

Text Mining Exercise week 4

By Joris van Vugt

Introduction

The goal of this assignment was to extract dates (in an arbitrary format) from articles from <http://www.biography.com/>. The dates then had to be converted to ISO 8601 standard (i.e. yyyy-mm-dd) and presented in a (sorted) table along with the sentence it was extracted from.

Approach

I used an article on [Mahatma Gandhi](#) to find patterns for dates. I concluded that the most frequently used format looks like "January 5, 2016", although some other formats are also used (E.g., no day or only a year). I used regular expressions to find these patterns in the text. The only adjustment I had to make to my NLP program from week 2 was better utf-8 support.

Results

I've tested my program on the article on [Barack Obama](#). I extracted 107 dates, all of which were converted to their ISO 8601 formats correctly. $\text{Precision} = 107 / 107 = 1$. Although the article on Obama is a lot longer than the one I used for finding patterns, it still worked very well. This is because the dates were formatted very consistently.

Discussion

Extracting dates is not very difficult when they are consistently and unambiguously formatted. However, some dates might not be very relevant. For example, the last date extracted, 2030, ("*President Obama praised the agreement for establishing the "enduring framework the world needs to solve the climate crisis" and pledged that the United States would cut its emissions more than 25 percent by 2030*"), is not about Obama's life and is a lot less relevant than "*In February 1990, Obama was elected the first African-American editor of the Harvard LawReview*".

Date	Sentence
1928	The first sitting American president to visit the island nation since 1928, Obama made the three-day visit accompanied by First Lady Michelle Obama and their daughters Malia and Sasha.
1961	Obama's visit was part of a larger program to establish greater cooperation between the two countries, the foundations of which were laid in late 2014, when Obama and Cuban president Raul Castro announced the normalizing of diplomatic relations for the first time

	since 1961.
1961-02-02	While studying at the University of Hawaii at Manoa, Obama Sr. met fellow student Ann Dunham, and they married on February 2, 1961.
1961-08-04	SynopsisBorn on August 4, 1961, in Honolulu, Hawaii, Barack Obama is the 44th and current president of the United States.
1961-08-04	Early LifeBarack Hussein Obama II was born on August 4, 1961, in Honolulu, Hawaii.
1964-03	Obama's parents officially separated several months later and ultimately divorced in March 1964, when their son was two.
1965	Soon after, Obama Sr. returned to Kenya.In 1965, Dunham married Lolo Soetoro, a University of Hawaii student from Indonesia.
1970	A year later, the family moved to Jakarta, Indonesia, where Obama's half-sister, Maya Soetoro Ng, was born in 1970.
1971	He later described how he struggled to reconcile social perceptions of his multiracial heritage with his own sense of self: "I noticed that there was nobody like me in the Sears, Roebuck Christmas catalog. . .and that Santa was a white man," he wrote. "I went into the bathroom and stood in front of the mirror with all my senses and limbs seemingly intact, looking as I had always looked, and wondered if something was wrong with me."Obama also struggled with the absence of his father, who he saw only once more after his parents divorced, when Obama Sr. visited Hawaii for a short time in 1971. "[My father] had left paradise, and nothing that my mother or grandparents told me could obviate that single, unassailable fact," he later reflected. "They couldn't describe what it might have been like had he stayed."Ten years later, in 1981, tragedy struck Obama Sr. when he lost both of his legs in a serious car accident.
1979	EducationWhile living with his grandparents, Obama enrolled in the esteemed Punahou Academy, He excelled in basketball and graduated with academic honors in 1979.
1981	He later described how he struggled to reconcile social perceptions of his multiracial heritage with his own sense of self: "I noticed that there was nobody like me in the Sears, Roebuck Christmas catalog. . .and that Santa was a white man," he wrote. "I went into the bathroom and stood in front of the mirror with all my senses and limbs seemingly intact, looking as I had always looked, and wondered if something was wrong with me."Obama also struggled with the absence of his father, who he saw only once more after his parents divorced, when Obama Sr. visited Hawaii for a short time in 1971. "[My father] had left paradise, and nothing that my mother or grandparents told me could obviate that single, unassailable fact," he later reflected. "They couldn't describe what it might have been like had he stayed."Ten years later, in 1981, tragedy struck Obama Sr. when he lost both of his legs in a serious car accident.
1982	In 1982, Obama Sr. was involved in yet another car accident while traveling in Nairobi.
1982-11-24	Obama Sr. died on November 24, 1982, when Obama was 21 years old. "At the time of his death, my father remained a myth to me," Obama later wrote, "both more and less than a man."After high school, Obama studied at Occidental College in Los Angeles for two years.
1983	He then transferred to Columbia University in New York City, graduating in 1983 with a degree in political science.
1985	After working in the business sector for two years, Obama moved to Chicago in 1985.

1988	He also visited relatives in Kenya, and paid an emotional visit to the graves of his biological father and paternal grandfather. "For a long time I sat between the two graves and wept," Obama wrote. "I saw that my life in America—the black life, the white life, the sense of abandonment I'd felt as a boy, the frustration and hope I'd witnessed in Chicago—all of it was connected with this small plot of earth an ocean away."Returning from Kenya with a sense of renewal, Obama entered Harvard Law School in 1988.
1990-02	In February 1990, Obama was elected the first African-American editor of the Harvard Law Review.
1991	He graduated magna cum laude from Harvard Law in 1991.After law school, Obama returned to Chicago to practice as a civil rights lawyer with the firm of Miner, Barnhill & Galland.
1992	He also taught constitutional law part-time at the University of Chicago Law School between 1992 and 2004—first as a lecturer and then as a professor—and helped organize voter registration drives during Bill Clinton's 1992 presidential campaign.
1992-10-03	On October 3, 1992, he and Michelle were married.
1995	Entry Into Illinois PoliticsObama published an autobiography, Dreams from My Father: A Story of Race and Inheritance, in 1995.
1996	He was elected to the Illinois State Senate in 1996 and to the U.S.
1996	He ran as a Democrat and won election in 1996.
1998	They moved to Kenwood, on Chicago's South Side, and welcomed two daughters several years later: Malia (born 1998) and Sasha (born 2001).
2000	As chairman of the Illinois Senate's Health and Human Services Committee Obama worked with law enforcement officials to require the videotaping of interrogations and confessions in all capital cases after a number of death-row inmates were found to be innocent.In 2000, Obama made an unsuccessful Democratic primary run for the U.S.
2001	They moved to Kenwood, on Chicago's South Side, and welcomed two daughters several years later: Malia (born 1998) and Sasha (born 2001).
2001	With the help of political consultant David Axelrod, Obama began assessing his prospects for a Senate win.Following the 9/11 attacks in 2001, Obama was an early opponent of President George W.
2002	Undeterred, he created a campaign committee in 2002 and began raising funds to run for a seat in the U.S.
2002-10	Obama was still a state senator when he spoke against a resolution authorizing the use of force against Iraq during a rally at Chicago's Federal Plaza in October 2002. "I am not opposed to all wars.
2003	"I'm opposed to dumb wars," he said. "What I am opposed to is the cynical attempt by Richard Perle and Paul Wolfowitz and other armchair, weekend warriors in this administration to shove their own ideological agendas down our throats, irrespective of the costs in lives lost and in hardships borne." Despite his protests, the Iraq War began in 2003.
2004	Senate in 2004.
2004	He also taught constitutional law part-time at the University of Chicago Law School between 1992 and 2004—first as a lecturer and then as a professor—and helped

	organize voter registration drives during Bill Clinton's 1992 presidential campaign.
2004	The book had a second printing in 2004 and was adapted for a children's version.
2004	Senate in 2004.
2004	In the 2004 Democratic primary, he defeated multimillionaire businessman Blair Hull and Illinois Comptroller Daniel Hynes with 52 percent of the vote.
2004	That summer, he was invited to deliver the keynote speech in support of John Kerry at the 2004 Democratic National Convention in Boston.
2004-06	However, Ryan withdrew from the race in June 2004 following public disclosure of unsubstantiated sexual deviancy allegations by his ex-wife, actress Jeri Ryan.In August 2004, diplomat and former presidential candidate Alan Keyes accepted the Republican nomination to replace Ryan.
2004-08	However, Ryan withdrew from the race in June 2004 following public disclosure of unsubstantiated sexual deviancy allegations by his ex-wife, actress Jeri Ryan.In August 2004, diplomat and former presidential candidate Alan Keyes accepted the Republican nomination to replace Ryan.
2004-11	In the November 2004 general election, Obama received 70 percent of the vote to Keyes' 27 percent, the largest electoral victory in Illinois history.
2005	President Obama called the plan the "single most important step that America has ever made in the fight against global climate change."The plan calls for aggressive Environmental Protection Agency regulations including requiring existing power plants to cut carbon dioxide emissions 32 percent from 2005 levels by 2030 and use more renewable energy sources like wind and solar power.
2005-01-03	Senate since Reconstruction.Sworn into office on January 3, 2005, Obama partnered with Republican Senator Richard Lugar of Indiana on a bill that expanded efforts to destroy weapons of mass destruction in Eastern Europe and Russia.
2006	The audiobook version of Dreams, narrated by Obama, received a Grammy Award for best spoken word album in 2006.Obama's advocacy work led him to run for a seat in the Illinois State Senate.
2006-10	Obama also spoke out for victims of Hurricane Katrina, pushed for alternative energy development and championed improved veterans' benefits.His second book, The Audacity of Hope: Thoughts on Reclaiming the American Dream, was published in October 2006.
2007-02	2008 Presidential ElectionIn February 2007, Obama made headlines when he announced his candidacy for the 2008 Democratic presidential nomination.
2008	First elected to the presidency in 2008, he won a second term in 2012.
2008	He was elected to the U.S. presidency in 2008, and won re-election in 2012 against Republican challenger Mitt Romney.
2008	2008 Presidential ElectionIn February 2007, Obama made headlines when he announced his candidacy for the 2008 Democratic presidential nomination.
2008	2012 Re-ElectionAs he did in 2008, during his campaign for a second presidential term, Obama focused on grassroots initiatives.
2008-06-03	On June 3, 2008, Obama became the Democratic Party's presumptive nominee after winning a sufficient number of pledged delegates during the primaries, and Clinton

	delivered her full support to Obama for the duration of his campaign.
2008-11-04	On November 4, 2008, Barack Obama defeated Republican presidential nominee John McCain, 52.9 percent to 45.7 percent, to win election as the 44th president of the United States and the first African-American to hold this office.
2009	For his efforts, the Nobel Committee in Norway awarded Obama the 2009 Nobel Peace Prize.
2009	Voting with the majority were two associate justices appointed by Obama Sonia Sotomayor (confirmed in 2009) and Elena Kagan (confirmed in 2010).
2009-01-20	Obama's inauguration took place on January 20, 2009. When Obama took office, he inherited a global economic recession, two ongoing foreign wars and the lowest-ever international favorability rating for the United States.
2009-04-29	First 100 Days Between Inauguration Day and April 29, 2009, the Obama administration took action on many fronts.
2010	2010 State of the Union On January 27, 2010, President Obama delivered his first State of the Union speech.
2010	House of Representatives in the 2010 mid-term elections, he signed the Budget Control Act of 2011 in an effort to rein in government spending and prevent the government from defaulting on its financial obligations.
2010	Voting with the majority were two associate justices appointed by Obama Sonia Sotomayor (confirmed in 2009) and Elena Kagan (confirmed in 2010).
2010-01-27	2010 State of the Union On January 27, 2010, President Obama delivered his first State of the Union speech.
2010-03	In spite of opposition from Congressional Republicans and the populist Tea Party movement, Obama signed his health care reform plan, known as the Affordable Care Act, into law in March 2010.
2010-08	He committed an additional 21,000 troops to Afghanistan and set an August 2010 date for withdrawal of nearly all U.S. troops from Iraq.
2011	House of Representatives in the 2010 mid-term elections, he signed the Budget Control Act of 2011 in an effort to rein in government spending and prevent the government from defaulting on its financial obligations.
2011	The act also called for the creation of a bipartisan committee to seek solutions to the country's fiscal issues, but the group failed to reach any agreement on how to solve these problems. Also in 2011, Obama signed a repeal of the military policy known as "Don't Ask, Don't Tell," which prevented openly gay troops from serving in the U.S.
2011-03	In March 2011, he approved U.S. participation in NATO airstrikes to support rebels fighting against the forces of Libyan dictator Muammar al-Qaddafi, and in May he also gave the green light to a covert operation in Pakistan that led to the killing of infamous al-Qaeda leader Osama bin Laden by a team of U.S.
2012	First elected to the presidency in 2008, he won a second term in 2012.
2012	He was elected to the U.S. presidency in 2008, and won re-election in 2012 against Republican challenger Mitt Romney.
	The next year, he met with constitutional law professor Laurence Tribe and their

2012	discussion so impressed Tribe, that when Obama asked to join his team as a research assistant, the professor agreed. ♦ The better he did at Harvard Law School and the more he impressed people, the more obvious it became that he could have had anything, said Professor Tribe in a 2012 interview with Frontline, ♦ but it was clear that he wanted to make a difference to people, to communities. ♦ That same year Obama joined the Chicago law firm of Sidley Austin as a summer associate and it was there he met Michelle Robinson, a young lawyer who was assigned to be his adviser.
2012	2012 Re-ElectionAs he did in 2008, during his campaign for a second presidential term, Obama focused on grassroots initiatives.
2012	And we'll remind the world just why it is that the United States of America is the greatest nation on Earth."In the 2012 election, Obama faced Republican opponent Mitt Romney and Romney's vice-presidential running mate, U.S.
2012	Prior to the bill's passage, in late 2012, tense negotiations between Republicans and Democrats over spending cuts and tax increases became a bitter political battle until Vice President Joe Biden managed to hammer out a deal with Republican Senate Minority Leader Mitch McConnell.
2012	Citing examples such as the 2012 mass shooting at Sandy Hook elementary school, the president shed tears as he called on Congress and the gun lobby to work with him to make the country safer.
2012	Citing examples such as the 2012 mass shooting at Sandy Hook elementary school in Connecticut, the president shed tears as he called on Congress and the gun lobby to work with him to make the country safer.
2012-05	President Obama, who became the first president to voice support for same-sex marriage in May 2012, praised the court for affirming "that the Constitution guarantees marriage equality.
2012-06	Obama gained a legal victory in June 2012 when the U.S.
2012-06	Celebrities such as Anna Wintour and Sarah Jessica Parker aided the president's campaign by hosting fund-raising events."I guarantee you, we will move this country forward," Obama stated in June 2012, at a campaign event in Maryland. "We will finish what we started.
2012-11-06	On November 6, 2012, Obama won a second four-year term as president by receiving nearly five million more votes than Romney and capturing more than 60 percent of the Electoral College.Nearly one month after President Obama's re-election, the nation endured one of its most tragic school shootings to date when 20 children and six adults were shot to death at the Sandy Hook Elementary School in Newtown, Connecticut, on December 14, 2012.
2012-12-14	On November 6, 2012, Obama won a second four-year term as president by receiving nearly five million more votes than Romney and capturing more than 60 percent of the Electoral College.Nearly one month after President Obama's re-election, the nation endured one of its most tragic school shootings to date when 20 children and six adults were shot to death at the Sandy Hook Elementary School in Newtown, Connecticut, on December 14, 2012.
2013	Managing Foreign CrisesThe fall of 2013 brought Obama additional challenges in the area of foreign relations.
	Are we really prepared to say that we're powerless in the face of such carnage, that the

2013-01-01	politics are too hard?"Obama achieved a major legislative victory on January 1, 2013, when the Republican-controlled House of Representatives approved a bipartisan agreement on tax increases and spending cuts, in an effort to avoid the looming fiscal cliff crisis (the Senate voted in favor of the bill earlier that day).
2013-01-21	Second TermBarack Obama officially began his second term on January 21, 2013, when U.S.
2013-04-15	Alicia Keys and Jamie Foxx also performed.After the inauguration, Obama led the nation through many challenges◆none more difficult, perhaps, than the terrorist bombings of the Boston Marathon on April 15, 2013, which killed three people and left more than 200 injured.
2013-07	Obama defended the NSA's email monitoring and telephone wiretapping during a visit to Germany that June. "We are not rifling through the emails of German citizens or American citizens or French citizens or anyone else,◆ he said. "The encroachment on privacy has been strictly limited." Obama stated that the program had helped stop roughly 50 threats.In early July 2013, President Obama made history when he joined former President George W.
2013-09	International ChallengesObama found himself grappling with an international crisis in late August and September 2013 when it was discovered that Syrian leader Bashar al-Assad had used chemical weapons against civilians.
2013-09-10	Obama then announced an alternative solution on September 10, 2013, by stating that if al-Assad agreed with the stipulations outlined in a proposal made by Russia to give up its chemical weapons, then a direct strike against the nation could be avoided.
2013-10	Domestic Policies and ProblemsObama found himself struggling on the domestic front in October 2013.
2013-10	In October 2013, German Chancellor Angela Merkel revealed that the NSA had been listening in to her cell phone calls. "Spying among friends is never acceptable," Merkel told a summit of European leaders.
2013-11	In the wake of these controversies, Obama saw his approval rating drop to a new low in November 2013.
2014	In response, Obama ordered sanctions targeting individuals and businesses considered by the U.S. government to be Ukraine agitators or involved in the Crimean crisis. "In 2014 we are well beyond the days when borders can be redrawn over the heads of democratic leaders," Obama stated.
2014	The president said the sanctions were taken in close coordination with European allies and gave the U.S. "the flexibility to adjust our response going forward based on Russia's actions.◆ In addition to the ongoing troubles in Ukraine, tensions between Israelis and Palestinians erupted into violence in Gaza during the summer of 2014.
2014	Presidency After 2014 ElectionsThat November, Obama had to cope with new challenges on the home front.
2014	Obama's visit was part of a larger program to establish greater cooperation between the two countries, the foundations of which were laid in late 2014, when Obama and Cuban president Raul Castro announced the normalizing of diplomatic relations for the first time since 1961.
2014-	Only 37 percent of Americans polled by CBS News approved of the job he was doing as president, while 57 percent disapproved of his handling of the job.Echoes of the Cold

02	War also returned after civil unrest and protests in the capital city of Kiev led to the downfall of Ukrainian President Viktor Yanukovich's administration in February 2014.
2014-08	In August 2014, Obama ordered the first airstrikes against the self-proclaimed Islamic State, also known as ISIS or ISIL, which had seized large swathes of Iraq and Syria and conducted high-profile beheadings of foreign hostages.
2015	In his 2015 State of the Union address, Obama declared that the nation was out of recession. "America, for all that we've endured; for all the grit and hard work required to come back . . . know this: The shadow of crisis has passed," he said.
2015	Supreme Court Victories The summer of 2015 brought two major U.S.
2015	2015 Paris Climate ConferenceIn November 2015, Obama further demonstrated his commitment to environmental issues as a primary player in the international COP21 summit held outside of Paris, France.
2015	According to a 2015 Gallup poll, most Americans favor some kind of stricter regulations of gun sales.
2015	According to a 2015 Gallup poll, most Americans favor some kind of stricter regulations of gun sales.Shortly after the press conference, on January 12, 2016, Barack Obama delivered what would be his final State of the Union address.
2015-07	Iran Nuclear DealIn July 2015, Obama announced that, after lengthy negotiations, the United States and five world powers had reached an agreement with Iran over its nuclear program.
2015-08	Clean Power PlanIn August 2015, the Obama administration announced The Clean Power Plan, a major climate change plan aimed at reducing greenhouse gas emissions, the first-ever national standards to limit carbon pollution from coal-burning power plants in the United States.
2015-11	2015 Paris Climate ConferenceIn November 2015, Obama further demonstrated his commitment to environmental issues as a primary player in the international COP21 summit held outside of Paris, France.
2016	Under the regulations, states will be allowed to create their own plans to reduce emissions and are required to submit initial plans by 2016 and final versions by 2018.
2016-01	Gun ControlEntering his final year as President of the United States, in early January 2016 Obama held a press conference to announce a new series of executive orders related to gun control.
2016-01	Final Year in OfficeEntering his final year as President of the United States, in early January 2016 Obama held a press conference to announce a new series of executive orders related to gun control.
2016-01-12	According to a 2015 Gallup poll, most Americans favor some kind of stricter regulations of gun sales.Shortly after the press conference, on January 12, 2016, Barack Obama delivered what would be his final State of the Union address.
2018	Under the regulations, states will be allowed to create their own plans to reduce emissions and are required to submit initial plans by 2016 and final versions by 2018.
2030	President Obama called the plan the "single most important step that America has ever made in the fight against global climate change."The plan calls for aggressive Environmental Protection Agency regulations including requiring existing power plants to cut carbon dioxide emissions 32 percent from 2005 levels by 2030 and use more

	renewable energy sources like wind and solar power.
2030	President Obama praised the agreement for establishing the ♦enduring framework the world needs to solve the climate crisis♦ and pledged that the United States would cut its emissions more than 25 percent by 2030.