

Métodos Computacionales para Políticas Públicas (MCP)

Santiago Matallana

santiago.matallana@gmail.com

2019 – II

Hoy (esta semana)

1. Presentación (de ustedes y mía)
2. Encuesta en línea
3. ¿En qué nos metimos?
4. Ejemplos de lo que aprenderemos
5. Metodología y evaluación
6. Cronograma
7. Tarea

Hoy (esta semana)

1. Presentación (de ustedes y mía)
2. Encuesta en línea
3. ¿En qué nos metimos?
4. Ejemplos de lo que aprenderemos
5. Metodología y evaluación
6. Cronograma
7. Tarea

Presentación (ustedes)

Nombre

Profesión

“¿Estudias o trabajas?”

¿Por qué está acá? ¿Qué espera?

Presentación (yo)

<https://co.linkedin.com/in/smatallana>



Presentación (yo)

En UChicago:

1. City Lab
2. Computer Science with Applications
3. Text Mining for Public Policy
4. Computation for Public Policy
5. GIS Applications in the Social Sciences
6. Machine Learning for Public Policy
7. Big Data

Y MOOCs:

1. The Data Scientist Toolbox – Johns Hopkins
2. Introduction to R – Johns Hopkins
3. R Programming – Johns Hopkins
4. Programming for Everybody (Python) – Michigan
5. Python Data Structures – Michigan
6. ...

Hoy (esta semana)

1. Presentación (de ustedes y mía)
2. Encuesta en línea
3. ¿En qué nos metimos?
4. Ejemplos de lo que aprenderemos
5. Metodología y evaluación
6. Cronograma
7. Tarea

Encuesta en línea

http://bit.ly/mcpp192_encuesta

Hoy (esta semana)

1. Presentación (de ustedes y mía)
2. Encuesta en línea (Quiz 1)
3. ¿En qué nos metimos?
4. Ejemplos de lo que aprenderemos
5. Metodología y evaluación
6. Tarea

¿En qué nos metimos?

“MCPP” es un curso sobre **herramientas computacionales**, útiles (o no) en problemas de política pública (y en muchos otros ámbitos).

El análisis y la toma de decisiones de política pública se basan crecientemente en **datos y evidencia**:

- recolectar
- almacenar
- procesar
- estructurar
- compartir
- visualizar
- analizar

Intersección reciente entre ciencias de la computación y políticas públicas (ciencias sociales).

¿En qué nos metimos?

Algunos ejemplos:

<https://capp.uchicago.edu/>

<https://dssg.uchicago.edu/>

https://idss.mit.edu/academics/ses_doc/

<http://gking.harvard.edu/files/gking/files/censored.pdf>

¿En qué nos metimos?

Es un **curso práctico**: vamos a **aprender haciendo**.

La idea de que los *policy-makers* simplemente le delegan el trabajo sucio (procesamiento y análisis de datos) a unos *nerds* (“el chino de los datos”) es una caricatura del pasado.

Al final del semestre, tendremos habilidades nuevas:

- **Python**
- Github
- *Web scraping*
- Análisis de textos
- Procesamiento de datos
- Visualización de datos
- R (algo)
- Análisis de redes sociales (algo)
- QGIS (algo)
- *Data Mining / Big Data / Data Science / Machine Learning* (conceptos)

Analítica/ciencia de datos: un mundo enorme. Haremos **renuncias**.

Hoy (esta semana)

1. Presentación (de ustedes y mía)
2. Encuesta en línea (Quiz 1)
3. ¿En qué nos metimos?
4. Ejemplos de lo que aprenderemos
5. Metodología y evaluación
6. Tarea

Ejemplos de lo que aprenderemos

**PRESIDENCIA DE LA REPÚBLICA**



English



InicioPresidenciaEl PresidenteNormativaSala de PrensaAtención Ciudadanía

Presidencia > Discursos > 2015

Sistema Informativo del Gobierno

2015

Enero

Febrero

Marzo

Abril

Mayo

Junio

Julio

Agosto

Septiembre

Octubre

Noviembre

Diciembre

AÑOS ANTERIORES

2015

2014 (Desde el 7 de agosto)

2014(Hasta el 6 de agosto)

2013

2012

2011

2010

SALA DE PRENSA

NOTICIAS

FOTOS

AUDIOS

VIDEOS

DISCURSOS

SIG

Discursos - Diciembre de 2015

Palabras del Presidente de la República, Juan Manuel Santos, durante la inauguración de la Variante de Tocancipá
Tocancipá, Cundinamarca, 8 dic (SIG).

Palabras del Presidente Juan Manuel Santos en la instalación del conversatorio 'Los dividendos de la paz'
Bogotá, 9 dic (SIG).

Palabras del Presidente Juan Manuel Santos en la presentación del 'Informe Nacional de Derechos Humanos 2015' y lanzamiento del 'Plan Nacional de Acción sobre Derechos Humanos y Empresas'
Bogotá, 9 dic (SIG).

Declaración del Presidente Juan Manuel Santos en el Día Internacional de Lucha contra la Corrupción
Bogotá, 9 dic (SIG).

Palabras del Presidente Juan Manuel Santos en la posesión de los magistrados del Consejo de Gobierno de la Rama Judicial
Bogotá, 11 dic (SIG).

Declaración del Presidente Juan Manuel Santos luego de inaugurar el Centro Regional de Atención y Reparación a Víctimas
Cúcuta, 11 dic (SIG).

Palabras del Presidente Juan Manuel Santos en la inspección de obras del Proyecto de Reconstrucción de Gramalote
Gramalote, Norte de Santander, 11 dic (SIG).

Declaración de los cancilleres de Colombia, María Ángela Holguín y de España, José Manuel García-Margallo, al término de la reunión bilateral
Cartagena, 12 dic (SIG).

Palabras del Presidente Juan Manuel Santos en la ceremonia de ascensos de la Fuerza Aérea Colombiana (FAC)
Cali, 7 dic (SIG).

Palabras del Presidente Juan Manuel Santos en la ceremonia de ascensos del Ejército Nacional
Bogotá, 7 dic (SIG).

Palabras del Presidente Juan Manuel Santos en la ceremonia de ascensos de la Policía Nacional
Bogotá, 7 dic (SIG).

2,600+

Ejemplos de lo que aprenderemos

```
25 # Constants
26
27 MONTHS = ['Enero', 'Febrero', 'Marzo', 'Abril', 'Mayo', 'Junio', 'Julio',\
28           'Agosto', 'Septiembre', 'Octubre', 'Noviembre', 'Diciembre']
29 YEARS = ['2010', '2011', '2012', '2013', '2014', '2015']
30
31
32 # In[3]:
33
34 def get_html(url):
35     """
36     Requests url and parses text. Returns BeautifulSoup object.
37     """
38     resp = requests.get(url).text
39     return BeautifulSoup(resp, "lxml")
40
41
42 # In[4]:
43
44 def get_links(soup, regex):
45     """
46     Extracts url links to speeches. Returns a list of urls of speeches.
47     """
48     links = soup.find_all('a')
49     return re.findall(regex, str(links))
50
51
52 # In[5]:
53
54 def links_speeches(main_urls, regex):
55     """
56     Makes a list with the links to all speeches.
57     """
58     speeches_urls = []
59     for url in main_urls:
60         soup = get_html(url)
61         urls = get_links(soup, regex)
62         speeches_urls = speeches_urls + urls
63     return speeches_urls
64
65
66 # In[6]:
67
68 def get_date(speech_url):
69     """
70     Returns year, month, and day from url
71     """
72     date = re.findall('Paginas/{(8)}', speech_url)[0]
73     year = date[0:4]
74     month = date[4:6]
75     day = date[6:8]
76     return year, month, day
```

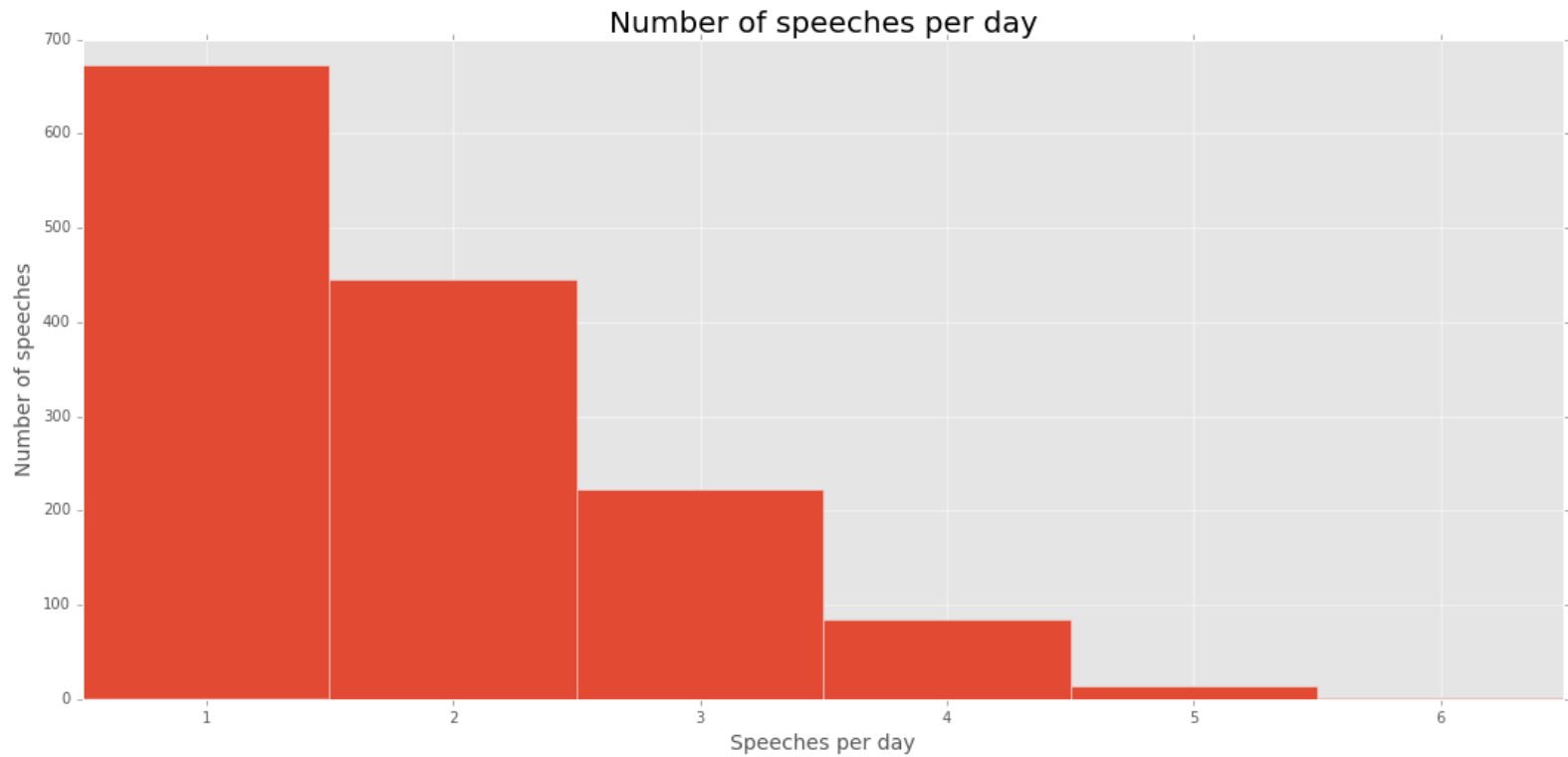
```
88 # In[8]:
89
90 def get_speech(soup):
91     """
92     Returns text of speech
93     """
94     return str(soup.find_all('p'))
95
96
97 # In[9]:
98
99 def get_location(soup):
100     """
101     Returns location of speech
102     """
103     text = soup.p.text
104     str1 = 'id="ctl00_PlaceHolderMain_ctl05_ControlWrapper_RichHtmlField" style'
105     str2 = 'id="ctl00_PlaceHolderMain_content_ControlWrapper_RichHtmlField" sty
106     if len(re.findall('(.+), [0-9].+', text)) != 0:
107         location = re.findall('(.+), [0-9].+', text)[0]
108     elif len(re.findall(str1 + '(.+), [0-9].+', str(soup))) != 0:
109         location = re.findall(str1 + '(.+), [0-9].+', str(soup))[0]
110     elif len(re.findall(str2 + '(.+), [0-9].+', str(soup))) != 0:
111         location = re.findall(str2 + '(.+), [0-9].+', str(soup))[0]
112     else:
113         location = 'N.A.'
114     location = location.strip('<strong>')
115     return location
116
117
118 # In[10]:
119
120 def remove_tags(text):
121     """
122     Strips text from html and other tags
123     """
124     return re.sub('<[^>+>|\\|\\|\\|\\u200b|\\n|\\xa0', '', text)
125
126
127 # In[12]:
128
129 def get_speech_elements(speech_url):
130     """
131     Returns
132     """
133     print('Speech url: ', speech_url)
134     soup = get_html(speech_url)
135     title = get_title(soup)
136     year, month, day = get_date(speech_url)
137     speech = remove_tags(get_speech(soup))
138     location = get_location(soup)
139     return [year, month, day, title, location, speech]
```

Ejemplos de lo que aprenderemos

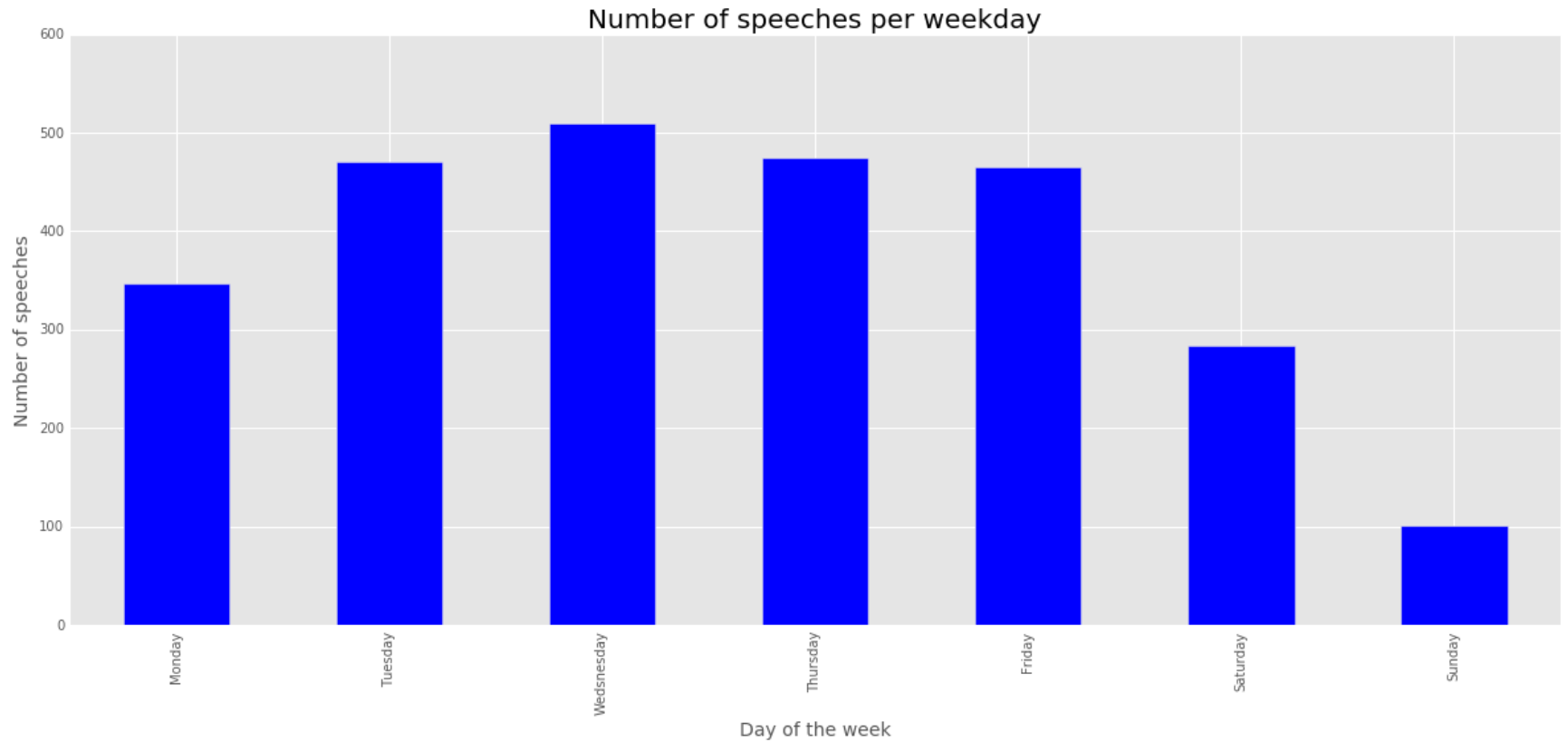
	Year	Month	Day	Title	Location	URL	Speech
0	2010	08	30	Palabras del Presidente Juan Manuel Santos Cal...	Bogotá	http://wsp.presidencia.gov.co/Prensa/2010/Agos...	Bogotá, 30 ago (SIG). “Muchos motivos nos llen...
1	2010	08	29	Palabras del Presidente de la República, Juan ...	Bogotá	http://wsp.presidencia.gov.co/Prensa/2010/Agos...	“Sólo la unidad del pueblo y la solidaridad de...
2	2010	08	28	Palabras del Presidente Juan Manuel Santos Cal...	Popayá	http://wsp.presidencia.gov.co/Prensa/2010/Agos...	Popayán, 28 ago (SIG). “Nuevamente ¡qué bueno ...

Speech	Date	no_punct	tokens	no_stopwords	nltk_Text
).- El siguiente es el discurso del presidente...	2010-08-07	el siguiente es el discurso del presidente ju...	[el, siguiente, es, el, discurso, del, preside...	[siguiente, discurso, presidente, juan, manuel...	(siguiente, discurso, presidente, juan, manuel...
“Quise venir de primero aquí a Guaranda, para ...	2010-08-08	quise venir de primero aquí a guaranda para da...	[quise, venir, de, primero, aquí, a, guaranda,...	[quise, venir, primero, aquí, guaranda, darle,...	(quise, venir, primero, aquí, guaranda, darle,...

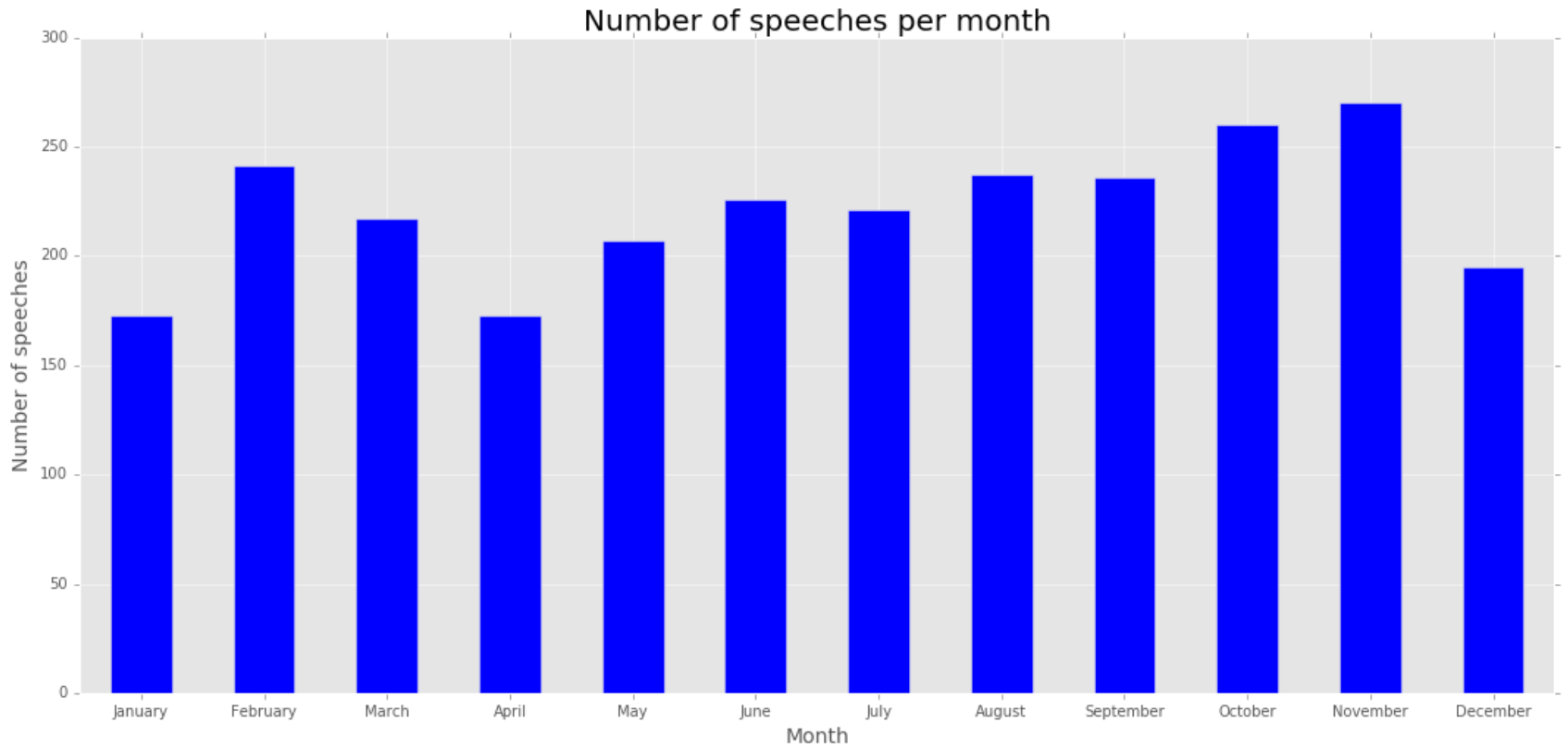
Ejemplos de lo que aprenderemos



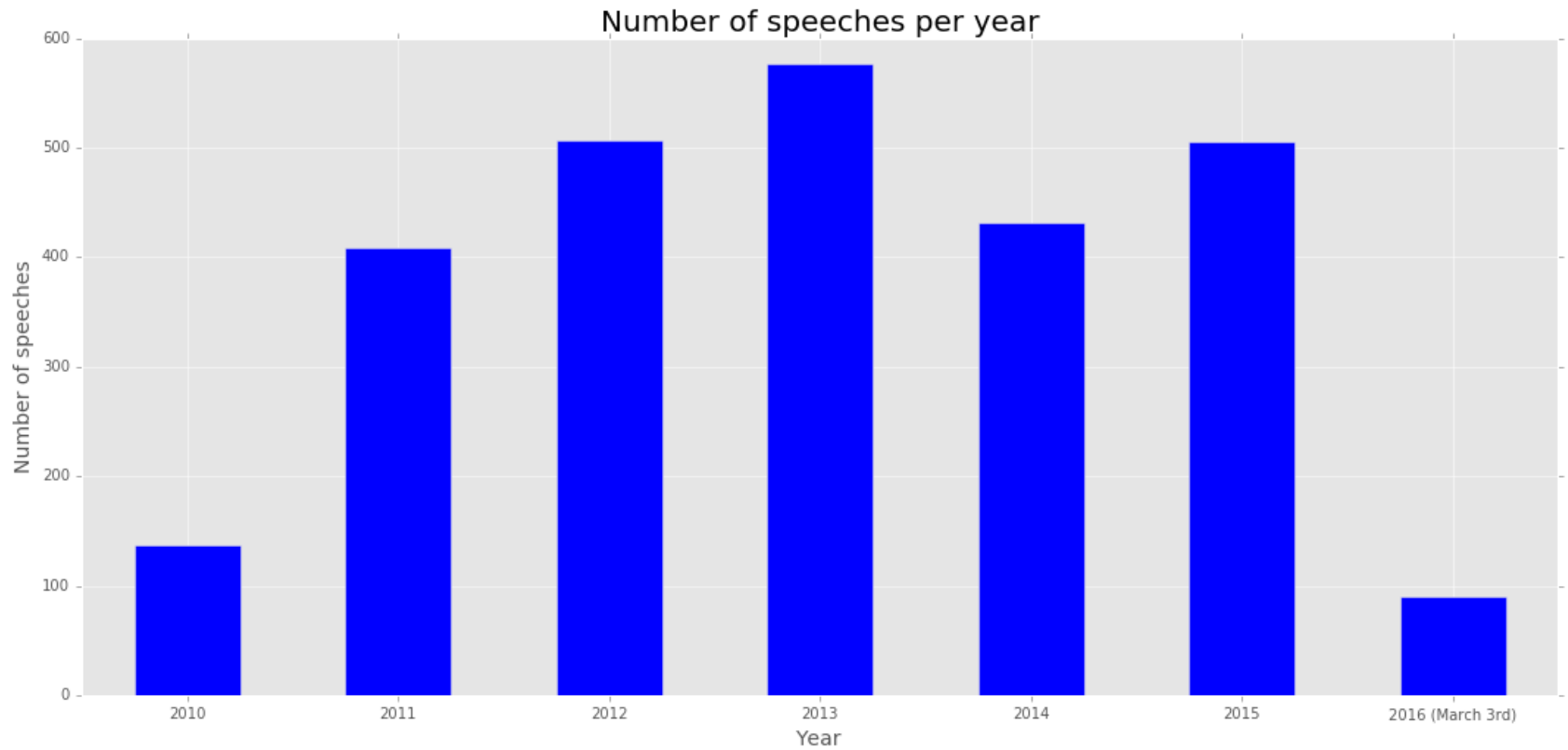
Ejemplos de lo que aprenderemos



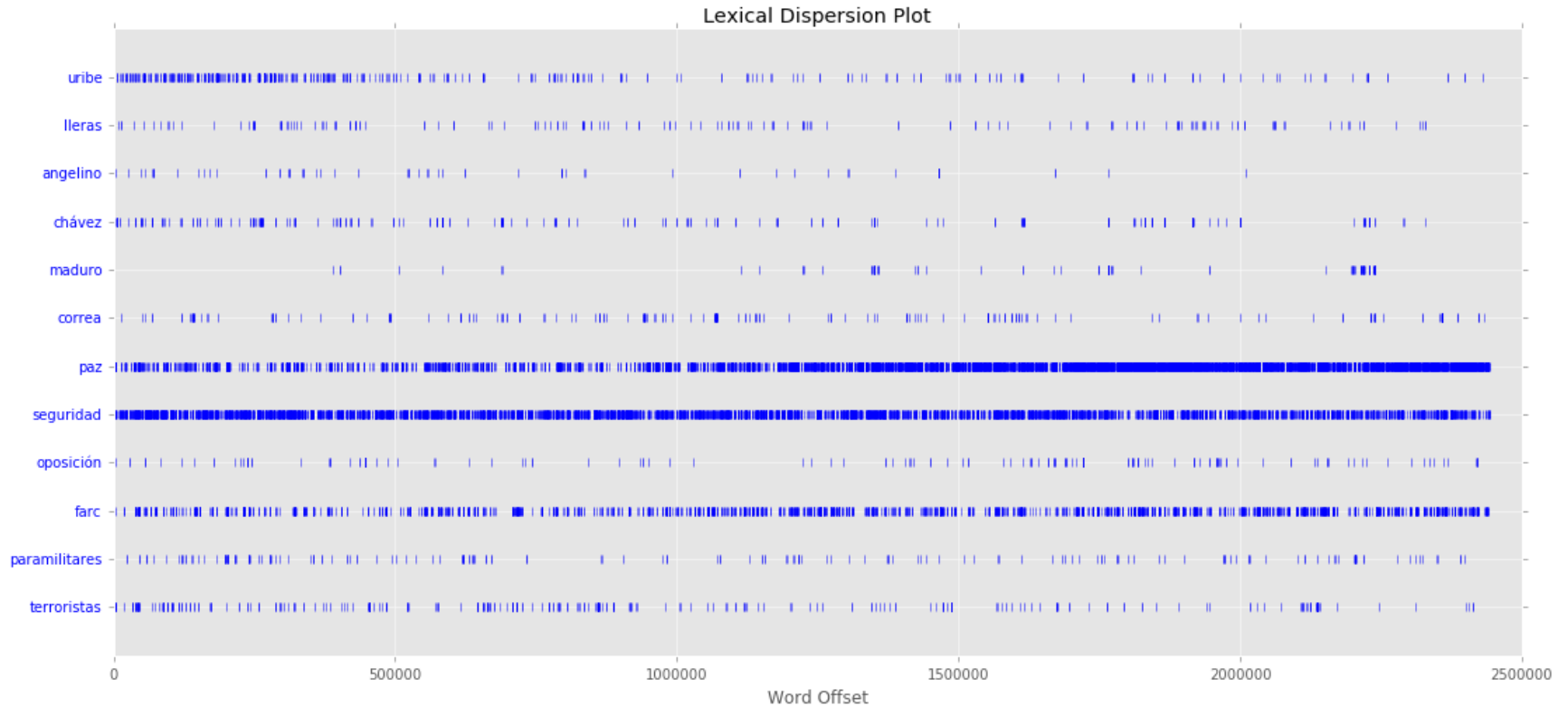
Ejemplos de lo que aprenderemos



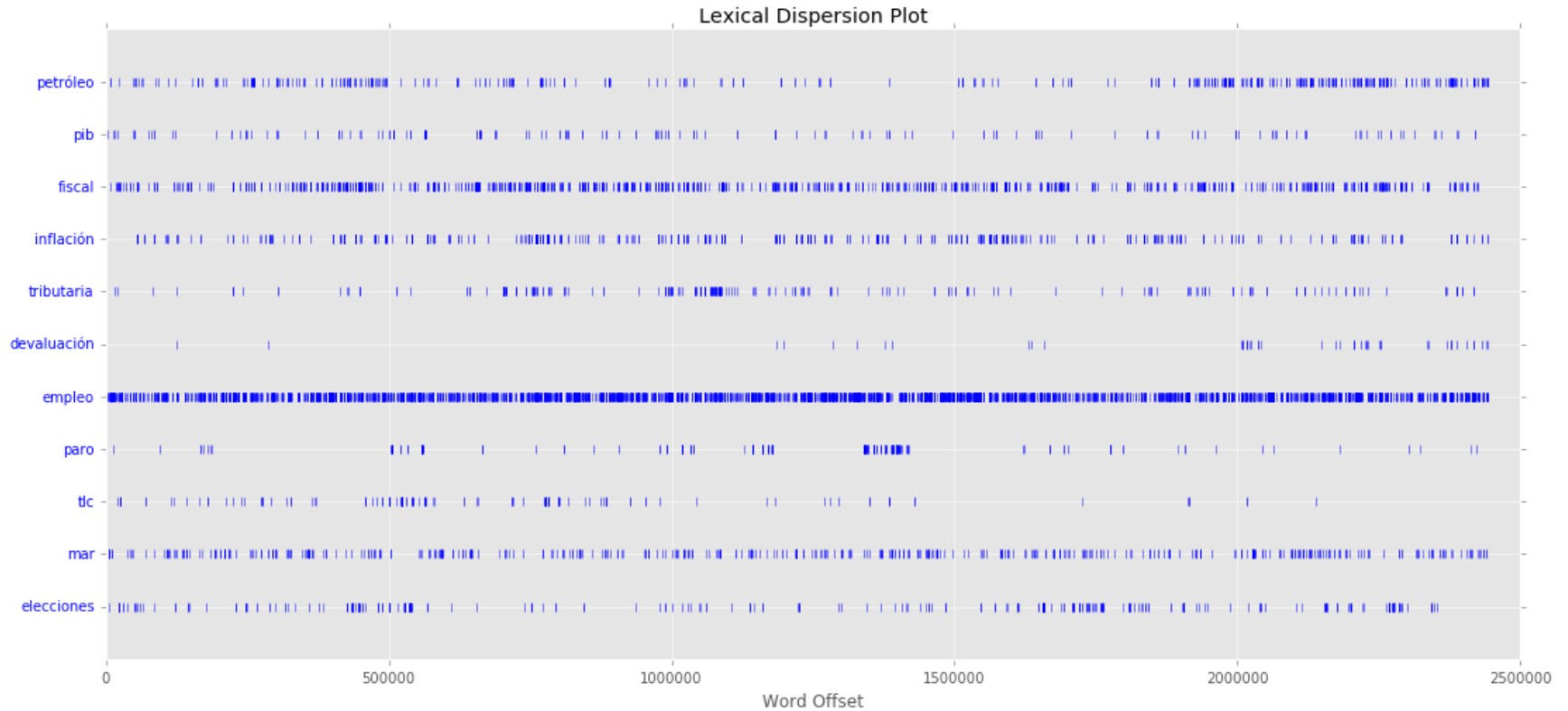
Ejemplos de lo que aprenderemos



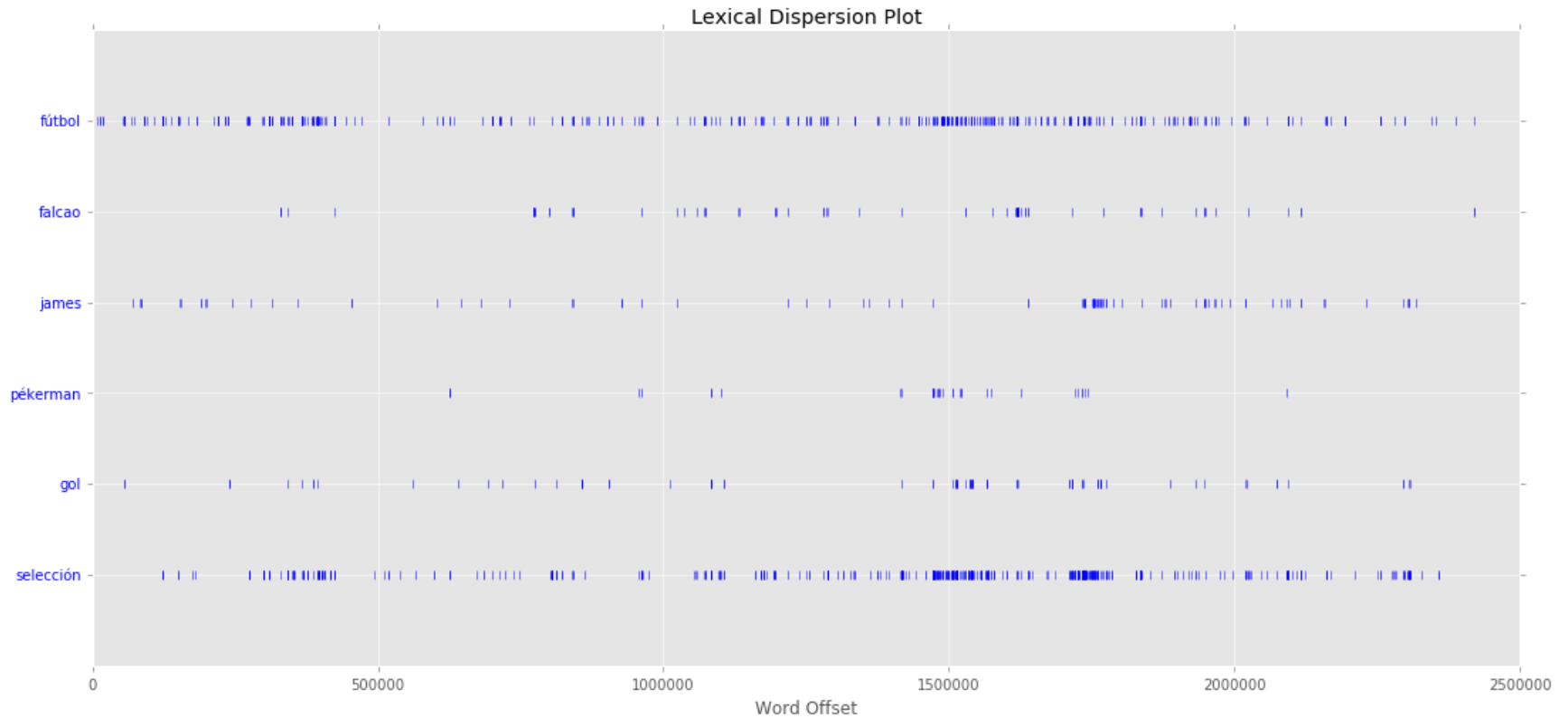
Ejemplos de lo que aprenderemos



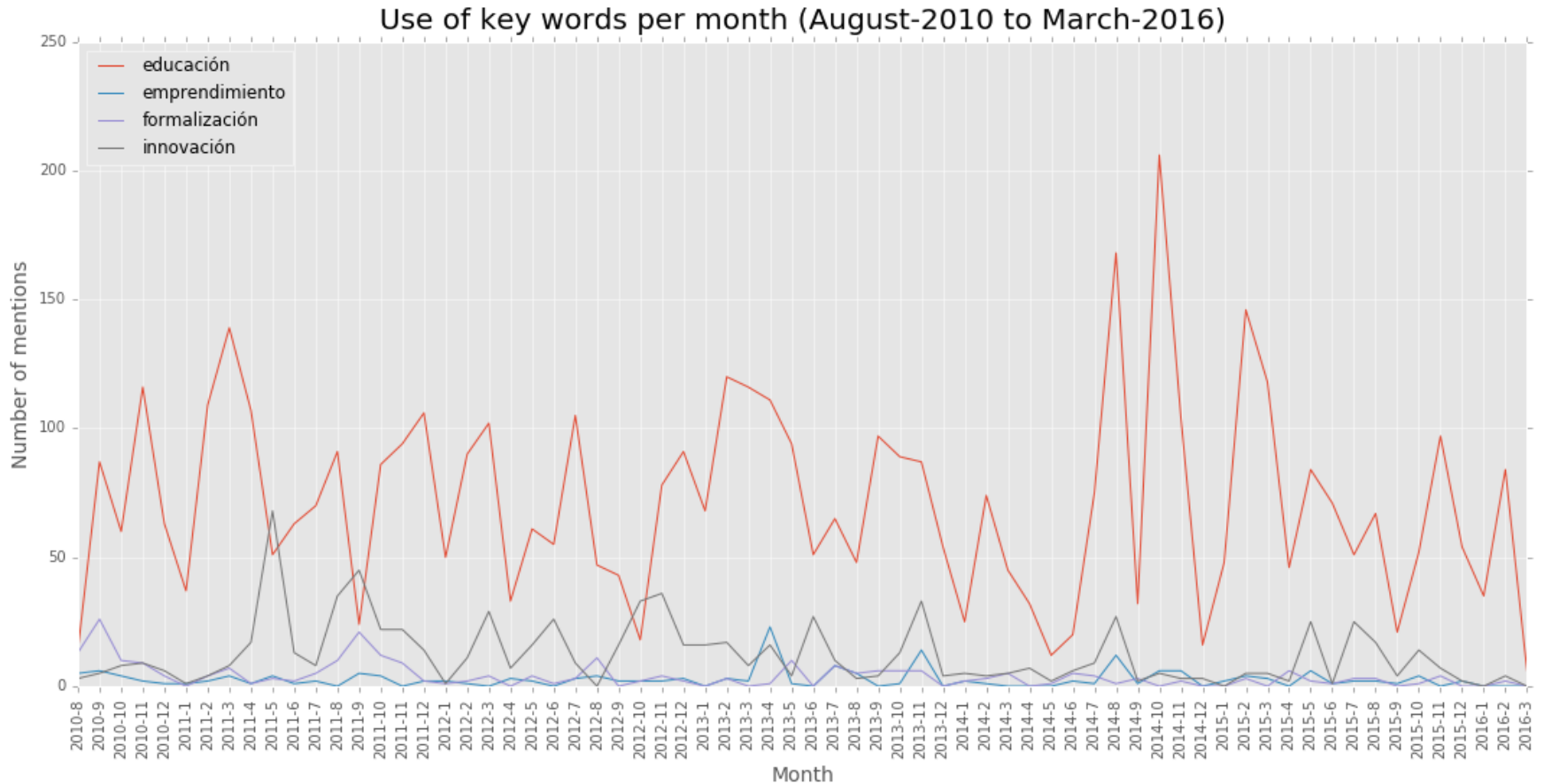
Ejemplos de lo que aprenderemos



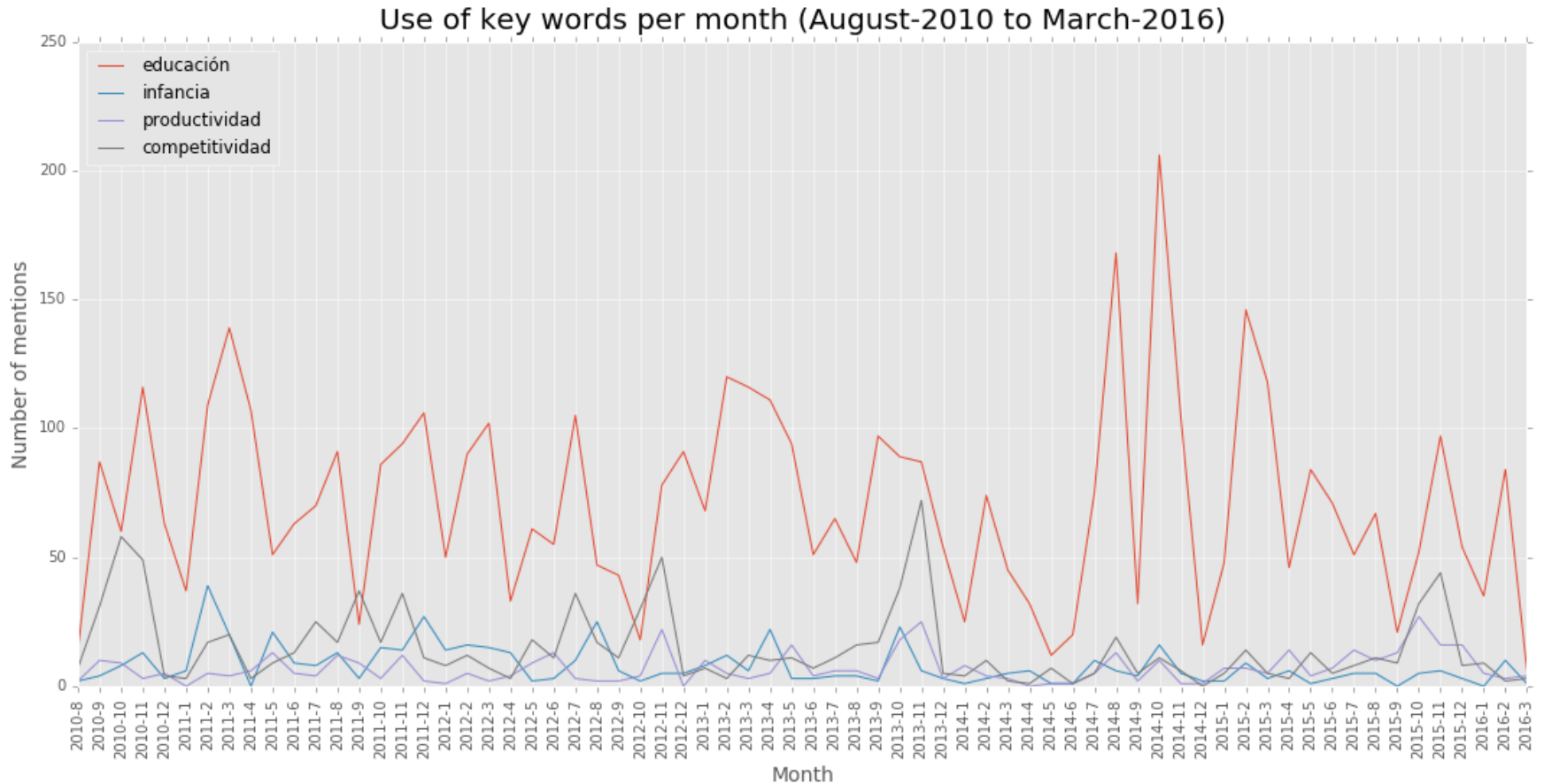
Ejemplos de lo que aprenderemos



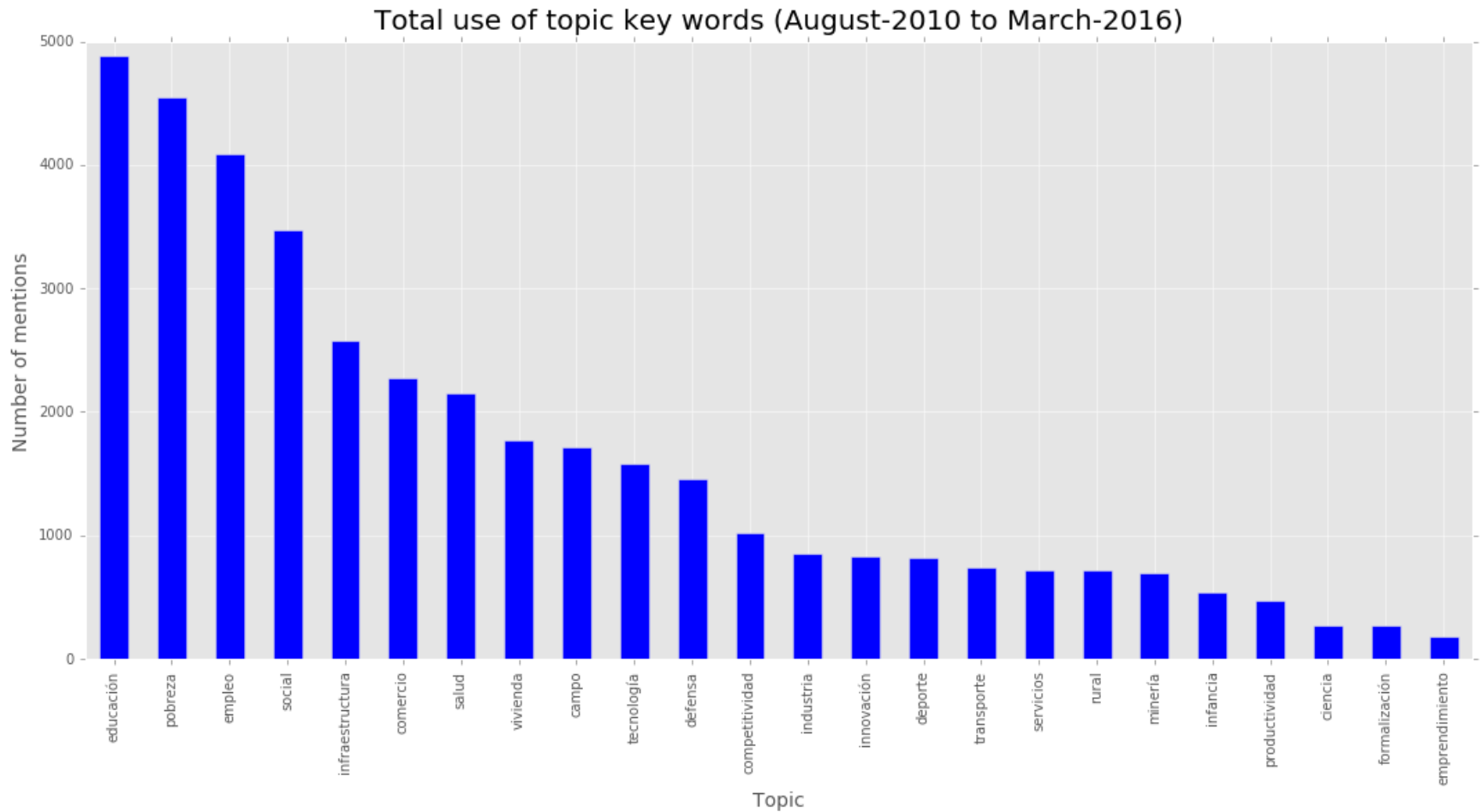
Ejemplos de lo que aprenderemos



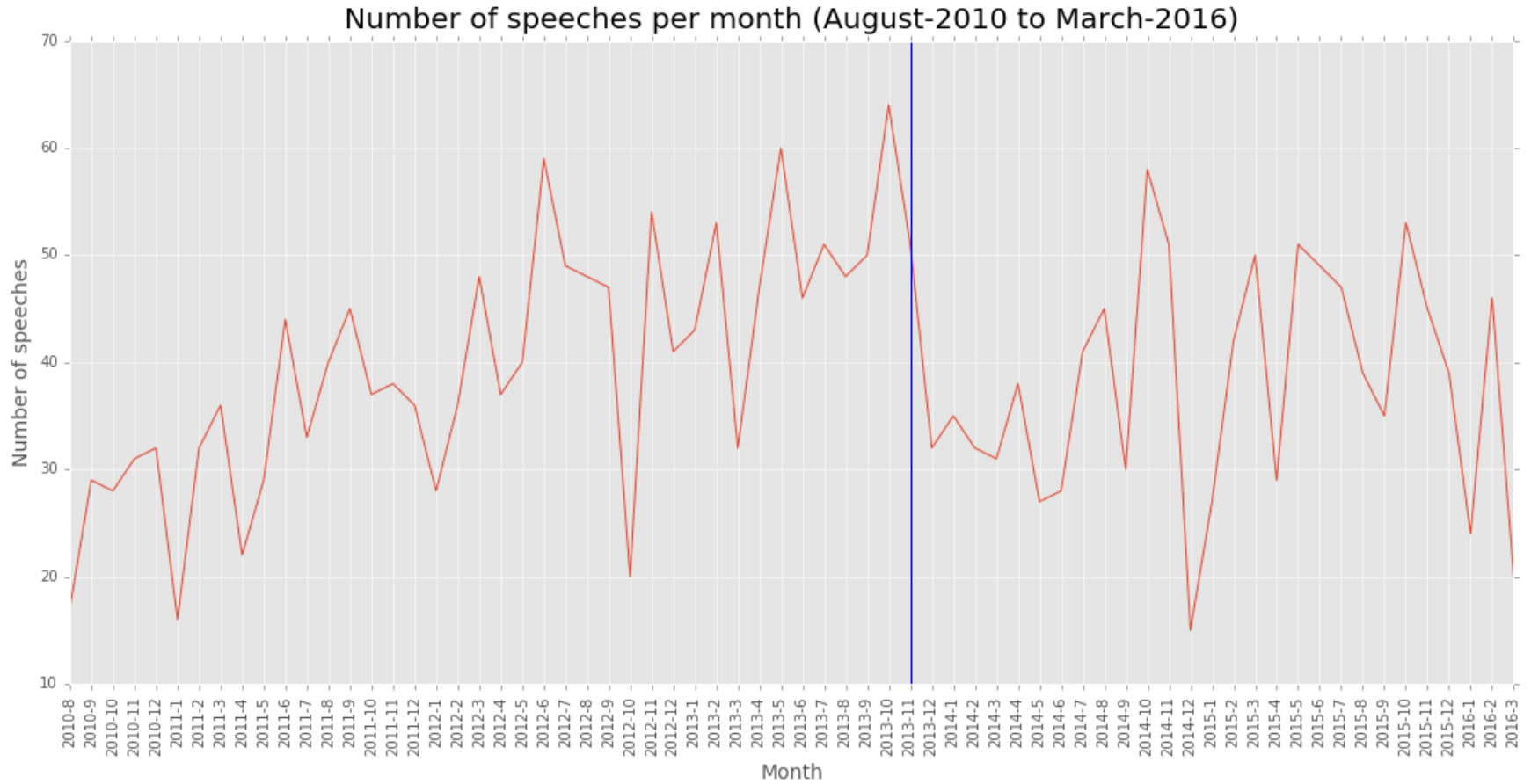
Ejemplos de lo que aprenderemos



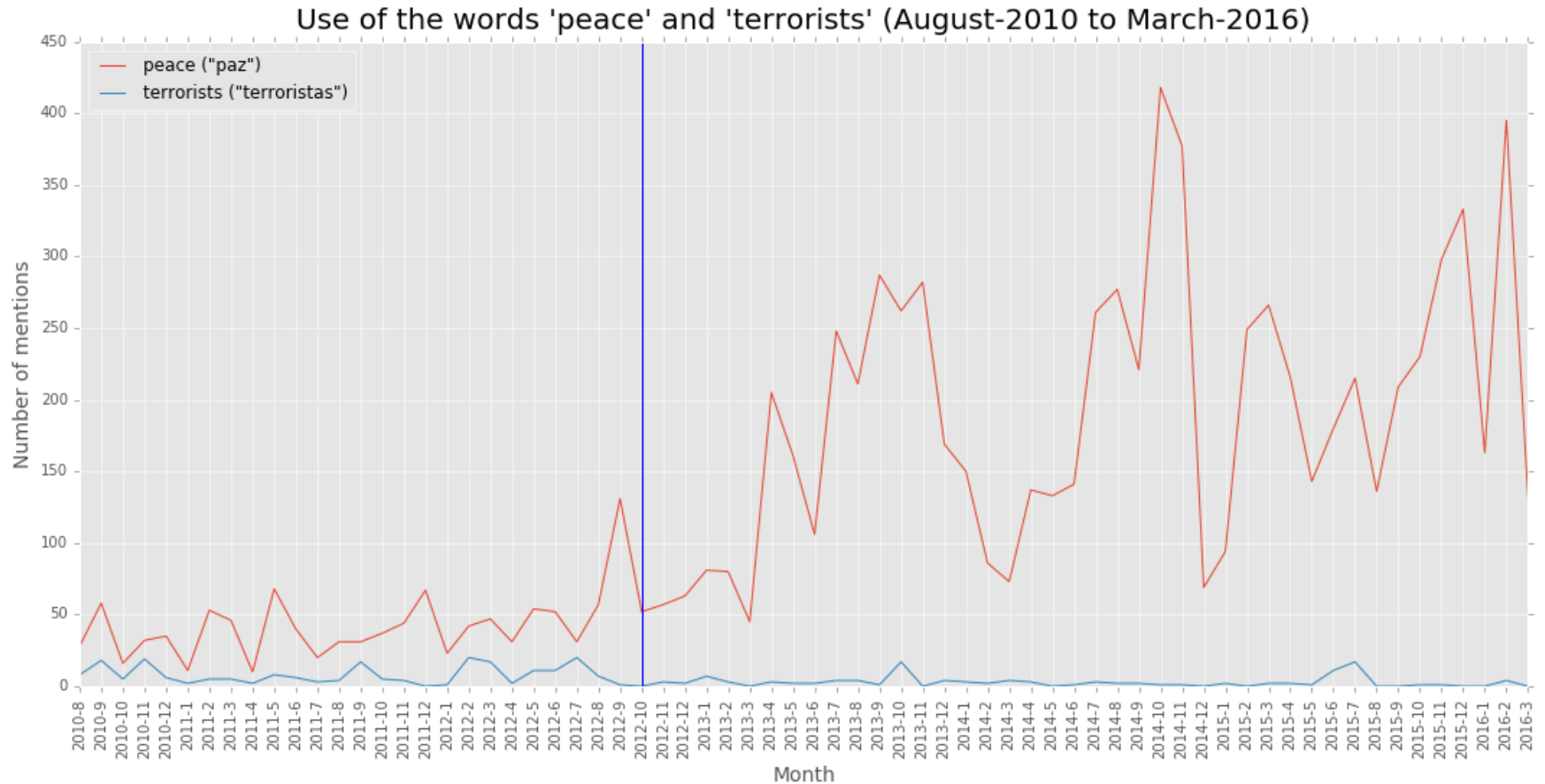
Ejemplos de lo que aprenderemos



Ejemplos de lo que aprenderemos



Ejemplos de lo que aprenderemos



Ejemplos de lo que aprenderemos



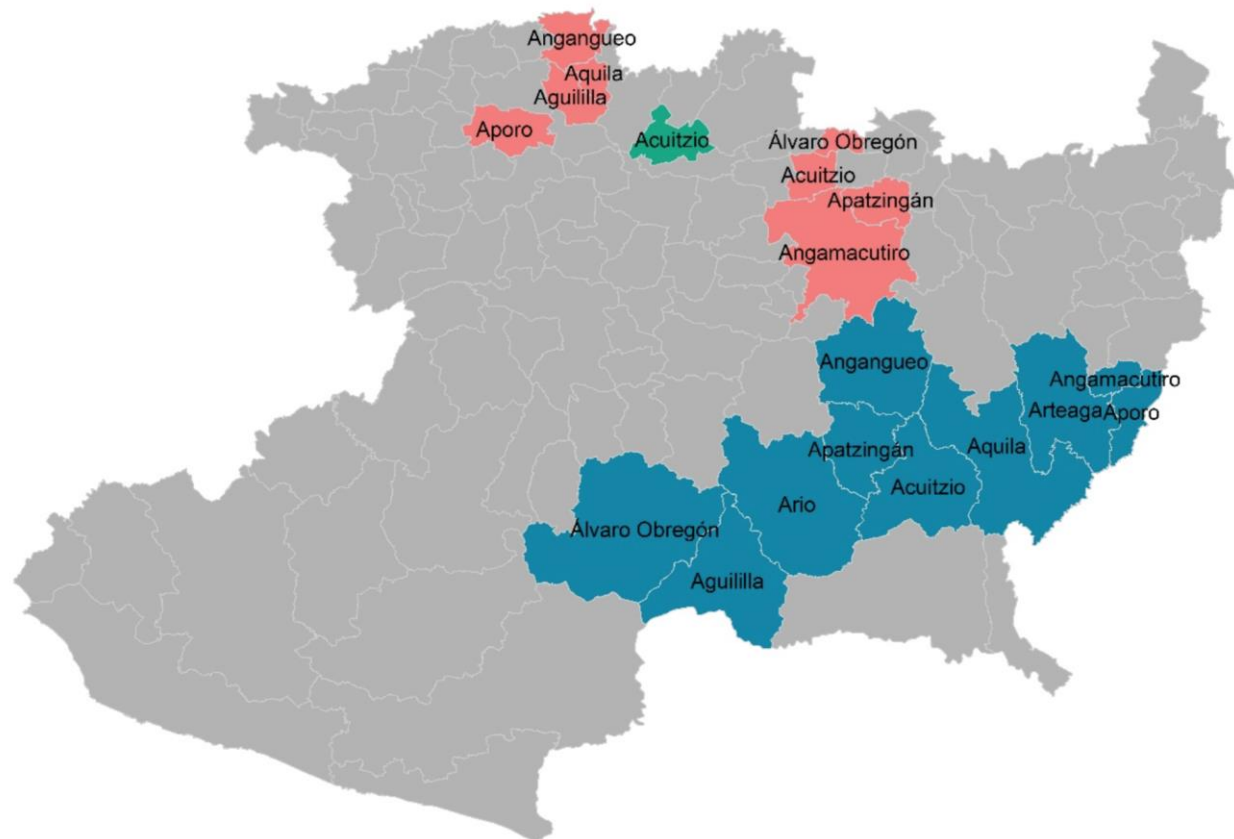
Ejemplos de lo que aprenderemos

Michoacán State in Mexico



Ejemplos de lo que aprenderemos

Poverty clusters in Michoacán



Legend

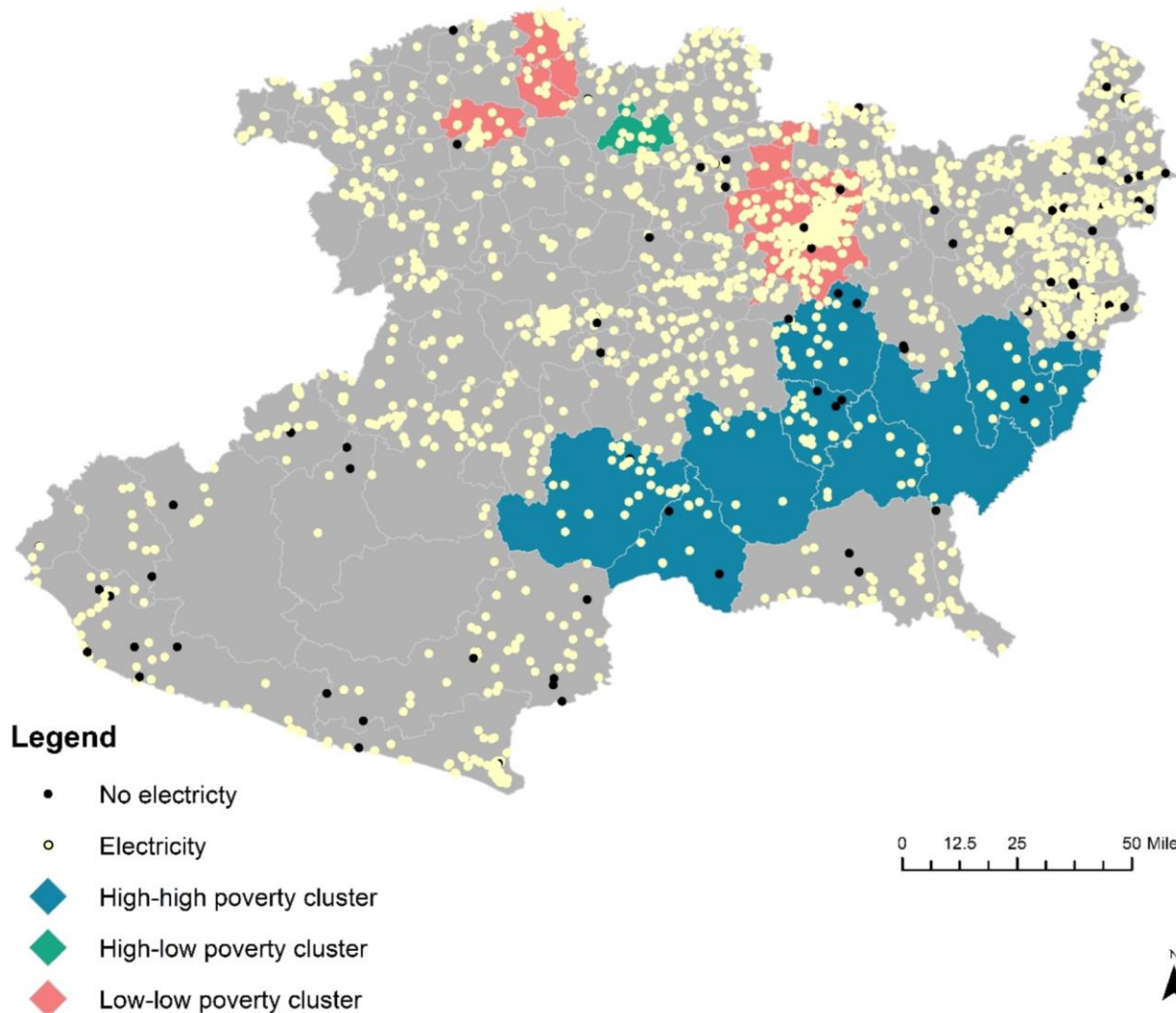
- ◆ High-high poverty cluster
- ◆ High-low poverty cluster
- ◆ Low-low poverty cluster

0 12.5 25 50 Miles



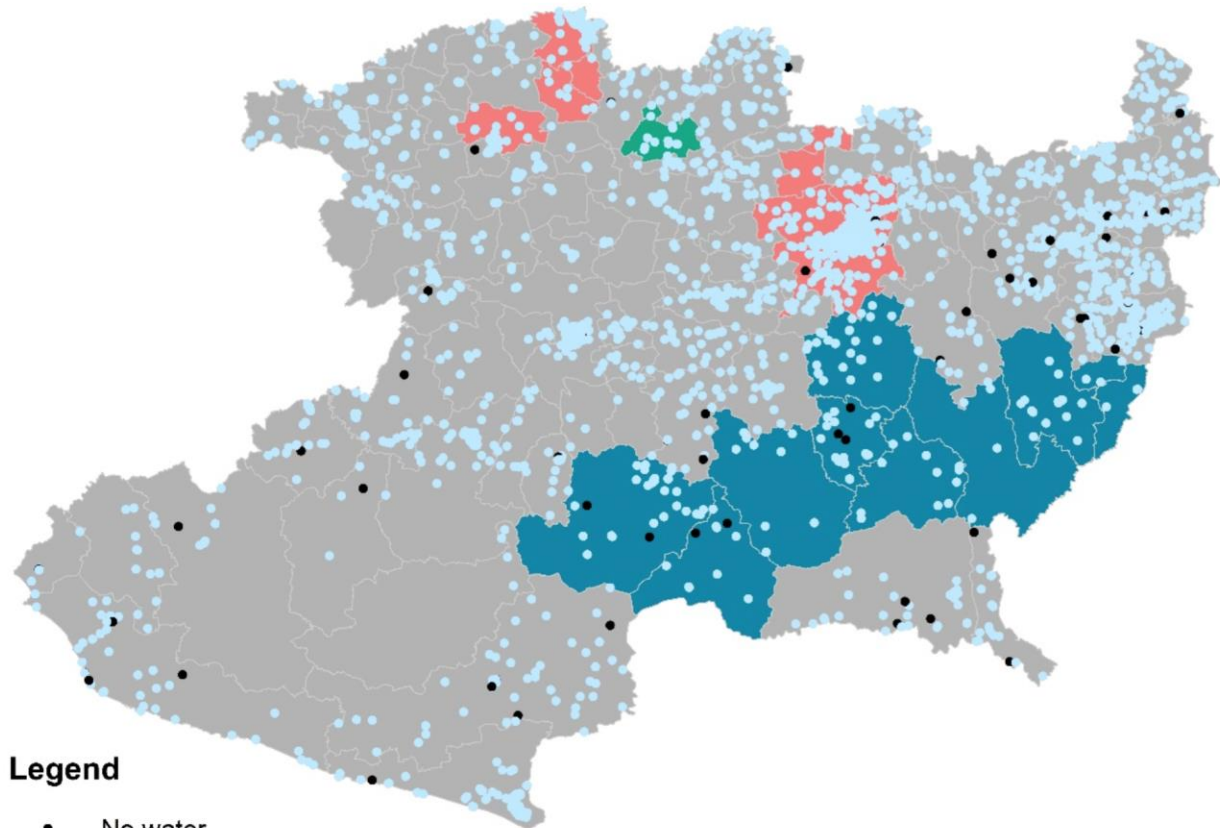
Ejemplos de lo que aprenderemos

Schools with and without Electricity



Ejemplos de lo que aprenderemos

Schools with and without Water



Legend

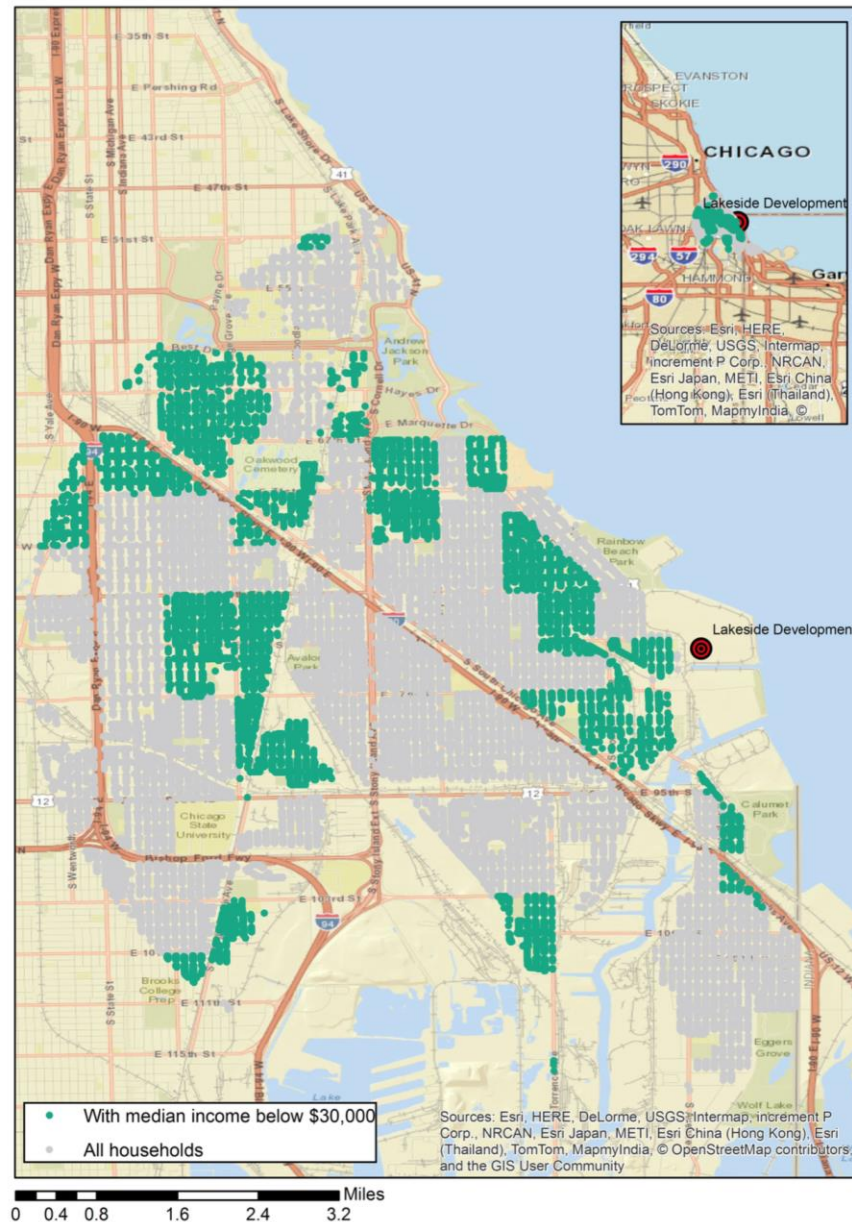
- No water
- Water
- ◆ High-high poverty cluster
- ◆ High-low poverty cluster
- ◆ Low-low poverty cluster

0 12.5 25 50 Miles

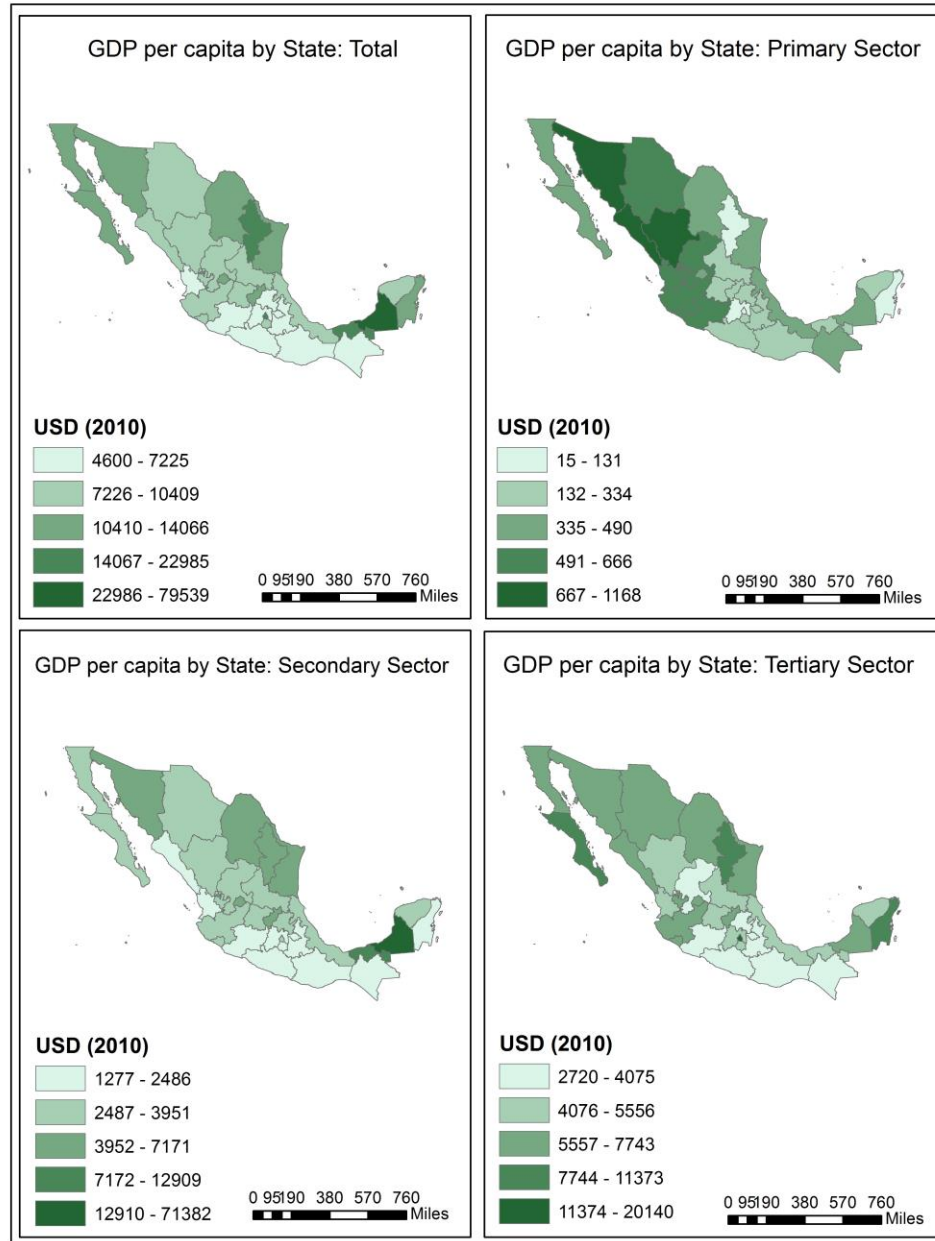


Ejemplos de lo que aprenderemos

Sample of households within 5 miles of Lakeside development

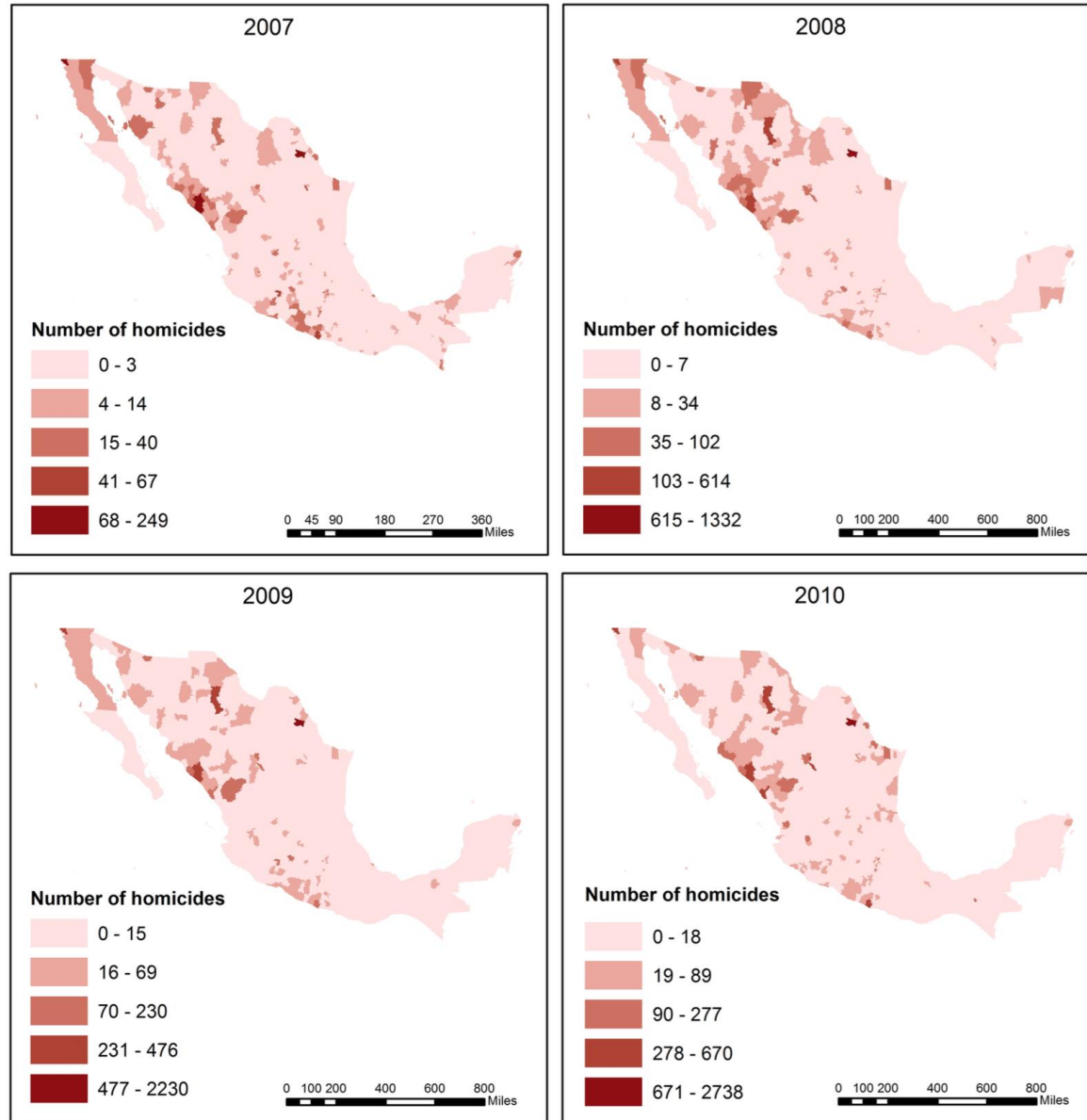


Ejemplos de lo que aprenderemos



Ejemplos de lo que aprenderemos

Number of homicides in Mexico: 2007 – 2010

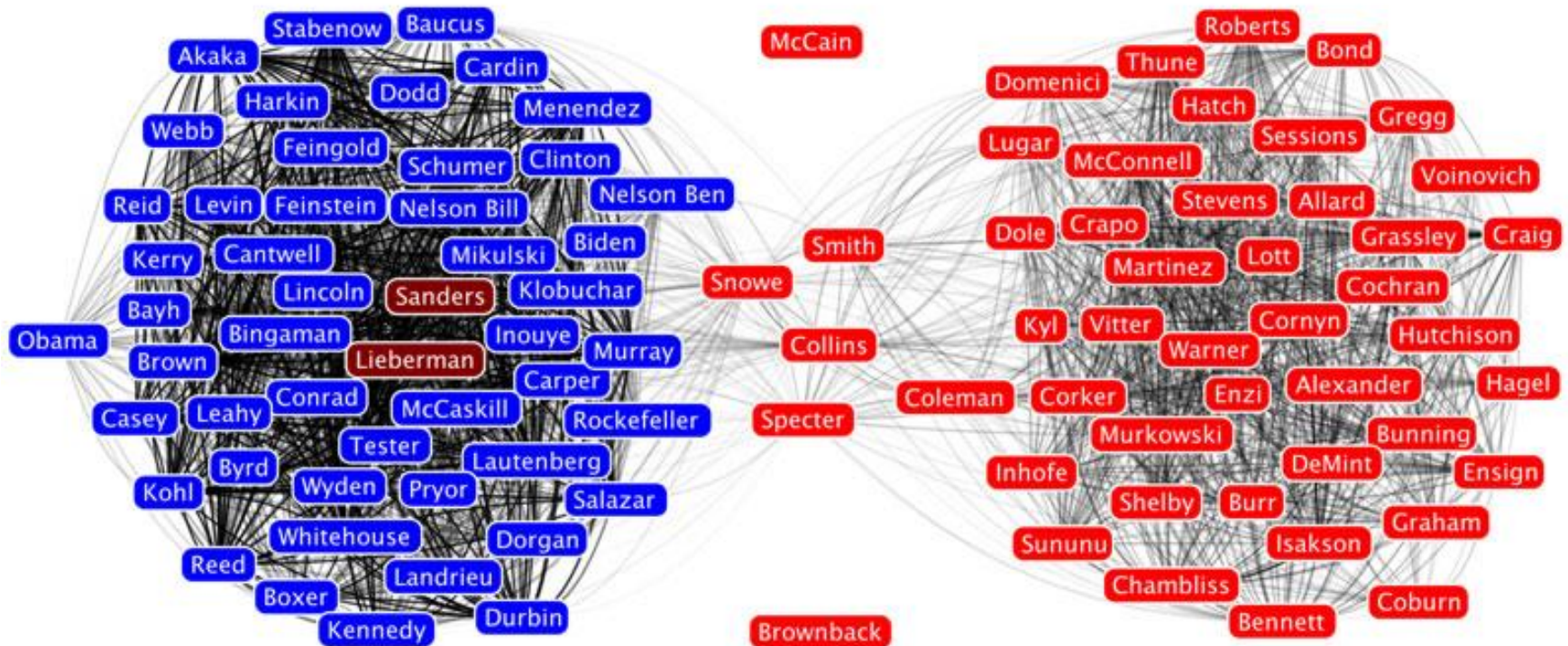


Ejemplos de lo que aprenderemos

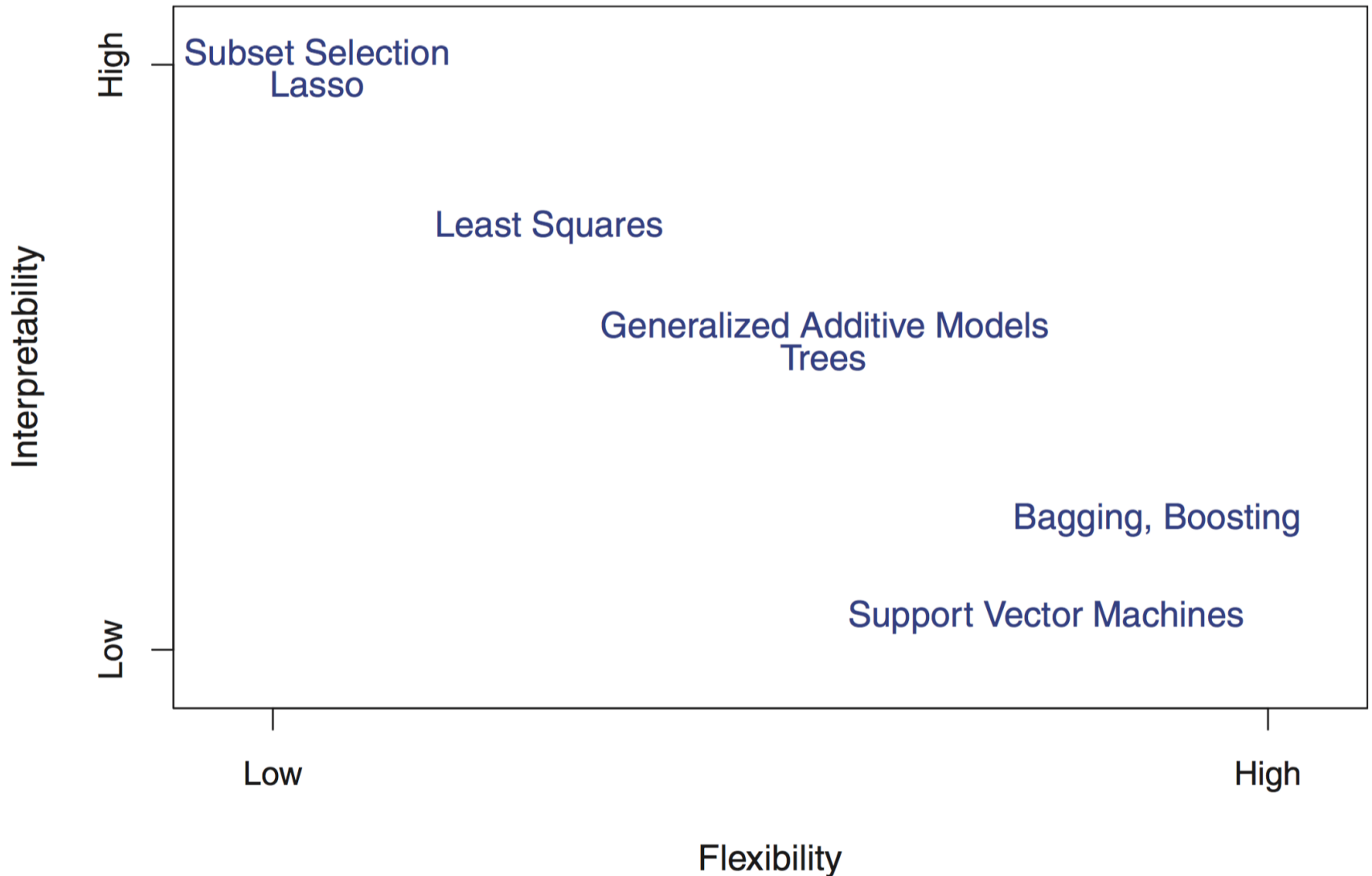
Tableau

<https://www.tableau.com/>

Ejemplos de lo que aprenderemos



Ejemplos de lo que aprenderemos



Ejemplos de lo que aprenderemos

*The **hacker culture** is a subculture of individuals who enjoy the intellectual challenge of creatively overcoming limitations of software systems to achieve novel and clever outcomes. The act of engaging in activities (such as programming or other media) in a spirit of playfulness and exploration is termed 'hacking'. However, **the defining characteristic of a hacker is not the activities performed themselves (e.g. programming), but the manner in which it is done: hacking entails some form of excellence, for example exploring the limits of what is possible, thereby doing something exciting and meaningful.***

General hacker attitudes:

- ***Access** to computers –and anything that might teach you something about the way the world works– **should be unlimited and total.***
- ***All information should be free.***
- ***Hackers should be judged by their hacking, not bogus criteria such as degrees, age, race, or position.***
- ***You can create art and beauty on a computer.***

(Wikipedia)

Hoy (esta semana)

1. Presentación (de ustedes y mía)
2. Encuesta en línea (Quiz 1)
3. ¿En qué nos metimos?
4. Ejemplos de lo que aprenderemos
5. Metodología y evaluación
6. Cronograma
7. Tarea

Metodología y evaluación

El curso será demandante.

Haré lo posible por diseñar tareas “útiles”, pero deberemos “comernos los vegetales”.

Para aprender, deberán hacer. No basta con observar.

Cometeré errores; los corregiré.

El programa del curso es una hoja de ruta. Haremos precisiones y ajustes sobre la marcha.

Valoro la retroalimentación constante –no solo al final del semestre–.

Disclaimer: uso frases y palabras en inglés, no por esnobismo sino por ignorancia/pereza.

Metodología y evaluación

“Aprender haciendo”:

- 8:00 ~ 9:30: Clase “magistral”
- 9:45 ~ 11:00: Clase + taller práctico / tarea
- 11:00 — 12:00: Monitoría: Daniel Rojas
(danielm.rojas@urosario.edu.co)

Calificaciones:

- Comentario semanal: 5 %
- Preguntas en clase (“quices”): 20 %
- Talleres: 25 %
- Examen parcial: 20 %
- Proyecto final: 30 %

Hoy (esta semana)

1. Presentación (de ustedes y mía)
2. Encuesta en línea (Quiz 1)
3. ¿En qué nos metimos?
4. Ejemplos de lo que aprenderemos
5. Metodología y evaluación
- 6. Cronograma**
7. Tarea

Cronograma

1. Presentación, introducción al curso
2. Introducción a computer science, Python, iPython, Jupyter Notebook, programación básica
3. Condicionales, *loops*
4. Listas, tuplas, *strings*
5. Funciones
6. Diccionarios, *functional programming*, funciones anónimas, *map()*, *filter()*, *reduce()*, *list comprehensions*
7. Trabajar con archivos de texto
8. Parcial
9. Numpy, matplotlib
10. Pandas
11. Expresiones regulares, *web scraping*
12. Más *web scraping*, análisis de texto
13. R, redes
14. QGIS, Tableau, Gephi
15. Big data, ciencia de datos
16. Presentaciones proyecto final

Tarea

GitHub

1. Cree una cuenta gratuita en <https://github.com/>.
2. Cree un repositorio y llámelo “MCPP_<primer nombre.primer apellido>” (e.g. mi repositorio se llamaría “MCPP_santiago.matallana”).
3. Añádanos a Daniel y a mí como “colaboradores” en su repositorio. Nuestros usuarios de GitHub son: “smatallana” y “rojasdaniel”.
4. Instale Github desktop en su computador y vincúlelo a su cuenta de GitHub: <https://desktop.github.com/>
5. Lea la guía de GitHub:
<https://guides.github.com/activities/hello-world/>

Tarea

Python

1. Instale Anaconda, con Python 3.7:
<https://www.anaconda.com/download/>
2. Lea:
 - Computational Thinking, Jeanette Wing
(<http://bit.ly/2jwFEPY>)
 - Zelle: 1.1-1.9, 2.2-2.5, 3.1, 8.4
(en el repositorio del curso)

Comentario semanal

- Envíe el comentario semanal: http://bit.ly/mcpp192_r1

¡Importante!

El repositorio del curso, donde podrán encontrar toda la información del mismo, es:

https://github.com/smatallana/mcpp_2019_2

(Si no funciona el link, cópielo y péguelo en la barra de direcciones)

El repositorio es privado.

Dele a Daniel su nombre de usuario de Github, para que podamos darle acceso al repositorio.