선거결과 예측을 위한 기초데이터 수집 (2016.3.23)

1. 공공 데이터 포탈 : <https://www.data.go.kr>

[투개표정보조회서비스](https://www.data.go.kr/subMain.jsp?param=REFUQUAxNTAwMDkwMA==#/L2NvbW0vY29tbW9uU2VhcmNoL29yZ2luRGF0YVNldC5kbyRAXjAxMm0xJEBecHVibGljRGF0YVBrPTE1MDAwOTAwJEBeYnJtQ2Q9T0MwMDAzJEBeZXhjZWxDb3VudD0wJEBeZG93bmxvYWRDb3VudD02OSRAXm9yZ0luZGV4PURBVEE=) 2014년 상반기 재보궐선거 개표결과

[조례제정 및 개폐창구를 위한 19세이상 주민 총수](https://www.data.go.kr/subMain.jsp?param=REFUQUAzMDM5NzQz#/L2NvbW0vY29tbW9uU2VhcmNoL29yZ2luRGF0YVNldC5kbyRAXjAxMm0xJEBecHVibGljRGF0YVBrPTMwMzk3NDMkQF5icm1DZD1PQzAwMDMkQF5leGNlbENvdW50PTAkQF5kb3dubG9hZENvdW50PTIkQF5vcmdJbmRleD1EQVRB) 연제구 조례 개정 및 개폐 청구를 위한 19세 이상 선거권이 있는 주민총수 및 연서대상 주민수 공고

[당선인정보조회서비스](https://www.data.go.kr/subMain.jsp?param=T1BFTkFQSUAxNTAwMDg2NA==#/L3B1YnIvdXNlL3ByaS9Jcm9zT3BlbkFwaURldGFpbC9vcGVuQXBpTGlzdFBhZ2UuZG8kQF4wMTJtMSRAXnB1YmxpY0RhdGFQaz0xNTAwMDg2NCRAXmJybUNkPU9DMDAwMyRAXnJlcXVlc3RDb3VudD05JEBeb3JnSW5kZXg9T1BFTkFQSQ==) 선거 기간에만 제공되는 서비스

[코드조회 서비스](https://www.data.go.kr/subMain.jsp?param=T1BFTkFQSUAxNTAwMDg5Nw==#/L3B1YnIvdXNlL3ByaS9Jcm9zT3BlbkFwaURldGFpbC9vcGVuQXBpTGlzdFBhZ2UuZG8kQF4wMTJtMSRAXnB1YmxpY0RhdGFQaz0xNTAwMDg5NyRAXmJybUNkPU9DMDAwMyRAXnJlcXVlc3RDb3VudD0xMCRAXm9yZ0luZGV4PU9QRU5BUEk=) 선거 기간에만 제공되는 서비스

[위치기반서비스](https://www.data.go.kr/subMain.jsp?param=T1BFTkFQSUAxNTAwMDg2OA==#/L3B1YnIvdXNlL3ByaS9Jcm9zT3BlbkFwaURldGFpbC9vcGVuQXBpTGlzdFBhZ2UuZG8kQF4wMTJtMSRAXnB1YmxpY0RhdGFQaz0xNTAwMDg2OCRAXmJybUNkPU9DMDAwMyRAXnJlcXVlc3RDb3VudD0xNCRAXm9yZ0luZGV4PU9QRU5BUEk=) 선거 기간에만 제공되는 서비스

[후보자정보 조회 서비스](https://www.data.go.kr/subMain.jsp?param=T1BFTkFQSUAxNTAwMDkwOA==#/L3B1YnIvdXNlL3ByaS9Jcm9zT3BlbkFwaURldGFpbC9vcGVuQXBpTGlzdFBhZ2UuZG8kQF4wMTJtMSRAXnB1YmxpY0RhdGFQaz0xNTAwMDkwOCRAXmJybUNkPU9DMDAwMyRAXnJlcXVlc3RDb3VudD0yOCRAXm9yZ0luZGV4PU9QRU5BUEk=) 선거 기간에만 제공되는 서비스

[부재자투표소 및 일반투표소 정보조회 서비스](https://www.data.go.kr/subMain.jsp?param=T1BFTkFQSUAxNTAwMDgzNg==#/L3B1YnIvdXNlL3ByaS9Jcm9zT3BlbkFwaURldGFpbC9vcGVuQXBpTGlzdFBhZ2UuZG8kQF4wMTJtMSRAXnB1YmxpY0RhdGFQaz0xNTAwMDgzNiRAXmJybUNkPU9DMDAwMyRAXnJlcXVlc3RDb3VudD0xMiRAXm9yZ0luZGV4PU9QRU5BUEk=) 선거 기간에만 제공되는 서비스

[투개표정보조회서비스](https://www.data.go.kr/subMain.jsp?param=T1BFTkFQSUAxNTAwMDkwMA==#/L3B1YnIvdXNlL3ByaS9Jcm9zT3BlbkFwaURldGFpbC9vcGVuQXBpTGlzdFBhZ2UuZG8kQF4wMTJtMSRAXnB1YmxpY0RhdGFQaz0xNTAwMDkwMCRAXmJybUNkPU9DMDAwMyRAXnJlcXVlc3RDb3VudD0xNCRAXm9yZ0luZGV4PU9QRU5BUEk=) 선거 기간에만 제공되는 서비스

[전지역 투표현황 조회](javascript:;)

[지역별 개표현황 조회(선거구)](javascript:;)

[전지역 개표현황 조회](javascript:;)

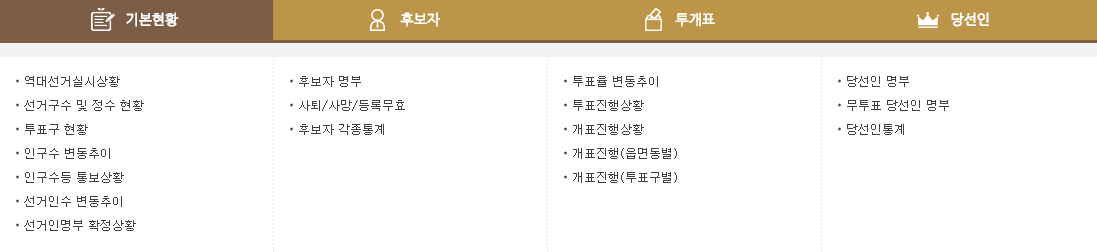
[지역별 개표현황 조회(읍면동)](javascript:;)

[지역별 투표현황 조회(선거구)](javascript:;)

2. 중앙선거관리위원회 통계관리시스템 http://info.nec.go.kr//

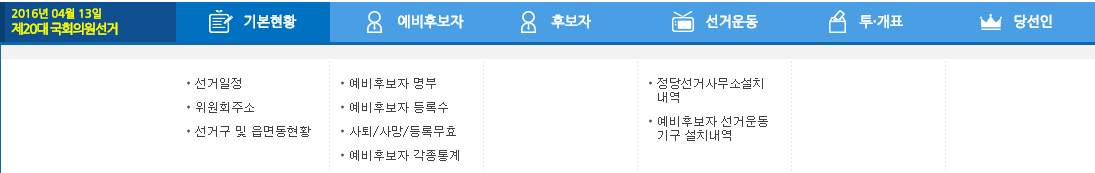
필요정보 : 국회의원 선거 1~20대 선거일정, 선거구, 후보자, 투개표, (당선자 – 투개표 결과임)

1차 시도 : Training Set (당장은 16~19대 선거데이터)



 국회의원선거 데이터 셋 만들기

Test Set 20대 예측해야 하는 데이터 : 투표율(각지역별), 득표수



크롤러를 위한 규칙확인

electionName :

electionType

[선거 타입 및 코드]

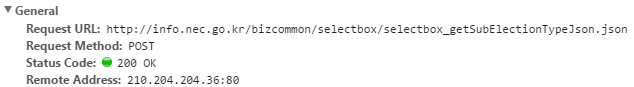
<http://info.nec.go.kr/main/showDocument.xhtml?electionId=0000000000&topMenuId=CP&secondMenuId=CPRI03>



여기서 국회의원 선거는 javascript:setElectionType(2);

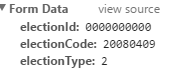
19대 선택시 : http://info.nec.go.kr/bizcommon/selectbox/selectbox\_getSubElectionTypeJson.json

요청 URL: <http://info.nec.go.kr/electioninfo/electionInfo_report.xhtml>



electionId=0000000000&electionCode=20080409&electionType=2

1. **electionId:** 0000000000
2. **electionCode:** 20120411 🡸== 19대 선거 일자!!!!
3. **electionType:2 <= 이건 국회의원 선거**

 **(18대 테스트)**

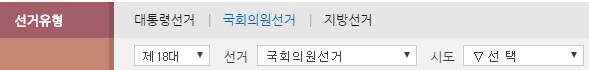
결과 {"jsonResult":{"header":{"result":"ok","errorMessage":"","errorCode":""},"body":[{"CODE":2,"NAME":"국회의원선거"},{"CODE":7,"NAME":"비례대표국회의원선거"}]}}



GET으로도 잘됨 OK…

<http://info.nec.go.kr/bizcommon/selectbox/selectbox_getSubElectionTypeJson.json?electionId=0000000000&electionCode=20080409&electionType=2>

* 이걸로 할 수 있는건 각 선거 19대 -> 국회의원선거코드가 2번인 것(2와 7이 있음)



제18대 >> 국회의원선거 선택한 경우 시도정보

http://info.nec.go.kr/bizcommon/selectbox/selectbox\_cityCodeBySgJson\_Old.json?electionId=0000000000&electionCode=20080409&electionType=2

해당 시도정보 획득

{"jsonResult":{"header":{"result":"ok","errorMessage":"","errorCode":""},"body":[{"CODE":**1100**,"NAME":"서울특별시"},{"CODE":2600,"NAME":"부산광역시"},{"CODE":2700,"NAME":"대구광역시"},{"CODE":2800,"NAME":"인천광역시"},{"CODE":2900,"NAME":"광주광역시"},{"CODE":3000,"NAME":"대전광역시"},{"CODE":3100,"NAME":"울산광역시"},{"CODE":4100,"NAME":"경기도"},{"CODE":4200,"NAME":"강원도"},{"CODE":4300,"NAME":"충청북도"},{"CODE":4400,"NAME":"충청남도"},{"CODE":4500,"NAME":"전라북도"},{"CODE":4600,"NAME":"전라남도"},{"CODE":4700,"NAME":"경상북도"},{"CODE":4800,"NAME":"경상남도"},{"CODE":4900,"NAME":"제주특별자치도"}]}}

http://info.nec.go.kr/bizcommon/selectbox/selectbox\_cityCodeBySgJson\_Old.json?electionId=0000000000&electionCode= 20120411&electionType=2 => 자동화 가능할 듯!!!

|  |
| --- |
| {"jsonResult":{"header":{"result":"ok","errorMessage":"","errorCode":""},"body":[{"CODE":1100,"NAME":"서울특별시"},{"CODE":2600,"NAME":"부산광역시"},{"CODE":2700,"NAME":"대구광역시"},{"CODE":2800,"NAME":"인천광역시"},{"CODE":2900,"NAME":"광주광역시"},{"CODE":3000,"NAME":"대전광역시"},{"CODE":3100,"NAME":"울산광역시"},{"CODE":5100,"NAME":"세종특별자치시"},{"CODE":4100,"NAME":"경기도"},{"CODE":4200,"NAME":"강원도"},{"CODE":4300,"NAME":"충청북도"},{"CODE":4400,"NAME":"충청남도"},{"CODE":4500,"NAME":"전라북도"},{"CODE":4600,"NAME":"전라남도"},{"CODE":4700,"NAME":"경상북도"},{"CODE":4800,"NAME":"경상남도"},{"CODE":4900,"NAME":"제주특별자치도"}]}} |

시도 아래 구 수준의 실제 선거구 정보

<http://info.nec.go.kr/bizcommon/selectbox/selectbox_getSggCityCodeJson_Old.json>? electionId=0000000000&electionCode=20080409&electionType=2&cityCode=1100

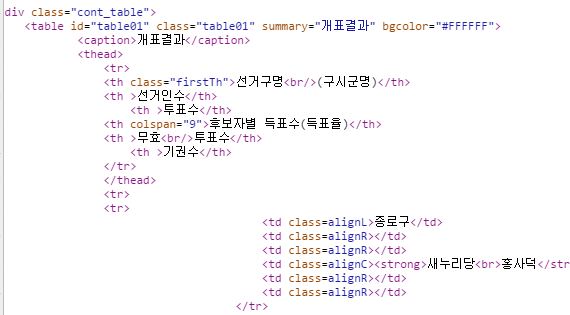
{"jsonResult":{"header":{"result":"ok","errorMessage":"","errorCode":""},"body":[{"CODE":2110101,"NAME":"종로구"},{"CODE":2110201,"NAME":"중구"},{"CODE":2110301,"NAME":"용산구"},{"CODE":2110401,"NAME":"성동구갑"},{"CODE":2110402,"NAME":"성동구을"},{"CODE":2110501,"NAME":"광진구갑"},{"CODE":2110502,"NAME":"광진구을"},{"CODE":2110601,"NAME":"동대문구갑"},{"CODE":2110602,"NAME":"동대문구을"},{"CODE":2110701,"NAME":"중랑구갑"},{"CODE":2110702,"NAME":"중랑구을"},{"CODE":2110801,"NAME":"성북구갑"},{"CODE":2110802,"NAME":"성북구을"},{"CODE":2110901,"NAME":"강북구갑"},{"CODE":2110902,"NAME":"강북구을"},{"CODE":2111001,"NAME":"도봉구갑"},{"CODE":2111002,"NAME":"도봉구을"},{"CODE":2111101,"NAME":"노원구갑"},{"CODE":2111102,"NAME":"노원구을"},{"CODE":2111103,"NAME":"노원구병"},{"CODE":2111201,"NAME":"은평구갑"},{"CODE":2111202,"NAME":"은평구을"},{"CODE":2111301,"NAME":"서대문구갑"},{"CODE":2111302,"NAME":"서대문구을"},{"CODE":2111401,"NAME":"마포구갑"},{"CODE":2111402,"NAME":"마포구을"},{"CODE":2111501,"NAME":"양천구갑"},{"CODE":2111502,"NAME":"양천구을"},{"CODE":2111601,"NAME":"강서구갑"},{"CODE":2111602,"NAME":"강서구을"},{"CODE":2111701,"NAME":"구로구갑"},{"CODE":2111702,"NAME":"구로구을"},{"CODE":2111801,"NAME":"금천구"},{"CODE":2111901,"NAME":"영등포구갑"},{"CODE":2111902,"NAME":"영등포구을"},{"CODE":2112001,"NAME":"동작구갑"},{"CODE":2112002,"NAME":"동작구을"},{"CODE":2112101,"NAME":"관악구갑"},{"CODE":2112102,"NAME":"관악구을"},{"CODE":2112201,"NAME":"서초구갑"},{"CODE":2112202,"NAME":"서초구을"},{"CODE":2112301,"NAME":"강남구갑"},{"CODE":2112302,"NAME":"강남구을"},{"CODE":2112401,"NAME":"송파구갑"},{"CODE":2112402,"NAME":"송파구을"},{"CODE":2112403,"NAME":"송파구병"},{"CODE":2112501,"NAME":"강동구갑"},{"CODE":2112502,"NAME":"강동구을"}]}}







electionId=0000000000&requestURI=%2FWEB-INF%2Fjsp%2Felectioninfo%2F0000000000%2Fvc%2Fvccp09.jsp&topMenuId=VC&secondMenuId=VCCP09&menuId=VCCP09&statementId=VCCP09\_%232&oldElectionType=1&**electionType=2&electionName=20120411&electionCode=2&cityCode=1100**&townCode=-1&sggCityCode=-1&x=35&y=12



이데이터는 HTML을 분석해서 가져와야 한다!!!!

여기서 각 선거구별 선거인수, 총투표자수 (일반투표수 + 부재자투표수) 확인가능

1차 목표 : 각 데이터의 XML 다운로드 기능(크롤링) - 선거회차(제X대) 시간대별 투표 선거구명

* 투/개표>투표진행현황 : 선거구 별 선거인수/투표자수
* 투/개표>개표진행상황 : 선거구 내 후보 별 득표수

2차 목표 : 데이터 분석용 트레이닝 셋 마련

if(electionType == '1'){

temp.append('<option value="20121219">제18대</option>');

temp.append('<option value="20071219">제17대</option>');

temp.append('<option value="20021219">제16대</option>');

temp.append('<option value="19971218">제15대</option>');

temp.append('<option value="19921218">제14대</option>');

temp.append('<option value="19871216">제13대</option>');

temp.append('<option value="19810225">제12대</option>');

temp.append('<option value="19800827">제11대</option>');

temp.append('<option value="19791206">제10대</option>');

temp.append('<option value="19780706">제9대</option>');

temp.append('<option value="19721223">제8대</option>');

temp.append('<option value="19710427">제7대</option>');

temp.append('<option value="19670503">제6대</option>');

temp.append('<option value="19631015">제5대</option>');

temp.append('<option value="19600315">제4대</option>');

temp.append('<option value="19560515">제3대</option>');

temp.append('<option value="19520805">제2대</option>');

temp.append('<option value="19480720">제1대</option>');

}

**if(electionType == '2'){ => 국회의원 선거**

temp.append('<option value="20120411">제19대</option>');

temp.append('<option value="20080409">제18대</option>');

temp.append('<option value="20040415">제17대</option>');

temp.append('<option value="20000413">제16대</option>');

temp.append('<option value="19960411">제15대</option>');

temp.append('<option value="19920324">제14대</option>');

temp.append('<option value="19880426">제13대</option>');

temp.append('<option value="19850212">제12대</option>');

temp.append('<option value="19810325">제11대</option>');

temp.append('<option value="19781212">제10대</option>');

temp.append('<option value="19730227">제9대</option>');

temp.append('<option value="19710525">제8대</option>');

temp.append('<option value="19670608">제7대</option>');

temp.append('<option value="19631126">제6대</option>');

temp.append('<option value="19600729">제5대</option>');

temp.append('<option value="19580502">제4대</option>');

temp.append('<option value="19540520">제3대</option>');

temp.append('<option value="19500530">제2대</option>');

temp.append('<option value="19480510">제1대</option>');

}

if(electionType == '4'){

temp.append('<option value="20140604">제6회</option>');

temp.append('<option value="20100602">제5회</option>');

temp.append('<option value="20060531">제4회</option>');

temp.append('<option value="20020613">제3회</option>');

temp.append('<option value="19980604">제2회</option>');

temp.append('<option value="19950627">제1회</option>');

}

if(electionType == '0'){

temp.append('<option value="20151028">2015년 하반기</option>');

temp.append('<option value="20150429">2015년 상반기</option>');

temp.append('<option value="20141029">2014년 하반기</option>');

temp.append('<option value="20140730">2014년 상반기</option>');

temp.append('<option value="20131204">2013년</option>');

temp.append('<option value="20131030">2013년 하반기</option>');

temp.append('<option value="20130626">2013년</option>');

temp.append('<option value="20130424">2013년 상반기</option>');

temp.append('<option value="20121031">2012년</option>');

temp.append('<option value="20121017">2012년</option>');

temp.append('<option value="20120627">2012년</option>');

temp.append('<option value="20111207">2011년</option>');

temp.append('<option value="20111116">2011년</option>');

temp.append('<option value="20111026">2011년 하반기</option>');

temp.append('<option value="20110824">2011년</option>');

temp.append('<option value="20110817">2011년</option>');

temp.append('<option value="20110427">2011년 상반기</option>');

temp.append('<option value="20101027">2010년 하반기</option>');

temp.append('<option value="20100728">2010년 상반기</option>');

temp.append('<option value="20091028">2009년 하반기</option>');

temp.append('<option value="20090826">2009년</option>');

temp.append('<option value="20090429">2009년 상반기</option>');

temp.append('<option value="20090408">경기도교육감</option>');

temp.append('<option value="20081217">대전교육감</option>');

temp.append('<option value="20081029">2008년 하반기</option>');

temp.append('<option value="20080730">서울교육감</option>');

temp.append('<option value="20080723">전북교육감</option>');

temp.append('<option value="20080625">충남교육감</option>');

temp.append('<option value="20080604">2008년 상반기</option>');

temp.append('<option value="20070425">2007년 상반기</option>');

temp.append('<option value="20070214">부산교육감</option>');

temp.append('<option value="20061025">2006년 하반기</option>');

temp.append('<option value="20060726">2006년 상반기</option>');

temp.append('<option value="20051026">2005년 하반기</option>');

temp.append('<option value="20050430">2005년 상반기</option>');

temp.append('<option value="20041030">2004년 하반기</option>');

temp.append('<option value="20040605">2004년 상반기</option>');

temp.append('<option value="20031030">2003년 하반기</option>');

temp.append('<option value="20030424">2003년 상반기</option>');

temp.append('<option value="20020808">2002년 상반기</option>');

}

if(electionType == '11'){

temp.append('<option value="20140604">제6회</option>');

temp.append('<option value="20100602">제5회</option>');

}