

Movie Project Visualize

December 8, 2017

```
In [1]: from bs4 import BeautifulSoup, NavigableString, Tag
        from datascience import *
        from collections import Counter
```

```
In [2]: data = Table.read_table('scripts_metadata.csv')
        data.show(5)
```

<IPython.core.display.HTML object>

```
In [3]: data = data.where('script_path', are.not_equal_to('nan'))
        data.show(5)
```

<IPython.core.display.HTML object>

```
In [4]: all_scripts = {}
```

```
    for fname in data['script_path']:

        print(fname)
        with open(fname, 'r') as f:
            raw = f.read()
            soup = BeautifulSoup(raw, 'html5lib')

        try:
            bolded = soup.find('td', {'class': 'scrtext'}).find_all('b') #find all bolded
            text = soup.find('td', {'class': 'scrtext'}).text
            b_text = [b.text.strip() for b in bolded]
            bolded_text = [b for b in b_text if len(b) > 0]
            sift_out = ['INT.', "EXT.", "-", "CONTINUED", ":", "PAN", "MUSIC"]
            characters = []
            scenes = []
            for c in bolded_text:
                character = True
                for s in sift_out:
                    if s in c:
                        character = False
```

```

        if character == True:
            characters.append(c)
        elif len(c) > 4:
            scenes.append(c)

characters = [c[0] for c in Counter(characters).most_common() if c[1] > 5]

scenes.extend([c[0] for c in Counter(characters).most_common() if c[1] <= 5 and c[0] != ''])

movie_name = fname.split('/')[1][:-5].replace(' Script', '')

all_scripts[movie_name] = {}
all_scripts[movie_name]['cast'] = characters
all_scripts[movie_name]['scenes'] = scenes
all_scripts[movie_name]['text'] = text

except:
    pass

```

```

scripts/10 Things I Hate About You Script.html
scripts/12 Script.html
scripts/12 and Holding Script.html
scripts/12 Monkeys Script.html
scripts/12 Years a Slave Script.html
scripts/127 Hours Script.html
scripts/1492: Conquest of Paradise Script.html
scripts/15 Minutes Script.html
scripts/17 Again Script.html
scripts/187 Script.html
scripts/2001: A Space Odyssey Script.html
scripts/2012 Script.html
scripts/25th Hour Script.html
scripts/30 Minutes or Less Script.html
scripts/42 Script.html
scripts/44 Inch Chest Script.html
scripts/48 Hrs. Script.html
scripts/50-50 Script.html
scripts/500 Days of Summer Script.html
scripts/8 Mile Script.html
scripts/8MM Script.html
scripts/9 Script.html
scripts/A Few Good Men Script.html
scripts/A Most Violent Year Script.html
scripts/A Serious Man Script.html
scripts/Above the Law Script.html
scripts/Absolute Power Script.html
scripts/Abyss, The Script.html

```

scripts/Ace Ventura: Pet Detective Script.html
scripts/Adaptation Script.html
scripts/Addams Family, The Script.html
scripts/Adjustment Bureau, The Script.html
scripts/Adventures of Buckaroo Banzai Across the Eighth Dimension, The Script.html
scripts/Affliction Script.html
scripts/After School Special Script.html
scripts/After.Life Script.html
scripts/Agnes of God Script.html
scripts/Air Force One Script.html
scripts/Airplane Script.html
scripts/Airplane 2: The Sequel Script.html
scripts/Aladdin Script.html
scripts/Ali Script.html
scripts/Alien Script.html
scripts/Alien 3 Script.html
scripts/Alien Nation Script.html
scripts/Alien vs. Predator Script.html
scripts/Alien: Resurrection Script.html
scripts/Aliens Script.html
scripts/All About Eve Script.html
scripts/All About Steve Script.html
scripts/All the King's Men Script.html
scripts/All the President's Men Script.html
scripts/Almost Famous Script.html
scripts/Alone in the Dark Script.html
scripts/Amadeus Script.html
scripts/Amelia Script.html
scripts/American Beauty Script.html
scripts/American Gangster Script.html
scripts/American Graffiti Script.html
scripts/American History X Script.html
scripts/American Hustle Script.html
scripts/American Madness Script.html
scripts/American Outlaws Script.html
scripts/American Pie Script.html
scripts/American President, The Script.html
scripts/American Psycho Script.html
scripts/American Shaolin: King of Kickboxers II Script.html
scripts/American Sniper Script.html
scripts/American Splendor Script.html
scripts/American Werewolf in London Script.html
scripts/American, The Script.html
scripts/Amityville Asylum, The Script.html
scripts/Amour Script.html
scripts/An Education Script.html
scripts/Analyze That Script.html
scripts/Analyze This Script.html

scripts/Anastasia Script.html
scripts/Angel Eyes Script.html
scripts/Angels & Demons Script.html
scripts/Anna Karenina Script.html
scripts/Annie Hall Script.html
scripts/Anniversary Party, The Script.html
scripts/Anonymous Script.html
scripts/Antitrust Script.html
scripts/Antz Script.html
scripts/Apartment, The Script.html
scripts/Apocalypse Now Script.html
scripts/Apollo 13 Script.html
scripts/April Fool's Day Script.html
scripts/Apt Pupil Script.html
scripts/Arbitrage Script.html
scripts/Arcade Script.html
scripts/Arctic Blue Script.html
scripts/Argo Script.html
scripts/Armageddon Script.html
scripts/Army of Darkness Script.html
scripts/Arsenic and Old Lace Script.html
scripts/Arthur Script.html
scripts/Artist, The Script.html
scripts/As Good As It Gets Script.html
scripts/Assassins Script.html
scripts/Assignment, The Script.html
scripts/At First Sight Script.html
scripts/August: Osage County Script.html
scripts/Austin Powers - International Man of Mystery Script.html
scripts/Austin Powers - The Spy Who Shagged Me Script.html
scripts/Authors Anonymous Script.html
scripts/Autumn in New York Script.html
scripts/Avatar Script.html
scripts/Avengers, The Script.html
scripts/Avengers, The (2012) Script.html
scripts/Avventura, L' (The Adventure) Script.html
scripts/Awakenings Script.html
scripts/Babel Script.html
scripts/Bachelor Party Script.html
scripts/Bachelor Party, The Script.html
scripts/Back to the Future Script.html
scripts/Back to the Future II & III Script.html
scripts/Back-up Plan, The Script.html
scripts/Backdraft Script.html
scripts/Bad Boys Script.html
scripts/Bad Country Script.html
scripts/Bad Day at Black Rock Script.html
scripts/Bad Dreams Script.html

scripts/Bad Lieutenant Script.html
scripts/Bad Santa Script.html
scripts/Bad Teacher Script.html
scripts/Badlands Script.html
scripts/Bamboozled Script.html
scripts/Barry Lyndon Script.html
scripts/Barton Fink Script.html
scripts/Basic Script.html
scripts/Basic Instinct Script.html
scripts/Basquiat Script.html
scripts/Batman Script.html
scripts/Batman 2 Script.html
scripts/Battle of Algiers, The Script.html
scripts/Battle of Shaker Heights, The Script.html
scripts/Battle: Los Angeles Script.html
scripts/Beach, The Script.html
scripts/Bean Script.html
scripts/Beasts of the Southern Wild Script.html
scripts/Beavis and Butt-head Do America Script.html
scripts/Beginners Script.html
scripts/Being Human Script.html
scripts/Being John Malkovich Script.html
scripts/Being There Script.html
scripts/Believer, The Script.html
scripts/Belle Script.html
scripts/Beloved Script.html
scripts/Benny & Joon Script.html
scripts/Best Exotic Marigold Hotel, The Script.html
scripts/Big Script.html
scripts/Big Blue, The Script.html
scripts/Big Eyes Script.html
scripts/Big Fish Script.html
scripts/Big Lebowski, The Script.html
scripts/Big White, The Script.html
scripts/Birdman Script.html
scripts/Birds, The Script.html
scripts/Birthday Girl Script.html
scripts/Black Dahlia, The Script.html
scripts/Black Rain Script.html
scripts/Black Snake Moan Script.html
scripts/Black Swan Script.html
scripts/Blade Script.html
scripts/Blade II Script.html
scripts/Blade Runner Script.html
scripts/Blade: Trinity Script.html
scripts/Blast from the Past, The Script.html
scripts/Blind Side, The Script.html
scripts/Bling Ring, The Script.html

scripts/Blood and Wine Script.html
scripts/Blood Simple Script.html
scripts/Blow Script.html
scripts/Blue Valentine Script.html
scripts/Blue Velvet Script.html
scripts/Blues Brothers, The Script.html
scripts/Bodies, Rest & Motion Script.html
scripts/Body Heat Script.html
scripts/Body of Evidence Script.html
scripts/Bodyguard Script.html
scripts/Bones Script.html
scripts/Bonfire of the Vanities Script.html
scripts/Bonnie and Clyde Script.html
scripts/Boogie Nights Script.html
scripts/Book of Eli, The Script.html
scripts/Boondock Saints 2: All Saints Day Script.html
scripts/Boondock Saints, The Script.html
scripts/Bottle Rocket Script.html
scripts/Bound Script.html
scripts/Bounty Hunter, The Script.html
scripts/Bourne Identity, The Script.html
scripts/Bourne Supremacy, The Script.html
scripts/Bourne Ultimatum, The Script.html
scripts/Box, The Script.html
scripts/Boyhood Script.html
scripts/Braveheart Script.html
scripts/Brazil Script.html
scripts/Break Script.html
scripts/Breakdown Script.html
scripts/Breakfast Club, The Script.html
scripts/Breaking Away Script.html
scripts/Brick Script.html
scripts/Bridesmaids Script.html
scripts/Bringing Out the Dead Script.html
scripts/Broadcast News Script.html
scripts/Broken Arrow Script.html
scripts/Broken Embraces Script.html
scripts/Brothers Bloom, The Script.html
scripts/Bruce Almighty Script.html
scripts/Buffy the Vampire Slayer Script.html
scripts/Bull Durham Script.html
scripts/Buried Script.html
scripts/Burlesque Script.html
scripts/Burn After Reading Script.html
scripts/Burning Annie Script.html
scripts/Butterfly Effect, The Script.html
scripts/Cable Guy Script.html
scripts/Candle to Water Script.html

scripts/Capote Script.html
scripts/Carrie Script.html
scripts/Cars 2 Script.html
scripts/Case 39 Script.html
scripts/Casino Script.html
scripts/Cast Away Script.html
scripts/Catch Me If You Can Script.html
scripts/Catwoman Script.html
scripts/Cecil B. Demented Script.html
scripts/Cedar Rapids Script.html
scripts/Celeste & Jesse Forever Script.html
scripts/Cell, The Script.html
scripts/Cellular Script.html
scripts/Change-Up, The Script.html
scripts/Changeling Script.html
scripts/Chaos Script.html
scripts/Charade Script.html
scripts/Charlie's Angels Script.html
scripts/Chasing Amy Script.html
scripts/Chasing Sleep Script.html
scripts/Cherry Falls Script.html
scripts/Chinatown Script.html
scripts/Christ Complex Script.html
scripts/Chronicle Script.html
scripts/Chronicles of Narnia: The Lion, the Witch and the Wardrobe Script.html
scripts/Cider House Rules, The Script.html
scripts/Cincinnati Kid, The Script.html
scripts/Cinema Paradiso Script.html
scripts/Cirque du Freak: The Vampire's Assistant Script.html
scripts/Citizen Kane Script.html
scripts/City of Joy Script.html
scripts/Clash of the Titans Script.html
scripts/Clerks Script.html
scripts/Cliffhanger Script.html
scripts/Clockwork Orange, A Script.html
scripts/Clueless Script.html
scripts/Cobb Script.html
scripts/Code of Silence Script.html
scripts/Cold Mountain Script.html
scripts/Collateral Script.html
scripts/Collateral Damage Script.html
scripts/Colombiana Script.html
scripts/Color of Night Script.html
scripts/Commando Script.html
scripts/Conan the Barbarian Script.html
scripts/Confessions of a Dangerous Mind Script.html
scripts/Confidence Script.html
scripts/Constantine Script.html

scripts/Cooler, The Script.html
scripts/Copycat Script.html
scripts/Coraline Script.html
scripts/Coriolanus Script.html
scripts/Courage Under Fire Script.html
scripts/Cowboys & Aliens Script.html
scripts/Cradle 2 the Grave Script.html
scripts/Crank Script.html
scripts/Crash Script.html
scripts/Crazy, Stupid, Love Script.html
scripts/Crazylove Script.html
scripts/Creation Script.html
scripts/Crime Spree Script.html
scripts/Croods, The Script.html
scripts/Crouching Tiger, Hidden Dragon Script.html
scripts/Croupier Script.html
scripts/Crow Salvation, The Script.html
scripts/Crow, The Script.html
scripts/Crow: City of Angels, The Script.html
scripts/Cruel Intentions Script.html
scripts/Crying Game Script.html
scripts/Cube Script.html
scripts/Curious Case of Benjamin Button, The Script.html
scripts/Custody Script.html
scripts/Dallas Buyers Club Script.html
scripts/Damned United, The Script.html
scripts/Dances with Wolves Script.html
scripts/Dark City Script.html
scripts/Dark Knight Rises, The Script.html
scripts/Dark Star Script.html
scripts/Darkman Script.html
scripts/Date Night Script.html
scripts/Dave Barry's Complete Guide to Guys Script.html
scripts/Dawn of the Dead Script.html
scripts/Day of the Dead Script.html
scripts/Day the Clown Cried, The Script.html
scripts/Day the Earth Stood Still, The Script.html
scripts/Days of Heaven Script.html
scripts/Dead Poets Society Script.html
scripts/Deadpool Script.html
scripts/Dear White People Script.html
scripts/Death at a Funeral Script.html
scripts/Death to Smoochy Script.html
scripts/Debt, The Script.html
scripts/Deception Script.html
scripts/Deep Cover Script.html
scripts/Deep Rising Script.html
scripts/Deer Hunter, The Script.html

scripts/Defiance Script.html
scripts/Departed, The Script.html
scripts/Descendants, The Script.html
scripts/Despicable Me 2 Script.html
scripts/Detroit Rock City Script.html
scripts/Devil in a Blue Dress Script.html
scripts/Devil's Advocate Script.html
scripts/Die Hard Script.html
scripts/Die Hard 2 Script.html
scripts/Diner Script.html
scripts/Distinguished Gentleman, The Script.html
scripts/Disturbia Script.html
scripts/Django Unchained Script.html
scripts/Do The Right Thing Script.html
scripts/Dog Day Afternoon Script.html
scripts/Dogma Script.html
scripts/Donnie Brasco Script.html
scripts/Doors, The Script.html
scripts/Double Indemnity Script.html
scripts/Drag Me to Hell Script.html
scripts/Dragonslayer Script.html
scripts/Drive Script.html
scripts/Drive Angry Script.html
scripts/Drop Dead Gorgeous Script.html
scripts/Dry White Season, A Script.html
scripts/Duck Soup Script.html
scripts/Dumb and Dumber Script.html
scripts/Dune Script.html
scripts/E.T. Script.html
scripts/Eagle Eye Script.html
scripts/Eastern Promises Script.html
scripts/Easy A Script.html
scripts/Ed TV Script.html
scripts/Ed Wood Script.html
scripts/Edward Scissorhands Script.html
scripts/Eight Legged Freaks Script.html
scripts/El Mariachi Script.html
scripts/Election Script.html
scripts/Elephant Man, The Script.html
scripts/Elizabeth: The Golden Age Script.html
scripts/Enemy of the State Script.html
scripts/English Patient, The Script.html
scripts/Enough Script.html
scripts/Entrapment Script.html
scripts/Equilibrium Script.html
scripts/Erik the Viking Script.html
scripts/Erin Brockovich Script.html
scripts/Escape From L.A. Script.html

scripts/Escape From New York Script.html
scripts/Eternal Sunshine of the Spotless Mind Script.html
scripts/Even Cowgirls Get the Blues Script.html
scripts/Event Horizon Script.html
scripts/Evil Dead Script.html
scripts/Evil Dead II: Dead by Dawn Script.html
scripts/Ex Machina Script.html
scripts/Excalibur Script.html
scripts/eXistenZ Script.html
scripts/Extract Script.html
scripts/Fabulous Baker Boys, The Script.html
scripts/Face Off Script.html
scripts/Fair Game Script.html
scripts/Family Man, The Script.html
scripts/Fantastic Four Script.html
scripts/Fantastic Mr Fox Script.html
scripts/Fargo Script.html
scripts/Fast Times at Ridgemont High Script.html
scripts/Fatal Instinct Script.html
scripts/Fault in Our Stars, The Script.html
scripts/Fear and Loathing in Las Vegas Script.html
scripts/Feast Script.html
scripts/Ferris Bueller's Day Off Script.html
scripts/Field of Dreams Script.html
scripts/Fifth Element, The Script.html
scripts/Fight Club Script.html
scripts/Fighter, The Script.html
scripts/Final Destination Script.html
scripts/Final Destination 2 Script.html
scripts/Finding Nemo Script.html
scripts/Five Easy Pieces Script.html
scripts/Flash Gordon Script.html
scripts/Fletch Script.html
scripts/Flight Script.html
scripts/Flintstones, The Script.html
scripts/Forrest Gump Script.html
scripts/Four Feathers Script.html
scripts/Four Rooms Script.html
scripts/Foxcatcher Script.html
scripts/Fracture Script.html
scripts/Frances Script.html
scripts/Frankenstein Script.html
scripts/Frankenweenie Script.html
scripts/Freaked Script.html
scripts/Freddy vs. Jason Script.html
scripts/French Connection, The Script.html
scripts/Frequency Script.html
scripts/Friday the 13th Script.html

scripts/Friday the 13th Part VIII: Jason Takes Manhattan Script.html
scripts/Fright Night Script.html
scripts/Fright Night (1985) Script.html
scripts/From Dusk Till Dawn Script.html
scripts/From Here to Eternity Script.html
scripts/Frozen Script.html
scripts/Frozen (Disney) Script.html
scripts/Frozen River Script.html
scripts/Fruitvale Station Script.html
scripts/Fugitive, The Script.html
scripts/Funny People Script.html
scripts/G.I. Jane Script.html
scripts/G.I. Joe: The Rise of Cobra Script.html
scripts/Game 6 Script.html
scripts/Game, The Script.html
scripts/Gamer Script.html
scripts/Gandhi Script.html
scripts/Gang Related Script.html
scripts/Gangs of New York Script.html
scripts/Garden State Script.html
scripts/Gattaca Script.html
scripts/Get Carter Script.html
scripts/Get Low Script.html
scripts/Get on Up Script.html
scripts/Get Shorty Script.html
scripts/Getaway, The Script.html
scripts/Ghost Script.html
scripts/Ghost and the Darkness, The Script.html
scripts/Ghost Rider Script.html
scripts/Ghost Ship Script.html
scripts/Ghost World Script.html
scripts/Ghostbusters Script.html
scripts/Ghostbusters 2 Script.html
scripts/Ginger Snaps Script.html
scripts/Girl with the Dragon Tattoo, The Script.html
scripts/Gladiator Script.html
scripts/Glengarry Glen Ross Script.html
scripts/Go Script.html
scripts/Godfather Script.html
scripts/Godfather Part II Script.html
scripts/Godfather Part III, The Script.html
scripts/Gods and Monsters Script.html
scripts/Godzilla Script.html
scripts/Gone in 60 Seconds Script.html
scripts/Good Girl, The Script.html
scripts/Good Will Hunting Script.html
scripts/Gothika Script.html
scripts/Graduate, The Script.html

scripts/Gran Torino Script.html
scripts/Grand Hotel Script.html
scripts/Grand Theft Parsons Script.html
scripts/Grapes of Wrath, The Script.html
scripts/Gravity Script.html
scripts/Green Mile, The Script.html
scripts/Gremlins Script.html
scripts/Gremlins 2 Script.html
scripts/Grifters, The Script.html
scripts/Grosse Pointe Blank Script.html
scripts/Groundhog Day Script.html
scripts/Grudge, The Script.html
scripts/Hackers Script.html
scripts/Hall Pass Script.html
scripts/Halloween: The Curse of Michael Myers Script.html
scripts/Hancock Script.html
scripts/Hangover, The Script.html
scripts/Hanna Script.html
scripts/Hannah and Her Sisters Script.html
scripts/Hannibal Script.html
scripts/Happy Birthday, Wanda June Script.html
scripts/Happy Feet Script.html
scripts/Hard Rain Script.html
scripts/Hard to Kill Script.html
scripts/Harold and Kumar Go to White Castle Script.html
scripts/Haunting, The Script.html
scripts/He's Just Not That Into You Script.html
scripts/Heat Script.html
scripts/Heathers Script.html
scripts/Heavenly Creatures Script.html
scripts/Heavy Metal Script.html
scripts/Hebrew Hammer, The Script.html
scripts/Heist Script.html
scripts/Hellbound: Hellraiser II Script.html
scripts/Hellboy Script.html
scripts/Hellboy 2: The Golden Army Script.html
scripts/Hellraiser Script.html
scripts/Hellraiser 3: Hell on Earth Script.html
scripts/Hellraiser: Deader Script.html
scripts/Hellraiser: Hellseeker Script.html
scripts/Help, The Script.html
scripts/Henry Fool Script.html
scripts/Henry's Crime Script.html
scripts/Her Script.html
scripts/Hesher Script.html
scripts/High Fidelity Script.html
scripts/Highlander Script.html
scripts/Highlander: Endgame Script.html

scripts/Hills Have Eyes, The Script.html
scripts/His Girl Friday Script.html
scripts/Hitchcock Script.html
scripts/Hitchhiker's Guide to the Galaxy, The Script.html
scripts/Hollow Man Script.html
scripts/Honeydripper Script.html
scripts/Horrible Bosses Script.html
scripts/Horse Whisperer, The Script.html
scripts/Hospital, The Script.html
scripts/Hostage Script.html
scripts/Hot Tub Time Machine Script.html
scripts/Hotel Rwanda Script.html
scripts/House of 1000 Corpses Script.html
scripts/How to Lose Friends & Alienate People Script.html
scripts/How to Train Your Dragon Script.html
scripts/How to Train Your Dragon 2 Script.html
scripts/Hudson Hawk Script.html
scripts/Hudsucker Proxy, The Script.html
scripts/Human Nature Script.html
scripts/Hunt for Red October, The Script.html
scripts/I Am Number Four Script.html
scripts/I am Sam Script.html
scripts/I Love You Phillip Morris Script.html
scripts/I Spit on Your Grave Script.html
scripts/I Still Know What You Did Last Summer Script.html
scripts/I'll Do Anything Script.html
scripts/I, Robot Script.html
scripts/Ice Storm, The Script.html
scripts/Ides of March, The Script.html
scripts/Imaginarium of Doctor Parnassus, The Script.html
scripts/In the Bedroom Script.html
scripts/In the Loop Script.html
scripts/Inception Script.html
scripts/Incredibles, The Script.html
scripts/Independence Day Script.html
scripts/Indiana Jones and the Last Crusade Script.html
scripts/Indiana Jones and the Raiders of the Lost Ark Script.html
scripts/Indiana Jones and the Temple of Doom Script.html
scripts/Indiana Jones IV Script.html
scripts/Informant, The Script.html
scripts/Inglourious Basterds Script.html
scripts/Insider, The Script.html
scripts/Insidious Script.html
scripts/Insomnia Script.html
scripts/Interstellar Script.html
scripts/Interview with the Vampire Script.html
scripts/Into the Wild Script.html
scripts/Into the Woods Script.html

scripts/Intolerable Cruelty Script.html
scripts/Inventing the Abbotts Script.html
scripts/Invention of Lying, The Script.html
scripts/Invictus Script.html
scripts/Iron Lady, The Script.html
scripts/Island, The Script.html
scripts/It Happened One Night Script.html
scripts/It Happened One Night Script.html
scripts/It's a Wonderful Life Script.html
scripts/It's Complicated Script.html
scripts/Italian Job, The Script.html
scripts/Jacket, The Script.html
scripts/Jackie Brown Script.html
scripts/Jacob's Ladder Script.html
scripts/Jade Script.html
scripts/Jane Eyre Script.html
scripts/Jason X Script.html
scripts/Jaws Script.html
scripts/Jaws 2 Script.html
scripts/Jay and Silent Bob Strike Back Script.html
scripts/Jennifer Eight Script.html
scripts/Jennifer's Body Script.html
scripts/Jerry Maguire Script.html
scripts/Jeux Interdits Script.html
scripts/JFK Script.html
scripts/Jimmy and Judy Script.html
scripts/John Q Script.html
scripts/Judge Dredd Script.html
scripts/Juno Script.html
scripts/Jurassic Park Script.html
scripts/Jurassic Park III Script.html
scripts/Jurassic Park: The Lost World Script.html
scripts/Kafka Script.html
scripts/Kalifornia Script.html
scripts/Kate & Leopold Script.html
scripts/Kids Script.html
scripts/Kids Are All Right, The Script.html
scripts/Kill Bill Volume 1 & 2 Script.html
scripts/Kill Your Darlings Script.html
scripts/Killing Zoe Script.html
scripts/King Kong Script.html
scripts/King of Comedy, The Script.html
scripts/King's Speech, The Script.html
scripts/Kingdom, The Script.html
scripts/Kiss of the Spider Woman Script.html
scripts/Klute Script.html
scripts/Knocked Up Script.html
scripts/Kramer vs Kramer Script.html

scripts/Kundun Script.html
scripts/Kung Fu Panda Script.html
scripts/L.A. Confidential Script.html
scripts/La La Land Script.html
scripts/Labor of Love Script.html
scripts/Labyrinth Script.html
scripts/Ladykillers, The Script.html
scripts/Lake Placid Script.html
scripts/Land of the Dead Script.html
scripts/Larry Crowne Script.html
scripts/Last Boy Scout, The Script.html
scripts/Last Chance Harvey Script.html
scripts/Last Flight, The Script.html
scripts/Last of the Mohicans, The Script.html
scripts/Last Samurai, The Script.html
scripts/Last Station, The Script.html
scripts/Last Tango in Paris Script.html
scripts/Law Abiding Citizen Script.html
scripts/Le Diable par la Queue Script.html
scripts/Leaving Las Vegas Script.html
scripts/Legally Blonde Script.html
scripts/Legend Script.html
scripts/Legion Script.html
scripts/Les Miserables Script.html
scripts/Les Tontons Flingueurs Script.html
scripts/Leviathan Script.html
scripts/Liar Liar Script.html
scripts/Life Script.html
scripts/Life As A House Script.html
scripts/Life of David Gale, The Script.html
scripts/Life of Pi Script.html
scripts/Light Sleeper Script.html
scripts/Limey, The Script.html
scripts/Limitless Script.html
scripts/Lincoln Script.html
scripts/Lincoln Lawyer, The Script.html
scripts/Lion King, The Script.html
scripts/Little Athens Script.html
scripts/Little Mermaid, The Script.html
scripts/Little Nicky Script.html
scripts/Living in Oblivion Script.html
scripts/Lock, Stock and Two Smoking Barrels Script.html
scripts/Logan's Run Script.html
scripts/Lone Star Script.html
scripts/Long Kiss Goodnight, The Script.html
scripts/Looper Script.html
scripts/Lord of Illusions Script.html
scripts/Lord of the Rings: Fellowship of the Ring, The Script.html

scripts/Lord of the Rings: Return of the King Script.html
scripts/Lord of the Rings: The Two Towers Script.html
scripts/Lord of War Script.html
scripts/Losers, The Script.html
scripts/Lost Highway Script.html
scripts/Lost Horizon Script.html
scripts/Lost in Space Script.html
scripts/Lost in Translation Script.html
scripts/Love and Basketball Script.html
scripts/Machete Script.html
scripts/Machine Gun Preacher Script.html
scripts/Mad Max 2: The Road Warrior Script.html
scripts/Made Script.html
scripts/Magnolia Script.html
scripts/Majestic, The (The Bijou) Script.html
scripts/Major League Script.html
scripts/Malcolm X Script.html
scripts/Malibu's Most Wanted Script.html
scripts/Man in the Iron Mask Script.html
scripts/Man On Fire Script.html
scripts/Man on the Moon Script.html
scripts/Man Trouble Script.html
scripts/Man Who Knew Too Much, The Script.html
scripts/Man Who Wasn't There, The Script.html
scripts/Manchurian Candidate, The Script.html
scripts/Manhattan Murder Mystery Script.html
scripts/Manhunter Script.html
scripts/Margaret Script.html
scripts/Margin Call Script.html
scripts/Margot at the Wedding Script.html
scripts/Mariachi, El Script.html
scripts/Marley & Me Script.html
scripts/Martha Marcy May Marlene Script.html
scripts/Marty Script.html
scripts/Mary Poppins Script.html
scripts/Mask, The Script.html
scripts/Master and Commander Script.html
scripts/Master, The Script.html
scripts/Matrix Reloaded, The Script.html
scripts/Matrix, The Script.html
scripts/Max Payne Script.html
scripts/Mean Streets Script.html
scripts/Mechanic, The Script.html
scripts/Meet Joe Black Script.html
scripts/Meet John Doe Script.html
scripts/Megamind Script.html
scripts/Memento Script.html
scripts/Men in Black Script.html

scripts/Men in Black 3 Script.html
scripts/Men Who Stare at Goats, The Script.html
scripts/Metro Script.html
scripts/Miami Vice Script.html
scripts/Midnight Cowboy Script.html
scripts/Midnight Express Script.html
scripts/Midnight in Paris Script.html
scripts/Mighty Morphin Power Rangers: The Movie Script.html
scripts/Milk Script.html
scripts/Miller's Crossing Script.html
scripts/Mimic Script.html
scripts/Mini's First Time Script.html
scripts/Minority Report Script.html
scripts/Miracle Worker, The Script.html
scripts/Mirrors Script.html
scripts/Misery Script.html
scripts/Mission Impossible Script.html
scripts/Mission Impossible II Script.html
scripts/Mission to Mars Script.html
scripts/Moneyball Script.html
scripts/Monkeybone Script.html
scripts/Monster's Ball Script.html
scripts/Monte Carlo Script.html
scripts/Moon Script.html
scripts/Moonrise Kingdom Script.html
scripts/Moonstruck Script.html
scripts/Mr Blandings Builds His Dream House Script.html
scripts/Mr Brooks Script.html
scripts/Mr Deeds Goes to Town Script.html
scripts/Mr. Holland's Opus Script.html
scripts/Mrs. Brown Script.html
scripts/Mud Script.html
scripts/Mulan Script.html
scripts/Mulholland Drive Script.html
scripts/Mumford Script.html
scripts/Mummy, The Script.html
scripts/Music of the Heart Script.html
scripts/Mute Witness Script.html
scripts/My Best Friend's Wedding Script.html
scripts/My Girl Script.html
scripts/My Mother Dreams the Satan's Disciples in New York Script.html
scripts/My Week with Marilyn Script.html
scripts/Mystery Men Script.html
scripts/Napoleon Dynamite Script.html
scripts/Nashville Script.html
scripts/Natural Born Killers Script.html
scripts/Never Been Kissed Script.html
scripts/Neverending Story, The Script.html

scripts/New York Minute Script.html
scripts/Newsies Script.html
scripts/Next Script.html
scripts/Next Friday Script.html
scripts/Next Three Days, The Script.html
scripts/Ni vu ni connu Script.html
scripts/Nick of Time Script.html
scripts/Night Time (The Poltergeist Treatment) Script.html
scripts/Nightbreed Script.html
scripts/Nightmare Before Christmas, The Script.html
scripts/Nightmare Before Christmas, The Script.html
scripts/Nightmare on Elm Street, A Script.html
scripts/Nightmare on Elm Street: The Final Chapter Script.html
scripts/Nine Script.html
scripts/Nines, The Script.html
scripts/Ninja Assassin Script.html
scripts/Ninotchka Script.html
scripts/Ninth Gate, The Script.html
scripts/No Country for Old Men Script.html
scripts/No Strings Attached Script.html
scripts/Notting Hill Script.html
scripts/Nurse Betty Script.html
scripts/O Brother Where Art Thou? Script.html
scripts/Oblivion Script.html
scripts/Observe and Report Script.html
scripts/Obsessed Script.html
scripts/Ocean's Eleven Script.html
scripts/Ocean's Twelve Script.html
scripts/Office Space Script.html
scripts/Officer and a Gentleman, An Script.html
scripts/Omega Man Script.html
scripts/One Flew Over the Cuckoo's Nest Script.html
scripts/Only God Forgives Script.html
scripts/Ordinary People Script.html
scripts/Orgy of the Dead Script.html
scripts/Orphan Script.html
scripts/Other Boleyn Girl, The Script.html
scripts/Out of Sight Script.html
scripts/Pacifier, The Script.html
scripts/Pandorum Script.html
scripts/Panic Room Script.html
scripts/Papadopoulos & Sons Script.html
scripts/ParaNorman Script.html
scripts/Pariah Script.html
scripts/Passion of Joan of Arc, The Script.html
scripts/Patriot, The Script.html
scripts/Paul Script.html
scripts/Pearl Harbor Script.html

scripts/Peeping Tom Script.html
scripts/Peggy Sue Got Married Script.html
scripts/Perfect Creature Script.html
scripts/Perfect World, A Script.html
scripts/Perks of Being a Wallflower, The Script.html
scripts/Pet Sematary Script.html
scripts/Pet Sematary II Script.html
scripts/Petulia Script.html
scripts/Philadelphia Script.html
scripts/Phone Booth Script.html
scripts/Pi Script.html
scripts/Pianist, The Script.html
scripts/Piano, The Script.html
scripts/Pineapple Express Script.html
scripts/Pirates of the Caribbean Script.html
scripts/Pirates of the Caribbean: Dead Man's Chest Script.html
scripts/Pitch Black Script.html
scripts/Planet of the Apes, The Script.html
scripts/Platinum Blonde Script.html
scripts/Platoon Script.html
scripts/Pleasantville Script.html
scripts/Point Break Script.html
scripts/Pokemon: Mewtwo Returns Script.html
scripts/Postman, The Script.html
scripts/Power of One, The Script.html
scripts/Precious Script.html
scripts/Predator Script.html
scripts/Pretty Woman Script.html
scripts/Pretty Woman (final script) Script.html
scripts/Pride and Prejudice Script.html
scripts/Priest Script.html
scripts/Princess Bride, The Script.html
scripts/Private Life of Sherlock Holmes, The Script.html
scripts/Producer, The Script.html
scripts/Program, The Script.html
scripts/Prom Night Script.html
scripts/Prometheus Script.html
scripts/Prophecy, The Script.html
scripts/Proposal, The Script.html
scripts/Psycho Script.html
scripts/Public Enemies Script.html
scripts/Pulp Fiction Script.html
scripts/Punch-Drunk Love Script.html
scripts/Purple Rain Script.html
scripts/Quantum Project Script.html
scripts/Queen of the Damned Script.html
scripts/Queen, The Script.html
scripts/Rachel Getting Married Script.html

scripts/Raging Bull Script.html
scripts/Raising Arizona Script.html
scripts/Rambling Rose Script.html
scripts/Rambo: First Blood II: The Mission Script.html
scripts/Reader, The Script.html
scripts/Real Genius Script.html
scripts/Rear Window Script.html
scripts/Rebel Without A Cause Script.html
scripts/Red Planet Script.html
scripts/Red Riding Hood Script.html
scripts/Reindeer Games Script.html
scripts/Relic, The Script.html
scripts/Remember Me Script.html
scripts/Replacements, The Script.html
scripts/Repo Man Script.html
scripts/Rescuers Down Under, The Script.html
scripts/Reservoir Dogs Script.html
scripts/Resident Evil Script.html
scripts/Return of the Apes Script.html
scripts/Revenant, The Script.html
scripts/Revolutionary Road Script.html
scripts/Ringu Script.html
scripts/Rise of the Guardians Script.html
scripts/Rise of the Planet of the Apes Script.html
scripts/RKO 281 Script.html
scripts/Road, The Script.html
scripts/Robin Hood: Prince of Thieves Script.html
scripts/Robocop Script.html
scripts/Rock, The Script.html
scripts/RocknRolla Script.html
scripts/Rocky Script.html
scripts/Rocky Horror Picture Show, The Script.html
scripts/Romeo & Juliet Script.html
scripts/Ronin Script.html
scripts/Room Script.html
scripts/Roommate, The Script.html
scripts/Roughshod Script.html
scripts/Ruins, The Script.html
scripts/Runaway Bride Script.html
scripts/Rush Script.html
scripts/Rush Hour Script.html
scripts/Rush Hour 2 Script.html
scripts/Rushmore Script.html
scripts/Rust and Bone Script.html
scripts/S. Darko Script.html
scripts/Saint, The Script.html
scripts/Salton Sea, The Script.html
scripts/Sandlot Kids, The Script.html

scripts/Save the Last Dance Script.html
scripts/Saving Mr. Banks Script.html
scripts/Saving Private Ryan Script.html
scripts/Saw Script.html
scripts/Scarface Script.html
scripts/Scary Movie 2 Script.html
scripts/Schindler's List Script.html
scripts/Scott Pilgrim vs the World Script.html
scripts/Scream Script.html
scripts/Scream 2 Script.html
scripts/Scream 3 Script.html
scripts/Se7en Script.html
scripts/Searchers, The Script.html
scripts/Semi-Pro Script.html
scripts/Sense and Sensibility Script.html
scripts/Serenity Script.html
scripts/Serial Mom Script.html
scripts/Sessions, The Script.html
scripts/Seventh Seal, The Script.html
scripts/Sex and the City Script.html
scripts/Sex, Lies and Videotape Script.html
scripts/Sexual Life Script.html
scripts/Shadow of the Vampire Script.html
scripts/Shakespeare in Love Script.html
scripts/Shallow Grave Script.html
scripts/Shame Script.html
scripts/Shampoo Script.html
scripts/Shawshank Redemption, The Script.html
scripts/She's Out of My League Script.html
scripts/Sherlock Holmes Script.html
scripts/Shifty Script.html
scripts/Shining, The Script.html
scripts/Shipping News, The Script.html
scripts/Shivers Script.html
scripts/Shrek Script.html
scripts/Shrek the Third Script.html
scripts/Sideways Script.html
scripts/Siege, The Script.html
scripts/Signs Script.html
scripts/Silence of the Lambs Script.html
scripts/Silver Bullet Script.html
scripts/Silver Linings Playbook Script.html
scripts/Simone Script.html
scripts/Single White Female Script.html
scripts/Sister Act Script.html
scripts/Six Degrees of Separation Script.html
scripts/Sixth Sense, The Script.html
scripts/Sleepless in Seattle Script.html

scripts/Sleepy Hollow Script.html
scripts/Sling Blade Script.html
scripts/Slither Script.html
scripts/Slumdog Millionaire Script.html
scripts/Smashed Script.html
scripts/Smokin' Aces Script.html
scripts/Snatch Script.html
scripts/Sneakers Script.html
scripts/Snow Falling On Cedars Script.html
scripts/Snow White and the Huntsman Script.html
scripts/So I Married an Axe Murderer Script.html
scripts/Social Network, The Script.html
scripts/Solaris Script.html
scripts/Soldier Script.html
scripts/Someone To Watch Over Me Script.html
scripts/Something's Gotta Give Script.html
scripts/Source Code Script.html
scripts/South Park Script.html
scripts/Spanglish Script.html
scripts/Spare Me Script.html
scripts/Spartan Script.html
scripts/Speed Script.html
scripts/Speed Racer Script.html
scripts/Sphere Script.html
scripts/Spider-Man Script.html
scripts/St. Elmo's Fire Script.html
scripts/Star Trek Script.html
scripts/Star Trek II: The Wrath of Khan Script.html
scripts/Star Trek: First Contact Script.html
scripts/Star Trek: Generations Script.html
scripts/Star Trek: Nemesis Script.html
scripts/Star Trek: The Motion Picture Script.html
scripts/Star Wars: A New Hope Script.html
scripts/Star Wars: Attack of the Clones Script.html
scripts/Star Wars: Return of the Jedi Script.html
scripts/Star Wars: Revenge of the Sith Script.html
scripts/Star Wars: The Empire Strikes Back Script.html
scripts/Star Wars: The Force Awakens Script.html
scripts/Star Wars: The Phantom Menace Script.html
scripts/Starman Script.html
scripts/Starship Troopers Script.html
scripts/State and Main Script.html
scripts/Station West Script.html
scripts/Stepmom Script.html
scripts/Sting, The Script.html
scripts/Stir of Echoes Script.html
scripts/Storytelling Script.html
scripts/Strange Days Script.html

scripts/Strangers on a Train Script.html
scripts/Stuntman, The Script.html
scripts/Sugar Script.html
scripts/Sugar and Spice Script.html
scripts/Sunset Blvd. Script.html
scripts/Sunshine Cleaning Script.html
scripts/Super 8 Script.html
scripts/Superbad Script.html
scripts/Superfights Script.html
scripts/Supergirl Script.html
scripts/Surfer King, The Script.html
scripts/Surrogates Script.html
scripts/Suspect Zero Script.html
scripts/Sweeney Todd: The Demon Barber of Fleet Street Script.html
scripts/Sweet Hereafter, The Script.html
scripts/Sweet Smell of Success Script.html
scripts/Swingers Script.html
scripts/Swordfish Script.html
scripts/Synecdoche, New York Script.html
scripts/Syriana Script.html
scripts/Take Shelter Script.html
scripts/Taking Lives Script.html
scripts/Taking of Pelham One Two Three, The Script.html
scripts/Taking Sides Script.html
scripts/Talented Mr. Ripley, The Script.html
scripts/Tall in the Saddle Script.html
scripts/Tamara Drewe Script.html
scripts/Taxi Driver Script.html
scripts/Ted Script.html
scripts/Terminator Script.html
scripts/Terminator 2: Judgement Day Script.html
scripts/Terminator Salvation Script.html
scripts/The Rage: Carrie 2 Script.html
scripts/Thelma & Louise Script.html
scripts/There's Something About Mary Script.html
scripts/They Script.html
scripts/Thing, The Script.html
scripts/Things My Father Never Taught Me, The Script.html
scripts/Thirteen Days Script.html
scripts/This Boy's Life Script.html
scripts/This is 40 Script.html
scripts/Thor Script.html
scripts/Three Kings Script.html
scripts/Three Kings (Spoils of War) Script.html
scripts/Three Men and a Baby Script.html
scripts/Three Musketeers, The Script.html
scripts/Thunderbirds Script.html
scripts/Thunderheart Script.html

scripts/Ticker Script.html
scripts/Timber Falls Script.html
scripts/Time Machine, The Script.html
scripts/Tin Cup Script.html
scripts/Tin Men Script.html
scripts/Tinker Tailor Soldier Spy Script.html
scripts/Titanic Script.html
scripts/TMNT Script.html
scripts/To Sleep with Anger Script.html
scripts/Tombstone Script.html
scripts/Tomorrow Never Dies Script.html
scripts/Top Gun Script.html
scripts/Total Recall Script.html
scripts/Tourist, The Script.html
scripts/Toy Story Script.html
scripts/Traffic Script.html
scripts/Training Day Script.html
scripts/Trainspotting Script.html
scripts/Transformers: The Movie Script.html
scripts/Tremors Script.html
scripts/Tristan and Isolde Script.html
scripts/TRON Script.html
scripts/TRON: Legacy Script.html
scripts/Tropic Thunder Script.html
scripts/True Grit Script.html
scripts/True Lies Script.html
scripts/True Romance Script.html
scripts/Truman Show, The Script.html
scripts/Twilight Script.html
scripts/Twilight: New Moon Script.html
scripts/Twin Peaks Script.html
scripts/Twins Script.html
scripts/Two For The Money Script.html
scripts/U Turn Script.html
scripts/Ugly Truth, The Script.html
scripts/Un Singe en Hiver Script.html
scripts/Unbreakable Script.html
scripts/Under Fire Script.html
scripts/Unknown Script.html
scripts/Up Script.html
scripts/Up in the Air Script.html
scripts/Usual Suspects, The Script.html
scripts/V for Vendetta Script.html
scripts/Valkyrie Script.html
scripts/Vanilla Sky Script.html
scripts/Verdict, The Script.html
scripts/Vertigo Script.html
scripts/Very Bad Things Script.html

scripts/Village, The Script.html
scripts/Virtuosity Script.html
scripts/Visitor, The Script.html
scripts/Wag the Dog Script.html
scripts/Walk to Remember, A Script.html
scripts/Walking Tall Script.html
scripts/Wall Street Script.html
scripts/Wall Street: Money Never Sleeps Script.html
scripts/Wall-E Script.html
scripts/Wanted Script.html
scripts/War Horse Script.html
scripts/War of the Worlds Script.html
scripts/Warm Springs Script.html
scripts/Warrior Script.html
scripts/Watchmen Script.html
scripts/Water for Elephants Script.html
scripts/Way Back, The Script.html
scripts/We Own the Night Script.html
scripts/What About Bob? Script.html
scripts/What Lies Beneath Script.html
scripts/When a Stranger Calls Script.html
scripts/When Harry Met Sally Script.html
scripts/While She Was Out Script.html
scripts/Whistleblower, The Script.html
scripts/White Christmas Script.html
scripts/White Jazz Script.html
scripts/White Ribbon, The Script.html
scripts/White Squall Script.html
scripts/Whiteout Script.html
scripts/Who Framed Roger Rabbit? Script.html
scripts/Who's Your Daddy Script.html
scripts/Wild At Heart Script.html
scripts/Wild Bunch, The Script.html
scripts/Wild Hogs Script.html
scripts/Wild Things Script.html
scripts/Wild Things: Diamonds in the Rough Script.html
scripts/Wild Wild West Script.html
scripts/Willow Script.html
scripts/Win Win Script.html
scripts/Wind Chill Script.html
scripts/Withnail and I Script.html
scripts/Witness Script.html
scripts/Wizard of Oz, The Script.html
scripts/Wolf of Wall Street, The Script.html
scripts/Wonder Boys Script.html
scripts/Woodsman, The Script.html
scripts/World is not Enough, The Script.html
scripts/Wrestler, The Script.html

```

scripts/X-Files: Fight the Future, The Script.html
scripts/X-Men Script.html
scripts/X-Men Origins: Wolverine Script.html
scripts/xXx Script.html
scripts/Year One Script.html
scripts/Yes Man Script.html
scripts/You Can Count On Me Script.html
scripts/You've Got Mail Script.html
scripts/Youth in Revolt Script.html
scripts/Zero Dark Thirty Script.html
scripts/Zerophilia Script.html

```

```
In [5]: all_scripts.keys()
```

```
Out[5]: dict_keys(['10 Things I Hate About You', '12', '12 and Holding', '12 Monkeys', '12 Year
```

```
In [6]: len(all_scripts.keys())
```

```
Out[6]: 1121
```

```

In [7]: import numpy as np
        import networkx as nx
        from lxml import etree
        import itertools
        from datascience import *
        import matplotlib.pyplot as plt
        from nltk.util import ngrams
        import re

def make_graph(c_dict):
    """
    This function accepts a dictionary with number of lines and scenes to create a
    NetworkX graph object
    """
    # setup graph object
    G = nx.Graph()

    # add nodes with attributes of number of lines and scenes
    for c in c_dict.keys():
        G.add_node(
            c,
            scenes=c_dict[c]
            #scenes=c_dict[c]['scenes']
        )

    # make edges by iterating over all combinations of nodes
    for (node1, data1), (node2, data2) in itertools.combinations(G.nodes(data=True), 2):

```

```

# count scenes together by getting union of their sets
scenes_together = len(set(data1['scenes']) & set(data2['scenes']))

if scenes_together:
    # add more weight for more scenes together
    G.add_edge(node1, node2, weight=scenes_together)

return G

```

```

In [8]: def gini(array):
    """Calculate the Gini coefficient of a numpy array."""
    # https://github.com/oliviaguest/gini
    array = np.sort(array) # values must be sorted
    index = np.arange(1, array.shape[0] + 1) # index per array element
    n = array.shape[0] # number of array elements
    return ((np.sum((2 * index - n - 1) * array)) / (n * np.sum(array))) #Gini coefficient

In [9]: titles = []
    gini_array = []
    good_movie_array = []
    bad_movie_array = []
    for i in all_scripts.keys():
        try:
            #print(i)
            title = i
            good_movie_array = np.append(good_movie_array, i)
            scene_index_list = []
            count = 0
            for scene in set(all_scripts[i]['scenes']):
                indices = [m.start() for m in re.finditer(scene, all_scripts[i]['text'])]
                scene_index_list.extend(indices)
            scene_texts = []
            for n in ngrams(sorted(scene_index_list), 2):
                scene_texts.append(all_scripts[i]['text'][n[0]:n[1]])
            cast_dict = {}
            for c in all_scripts[i]['cast']:
                cast_dict[c] = []
                for i, scene in enumerate(scene_texts):
                    if scene.count(c) > 0:
                        cast_dict[c].append(i)

            G = make_graph(cast_dict)

            network_tab = Table()
            network_tab.append_column(label="Characters", values=[c for c in sorted(cast_dict.keys())])
            dc = [x[1] for x in sorted(nx.degree_centrality(G).items(), key=lambda x: x[0])]
            network_tab.append_column(label="Degree Centrality", values=dc)
            bc = [x[1] for x in sorted(nx.betweenness_centrality(G).items(), key=lambda x:

```

```

network_tab.append_column(label="Betweenness Centrality", values=bc)
ec = [x[1] for x in sorted(nx.eigenvector_centrality(G).items(), key=lambda x:
network_tab.append_column(label="Eigenvector Centrality", values=ec)
network_tab.append_column(label="Title", values=title)
#network_tab.show(5)
#print('The Gini Coefficient of this movie is', gini(network_tab.column('Eigen
gini_array = np.append(gini_array, gini(network_tab.column('Eigenvector Centra
titles = np.append(titles, title)
except:
    bad_movie_array = np.append(bad_movie_array, i)
pass

```

```

/srv/app/venv/lib/python3.6/site-packages/ipykernel_launcher.py:16: DeprecationWarning: generate
app.launch_new_instance()
/srv/app/venv/lib/python3.6/site-packages/ipykernel_launcher.py:7: RuntimeWarning: invalid val
import sys

```

```
In [10]: len(gini_array)
```

```
Out[10]: 909
```

```
In [11]: len(titles)
```

```
Out[11]: 909
```

```

In [12]: movie_script_title = []
        for i in titles:
            thing = str(i) + ' Script'
            movie_script_title = np.append(movie_script_title, thing)

        len(movie_script_title)

```

```
Out[12]: 909
```

```
In [13]: gini_table = Table().with_column("Title", movie_script_title).with_column('Gini', gini
gini_table.num_rows
```

```
Out[13]: 909
```

```

In [14]: joined_table = data.join('title', gini_table, "Title")
        joined_table.show(200)

```

```
<IPython.core.display.HTML object>
```

```
In [15]: final_table = joined_table.sort('Gini', descending=True).exclude(np.arange(0,20))
```

```
In [16]: genres = make_array('Action', 'Adventure', 'Animation', 'Comedy', 'Crime', 'Drama', 'I
```

```

In [17]: genre_ginis = []
        for i in genres:
            table = final_table.where('Genres', are.containing(i))
            ginis = table.column('Gini')
            new_gini = gini(ginis)
            genre_ginis = np.append(genre_ginis, new_gini)

In [18]: genre_gini_averages = []
        for i in genres:
            table = final_table.where('Genres', are.containing(i))
            ginis = table.column('Gini')
            average = np.mean(ginis)
            genre_gini_averages = np.append(genre_gini_averages, average)

In [19]: final_table_two = Table().with_column('Genres', genres).with_column('Variation', genres)

In [20]: final_table_two.show()

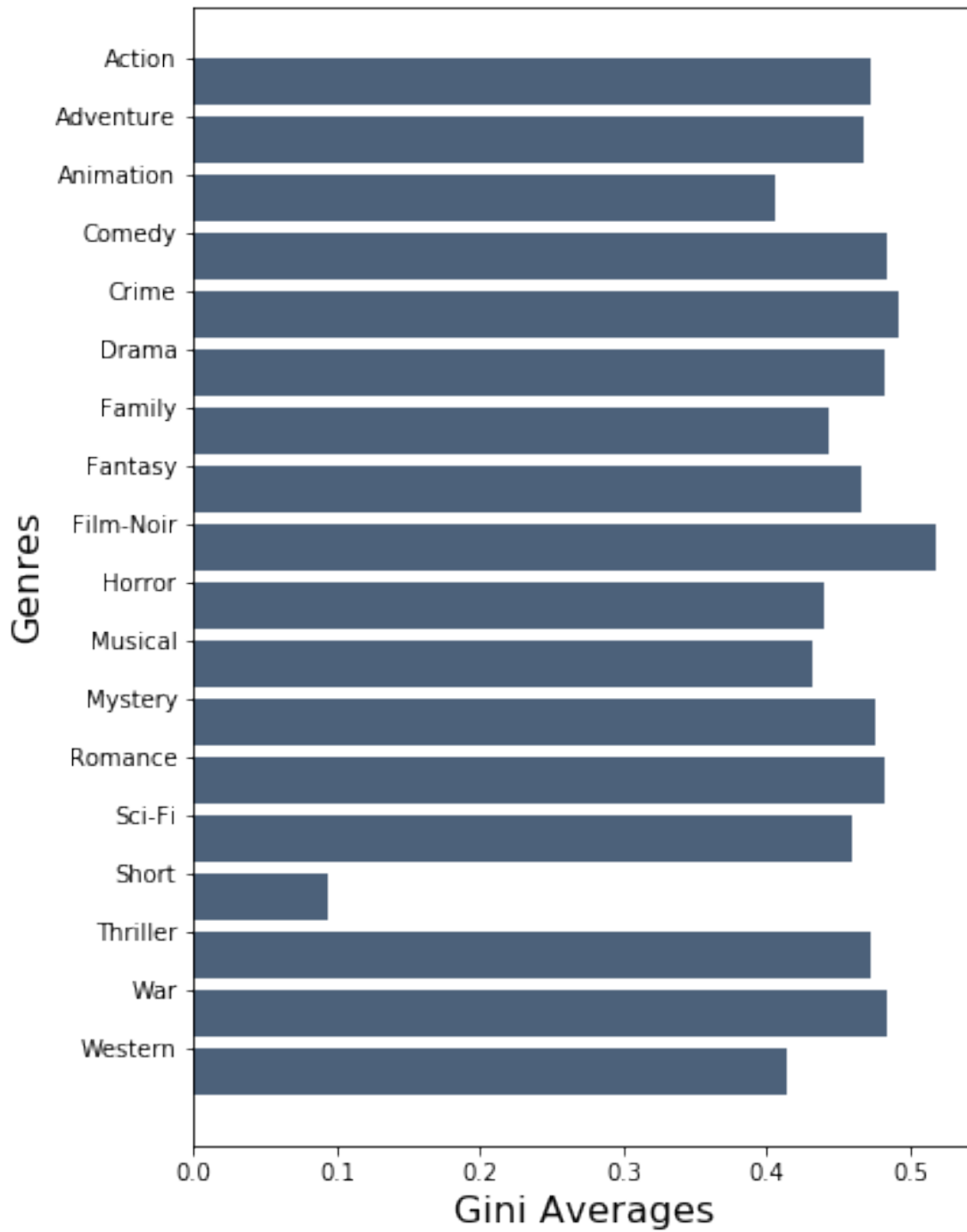
<IPython.core.display.HTML object>

In [21]: import matplotlib.pyplot as plt

        %matplotlib inline

        final_table_two.barh("Genres", "Gini Averages")

```



```
In [22]: final_table.show(200)
```

```
<IPython.core.display.HTML object>
```

```
In [23]: int(final_table.column('Script Date').item(0).replace(': January ', ''))
```

Out[23]: 2001

In [24]: len(final_table.column('Script Date'))

Out[24]: 891

```
In [25]: new_dates = []
        for i in final_table.column('Script Date'):
            if ":" in i:
                new_dates = np.append(new_dates, i.replace(":", " "))
            else:
                new_dates = np.append(new_dates, i)
```

In [26]: new_dates

```
Out[26]: array(['January 2001', 'nan', 'March 1975', 'nan', 'April 1985',
                'February 1998', 'nan', ' 1997', 'September 2000', 'May 2003',
                'May 1998', 'nan', 'March 1939', 'January 1976', 'nan', 'June 1997',
                'February 1989', 'nan', 'May 2009', 'March 2002', 'nan',
                'March 2000', 'April 2009', 'June 1998', 'nan', 'March 1996', 'nan',
                'nan', 'nan', ' 1986', 'May 2012', 'February 2004', 'nan',
                'April 2010', 'nan', 'May 1949', 'nan', 'nan', 'June 2002',
                'May 1997', 'nan', 'June 1979', 'nan', 'October 2004', 'nan', 'nan',
                'nan', 'nan', 'nan', 'nan', 'November 1986', 'May 2003',
                'March 1999', ' 1982', 'July 1985', 'February 2002', 'nan',
                'November 1998', 'October 1995', 'February 1972', 'December 1975',
                'March 2004', ' 1992', 'nan', 'nan', 'October 1999', 'March 1998',
                'nan', 'March 1988', 'June 2008', 'nan', 'June 1994',
                'December 2001', 'nan', 'nan', 'October 1996', 'June 2007', 'nan',
                'March 1962', 'July 1991', 'nan', 'nan', 'January 1992', ' 1975',
                'nan', 'July 2009', 'March 2004', 'January 1979', 'August 2003',
                'nan', 'January 1992', 'February 1994', 'August 1997', ' 1979',
                'March 2007', 'March 2007', 'July 1991', 'February 2008',
                'October 2008', 'February 2009', 'nan', 'nan', 'nan', 'nan', 'nan',
                'September 2007', ' 1991', 'October 2007', 'nan', 'nan',
                'December 1981', 'December 1976', 'July 1992', 'nan', 'June 1995',
                'April 1991', 'August 2006', 'nan', 'October 1987', 'March 2013',
                'nan', 'September 2004', 'July 1998', 'July 1993', 'December 2001',
                'September 2008', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan',
                'May 1993', 'September 2007', 'June 1998', 'nan', 'September 1997',
                'nan', 'nan', 'June 1990', 'nan', 'January 2004', 'nan',
                'October 2007', 'nan', 'February 1990', 'nan', 'nan', 'nan',
                ' 1969', 'November 1993', 'August 1995', 'February 2003', 'nan',
                'August 2004', 'nan', 'February 1998', 'October 1997',
                'January 2001', 'nan', 'nan', 'November 2007', 'February 1998',
                'September 1990', 'September 2007', 'nan', 'February 1982', 'nan',
                'nan', 'November 2000', 'nan', 'September 2007', 'September 1990',
                ' 1999', 'nan', 'July 2006', 'July 1990', 'July 1996', 'nan',
                'March 1994', 'November 1997', 'nan', 'September 1973', 'nan',
```

'September 2010', 'March 2008', 'September 2009', 'August 1988',
 'nan', 'nan', 'nan', 'March 1998', 'August 2001', 'December 2009',
 'July 2007', 'nan', 'nan', 'nan', 'nan', 'April 2008',
 'February 1999', 'January 2008', 'April 2000', 'nan', ' ' 1997',
 'nan', 'December 2011', 'October 1996', 'February 1951',
 'October 1992', 'nan', 'nan', 'nan', 'May 2004', 'nan', 'May 2011',
 'October 1999', 'March 1967', 'nan', 'nan', 'May 2012',
 'February 1981', 'nan', 'June 2006', ' ' 1993', 'nan', 'nan', 'nan',
 'nan', 'nan', 'nan', ' ' 1997', 'November 2005', 'August 1994', 'nan',
 'December 1998', 'February 2007', 'February 1992', 'nan', 'nan',
 'September 1990', 'March 2009', 'nan', 'nan', 'February 2007',
 'August 2009', 'May 2009', 'October 2007', 'nan', 'nan', 'nan',
 ' ' 1998', 'August 2011', 'nan', 'September 2007', 'nan', 'nan',
 'nan', 'nan', 'March 1995', 'March 2008', 'nan', 'nan', ' ' 1984',
 'November 2007', 'December 1994', 'May 2004', 'nan',
 'February 2008', 'nan', 'August 2008', 'nan', 'January 1997',
 'May 2007', 'nan', 'nan', 'nan', 'July 1990', 'December 1983',
 'March 2000', 'nan', ' ' 1999', 'February 2008', 'March 1936', 'nan',
 'March 1999', 'May 2012', 'April 1996', 'February 2004', 'nan',
 'nan', 'nan', 'April 1985', 'July 1995', 'nan', 'September 1998',
 'August 2007', 'July 2008', 'May 1995', 'nan', 'July 2006', ' ' 1998',
 'nan', 'February 2003', 'January 2007', 'February 1997',
 'November 1998', 'March 1995', 'nan', 'February 2002', 'May 2010',
 'nan', 'nan', 'nan', 'August 2007', 'November 2007', 'July 2006',
 'nan', 'nan', 'nan', 'nan', ' ' 1997', 'August 2009', 'nan',
 'July 2012', 'November 2006', 'February 2003', 'nan', 'nan',
 'September 1995', 'August 2009', 'October 1994', 'nan',
 'August 1999', 'nan', 'nan', 'July 2009', 'November 2006',
 'July 1996', 'nan', 'November 1986', 'nan', 'September 2011',
 'February 2007', ' ' 1999', 'November 1999', 'March 1989',
 'October 2011', 'nan', 'nan', 'nan', 'October 2007',
 'December 1994', 'March 2004', 'nan', 'January 1983',
 'September 2011', 'April 1997', 'April 2000', 'July 2000',
 'March 2004', 'November 2009', 'nan', 'August 1993', 'nan',
 'June 2007', 'July 2010', 'nan', 'July 1997', 'nan', 'nan', ' ' 1998',
 'nan', 'August 2007', 'January 1999', 'May 2002', 'December 2006',
 'July 2012', 'January 2000', 'nan', 'nan', 'nan', 'nan', 'nan',
 'November 1997', 'January 1993', 'July 1998', 'February 1995',
 'nan', 'June 1999', 'nan', 'May 2007', 'nan', 'nan', 'nan', 'nan',
 'April 2009', 'nan', 'nan', 'March 2006', 'nan', 'nan', 'nan',
 'February 1998', 'nan', 'nan', 'nan', 'September 2007', 'May 1997',
 'nan', 'July 2000', 'February 2010', ' ' 1987', 'October 2005',
 'October 2009', 'nan', 'nan', 'March 1998', 'nan', ' ' 1996', 'nan',
 'April 2004', 'February 2002', 'May 2001', 'September 2007',
 'February 2005', ' ' 1987', ' ' 1991', 'December 2006', 'nan',
 'February 2002', 'nan', 'January 1997', 'March 1997', 'nan',
 'July 2012', 'June 1991', 'May 2008', 'nan', 'nan', 'February 2007',
 'May 1994', 'nan', 'nan', 'May 1997', 'May 1994', 'nan', 'nan',

'nan', 'January 2001', 'nan', ' 1993', 'nan', 'nan', ' 1973',
 'November 1999', 'May 2001', 'nan', 'nan', 'March 2012', 'nan',
 'nan', 'April 1983', ' 1993', 'nan', 'February 1992', 'July 2007',
 'nan', 'January 1996', 'June 2007', 'nan', 'nan', 'June 1995',
 'December 1953', 'nan', 'nan', ' 1991', 'nan', 'nan', 'nan', 'nan',
 'February 1989', 'June 2002', 'December 2008', 'May 1994', 'nan',
 'nan', 'nan', 'August 2005', 'nan', 'nan', 'nan', 'nan', 'nan',
 'nan', 'June 1988', 'nan', 'January 2010', 'July 2000', 'nan',
 'nan', 'July 2008', 'March 2003', 'nan', ' 1995', 'nan', 'nan',
 'nan', 'nan', 'nan', 'March 2004', ' 1997', 'December 2009',
 'October 1990', ' 1994', 'nan', 'September 1987', 'September 2006',
 'July 1991', 'July 2002', 'nan', 'March 1999', 'nan', 'nan', 'nan',
 'nan', ' 1998', 'June 1997', ' 1995', 'nan', 'nan', 'July 2006',
 'June 2009', ' 1991', 'March 1983', 'nan', 'nan', 'nan', 'nan',
 'September 1992', 'nan', 'nan', 'October 1997', 'July 1997', 'nan',
 'nan', 'nan', 'November 2006', 'nan', 'March 1971', 'nan',
 'October 1962', 'nan', 'nan', 'May 1999', 'August 2002',
 'December 2000', 'nan', 'January 1991', 'December 2006',
 'August 2000', 'nan', 'nan', 'April 1981', 'August 2004', 'nan',
 'March 1999', 'nan', 'March 1994', 'nan', 'nan', 'November 1992',
 'April 1975', 'nan', 'September 2013', 'nan', ' 1999', ' 1998',
 ' 1997', 'June 2006', 'nan', 'nan', 'April 2005', 'nan', 'nan',
 'nan', 'November 1990', 'nan', 'nan', 'October 2009', 'May 1998',
 'nan', 'nan', 'March 1997', 'nan', 'nan', 'nan', 'February 1996',
 'January 2007', 'nan', 'June 2006', 'July 2005', 'January 2011',
 ' 1995', 'June 1996', 'November 1987', 'February 1999',
 'September 2010', 'January 1976', 'February 2000', 'April 2008',
 'August 1952', 'nan', 'nan', 'March 1996', 'nan', 'nan', 'nan',
 'March 2007', 'nan', 'September 1998', 'nan', 'July 1995',
 'October 1995', 'October 2009', 'nan', ' 1978', ' 1997', 'nan',
 'nan', 'nan', 'August 2009', ' 1996', 'May 2011', 'nan',
 'December 1995', 'March 1993', 'June 1998', 'nan', 'October 2003',
 'October 2009', 'May 1985', 'August 1995', 'January 1999',
 'February 1993', 'nan', 'March 1995', 'June 1947', 'February 1988',
 'January 2005', 'nan', 'nan', 'nan', ' 1989', 'May 2012', 'nan',
 'nan', 'February 2009', 'July 1953', 'January 2002', 'October 1990',
 'June 1997', 'December 1992', ' 1996', 'December 2011', 'May 2011',
 'June 1986', 'September 1995', 'November 1992', 'April 2004', 'nan',
 'nan', 'nan', 'June 2007', 'nan', 'April 2004', 'nan', 'July 2008',
 'nan', 'nan', 'nan', 'September 2007', 'nan', 'September 2012',
 'August 2003', 'nan', 'January 2010', 'nan', 'May 1996', ' 1981',
 'December 1959', 'July 2002', 'October 1950', 'nan', ' 1999', 'nan',
 'November 1996', 'November 1997', 'nan', 'nan', 'September 2004',
 'nan', ' 1981', 'February 1998', 'January 1997', 'March 2000',
 'April 1985', 'nan', 'November 2011', 'November 1990', 'nan',
 'November 2009', 'December 1983', 'April 2009', 'nan', ' 1997',
 'nan', 'nan', 'nan', 'June 1994', 'June 2008', 'February 2000',
 'March 1987', 'October 1988', 'nan', 'nan', 'June 1990',

```

'September 1999', 'June 2008', 'nan', 'nan', 'February 2005',
'November 1995', 'May 2009', 'October 2007', 'nan', 'nan', 'nan',
'November 1997', 'nan', 'nan', 'October 2007', 'nan', 'August 1985',
'May 2012', 'nan', 'nan', 'nan', 'April 1999', 'nan', ' 1989',
'March 2008', 'May 2004', 'March 1967', 'July 2009', 'nan',
'June 2003', 'November 2009', 'April 1999', 'nan', 'nan',
'September 1991', ' 1997', 'nan', 'March 1997', 'May 2004', 'nan',
'October 1998', 'March 1979', 'January 2010', 'November 2015',
'May 1994', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan',
'September 1998', 'October 2001', 'April 1992', 'February 2003',
'nan', 'October 2007', 'January 2012', 'nan', 'September 1984',
'nan', 'nan', 'August 1991', 'May 1997', 'December 1986', 'nan',
'nan', 'nan', 'nan', 'nan', 'nan', 'October 2000', 'nan', 'nan',
'December 2012', 'September 2007', 'September 1996', 'nan', 'nan',
' 1985', 'March 1994', 'January 2007', 'March 1998',
'February 2006', 'nan', 'nan', 'nan', 'February 1999', 'nan',
'January 1992', 'nan', 'January 1996', 'nan', 'October 1994',
'January 1999', 'March 2004', 'nan', 'April 2009', 'nan', 'nan',
'nan', 'December 1994', 'nan', ' 1981', 'nan', 'nan', 'nan', 'nan',
'January 2001', 'April 1991', 'January 1989', 'January 1991',
'January 1987', 'December 2012', ' 1991', 'nan', 'July 1992', 'nan',
'July 1980', 'nan', 'March 2000', 'November 2002', 'nan',
'June 1977', 'nan', 'March 1981', 'August 2006', 'June 2007',
'October 1997', 'October 2009', 'nan', 'nan', 'nan', 'January 2001',
'August 1991', 'August 1991', 'nan', 'January 2004', 'May 2014',
'September 1987', 'nan', ' 1986', ' 2000', 'April 1981', 'nan',
'May 2003', 'nan', 'June 1974', 'nan', 'nan', 'nan', 'January 1991',
'nan', ' 1996', ' 1991', 'nan', 'May 1988', 'nan', 'nan',
'June 1978', 'nan', ' 1998', 'nan', 'nan'],
dtype='<U32')

```

```

In [27]: better_dates = []
for i in new_dates:
    if "January " in i:
        better_dates = np.append(better_dates, i.replace('January ', ""))
    elif "February " in i:
        better_dates = np.append(better_dates, i.replace("February ", ""))
    elif "March " in i:
        better_dates = np.append(better_dates, i.replace("March ", ""))
    elif "April " in i:
        better_dates = np.append(better_dates, i.replace("April ", ""))
    elif "May " in i:
        better_dates = np.append(better_dates, i.replace("May ", ""))
    elif "June " in i:
        better_dates = np.append(better_dates, i.replace("June ", ""))
    elif "July " in i:
        better_dates = np.append(better_dates, i.replace("July ", ""))
    elif "August " in i:

```

```

        better_dates = np.append(better_dates, i.replace("August ", ""))
    elif "September " in i:
        better_dates = np.append(better_dates, i.replace("September ", ""))
    elif "October " in i:
        better_dates = np.append(better_dates, i.replace("October ", ""))
    elif "November " in i:
        better_dates = np.append(better_dates, i.replace("November ", ""))
    elif "December " in i:
        better_dates = np.append(better_dates, i.replace("December ", ""))
    else:
        better_dates = np.append(better_dates, i)

```

```

In [28]: corrected_dates = final_table.with_column("better dates", better_dates)
almost_finished = corrected_dates.where('better dates', are.not_containing('nan'))
int_dates = almost_finished.apply(int, 'better dates')
year_table = almost_finished.with_column('Year', int_dates).select('title', 'Genres',
year_table.show(20)

```

<IPython.core.display.HTML object>

```

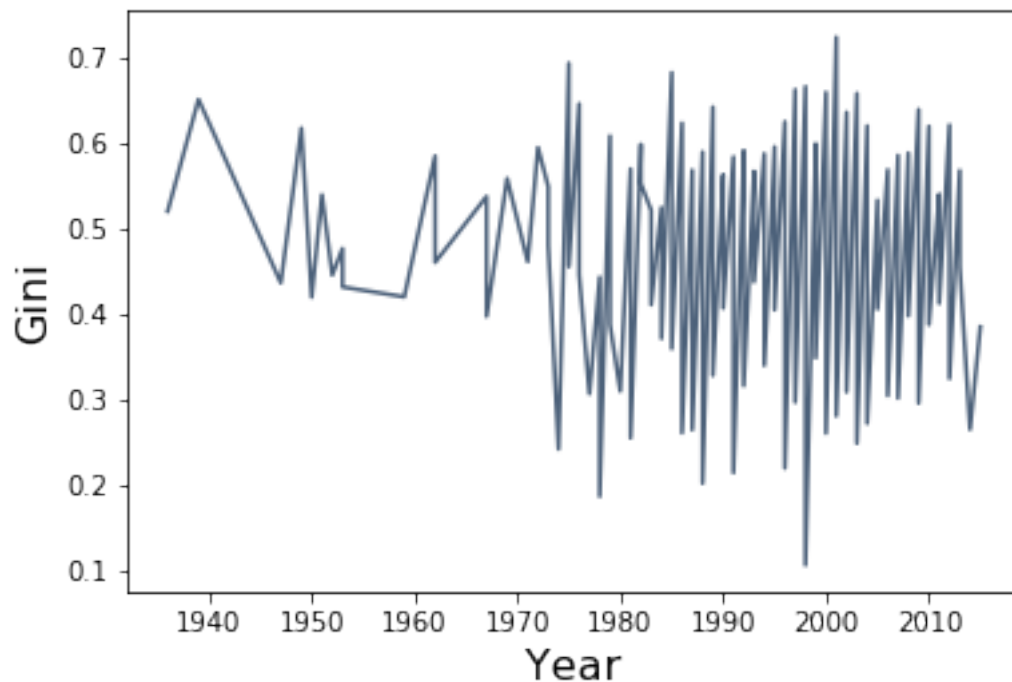
In [29]: better_year = year_table.sort('Year')

```

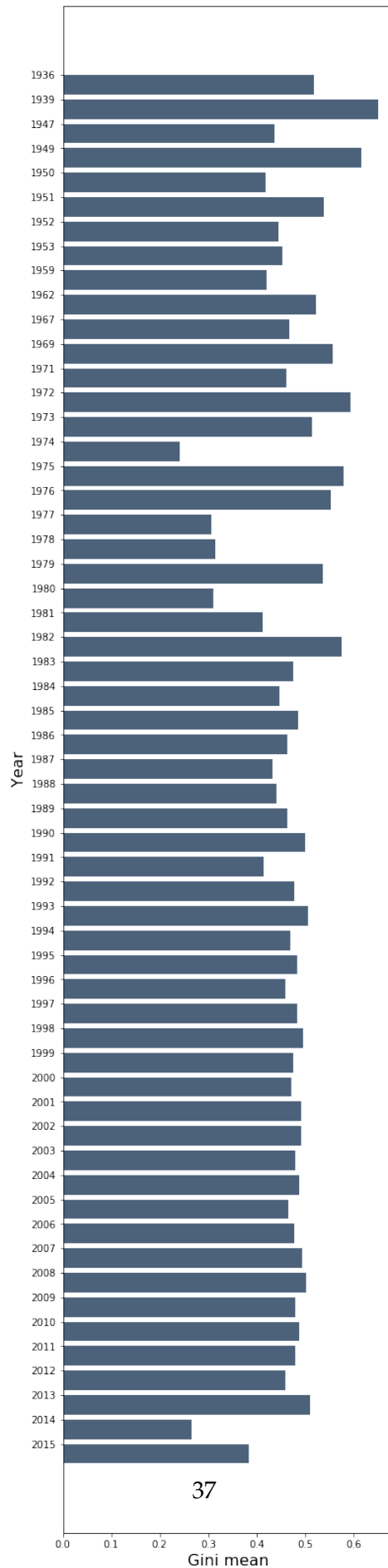
```

In [30]: better_year.plot('Year', "Gini")

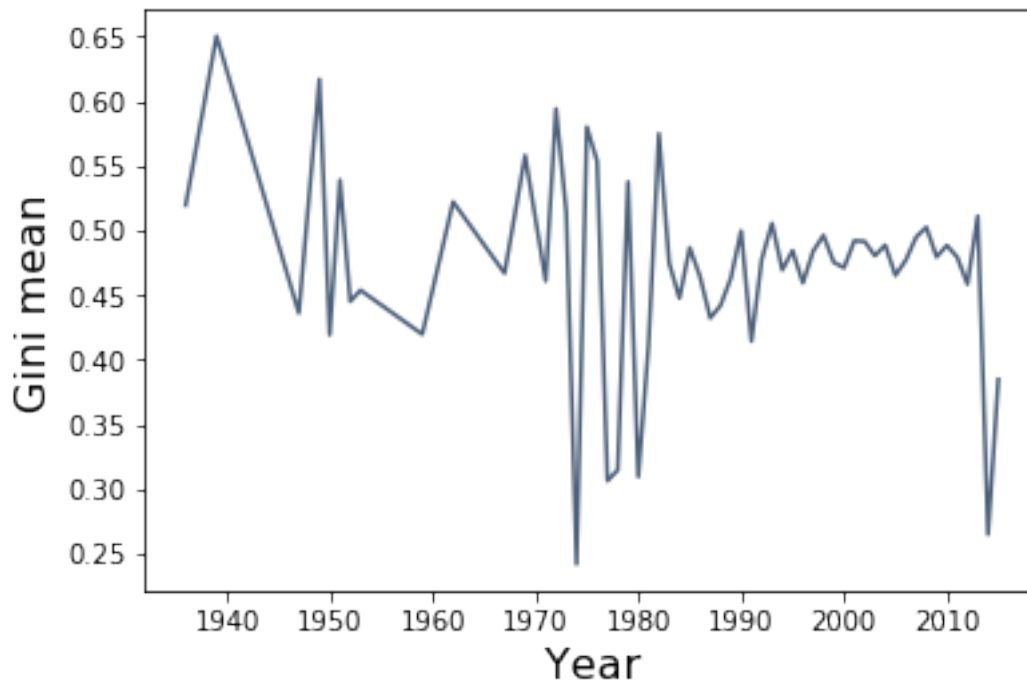
```



```
In [31]: year_avg_gini = better_year.group("Year", collect=np.mean).select('Year', "Gini mean")
        year_avg_gini.barh("Year", "Gini mean")
```



```
In [32]: year_avg_gini.plot('Year')
```



```
In [33]: int_and_str_years = almost_finished.with_column('Year', int_dates)
decades = ['193', '194', '195', '196', '197', '198', '199', '200', '201']
gini_by_decade = []

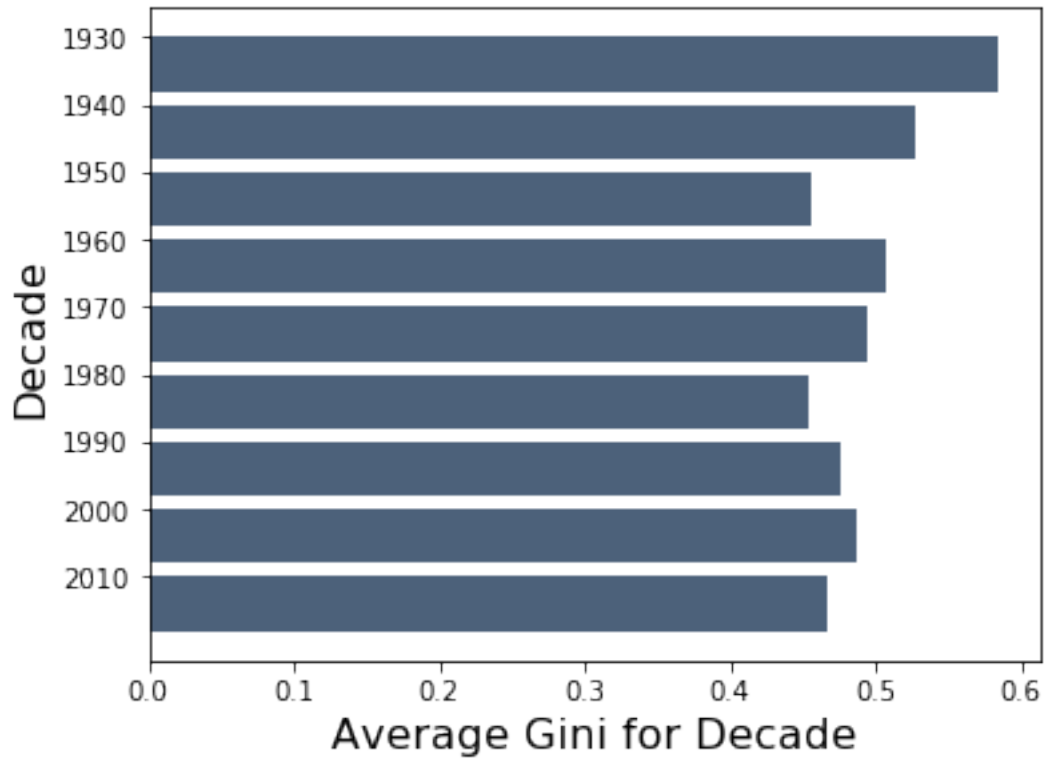
for i in decades:
    avg_gini = np.mean(int_and_str_years.where('better dates', are.containing(i)).col)
    gini_by_decade = np.append(gini_by_decade, avg_gini)

In [34]: pretty_decades = [1930, 1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010]
decade_table = Table().with_column("Decade", pretty_decades).with_column("Average Gini", gini_by_decade)
decade_table
```

```
Out[34]: Decade | Average Gini for Decade
1930      | 0.58501
1940      | 0.526555
1950      | 0.455203
1960      | 0.507279
1970      | 0.494391
1980      | 0.453156
1990      | 0.476261
```

```
2000 | 0.486237
2010 | 0.467694
```

```
In [35]: decade_table.barh('Decade')
```

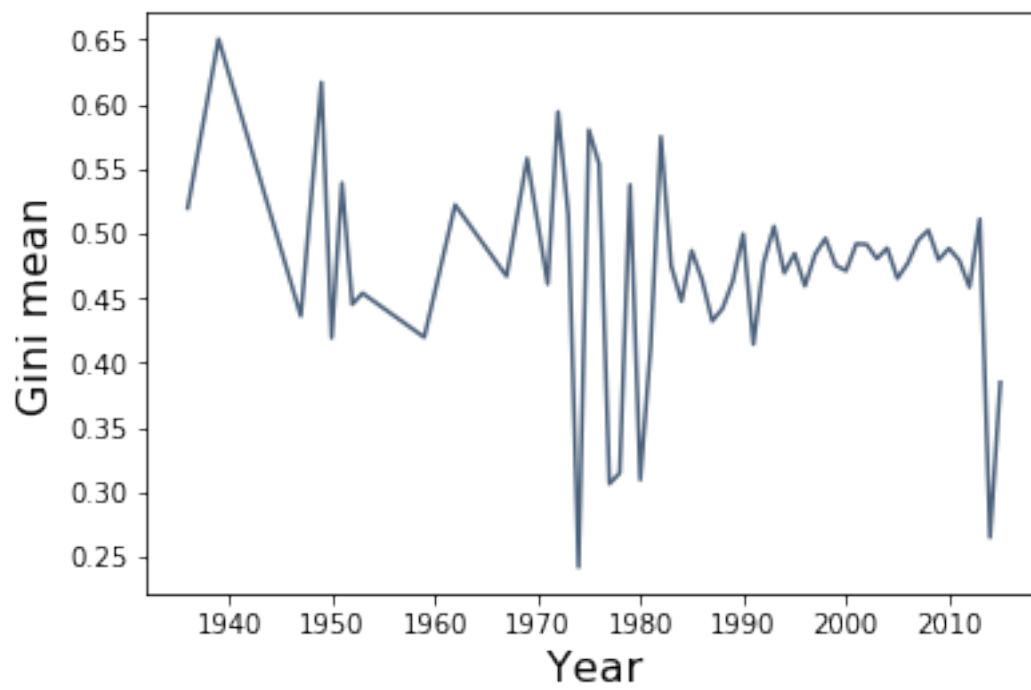


```
In [36]: from pylab import *
         decade_table.plot('Decade')
         ylim([0.42,0.6])
```

```
Out [36]: (0.42, 0.6)
```



```
In [37]: year_avg_gini.plot('Year', "Gini mean")
```




```

In [38]: #better_year.pivot('Year', 'Genres', values="Gini", collect=np.mean)

In [39]: #year_table.sort('Year').barh("Year", "Gini")

In [ ]: correct_titles = []

        for item in list(good_movie_array):
            if item not in list(bad_movie_array):
                correct_titles = np.append(correct_titles, item)

        len(correct_titles)

In [ ]: list(correct_titles).index('Magnolia')

In [ ]: good = Table().with_column('good', good_movie_array)
        bad = Table().with_column('bad', bad_movie_array)

        joined = good.join('good', bad, 'bad')
        joined.num_rows

In [ ]: table = Table().with_column('Title', good_movie_array)
        correct = table.where(table.column(0), are.not_contained_in(bad_movie_array))
        correct.num_rows

In [ ]: correct_data = data.where('title', are.not_contained_in(movie_script_title_bad))
        correct_data.num_rows

In [ ]: genres = make_array('Action', 'Adventure', 'Animation', 'Comedy', 'Crime', 'Drama', 'F

In [ ]: len(genres)

```