some of your plans to learn. For example, try to specify:

a) **When and where** do you plan to spend time engaging the course content?
b) **What specific steps** you will take to ensure you complete the required course work?
c) **How will you respond** to obstacles that you might encounter during the course?

The Plans People Make

“I'm going to specifically set aside time to focus on coursework. Set study hours for myself. Create an environment with no distractions”

“I will study at home and office. Watch videos, study materials and contribute to discussions. Post questions to staff and co-participants”

“I'm gonna do whatever, whenever.”

Topics in Plan-Making

Topics in Plan-Making

“I will study at home and office. Watch videos, study materials and contribute to discussions. Post questions to staff and co-participants”

Topics in Plan-Making

“I will study at home and office.
Watch videos, study materials and
contribute to discussions. Post
questions to staff and co-participants”

“ Reading about the topics covered.
Watching the videos and lectures.
Search for the topics individually
to completely understand it”

"read course materials, complete
quiz and assignments, video lectures,
read more bibliography,
participate on the discussion board”

Topics in Plan-Making

“I will study at home and office. **Watch** **videos**, **study** materials and contribute to discussions. Post **questions** to staff and **participants**”

“ **Reading** about the **topics** covered. **Watching** the **videos** and **lectures**. Search for the **topics** individually to completely **understand** it”

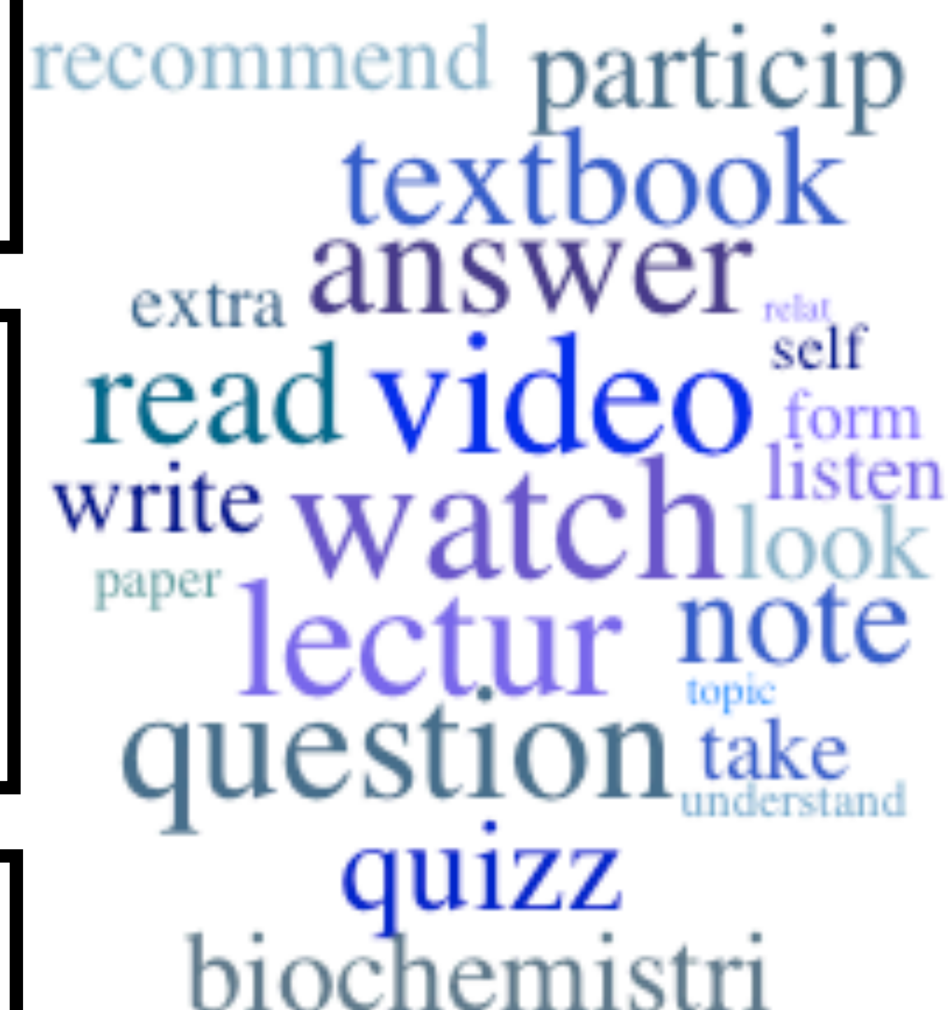
"**read** course materials, complete **quiz** and assignments, **video lectures**, **read** more bibliography, **participate** on the discussion board”

Topics in Plan-Making

"I will study at home and office. **Watch videos**, **study** materials and contribute to discussions. Post **questions** to staff and **participants**"

" **Reading** about the **topics** covered. **Watching** the **videos** and **lectures**. Search for the **topics** individually to completely **understand** it"

"**read** course materials, complete **quiz** and assignments, **video lectures**, **read** more bibliography, **participate** on the discussion board"



Topics in Plan-Making

“I'm going to specifically set aside time to focus on coursework. Set study hours for myself. Create an environment with no distractions”

Topics in Plan-Making

“I'm going to specifically set aside time to focus on coursework. Set study hours for myself. Create an environment with no distractions”

“Set aside specific days and times during the week for study. Use weekends to catch up - if needed. Eliminate distractions - no phones, television during study time”

"Every other day, practice during lunch and T.A period. Set reminders on my phone of specific deadlines”

Topics in Plan-Making

"I'm going to **specifically set aside time** to focus on coursework. Set study hours for myself. Create an environment **with no distractions**"

"**Set aside specific days** and **times** during the week for study. Use weekends to **catch up** - if **needed.** Eliminate **distractions** - no **phones**, **television** during study **time**"

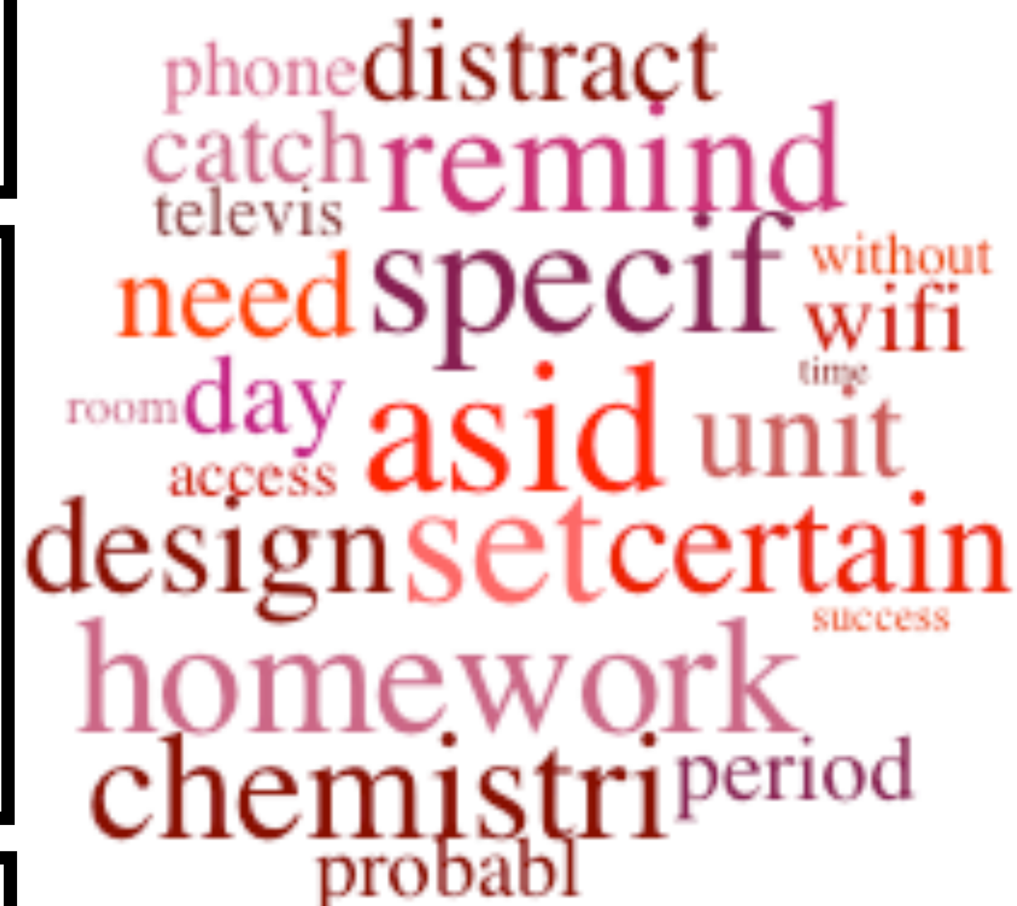
"Every other **day**, practice during lunch and T.A period. **Set reminders** on my **phone** of **specific** deadlines"

Topics in Plan-Making

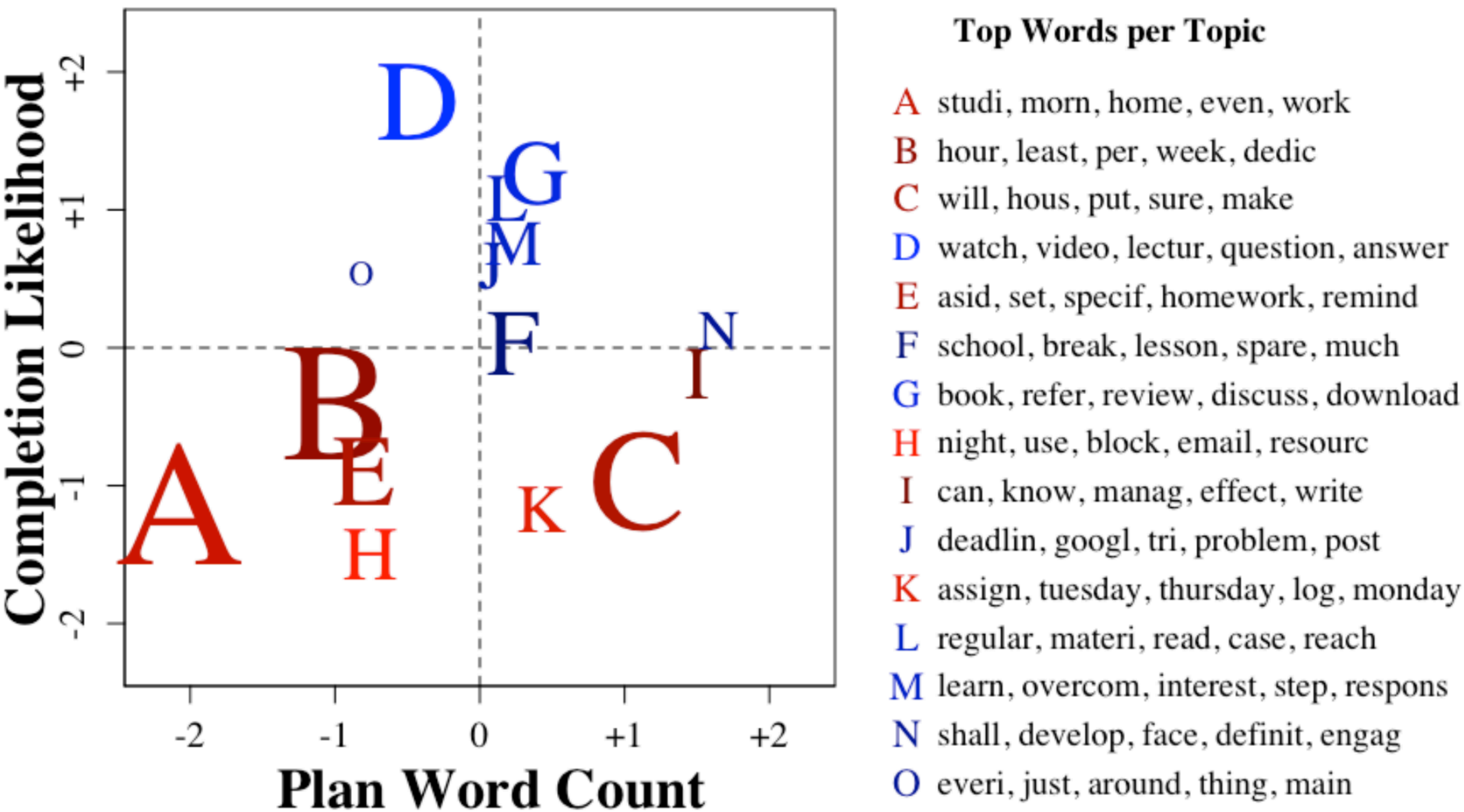
"I'm going to **specifically set aside time** to focus on coursework. Set study hours for myself. Create an environment **with no distractions**"

"**Set aside specific days** and **times** during the week for study. Use weekends to **catch up** - if **needed.** Eliminate **distractions** - no **phones**, **television** during study **time**"

"Every other **day**, practice during lunch and T.A period. **Set reminders** on my **phone** of **specific** deadlines"

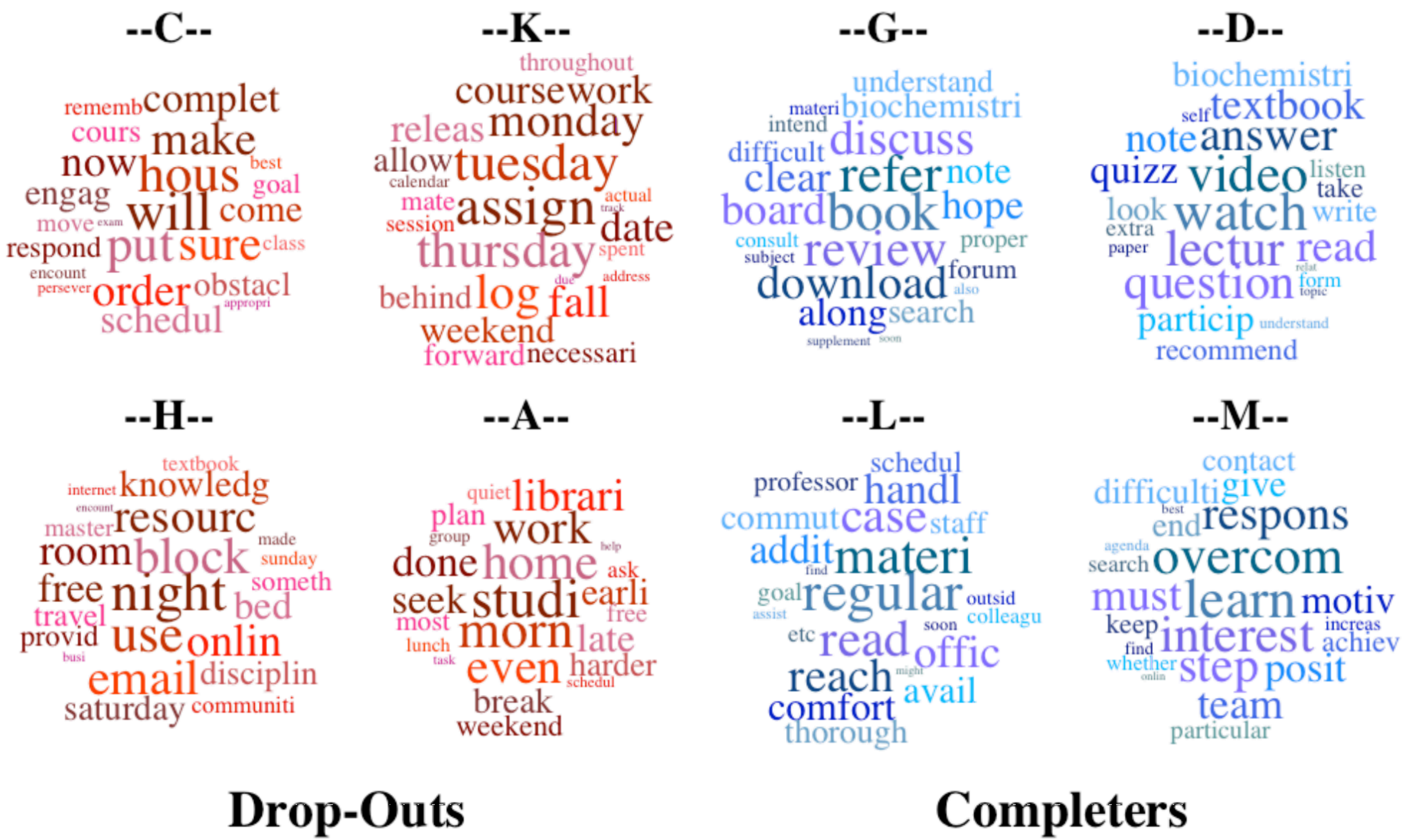


Topic Models



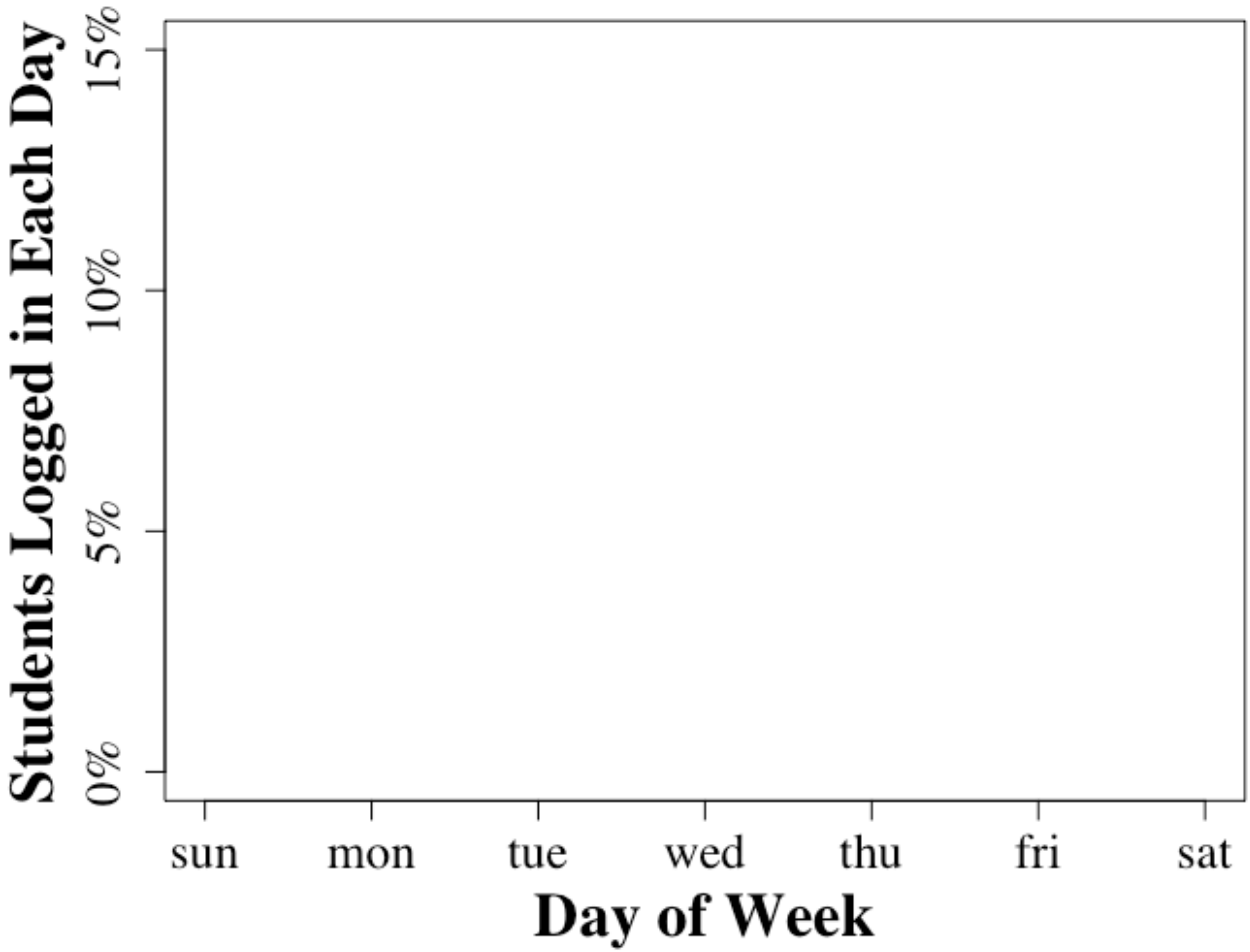
(Yeomans & Reich, 2017)

Topic Models

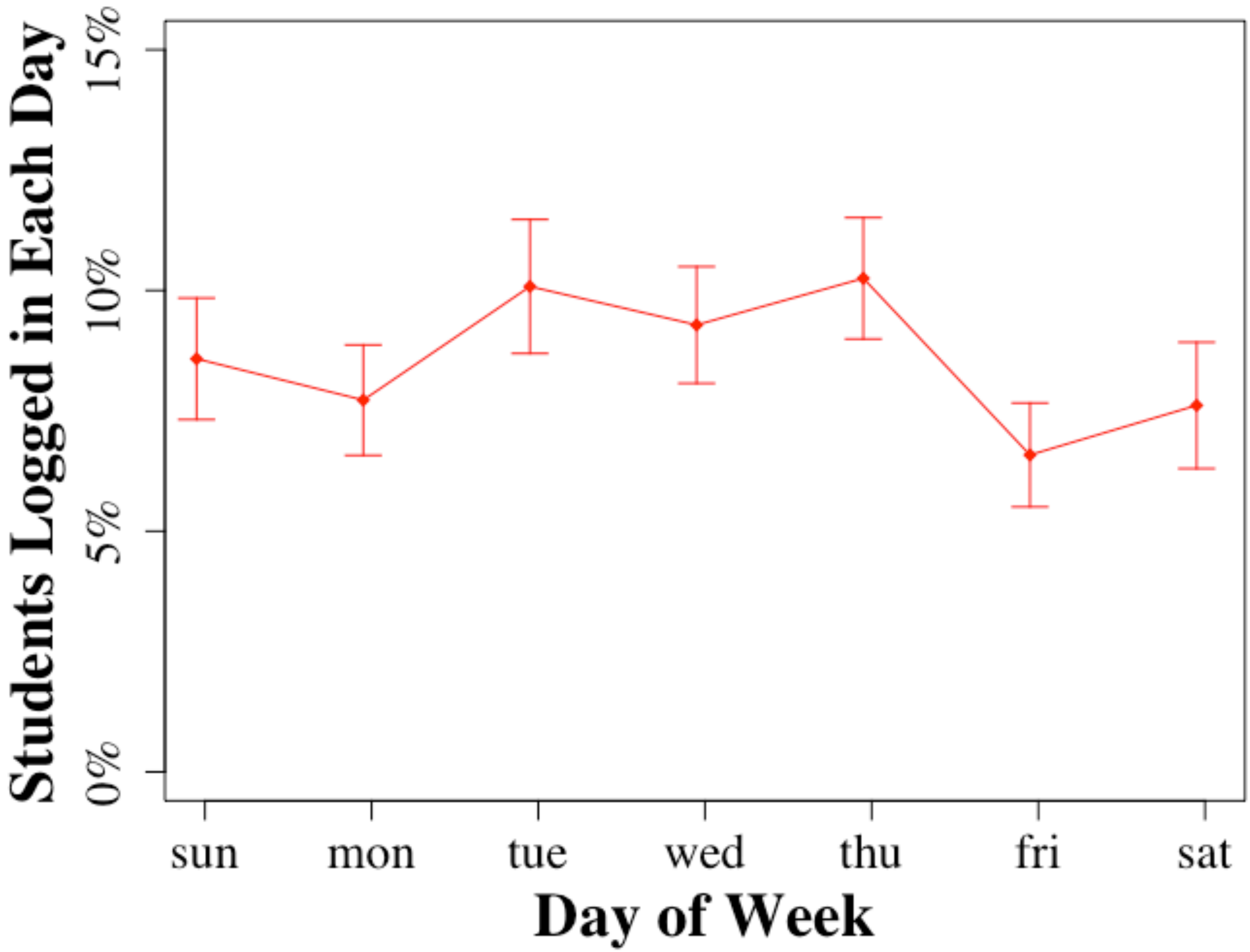


(Yeomans & Reich, 2017)

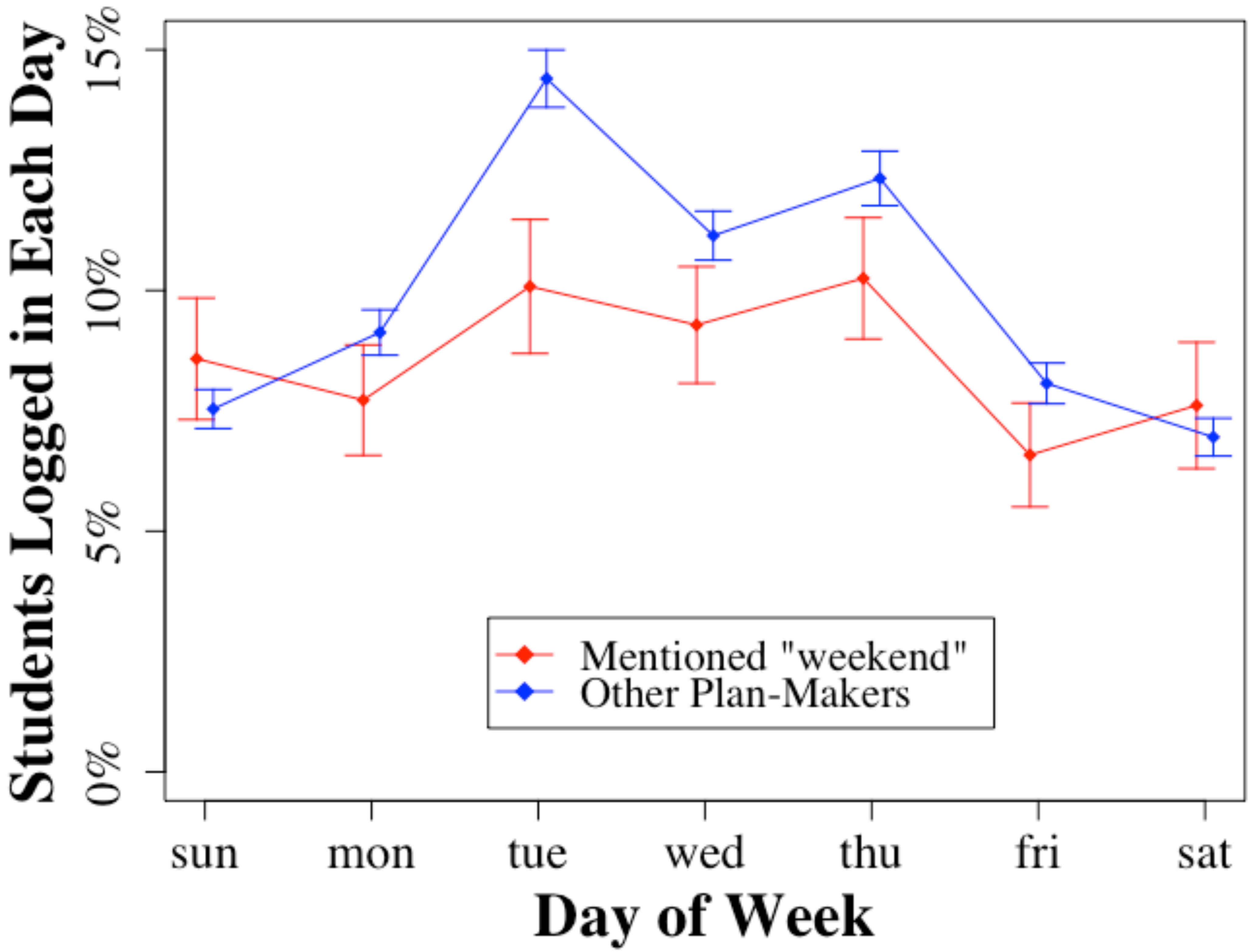
Planning for the Weekend



Planning for the Weekend



Planning for the Weekend



Topic Models

Topic Models

Advantages

Unsupervised - learns within-domain
Interpretable - improves reading
Focuses on main nouns/verbs that define domain
“Generative” - simulates writing process

Topic Models

Advantages

Unsupervised - learns within-domain
Interpretable - improves reading
Focuses on main nouns/verbs that define domain
“Generative” - simulates writing process

Limitations

How to set number of topics?
Topics may not capture quantities of interest
Removes stop words, word order, etc…

Topic Models

Advantages

Unsupervised - learns within-domain
Interpretable - improves reading
Focuses on main nouns/verbs that define domain
“Generative” - simulates writing process

Limitations

How to set number of topics?
Topics may not capture quantities of interest
Removes stop words, word order, etc…

Recommendations

Use stm (but estimate covariates after!)
Best for longer, single-author documents (e.g. news articles)
Much better for interpretation than estimation