

NLP

Find your favorite news source and grab the article text.

1. Show the most common words in the article.
2. Show the most common words under a part of speech. (i.e. NOUN: {'Bob':12, 'Alice':4,})
3. Find a subject/object relationship through the dependency parser in any sentence.
4. Show the most common Entities and their types.
5. Find Entites and their dependency (hint: entity.root.head)
6. Find the most similar words in the article

Note: Yes, the notebook from the video is not provided, I leave it to you to make your own :) it's your final assignment for the semester. Enjoy!

```
In [1]: import spacy
nlp = spacy.load("en_core_web_sm")

In [2]: text = """Posted February 15, 2023 at 11:20 am by Kyle K. Moore
Five principles for making state and local reparations plans reparative
We are still living in the aftermath of 2020's overlapping crises of racial injustice, our nation's polycrisis. Between the emerge
This confluence of tragedies brought awareness of the Black American condition to a new generation. It also reinvigorated interest
This renewed awareness and interest in addressing racial disparities brought attention to arguably the only structural solution to
William Darity and Kirsten Mullen's 2020 book, From Here to Equality: Reparations for Black Americans in the Twenty-First Century,
Despite the raised awareness around the issue of reparations, the political will to enact a federal plan is currently lacking. H.R
In the absence of a federal plan, various states and localities have convened their own commissions and, in some cases, enacted th
As more states and localities attempt to design reparations plans, it is important that they keep certain principles in mind to en
Darity and Mullen identify three criteria for an effective federal reparations plan. The plan must include acknowledgement and apo
In addition to acknowledgement and redress, I suggest that any sub-federal reparations plan that aims to be effective should:

Specify what harms are being addressed and who will benefit.
Stay within its capacity to provide redress for its harm—and avoid absolving the federal government from its responsibility.
Commit to structural change designed to prevent future racial injustice.
Sub-federal reparations plans should acknowledge and apologize for the harm done

Acknowledgement and apology of the atrocities committed are essential to making any reparations plan legitimate, whether federal o
For example, if the cities of Wilmington, North Carolina, or Tulsa, Oklahoma, developed reparations programs centered around the m
Sub-federal reparations plans should include material redress to the beneficiaries

Any plan for Black American reparations needs to include material redress. For example, if the harm being rectified is an historic
Sub-federal reparations plans need to specify what harms are being addressed and who will benefit

For a plan to be considered reparative, it must explicitly outline the incident or historical precedent being addressed and the vi
Sub-federal reparations plans should not attempt to absolve the federal government in its responsibility to provide redress for it
The idea that a sub-federal reparations plan should stay within an appropriate scope for its capacity is important for more reason
State and local reparations plans should be grounded in an historical analysis of what institutional atrocities have taken place a
Sub-federal reparations plans should include structural change and a commitment to ongoing vigilance against future racial injusti
In the spirit of the “closure” aspect of the criteria put forward by Darity and Mullen, I suggest that sub-federal reparations pla
Sub-federal reparations plans like the ones being explored in Providence and Evanston represent an important step forward in our c

In [3]: processed_text = nlp(text)

In [4]: processed_text
```

Out[4]: Posted February 15, 2023 at 11:20 am by Kyle K. Moore

Five principles for making state and local reparations plans reparative

We are still living in the aftermath of 2020's overlapping crises of racial injustice, our nation's polycrisis. Between the emergence of the COVID-19 pandemic, the ensuing economic recession, and the public police murder of George Floyd, we saw a harsh truth about the structure of American political economy: White supremacy has shaped our institutions such that their outcome is consistent Black precarity and premature death.

This confluence of tragedies brought awareness of the Black American condition to a new generation. It also reinvigorated interest among academics and policymakers to finally do something about the problem of racial disparities (though activists and community organizers largely never lost interest in this).

This renewed awareness and interest in addressing racial disparities brought attention to arguably the only structural solution to persistent Black-white economic and social disparities, one that we have put off as a country for generations: reparations for slavery, Reconstruction, Jim Crow, and mass incarceration.

William Darity and Kirsten Mullen's 2020 book, *From Here to Equality: Reparations for Black Americans in the Twenty-First Century*, outlines a clear strategy for implementing a reparations plan that would close the Black-white racial wealth gap in America—a foundational source of inequality across several social and economic domains. Darity and Mullen call for a federal reparations plan, on the grounds that the federal government is the only institution with the means to enact a transfer at a large enough scale to close the racial wealth gap, and is further the culpable party for allowing the institutional atrocities being rectified to exist in the first place.

Despite the raised awareness around the issue of reparations, the political will to enact a federal plan is currently lacking. H. R. 40, a bill to establish the Commission to Study and Develop Reparations Proposals for African Americans, is an important first step toward passing such a federal plan, but has yet to be brought to the House floor for a vote in its 32-year history.

In the absence of a federal plan, various states and localities have convened their own commissions and, in some cases, enacted their own reparations-style plans at the sub-federal level (i.e., below the federal level). California's Reparations Task Force represents the largest state effort to investigate reparations thus far, but cities like Evanston, Illinois, St. Paul, Minnesota, and Providence, Rhode Island have all investigated and worked toward creating and implementing programs under the auspices of reparations.

As more states and localities attempt to design reparations plans, it is important that they keep certain principles in mind to ensure that the plans are effective as reparations. There is an important difference between policy that is simply "good" for Black people or people of color more broadly, and policy that is reparative.

Darity and Mullen identify three criteria for an effective federal reparations plan. The plan must include acknowledgement and apology for the harm committed; it must provide material redress for that harm; and it must bring closure through a mutual understanding between the beneficiaries (white Americans) and victims (Black Americans) of the oppressive systems that were implemented. In this post, I expand these criteria slightly to provide guidelines for states and localities that are attempting to create their own reparations-style plans in the absence of a federal effort.

In addition to acknowledgement and redress, I suggest that any sub-federal reparations plan that aims to be effective should:

Specify what harms are being addressed and who will benefit.

Stay within its capacity to provide redress for its harm—and avoid absolving the federal government from its responsibility.

Commit to structural change designed to prevent future racial injustice.

Sub-federal reparations plans should acknowledge and apologize for the harm done

Acknowledgement and apology of the atrocities committed are essential to making any reparations plan legitimate, whether federal or sub-federal. Attempting to rectify harms committed at an institutional level without speaking directly to the history of those harms would be institutional gaslighting and likely lead to poorly designed policy. Acknowledgment and apology are also the preconditions for reconciliation, closure, and a future commitment to avoid causing harm in the future.

For example, if the cities of Wilmington, North Carolina, or Tulsa, Oklahoma, developed reparations programs centered around the massacres that took place in those cities in 1898 and 1921, respectively, those plans would need to include explicit acknowledgment of and apology for those massacres. An indirect compensation plan designed to disproportionately benefit Black residents would not suffice for legitimate reparations.

Sub-federal reparations plans should include material redress to the beneficiaries

Any plan for Black American reparations needs to include material redress. For example, if the harm being rectified is an historical exclusion from a local housing program, the descendants of those excluded should be the beneficiaries of a new housing program and be materially compensated for the lost potential wealth from having been excluded. Darity and Mullen are explicit that a federal reparations plan should include direct transfers of wealth from the federal government to the descendants of Black American slaves in an amount sufficient to compensate for both the unpaid labor and the lost appreciation of the fruits of that labor. The current racial wealth gap reflects the cumulative effect of Black Americans' experience with economic racism over time, and closing that gap would be the way to redress that experience. Plans at the local and state level do not have the capacity to do this and should instead focus on addressing disparities that are within their scope.

Sub-federal reparations plans need to specify what harms are being addressed and who will benefit

For a plan to be considered reparative, it must explicitly outline the incident or historical precedent being addressed and the victims that are being compensated through the plan, in as much specificity as possible. A universal plan (that is, one that would benefit everyone) that would disproportionately benefit the descendants of those impacted by Jim Crow laws could not be considered reparative—though such policies may be worth pursuing on their own merits. The fact that a policy benefits the disadvantaged is not enough on its own to consider it a part of an agenda for racial justice, nor enough to make that policy reparative. The beneficiaries of reparations must be the descendants of the aggrieved parties—in the cases of American chattel slavery and Jim Crow, for example, those beneficiaries would be Black Americans. At the sub-federal level, reparations plans should specify the injustices committed by the state or locality, and redress should be provided to the victims of those specific policies and their descendants.

Sub-federal reparations plans should not attempt to absolve the federal government in its responsibility to provide redress for its harm

The idea that a sub-federal reparations plan should stay within an appropriate scope for its capacity is important for more reasons

s than just feasibility. It also ties into which parties are ultimately culpable for what institutional harms have been committed. For example, it would be inappropriate for a reparations plan at the city level to attempt to circumvent the federal government and provide reparations to its Black residents for chattel slavery, because the city government is not the culpable party for legalizing and supporting chattel slavery. Only the federal government can play that role.

State and local reparations plans should be grounded in an historical analysis of what institutional atrocities have taken place at that level. This similarly applies to the statistics being cited for developing state and local reparations plans. For example, citing the national racial wealth gap as part of the acknowledgement component of a city-level reparations plan would be inappropriate. Instead, that city should probe its own history and account for its own legacy of racial injustice.

Sub-federal reparations plans should include structural change and a commitment to ongoing vigilance against future racial injustice

In the spirit of the “closure” aspect of the criteria put forward by Darity and Mullen, I suggest that sub-federal reparations plans should also commit to making racial justice an ethic and practice, rather than a one-time occurrence. This would mean including elements of structural change—permanent or ongoing laws and institutions designed to prevent these atrocities from happening in the future or to stop these disparities from re-opening in unjust ways in the plan’s aftermath. These commitments would also serve to protect against the backlash to racial progress we have so often seen in American history—when reactionaries attempt to roll back policies designed to achieve some level of equity. If we have seen these periods of backlash in response to raised awareness of racial disparities and symbolic progress, we can expect them to follow successful reparations plans at any level. Policymakers should be aware and prepared for this.

Sub-federal reparations plans like the ones being explored in Providence and Evanston represent an important step forward in our collective recognition that Black Americans deserve redress for the harms inflicted by government institutions against them throughout this country’s history. However, it is important to reiterate that these plans are not and cannot be substitutes for a federal reparations effort. As Darity and Mullen point out in their work, the federal government is the culpable party for the racial wealth gap and the long-term national consequences of chattel slavery, Reconstruction, Jim Crow, and mass incarceration for Black Americans. The federal government is also the one with the requisite resources to meet the reparative challenge. But as long as state and local reparations plans are a part of the national conversation on achieving justice and equity for Black Americans, these guidelines should be a useful framework for making plans that are effective and, most importantly, reparative.

Show the most common words in the article.

```
In [5]: words = [token.text
              for token in processed_text
              if not token.is_stop and not token.is_punct]

In [6]: from collections import Counter
        common_words = Counter(words)

In [7]: del common_words["\n"], common_words["\n\n"], common_words["\n\n\n"]

In [8]: sorted(common_words.items(), key=lambda x:x[1], reverse=True)
```

```
Out[8]: [('reparations', 36),
('federal', 30),
('plans', 20),
('plan', 19),
('Black', 16),
('racial', 15),
('government', 10),
('level', 10),
('Americans', 9),
('redress', 9),
('reparative', 7),
('wealth', 7),
('harm', 7),
('state', 6),
('local', 6),
('American', 6),
('disparities', 6),
('Darity', 6),
('Mullen', 6),
('gap', 6),
('important', 6),
('sub', 6),
('include', 6),
('harms', 6),
('Sub', 6),
('slavery', 5),
('institutional', 5),
('history', 5),
('policy', 5),
('committed', 5),
('provide', 5),
('beneficiaries', 5),
('benefit', 5),
('designed', 5),
('future', 5),
('example', 5),
('descendants', 5),
('making', 4),
('injustice', 4),
('economic', 4),
('awareness', 4),
('structural', 4),
('Jim', 4),
('Crow', 4),
('culpable', 4),
('atrocities', 4),
('attempt', 4),
('effective', 4),
('acknowledgement', 4),
('apology', 4),
('chattel', 4),
('city', 4),
('institutions', 3),
('brought', 3),
('interest', 3),
('lost', 3),
('white', 3),
('Reparations', 3),
('party', 3),
('place', 3),
('states', 3),
('localities', 3),
('effort', 3),
('cities', 3),
('criteria', 3),
('material', 3),
('closure', 3),
('victims', 3),
('addressed', 3),
('capacity', 3),
('change', 3),
('historical', 3),
('policies', 3),
('justice', 3),
('national', 3),
('principles', 2),
('aftermath', 2),
('2020', 2),
('political', 2),
('new', 2),
('addressing', 2),
('social', 2),
('country', 2),
('Reconstruction', 2),
('mass', 2),
('incarceration', 2),
```

('implementing', 2),
('close', 2),
('enact', 2),
('rectified', 2),
('raised', 2),
('step', 2),
('absence', 2),
('cases', 2),
('style', 2),
('like', 2),
('Evanston', 2),
('Providence', 2),
('programs', 2),
('people', 2),
('guidelines', 2),
('suggest', 2),
('avoid', 2),
('responsibility', 2),
('prevent', 2),
('legitimate', 2),
('commitment', 2),
('massacres', 2),
('need', 2),
('explicit', 2),
('disproportionately', 2),
('residents', 2),
('housing', 2),
('program', 2),
('excluded', 2),
('compensated', 2),
('labor', 2),
('experience', 2),
('time', 2),
('scope', 2),
('specify', 2),
('considered', 2),
('laws', 2),
('parties', 2),
('inappropriate', 2),
('ongoing', 2),
('forward', 2),
('backlash', 2),
('progress', 2),
('seen', 2),
('equity', 2),
('long', 2),
('Posted', 1),
('February', 1),
('15', 1),
('2023', 1),
('11:20', 1),
('Kyle', 1),
('K.', 1),
('Moore', 1),
('living', 1),
('overlapping', 1),
('crises', 1),
('nation', 1),
('polycrisis', 1),
('emergence', 1),
('COVID-19', 1),
('pandemic', 1),
('ensuing', 1),
('recession', 1),
('public', 1),
('police', 1),
('murder', 1),
('George', 1),
('Floyd', 1),
('saw', 1),
('harsh', 1),
('truth', 1),
('structure', 1),
('economy', 1),
('White', 1),
('supremacy', 1),
('shaped', 1),
('outcome', 1),
('consistent', 1),
('precarity', 1),
('premature', 1),
('death', 1),
('confluence', 1),
('tragedies', 1),
('condition', 1),
('generation', 1),

('reinvigorated', 1),
('academics', 1),
('policymakers', 1),
('finally', 1),
('problem', 1),
('activists', 1),
('community', 1),
('organizers', 1),
('largely', 1),
('renewed', 1),
('attention', 1),
('arguably', 1),
('solution', 1),
('persistent', 1),
('generations', 1),
('William', 1),
('Kirsten', 1),
('book', 1),
('Equality', 1),
('Century', 1),
('outlines', 1),
('clear', 1),
('strategy', 1),
('America', 1),
('foundational', 1),
('source', 1),
('inequality', 1),
('domains', 1),
('grounds', 1),
('institution', 1),
('means', 1),
('transfer', 1),
('large', 1),
('scale', 1),
('allowing', 1),
('exist', 1),
('Despite', 1),
('issue', 1),
('currently', 1),
('lacking', 1),
('H.R.', 1),
('40', 1),
('bill', 1),
('establish', 1),
('Commission', 1),
('Study', 1),
('Develop', 1),
('Proposals', 1),
('African', 1),
('passing', 1),
('House', 1),
('floor', 1),
('vote', 1),
('32', 1),
('year', 1),
('convened', 1),
('commissions', 1),
('enacted', 1),
('i.e', 1),
('California', 1),
('Task', 1),
('Force', 1),
('represents', 1),
('largest', 1),
('investigate', 1),
('far', 1),
('Illinois', 1),
('St.', 1),
('Paul', 1),
('Minnesota', 1),
('Rhode', 1),
('Island', 1),
('investigated', 1),
('worked', 1),
('creating', 1),
('auspices', 1),
('design', 1),
('certain', 1),
('mind', 1),
('ensure', 1),
('difference', 1),
('simply', 1),
('good', 1),
('color', 1),
('broadly', 1),
('identify', 1),

('bring', 1),
('mutual', 1),
('understanding', 1),
('oppressive', 1),
('systems', 1),
('implemented', 1),
('post', 1),
('expand', 1),
('slightly', 1),
('attempting', 1),
('create', 1),
('addition', 1),
('aims', 1),
('Specify', 1),
('Stay', 1),
('absolving', 1),
('Commit', 1),
('acknowledge', 1),
('apologize', 1),
('Acknowledgement', 1),
('essential', 1),
('Attempting', 1),
('rectify', 1),
('speaking', 1),
('directly', 1),
('gaslighting', 1),
('likely', 1),
('lead', 1),
('poorly', 1),
('Acknowledgment', 1),
('preconditions', 1),
('reconciliation', 1),
('causing', 1),
('Wilmington', 1),
('North', 1),
('Carolina', 1),
('Tulsa', 1),
('Oklahoma', 1),
('developed', 1),
('centered', 1),
('took', 1),
('1898', 1),
('1921', 1),
('respectively', 1),
('indirect', 1),
('compensation', 1),
('suffice', 1),
('needs', 1),
('exclusion', 1),
('materially', 1),
('potential', 1),
('having', 1),
('direct', 1),
('transfers', 1),
('slaves', 1),
('sufficient', 1),
('compensate', 1),
('unpaid', 1),
('appreciation', 1),
('fruits', 1),
('current', 1),
('reflects', 1),
('cumulative', 1),
('effect', 1),
('racism', 1),
('closing', 1),
('way', 1),
('Plans', 1),
('instead', 1),
('focus', 1),
('explicitly', 1),
('outline', 1),
('incident', 1),
('precedent', 1),
('specificity', 1),
('possible', 1),
('universal', 1),
('impacted', 1),
('worth', 1),
('pursuing', 1),
('merits', 1),
('fact', 1),
('benefits', 1),
('disadvantaged', 1),
('consider', 1),
('agenda', 1),

('aggrieved', 1),
('injustices', 1),
('locality', 1),
('provided', 1),
('specific', 1),
('absolve', 1),
('idea', 1),
('stay', 1),
('appropriate', 1),
('reasons', 1),
('feasibility', 1),
('ties', 1),
('ultimately', 1),
('circumvent', 1),
('legalizing', 1),
('supporting', 1),
('play', 1),
('role', 1),
('State', 1),
('grounded', 1),
('analysis', 1),
('taken', 1),
('similarly', 1),
('applies', 1),
('statistics', 1),
('cited', 1),
('developing', 1),
('citing', 1),
('component', 1),
('Instead', 1),
('probe', 1),
('account', 1),
('legacy', 1),
('vigilance', 1),
('spirit', 1),
('aspect', 1),
('commit', 1),
('ethic', 1),
('practice', 1),
('occurrence', 1),
('mean', 1),
('including', 1),
('elements', 1),
('permanent', 1),
('happening', 1),
('stop', 1),
('opening', 1),
('unjust', 1),
('ways', 1),
('commitments', 1),
('serve', 1),
('protect', 1),
('reactionaries', 1),
('roll', 1),
('achieve', 1),
('periods', 1),
('response', 1),
('symbolic', 1),
('expect', 1),
('follow', 1),
('successful', 1),
('Policymakers', 1),
('aware', 1),
('prepared', 1),
('ones', 1),
('explored', 1),
('represent', 1),
('collective', 1),
('recognition', 1),
('deserve', 1),
('inflicted', 1),
('reiterate', 1),
('substitutes', 1),
('point', 1),
('work', 1),
('term', 1),
('consequences', 1),
('requisite', 1),
('resources', 1),
('meet', 1),
('challenge', 1),
('conversation', 1),
('achieving', 1),
('useful', 1),
('framework', 1),
('importantly', 1)]

Show the most common words under a part of speech. (i.e. NOUN: {'Bob':12, 'Alice':4,})

```
In [9]: pos = []
        for token in processed_text:
            pos.append(token.pos_)
```

```
In [10]: import numpy as np
         pos_unique = np.unique(pos)
         pos_unique
```

```
Out[10]: array(['ADJ', 'ADP', 'ADV', 'AUX', 'CCONJ', 'DET', 'NOUN', 'NUM', 'PART',
                'PRON', 'PROPN', 'PUNCT', 'SCONJ', 'SPACE', 'VERB'], dtype='<U5')
```

```
In [11]: for pos in pos_unique:
         items = []
         for token in processed_text:
             if token.pos_ == pos:
                 items.append(token.text)
         ITEMS = Counter(items)
         ITEMS = sorted(ITEMS.items(), key=lambda x:x[1], reverse=True)
         print(pos, ITEMS, "\n")
```

ADJ [(('federal', 30), ('racial', 15), ('-', 13), ('Black', 9), ('own', 7), ('local', 6), ('American', 6), ('important', 6), ('sub', 6), ('institutional', 5), ('reparative', 5), ('economic', 4), ('structural', 4), ('culpable', 4), ('effective', 4), ('white', 3), ('enough', 3), ('future', 3), ('historical', 3), ('national', 3), ('political', 2), ('such', 2), ('new', 2), ('only', 2), ('social', 2), ('mass', 2), ('first', 2), ('more', 2), ('legitimate', 2), ('explicit', 2), ('material', 2), ('inappropriate', 2), ('ongoing', 2), ('pandemic', 1), ('public', 1), ('harsh', 1), ('White', 1), ('consistent', 1), ('premature', 1), ('persistent', 1), ('clean', 1), ('foundational', 1), ('several', 1), ('large', 1), ('further', 1), ('various', 1), ('largest', 1), ('certain', 1), ('good', 1), ('mutual', 1), ('oppressive', 1), ('essential', 1), ('indirect', 1), ('potential', 1), ('direct', 1), ('sufficient', 1), ('unpaid', 1), ('current', 1), ('cumulative', 1), ('much', 1), ('possible', 1), ('universal', 1), ('worth', 1), ('disadvantaged', 1), ('aggrieved', 1), ('specific', 1), ('appropriate', 1), ('ethic', 1), ('permanent', 1), ('unjust', 1), ('symbolic', 1), ('successful', 1), ('aware', 1), ('collective', 1), ('long', 1), ('requisite', 1), ('useful', 1)]

ADP [(('of', 44), ('for', 38), ('in', 26), ('to', 19), ('at', 8), ('from', 6), ('by', 5), ('on', 5), ('For', 5), ('as', 4), ('In', 4), ('with', 3), ('within', 3), ('against', 3), ('about', 2), ('around', 2), ('toward', 2), ('like', 2), ('between', 2), ('through', 2), ('than', 2), ('Between', 1), ('among', 1), ('off', 1), ('From', 1), ('across', 1), ('i.e.', 1), ('below', 1), ('under', 1), ('without', 1), ('over', 1), ('At', 1), ('into', 1), ('back', 1), ('throughout', 1), ('out', 1)]

ADV [(('also', 6), ('disproportionately', 2), ('as', 2), ('forward', 2), ('still', 1), ('finally', 1), ('largely', 1), ('never', 1), ('arguably', 1), ('Here', 1), ('currently', 1), ('yet', 1), ('thus', 1), ('far', 1), ('simply', 1), ('more', 1), ('broadly', 1), ('slightly', 1), ('directly', 1), ('likely', 1), ('poorly', 1), ('respectively', 1), ('materially', 1), ('instead', 1), ('explicitly', 1), ('is', 1), ('just', 1), ('ultimately', 1), ('Only', 1), ('similarly', 1), ('Instead', 1), ('rather', 1), ('so', 1), ('often', 1), ('However', 1), ('long', 1), ('most', 1), ('importantly', 1)]

AUX [(('be', 18), ('should', 16), ('is', 15), ('are', 14), ('would', 12), ('being', 8), ('have', 7), ('must', 5), ('can', 3), ('will', 2), ('been', 2), ('do', 2), ('has', 1), ('were', 1), ('having', 1), ('could', 1), ('may', 1)]

CCONJ [(('and', 65), ('or', 7), ('but', 2), ('both', 1), ('nor', 1), ('But', 1)]

DET [(('the', 93), ('a', 31), ('an', 11), ('that', 7), ('The', 6), ('these', 6), ('those', 6), ('any', 3), ('This', 2), ('some', 2), ('this', 2), ('such', 1), ('An', 1), ('Any', 1), ('A', 1), ('These', 1)]

NOUN [(('reparations', 36), ('plans', 20), ('plan', 18), ('government', 10), ('level', 10), ('redress', 8), ('wealth', 7), ('harm', 7), ('state', 6), ('disparities', 6), ('gap', 6), ('harms', 6), ('slavery', 5), ('history', 5), ('policy', 5), ('beneficiaries', 5), ('example', 5), ('descendants', 5), ('injustice', 4), ('awareness', 4), ('atrocities', 4), ('acknowledgement', 4), ('apology', 4), ('chattel', 4), ('city', 4), ('institutions', 3), ('interest', 3), ('party', 3), ('place', 3), ('states', 3), ('localities', 3), ('effort', 3), ('cities', 3), ('criteria', 3), ('closure', 3), ('victims', 3), ('capacity', 3), ('change', 3), ('policies', 3), ('part', 3), ('justice', 3), ('principles', 2), ('aftermath', 2), ('one', 2), ('country', 2), ('incarceration', 2), ('Darity', 2), ('step', 2), ('absence', 2), ('cases', 2), ('style', 2), ('programs', 2), ('people', 2), ('guidelines', 2), ('responsibility', 2), ('commitment', 2), ('future', 2), ('massacres', 2), ('residents', 2), ('housing', 2), ('program', 2), ('labor', 2), ('experience', 2), ('time', 2), ('scope', 2), ('laws', 2), ('parties', 2), ('backlash', 2), ('progress', 2), ('equity', 2), ('am', 1), ('crises', 1), ('nation', 1), ('polycrisis', 1), ('emergence', 1), ('recession', 1), ('police', 1), ('murder', 1), ('truth', 1), ('structure', 1), ('economy', 1), ('supremacy', 1), ('outcome', 1), ('precarity', 1), ('death', 1), ('confluence', 1), ('tragedies', 1), ('condition', 1), ('generation', 1), ('academics', 1), ('policymakers', 1), ('problem', 1), ('activists', 1), ('community', 1), ('organizers', 1), ('attention', 1), ('solution', 1), ('generations', 1), ('book', 1), ('Equality', 1), ('Reparations', 1), ('strategy', 1), ('source', 1), ('inequality', 1), ('domains', 1), ('call', 1), ('grounds', 1), ('institution', 1), ('means', 1), ('transfer', 1), ('scale', 1), ('issue', 1), ('will', 1), ('bill', 1), ('floor', 1), ('vote', 1), ('year', 1), ('commissions', 1), ('auspices', 1), ('mind', 1), ('difference', 1), ('color', 1), ('material', 1), ('understanding', 1), ('systems', 1), ('post', 1), ('addition', 1), ('gaslighting', 1), ('Acknowledgment', 1), ('preconditions', 1), ('reconciliation', 1), ('compensation', 1), ('exclusion', 1), ('transfers', 1), ('slaves', 1), ('amount', 1), ('appreciation', 1), ('fruits', 1), ('effect', 1), ('racism', 1), ('way', 1), ('Plans', 1), ('incident', 1), ('precedent', 1), ('specificity', 1), ('merits', 1), ('fact', 1), ('agenda', 1), ('reparative', 1), ('injustices', 1), ('locality', 1), ('idea', 1), ('reasons', 1), ('feasibility', 1), ('role', 1), ('analysis', 1), ('statistics', 1), ('component', 1), ('account', 1), ('legacy', 1), ('vigilance', 1), ('spirit', 1), ('aspect', 1), ('practice', 1), ('occurrence', 1), ('elements', 1), ('re', 1), ('opening', 1), ('ways', 1), ('commitments', 1), ('reactionaries', 1), ('periods', 1), ('response', 1), ('Policymakers', 1), ('ones', 1), ('recognition', 1), ('substitutes', 1), ('point', 1), ('work', 1), ('term', 1), ('consequences', 1), ('resources', 1), ('challenge', 1), ('conversation', 1), ('framework', 1)]

NUM [(('2020', 2), ('one', 2), ('15', 1), ('2023', 1), ('11:20', 1), ('Five', 1), ('40', 1), ('32', 1), ('three', 1), ('1898', 1), ('1921', 1)]

PART [(('to', 40), ('not', 8), ('s', 6), ('', 1)]

PRON [(('that', 14), ('its', 11), ('their', 8), ('it', 7), ('we', 5), ('what', 4), ('our', 3), ('this', 3), ('I', 3), ('It', 2), ('who', 2), ('those', 2), ('This', 2), ('them', 2), ('We', 1), ('something', 1), ('all', 1), ('they', 1), ('There', 1), ('everyone', 1), ('which', 1)]

PROPN [(('Americans', 9), ('Black', 7), ('Mullen', 6), ('Sub', 6), ('Jim', 4), ('Crow', 4), ('Darity', 4), ('Reconstruction', 2), ('Reparations', 2), ('Evanston', 2), ('Providence', 2), ('February', 1), ('Kyle', 1), ('K.', 1), ('Moore', 1), ('COVID-19', 1), ('George', 1), ('Floyd', 1), ('William', 1), ('Kirsten', 1), ('Twenty', 1), ('First', 1), ('Century', 1), ('America', 1), ('H.R.', 1), ('Commission', 1), ('Develop', 1), ('Proposals', 1), ('African', 1), ('House', 1), ('California', 1), ('Task', 1), ('Force', 1), ('Illinois', 1), ('St.', 1), ('Paul', 1), ('Minnesota', 1), ('Rhode', 1), ('Island', 1), ('Acknowledgement', 1), ('Wilmington', 1), ('North', 1), ('Carolina', 1), ('Tulsa', 1), ('Oklahoma', 1), ('State', 1)]

PUNCT [((',', 67), ('.', 52), ('-', 9), ('-', 6), ('(', 5), (')', 5), (':', 4), ('"', 2), ('"', 2), (';', 2)]

SCONJ [(('that', 12), ('though', 2), ('As', 2), ('for', 2), ('if', 2), ('Despite', 1), ('whether', 1), ('because', 1), ('when', 1), ('If', 1), ('as', 1)]

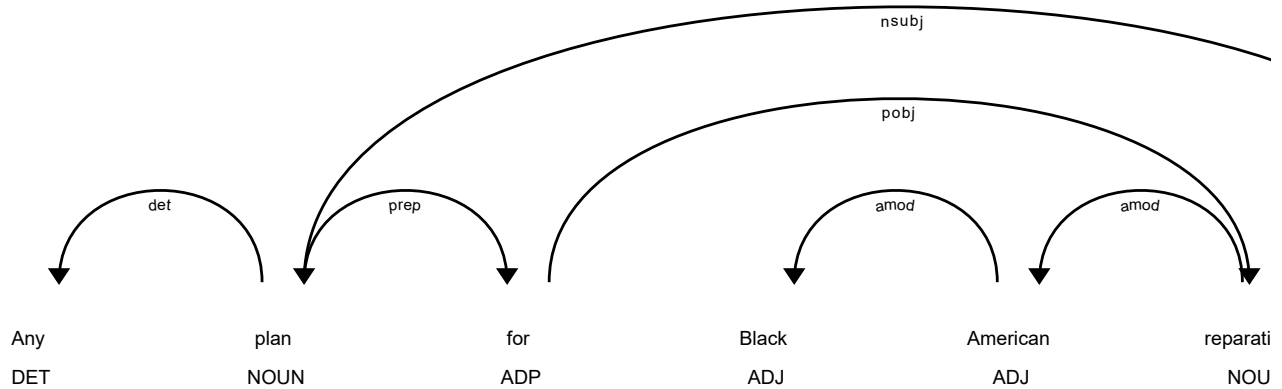
SPACE [(('\n\n', 20), ('\n', 5), ('\n\n\n', 1)]

VERB [(('include', 6), ('committed', 5), ('provide', 5), ('benefit', 5), ('designed', 5), ('making', 4), ('attempt', 4), ('brought', 3), ('lost', 3), ('addressed', 3), ('addressing', 2), ('put', 2), ('implementing', 2), ('close', 2), ('enact', 2), ('rectified', 2), ('raised', 2), ('suggest', 2), ('avoid', 2), ('prevent', 2), ('need', 2), ('excluded', 2), ('compensated', 2), ('specify', 2), ('considered', 2), ('seen', 2), ('Posted', 1), ('reparative', 1), ('living', 1), ('overlapping', 1), ('ensuing', 1), ('saw', 1), ('shaped', 1), ('reinvigorated', 1), ('do', 1), ('renewed', 1), ('outlines', 1), ('allowing', 1), ('exist', 1), ('lacking', 1), ('establish', 1), ('Study', 1), ('passing', 1), ('has', 1), ('convened', 1), ('enacted', 1), ('represents', 1), ('investigate', 1), ('investigated', 1), ('worked', 1), ('creating', 1), ('design', 1), ('keep', 1), ('ensure', 1), ('is', 1), ('identify', 1), ('bring', 1), ('implemented', 1), ('expand', 1), ('attempting', 1), ('create', 1), ('plan', 1), ('aims', 1), ('Specify', 1),

```
(('Stay', 1), ('absolving', 1), ('Commit', 1), ('acknowledge', 1), ('apologize', 1), ('done', 1), ('Attempting', 1), ('rectify', 1), ('speaking', 1), ('lead', 1), ('causing', 1), ('developed', 1), ('centered', 1), ('took', 1), ('suffice', 1), ('needs', 1), ('compensate', 1), ('reflects', 1), ('closing', 1), ('redress', 1), ('have', 1), ('focus', 1), ('outline', 1), ('impacted', 1), ('pursuing', 1), ('benefits', 1), ('consider', 1), ('make', 1), ('provided', 1), ('absolve', 1), ('stay', 1), ('ties', 1), ('circumvent', 1), ('legalizing', 1), ('supporting', 1), ('play', 1), ('grounded', 1), ('taken', 1), ('applies', 1), ('cited', 1), ('developing', 1), ('citing', 1), ('probe', 1), ('commit', 1), ('mean', 1), ('including', 1), ('happening', 1), ('stop', 1), ('serve', 1), ('protect', 1), ('roll', 1), ('achieve', 1), ('expect', 1), ('follow', 1), ('prepared', 1), ('explored', 1), ('represent', 1), ('deserve', 1), ('inflicted', 1), ('reiterate', 1), ('meet', 1), ('achieving', 1))]
```

Find a subject/object relationship through the dependency parser in any sentence.

```
In [12]: sentence = """Any plan for Black American reparations needs to include material redress."""
sentence = nlp(sentence)
spacy.displacy.render(sentence, style="dep")
```



Show the most common Entities and their types.

```
In [13]: count = Counter()
for ent in processed_text.ents:
    count[f"{ent.text} - {ent.label_}"] += 1
count = sorted(count.items(), key=lambda x:x[1], reverse=True)

print(count)

[('Black Americans - NORP', 6), ('Mullen - PERSON', 5), ('Jim Crow - PERSON', 4), ('American - NORP', 3), ('Black American - NORP', 3), ('2020 - DATE', 2), ('Reconstruction - ORG', 2), ('first - ORDINAL', 2), ('Evanston - ORG', 2), ('Providence - GPE', 2), ('February 15, 2023 - DATE', 1), ('11:20 am - TIME', 1), ('Kyle K. Moore - PERSON', 1), ('Five - CARDINAL', 1), ('COVID-19 - ORG', 1), ('George Floyd - PERSON', 1), ('William Darity - PERSON', 1), ('Kirsten Mullen's - PERSON', 1), ('America - GPE', 1), ('H.R. 40 - PERSON', 1), ('African Americans - NORP', 1), ('House - ORG', 1), ('32-year - DATE', 1), ('i.e - NORP', 1), ('California - GPE', 1), ('Reparations Task Force - ORG', 1), ('Illinois - GPE', 1), ('St. Paul - GPE', 1), ('Minnesota - GPE', 1), ('Rhode Island - GPE', 1), ('three - CARDINAL', 1), ('Americans - NORP', 1), ('Wilmington - GPE', 1), ('North Carolina - GPE', 1), ('Tulsa - GPE', 1), ('Oklahoma - GPE', 1), ('1898 - DATE', 1), ('1921 - DATE', 1), ('Darity - ORG', 1), ('one - CARDINAL', 1)]
```

Find Entities and their dependency (hint: entity.root.head)

```
In [14]: for entity in processed_text.ents:
print(entity, " - ", entity.root.head)
```

February 15, 2023 - Posted
 11:20 am - at
 Kyle K. Moore - Five
 Five - principles
 2020 - of
 COVID-19 - of
 George Floyd - of
 American - economy
 Black American - condition
 Reconstruction - slavery
 Jim Crow - Reconstruction
 William Darity - outlines
 Kirsten Mullen's - book
 2020 - book
 America - in
 Mullen - Darity
 first - place
 H.R. 40 - is
 African Americans - for
 first - step
 House - floor
 32-year - history
 i.e - level
 California - Force
 Reparations Task Force - represents
 Evanston - like
 Illinois - Evanston
 St. Paul - Illinois
 Minnesota - Paul
 Providence - Paul
 Rhode Island - Providence
 Mullen - Darity
 three - criteria
 Americans - beneficiaries
 Black Americans - victims
 Wilmington - of
 North Carolina - Wilmington
 Tulsa - Wilmington
 Oklahoma - Tulsa
 1898 - in
 1921 - 1898
 Black American - reparations
 Mullen - Darity
 Black American - slaves
 Black Americans - experience
 Jim Crow - laws
 American - slavery
 Jim Crow - slavery
 Black Americans - be
 Darity - by
 Mullen - Darity
 one - time
 American - history
 Providence - in
 Evanston - Providence
 Black Americans - deserve
 Mullen - Darity
 Reconstruction - slavery
 Jim Crow - Reconstruction
 Black Americans - for
 Black Americans - for

Find the most similar words in the article

```

In [84]: nlp_lg = spacy.load('en_core_web_lg')

doc = nlp_lg(text)

allowed_pos = {"NOUN", "VERB", "ADJ", "ADV"}

similarities = []
seen_words = set()

for token1 in doc:
    for token2 in doc:
        if (
            token1.pos_ in allowed_pos
            and token2.pos_ in allowed_pos
            and token1.text != token2.text
            and token1.lemma_ != token2.lemma_
            and token2.lemma_ not in seen_words
        ):
            similarity = token1.similarity(token2)
            seen_words.add(token1.lemma_)
            if similarity > 0.8:

```

```

        similarities.append((token1.text, token2.text, similarity))

similarities.sort(key=lambda x: x[2], reverse=True)

comparisons = []
for token1, token2, similarity in similarities:
    new = str(token1) + " and " + str(token2) + " have similarity: " + str(similarity)
    comparisons.append(new)
print(np.unique(comparisons))

```

C:\Users\rmjol\AppData\Local\Temp\ipykernel_6100\1510305296.py:19: UserWarning: [W008] Evaluating Token.similarity based on empty vectors.

```

    similarity = token1.similarity(token2)
['Reparations and preconditions have similarity: 0.8225531578063965'
'acknowledgement and Acknowledgment have similarity: 0.948468029499054'
'acknowledgement and acknowledge have similarity: 0.855330228805542'
'acknowledgement and recognition have similarity: 0.8127683997154236'
'closing and opening have similarity: 0.8243698477745056'
'compensation and compensated have similarity: 0.8224815130233765'
'component and elements have similarity: 0.8131995797157288'
'condition and preconditions have similarity: 0.86434006690979'
'economic and economy have similarity: 0.8494637608528137'
'explicit and explicitly have similarity: 0.864477276802063'
'foundational and institutional have similarity: 0.8164942860603333'
'implementing and developing have similarity: 0.8269399404525757'
'implementing and supporting have similarity: 0.8147599101066589'
'important and essential have similarity: 0.8057520985603333'
'important and importantly have similarity: 0.8322049379348755'
'indirect and direct have similarity: 0.8343405723571777'
'injustice and justice have similarity: 0.816523551940918'
'institution and institutional have similarity: 0.9061031937599182'
'institutions and commissions have similarity: 0.8001397252082825'
'institutions and institutional have similarity: 0.9142367243766785'
'nation and country have similarity: 0.8008629679679871'
'nation and national have similarity: 0.8400529623031616'
'reparations and compensation have similarity: 0.8135786652565002'
'reparations and preconditions have similarity: 0.8251110315322876'
'reparations and reconciliation have similarity: 0.8034963011741638'
'reparations and reparative have similarity: 0.8499647974967957'
'several and various have similarity: 0.8329846262931824'
'specificity and specific have similarity: 0.8078683614730835'
'specify and specific have similarity: 0.8148546814918518'
'structure and structural have similarity: 0.8873292803764343']

```

