

Predicting State Legislative Agendas

• • •
Mark Mummert

Goal:

Use topic modeling and regression to predict the distribution of state legislative agendas by topic.

Outline

Acquiring Data

Data Exploration

Topic Modeling with
LDA

Modeling bill
introduction numbers
—

Data Acquisition and Exploration

Data Acquisition:

API Requests from
LegiScan

Title

Description

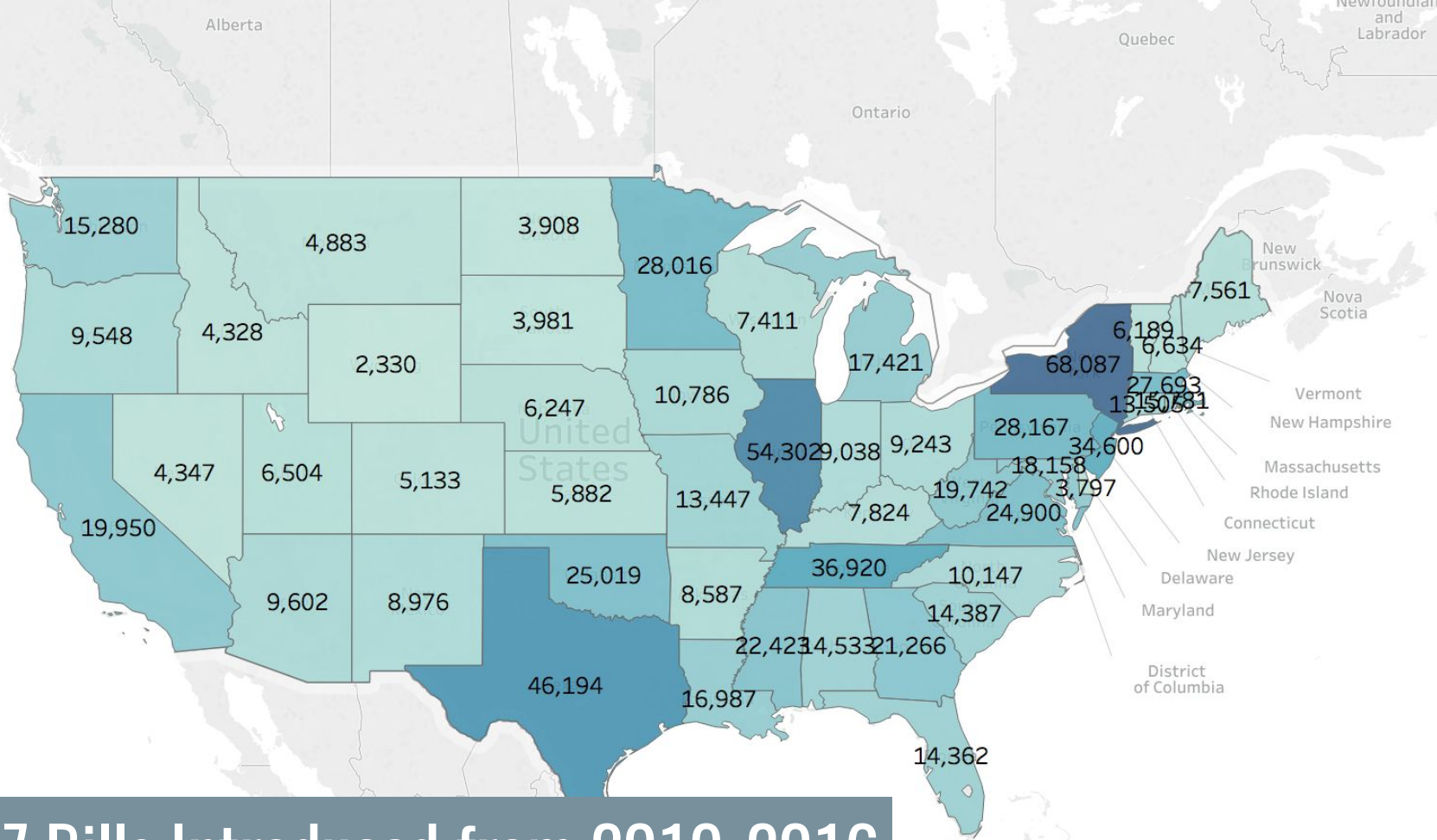
Last Action

Date of Last Action

Session ID

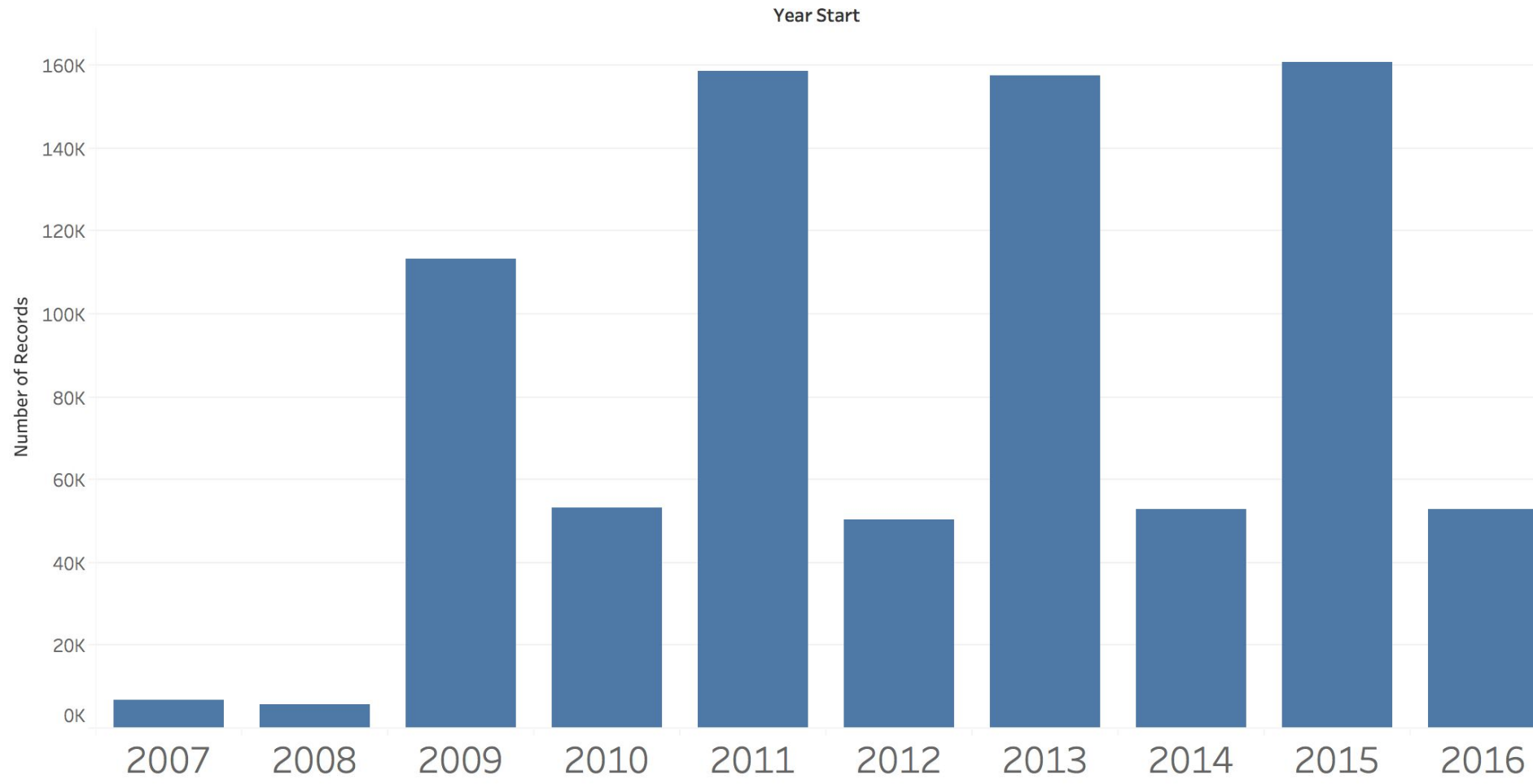
State

Session Year



812,307 Bills Introduced from 2010-2016

Bills Introduced Per Year



Topic Modeling with LDA

What is LDA?

- Bayesian model that produces topics from text
- Assign each topic based on key words most associated with the topic
- Unsupervised
 - Input number of topics
 - Can be hard to 'fit'

Implementation

'An act to add Article 3.5 (commencing with Section 221.2) to Chapter 2 of Part 1 of Division 1 of Title 1 of the Education Code, relating to schools.'



[u'act',
u'add',
u'article',
u'35',
u'commencing',
u'section',
u'2212',
u'chapter',
u'part',
u'division',
u'title',
u'education',
u'code',
u'relating',
u'school']



[(143, 1),
(258, 1),
(304, 1),
(607, 1),
(650, 1),
(697, 1),
(704, 1),
(908, 1),
(1000, 1),
(1014, 1),
(1207, 1),
(1222, 1),
(1231, 1),
(1277, 1),
(18854, 1)]

LDA Results - Topic Results

Education

Resolution
School
Commending

Transpo

Fund
County
Transportation

???

Program
Service
Child

???

Providing
Act
Related

Education

School
Amends
Act

???

Title

Chapter
tca

Vehicle Regs

Vehicle
Person
License

Taxes

Tax
Year
Property

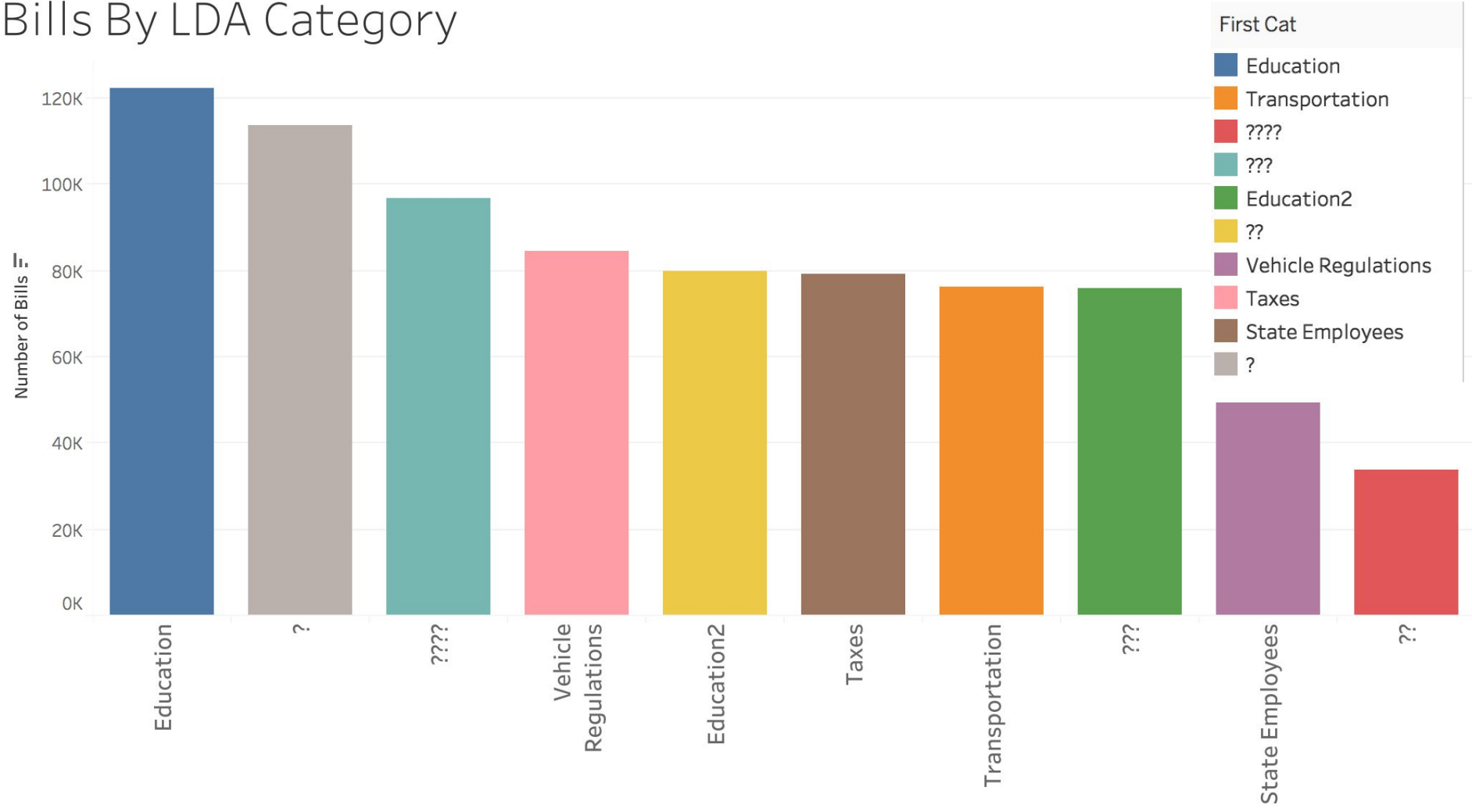
State Workers

Public
Employee
System

???

Provides
Requires
Certain

Bills By LDA Category



Predicting Distribution of Agendas

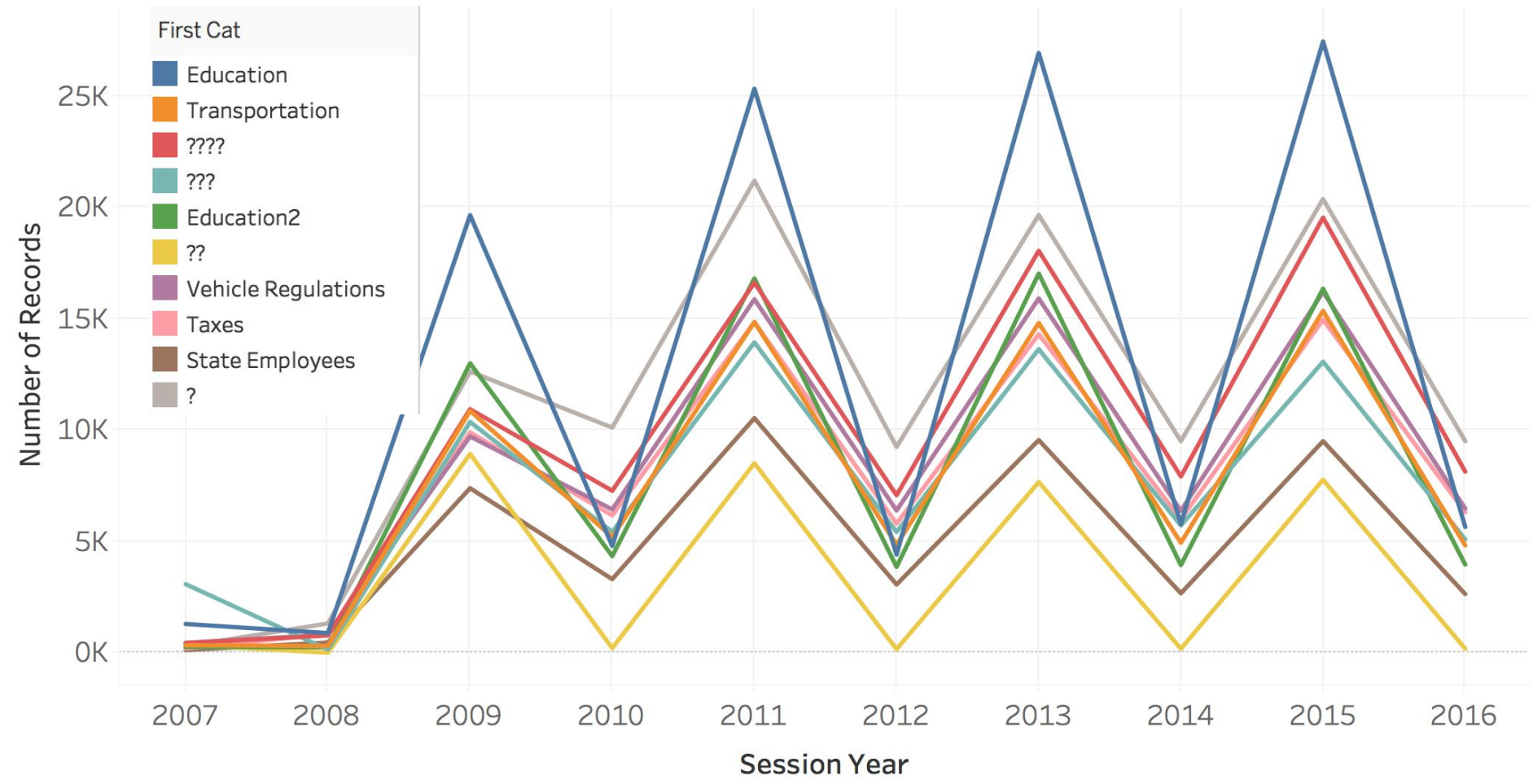
Modeling the Legislative Agenda

Goal: To predict the number of bills in each category based on the number from two years ago.

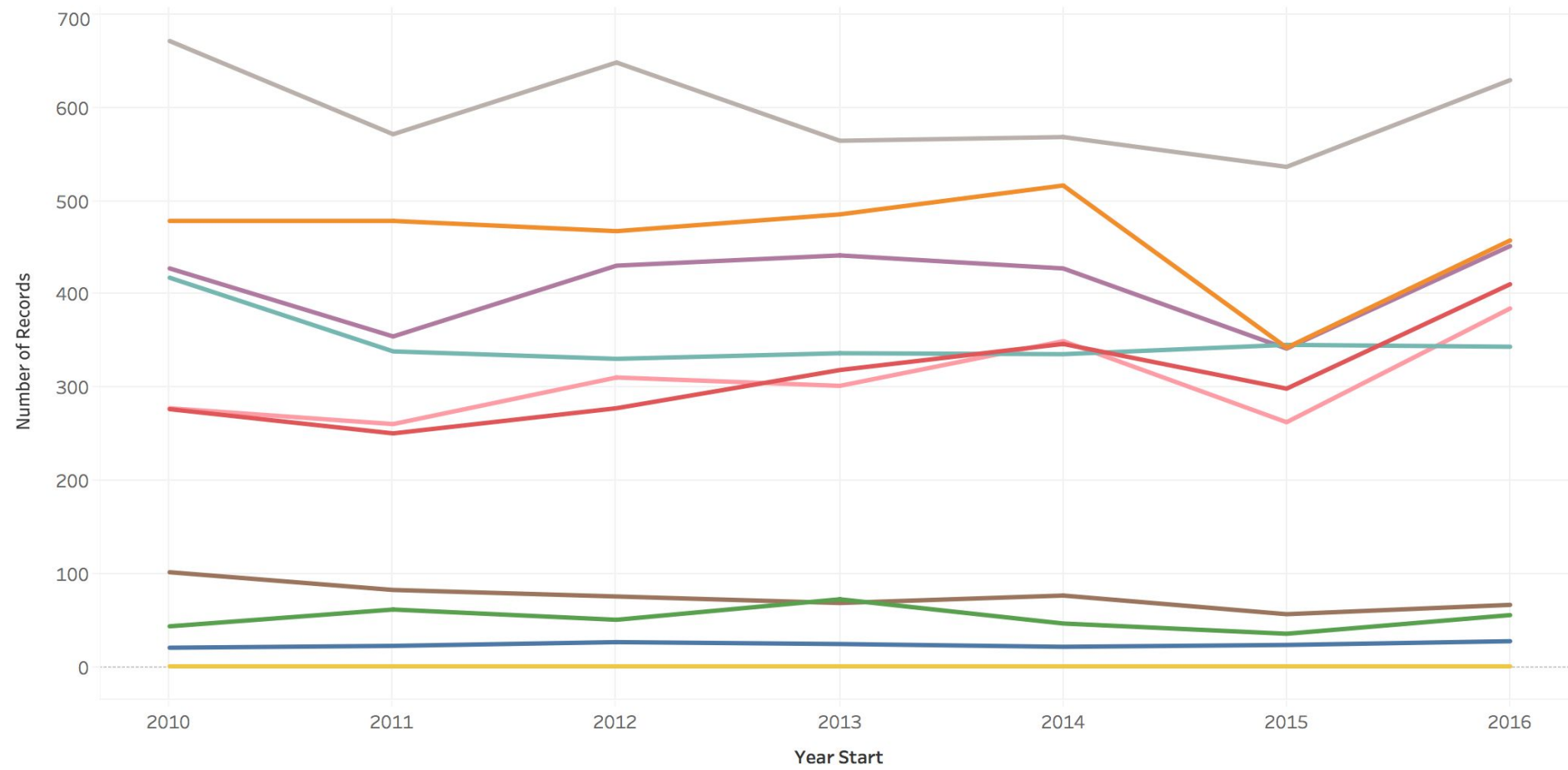
Model Selection

- Poisson and Negative Binomial Regression
 - Bad performance - the distribution did not fit the model
- Linear Regression performed best
 - R^2 .64

Bill Subject by Year



Bills in Maryland by Category



Conclusions

- LDA categories are not always interpretable
- No significant change in topic distribution by year
 - Relatively easy to predict

Next Steps/Future Goals

- Acquire full text for topic modeling
- Focus on a specific state
- Examine bill similarities across multiple states
- Conduct more ‘fitting’ with LDA

Questions