

nlp_tp_pe

December 5, 2019

1 Ejercicio Procesamiento del Lenguaje Natural

1.1 Analizaremos los titulares del diario Perfil

2 Que vamos a hacer:

1 - Cargar los datos

- 2 - Limpiar datos
- 3 - Analisis Exploratorio
- 4 - Anáisis de Sentimiento

```
[31]: %%time
      # importamos librerias utiles
      import requests
      import pickle
      from time import sleep
```

CPU times: user 285 ms, sys: 7.86 ms, total: 293 ms

Wall time: 291 ms

2.1 Perfil

3 1 - Cargar los Datos

```
[32]: %%time
      # lo metemos en un Panda dataframe
      import pandas as pd
      pd.set_option('max_colwidth',150)
      import pickle
      # Cargamos el csv de los titulares de La Nacion en un dataframe

      data_pe = pd.read_csv('pe_por_year_2.csv',index_col=0 )
      data_pe.index.name=None
      data_pe
```

CPU times: user 28.7 ms, sys: 4.13 ms, total: 32.9 ms
Wall time: 110 ms

[32]:

Texto

2006 córdoba alianza delasotista se impone en marcos juarez el reclutamiento docente cae en picada relatos de una obsesión ya no podía entrar al aul...
2007 la historia se repite más de cien heridos con quemaduras de pirotecnia y lesiones en los ojos el primer bebé del año nació en buenos aires quemar...
2008 atentados y operativos de seguridad oscurecen festejos el primer bebé porteño del año nació en parque patricios un motociclista perdió la vida al ...
2009 balas perdidas corchos riñas callejeras e incidentes con pirotecnia en los festejos de fin de año cromañon en bangkok muertos en una discoteca...
2010 peajes bien chévere los placeres mínimos de cada día para estar saludables mateo ramos se postula para ser el primer bebé de voraz incendio en be...
2011 toman otro predio en villa lugano durante los festejos de año nuevo hoy asume dilma rousseff en brasil en santiago del estero el arrancó con un s...
2012 murió el gobernador de río negro quién asume el gobierno en río negro primeros detalles y repercusiones de la confusa muerte de soria carlos soria...
2013 pablo echarri la militancia me llegó tarde a los infografía feriados encuentran muerto al andinista perdido en el aconcagua boudou pasó el...
2014 no creo que haya amor creo que hay propiedad privada y no me la toca nadie la verdad mexicana no es un lujo es algo que necesito la conferenci...
2015 caso lola aseguran que es extremadamente complicado choque y vuelco fatal en camino del buen ayre boudou el embajador en la reasunción de di...
2016 la búsqueda de los tres prófugos se concentra en zona sur el papa francisco ofició la primera misa del año y alertó sobre la opresión isis planeab...

[69]:

```
%%time  
# Veamos uno de los contenidos  
data_pe.Texto.loc[2008]
```

CPU times: user 622 µs, sys: 14 µs, total: 636 µs
Wall time: 650 µs

[69]:

'atentados y operativos de seguridad oscurecen festejos el primer bebé porteño del año nació en parque patricios un motociclista perdió la vida al chocar con un patrullero fue a comprar hielo para la cena y murió electrocutado córdoba siete adultos y cinco niños con heridas leves un joven mató a sus ex suegros mientras festejaban asambleístas de gualaguaychú cenaron en familia dos hombres asesinados a cuchillazos en córdoba el año empieza con más calor y probabilidad de lluvias martin palermo \x93nunca soñé con ganar tantas cosas\x94 benedicto xvi volvió a defender a la familia natural los medios internacionales se hacen eco del fracaso al menos personas quemadas vivas en una iglesia nuevas pistas en el caso maddie y apuntan a marruecos lindsay lohan fue elegida la peor actriz del año quedaron heridos luego de chocar y les robaron ni los funerales se salvan de los atentados en irak boca define el pase de banega al valencia de

españa las ventas minoristas aumentaron en diciembre captores de la enfermera
 huyeron a zona montañosa tres noches con insomnio aumentan riesgo de diabetes en
 españa aseguran que saviola se va del madrid harry potter una historia de
 nunca acabar salen a controlar el trabajo en negro en la costa atlántica
 murió al caer por el hueco de un ascensor del hospital el hijo de benazir bhutto
 tiene club de fans en internet tomarán prueba de adn a la abuela de emmanuel
 Chávez acusó a uribe de dinamitar la operación humanitaria tevez brilló en una
 nueva victoria del manchester entra en vigencia el aumento de los aportes a las
 afjp caso lópez arslanián le apuntó a fuerzas represoras la argentina alcanzó
 un récord de trasplantes en año nuevo trágico cinco muertos en accidentes
 viales regreso sin gloria kirchner se reencontró con cristina en el calafate
 sanz el menos querido por la prensa hispana de ee uu nueva moneda chavista
 para frenar la inflación el gobierno volvió a pedir la extradición de antonini
 aysa dice haber solucionado la falta de agua pero llueven las quejas la uom no
 se pone a la sombra de hugo moyano año nuevo francés sólo quemaron vehículos
 hallaron el cadáver de una joven en dique los molinos un nene de años muerto
 por bala perdida casi dos centenares de heridos por pirotecnia y corchos en
 rosario atacan a guardavidas por joven ahogado chipre y malta ingresaron a la
 unión europea de la brasa al horno anticipan máxima de grados matan a joven a
 cuchilladas en festejo de año nuevo publican libro con cartas de betancourt a su
 familia punta del este empezó la temporada pero de robos dicen que fueron
 contratados por telerman cuando ya había perdido alerta por volcán chileno la
 lava avanza hacia argentina censura por dos hombres desnudos en una cama nicole
 kidman anunció a su familia que está embarazada escándalo sexual involucra a un
 futbolista de arsenal angelina jolie fue elegida la más humanitaria de choque
 e incendio en escobar un muerto boqueteros aprovecharon festejos para robar
 corralón la esposa lo acusa de adúltero en su propio programa pakistán aumenta
 la violencia tras aplazo de elecciones el alcohol hace más daño a mujeres que a
 hombres harry potter deja la magia y protagoniza una historia real la justicia
 investiga a miceli por su custodia extra falleció willy sommerfeld el
 pianista del cine mudo conflicto judicial por una causa contra anarquistas entre
 los refuerzos tal vez haya que buscar un presidente del club con un decreto
 ascendió a secretario de estado al médico familiar wanda nara chocó en la x
 que le regaló maxi lópez paul mccartney fue operado de urgencia del corazón
 repudian nombramiento de funcionario de la dictadura pilar le dijo a taiana que
 está \x93cansada pero bien de ánimo\x94 montenegro quiero tener revancha ya
 mismo pidieron que bajen el volumen y terminaron a las piñas tenis argentina
 no pudo recuperarse ante taiwán caso acro rechazan exención de prisión a un
 barra nueva pista fortalece la teoría de uribe sobre emmanuel médicos sin
 fronteras confirma la liberación y expresa su \x93alivio\x94 para colombia las
 farc \x93engañaron\x94 a hugo Chávez macri ratificó que seguirá adelante con la
 revisión de contratos guerra de municipales contra macri no somos ñoquis
 cristina decidió tomarse vacaciones en el calafate el viaje de kirchner debió
 discutirse en el congreso castromán un nombre notable para el boca de ischia
 una que se va y otra que se queda febres postergaron la resolución procesal de
 familiares randazzo la oposición tiene el vuelo de las gallinas duras
 críticas de la oposición por el fracaso en colombia calleri puede volvió a

ganar y está en cuartos en doha regresa pilar bauzá la enfermera liberada en somalia histórico el petróleo alcanzó los u s por primera vez eeuu la onu no debe investigar asesinato de bhutto nube de cenizas en neuquén por volcán chileno manuel contepomi duro con los dirigentes del rugby del neandertal al hombre gracias a un cambio climático con una nueva saga buscan reemplazar a harry potter por un embargo peligra la pretemporada de racing responderemos en la justicia prometió el sutecha el agobiante calor puso en alerta el sistema eléctrico los docentes de amet chochos con la intervención copa libertadores sepa cuándo debutará su equipo confirmaron que la recaudación de fue récord por qué abandonar el modelo productivo para pasar a un verdadero modelo de prosperidad terminamos siendo empleados de Chávez dijo gerardo morales industriales comenzaron con el plan de uso racional de la energía despedidos del casino no pudieron llegar al hipódromo la española ence empieza su pastera en colonia internet gratis a cielo abierto en rio de janeiro rosario bajo ataque de palometas feroces cayó la demanda laboral en diciembre scioli destina millones para refaccionar escuelas este river ya es cosa de locos llega sebastián abreu emmanuel los análisis de adn estarían en horas la lluvia trajo en la noche un desesperado alivio dos heridos de bala en confuso episodio en constitución empieza la carrera hacia la casa blanca en iowa los argentinos creen que aumentará la inflación al menos desplazados por la violencia en kenia tras la intervención los municipales retoman las medidas de fuerza aseguran que el precio del petróleo seguirá subiendo récord en la madrugada heridos en cuatro choques intentaron robar a tres cuadas de la casa rosada china quiere generalizar el uso de la inyección letal sean penn presidirá el jurado del festival de cannes penélope cruz y sus hermanos en un video bien caliente la suba del abl también afectará a los inquilinos el abogado de la familia febres confía en la liberación paseadores indemnizarán a dueña por muerte de perra la gente tomó menos en las fiestas que el resto del año cómo aprender a cocinar jugando a los videos santa fe binner promulgó la ley de talles sigue el éxodo en boca también se va krupoviesia revelan que dos de las spices sufrieron bulimia miami hace una ficción de la crisis hipotecaria de ee uu otra nena resultó herida con una bala perdida exhibirán una versión restaurada de \x93un perro andaluz\x94 denuncian que paglieri llega al indec con una patota y tiene custodia comenzó la era del cholo simeone en river las fotos del cholo simeone el beckham argentino vence plazo para que la justicia defina el caso febres hasta lópez murphy critica la política tributaria de macri dieron de alta a la modelo mariana de melo abogado de municipales asegura que no pueden intervenir el gremio duras críticas al filme sobre el asesino de lennon alerta en chile por posibles nuevos aluviones volcánicos uruguay pide embargo de red internacional de narcos malvinas reclamo argentino a años de la ocupación la cgt respaldó el plan de acción de los municipales calleri y acasuso eliminados en doha y adelaída confirman el procesamiento de madonna quiroz de vido el cambio de huso horario está funcionando perfectamente desplazan a otro empleado incómodo del indec kaká donó a su iglesia el trofeo a mejor jugador de policía mató a joven mapuche en protesta por tierras el kirchnerista santamaría criticó la intervención de macri a obsba Perú destituyeron a profesores por abuso sexual radicales contra kirchner por su presencia en

colombia las farc prometen más sangre y anuncian ofensiva general patovica dejó internado a un turista en mar del plata boca va por un ex river el paraguay julio cáceres malformaciones genéticas irían de la mano con el cáncer botnia diario uruguayo publicó informe viejo de la uba un bielsa auténtico en chile excluyó a tres figuras por teléfonos y electrodomésticos las mayores quejas pese al conflicto gremial en regresa la ficción scioli suspendió por días el impuesto a la riqueza verón dolido con simeone hay que ir con la verdad tras los cacerolazos aseguran que volvió el servicio la defensa de astiz apela a tratados sobre dd hh el petróleo cruzó otra vez los míticos dólares daia invitó a maradona al museo del holocausto ¿los cometas ayudaron a la formación de la atmósfera bonafini estamos con las farc con kirchner y Chávez d elía trató de \x93fascistas\x94 a dirigentes judíos que criticaron a maradona ahora las chicas podrán disfrutar del primer slip open tabaré ordenó mantener abierto el paso fonterizo de fray bentos ¿de que se ufanan en la casa rosada un concurso del central para descubrir al nuevo lousteau el garante brasileño trató de darle una mano a kirchner deuda externa privada supera los us millones cacerolazos por el combo verano falta de luz y agua ya podemos preocuparnos por el invierno bolivia no cubrirá demanda de gas por tercera vez abogados abandonan a britney spears durante los accidentes dejaron muertos por día el seleccionado de voleibol venció bien a uruguay caos por un paro y bloqueo de la terminal en san luis los servicios porteños con un cronograma de emergencia obama y huckabee ganaron las primarias en iowa suspenden días plazos de escrituración de inmuebles tres muertos en un choque frontal con un trailer cayó un narco acusado de enviar cocaína a españa profunda reestructuración en el gabinete de Chávez fue a probarse un vestido de fiesta y se le cayó el piso empleados del casino cumplen días encadenados intoxicada britney fue hospitalizada en los angeles muere periodista tucumano calcinado en un accidente un muerto al chocar con un camión en san nicolás servicios afectados por el paro de sutecba cinco adolescentes atropellados por un movil policial sin ginóbili los spurs cayeron ante denver nuggets ever banega al valencia \x93me pasó todo muy rápido\x94 cómo enfrentará el gobierno la crisis energética kenia manifestación opositora exige nuevas elecciones prorrogan la emergencia penitenciaria bonaerense apoyo de camioneros decretarían paro y no habría recolección de residuos el príncipe william quiere convertirse en piloto secuestran autos cuyos conductores estaban ebrios nuevo juicio oral para el represor jorge el tigre acosta remarkan la responsabilidad institucional de la prefectura por estrictas explota la demanda de niñeras británicas moyano acusó a macri de llevar a cabo un \x93genocidio laboral\x94 denuncian aumentos ilegales del boleto de colectivo depp la estrella más taquillera de hollywood del el gobierno porteño analiza una respuesta oficial kravetz \x93macri tiene de rehenes a los vecinos de la ciudad\x94 multan a metrovías por la lesión de un niño en el subte guardia se suicida tras espectacular fuga de dos presos dicen que el estará entre los años más calurosos obsba impiden el ingreso del interventor designado por macri roban un auto se tirotean con la policía y lo abandonan schlenker busca testigos de elite para que declaren murió otro represor de la dictadura tres en menos de un mes febres liberaron a familiares y acusaron a prefectos pedido para que la justicia trabee el pase de banega dos

años de suspensión para martina hingis por doping assembleístas de colón vuelven a cortar el puente reencuentro emotivo de pilar bauzá con su familia moyano garantizó la recolección de basura en la ciudad por despidos macri sufrió su primera huelga general de municipales cada uno de los chinos y cada uno de nosotros cancelan el lisboa dakar por amenazas de al Qaeda latinoamérica y sus conflictos servicios quejas por aumentos son el tema del verano la rioja escalofriante un aborto cada dos nacimientos suazo no será de independiente se queda en México el paco crea \x93discapacitados intelectuales monstruosos\x94 botnia paró por horas pero los olores no pingüino de la triste figura la fiscalía general señaló que el niño de bogotá es emmanuel alex ferguson y rooney se sumaron a la tevezmania se viene el live earth latino con roger waters nuevos paros de sutecha para martes miércoles y jueves insólita lluvia de iguanas en miami por ola de frío otra discográfica accede a vender música por internet argentina quinta en el top de ligas de fútbol del mundo la echaron de la pileta porque no quiso tirarse en malla macri convocó a discutir la intervención de obsba fuerte caída en los principales mercados internacionales por una grave lesión insúa estará parado cinco meses no es mi sueño el peor fútbol del mundo ii el cumple de juancito las niñas del maíz ¿qué plan capítulo incluso blindaje y exposición país de plomo el caso febres y el fantasma de röhm el impuesto a la distancia y la falta de integración regalos de reyes en el condominio de materia gris no es mi sueño estrellas de hollywood boicotarán los globos de oro analistas hablan de la \x93pérdida de credibilidad\x94 de Chávez y las farc boca denuncias cruzadas en la interna de los dirigentes mariana de melo cree que podrá debutar en carlos paz moyano aseguró que despedidos \x93no son ñoquis\x94 y quedarán sin obra social punta bolocco se prepara para desfilas con giordano el fue un año récord en el secuestro de droga vislumbran la \x93perdición\x94 con un presidente creacionista el primer arrepentido del gobierno de cristina k el cambio de hora suma riesgos a la exposición al sol kenia crisis humanitaria y lento retorno a la calma los abuelos ya se animan a manejar nuevas tecnologías banega llegó al valencia y prometió \x93entrega y sacrificio\x94 el rey juan carlos festeja sus años en la intimidad cinco cruceros arriban a buenos aires el fin de semana venezuela ajusta su economía con el bolívar fuerte denuncian rechazo de donativo de un \x93matrimonio mixto\x94 derrota política las farc también admitieron que el chico es emmanuel crimen de acro detuvieron a otro barrabrava de river preventiva a Juárez por crímenes de lesa humanidad tras el esperado regreso habló el padre de pilar bauzá buscan a chico de años perdido en el río de la plata abogado de familia febres dice que no hay pruebas caso febres zona liberada y crimen mafioso carla bruni una novia que trae problemas de protocolo el carnaval de gualaguaychú con todo contra botnia valijagate en plena feria allanaron la estatal enarsa caruso técnicos no hablan para conservar el trabajo bus turístico no circula y preparan demanda millonaria detuvieron a un hombre por abuso de hijastra de años para López Murphy la gestión de macri tiene luces y sombras interventor de obsba denunció robo de documentación este año las ventas de reyes crecieron un por ciento empresas de combustibles reclaman \x93medidas urgentes\x94 afirman que el de los despedidos tiene años de antigüedad reyes magos argentinos y la lamparita

no se prende un impacto más político que deportivo un papel protagónico la moral del fútbol apocalypse now de primera las farc se arriesgan a una derrota bienvenidos a la economía real comparaciones ociosas desde la mesa de saldos el regreso de los mitos la ciudad trabaja la lista de lópez mayoría de los ex policías sigue libre celina rucci se confiesa sexo política y dictadura otro atentado suicida en irak muertos y heridos las playas del sur marcan la tendencia de la moda el condominio de antonini wilson en key biscayne guillermo montenegro el ex juez del caso skanska no cree que existan zonas liberadas por la policía federal oreiro recibió una ovación y bolocco zafó del papelón asalto tiroteo y muerte en el barrio de balvanera por internet al qaeda llama a recibir bush con bombas el papa abogó por una distribución justa de las riquezas la inflación de diciembre sería del por ciento los reyes magos gastaron más efectivo que papá noel psicóloga recomendó a pilar que no hable del secuestro piden a perfil una grabación a piquetero kirchnerista yesica bopp mata con los puños y su belleza son dos los chicos muertos por temporal en calamuchita los porteños van a derretirse hasta el miércoles macri filmó con cámaras ocultas el vaciamiento de la obra social se suma otro gremio al paro de los municipales decomisan un millonario cargamento de marihuana sarkozy y carla bruni ya tienen fecha de casamiento cobos elogia a cristina y bromea que tiene poco trabajo estaciones independientes afirman que falta combustible cabrera hace agua en el campeonato de maestros barack obama tiene leve ventaja en las encuestas la educación y sus noticias a largo plazo youznhy aplastó a un nadal que no podía con su alma Chávez lo importante es que emmanuel está libre a mitad de año cierran el relleno sanitario de ensenada macri apuesta fuerte en la pelea con municipales si la ve en el espejo retrovisor mejor déjela pasar con castromán el boca de ischia puso primera confirman que a emmanuel tendrán que operarlo dos muertos por temporal de lluvia y nieve en ee uu cristina reaparecerá con un anuncio ambiental agüero sigue haciendo maravillas en el atlético levantaron el corte del paso fronterizo en colón en diputados el oficialismo no necesitará ayuda cinco personas murieron en accidentes de tránsito buenos aires epicentro del tenis mundial en febrero los cruceros de lujo hacen cola en el puerto porteño los trabajadores del casino pidieron ayuda a cristina adolescente mató de un disparo en la cabeza al padre por lesiones aimar estará tres meses parado y heinze uno un preso muerto en una celda de castigo vóley argentina ganó y quedó a un pasito de los juegos vendado y todo ginóbili volvió a brillar con los spurs comienza esta semana el megacanaje de lamparitas yacyretá batió su récord de generación energética matan a un jefe de las milicias sunitas en bagdad juzgan a un argentino por agresión a guardia español valijagate formalizan en miami las acusaciones contra dos detenidos dos muertos y seis heridos en un vuelco en rosario scioli también revisará contratos y echará a ñoquis bonaerenses francisco casavella ganó el premio nadal de novela britney sufre de trastorno bipolar y depresión posparto ya no es un tema para el gobierno dijo alberto fernández peguen tranquilos en el fútbol fracturar no es un delito investigan robo de joyas a la hija de hugo moyano mar del plata los turistas se conforman con cinco días denuncian a un mercado negro que trafica monedas no es solución tomar a los vecinos como rehenes había dicho mauricio atan roban y abusan de dos primas turistas consejos para prevenir los

golpes de calor a natalie portman le agarró un ataque de maternidad córdoba la capacidad hotelera está casi cubierta un chico mató a golpes a un bebé porque lloraba un ovejero alemán le mordió la cola a germán denis robaron cuatro cuatriciclos en garages y casas de pinamar miceli ahora dice que todo fue armado por el caso greco la justicia procesó a miceli por la bolsa de dinero salió del neuropsiquiátrico y mató al esposo a puñaladas trágico choque múltiple muertos y heridos el sutecha levantó las protestas de esta semana el affaire maradona irán llegó al gobierno de israel genta si algún director o empleado cometió delito debe ser sancionado sancionaron a la línea de micros por cobrar de más extienden hasta el viernes el plazo a monotributistas amenaza de bomba extraño desalojo de empleados y asueto mclaren presentó en alemania la nueva flecha de plata bill gates deja la dirección de microsoft vea su genial video de despedida pidiendo trabajo a celebridades el gobierno prohibió exportar nafta y gasoil reclaman medidas contundentes para asegurar suministro de gas con de térmica la ciudad trata de no evaporarse ate no levanta su medida de fuerza paro de horas el miércoles táctica y estrategia de macri golpear primero gobernar después caso antonini dos detenidos se declararon inocentes indec la inflación de diciembre fue del por ciento la primera reacción argentina fue de asombro y sorpresa más vale tarde que nunca abrieron las colonias porteñas colombia recriminó duramente a néstor kirchner y los garantes cabral el segundo refuerzo para el river de simeone en la provincia habrá mil escuelas con internet se pegó las manos a la cama para no ir a la escuela reír minutos por día es bueno para el corazón león gieco compartirá escenario con bob dylan cristina vuelve a la rosada y decide si se va a españa el sol compensa falta de vitamina d y evita el cáncer estalla en israel un cohete disparado desde líbano alimentos en mal estado en las colonias de vacaciones matan a un ministro de sri lanka en atentado separatista afirman que la sangre hallada en el auto es de maddie la mamá de años más sensual del planeta scioli prorrogó la emergencia vial por seis meses la temporada está de moda en mar del plata dictan prisión preventiva para el piquetero kirchnerista abortos histórico paro en las clínicas españolas balean a un niño de años en una pelea entre barras grassi se defendió por tv me acusan con mentiras misterio sobre la posible boda entre sarkozy y bruni cancelan los golden globe por conflicto de guionistas los hermanos coen son los favoritos para los oscar multan a empresas de colectivos por cobrar de más aseguran que las pruebas contra miceli son \x93suficientes\x94 boca el único argentino en el top ten futbolero el mouse y el teclado ya tienen fecha de vencimiento santiago la provincia de los motines hay heridos argentinos se alzan con un premio literario en la web turismo fatal muertos en dos choques en la autovía y en la ruta resse whitherspoon se negó a filmar escenas de sexo colonias de vacaciones rechazan denuncias sobre alimentos en mal estado una multitud venera al gauchito gil en su santuario crece la venta de protectores solares en supermercados mondino el estado sigue de vacaciones con la seguridad vial maradona denunció el robo de objetos de su museo protesta mapuche en la embajada de chile otro obrero muerto en una obra en construcción embajador colombiano sumó críticas contra los garantes el calor del sauna callejero llegó a los grados ischia aseguró que boca no se retira del mercado evoluciona

favorablemente el escritor mario benedetti freddy grisales tiene todo arreglado con independiente afirman que despedidos tendrán prioridad en los concursos bariloche extranjero quemó a su mujer e intentó matarse alertan sobre ingesta de comidas rápidas en vacaciones habla pilar bauzá no voy a regresar a somalia macri acordó con el gremio y confirmó los despidos para el kirchnerismo el acuerdo no contempla los despidos titular del coprosede admitió que conoce a barras argentina y venezuela economías que crecerán menos familiares de caídos en malvinas quieren ver a cristina colombia no quiere garantes que hagan show mediático el verdugo de diego uno de los más recios de la historia victoria beckham y amy winehouse las peor vestidas verificaron la falta de combustibles y precios abusivos afirman que dar de mamar disminuye ansiedad del bebé en merlo san luis el balance era simple ni una gota de nafta o gaosil en salud de países del primer mundo ee uu es el peor por primera vez en el merval no cerró en baja aunque no aparece quieren llevar a maddie al cine en el primer lugar escrutado ganaron obama y el republicano mccain una familia a juicio oral por un crimen escalofriante en apenas días de ya murieron personas en accidentes cuatro detenidos por el robo de computadoras en ezeiza esperan precio récord en subasta de una obra de bacon no hay fiesta de globos de oro por huelga de guionistas cristina volvió con de térmica y se recluyó en olivos se lesionó nalbandián y no jugaría en australia raptan golpean y violan a adolescente en hurlingham los usuarios protestan por el regreso de los apagones una cerda fluorescente tuvo crías que también brillan george w bush llegó a israel para una visita histórica la policía israelí toma las calles de jerusalén para la visita de bush neri cardozo reemplazaría a insúa en el américa ¿piquetes en la alfombra roja de los premios oscar adolescentes golpean a mujer y lo filman con celulares procesaron a dos empresarios de la cámara de tickets indagarán a banqueros por el \x93corralito\x94 de simone de beauvoir el nombre del feminismo enjambre de abejas casi mata a dos personas en salta belgrano roban un edificio por cuarta vez en meses sepa qué hábitos adquirir para vivir años más saramago fue internado por una insuficiencia respiratoria carrió criticó la lógica de alfonsín que usó macri al pactar con municipales un guardaespaldas para que banega se porte bien carla bruni ya se instaló en la residencia presidencial scorsese estrena un film sobre los rolling stones controlan precios en estaciones de servicio de la costa el colmo del bombero quedarse sin presión de agua una mujer casi se tira desde un edificio del gcba randazzo \x93carrió tiene un problema psicológico que no le permite avanzar\x94 duro balance hizo carrió del primer mes de cristina robbie williams anunció que se retira de la música nina no está en buenas condiciones psicológicas el intendente dijo que uno de los conductores se quedó dormido otra tragedia con turistas en una ruta muertos hebe de bonafini se solidarizó con los despedidos por macri cayó avioneta en el conurbano murieron padre e hijo a ate no le queda claro el acuerdo firmado entre los municipales y macri caranta se lesionó y se pierde los partidos de verano devuelven a museo brasileño obra robada de picasso denuncian que el gobierno programa cortes encubiertos detienen a prófugo por martirio de niño en rito satánico el líder de la barra de boca puede ir a la playa tranquilo bush dijo que irán es una amenaza para la paz mundial carlos el indio solari volverá a sonar en córdoba luego de tres días de sofocante

calor en buenos aires anuncian un alerta meteorológico por fuertes tormentas
robot responde consultas sexuales de jóvenes por chat hillary clinton logró su
primera victoria en las internas desde nueva york blumberg respaldó a daniel
scioli asambleístas de colón cortan por tiempo indeterminado iría a juicio por
tener plantas de marihuana en su placard especialistas españoles certificaron la
identidad de emmanuel el bloque k quiere trabar en la legislatura la
intervención de obsba afirmó Chávez que ya tiene las coordenadas para ir a
buscar a las rehenes clara rojas y consuelo gonzález el desafío de mauricio
macri detectan misteriosa fuente de rayos x en galaxia vecina la cruz roja
corroboró que el rescate de rehenes se hará mañana difunden recaudos para
aquellos que vayan a brasil presidenta part time rápida respuesta colombiana
autorizando la operación mauricio macri no revisaremos ni un sólo contrato
hace días alberto f negaba que hubiera crisis ni mucho menos por primera vez
los springboks tienen un técnico negro presentaron el programa de saneamiento
del río salí dulce cristina fernández reconoció que hubo mil cortes de luz
mercedes marcó del pont será titular de banco nación la cinematográfica historia
de emmanuel pensamos que era niña la mona festeja su cumple regalando
;aguinaldos argentina está en contacto permanente pero no mandan a nadie
niegan en brasil que haya un brote de fiebre amarilla caso antonini difundirán
pruebas fílmicas y sonoras encara otra pelea tan dura como la de los gremios
tapar baches brasileños afirman que el dinero genera felicidad argentina sin
representantes pero en contacto permanente alerta por tormentas y granizo en
gran parte del país otro atentado suicida en pakistán muertos máxima se cotiza
cobra más del doble que bush alberto fernández contradice a cristina \x93no hay
problemas de energía\x94 prohíben las ojotas y bermudas en un cementerio
chocaron dos colectivos pasajeros con heridas abuso sexual odisea de turistas
argentinos en brasil desvalijan a un grupo de adolescentes en villa gesell
mónaco venció a llodrá y ya está en las semifinales el padre de madeleine quiere
volver el tiempo atrás más de mil pasajeros quedaron varados en ezeiza temporada
trágica otros dos muertos en choques tiempo loco frío y agua nieve en
bariloche violenta represión a manifestantes mapuches en chile asesinó a su
novia de años y se suicidó con la misma arma estudio español confirma que el
chico que está en bogotá es emmanuel joven mató a su novia de años por disputa
familiar afirman que es una barbaridad usar a empleados para datos privados de
ucrania con amor para el agente james bond la justicia ordenó suspender la
intervención de obsba cortes de calles contra los cortes de luz hacen un
piquete en flores en sutecba habían hablado de fallo contundente el bafici
abrió su convocatoria a directores argentinos américa latina millones de niños
sin agua potable defensoría del pueblo confirmó falta de higiene y seguridad en
las colonias la reina no quería que diana luchara contra el sida paulina rubio
está de luto por la muerte de su perra en el gobierno se sienten partícipes
del rescate de los rehenes genta le pidió a macri que respete el fallo que
frenó la intervención guillermo cañas descartado para el abierto de australia
carlos bilardo grondona está en la fifa gracias a mi el colombiano del
escándalo sexual se fue de arsenal sale a la calle el auto más barato del mundo
u s ahora esperan por la liberación de ingrid betancourt el macrismo piensa en
un juicio político a la jueza que frenó la intervención elisa carrió advirtió

sobre los \x93papelones\x94 internacionales de cristina descubren cómo crear células madre sin dañar embrión al final cristina fernández no viajará a españa estados unidos felicitó a uribe y a desgano mencionó a Chávez raposo alan schlenker quería ser presidente de river gustavo cerati no participará del live earth latino acusaciones cruzadas entre aerolíneas argentinas y los gremios murió jorge isaac anaya impulsor de malvinas los porteros serán los nuevos informantes del gobierno las rescatadas trajeron pruebas de vida de otros ocho rehenes denuncias y desmentidas por las colonias de vacaciones accidentes y crisis energética un camino al infierno mujer que quiso suicidarse no había pedido un crédito clara rojas ahora sólo quiere recuperar a emmanuel en el ranking top de dt a basile lo sigue una mujer indec los salarios aumentaron un en un año \x93no es ni más ni menos que cumplir con la ley\x94 dijo mauricio cristina fernández felicitó a Chávez por la liberación contaron dramáticos detalles de la vida en cautiverio a maionica le denegaron la libertad por peligroso para la comunidad luego de largo silencio uribe agradeció a Chávez y a fidel castro reapareció uribe y agradeció a Chávez y a fidel rodrigo archubi arregló su incorporación a river condenado por intentar chantajear a tom cruise clara rojas y consuelo gonzález las rehenes liberadas volvieron a vivir vence el plazo para el reempadronamiento en la afip mónaco no pudo con kohlschreiber y quedó eliminado ee uu guionistas en huelga eligieron lo mejor de bush presagia un acuerdo de paz palestino israelí brigadas católicas por la virginidad y contra los gays Canadá prohíben a los homosexuales donar órganos desbarataron banda de ladrones de autopartes con apoyo de carrió macri logró que la legislatura apruebe la intervención obesidad en EEUU consecuencia de un estilo de vida policia se quedó dormido y disparó contra su superior roban documentación y golpean a dirigente de hijos descubren proteínas que hacen propagar al sida bush deberíamos haber bombardeado auschwitz murió el primer hombre que conquistó el everest michetti espera avanzar con la intervención de obsba padres de maddie enojados con agencia de dobles johnny depp en el estreno mundial de su nueva película los porteros no están obligados a informar sobre aires acondicionados arsenal en colombia justifican el plantón de castrillón un turista ebrio agredió a golpes a un agente municipal chile temen explosión de magnitudes en volcán llaima \x93cuando solucionemos estos problemas vendrán otros\x94 ocaña calificó a carrió su ex aliada de \x93mezquina\x94 las farc piden un territorio desmilitarizado para negociar clara rojas tres años sin noticias directas de ingrid betancourt tras años de casado se divorció el \x93hombre araña\x94 una página web junta firmas para reelección de tabaré racing y san lorenzo abren el juego veraniego en salta con los ojos en el una droga revirtió en minutos síntomas del alzheimer el gobierno garantizó la realización del carnaval porteño el gobierno desmintió el regreso de d elía a la función pública maradona le regaló una camiseta al presidente iraní las películas que competirán por los cóndor de plata las rehenes liberadas por las farc vendrán a la argentina tensión e incidentes en reapertura del casino flotante mar del plata en segundos roban mil del banco simeone confirmó que ortega será el capitán de river politica y economia en el arranque de Chávez no se rinde insiste con la reelección indefinida cristina vs Néstor Aníbal fernández aseguró que a febres lo asesinaron\x94 por primera vez en un siglo

cayó nieve sobre bagdad Chávez pidió no catalogar de guerrilleros a las farc
naomi campbell me he convertido en fan de cristina el birmingham inglés casi
deja a boca sin castromán un reclamo de pasajeros paralizó los vuelos en ezeiza
grisales da como un hecho su llegada a independiente la soberbia mata luz en la
oscuridad entre mejores plumas el campo nuestro lo demás es silencio en latín
también un héroe de nuestro tiempo morir en cisjordania hágase la luz cuántos
tiroteos podrá enfrentar la argentina blindada por la soja petroleras acordaron
volver a exportar y bajar las naftas racing arrancó el con el pie derecho
consuelo gonzález dedicará su energía al canje humanitario el de los
argentinos puede tener ciudadanía italiana naomi campbell se declaró fan de
cristina kirchner uzcátegui specht asegura que antonini estuvo en la rosada
reverten daños en el hígado producidos por cirrosis colombia no cambiará
estatus de terroristas a las farc en familia tinelli descansa en punta y
planifica el bush dice que está en buen camino para reducir tropas afirman
que maddie fue violada y arrojada a un lago la volpe será entrenador de los
rayados de monterrey nocioni recupera nivel y scola es revelación de la nba ex
beatle ringo starr estrella de la fiesta de liverpool aseguran que quien
traía la valija no era antonini wilson es muy preocupante la situación
sanitaria en santa fé denuncian graves faltas de recolectoras de residuos a los
años murió el poeta español angel gonzález a partir del lunes se verificaría la
baja en la nafta clara rojas dijo que el padre de emmanuel \x93habría podido
morir\x94 farc el gobierno tomó distancia del pedido de Chávez las mascotas
argentinas cada vez más \x93lookeadas\x94 sarkozy estresado novia embarazada
e hijo músico tba debe reestablecer el tren blanco de los cartoneros
adolescentes en pinamar seducción descontrolada hubo forcejeos entre un
diputado uruguayo y un ambientalista grietas en la alianza entre socialistas y
elisa carrió empleados y apostadores en catamarán al casino flotante el yabrán
de kirchner manejará el hipódromo tiger woods se encamina a los u s millones
clara y consuelo denuncian delitos de lesa humanidad tevez otra vez genio y
figura en el manchester united guerra de marcas y estándares en la electrónica
pasajeros furiosos atacaron mostradores de aerolíneas en ezeiza colón levanta
corte de ruta pero sigue con su lucha assembleístas de concordia evalúan cortar
único paso libre a escapar de buenos aires vuelve la ola de calor florencia
kirchner en un cinco estrellas de pinamar ¿por quién doblan las campanas la
ofuscada señora presidenta como nosotros pero distintos la gran ruscullada
elogio del amor insensato hillary recrea la coalición el método k de resolución
de conflictos el sr bellatin y su psicoanalista una relectura de la derecha la
historia de las farc una noche en el casino la fiebre los chicos de simeone se
durmieron y les empataron boca suma su último refuerzo el paraguayó cárceles
jessica alba recomienda a las chicas seguir vírgenes hagan sus apuestas que
nalbandián es favorito los prostíbulos de punta del este quedaron sin secretos
se acabó la tregua vuelve el calor sofocante no videntes controlan la calidad
de los alimentos cuestión de pesos asaltaron a alberto cormillot el sueño del
hijo de bin laden es vivir en gran breaña ¿madonna es el nuevo símbolo de la
mujer moderna robo del siglo un acusado teme terminar como chabán guinzburg
kirchner y cristina dan para hacer humor lavagna argentina hace de furgón de
cola de Chávez irán se ríe de estados unidos han hecho el ridículo el

titular de enarsa figura como empleado de petrolera hedonismo espiritual la nueva tendencia de los jóvenes corrupción en interpol renunció su presidente las nuevas mimadas del cine argentino los mitos que agravan los trastornos del lenguaje richard burton también se rindió ante marilyn monroe el día después de los incidentes continúan las demoras en los vuelos internacionales que parten desde ezeiza asambleístas de concordia analizan cortar el puente luis delía se sacrifica y no volvería al gobierno mamitas nicole richie le ganó a christina aguillera blejer va a ser difícil que los precios no suban de la rúa se arrepiente de haberse ido en helicóptero antonini aportó grabaciones de audio y ocho de video de las reuniones las privatizadas en la mira river ya conoce a todos sus rivales para la libertadores la suerte de de vido depende del silencio de un abogado detenido la basura puede costarle una fortuna a nápoles en ezeiza reprograman lentamente los vuelos consuelo gonzález puso límites a los elogios de Chávez a las farc muchas construcciones sin declarar y patentes impagas mariano pavone y diego milito marcaron por partida doble miceli sigue en contacto con funcionarios de cristina scioli de festejo con famosos políticos y empresarios yo no apoyo a las farc apoyo la paz aclaró Chávez triunfo de greenpeace frente a cazadores de ballenas subastan más de autos de colección en ee uu los accidentes de tránsito que al gobierno parecen no interesarle costaron hoy las vidas de otros argentinos en el debut de banega el rey de la noche fue agüero mardel despidió la primera quincena con sol y sirenas murió el bajista de babasónicos gabo mannelli el estremecedor abrazo de clara con su hijo emmanuel exigen quitarle la nacionalidad israelí a barenboim todos los ganadores de una noche sin sorpresas ni glamour calleri fue la única alegría de la legión argentina bardem se quedó con el globo de oro y va por el oscar trágico recambio de temporada trece muertos en córdoba ¿cuál es el valor de la eficacia cientos de prostitutas rezan por el dictador suharto golpean y asaltan a una pareja de ancianos en caballito le prohíben correr con sus prótesis de carbono irá a juicio oral por disparar al aire en navidad dodi al fayed no fue el último gran amor de diana hawking no hay vida extraterrestre cerca de la tierra estuvieron cuatro horas varados por desperfecto en tren reclaman que la seguridad vial sea política de estado en las estaciones de servicio se niegan a bajar precios cuenta regresiva para pasarse al sistema de reparto boca el nuevo millonario del fútbol argentino björk enloqueció y atacó a un fotógrafo controlaron un incendio en el hospital militar en aeroparque tantos problemas como en ezeiza millonario premio si tiene poderes paranormales choque de trenes al menos personas heridas los pitufos cumplen medio siglo e invaden europa comienzan la construcción de la pastera de ence en colonia españa rodríguez zapatero disolvió el parlamento guillermo espera callado pero no perdona a román clara rojas ya disfruta de emmanuel soy la mujer más feliz del mundo en las montañas las farc secuestraron a más turistas sarkozy ya se habría casado con carla bruni el nuevo fallo de la jueza que desató la furia del macrismo vuelven los reclamos por botnia frente a buquebús ni un solo argentino en el top ten mundial de arqueros el deshielo antártico se aceleró un en años en los juegos olímpicos chinos checho será el coco amia interpol localizó a uno de los iraníes acusados los gastronómicos se quejan del cambio de horario en Perú repsol y petrobras hallaron enorme yacimiento se reempadronaron más de

cien mil monotributistas como en la antigüedad el papa ofició misa de espaldas el louvre de paris el museo más visitado del mundo uruguay protesta ante cristina y la haya por cortes detuvieron a prófugo acusado de matar a sus ex suegros una nena de meses murió atropellada por un trencito casino sigue la tensión entre prefectura y despedidos dormir minutos de siesta ayuda a la memoria victorias de nalbandián y múnaco en australia el debut de ischia en boca será con los suplentes la sonda messenger explora con éxito mercurio luego de larga reclusión antonini reapareció haciendo shopping en florida promueven alquiler de bicicletas para aliviar el tránsito siguen las demoras con los vuelos de aerolíneas colón corta el puente a paysandú en solidaridad a la protesta en buquebús sorpresiva visita de condoliezza rice a irak paulo coelho mensajero de la paz de la onu siete palestinos muertos por incursión militar israelí hartfield duró poco más de una hora ante federer afirman que expedición de colón llevó la sífilis a europa marche preso stallone no afloja con los esteroides se acabó el misterio revelan quién fue la mona lisa kurt cobain ya tiene quien lo interprete el iraní localizado en Bélgica no es prófugo de la amia fiebre amarilla no se vacune si viaja a brasil el hijo de nicolas sarkozy amenazado de muerte shopping debe indemnizar por el robo de un auto scarlett johansson debuta como directora córdoba dos jóvenes murieron en accidentes de tránsito consuelo gonzález entregó las desgarradoras pruebas de vida cristina fernández suspendió su visita a entre ríos ramón enigmático quiere una joyita para san lorenzo embarazó a una nena pero fue condenado por abuso botnia asambleístas no pudieron llegar a buquebus el papa no asistirá a un acto por protestas en su contra simularon fusilarlo para que les entregue todo el dinero se negó a declarar el acusado de matar a ex suegros escritores lanzan sus obras por un youtube literario en el casino aseguran que no reincorporarán a nadie dicen que casi todas las estaciones de servicio retrotrajeron precios marcha atrás en el caso por el crimen de gonzalo acro ee uu aprobó venta de derivados de animales clonados el gobierno porteño ya apeló los amparos de sutecba gobierno porteño apeló límites a intervención en obsba denuncian desinterés del gobierno por rehén argentino de las farc panamericana choque múltiple dejó al menos heridos raúl castells macri es un niño bien de la oligarquía docentes porteños denunciaron más de cien despidos tras el calor insoportable se vienen las tormentas bignone hay que terminar el nürnberg argentino farc ¿grupo beligerante o terrorista preocupa falta de evolución en la salud de benedetti cancellería desmiente que no hayan colaborado con la familia gustavo kuerten confirmó que no jugará más al tenis caso antonini wilson el juicio empieza el de marzo un incendio destruyó la fábrica de pastas la salteña anibal f dijo que macri hace lo mismo que el gobierno levantan la prohibición de exportar combustibles jaime compra nuevos trenes para la línea san martín superávit primario de superó los millones clara rojas y consuelo gonzález vienen a la argentina el merval se sumó al mal clima global cayó un presentan la notebook más delgada del mundo independiente venció a boca en mar del plata se extiende el alerta de tormentas podría caer granizo agustín calleri también se despidió de australia tras el martes negro se derrumbó bolsa de hong kong educación buscan el perfil de un leonardo da vinci aumentan los robos cometidos por adolescentes brutal atentado contra un micro escolar muertos el

fuerte temporal provocó la caída de árboles y carteles hallan en uruguay el fósil de una rata gigante tras el arresto de su esposa dimite un ministro italiano ladrones aprovecharon el caos que provocó el fuego paul organizará exposición de fotos de linda mccartney aníbal fernández no aplicará la fuerza en las protestas araceli me separé con el amor con el que me casé casino ex empleados piden la intervención del gobierno las chicas defienden su derecho a nadar en topless luis miguel debe pagar mil dólares a un músico scioli no quiere gastar un peso en nuevos empleados los stones y scorsese irán al estreno de \x93shine a light\x94 adivine dónde cayó un chileno con pedido de captura la enigmática internación de gwyneth paltrow manu chao enamorado de los pacientes del borda alertan que la baja de contratos va a afectar al sistema educativo presentan software que decodifica el ladrido de perros homenaje al primer hombre que escaló el everest ischia dio la cara \x93las cosas van a ir saliendo mejor\x94 dejaron a pie a todos los jugadores de river parís se prepara para el último desfile de valentino comprador personal servicio de lujo para señoras bien fidel \x93no tengo capacidad física para hablar en público\x94 una anciana murió de un paro cardíaco al ser asaltada schlenker volvió a desvincularse del asesinato de acro fabrican robot que activa sus piernas a distancia por el caso de jorge gillanders miller el gobierno se siente agraviado sin oscar xxy no podrá competir por una estatuilla la inflación en estados unidos llegó al nivel más alto desde aníbal ibarra desmintió haberse reunido con la jueza liberatori la jueza aceptó un recurso de amparo de municipales cesanteados acreedores se reagrupan para protestar contra argentina veteranos de malvinas molestos por rompehielos inglés mitt romney ganó la interna republicana en michigan el martes cristina recibirá a moyano en casa de gobierno chocaron dos colectivos en mataderos hay heridos las piernas largas provocan mayor atracción sexual américa latina poco tentadora para inversores los argentinos leen menos de cinco libros al año gobierno porteño perdonó multas de tránsito si el vino es caro el cerebro dice que es más rico mujeres con cáncer de mama sufren discriminación sonríe el barcelona lionel messi vuelve a las canchas visión española una licitación cortísima lejos de las fórmulas habituales macri calificó de disparate el último fallo de la jueza liberatori sólo en los últimos dos años hubo ocho descarrilamientos confesó un acusado por crimen de niño en rito satánico diciembre en rojo por primera vez en tres años millones leandro lázzaro abandonó tigre y se fue a estudiantes cristina fernández anunció el tren bala a córdoba el pro pide juicio político la jueza que no la llame la prensa el gobierno defiende el resultado de cuentas públicas el kun agüero e ibagaza brillaron en españa hombre araña asalta y viola a una joven en núñez sarkozy pidió que argentina redoble esfuerzos para liberar más rehenes club alemán subastó en internet el puesto de técnico paritarias moyano no le cree a la inflación del indec accidentes de tránsito provocaron otras seis muertes rutas fatales trágico choque en atalaya nalbandian y mónaco avanzan en el abierto de australia para la onu la condición de los rehenes es infame fusilaron a un taxista cuando volvía a su casa israel prometió mantener presión sobre hamas en gaza joven violento es enviado a siberia \x93para que aprenda\x94 reconstruyen operativo de rastrillaje por julio lópez exhaustiva búsqueda en españa de nena desaparecida johnny depp donó un

millón de dólares a un hospital llevarán al cine otra obra de garcía márquez un tren lleno de turistas estuvo dos horas varado la versión de aerolíneas es que no hubo cancelaciones acusan a macri de \x93apretar\x94 a liberatori y lo tratan de papelonero caso acro rechazan la excarcelación de alan schlenker suplementos de calcio pueden aumentar riesgo cardíaco el procurador porteño está sorprendido por los ataques de la jueza eddie murphy no duró ni dos semanas casado muertes y nuevos malvados sorprenden a los héroes casino los despedidos desmienten reincorporaciones el sindicato cuestiona el campamento frente al casino encuentran cuadernos con bocetos que serían de van gogh basta de misterios con las obras de matisse y picasso caso lópez habrá careo entre testigos del rastrillaje los stones dejan a su discográfica cerca de la quiebra temporada de robos asaltan a familias en miramar control de las protestas en las calles nueva advertencia a castells prueban nueva terapia contra los dolores crónicos el gobierno porteño recusó a la jueza elena liberatori grandes firmas de la moda italiana al mercado bursátil avalan la prisión por tenencia de drogas para consumo ramón díaz no se guarda nada para enfrentar a boca denuncian al hombre araña por otras violaciones en zona de núñez una web para que le devuelvan los objetos que prestó impresionante aterrizaje de emergencia en inglaterra buena sintonía binner logró acuerdos con el gobierno en europa quieren llenar la cancha con más árbitros para el ministro de justicia bonaerense bajó el índice de homicidios san luis anunció que retrasará los relojes una hora regresan al trabajo algunos municipales despedidos vuelven acusaciones cruzadas entre bogotá y caracas nacieron dos siameses unidos por el hígado en rosario el tren bala el optimismo de cristina y el tango siempre hay un rincón con grabaciones de los beatles bush y bernanke contra la recesión en estados unidos crearon embriones humanos por clonación de células madres de plaza de mayo \x93desembarcarán\x94 en la esma volverían a cortar los tres pasos fronterizos con uruguay no hay que analizar sólo esos bienes sino los de empresarios amigos tres heridos por la inédita caída de un puente en ezeiza el chelsea quiere a messi y \x80 millones por drogba la confianza en la economía cayó durante este mes los alimentos con cobre reducen estrías y osteoporosis riquelme mascherano y demichelis un gol olímpico expreso imaginario cárceles en azul y oro la fuerza aérea boquense denuncian a titular de enarsa por aportes de petrolera confiscaron toneladas de juguetes chinos truchos kirchner salió de la rosada con millones más de los que entró vengo en nombre mío y del compañero presidente récord de exportaciones pero caída del superávit comercial indec echan por la fuerza a un delegado de ate hipotecarán la destilería de shell en dock sud gobierno exigió a aerolíneas que normalice servicios indignación en hurlingham por el asesinato de un almacenero francia tiente con más megaobras a de vido y jaime los nuevos millones del patrimonio k la oposición exigió una investigación después de un impasse vuelve la lluvia a la ciudad gaza peligra diálogo de paz por escalada de violencia asaltan a mano armada una casa de fotos en san clemente cariló roban dinero y objetos de valor a una familia murió el legendario campeón de ajedrez bobby fischer creen que el asesino no siguió disparando porque se le trabó el arma despidieron los restos de la mujer asesinada en el country nadal avanza a paso firme en el abierto de australia control de alcoholemia en la costa secuestran autos afirman que el índice de

robos es como el del año pasado el gobierno porteño dice que identificará a los responsables de la marcha asaltaron oficinas privadas de alberto fernández la cia acusa a al Qaeda del crimen de bhutto el gobierno nacional se distancia de macri pero también de castells ricardo mollo homenajeará a yupanqui en cosquín científicos hallan una clave de la leucemia infantil show de moda termina con streap tease y un detenido javier bardem candidato al oscar y actor de moda crimen del country el ex apuntó contra supuesto amante archubi tuvo su primera práctica como jugador de river la mayoría de los porteños aprueba los matrimonios gay murió uno de los bebés siameses nacidos en rosario más de millón de personas abandonaron las afjp incidentes con despedidos por el intendente de la plata bush instó al congreso a aprobar un estímulo económico el vaticano condenó la creación de embriones humanos el sindicalismo también quiere a kirchner presidente castells movilizó a piqueteros sin autorización municipal lindsay lohan fue condenada a trabajar en una morgue cristina frente al \x93cambio climático\x94 ¿qué es la realidad en materia energética protesta piqueterizada hillary clinton habló sobre la infidelidad de su marido liquidación de temporada ischia aceptó que habrá bajas cristina prometió ayuda económica a familiares de ex combatientes el gobierno anunció millonaria exportación a venezuela continúa la pulseada por el traspaso de la policía george clooney nuevo embajador de onu para la paz sin vueltas murió el inventor del hula hula y el frisbee bolivia pidió paciencia para garantizar el envío de gas labran infracciones contra piqueteros por llevar capucha y palos alfonsín falta diálogo entre oficialismo y oposición la fábula del pastor y el lobo y la histeria social de nunca acabar no sorprende pero duele suspiros por beatrice el mal absoluto el pasquín trotskista la ruta a ninguna parte una lección ejemplar ¿y yo qué tengo que ver de la favela a copacabana diferencia de tres meses peronistas de todo el país insisten en que kirchner será el presidente del pj consuelo recibida como heroína en su pueblo natal festividad chiita enlutada por al menos muertos temporada de robos al menos asaltos en la costa asambleístas intentan cortar los tres pasos a uruguay los diez minutos finales le dieron el triunfo a river latinoamérica pisa fuerte en el festival de sundance ¿ovnis en córdoba o una aeronave de la nasa israel avanza con gran ofensiva contra hamas en gaza los lujos millonarios de abramovich anclaron en punta federer ganó pero transpiró más de lo previsto juan mónaco eliminado por berdych en australia pakistán detienen a sospechoso de atentado a bhutto es sudafricano le robaron todo y duerme en la calle peloso dijo que el mijd no pedirá permiso para marchar san luis ratificó cambio de hora y criticó al gobierno secuestran mil kilos de cocaína hay veinte detenidos otra prueba electoral para demócratas y republicanos brasil paranoia por fiebre amarilla causa problemas aparato ayuda a caminar a enfermos de parkinson un español es el nuevo \x93papa negro\x94 de los jesuitas \x93todavía no encontré al padre de mis hijos\x94 messi está listo para jugar y hace ilusionar al barcelona abuelos contratan strippers para divertirse en sus fiestas acusan a liberatori de querer reemplazar a macri mario vargas llosa internado por problemas cardíacos el manchester united de tevez sigue firme en inglaterra uruguay comparó cortes de ruta con las torres gemelas hágalo usted mismo edite su propio video de r e m vidas paralelas cristina y macri apagan incendios limitan el uso de

videocámaras en espacios públicos rehenes acusan a Chávez de armar toda la escena como maradona el señor barriga se operó la panza dispuesto a competir rodríguez saá pidió \x93un tribunal imparcial\x94 operativo k en puerto madero para controlar el pj bellos hábitos nicole kidman pensó en ser monja más accidentes de peatones por uso de celulares y hillary derrotó a obama en nevada gracias a hispanos desear o querer una bala para el tren bobby fischer úselo y tírelo trágica predilección elija usted la forma nietzsche te coloca en europa pibe qué son realmente las farc se acabó el colchón malditos poetas sobre la crítica y la tradición la mujer ilustrada piedra libre solidaridad en la región las primarias más inciertas la inclinación autobiográfica boca y san lorenzo quedaron a mano un desconocido nalandian fue eliminado en australia detienen a ex represor que fue chofer de aníbal gordon la comunidad homosexual ya tiene su propio geriátrico los cubanos van a las urnas para renovar el parlamento marcha en apoyo al papa tras polémica con universidad catherine fulop Chávez compró al presidente kirchner el paraíso en el que ricos y famosos andan desnudos calificación negativa para la política de empleo los nocheros brillaron en la primera luna de cosquín la sigen denuncia irregularidades en enarsa a sean penn ya no le divierte la actuación la única testigo del crimen del country rompe el silencio díaz bancalari quiere una amnistía total en el pj para asambleístas quieren limpiar la imagen de botnia para castells decir que el gobierno es de izquierda es una truchada porque los kirchner son conservadores duelos argentinos en el máximo nivel de la nba sólo repartieron lámparas de bajo consumo cóppola en otra época hubiese seducido a cristina un contrato nefasto moreno pagó una cifra millonaria por un obsoleto satélite que funcionó solamente días carrió el conflicto gremial que se viene será el más grande de la historia en el caso de la valija el gobierno no ha investigado nada kirchner destruyó el pj que está desactivado y en manos de una jueza no sé si macri va a tener la fuerza política suficiente para hacer cambios eduardo menem el país lo gobierna una diarquía además de evasores montoya vigila residuos tóxicos ibarra sufrió rotura de menisco de rodilla y será operado el nuevo idenikit del posible secuestrador de maddie polémica por colegio de la salle proyecto de ibarra vuelven en febrero los candentes debates legislativos hugo Chávez amenaza enviar al ejército a la frontera hillary y mccain ganaron en nevada pero la lucha por candidaturas a la casa blanca se mantiene en suspenso el real madrid fue demasiado para el kun agüero israel dice que hassan nasrallah \x93debe ser eliminado\x94 en México pekerman le ganó el duelo a brindisi roban millones en casa de un abogado santiagueño un incendio avanza en la zona de sierra de la ventana el nou camp ovacionó la vuelta de lionel messi cambiasso y cruz figuras en el puntero italiano colón levantó su corte de fin de semana los cubanos como siempre votaron a fidel castro automovilista mató a dos chicos e intentó escapar león gieco iluminó la segunda luna de cosquín con oliver stone bush seguirá los pasos de nixon denuncian conspiración para matar al presidente correa ya no es noticia turistas quedaron varados en un tren el congreso investigaría el patrimonio de kirchner parece que los políticos descubrieron que hay ñoquis independiente tuvo una noche incomparable javier calamaro tiene que aprender a conducir bolsas de asia y europa descreen del plan de bush fracasa plan del gobierno para rebajar las naftas river no

duerme pensando en boca otra vez un apagón afectó a miles de usuarios hipnosis publicitaria no es verdad aunque usted lo crea johan cruyf criticó a messi y pidió que lo sancionen el encuentro menos pensado de susana giménez la magia volvió al carnaval de gualeguaychú picanean a una chica y está muy grave a rosana no le interesaba el dinero de su ex marido asaltan al puma manuel contempomi en san isidro super federer sigue adelante en el horno australiano taiana destapó un agujero negro el ingreso de autos diplomáticos sigue el incendio de campos en sierra de la ventana llevará tiempo emerger de la crisis y la volatilidad cuidado con el café duplica el riesgo de aborto el guerrero bush homenajeó al pacifista luther king la bellísima lindsay lohan nominada como peor actriz gallardo no viene a san lorenzo e ira a ee uu para el titular de aduanas es contrabando agravado se vende el mayor proveedor de fotos del mundo \x93la gente no está comprando la historia del rescate estadounidense\x94 la economía estadounidense arrastró hacia abajo a las bolsas mundiales el merval cayó más del por ciento un joven murió ahogado en una pileta de versalles tres diputados asistieron sólo a tres de veinticinco sesiones en menem fue una estatua legislativa en llega el último libro de harry potter en español david mónaco calleri y acasuso poker de ases nadal arrollador ganó fácil y ya está en semifinales cumbre en balcarce cristina recibe a moyano roban en monedas de un supermercado chino boca enfrenta a racing con mayoría de suplentes pena de años a cura que asesinó a su hija y su mujer israel alivia por un día el bloqueo en la franja de gaza es palabra de carlos bianchi \x93boca es el mejor\x94 cuatro jóvenes armados en el golf club de cariló hillary y obama se dieron con todo en debate demócrata lo estaban velando despertó y comenzó a hablar los ingleses pronostican que perderán por paliza la aduana se despega del contrabando fue cancillería descubren un virus que provocaría el cáncer de piel pret a couture la opción de moda para los impacientes terrible hoy morirán mil chicos menores de años catamarca quiere seguir los pasos de san luis cero en geografía para la actriz charlize theron es la historia de un amor como no hay otro igual coppola hollywood fomenta el lavado de cerebro\x94 divididos fascinados pese al rechazo del público ee uu baja las tasas para frenar la crisis financiera mientras las bolsas europeas y asiáticas vuelven a caer habló el jardinero y aseguró que \x93no tiene nada que ver\x94 \x93no country\x94 y \x93blood\x94 lideran nominaciones a los oscar caso lópez los testigos mantuvieron dichos en el careo \x93no\x94 de gracián a san lorenzo \x93quiero jugar en boca\x94 zapatero promete medidas a favor de los inmigrantes el ex marido de rosana tuvo un pico de presión y está sedado galliano más testigos hablan de un hombre que huía de el remanso greenpeace denunció utilización de lamparitas comunes macri y sutecba negociarán salarios la próxima semana otra tragedia tres muertos por un vuelco en córdoba cristina también estuvo con los pobres visitó la fundación carasucias ee uu redujo las tasas para frenar la crisis financiera carla bruni desmiente su casamiento y posa desnuda vishing nueva herramienta de piratas informáticos servini dejó las vacaciones y resolverá si la deriva a la corte reparto de gas evo quiere una reunión con lula ser o no ser generoso tan sólo una cuestión de genes cristina recibió a moyano y a la cúpula cegetista mensaje a barrionuevo no se puede ser traidor a compañeros no habrá salvataje sino moratorias para obras sociales

escándalo en rosario un concejal golpeó a empleada una foto que lleva tranquilidad a ambas orillas fray bentos una ciudad aletargada sin el movimiento de meses atrás botnia lamenta no contar con recurso humano argentino no hace mucho lo habían visto en el subte sucio y desaliñado tevez entre los tres jugadores más feos del mundo una cadena de romances signó sus últimos meses la respuesta en gualeguaychú fue ardiente quemaron una bandera una muerte que trae el recuerdo de river phoenix degollaron a una contadora en gualeguay moyano \x93cristina nos dejó satisfechos a todos\x94 boca más juniors que nunca empató con racing tenía esa cosa misteriosa que ilumina la pantalla dijo mel gibson el campeón lanús fue demasiado para vélez alivio financiero en la bolsa de hong kong y en europa independiente enfrenta a un san lorenzo de ensayo contrabando la larga lista de los diplomáticos sospechados subastan piezas de historia natural en parís novak djokovic castigó a ferrer y es semifinalista una atípica reunión de ejecutivos políticos y artistas tras la crisis bursátil elite económica discute en davos los problemas en los ojos no pueden esperar miles de palestinos se lanzan a egipto para abastecerse hallan en china una calavera humana de años araceli y mazzei protagonistas del romance del año federer superó a blake y ya no hay quien lo destrone china cerró sitios web pornográficos en balean el automóvil del intendente de almirante brown reclaman prohibir el popular videojuego counter strike un taxista fue asesinado a puñaladas en la ruta la extraña adicción de la actriz christina ricci desastre en bolivia muertos por inundaciones asalta a turistas en mar del plata y antes de irse pide perdón de rodillas asaltan a una familia que descansaba en chapadmalal diane keaton no tiene un hombre que la mime apareció un hombre capaz de ganarle a google morir antes de los las jóvenes promesas que se convirtieron en mito detienen a un hombre por el robo del corazón de esquí crimen de la contadora detienen a empleada doméstica para vivir más pruebe discutiendo con su pareja encuentran un mensaje de helena christensen en el contestador diana había averiguado para casarse con un musulmán hallan muerto en manhattan al actor heath ledger hoy las habituales críticas presidenciales a la prensa fueron matinales la sorpresiva muerte obligó a cambiar la promoción de batman fiscal descarta que arce haya sido el autor material la lista de sospechosos del viudo incluye a un heladero nació la nueva biznieta de mirtha legrand vecinos de santa ro'

4 3 - Limpiar los Datos

```
[34]: %%time
# Aplicaremos varios rounds de limpieza
import re
import string

def clean_text_round1(text):
    '''Elimina todo lo que no es texto'''
    text = text.lower()
```

```

text = re.sub('\[.*?\]\%', ' ', text)
text = re.sub('[%s]' % re.escape(string.punctuation), ' ', text)
text = re.sub('\w*\d\w*', '', text)
return text

```

```

round1 = lambda x: clean_text_round1(x)

```

CPU times: user 338 µs, sys: 0 ns, total: 338 µs

Wall time: 78.2 µs

```

[35]: %%time
      # vemos la primer limpieza
      data_pe_clean = pd.DataFrame(data_pe.Texto.apply(round1))
      data_pe_clean

```

CPU times: user 127 ms, sys: 39 µs, total: 127 ms

Wall time: 126 ms

```

[35]:

```

	Texto
2006	córdoba alianza delasotista se impone en marcos juarez el reclutamiento docente cae en picada relatos de una obsesión ya no podía entrar al aul...
2007	la historia se repite más de cien heridos con quemaduras de pirotecnia y lesiones en los ojos el primer bebé del año nació en buenos aires quemar...
2008	atentados y operativos de seguridad oscurecen festejos el primer bebé porteño del año nació en parque patricios un motociclista perdió la vida al ...
2009	balas perdidas corchos riñas callejeras e incidentes con pirotecnia en los festejos de fin de año cromañon en bangkok muertos en una discoteca...
2010	peajes bien chévere los placeres mínimos de cada día para estar saludables mateo ramos se postula para ser el primer bebé de voraz incendio en be...
2011	toman otro predio en villa lugano durante los festejos de año nuevo hoy asume dilma rousseff en brasil en santiago del estero el arrancó con un s...
2012	murió el gobernador de río negro quién asume el gobierno en río negro primeros detalles y repercusiones de la confusa muerte de soria carlos soria...
2013	pablo echarri la militancia me llegó tarde a los infografía feriados encuentran muerto al andinista perdido en el aconcagua boudou pasó el...
2014	no creo que haya amor creo que hay propiedad privada y no me la toca nadie la verdad mexicana no es un lujo es algo que necesito la conferenci...
2015	caso lola aseguran que es extremadamente complicado choque y vuelco fatal en camino del buen ayre boudou el embajador en la reasunción de di...
2016	la búsqueda de los tres prófugos se concentra en zona sur el papa francisco ofició la primera misa del año y alertó sobre la opresión isis planeab...

```

[36]: %%time
      # Segundo round
      def clean_text_round2(text):

```

```

'''Removemos otros caracteres adicionales'''
text = re.sub('[\'\"...«»]', '', text)
text = re.sub('\n', ' ', text)
return text

round2 = lambda x: clean_text_round2(x)

```

CPU times: user 33 µs, sys: 1e+03 ns, total: 34 µs
Wall time: 48.4 µs

```

[37]: %%time
# veamos como queda
data_pe_clean = pd.DataFrame(data_pe_clean.Texto.apply(round2))
data_pe_clean

```

CPU times: user 17.6 ms, sys: 34 µs, total: 17.7 ms
Wall time: 16.5 ms

```

[37]:

```

	Texto
2006	córdoba alianza delasotista se impone en marcos juarez el reclutamiento docente cae en picada relatos de una obsesión ya no podía entrar al aul...
2007	la historia se repite más de cien heridos con quemaduras de pirotecnia y lesiones en los ojos el primer bebé del año nació en buenos aires quemar...
2008	atentados y operativos de seguridad oscurecen festejos el primer bebé porteño del año nació en parque patricios un motociclista perdió la vida al ...
2009	balas perdidas corchos riñas callejeras e incidentes con pirotecnia en los festejos de fin de año cromañon en bangkok muertos en una discoteca...
2010	peajes bien chévere los placeres mínimos de cada día para estar saludables mateo ramos se postula para ser el primer bebé de voraz incendio en be...
2011	toman otro predio en villa lugano durante los festejos de año nuevo hoy asume dilma rousseff en brasil en santiago del estero el arrancó con un s...
2012	murió el gobernador de río negro quién asume el gobierno en río negro primeros detalles y repercusiones de la confusa muerte de soria carlos soria...
2013	pablo echarri la militancia me llegó tarde a los infografía feriados encuentran muerto al andinista perdido en el aconcagua boudou pasó el...
2014	no creo que haya amor creo que hay propiedad privada y no me la toca nadie la verdad mexicana no es un lujo es algo que necesito la conferenci...
2015	caso lola aseguran que es extremadamente complicado choque y vuelco fatal en camino del buen ayre boudou el embajador en la reasunción de di...
2016	la búsqueda de los tres prófugos se concentra en zona sur el papa francisco ofició la primera misa del año y alertó sobre la opresión isis planeab...

```

[38]: %%time
# Como no tenemos un Lemmatizer en español, hacemos manualmente algunas
↪ conversiones

```

```
# OJO: esto realmente no se hace a mano!!!

def detectadas(palabra):
    eliminar_s =
    →('libreros', 'textos', 'papelitos', 'monedas', 'páginas', 'anécdotas', 'perros', 'cuadernos', 'blog
        )
    →('revistas', 'caballos', 'vecinos', 'madres', 'puntos', 'ricos', 'libros')
    if palabra in eliminar_s :
        return palabra[:-1]
    eliminar_es =
    →('mundiales', 'lectores', 'campeones', 'maníes', 'ustedes', 'autores')
    if palabra in eliminar_es:
        return palabra[:-2]
    return palabra

def clean_text_round3(text):
    """
    return " ".join([detectadas(word) for word in text.split()])

round3 = lambda x: clean_text_round3(x)
```

CPU times: user 17 µs, sys: 0 ns, total: 17 µs

Wall time: 21.7 µs

```
[39]: %%time
#vemos como queda
data_pe_clean = pd.DataFrame(data_pe_clean.Texto.apply(round3))
data_pe_clean
```

CPU times: user 87 ms, sys: 53 µs, total: 87.1 ms

Wall time: 89.5 ms

```
[39]:
```

	Texto
2006	córdoba alianza delasotista se impone en marcos juarez el reclutamiento docente cae en picada relatos de una obsesión ya no podía entrar al aula e...
2007	la historia se repite más de cien heridos con quemaduras de pirotecnia y lesiones en los ojos el primer bebé del año nació en buenos aires quemaro...
2008	atentados y operativos de seguridad oscurecen festejos el primer bebé porteño del año nació en parque patricios un motociclista perdió la vida al ...
2009	balas perdidas corchos riñas callejeras e incidentes con pirotecnia en los festejos de fin de año cromañon en bangkok muertos en una discoteca mur...
2010	peajes bien chévere los placeres mínimos de cada día para estar saludables mateo ramos se postula para ser el primer bebé de voraz incendio en ber...
2011	toman otro predio en villa lugano durante los festejos de año nuevo hoy asume dilma rousseff en brasil en santiago del estero el arrancó con un si...
2012	murió el gobernador de río negro quién asume el gobierno en río negro primeros detalles y repercusiones de la confusa muerte de soria carlos soria...

2013 pablo echarri la militancia me llegó tarde a los infografía feriado
 encuentran muerto al andinista perdido en el aconcagua boudou pasó el año nue...
 2014 no creo que haya amor creo que hay propiedad privada y no me la toca nadie
 la verdad mexicana no es un lujo es algo que necesito la conferencia de...
 2015 caso lola aseguran que es extremadamente complicado choque y vuelco fatal
 en camino del buen ayre boudou el embajador en la reasunción de Dilma lo...
 2016 la búsqueda de los tres prófugos se concentra en zona sur el papa
 francisco ofició la primera misa del año y alertó sobre la opresión isis
 planeab...

```
[40]: %%time
# Esto es un nuevo campo por si quisieramos agregar alguna info adicional a
→ cada año
# Nuestro caso repetimos los años, nos servirá para alguna visualización
full_names =
→ ['2006', '2007', '2008', '2009', '2010', '2011', '2012', '2013', '2014', '2015', '2016']

data_pe['full_name'] = full_names
data_pe
```

CPU times: user 3.25 ms, sys: 72 µs, total: 3.32 ms

Wall time: 2.52 ms

```
[40]:
```

Texto \

2006 córdoba alianza del asotista se impone en marcos juarez el reclutamiento
 docente cae en picada relatos de una obsesión ya no podía entrar al aul...
 2007 la historia se repite más de cien heridos con quemaduras de pirotecnia y
 lesiones en los ojos el primer bebé del año nació en buenos aires quemar...
 2008 atentados y operativos de seguridad oscurecen festejos el primer bebé
 porteño del año nació en parque patricios un motociclista perdió la vida al ...
 2009 balas perdidas corchos riñas callejeras e incidentes con pirotecnia en
 los festejos de fin de año cromañon en bangkok muertos en una discoteca...
 2010 peajes bien chévere los placeres mínimos de cada día para estar saludables
 mateo ramos se postula para ser el primer bebé de voraz incendio en be...
 2011 toman otro predio en villa lugano durante los festejos de año nuevo hoy
 asume Dilma Rousseff en Brasil en Santiago del Estero el arrancó con un s...
 2012 murió el gobernador de Río Negro quién asume el gobierno en Río Negro
 primeros detalles y repercusiones de la confusa muerte de Soria Carlos Soria...
 2013 pablo echarri la militancia me llegó tarde a los infografía
 feriado encuentran muerto al andinista perdido en el aconcagua boudou pasó
 el...
 2014 no creo que haya amor creo que hay propiedad privada y no me la toca
 nadie la verdad mexicana no es un lujo es algo que necesito la conferenci...
 2015 caso lola aseguran que es extremadamente complicado choque y vuelco
 fatal en camino del buen ayre boudou el embajador en la reasunción de di...
 2016 la búsqueda de los tres prófugos se concentra en zona sur el papa
 Francisco ofició la primera misa del año y alertó sobre la opresión isis

planeab...

	full_name
2006	2006
2007	2007
2008	2008
2009	2009
2010	2010
2011	2011
2012	2012
2013	2013
2014	2014
2015	2015
2016	2016

```
[41]: %%time
# Hacemos el pickle para usar más adelante
data_pe.to_pickle("corpus_pe.pkl")
```

CPU times: user 9.97 ms, sys: 4.07 ms, total: 14 ms
Wall time: 11.4 ms

```
[42]: %%time
data_pe.Texto[0:255]
```

CPU times: user 1.91 ms, sys: 43 µs, total: 1.95 ms
Wall time: 1.59 ms

```
[42]: 2006    córdoba  alianza delasotista se impone en marcos juarez el reclutamiento
docente cae en picada relatos de una obsesión  ya no podía entrar al aul...
2007    la historia se repite  más de cien heridos con quemaduras de pirotecnia
y lesiones en los ojos el primer bebé del año nació en buenos aires quemar...
2008    atentados y operativos de seguridad oscurecen festejos el primer bebé
porteño del año nació en parque patricios un motociclista perdió la vida al ...
2009    balas perdidas  corchos  riñas callejeras e incidentes con pirotecnia en
los festejos de fin de año cromañon en bangkok  muertos en una discoteca...
2010    peajes bien chévere los placeres mínimos de cada día para estar
saludables mateo ramos se postula para ser el primer bebé de  voraz incendio en
be...
2011    toman otro predio en villa lugano durante los festejos de año nuevo hoy
asume dilma rousseff en brasil en santiago del estero el  arrancó con un s...
2012    murió el gobernador de río negro quién asume el gobierno en río negro
primeros detalles y repercusiones de la confusa muerte de soria carlos soria...
2013    pablo echarri  la militancia me llegó tarde  a los  infografía
feriados  encuentran muerto al andinista perdido en el aconcagua boudou pasó
el...
2014    no creo que haya amor  creo que hay propiedad privada y no me la toca
```

nadie la verdad mexicana no es un lujo es algo que necesito la conferenci...
 2015 caso lola aseguran que es extremadamente complicado choque y vuelco
 fatal en camino del buen ayre boudou el embajador en la reasunción de di...
 2016 la búsqueda de los tres prófugos se concentra en zona sur el papa
 francisco ofició la primera misa del año y alertó sobre la opresión isis
 planeab...
 Name: Texto, dtype: object

```
[43]: %%time
# Creamos una matriz de palabras y quitamos las palabras mas comunes del
↳ castellano
from sklearn.feature_extraction.text import CountVectorizer

# En spanish.txt tenemos una lista de palabras comunes
with open('spanish.txt') as f:
    lines = f.read().splitlines()

cv = CountVectorizer(stop_words=lines)
data_pe_cv = cv.fit_transform(data_pe_clean.Texto)
data_pe_dtm = pd.DataFrame(data_pe_cv.toarray(), columns=cv.get_feature_names())
data_pe_dtm.index = data_pe_clean.index
data_pe_dtm
```

CPU times: user 132 ms, sys: 7.31 ms, total: 139 ms
 Wall time: 138 ms

```
[43]:
```

	aaa	abajo	abal	aballay	abanadonado	abanderada	abandona	\
2006	0	0	0	0	0	0	0	
2007	1	0	0	0	0	0	1	
2008	0	1	0	0	0	0	0	
2009	0	1	0	0	0	0	0	
2010	0	0	0	0	0	0	2	
2011	0	0	4	0	0	0	0	
2012	0	0	2	1	0	1	1	
2013	0	1	2	0	0	0	1	
2014	0	1	0	0	0	0	0	
2015	0	1	1	0	1	0	0	
2016	0	0	0	0	0	0	0	

	abandonada	abandonadas	abandonado	...	últimas	último	últimos	\
2006	1	0	0	...	1	4	1	
2007	0	0	0	...	0	1	0	
2008	0	0	0	...	0	5	2	
2009	0	0	0	...	1	3	0	
2010	0	0	0	...	1	5	3	
2011	0	0	1	...	1	2	2	
2012	0	1	0	...	3	3	1	

2013	0	0	0 ...	1	1	2
2014	0	0	1 ...	0	1	0
2015	0	0	0 ...	0	6	1
2016	0	0	0 ...	2	3	1

	única	únicamente	únicas	único	úselo	útero	útil
2006	0	0	1	3	0	0	0
2007	1	0	0	1	0	1	0
2008	2	0	0	2	1	0	0
2009	2	0	0	0	0	0	0
2010	1	0	1	2	0	0	0
2011	1	1	0	2	0	0	1
2012	1	0	0	2	0	0	0
2013	0	0	0	2	0	0	0
2014	3	0	0	1	0	0	0
2015	2	0	0	0	0	0	0
2016	0	0	0	1	0	0	0

[11 rows x 15702 columns]

```
[44]: %%time
# Lo guardamos como pickle
data_pe_dtm.to_pickle("dtm_pe.pkl")
```

CPU times: user 13.4 ms, sys: 4.15 ms, total: 17.5 ms
Wall time: 38.2 ms

```
[45]: %%time
# Lo guardamos como pickle también
data_pe_clean.to_pickle('data_pe_clean.pkl')
pickle.dump(cv, open("cv_pe.pkl", "wb"))
```

CPU times: user 60.6 ms, sys: 16 ms, total: 76.5 ms
Wall time: 81.9 ms

5 4 - Análisis Exploratorio

```
[46]: %%time
# Vemos la matriz de palabras
import pandas as pd

data_pe = pd.read_pickle('dtm_pe.pkl')
data_pe = data_pe.transpose()
data_pe.head()
```

CPU times: user 23.5 ms, sys: 78 µs, total: 23.6 ms
Wall time: 21.1 ms

```
[46]:
```

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
aaa	0	1	0	0	0	0	0	0	0	0	0
abajo	0	0	1	1	0	0	0	1	1	1	0
abal	0	0	0	0	0	4	2	2	0	1	0
aballay	0	0	0	0	0	0	1	0	0	0	0
abanadonado	0	0	0	0	0	0	0	0	0	1	0

```
[47]: %%time
# Buscamos las palabras mas comunes por año
top_dict = {}
for c in data_pe.columns:
    top = data_pe[c].sort_values(ascending=False).head(30)
    top_dict[c]= list(zip(top.index, top.values))

top_dict
```

CPU times: user 29.2 ms, sys: 22 µs, total: 29.2 ms
Wall time: 28.1 ms

```
[47]: {2006: [('kirchner', 44),
('argentina', 42),
('gobierno', 29),
('quiere', 24),
('dos', 22),
('ee', 18),
('uu', 18),
('tres', 17),
('años', 17),
('va', 17),
('ser', 16),
('fernández', 16),
('presidente', 16),
('san', 16),
('ex', 15),
('boca', 15),
('chávez', 14),
('país', 14),
('sigue', 14),
('nuevo', 14),
('final', 14),
('ahora', 13),
('lula', 13),
('ciento', 13),
('river', 12),
('crisis', 12),
```

('triunfo', 12),
 ('busca', 12),
 ('si', 11),
 ('mujer', 11)],
 2007: [('muertos', 33),
 ('ee', 29),
 ('uu', 29),
 ('chávez', 28),
 ('kirchner', 28),
 ('gobierno', 28),
 ('gerez', 26),
 ('ex', 22),
 ('año', 21),
 ('gran', 18),
 ('dos', 18),
 ('argentina', 18),
 ('uruguay', 18),
 ('isabel', 17),
 ('bush', 17),
 ('caso', 16),
 ('plata', 16),
 ('murió', 15),
 ('años', 15),
 ('saddam', 15),
 ('mujer', 15),
 ('irak', 15),
 ('madrid', 15),
 ('policía', 14),
 ('ahora', 14),
 ('san', 13),
 ('gago', 13),
 ('españa', 13),
 ('muerte', 13),
 ('triple', 13)],
 2008: [('gobierno', 34),
 ('cristina', 33),
 ('macri', 30),
 ('dos', 26),
 ('años', 22),
 ('chávez', 21),
 ('boca', 21),
 ('muertos', 18),
 ('farc', 18),
 ('caso', 17),
 ('kirchner', 17),
 ('tres', 17),
 ('argentina', 16),

```

('san', 15),
('ex', 14),
('emmanuel', 14),
('año', 14),
('river', 14),
('menos', 14),
('primera', 13),
('bush', 13),
('quiere', 13),
('murió', 13),
('vez', 12),
('familia', 12),
('intervención', 12),
('nueva', 12),
('rehenes', 11),
('nuevo', 11),
('sigue', 11)],
2009: [('cristina', 62),
('obama', 56),
('kirchner', 40),
('gobierno', 37),
('gaza', 35),
('dakar', 29),
('campo', 28),
('bergara', 27),
('israel', 26),
('nuevo', 23),
('moreno', 17),
('casa', 17),
('argentino', 16),
('cómo', 16),
('dice', 16),
('chávez', 16),
('argentina', 16),
('scioli', 16),
('ser', 16),
('años', 15),
('boca', 15),
('crisis', 14),
('israelí', 14),
('ruta', 14),
('néstor', 14),
('maradona', 14),
('dos', 14),
('castro', 14),
('día', 13),
('muertos', 13)],

```

2010: [('redrado', 91),
 ('cristina', 73),
 ('kirchner', 72),
 ('gobierno', 47),
 ('cobos', 39),
 ('central', 33),
 ('sandro', 31),
 ('anibal', 27),
 ('banco', 21),
 ('murió', 21),
 ('macri', 20),
 ('justicia', 20),
 ('haití', 20),
 ('chávez', 19),
 ('oposición', 19),
 ('cómo', 19),
 ('boca', 18),
 ('hiena', 17),
 ('argentina', 17),
 ('dólares', 17),
 ('dakar', 17),
 ('reservas', 17),
 ('boudou', 16),
 ('mujer', 16),
 ('bcra', 15),
 ('quiere', 15),
 ('si', 15),
 ('fernández', 15),
 ('millones', 15),
 ('muertos', 14)],
 2011: [('cristina', 49),
 ('cfk', 37),
 ('gobierno', 29),
 ('kirchner', 28),
 ('argentina', 25),
 ('nuevo', 24),
 ('macri', 23),
 ('anibal', 23),
 ('murió', 22),
 ('brasil', 22),
 ('egipto', 22),
 ('muerte', 21),
 ('punta', 18),
 ('mubarak', 18),
 ('dos', 18),
 ('plata', 18),
 ('año', 17),

```

('si', 17),
('juliá', 17),
('dakar', 17),
('cocaína', 17),
('mar', 16),
('casa', 16),
('ser', 15),
('candidato', 15),
('años', 15),
('scioli', 14),
('cobos', 14),
('muertos', 14),
('venegas', 14)],
2012: [('gobierno', 40),
('cristina', 39),
('soria', 37),
('cfk', 29),
('cómo', 28),
('años', 25),
('argentina', 25),
('malvinas', 23),
('moyano', 22),
('murió', 21),
('boudou', 21),
('macri', 21),
('maestra', 20),
('scioli', 19),
('muerte', 18),
('dos', 17),
('subte', 17),
('millones', 17),
('ex', 17),
('nuevo', 15),
('intendente', 15),
('crucero', 15),
('tras', 15),
('moreno', 14),
('kirchner', 14),
('susana', 14),
('oyarbide', 14),
('juez', 13),
('carlos', 13),
('operación', 13)],
2013: [('cristina', 55),
('cfk', 50),
('gobierno', 41),
('cómo', 31),

```


('scioli', 27),
 ('video', 23),
 ('macri', 23),
 ('argentina', 22),
 ('tras', 20),
 ('años', 20),
 ('intendente', 20),
 ('chávez', 20),
 ('once', 20),
 ('fragata', 19),
 ('darín', 19),
 ('tragedia', 18),
 ('boudou', 17),
 ('dólar', 17),
 ('mar', 16),
 ('murió', 15),
 ('dos', 15),
 ('punta', 15),
 ('subte', 14),
 ('irán', 13),
 ('brasil', 13),
 ('plata', 13),
 ('foto', 12),
 ('buenos', 12),
 ('sota', 12),
 ('muertos', 12)],
 2014: [('cristina', 54),
 ('gobierno', 45),
 ('argentina', 28),
 ('nuevo', 26),
 ('capitanich', 25),
 ('echegaray', 24),
 ('dos', 22),
 ('scioli', 22),
 ('cfk', 22),
 ('dólar', 19),
 ('papa', 18),
 ('brasil', 17),
 ('macri', 17),
 ('gesell', 17),
 ('dólares', 17),
 ('villa', 17),
 ('cómo', 17),
 ('luz', 16),
 ('mar', 16),
 ('caso', 15),
 ('rayo', 15),

```

('boudou', 15),
('tras', 15),
('massa', 15),
('plata', 14),
('murió', 14),
('tres', 13),
('mil', 13),
('años', 12),
('quiere', 12)],
2015: [('nisman', 223),
('muerte', 74),
('cristina', 62),
('lola', 54),
('caso', 49),
('fiscal', 44),
('cfk', 38),
('amia', 35),
('denuncia', 32),
('charlie', 30),
('gobierno', 30),
('hebdo', 29),
('scioli', 28),
('argentina', 27),
('cómo', 24),
('lagomarsino', 22),
('justicia', 21),
('parís', 21),
('dos', 19),
('gils', 19),
('carbó', 19),
('atentado', 18),
('timerman', 18),
('francia', 17),
('nuevo', 17),
('tras', 17),
('si', 16),
('país', 15),
('clarín', 15),
('denuncian', 14)],
2016: [('macri', 95),
('lanatta', 55),
('prófugos', 53),
('fuga', 40),
('sala', 35),
('gobierno', 34),
('victor', 33),
('nisman', 32),

```

```
(('triple', 30),
 ('milagro', 29),
 ('hugo', 27),
 ('seguridad', 24),
 ('santa', 22),
 ('dos', 22),
 ('video', 21),
 ('davos', 20),
 ('corradi', 20),
 ('captura', 20),
 ('morales', 20),
 ('schillaci', 20),
 ('año', 19),
 ('pérez', 19),
 ('fe', 18),
 ('tres', 18),
 ('martín', 18),
 ('carrió', 17),
 ('tras', 17),
 ('aníbal', 17),
 ('nuevo', 16),
 ('bullrich', 16)])}
```

```
[48]: %%time
# Las mostramos un poco mas ordenadas
for anio, top_words in top_dict.items():
    print(anio)
    print(', '.join([word for word, count in top_words[0:14]]))
    print('---')
```

2006

kirchner, argentina, gobierno, quiere, dos, ee, uu, tres, años, va, ser,
fernández, presidente, san

2007

muertos, ee, uu, Chávez, kirchner, gobierno, gerez, ex, año, gran, dos,
argentina, uruguay, isabel

2008

gobierno, cristina, macri, dos, años, Chávez, boca, muertos, farc, caso,
kirchner, tres, argentina, san

2009

cristina, obama, kirchner, gobierno, gaza, dakar, campo, bergara, israel, nuevo,
moreno, casa, argentino, cómo

2010

```

redrado, cristina, kirchner, gobierno, cobos, central, sandro, aníbal, banco,
murió, macri, justicia, haití, chávez
---
2011
cristina, cfk, gobierno, kirchner, argentina, nuevo, macri, aníbal, murió,
brasil, egipto, muerte, punta, mubarak
---
2012
gobierno, cristina, soria, cfk, cómo, años, argentina, malvinas, moyano, murió,
boudou, macri, maestra, scioli
---
2013
cristina, cfk, gobierno, cómo, scioli, video, macri, argentina, tras, años,
intendente, chávez, once, fragata
---
2014
cristina, gobierno, argentina, nuevo, capitanich, echegaray, dos, scioli, cfk,
dólar, papa, brasil, macri, gesell
---
2015
nisman, muerte, cristina, lola, caso, fiscal, cfk, amia, denuncia, charlie,
gobierno, hebdo, scioli, argentina
---
2016
macri, lanatta, prófugos, fuga, sala, gobierno, victor, nisman, triple, milagro,
hugo, seguridad, santa, dos
---
CPU times: user 0 ns, sys: 3.35 ms, total: 3.35 ms
Wall time: 2.22 ms

```

```

[49]: %%time
      # Agregamos esas palabras a la lista de stopwords
      from collections import Counter

      words = []
      for anio in data_pe.columns:
          top = [word for (word, count) in top_dict[anio]]
          for t in top:
              words.append(t)

      words

```

```

CPU times: user 696 µs, sys: 16 µs, total: 712 µs
Wall time: 723 µs

```

```

[49]: ['kirchner',
      'argentina',

```

'gobierno',
'quiere',
'dos',
'ee',
'uu',
'tres',
'años',
'va',
'ser',
'fernández',
'presidente',
'san',
'ex',
'boca',
'chávez',
'país',
'sigue',
'nuevo',
'final',
'ahora',
'lula',
'ciento',
'river',
'crisis',
'triunfo',
'busca',
'si',
'mujer',
'muertos',
'ee',
'uu',
'chávez',
'kirchner',
'gobierno',
'gerez',
'ex',
'año',
'gran',
'dos',
'argentina',
'uruguay',
'isabel',
'bush',
'caso',
'plata',
'murió',
'años',

'saddam',
'mujer',
'irak',
'madrid',
'policía',
'ahora',
'san',
'gago',
'españa',
'muerte',
'triple',
'gobierno',
'cristina',
'macri',
'dos',
'años',
'chávez',
'boca',
'muertos',
'farc',
'caso',
'kirchner',
'tres',
'argentina',
'san',
'ex',
'emmanuel',
'año',
'river',
'menos',
'primera',
'bush',
'quiere',
'murió',
'vez',
'familia',
'intervención',
'nueva',
'rehenes',
'nuevo',
'sigue',
'cristina',
'obama',
'kirchner',
'gobierno',
'gaza',
'dakar',

'campo',
'bergara',
'israel',
'nuevo',
'moreno',
'casa',
'argentino',
'cómo',
'dice',
'chávez',
'argentina',
'scioli',
'ser',
'años',
'boca',
'crisis',
'israelí',
'ruta',
'néstor',
'maradona',
'dos',
'castro',
'día',
'muertos',
'redrado',
'cristina',
'kirchner',
'gobierno',
'cobos',
'central',
'sandro',
'aníbal',
'banco',
'murió',
'macri',
'justicia',
'haití',
'chávez',
'oposición',
'cómo',
'boca',
'hiena',
'argentina',
'dólares',
'dakar',
'reservas',
'boudou',

'mujer',
'bcra',
'quiere',
'si',
'fernández',
'millones',
'muertos',
'cristina',
'cfk',
'gobierno',
'kirchner',
'argentina',
'nuevo',
'macri',
'aníbal',
'murió',
'brasil',
'egipto',
'muerte',
'punta',
'mubarak',
'dos',
'plata',
'año',
'si',
'juliá',
'dakar',
'cocaína',
'mar',
'casa',
'ser',
'candidato',
'años',
'scioli',
'cobos',
'muertos',
'venegas',
'gobierno',
'cristina',
'soria',
'cfk',
'cómo',
'años',
'argentina',
'malvinas',
'moyano',
'murió',

'boudou',
'macri',
'maestra',
'scioli',
'muerte',
'dos',
'subte',
'millones',
'ex',
'nuevo',
'intendente',
'crucero',
'tras',
'moreno',
'kirchner',
'susana',
'oyarbide',
'juez',
'carlos',
'operación',
'cristina',
'cfk',
'gobierno',
'cómo',
'scioli',
'video',
'macri',
'argentina',
'tras',
'años',
'intendente',
'chávez',
'once',
'fragata',
'darín',
'tragedia',
'boudou',
'dólar',
'mar',
'murió',
'dos',
'punta',
'subte',
'irán',
'brasil',
'plata',
'foto',

'buenos',
'sota',
'muertos',
'cristina',
'gobierno',
'argentina',
'nuevo',
'capitanich',
'echegaray',
'dos',
'scioli',
'cfk',
'dólar',
'papa',
'brasil',
'macri',
'gesell',
'dólares',
'villa',
'cómo',
'luz',
'mar',
'caso',
'rayo',
'boudou',
'tras',
'massa',
'plata',
'murió',
'tres',
'mil',
'años',
'quiere',
'nisman',
'muerte',
'cristina',
'lola',
'caso',
'fiscal',
'cfk',
'amia',
'denuncia',
'charlie',
'gobierno',
'hebdo',
'scioli',
'argentina',

'cómo',
'lagomarsino',
'justicia',
'parís',
'dos',
'gils',
'carbó',
'atentado',
'timerman',
'francia',
'nuevo',
'tras',
'si',
'país',
'clarín',
'denuncian',
'macri',
'lanatta',
'prófugos',
'fuga',
'sala',
'gobierno',
'victor',
'nisman',
'triple',
'milagro',
'hugo',
'seguridad',
'santa',
'dos',
'video',
'davos',
'corradi',
'captura',
'morales',
'schillaci',
'año',
'pérez',
'fe',
'tres',
'martín',
'carrió',
'tras',
'aníbal',
'nuevo',
'bullrich']

```
[50]: Counter(words).most_common()
```

```
[50]: [('gobierno', 11),
      ('argentina', 10),
      ('dos', 10),
      ('años', 8),
      ('nuevo', 8),
      ('cristina', 8),
      ('kirchner', 7),
      ('murió', 7),
      ('macri', 7),
      ('chávez', 6),
      ('muertos', 6),
      ('cómo', 6),
      ('scioli', 6),
      ('cfk', 5),
      ('tras', 5),
      ('quiere', 4),
      ('tres', 4),
      ('ex', 4),
      ('boca', 4),
      ('si', 4),
      ('año', 4),
      ('caso', 4),
      ('plata', 4),
      ('muerte', 4),
      ('boudou', 4),
      ('ser', 3),
      ('san', 3),
      ('mujer', 3),
      ('dakar', 3),
      ('aníbal', 3),
      ('brasil', 3),
      ('mar', 3),
      ('ee', 2),
      ('uu', 2),
      ('fernández', 2),
      ('país', 2),
      ('sigue', 2),
      ('ahora', 2),
      ('river', 2),
      ('crisis', 2),
      ('bush', 2),
      ('triple', 2),
      ('moreno', 2),
      ('casa', 2),
      ('cobos', 2),
```

('justicia', 2),
('dólares', 2),
('millones', 2),
('punta', 2),
('subte', 2),
('intendente', 2),
('video', 2),
('dólar', 2),
('nisman', 2),
('va', 1),
('presidente', 1),
('final', 1),
('lula', 1),
('ciento', 1),
('triunfo', 1),
('busca', 1),
('gerez', 1),
('gran', 1),
('uruguay', 1),
('isabel', 1),
('saddam', 1),
('irak', 1),
('madrid', 1),
('policía', 1),
('gago', 1),
('españa', 1),
('farc', 1),
('emmanuel', 1),
('menos', 1),
('primera', 1),
('vez', 1),
('familia', 1),
('intervención', 1),
('nueva', 1),
('rehenes', 1),
('obama', 1),
('gaza', 1),
('campo', 1),
('bergara', 1),
('israel', 1),
('argentino', 1),
('dice', 1),
('israelí', 1),
('ruta', 1),
('néstor', 1),
('maradona', 1),
('castro', 1),

('día', 1),
('redrado', 1),
('central', 1),
('sandro', 1),
('banco', 1),
('haití', 1),
('oposición', 1),
('hiena', 1),
('reservas', 1),
('bcra', 1),
('egipto', 1),
('mubarak', 1),
('juliá', 1),
('cocaína', 1),
('candidato', 1),
('venegas', 1),
('soria', 1),
('malvinas', 1),
('moyano', 1),
('maestra', 1),
('crucero', 1),
('susana', 1),
('oyarbide', 1),
('juez', 1),
('carlos', 1),
('operación', 1),
('once', 1),
('fragata', 1),
('darín', 1),
('tragedia', 1),
('irán', 1),
('foto', 1),
('buenos', 1),
('sota', 1),
('capitanich', 1),
('echegaray', 1),
('papa', 1),
('gesell', 1),
('villa', 1),
('luz', 1),
('rayo', 1),
('massa', 1),
('mil', 1),
('lola', 1),
('fiscal', 1),
('amia', 1),
('denuncia', 1),

```
( 'charlie', 1),
( 'hebdo', 1),
( 'lagomarsino', 1),
( 'parís', 1),
( 'gils', 1),
( 'carbó', 1),
( 'atentado', 1),
( 'timerman', 1),
( 'francia', 1),
( 'clarín', 1),
( 'denuncian', 1),
( 'lanatta', 1),
( 'prófugos', 1),
( 'fuga', 1),
( 'sala', 1),
( 'victor', 1),
( 'milagro', 1),
( 'hugo', 1),
( 'seguridad', 1),
( 'santa', 1),
( 'davos', 1),
( 'corradi', 1),
( 'captura', 1),
( 'morales', 1),
( 'schillaci', 1),
( 'pérez', 1),
( 'fe', 1),
( 'martín', 1),
( 'carrió', 1),
( 'bullrich', 1)]
```

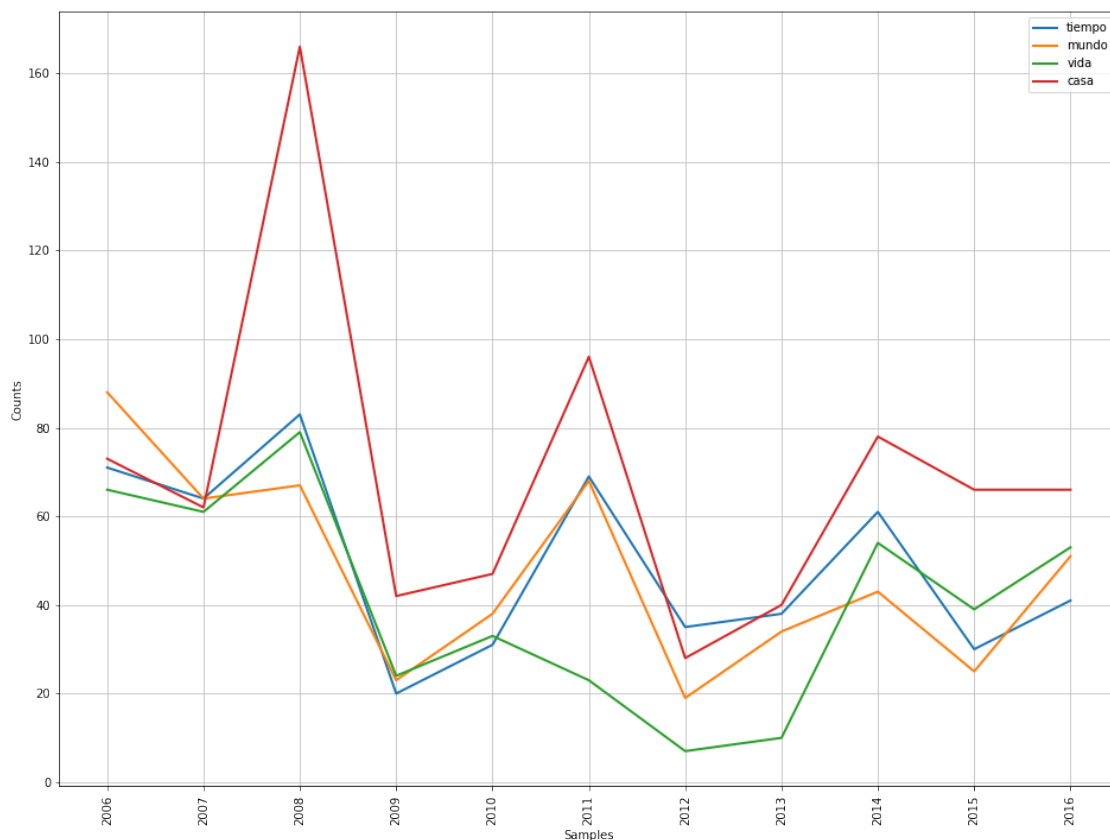
```
[51]: %%time
# Las mas repetidas las descartaremos
add_stop_words = [word for word, count in Counter(words).most_common() if count_
→> 6]
add_stop_words
```

CPU times: user 275 µs, sys: 0 ns, total: 275 µs
Wall time: 283 µs

```
[51]: ['gobierno',
      'argentina',
      'dos',
      'años',
      'nuevo',
      'cristina',
      'kirchner',
```

```
'murió',  
'macri']
```

```
[52]: %%time  
# A partir de unos textos genericos, creamos un corpus linguistico (ejemplos de uso)  
# usamos un blog  
anios = ['2006', '2007', '2008', '2009', '2010', '2011', '2012', '2013', '2014', '2015', '2016']  
import nltk  
from nltk.corpus import PlaintextCorpusReader  
corpus_root = '/home/administrador/git/python/tp_ds/blog/'  
wordlists = PlaintextCorpusReader(corpus_root, '.*', encoding='latin-1')  
  
cfd = nltk.ConditionalFreqDist(  
    (word, genre)  
    for genre in anios  
    for w in wordlists.words(genre + '.txt')  
    for word in ['casa', 'mundo', 'tiempo', 'vida']  
    if w.lower().startswith(word) )  
cfd.plot()
```



CPU times: user 816 ms, sys: 0 ns, total: 816 ms
Wall time: 813 ms

[52]: <matplotlib.axes._subplots.AxesSubplot at 0x7eff639fd3d0>

```
[53]: %%time
# Como no alcanzan las palabras del diccionario, debemos agregar mas stopwords
# extraidas manualmente del texto
from sklearn.feature_extraction import text
from sklearn.feature_extraction.text import CountVectorizer

# Leemos los textos limpios
data_pe_clean = pd.read_pickle('data_pe_clean.pkl')

# Agregamos algunas stopwords de nuestra cosecha
with open('spanish.txt') as f:
    stop_words = f.read().splitlines()
for pal in add_stop_words:
    stop_words.append(pal)
more_stop_words=[ 'alex', 'andr s', 'asi', 'andres', 'as ', 'a o', 'alejandro', 'alfonso', 'all ', 'algu
                'basdala', 'bernardo', 'bien',
                 
      'cosa', 'cosas', 'costoya', 'costa', 'cinco', 'celoni', 'cuatro', '  mo', 'casi', 'colo', 'caprio', '  
                'dos', 'dice', 'decir', 'd as', 'dije', 'digo', 'diez',
                ' sa', ' sas', ' se', ' sos', ' sta', ' stas', ' ste', ' sto',  
      ' stos',
                'fernando', 'fenwick',
                'gel s', 'gente',
                'hornby', 'hoy', 'horacio', 'horas', 'har ', 'hans', 'hac a', 'haber',
                'iveta',
                'jes s', 'jorge', 'juan',
                'karen',
                'lucas', 'luego', 'luis',
                 
      'mirta', 'mientras', 'menos', 'm nica', 'medio', 'mil', 'moncho', 'momento', 'ma ana', 'mejor',
                'narc s', 'n mero', 'noche', 'nadie',
                'ojos',
                'primer', 'primera', 'pase', 'pablo', 'pepe', 'pack', 'peter',  
      'pues', 'prieto', 'politto', 'pol', 'paola', 'puede', 'pr ximo', 'podr n', 'pod a',
                'quiz ', 'quiz s', 'quince', 'qu en', 'quiero',
                'rato',
                 
      's lo', 'solamente', 'sakhan', ' eredova', 'seis', ' eredov ', 'seselovsky', 'solo', 'salas', 'sant'
                'tres', 'tan', 'todas', 'trece', 'toda', 'todav a', 'tarde', 'tener',
```

```

        'uno', 'usted',
        'veces', 'ver', 've', 'vos', 'va', 'voy',
        'waiser', 'woung'
    ]
for pal in more_stop_words:
    stop_words.append(pal)

cv = CountVectorizer(stop_words=stop_words)
data_pe_cv = cv.fit_transform(data_pe_clean.Texto)
data_pe_stop = pd.DataFrame(data_pe_cv.toarray(), columns=cv.
    ↳get_feature_names())
data_pe_stop.index = data_pe_clean.index

# guardamos por las dudas
import pickle
pickle.dump(cv, open("cv_pe_stop.pkl", "wb"))
data_pe_stop.to_pickle("dtm_pe_stop.pkl")

```

CPU times: user 188 ms, sys: 0 ns, total: 188 ms
 Wall time: 188 ms

```

[54]: %%time
# Terminal / Anaconda Prompt: conda install -c conda-forge wordcloud
from wordcloud import WordCloud

wc = WordCloud(stopwords=stop_words, background_color="white", colormap="Dark2",
    max_font_size=150, random_state=42)

```

CPU times: user 693 µs, sys: 16 µs, total: 709 µs
 Wall time: 512 µs

```

[55]: %%time
# Dibujamos las nubes de palabras
import matplotlib.pyplot as plt

plt.rcParams['figure.figsize'] = [16,12]

#anios =
↳ ['2004', '2005', '2006', '2007', '2008', '2009', '2010', '2011', '2012', '2013', '2014', '2015']

# Create subplots for each anio
for index, anio in enumerate(data_pe.columns):
    wc.generate(data_pe_clean.Texto[anio])
    plt.subplot(4, 3, index+1)
    plt.imshow(wc, interpolation="bilinear")
    plt.axis("off")

```

```
plt.show()
```



6 5 - Análisis de Sentimiento

```
[56]: %%time
      # Leeremos el corpus que aún preserva el orden de las palabras
      import pandas as pd

      data_pe = pd.read_pickle('corpus_pe.pkl')
      data_pe
```

51

[56]:

Texto \

2006 córdoba alianza delasotista se impone en marcos juarez el reclutamiento docente cae en picada relatos de una obsesión ya no podía entrar al aul...
2007 la historia se repite más de cien heridos con quemaduras de pirotecnia y lesiones en los ojos el primer bebé del año nació en buenos aires quemar...
2008 atentados y operativos de seguridad oscurecen festejos el primer bebé porteño del año nació en parque patricios un motociclista perdió la vida al ...
2009 balas perdidas corchos riñas callejeras e incidentes con pirotecnia en los festejos de fin de año cromañon en bangkok muertos en una discoteca...
2010 peajes bien chévere los placeres mínimos de cada día para estar saludables mateo ramos se postula para ser el primer bebé de voraz incendio en be...
2011 toman otro predio en villa lugano durante los festejos de año nuevo hoy asume dilma rousseff en brasil en santiago del estero el arrancó con un s...
2012 murió el gobernador de río negro quién asume el gobierno en río negro primeros detalles y repercusiones de la confusa muerte de soria carlos soria...
2013 pablo echarri la militancia me llegó tarde a los infografía feriados encuentran muerto al andinista perdido en el aconcagua boudou pasó el...
2014 no creo que haya amor creo que hay propiedad privada y no me la toca nadie la verdad mexicana no es un lujo es algo que necesito la conferenci...
2015 caso lola aseguran que es extremadamente complicado choque y vuelco fatal en camino del buen ayre boudou el embajador en la reasunción de di...
2016 la búsqueda de los tres prófugos se concentra en zona sur el papa francisco ofició la primera misa del año y alertó sobre la opresión isis planeab...

	full_name
2006	2006
2007	2007
2008	2008
2009	2009
2010	2010
2011	2011
2012	2012
2013	2013
2014	2014
2015	2015
2016	2016

[57]:

```
%%time
# Creamos funciones auxiliares para procesar el texto
# Terminal / Anaconda Navigator: conda install -c conda-forge textblob
from textblob import TextBlob

pol = lambda x: TextBlob(x).sentiment.polarity
pol2 = lambda x: x.sentiment.polarity
sub = lambda x: TextBlob(x).sentiment.subjectivity
```

```

sub2 = lambda x: x.sentiment.subjectivity

# Realmente lo traducimos al inglés pues el analisis de sentimiento de TextBlob
↳no funciona en Español :(
traducir = lambda x: TextBlob(x).translate(to="en")

data_pe['blob_en'] = data_pe['Texto'].apply(traducir)
data_pe['polarity'] = data_pe['blob_en'].apply(pol2)
data_pe['subjectivity'] = data_pe['blob_en'].apply(sub2)
data_pe

```

CPU times: user 2.06 s, sys: 12.2 ms, total: 2.07 s
Wall time: 15.3 s

[57]:

	Texto \
2006	córdoba alianza delasotista se impone en marcos juarez el reclutamiento docente cae en picada relatos de una obsesión ya no podía entrar al aul...
2007	la historia se repite más de cien heridos con quemaduras de pirotecnia y lesiones en los ojos el primer bebé del año nació en buenos aires quemar...
2008	atentados y operativos de seguridad oscurecen festejos el primer bebé porteño del año nació en parque patricios un motociclista perdió la vida al ...
2009	balas perdidas corchos riñas callejeras e incidentes con pirotecnia en los festejos de fin de año cromañon en bangkok muertos en una discoteca...
2010	peajes bien chévere los placeres mínimos de cada día para estar saludables mateo ramos se postula para ser el primer bebé de voraz incendio en be...
2011	toman otro predio en villa lugano durante los festejos de año nuevo hoy asume dilma rousseff en brasil en santiago del estero el arrancó con un s...
2012	murió el gobernador de río negro quién asume el gobierno en río negro primeros detalles y repercusiones de la confusa muerte de soria carlos soria...
2013	pablo echarri la militancia me llegó tarde a los infografía feriados encuentran muerto al andinista perdido en el aconcagua boudou pasó el...
2014	no creo que haya amor creo que hay propiedad privada y no me la toca nadie la verdad mexicana no es un lujo es algo que necesito la conferenci...
2015	caso lola aseguran que es extremadamente complicado choque y vuelco fatal en camino del buen ayre boudou el embajador en la reasunción de di...
2016	la búsqueda de los tres prófugos se concentra en zona sur el papa francisco ofició la primera misa del año y alertó sobre la opresión isis planeab...

	full_name \
2006	2006
2007	2007
2008	2008
2009	2009
2010	2010
2011	2011

2012	2012
2013	2013
2014	2014
2015	2015
2016	2016

```

blob_en \
2006 (c, o, r, d, o, b, a, , d, e, l, a, s, o, t, i, s, t, a, , a, l, l, i,
a, n, c, e, , i, m, p, o, s, e, s, , i, t, s, e, l, f, , o, n, , m, a...
2007 (h, i, s, t, o, r, y, , r, e, p, e, a, t, s, , m, o, r, e, , t, h, a,
n, , o, n, e, , h, u, n, d, r, e, d, , w, o, u, n, d, e, d, , w, i, t...
2008 (a, t, t, a, c, k, s, , a, n, d, , s, e, c, u, r, i, t, y, , o, p, e,
r, a, t, i, o, n, s, , o, b, s, c, u, r, e, , c, e, l, e, b, r, a, t, i...
2009 (l, o, s, t, , b, u, l, l, e, t, s, , c, o, r, k, s, , s, t, r, e, e,
t, , q, u, a, r, r, e, l, s, , a, n, d, , i, n, c, i, d, e, n, t, s, ...
2010 (t, o, l, l, s, , w, e, l, l, , c, o, o, l, , t, h, e, , m, i, n, i,
m, u, m, , p, l, e, a, s, u, r, e, s, , o, f, , e, a, c, h, , d, a, y...
2011 (t, h, e, y, , t, a, k, e, , a, n, o, t, h, e, r, , p, r, o, p, e, r,
t, y, , i, n, , V, i, l, l, a, , L, u, g, a, n, o, , d, u, r, i, n, g...
2012 (T, h, e, , g, o, v, e, r, n, o, r, , o, f, , R, í, o, , N, e, g, r,
o, , d, i, e, d, , , w, h, o, , a, s, s, u, m, e, s, , t, h, e, , g...
2013 (p, a, u, l, , e, c, h, a, r, i, , t, h, e, , m, i, l, i, t, a, n, c,
y, , I, , w, a, s, , l, a, t, e, , t, o, , t, h, e, , h, o, l, i, d...
2014 (I, , d, o, , n, o, t, , t, h, i, n, k, , t, h, e, r, e, , i, s, ,
l, o, v, e, , I, , t, h, i, n, k, , t, h, e, r, e, , i, s, , p, r, i...
2015 (l, o, l, a, , c, a, s, e, , a, s, s, u, r, e, , t, h, a, t, , i, t,
, i, s, , e, x, t, r, e, m, e, l, y, , c, o, m, p, l, i, c, a, t, e, d...
2016 (T, h, e, , s, e, a, r, c, h, , f, o, r, , t, h, e, , t, h, r, e, e,
, f, u, g, i, t, i, v, e, s, , i, s, , c, o, n, c, e, n, t, r, a, t, e...

```

	polarity	subjectivity
2006	0.096154	0.441835
2007	0.064330	0.408982
2008	0.064545	0.416332
2009	0.063941	0.408875
2010	0.067944	0.413715
2011	0.063373	0.415941
2012	0.057627	0.399632
2013	0.111451	0.419632
2014	0.047445	0.438419
2015	0.039250	0.408422
2016	0.014164	0.393963

```

[60]: %%time
# Dibujemos los resultados
import matplotlib.pyplot as plt

```

```

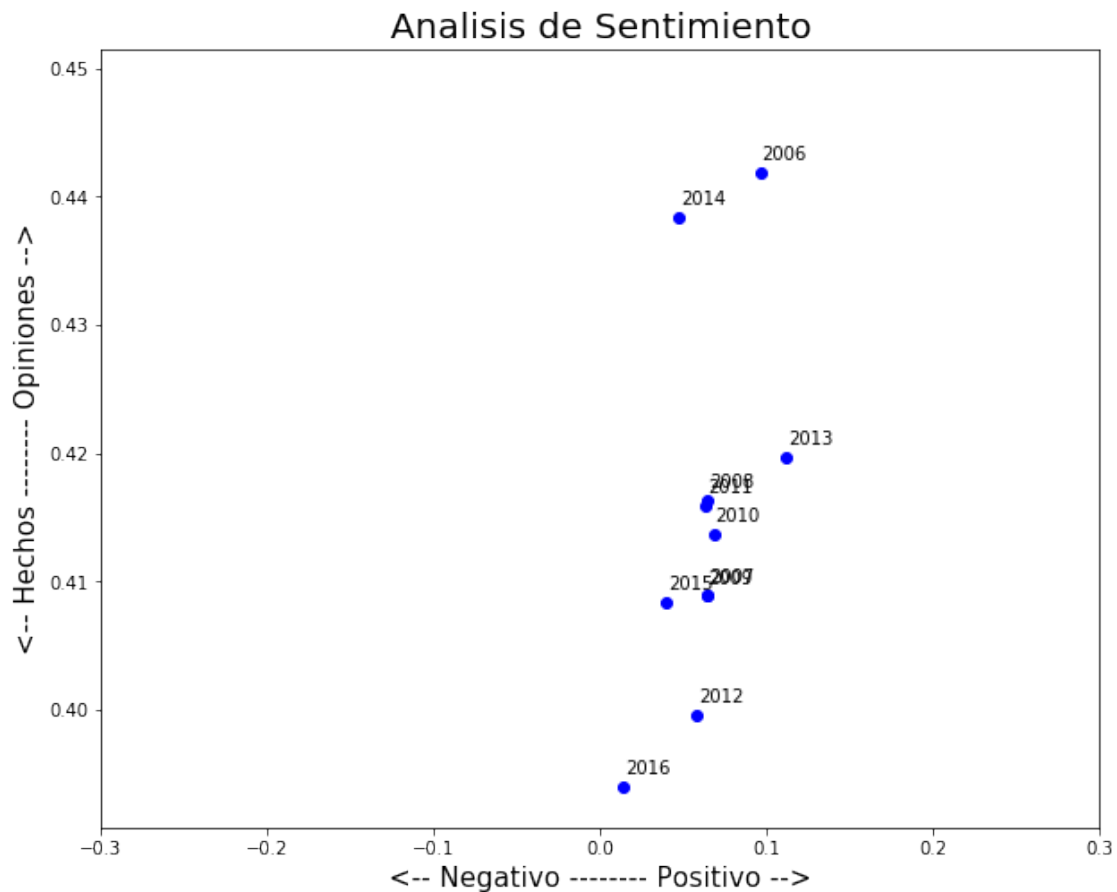
plt.rcParams['figure.figsize'] = [10, 8]

for index, anio in enumerate(data_pe.index):
    x = data_pe.polarity.loc[anio]
    y = data_pe.subjectivity.loc[anio]
    plt.scatter(x, y, color='blue')
    #plt.text(x+.001, y+.001, data['full_names'][index], fontsize=10)
    plt.text(x+.001, y+.001, anio, fontsize=10)
    #plt.xlim(-0.051, 0.152)
    plt.xlim(-0.3, 0.3)

plt.title('Análisis de Sentimiento', fontsize=20)
plt.xlabel('<-- Negativo ----- Positivo -->', fontsize=15)
plt.ylabel('<-- Hechos ----- Opiniones -->', fontsize=15)

plt.show()

```



CPU times: user 198 ms, sys: 4.04 ms, total: 202 ms
Wall time: 200 ms

6.1 Variacion del sentimiento con el tiempo

```
[61]: %%time
# Partimos cada año en 12 piezas correspondiente a los meses
import numpy as np
import math

def split_text(text, n=12):

    length = len(text)
    size = math.floor(length / n)
    start = np.arange(0, length, size)

    split_list = []
    for piece in range(n):
        split_list.append(text[start[piece]:start[piece]+size])
    return split_list
```

CPU times: user 23 μ s, sys: 1e+03 ns, total: 24 μ s
Wall time: 33.9 μ s

```
[62]: # Veamos que datos tenemos
data_pe
```

```
[62]:
```

	Texto \
2006	córdoba alianza delasotista se impone en marcos juarez el reclutamiento docente cae en picada relatos de una obsesión ya no podía entrar al aul...
2007	la historia se repite más de cien heridos con quemaduras de pirotecnia y lesiones en los ojos el primer bebé del año nació en buenos aires quemar...
2008	atentados y operativos de seguridad oscurecen festejos el primer bebé porteño del año nació en parque patricios un motociclista perdió la vida al ...
2009	balas perdidas corchos riñas callejeras e incidentes con pirotecnia en los festejos de fin de año cromañon en bangkok muertos en una discoteca...
2010	peajes bien chévere los placeres mínimos de cada día para estar saludables mateo ramos se postula para ser el primer bebé de voraz incendio en be...
2011	toman otro predio en villa lugano durante los festejos de año nuevo hoy asume dilma rousseff en brasil en santiago del estero el arrancó con un s...
2012	murió el gobernador de río negro quién asume el gobierno en río negro primeros detalles y repercusiones de la confusa muerte de soria carlos soria...
2013	pablo echarri la militancia me llegó tarde a los infografía feriados encuentran muerto al andinista perdido en el aconcagua boudou pasó el...
2014	no creo que haya amor creo que hay propiedad privada y no me la toca nadie la verdad mexicana no es un lujo es algo que necesito la conferenci...
2015	caso lola aseguran que es extremadamente complicado choque y vuelco fatal en camino del buen ayre boudou el embajador en la reasunción de di...
2016	la búsqueda de los tres prófugos se concentra en zona sur el papa

francisco ofició la primera misa del año y alertó sobre la opresión isis
planeab...

	full_name \		blob_en \		
2006	2006				
2007	2007				
2008	2008				
2009	2009				
2010	2010				
2011	2011				
2012	2012				
2013	2013				
2014	2014				
2015	2015				
2016	2016				
2006	(c, o, r, d, o, b, a, , d, e, l, a, s, o, t, i, s, t, a, , a, l, l, i, a, n, c, e, , i, m, p, o, s, e, s, , i, t, s, e, l, f, , o, n, , m, a...				
2007	(h, i, s, t, o, r, y, , r, e, p, e, a, t, s, , m, o, r, e, , t, h, a, n, , o, n, e, , h, u, n, d, r, e, d, , w, o, u, n, d, e, d, , w, i, t...				
2008	(a, t, t, a, c, k, s, , a, n, d, , s, e, c, u, r, i, t, y, , o, p, e, r, a, t, i, o, n, s, , o, b, s, c, u, r, e, , c, e, l, e, b, r, a, t, i...				
2009	(l, o, s, t, , b, u, l, l, e, t, s, , c, o, r, k, s, , s, t, r, e, e, t, , q, u, a, r, r, e, l, s, , a, n, d, , i, n, c, i, d, e, n, t, s, ...				
2010	(t, o, l, l, s, , w, e, l, l, , c, o, o, l, , t, h, e, , m, i, n, i, m, u, m, , p, l, e, a, s, u, r, e, s, , o, f, , e, a, c, h, , d, a, y...				
2011	(t, h, e, y, , t, a, k, e, , a, n, o, t, h, e, r, , p, r, o, p, e, r, t, y, , i, n, , V, i, l, l, a, , L, u, g, a, n, o, , d, u, r, i, n, g...				
2012	(T, h, e, , g, o, v, e, r, n, o, r, , o, f, , R, í, o, , N, e, g, r, o, , d, i, e, d, , , w, h, o, , a, s, s, u, m, e, s, , t, h, e, , g...				
2013	(p, a, u, l, , e, c, h, a, r, i, , t, h, e, , m, i, l, i, t, a, n, c, y, , I, , w, a, s, , l, a, t, e, , t, o, , t, h, e, , h, o, l, i, d...				
2014	(I, , d, o, , n, o, t, , t, h, i, n, k, , t, h, e, r, e, , i, s, , l, o, v, e, , I, , t, h, i, n, k, , t, h, e, r, e, , i, s, , p, r, i...				
2015	(l, o, l, a, , c, a, s, e, , a, s, s, u, r, e, , t, h, a, t, , i, t, , i, s, , e, x, t, r, e, m, e, l, y, , c, o, m, p, l, i, c, a, t, e, d...				
2016	(T, h, e, , s, e, a, r, c, h, , f, o, r, , t, h, e, , t, h, r, e, e, , f, u, g, i, t, i, v, e, s, , i, s, , c, o, n, c, e, n, t, r, a, t, e...				
	polarity	subjectivity			
2006	0.096154	0.441835			
2007	0.064330	0.408982			
2008	0.064545	0.416332			
2009	0.063941	0.408875			
2010	0.067944	0.413715			
2011	0.063373	0.415941			

2012	0.057627	0.399632
2013	0.111451	0.419632
2014	0.047445	0.438419
2015	0.039250	0.408422
2016	0.014164	0.393963

```
[63]: %%time
# Juntamos todos los textos
list_pieces = []
for t in data_pe.blob_en:#transcript:
    split = split_text(t,12)
    list_pieces.append(split)
#list_pieces
```

CPU times: user 22.2 ms, sys: 45 µs, total: 22.2 ms
Wall time: 21.6 ms

```
[64]: # Cada lista tiene n elementos
len(list_pieces)
```

[64]: 11

```
[65]: # Ahora lo partimos en piezas de 10
len(list_pieces[0])
```

[65]: 12

```
[66]: %%time
# Calculamos la polaridad de cada pieza de texto

polarity_transcript = []
for lp in list_pieces:
    polarity_piece = []
    for p in lp:
        #polarity_piece.append(TextBlob(p).translate(to="en").sentiment.
        ↪polarity)
        polarity_piece.append(p.sentiment.polarity)
    polarity_transcript.append(polarity_piece)

polarity_transcript
```

CPU times: user 421 ms, sys: 0 ns, total: 421 ms
Wall time: 418 ms

```
[66]: [[0.11719095719095719,
0.09546678276186475,
0.03472222222222222,
```

0.11645694230439996,
 0.03725399152818508,
 0.26444444444444444,
 0.0,
 0.025,
 0.16666666666666666,
 0.07619047619047618,
 0.18939393939393936,
 0.15],
 [0.07450394354503945,
 0.03789120715350223,
 0.057879453879453875,
 0.11409932659932662,
 0.02393939393939394,
 0.059554730983302405,
 0.19333333333333333,
 0.15,
 0.06666666666666667,
 0.11079545454545453,
 0.07777777777777778,
 0.0],
 [0.09597378941641237,
 0.04980996404725217,
 0.10604031385281387,
 0.053415736310473146,
 0.010527111777111776,
 0.04938973063973063,
 0.1,
 0.25,
 0.0050000000000000044,
 0.18333333333333335,
 0.23333333333333334,
 -0.050000000000000002],
 [0.021179981634527084,
 0.060296717171717175,
 0.08908481365377917,
 0.04274170274170274,
 0.07059483726150392,
 0.1035240800865801,
 0.13999999999999999,
 0.0,
 0.15,
 -0.1,
 0.100000000000000002,
 0.1875],
 [0.05464494442755312,
 0.04303977272727272,

0.1263104731525784,
0.011816524316524312,
0.08113780663780662,
0.03827875023527196,
0.19285714285714287,
0.0875,
0.23125,
0.10000000000000002,
0.13541666666666669,
0.08333333333333333],
[0.09385930735930738,
0.10505832130832131,
0.033326318742985396,
-0.014818882098293865,
0.14116573902288188,
0.009649122807017544,
-0.22666666666666663,
-0.4,
0.0,
0.1875,
0.29166666666666663,
0.19999999999999998],
[0.005075757575757573,
0.11388694638694638,
0.03756221719457015,
0.11045321637426896,
-0.0011278195488721885,
0.1877536231884058,
0.09999999999999999,
0.07142857142857142,
-0.4375,
0.20833333333333334,
0.125,
-0.07539682539682539],
[0.07891729797979799,
0.15084670231729055,
0.13376623376623376,
0.16667036215816708,
0.030364953886693014,
0.08373917748917749,
0.0,
0.2,
0.125,
0.10000000000000002,
0.17291666666666666,
0.04583333333333334],
[0.02257025761124121,

```

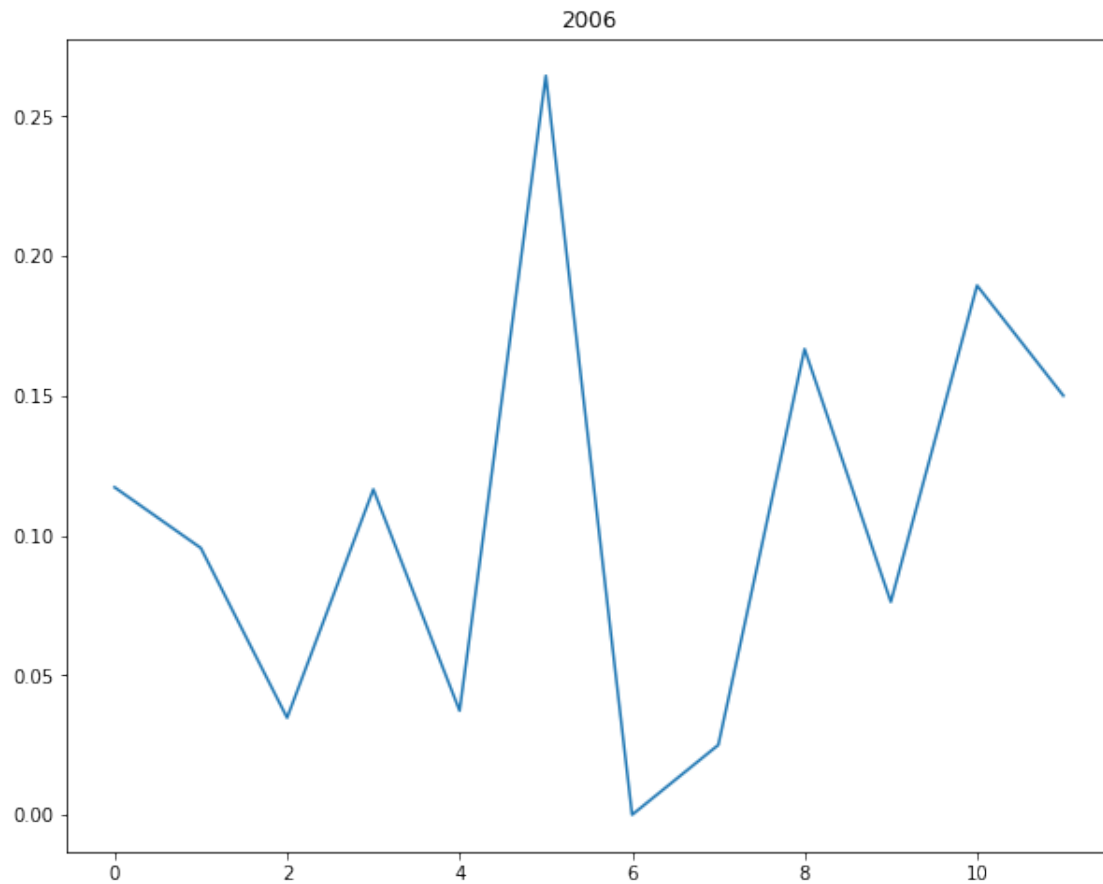
0.17694895382395381,
0.05009295078444016,
-0.05666502689229962,
0.0675273139558854,
0.01799242424242424,
0.0,
0.06666666666666667,
0.15,
0.1,
0.07333333333333333,
0.0],
[0.07470835111680182,
0.07891077441077443,
-0.01103476626732441,
0.06318928198238541,
0.01890782828282826,
-0.10434782608695647,
0.21666666666666667,
0.06818181818181818,
0.04166666666666667,
0.09659090909090909,
0.0,
0.0],
[0.030980693413729114,
-0.025694444444444443,
0.03232022607022609,
0.02927143778207608,
0.0216892955529319,
-0.06984126984126984,
0.08611111111111112,
0.08333333333333333,
0.03409090909090909,
0.0,
-0.01666666666666667,
0.1238095238095238]]

```

```

[67]: %%time
# Mostramos solo el etalle del primer año
plt.plot(polarity_transcript[0])
plt.title(data_pe['full_name'].index[0])
plt.show()

```



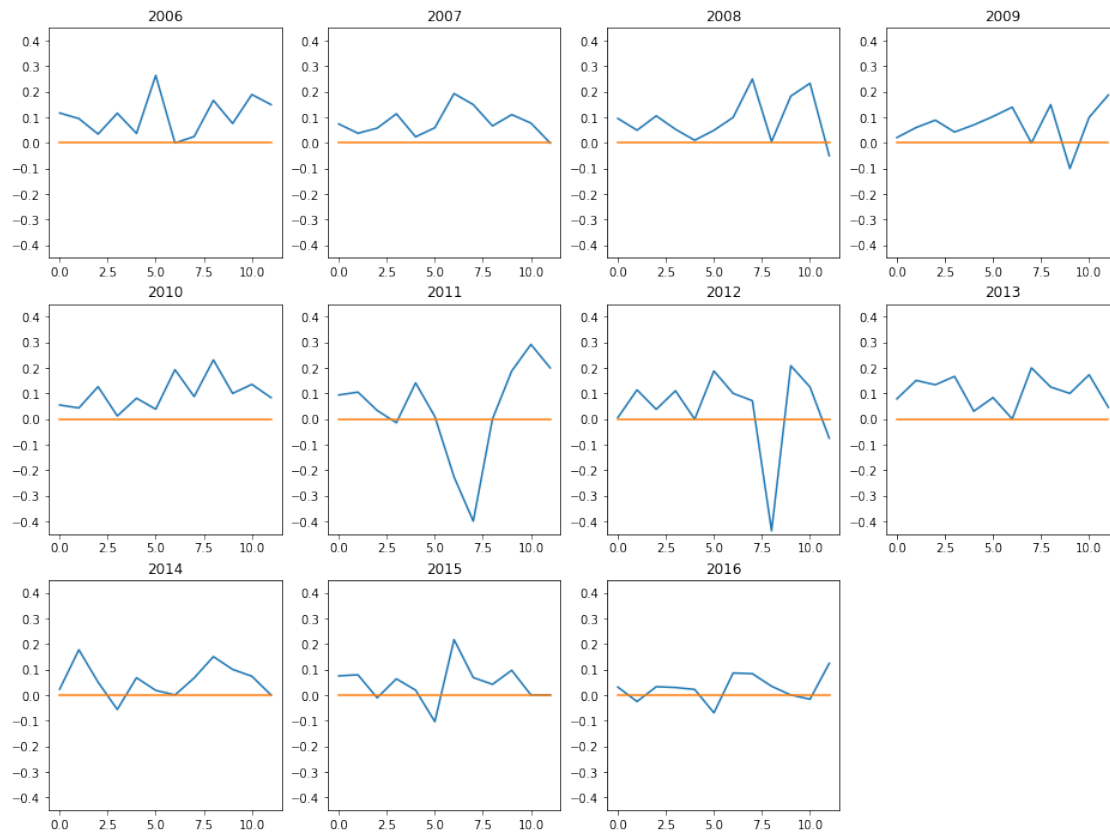
CPU times: user 147 ms, sys: 3.96 ms, total: 151 ms

Wall time: 149 ms

```
[68]: %%time
# Hacemos todos los años
plt.rcParams['figure.figsize'] = [16, 12]

for index, anio in enumerate(data_pe.index):
    plt.subplot(3, 4, index+1)
    plt.plot(polarity_transcript[index])
    plt.plot(np.arange(0,12), np.zeros(12))
    #plt.title(data['full_name'][index])
    plt.title(anio)
    plt.ylim(ymin=-.45, ymax=.45)

plt.show()
```



CPU times: user 1.33 s, sys: 7.51 ms, total: 1.33 s
Wall time: 1.33 s