0. Intro

p.1 안녕하세요. 코사다마의 컨퍼런스 팀 a입니다. 저희 팀은 코로나 바이러스가 여타 의제들을 휩쓸어가는 상황에 위기감을 느꼈습니다. 이에 따라 국민적 공론장이라고 할 수 있는 청와대 청원 중 인권/성평등 카테고리를 분석하여 최근 여론을 파악하고자 했습니다. 청와대 청원 분석에서 성범죄에 대한 불안과 처벌에 대한 불만을 읽을 수 있었고, 이에 따라 성범죄 데이터를 분석했습니다.

0-1. Index

pg.2 청와대 청원 분석은 사이트에 크롤링을 통해 어떤 의견이 제기 되었는지 데이터를 수집하고, 워드클라우드 기법을 사용해 키워드 분석을 진행했습니다.

범죄와 형사사법 통계정보 DB에서 2014년부터 최근까지 성범죄 데이터를 분석했습니다. 성범죄가 어떤 양상을 띄고, 피해자와 가해자의 특징은 어떠하며, 어떤 장소와 지역에서 많이 일어나는지 범죄 분석을 해 보았습니다.

또 인구와 면적 대비 성범죄 사이에 어떤 상관관계가 있는지 분석했습니다.

1. 청와대 청원분석

pg.3 저희 팀은 국민의 목소리를 파악하기 위해 청와대 청원 분석을 진행했습니다. 먼저, 전체 국민청원의 17년 8월부터 ~ 20년 8월의 데이터를 모아본 뒤 가장 많은 동의를 얻은 청원을 간략하게 분석했습니다.

이후 인권과 성평등 카테고리 청원 중 만명이 넘는 추천을 받은 것을 따로 수집하여 추천 순 청원 분석해보고 청원 내용의 최다 빈출 단어를 모아 워드 클라우드로 시각화 해보았습니다.

pg.4

1.전체적인 청원 분석

우선, 국민청원 모든 카테고리를 기반한 간략한 분석 내용입니다. 기간 내 가장 참여인원이 많았던 청원은 카테고리 안전/환경의 텔레그램 n번방 용의자의 신상 공개 및 포토라인을 세워달라는 청원으로 270만명이동의했고, 또 다른 n번방 가입자 전원의 신상공개 청원이 202만 명으로 두번째로 청원이 많았습니다. n번방 사건에 대한 국민들의 적극적 의견 표출이라고 보여집니다.

pg.5

한편, 전체 청원 중 인권/성평등 카테고리의 청원이 가장 많았고,

pg.6

년도 별로 살펴보더라도 이에 대한 청원이 꾸준히 많았던 것을 볼 수 있습니다. 그만큼 국민들이 인권과 성평등에 대한 지속적인 관심을 보인 것으로 분석할 수 있습니다.

pg.7

2. 연도별 추천순 인권/성평등 청원 분석

다음으로 전반적인 사회 이슈와 시각을 알아보기 위해 2017년도부터 2020년까지 연도별 추천순 인권/성평등 청원을 분석해보았습니다. 먼저 17년도에는

17년도

1. **청소년보호법 폐지 청원**이 가장 많은 동의를 얻었고 뒤이어
2. **낙태죄 폐지, 자연유산 유도약 합법화**
3. **해외 사이트를 기반으로 한 무분별한 일반인 모욕 사진의 유포 처벌**
4. **소방관 순직 심사 관련 제도 개선** 청원

순으로 많은 동의를 얻었습니다. 실제로 이 청원에 대한 반향으로 **2019년 4월 11일 낙태법 헌법불합치 결정되었습니다.**

pg.8

18년도

1. **홍대 누드크로키 몰카 사건**을 계기로 한 불법촬영에 대해 **가해자의 성별과 상관 없는 동등하고 강화된 처벌, 그리고 피해자 방지 대책 요구**하는 청원이 가장 많은 동의를 얻었습니다.

pg.9 == > 불법 촬영에 대한 이슈가 공론화 된 해였으나 전년보다 사건 발생 건수가 높았습니다.

pg.10 또한 그 이후에도 카메라, 통신매체 이용 성범죄는 비슷한 추이를 보였습니다. 사회적으로 큰 변화는 없었던 것으로 보여집니다.

1. pg.11 뒤이어 특정 사건에 대한 억울함을 호소하는 청원글과
2. 이수역 폭행사건 처벌에 대한 청원이 그 뒤를 이었습니다. 이수역 폭행 사건은 이수역 인근 한 주점에서 남녀가 서로 폭행을 한 혐의로 재판에 넘겨진 사건입니다. 당시 양측 주장이 맞서면서 젠더 갈등 이슈로 부각되었습니다.

pg.11

19년도

1. 특정 사건 성폭행범 강력 처벌에 대한 청원
2. 성범죄에 대한 피해자, 목격자, 증언자에 대한 보호 시설 및 신변 보호 강화
3. 가해자 중심적인 성범죄의 양형 기준 재정비 청원
4. 리얼돌 수입 및 판매 금지 청원

순으로 많은 청원 동의를 얻엇습니다. 19년도에는 전체적으로 성폭행과 성범죄에 대한 가해자 처벌 강화를 요구하는 청원이 많은 동의를 얻은 것으로 보입니다.

pg.12

20년도 -

1. 강 모 서울고등법원 판사의 대법관 후보 자격 박탈 청원
2. 전 서울시장의 서울특별시장 진행 반대
3. 학교 내 성폭력 처벌 강화
4. 서울퀴어문화축제 반대

20년도에는 성범죄 처벌 강화에 대한 직접적이고, 간접적인 청원에 동의가 많았습니다. 한편 퀴어문화축제 반대 청원도 많은 동의를 얻어 권리 신장과 혐오가 공존하는 사회의 모습을 볼 수 있었습니다.

pg.13

4. 연도별 청원 내용 워드클라우드 분석

마지막으로 연도별 청원 내용 최다 빈출단어를 시각화한 워드클라우드를 살펴보겠습니다.

전체

우선 모든 성평등/인권 청원 내용에 대한 분석을 살펴봅니다.

최다 빈출 단어는 피해자,여성,사건,가해자였습니다.

pg.14

17년도

생각,여성,사건,청소년,피해자,경찰/경찰관,낙태,낙태죄

pg.15

18년도

생각,사건,여성,청소년,경찰,아이,낙태

pg.16

19년도

가해자,사람, 사건, 아이, 경찰, 피해자,여성, 범죄, 처벌, 저희

pg.17

20년도

피해자, 사건, 가해자, 성,진술, 수사, 국민, 검사, 처벌, 증거, 청원, 성폭행, 번방

pg.18

모든 년도에는 여성, 사건, 피해자라는 단어가 포함되어 있었습니다.

pg.19 성범죄 이외에도 폭력이나 무기를 사용한 범죄인 강력 범죄에 대한 여성 피해자 비율이 87%~91%의 값을 주기적으로 보이고 있습니다. 이처럼 여성에 대한 위험이 높아지는 사회에서 여성의 안전에 대한 목소리를 청와대 청원에서 확인할 수 있었습니다.

pg. 18 한편 17년도와 18년도에 보였던 낙태, 낙태죄라는 단어들은 낙태법 헌법불합치 결정이 난 19년도 이후부터는 빈출되지 않아 사회의 변화를 알 수 있었습니다. 또 20년에는 n번방 사건이 사회에 큰 충격을 주었던만큼 청원 내용에도 n번방의 번방이라는 단어가 등장하는 것을 볼 수 있습니다. 17, 18년도에는 청소년에 대한 청원 내용이 많았으나 19,20 년도에는 처벌이라는 단어가 생겨 청원 내용에 많이 등장했음을 알 수 있습니다.

pg.20 덧붙여 2019년 성폭력안전실태조사에서도 성폭력 감소를 위해 요구되는 정책에 대한 설문에 가해자 신속 검거와 가해자 처벌 강화가 많은 응답을 얻었습니다.

저희는 이를 통해 **낙태죄에 대한 헌법적 사회 변화**와, 높아진 **범죄 가해자 처벌 강화 요구**를 확인할 수 있었습니다.

2. 성범죄란

pg.21 성범죄란 성과 관련되는 범죄로서, 강간과 추행, 성매매, 아청법위반 (아동청소년의 성보호에 관한 법률 위반), 카메라등이용촬영죄, 성적목적공공장소침입, 통신매체이용음란죄, 유사강간 등으로 나뉩니다.

p22 앞으로 관련 용어를 수월하게 사용하기 위해서 각각 설명을 드리겠습니다. 먼저

**강간**은 폭행이나 협박에 의해 상대방의 반항을 곤란하게 하고 사람을 간음 하는 것을 의미합니다.

**추행**은 성욕의 흥분, 자극 또는 만족을 목적으로 하는 행위로. 성적 수치심이나 혐오의 감정을 느끼게 하는 일체의 행위를 의미합니다.

**성매매**는 금품이나 그 밖의 재산상의 이익을 주고받거나 주고받기로 약속하고 성교 행위 혹은 유사성교행위를 하는 것, 또는 그 상대방이 되는 것입니다.

**아청법위반(아동청소년의 성보호에 관한 법률 위반)**은 아동과 청소년의 성을 사는 행위, 성매매를 조장하는 모든 매개 행위, 아동-청소년에 대한 성폭력 행위자를 강력하게 처벌하고 대상 아동-청소년을 보호-구제하는 법입니다.

**카메라 등 이용촬영죄**란 성적 수치심 또는 욕망을 유발할 수 있는 신체를 촬영했거나, 이 복제물을 반포하는 행위를 의미합니다.

**성적목적공공장소침입**이란 성적 욕망을 만족시킬 목적으로 공중 화장실 및 목욕장 등 공공장소에 침입하거나 이 장소에서 퇴거의 요구를 받고 응하지 않는 것을 의미합니다.

**유사강간**이란 강간과 유사하게 강제적으로 맺는 유사 성행위를 의미합니다. 2012년 이전까지 강간죄는 질내삽입만을 대상으로 했습니다. 이때까지 항문이나 구강을 통한 강간은 처벌되지 못하고 강제추행으로 처벌받았습니다.

3. 성범죄 데이터분석

p23 청와대 청원에서 알아봤듯이 여성의 안전, 그 중에서 성범죄로부터의 안전은 매우 중요한 이슈로 떠올랐습니다. 그래서 저희는 조금 더 구체적으로 성범죄가 어떻게 일어나고 있고, 주요 피해자는 주로 어떤 사람들이며, 가해자는 어떤 특성을 가졌는지, 주로 어떤 장소에서 발생하는지 분석해 보았습니다. 데이터는 검찰, 경찰, 사법부 전체를 아우르는 ‘범죄와 형사사법 통계정보’ DB를 이용했습니다.

3-1. 성범죄 데이터 분석 - 성범죄 추이

p.24 우선 성범죄 추이를 알아보기 위해 분기별로 저장된 성범죄 데이터를 시각화 했습니다. 전체적으로 여름에 늘고, 겨울에 줄어드는 계절성 주기를 보였습니다.

p25연도별로 성범죄를 따로 떼어서 보면 강간 범죄의 발생건수는 증가하는 추세이고,

p26강제추행도 증가하는 추세를 확인할 수 있습니다.

p27아청법의 경우 2017년에 일시적으로 줄었다가 2018년에 반등했습니다.

p28카메라등을 이용한 촬영 범죄의 경우 발생 건수 추이가 일정하지 않습니다. 범죄 포착이나 검거가 어렵기 때문이라고 추측됩니다. 즉, 성범죄는 증가하는 추세이며 여름에 많이 일어나는 것을 확인했습니다.

3-2. 성범죄 데이터 분석 - 피해자 특성

p.29 다음으로는 피해자 특성을 알아보겠습니다. 94.32퍼센트, 그리고 5.68퍼센트. 각각 어떤 수치일까요? 2013년부터 2018년까지 일어난 전체 성범죄 중 여성이 피해자인 비율이 94.32퍼센트고, 5.68퍼센트는 남성이 피해자인 비율로 나타났습니다. 여성이 더 많을 것이라고 예상되긴 했으나 약 94대 6이라는 압도적인 수치를 보여주었습니다

p30 범죄별로는 강간 피해자의 성비가 약 99.4대 0.6, 강제추행의 경우 92.8대 6.8의 비율을 보여주었습니다.

p31 왼쪽 그래프인 유사 강간은 여성이 86, 남성이 14퍼센트입니다. 2012년 이후 질내 삽입 이외의 강간도 유사강간으로 보다 폭넓게 해석되면서 남성 피해자가 보다 통계적으로 기록되었습니다. 범죄와 형사사법 통계정보 DB에서 강간, 강제 추행, 유사강간을 제외한 나머지 성범죄를 뭉뚱그려 기타 범죄로 데이터를 제공했습니다. 여기에는 성매매, 아청법위반, 카메라 촬영죄, 성적목적 장소침입 등이 포함됩니다. 기타 범죄 또한 여성이 97퍼센트, 남성이 3퍼센트를 기록했습니다.

p32 다음은 각 범죄별 남여 피해자의 연령을 히트맵으로 표현해봤습니다. 히트맵이란 데이터 배열을 색의 채도를 이용하여 표현한 그래프를 의미합니다. x축은 연도, y축은 피해자 연령 이며 아래로 갈수록 연도가 증가합니다. 흰색일수록 숫자가 많다는 것을 의미합니다.

따라서 기타 범죄를 제외하고 범죄와 성에 관계 없이 20대 피해자가 가장 많았고, 그 다음이 10대 후반인 양상을 보여줍니다. 기타 범죄에는 성매매, 아청법위반, 카메라 촬영죄, 음란물 유포, 성적목적 장소침입 등의 범죄가 포함되는데, 10대 후반을 중심으로 10대 초반과 20대가 포함되었습니다. 이를 통해 주로 10대에서 20대 여성이 성범죄의 대상이 되는 것을 알 수 있었습니다.

다만 남성보다 여성 피해자가 많기 때문에 같은 밝은 색이라고 해도 여성 피해자의 숫자가 많습니다.

즉 성범죄 피해자의 대부분은 여성이며, 그중에서도 10대와 20대가 주요 피해자라는 것을 알 수 있었습니다.

3-3. 성범죄 데이터 분석 - 위치, 장소적 특성

p44다음으로 성범죄 전반이 어떤 지역과 장소에서 많이 일어나는지 분석했습니다. 저희는 히트맵으로 데이터를 시각화해보았습니다.

히트맵을 보면 노상이 제일 많고, 숙박업소/목욕탕, 아파트연립다세대, 유흥업소, 단독주택에서 성범죄가 많이 일어나는 것으로 나타났습니다. 대중교통이라고 할 수 있는 역대합실, 지하철, 기타교통수단에서도 비교적 성범죄가 많이 일어났습니다.

p45 그중에서도 아청법에 위배되는 성매수의 경우 숙박업소-목욕탕, 단독주택, 아파트에서 많이 일어났으며 비교적 범행장소가 한정적인 것으로 보입니다.

p46 음란물을 시청 및 소유한 장소는 단독주택, 아파트, 사무실로 즉, 주로 생활하는 장소에서 일어나는 경우가 총 36.2%를 차지했습니다. 하지만 기타 장소에서 62퍼센트가 발생하였습니다. 이 수치로 짐작할 때, 미디어 관련 성범죄가 지금까지 구별해 온 고전적인 범행장소 바깥에서 이루어지며, 범행이 고도화 되고 있다고 추측됩니다.

4. 지역별 인구/면적 대비 성범죄 비율

1. 인구

p47다음은 지역별 인구와 면적 대비 성범죄 비율을 비교해보겠습니다.

p48 1. 지역별 성범죄 발생 건수

지역별 성범죄 발생 건수를 히트맵으로 시각화 해보았습니다. 가장 발생건수가 많은 지역은 서울이었으며 경기, 인천, 부산 순으로 많았습니다.

p49 2.인구대비 성범죄 발생율

인구대비 성범죄 발생율 역시 서울 지역에서 가장 높았습니다. . 그 다음으로 높은 곳은 제주, 인천, 대전 순이었습니다.

3.인구와 성폭력 발생 간의 상관관계

p.50 저희 팀은 인구와 성폭력 발생 간의 상관관계에 대해서도 알아보았습니다. 상관계수는 약 0.322로 약한 양의 상관관계를 보이는 것을 확인할 수 있었습니다. 서울이 범죄 건수와 인구가 가장 많아 이상치로 보였습니다.

p.51 이상치로 나타난 서울의 구별 성폭력건수와 인구수 간의 상관관계를 조사한 결과, 상관계수가 0.15로 낮게 나타났습니다. 이는 성폭력 발생 건수에는 인구 수와 더불어, 다른 변수가 요인으로 작용했다고 볼 수 있습니다.

1. 지역별 성범죄

* pg52. 2014년, 2015년

다음은 지역별 성범죄 발생의 지도 시각화입니다. 데이터가 없는 경기도 광주, 울릉도 등의 지역들은 제외하였으며, 지도에서 보이는 파란색이 짙은 지역일수록 성범죄 발생이 많은 곳입니다.

2014년의 경우 ‘서울 강남’과 ‘수원’, ‘성남’ 지역에서 성범죄 발생이 가장 많았고, 상위 10개 지역 중 6개가 서울 지역이었습니다. ‘인천 동구’와 ‘부산 강서구’, ‘김천’, ‘부산 기장군’은 50건 미만으로 범죄가 가장 적었습니다. 2015년의 경우에도 유사한 결과를 보였지만 상위 10개 지역에서 2개의 지역을 제외하고 모두 성범죄 발생이 증가했습니다. 2014년과 비교했을 때 ‘서울 강남’ 지역은 범죄 발생이 증가한 반면, ‘부산 남구’와와 ‘안양’은 줄었습니다.

* pg53. 2016년, 2017년

2016년의 경우에도 전반적으로 유사한 색을 보입니다. 하지만 ‘서울 마포’와 ‘안산’ 지역이 2014년부터 증가하면서 범죄 발생 상위 10개 지역에 들어섰으며, ‘인천 남동’과 ‘인천 부평’ 지역도 증가했습니다. 반면 ‘여수’, ‘대구 중구’, ‘서울 양천’ 지역은 작년 대비 감소했습니다. 2017년에는 ‘인천 부평’ 지역이 4번째로 높았고, ‘서울 강남’은 2016년에 잠시 낮아졌다가 다시 1위로 복귀했습니다.

* pg54. 2018

2018년 역시 큰 변동 없이 유사합니다. ‘서울특별시’는 여전히 상위에 있었고, 반면 ‘광주광역시’는 2014년부터 전반적으로 성범죄 발생이 줄어들며 색이 옅어졌습니다. ‘인천’ 지역은 전국에서 가장 낮은 ‘동구’와 높은 발생수를 보이는 ‘부평구’, ‘남동구’가 공존하며 상반되는 색을 보여주고 있습니다. ‘부산’ 지역은 모든 지역에서 150건 미만의 발생수를 보여주고 있습니다. 참고로 중앙대학교가 있는 ‘서울 동작구’는 396건으로 상당히 높은 수치를 보여줬습니다.

5. 성범죄 억제 효과가 있을까? CCTV, 기타 데이터에 대하여

p55. 저희는 CCTV를 성범죄 억제 요인 중 하나라고 추정하고 조사했습니다.5만개에 달하는 공적 cctv데이터를 공공데이터 포탈에서 찾아 정말 유의미한 분석이 될 것으로 예상했습니다. 하지만 api에 결측 데이터가 너무 많아서 분석에 실패했습니다.

ㅠㅠㅠㅠㅠ

p.56먼저 앞서 지역별 시각화처럼 데이터를 정제 하려고 했으나, 일단 cctv의 위치가 시도별로 구분되어있지 않았습니다. 그 후에 이 데이터 안에서 위치를 특정할 수 있는 모든 수단을 강구해봤지만 실패했습니다.

p57 이후 지번주소와 도로명 주소로 시,도를 구분하려고 했으나 결측치가 많았고 시, 도가 아니라 구부터 작성된 데이터가 많았습니다.

p58어쩔 수 없이 관리기관, 제공기관명으로 찾으려 했지만 시도별로 관리하지 않는 cctv가 많아서 위치를 특정할 수 없었습니다.

p59위도와 경도 데이터는 결측치 없이 5만개의 데이터가 존재해서 구글맵스의 reverse\_geocode를 사용해 각 시 도를 알아보려고 했습니다. 하지만 list index out of range 오류가 발생했습니다. 즉, 범위 내에 없는 위도와 경도가 많이 포함되어있었습니다.

p60 위도, 경도의 결측치를 따로 모아 그 부분만 다른 칼럼들로 주소를 확인하고자 했으나 아예 값이 없는 것도 아니었고, 공백도 아니었기 때문에 자료를 판별할 수 없었습니다.

p61이렇게 데이터의 질이 떨어지게 된 이유는 각 기관에서 데이터가 취합되는 형태를 띄기 때문으로 추측됩니다. 공공데이터 개방 표준 고시에 따라서 제공하였지만 부정확한 정보가 너무 많았고, 일정한 형식이 갖춰지지 않았으며 위치별로 여러 개의 cctv가 존재하는 것을 고려했을 때 훨씬 많은 데이터가 제대로 집적되지 않고있습니다.

6. p.62한계, 아쉬웠던 점, 더 나아가 발전시키고 싶은 주제

* 저희의 원래 목표는 현재의 성범죄 현황을 분석해보고 시행되고 있는 제도들 (cctv, 여성 안심 정책 - 안심택배, 스카우트 제도 등) 의 실효성에 대해 알아보는 것이었습니다. cctv의 경우 면적 대비 카메라가 커버하고 있는 비율, 개수를 통해 지역 별 격차가 존재하는지와 이것이 성범죄율과 상관관계가 있는지 알아보고자 했습니다. 그러나 cctv의 데이터의 경우 일정하지 못한 양식과 데이터가 다수 존재하여 전처리가 불가능했고 여성 안심 정책의 경우 표면적인 통계 수치만 나와 있어 시군구별 범죄 예방 효과를 확인하기 어려웠습니다. 또 자료의 대부분이 2018년 정도까지만 존재하여 최근 변화를 살피지 못한것도 아쉬웠던 부분입니다.
* 청와대 청원은 전반적인 사회 이슈나 여론의 변화를 파악할 수 있다는데에 기여가 크다고 볼 수 있습니다. 그러나 청원에 대한 답변이 모호한 경우와, 예정 되어 있는 법안과 정책들이 실제로 의결되고 시행되고 있는지에 대해서는 면밀하게 지켜봐야 할 필요가 있어보입니다. 추천 순 청원 분석을 통해 답변된 청원들 중 답변에 제시된 법안이나 정책들이 강화되고 실제 시행 되었는지 분석해보는 것도 좋은 분석 주제가 될 것 같습니다. 또한 다른 카테고리에도 성범죄 및 성희롱과 관련된 청원이 포함되어 있어 다음에는 이도 추가하여 보면 더 나은 분석이 될 것 같습니다.

타임시이즈

수고하셨어요~