

Counting and Apportioning Legal History with Natural Language Processing

Does
law
evolve?

Supreme
Court
Opinions,
1791-2019

Strategy
and
Methods

Findings of
Fact and
Conclusions
of Law

Next
Steps

2019/08/01
LORA JOHNS

Topic Modeling the Supreme Court Opinion Corpus

- This project used unsupervised learning to track the Supreme Court's evolving Fourteenth Amendment jurisprudence from its enactment to the present.
- I processed all 64,000+ of the Court's opinions and wrote custom NLP models to account for the unique challenges of working with legal language over a long time period.
- The models extracted topics that accurately reflect the history of the Amendment, correctly correlate with other highly relevant domain knowledge, and show that the Court's interpretation has not remained static over time.

Why do
this?

"The issue presented is whether the Federal Constitution confers a fundamental right upon homosexuals to engage in sodomy, and hence invalidates the laws of the many States that still make such conduct illegal, and have done so for a very long time. "

Bowers v. Hardwick, 478 U.S. 186 (1986)
(White, J.)

Quantifying the
Constitution's change in
meaning through the words
of the Supreme Court

"The petitioners are entitled to respect for their private lives. The State cannot demean their existence or control their destiny by making their private sexual conduct a crime. Their right to liberty under the Due Process Clause gives them the full right to engage in their conduct without intervention of the government."

"Bowers was not correct when it was decided, and it is not correct today. It ought not to remain binding precedent. Bowers v. Hardwick should be and now is overruled."

Lawrence v. Texas, 539 U.S. 558 (2003)
(Kennedy, J.)

Counting and Apportioning Legal History with Natural Language Processing

Does
law
evolve?

Supreme
Court
Opinions,
1791-2019

Strategy
and
Methods

Findings of
Fact and
Conclusions
of Law

Next
Steps

2019/08/01
LORA JOHNS

The Data

- Web scraping and API mining from CourtListener, Oyez, the Case Law Access Project, and the Supreme Court Database yielded >64,000 opinions and metadata
- The long time horizon, the vast quantity of text, and the complexities of legal language posed unique challenges for preprocessing

Scraping

- Scrapy
- DjangoSpider
- API
- BeautifulSoup
- SQL

Organizing

- Django schemata
- NoSQL database structures
- JSON

Processing

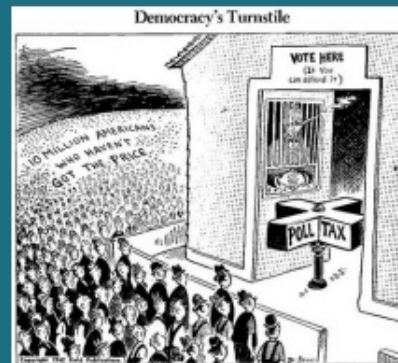
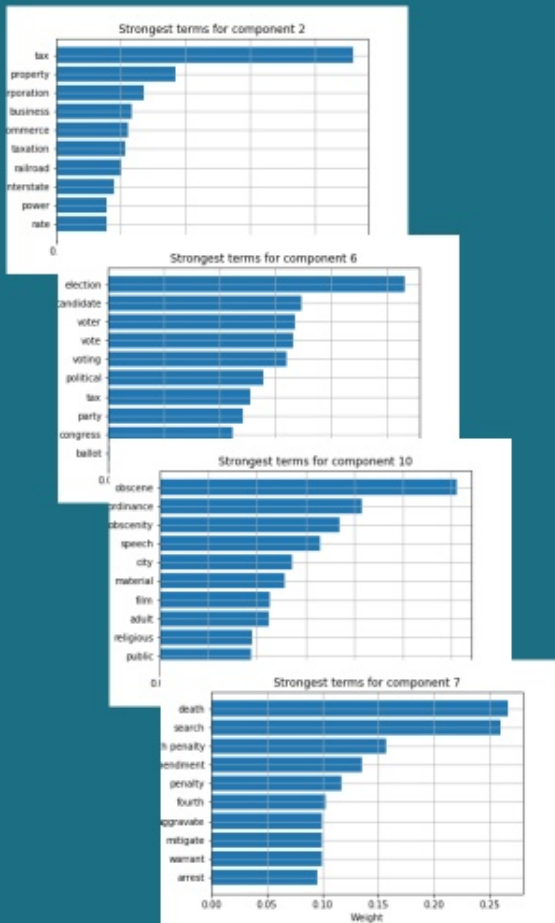
- Custom sklearn pipelines with LexNLP
- OOP customization of function classes

Notate
Bene

Using the
Corpus

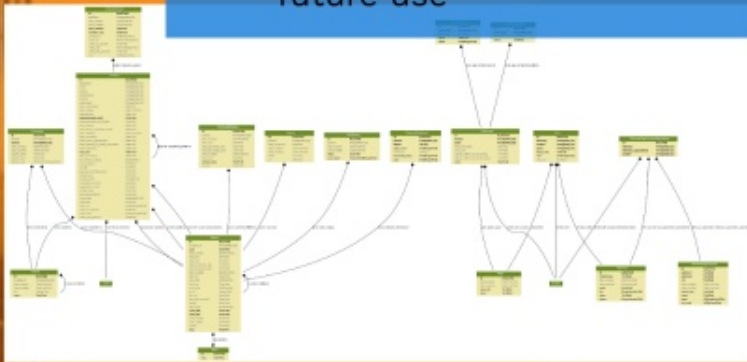
Latent Semantic Analysis

Unsupervised learning using LSA on the corpus as a whole produced coherent topics.



Data Structures

- Original underlying relational structure was poorly designed, resulting in null keys and corrupted API output
- Corrected, connected, and joined 128,000+ such files for future use



Legal Citations

```
# CLEAN OUT HEADS, APPENDICES AND TAILS
# Preliminary regex

def find_targets(string):
    amendments = re.compile("([first|second|third|fourth|fifth|sixth|seventh|eighth|ninth|tenth|eleventh|twelfth|thirteenth|fourteenth|fifteenth|sixteenth|seventeenth|eighteenth|nineteenth|twentieth|twenty-first|twenty-second|twenty-third|twenty-fourth|twenty-fifth|twenty-sixth|twenty-seventh|twenty-eighth|twenty-ninth|thirtieth|thirtieth-first|thirtieth-second|thirtieth-third|thirtieth-fourth|thirtieth-fifth|thirtieth-sixth|thirtieth-seventh|thirtieth-eighth|thirtieth-ninth|thirtieth-tenth|thirtieth-eleventh|thirtieth-twelfth|thirtieth-thirteenth|thirtieth-fourteenth|thirtieth-fifteenth|thirtieth-sixteenth|thirtieth-seventeenth|thirtieth-eighteenth|thirtieth-nineteenth|thirtieth-twentieth|thirtieth-twenty-first|thirtieth-twenty-second|thirtieth-twenty-third|thirtieth-twenty-fourth|thirtieth-twenty-fifth|thirtieth-twenty-sixth|thirtieth-twenty-seventh|thirtieth-twenty-eighth|thirtieth-twenty-ninth|thirtieth-thirtieth])")
    matches = re.findall(amendments, re.B)
    return(matches)

def compile_regex():
    f_regex = r"([0-9]+) ([P, Supp, 2d, 3d, 4d, 5d, 6d, 7d, 8d, 9d, 10d, 11d, 12d, 13d, 14d, 15d, 16d, 17d, 18d, 19d, 20d, 21d, 22d, 23d, 24d, 25d, 26d, 27d, 28d, 29d, 30d, 31d, 32d, 33d, 34d, 35d, 36d, 37d, 38d, 39d, 40d, 41d, 42d, 43d, 44d, 45d, 46d, 47d, 48d, 49d, 50d, 51d, 52d, 53d, 54d, 55d, 56d, 57d, 58d, 59d, 60d, 61d, 62d, 63d, 64d, 65d, 66d, 67d, 68d, 69d, 70d, 71d, 72d, 73d, 74d, 75d, 76d, 77d, 78d, 79d, 80d, 81d, 82d, 83d, 84d, 85d, 86d, 87d, 88d, 89d, 90d, 91d, 92d, 93d, 94d, 95d, 96d, 97d, 98d, 99d, 100d, 101d, 102d, 103d, 104d, 105d, 106d, 107d, 108d, 109d, 110d, 111d, 112d, 113d, 114d, 115d, 116d, 117d, 118d, 119d, 120d, 121d, 122d, 123d, 124d, 125d, 126d, 127d, 128d, 129d, 130d, 131d, 132d, 133d, 134d, 135d, 136d, 137d, 138d, 139d, 140d, 141d, 142d, 143d, 144d, 145d, 146d, 147d, 148d, 149d, 150d, 151d, 152d, 153d, 154d, 155d, 156d, 157d, 158d, 159d, 160d, 161d, 162d, 163d, 164d, 165d, 166d, 167d, 168d, 169d, 170d, 171d, 172d, 173d, 174d, 175d, 176d, 177d, 178d, 179d, 180d, 181d, 182d, 183d, 184d, 185d, 186d, 187d, 188d, 189d, 190d, 191d, 192d, 193d, 194d, 195d, 196d, 197d, 198d, 199d, 200d, 201d, 202d, 203d, 204d, 205d, 206d, 207d, 208d, 209d, 210d, 211d, 212d, 213d, 214d, 215d, 216d, 217d, 218d, 219d, 220d, 221d, 222d, 223d, 224d, 225d, 226d, 227d, 228d, 229d, 230d, 231d, 232d, 233d, 234d, 235d, 236d, 237d, 238d, 239d, 240d, 241d, 242d, 243d, 244d, 245d, 246d, 247d, 248d, 249d, 250d, 251d, 252d, 253d, 254d, 255d, 256d, 257d, 258d, 259d, 260d, 261d, 262d, 263d, 264d, 265d, 266d, 267d, 268d, 269d, 270d, 271d, 272d, 273d, 274d, 275d, 276d, 277d, 278d, 279d, 280d, 281d, 282d, 283d, 284d, 285d, 286d, 287d, 288d, 289d, 290d, 291d, 292d, 293d, 294d, 295d, 296d, 297d, 298d, 299d, 300d, 301d, 302d, 303d, 304d, 305d, 306d, 307d, 308d, 309d, 310d, 311d, 312d, 313d, 314d, 315d, 316d, 317d, 318d, 319d, 320d, 321d, 322d, 323d, 324d, 325d, 326d, 327d, 328d, 329d, 330d, 331d, 332d, 333d, 334d, 335d, 336d, 337d, 338d, 339d, 340d, 341d, 342d, 343d, 344d, 345d, 346d, 347d, 348d, 349d, 350d, 351d, 352d, 353d, 354d, 355d, 356d, 357d, 358d, 359d, 360d, 361d, 362d, 363d, 364d, 365d, 366d, 367d, 368d, 369d, 370d, 371d, 372d, 373d, 374d, 375d, 376d, 377d, 378d, 379d, 380d, 381d, 382d, 383d, 384d, 385d, 386d, 387d, 388d, 389d, 390d, 391d, 392d, 393d, 394d, 395d, 396d, 397d, 398d, 399d, 400d, 401d, 402d, 403d, 404d, 405d, 406d, 407d, 408d, 409d, 410d, 411d, 412d, 413d, 414d, 415d, 416d, 417d, 418d, 419d, 420d, 421d, 422d, 423d, 424d, 425d, 426d, 427d, 428d, 429d, 430d, 431d, 432d, 433d, 434d, 435d, 436d, 437d, 438d, 439d, 440d, 441d, 442d, 443d, 444d, 445d, 446d, 447d, 448d, 449d, 450d, 451d, 452d, 453d, 454d, 455d, 456d, 457d, 458d, 459d, 460d, 461d, 462d, 463d, 464d, 465d, 466d, 467d, 468d, 469d, 470d, 471d, 472d, 473d, 474d, 475d, 476d, 477d, 478d, 479d, 480d, 481d, 482d, 483d, 484d, 485d, 486d, 487d, 488d, 489d, 490d, 491d, 492d, 493d, 494d, 495d, 496d, 497d, 498d, 499d, 500d, 501d, 502d, 503d, 504d, 505d, 506d, 507d, 508d, 509d, 510d, 511d, 512d, 513d, 514d, 515d, 516d, 517d, 518d, 519d, 520d, 521d, 522d, 523d, 524d, 525d, 526d, 527d, 528d, 529d, 530d, 531d, 532d, 533d, 534d, 535d, 536d, 537d, 538d, 539d, 540d, 541d, 542d, 543d, 544d, 545d, 546d, 547d, 548d, 549d, 550d, 551d, 552d, 553d, 554d, 555d, 556d, 557d, 558d, 559d, 560d, 561d, 562d, 563d, 564d, 565d, 566d, 567d, 568d, 569d, 570d, 571d, 572d, 573d, 574d, 575d, 576d, 577d, 578d, 579d, 580d, 581d, 582d, 583d, 584d, 585d, 586d, 587d, 588d, 589d, 590d, 591d, 592d, 593d, 594d, 595d, 596d, 597d, 598d, 599d, 600d, 601d, 602d, 603d, 604d, 605d, 606d, 607d, 608d, 609d, 610d, 611d, 612d, 613d, 614d, 615d, 616d, 617d, 618d, 619d, 620d, 621d, 622d, 623d, 624d, 625d, 626d, 627d, 628d, 629d, 630d, 631d, 632d, 633d, 634d, 635d, 636d, 637d, 638d, 639d, 640d, 641d, 642d, 643d, 644d, 645d, 646d, 647d, 648d, 649d, 650d, 651d, 652d, 653d, 654d, 655d, 656d, 657d, 658d, 659d, 660d, 661d, 662d, 663d, 664d, 665d, 666d, 667d, 668d, 669d, 670d, 671d, 672d, 673d, 674d, 675d, 676d, 677d, 678d, 679d, 680d, 681d, 682d, 683d, 684d, 685d, 686d, 687d, 688d, 689d, 690d, 691d, 692d, 693d, 694d, 695d, 696d, 697d, 698d, 699d, 700d, 701d, 702d, 703d, 704d, 705d, 706d, 707d, 708d, 709d, 710d, 711d, 712d, 713d, 714d, 715d, 716d, 717d, 718d, 719d, 720d, 721d, 722d, 723d, 724d, 725d, 726d, 727d, 728d, 729d, 730d, 731d, 732d, 733d, 734d, 735d, 736d, 737d, 738d, 739d, 740d, 741d, 742d, 743d, 744d, 745d, 746d, 747d, 748d, 749d, 750d, 751d, 752d, 753d, 754d, 755d, 756d, 757d, 758d, 759d, 760d, 761d, 762d, 763d, 764d, 765d, 766d, 767d, 768d, 769d, 770d, 771d, 772d, 773d, 774d, 775d, 776d, 777d, 778d, 779d, 780d, 781d, 782d, 783d, 784d, 785d, 786d, 787d, 788d, 789d, 790d, 791d, 792d, 793d, 794d, 795d, 796d, 797d, 798d, 799d, 800d, 801d, 802d, 803d, 804d, 805d, 806d, 807d, 808d, 809d, 810d, 811d, 812d, 813d, 814d, 815d, 816d, 817d, 818d, 819d, 820d, 821d, 822d, 823d, 824d, 825d, 826d, 827d, 828d, 829d, 830d, 831d, 832d, 833d, 834d, 835d, 836d, 837d, 838d, 839d, 840d, 841d, 842d, 843d, 844d, 845d, 846d, 847d, 848d, 849d, 850d, 851d, 852d, 853d, 854d, 855d, 856d, 857d, 858d, 859d, 860d, 861d, 862d, 863d, 864d, 865d, 866d, 867d, 868d, 869d, 870d, 871d, 872d, 873d, 874d, 875d, 876d, 877d, 878d, 879d, 880d, 881d, 882d, 883d, 884d, 885d, 886d, 887d, 888d, 889d, 890d, 891d, 892d, 893d, 894d, 895d, 896d, 897d, 898d, 899d, 900d, 901d, 902d, 903d, 904d, 905d, 906d, 907d, 908d, 909d, 910d, 911d, 912d, 913d, 914d, 915d, 916d, 917d, 918d, 919d, 920d, 921d, 922d, 923d, 924d, 925d, 926d, 927d, 928d, 929d, 930d, 931d, 932d, 933d, 934d, 935d, 936d, 937d, 938d, 939d, 940d, 941d, 942d, 943d, 944d, 945d, 946d, 947d, 948d, 949d, 950d, 951d, 952d, 953d, 954d, 955d, 956d, 957d, 958d, 959d, 960d, 961d, 962d, 963d, 964d, 965d, 966d, 967d, 968d, 969d, 970d, 971d, 972d, 973d, 974d, 975d, 976d, 977d, 978d, 979d, 980d, 981d, 982d, 983d, 984d, 985d, 986d, 987d, 988d, 989d, 990d, 991d, 992d, 993d, 994d, 995d, 996d, 997d, 998d, 999d, 1000d, 1001d, 1002d, 1003d, 1004d, 1005d, 1006d, 1007d, 1008d, 1009d, 1010d, 1011d, 1012d, 1013d, 1014d, 1015d, 1016d, 1017d, 1018d, 1019d, 1020d, 1021d, 1022d, 1023d, 1024d, 1025d, 1026d, 1027d, 1028d, 1029d, 1030d, 1031d, 1032d, 1033d, 1034d, 1035d, 1036d, 1037d, 1038d, 1039d, 1040d, 1041d, 1042d, 1043d, 1044d, 1045d, 1046d, 1047d, 1048d, 1049d, 1050d, 1051d, 1052d, 1053d, 1054d, 1055d, 1056d, 1057d, 1058d, 1059d, 1060d, 1061d, 1062d, 1063d, 1064d, 1065d, 1066d, 1067d, 1068d, 1069d, 1070d, 1071d, 1072d, 1073d, 1074d, 1075d, 1076d, 1077d, 1078d, 1079d, 1080d, 1081d, 1082d, 1083d, 1084d, 1085d, 1086d, 1087d, 1088d, 1089d, 1090d, 1091d, 1092d, 1093d, 1094d, 1095d, 1096d, 1097d, 1098d, 1099d, 1100d, 1101d, 1102d, 1103d, 1104d, 1105d, 1106d, 1107d, 1108d, 1109d, 1110d, 1111d, 1112d, 1113d, 1114d, 1115d, 1116d, 1117d, 1118d, 1119d, 1120d, 1121d, 1122d, 1123d, 1124d, 1125d, 1126d, 1127d, 1128d, 1129d, 1130d, 1131d, 1132d, 1133d, 1134d, 1135d, 1136d, 1137d, 1138d, 1139d, 1140d, 1141d, 1142d, 1143d, 1144d, 1145d, 1146d, 1147d, 1148d, 1149d, 1150d, 1151d, 1152d, 1153d, 1154d, 1155d, 1156d, 1157d, 1158d, 1159d, 1160d, 1161d, 1162d, 1163d, 1164d, 1165d, 1166d, 1167d, 1168d, 1169d, 1170d, 1171d, 1172d, 1173d, 1174d, 1175d, 1176d, 1177d, 1178d, 1179d, 1180d, 1181d, 1182d, 1183d, 1184d, 1185d, 1186d, 1187d, 1188d, 1189d, 1190d, 1191d, 1192d, 1193d, 1194d, 1195d, 1196d, 1197d, 1198d, 1199d, 1200d, 1201d, 1202d, 1203d, 1204d, 1205d, 1206d, 1207d, 1208d, 1209d, 1210d, 1211d, 1212d, 1213d, 1214d, 1215d, 1216d, 1217d, 1218d, 1219d, 1220d, 1221d, 1222d, 1223d, 1224d, 1225d, 1226d, 1227d, 1228d, 1229d, 1230d, 1231d, 1232d, 1233d, 1234d, 1235d, 1236d, 1237d, 1238d, 1239d, 1240d, 1241d, 1242d, 1243d, 1244d, 1245d, 1246d, 1247d, 1248d, 1249d, 1250d, 1251d, 1252d, 1253d, 1254d, 1255d, 1256d, 1257d, 1258d, 1259d, 1260d, 1261d, 1262d, 1263d, 1264d, 1265d, 1266d, 1267d, 1268d, 1269d, 1270d, 1271d, 1272d, 1273d, 1274d, 1275d, 1276d, 1277d, 1278d, 1279d, 1280d, 1281d, 1282d, 1283d, 1284d, 1285d, 1286d, 1287d, 1288d, 1289d, 1290d, 1291d, 1292d, 1293d, 1294d, 1295d, 1296d, 1297d, 1298d, 1299d, 1300d, 1301d, 1302d, 1303d, 1304d, 1305d, 1306d, 1307d, 1308d, 1309d, 1310d, 1311d, 1312d, 1313d, 1314d, 1315d, 1316d, 1317d, 1318d, 1319d, 1320d, 1321d, 1322d, 1323d, 1324d, 1325d, 1326d, 1327d, 1328d, 1329d, 1330d, 1331d, 1332d, 1333d, 1334d, 1335d, 1336d, 1337d, 1338d, 1339d, 1340d, 1341d, 1342d, 1343d, 1344d, 1345d, 1346d, 1347d, 1348d, 1349d, 1350d, 1351d, 1352d, 1353d, 1354d, 1355d, 1356d, 1357d, 1358d, 1359d, 1360d, 1361d, 1362d, 1363d, 1364d, 1365d, 1366d, 1367d, 1368d, 1369d, 1370d, 1371d, 1372d, 1373d, 1374d, 1375d, 1376d, 1377d, 1378d, 1379d, 1380d, 1381d, 1382d, 1383d, 1384d, 1385d, 1386d, 1387d, 1388d, 1389d, 1390d, 1391d, 1392d, 1393d, 1394d, 1395d, 1396d, 1397d, 1398d, 1399d, 1400d, 1401d, 1402d, 1403d, 1404d, 1405d, 1406d, 1407d, 1408d, 1409d, 1410d, 1411d, 1412d, 1413d, 1414d, 1415d, 1416d, 1417d, 1418d, 1419d, 1420d, 1421d, 1422d, 1423d, 1424d, 1425d, 1426d, 1427d, 1428d, 1429d, 1430d, 1431d, 1432d, 1433d, 1434d, 1435d, 1436d, 1437d, 1438d, 1439d, 1440d, 1441d, 1442d, 1443d, 1444d, 1445d, 1446d, 1447d, 1448d, 1449d, 1450d, 1451d, 1452d, 1453d, 1454d, 1455d, 1456d, 1457d, 1458d, 1459d, 1460d, 1461d, 1462d, 1463d, 1464d, 1465d, 1466d, 1467d, 1468d, 1469d, 1470d, 1471d, 1472d, 1473d, 1474d, 1475d, 1476d, 1477d, 1478d, 1479d, 1480d, 1481d, 1482d, 1483d, 1484d, 1485d, 1486d, 1487d, 1488d, 1489d, 1490d, 1491d, 1492d, 1493d, 1494d, 1495d, 1496d, 1497d, 1498d, 1499d, 1500d, 1501d, 1502d, 1503d, 1504d, 1505d, 1506d, 1507d, 1508d, 1509d, 1510d, 1511d, 1512d, 1513d, 1514d, 1515d, 1516d, 1517d, 1518d, 1519d, 1520d, 1521d, 1522d, 1523d, 1524d, 1525d, 1526d, 1527d, 1528d, 1529d, 1530d, 1531d, 1532d, 1533d, 1534d, 1535d, 1536d, 1537d, 1538d, 1539d, 1540d, 1541d, 1542d, 1543d, 1544d, 1545d, 1546d, 1547d, 1548d, 1549d, 1550d, 1551d, 1552d, 1553d, 1554d, 1555d, 1556d, 1557d, 1558d, 1559d, 1560d, 1561d, 1562d, 1563d, 1564d, 1565d, 1566d, 1567d, 1568d, 1569d, 1570d, 1571d, 1572d, 1573d, 1574d, 1575d, 1576d, 1577d, 1578d, 1579d, 1580d, 1581d, 1582d, 1583d, 1584d, 1585d, 1586d, 1587d, 1588d, 1589d, 1590d, 1591d, 1592d, 1593d, 1594d, 1595d, 1596d, 1597d, 1598d, 1599d, 1600d, 1601d, 1602d, 1603d, 1604d, 1605d, 1606d, 1607d, 1608d, 1609d, 1610d, 1611d, 1612d, 1613d, 1614d, 1615d, 1616d, 1617d, 1618d, 1619d, 1620d, 1621d, 1622d, 1623d, 1624d, 1625d, 1626d, 1627d, 1628d, 1629d, 1630d, 1631d, 1632d, 1633d, 1634d, 1635d, 1636d, 1637d, 1638d, 1639d, 1640d, 1641d, 1642d, 1643d, 1644d, 1645d, 1646d, 1647d, 1648d, 1649d, 1650d, 1651d, 1652d, 1653d, 1654d, 1655d, 1656d, 1657d, 1658d, 1659d, 1660d, 1661d, 1662d, 1663d, 1664d, 1665d, 1666d, 1667d, 1668d, 1669d, 1670d, 1671d, 1672d, 1673d, 1674d, 1675d, 1676d, 1677d, 1678d, 1679d, 1680d, 1681d, 1682d, 1683d, 1684d, 1685d, 1686d, 1687d, 1688d, 1689d, 1690d, 1691d, 1692d, 1693d, 1694d, 1695d, 1696d, 1697d, 1698d, 1699d, 1700d, 1701d, 1702d, 1703d, 1704d, 1705d, 1706d, 1707d, 1708d, 1709d, 1710d, 1711d, 1712d, 1713d, 1714d, 1715d, 1716d, 1717d, 1718d, 1719d, 1720d, 1721d, 1722d, 1723d, 1724d, 1725d, 1726d, 1727d, 1728d, 1729d, 1730d, 1731d, 1732d, 1733d, 1734d, 1735d, 1736d, 1737d, 1738d, 1739d, 1740d, 1741d, 1742d, 1743d, 1744d, 1745d, 1746d, 1747d, 1748d, 1749d, 1750d, 1751d, 1752d, 1753d, 1754d, 1755d, 1756d, 1757d, 1758d, 1759d, 1760d, 1761d, 1762d, 1763d, 1764d, 1765d, 1766d, 1767d, 1768d, 1769d, 1770d, 1771d, 1772d, 1773d, 1774d, 1775d, 1776d, 1777d, 1778d, 1779d, 1780d, 1781d, 1782d, 1783d, 1784d, 1785d, 1786d, 1787d, 1788d, 1789d, 1790d, 1791d, 1792d, 1793d, 1794d, 1795d, 1796d, 1797d, 1798d, 1799d, 1800d, 1801d, 1802d, 1803d, 1804d, 1805d, 1806d, 1807d, 1808d, 1809d, 1810d, 1811d, 1812d, 1813d, 1814d, 1815d, 1816d, 1817d, 1818d, 1819d, 1820d, 1821d, 1822d, 1823d, 1824d, 1825d, 1826d, 1827d, 1828d, 1829d, 1830d, 1831d, 1832d, 1833
```

Counting and Apportioning Legal History with Natural Language Processing

Does
law
evolve?

Supreme
Court
Opinions,
1791-2019

Strategy
and
Methods

Findings of
Fact and
Conclusions
of Law

Next
Steps

2019/08/01
LORA JOHNS

Methods of analysis

Objectives



Tools and libraries:

- Scikit-learn, numpy, pandas
- Matplotlib, plotly, Plotly Express, OpenCV
- SpaCy, NLTK
- Django, Scrappy, REST API
- SQL, MongoDB, neo4j

Topic Modeling

Objectives

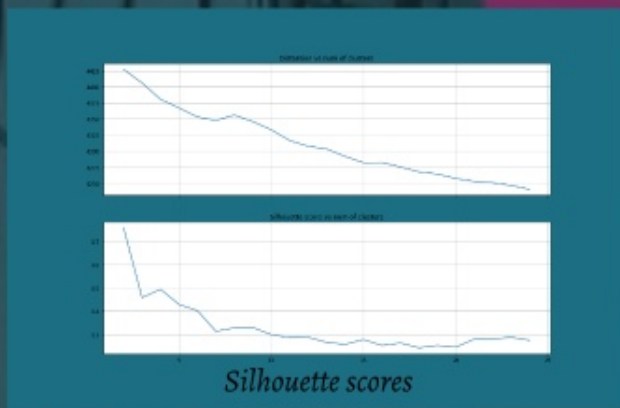
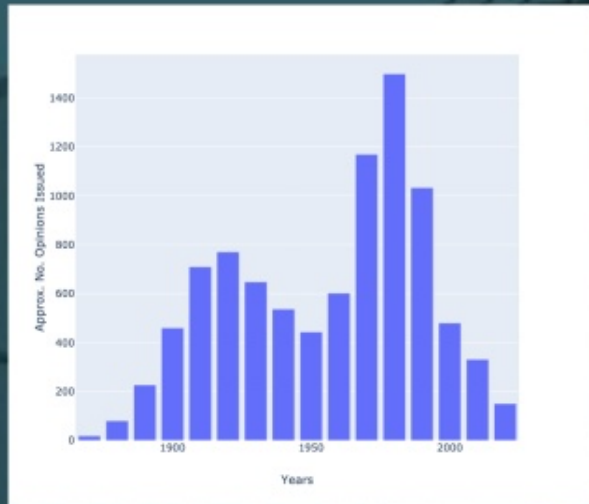
Extract topics from case law relevant to a broad legal issue (i.e., the 14th Amend.)

Obtain a picture of the whole corpus ✓

Obtain a picture of the whole corpus ✓

Fit the model to the corpus with a rolling window ✓

Topic extraction workflow



- Make TF-IDF matrix, bag of words, and extract features
- Perform LSA and singular value decomposition
- K-means test with silhouette method
- Create rolling window function class
- Evaluate and visualize results
- Revise and apply to additional document classes
- Expand and refine methods

Counting and Apportioning Legal History with Natural Language Processing

Does
law
evolve?

Supreme
Court
Opinions,
1791-2019

Strategy
and
Methods

Findings of
Fact and
Conclusions
of Law

Next
Steps

2019/08/01
LORA JOHNS

Evolving 14th Amendment Jurisprudence

Visualized through k-means clusters



The 14th
Amendment

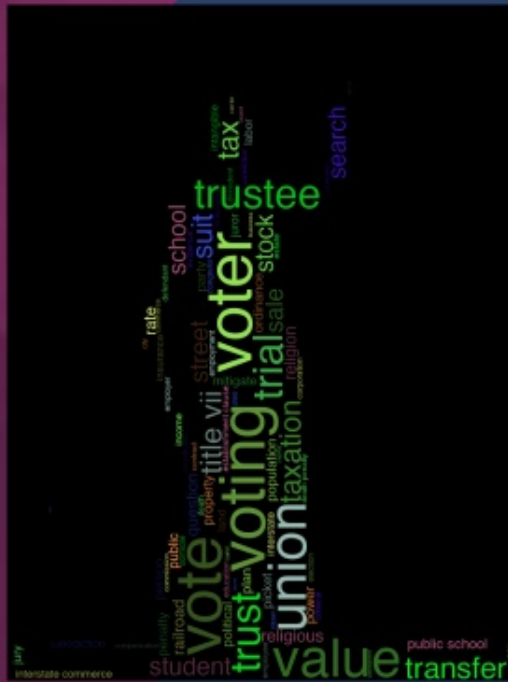
Distinct
Eras

Topics: The Corpus as a Whole

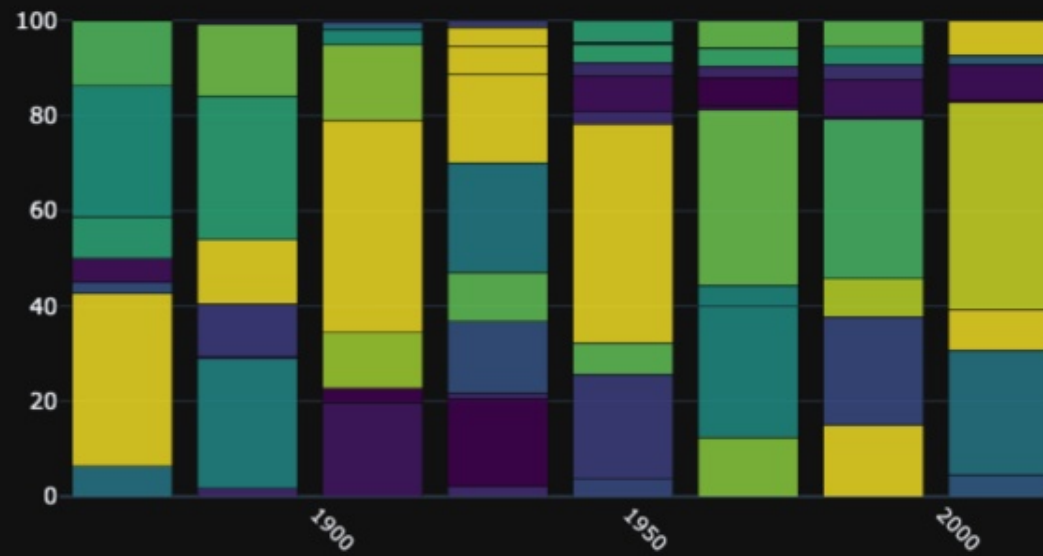
- Mini-batch k-means pipeline, using TF-IDF, with n-grams up to 2 words and 10 clusters
- Metrics: elbow plot, silhouette score
- Additional hyperparameter tuning with sklearn
- Additionally, I included the size of the time window and span of overlap in the training pipeline

Top Terms
over Time

Top
Topics
over Time



Percentage of cases per topic, 1791-Present



Topics through American History

Counting and Apportioning Legal History with Natural Language Processing

Does
law
evolve?

Supreme
Court
Opinions,
1791-2019

Strategy
and
Methods

Findings of
Fact and
Conclusions
of Law

Next
Steps

2019/08/01
LORA JOHNS

Next Steps

- Expand the analysis to more documents and corpora
- Incorporate additional analyses and refine those already used
- Build a graph database to contain the vast stores of information acquired
- Use the semantic analysis uncovered to build more relevant legal research tools

Counting and Apportioning Legal History with Natural Language Processing

Does
law
evolve?

Supreme
Court
Opinions,
1791-2019

Strategy
and
Methods

Findings of
Fact and
Conclusions
of Law

Next
Steps

2019/08/01
LORA JOHNS