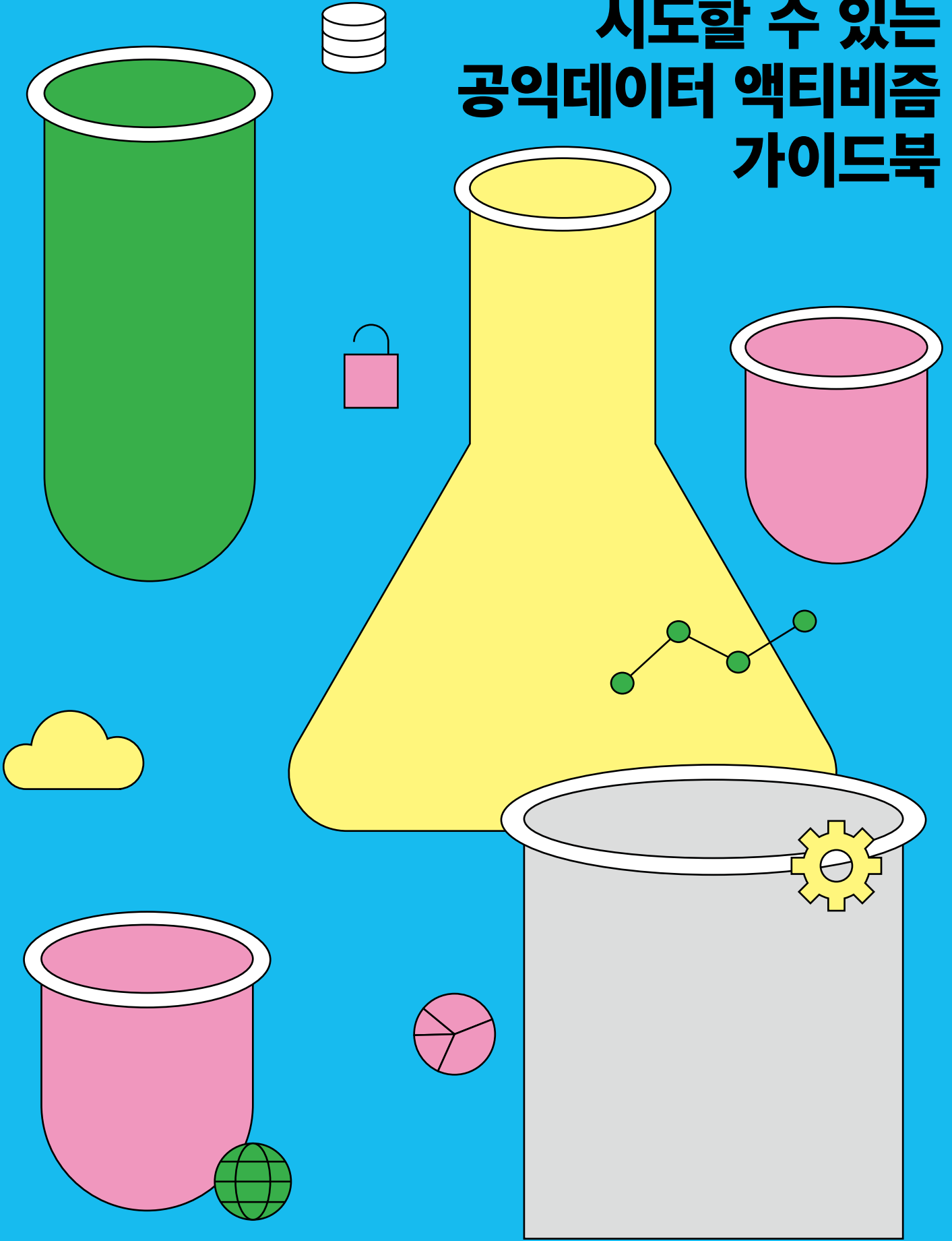


누구나 시작하고 시도할 수 있는 공익데이터 액티비즘 가이드북



들어가며

디지털 시대에 살고 있다고들 하죠. 일상 속 장면만 떠올려도, 디지털 기술과 데이터에 둘러싸인 나를 발견합니다. 비대면 업무로 화상회의를 진행하고, 휴대전화로 은행 업무를 보며, 스마트폰 어플리케이션과 온라인 플랫폼을 이용해 음식을 배달시키거나 필요한 서비스를 예약하기도 합니다. 생각과 정서를 SNS에 글이나 이미지로 남기며 사람들과 교류하는 것에도 너무나 익숙하죠. 디지털 기술의 활용은 더이상 선택값이 아닌 기본값인 시대가 되었습니다.

조금 더 들어가볼까요?

친구에게 들었던 드라마 정보가 유튜브 추천 목록에 떠서 놀란 적 있나요?
코로나19 마스크 재고 현황과 확진자 이동 동선을 체크해본 적 있나요?
우리 동네 맛집 리스트업, 어떻게 체크하고 관리하는지 궁금하다고요?

이런 데이터는 누가 어떻게 모아서, 어떻게 관리하는 걸까요? 가끔은 우리가 데이터의 참여자가 아닌 관찰자 같다는 생각도 듭니다. 데이터를 읽고 활용하는 것은 전문가들만의 영역일까요? 넘쳐나는 방대한 정보와 알 수 없는 알고리즘에 이끌려 가끔 알다가도 모를 세상에 살고 있는 듯한 느낌도 받습니다. 알고리즘이 추천하는 서비스와 광고에 노출되어 소비하듯, 우리는 수동적으로 대세에 맞춰 그대로 기술의 행동 유도성을 따르곤 합니다.

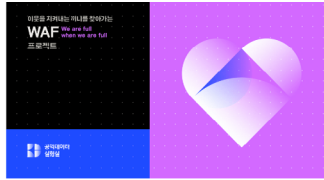
뭔가 단단히 잘못된 것 같긴 한데, 나만 이런 건지 고민했던 적 혹시 있으신가요? 디지털 시대이기에 더 많은, 더 나은 민주주의가 필요한지도 모릅니다. 빠뜨리는 ‘믿을 수 있는 데이터를 투명하게 공개한다면’이라는 질문에서 그 실마리를 찾았습니다.

시민을 위한, 시민에 의한 데이터란 어떻게 가능할까요? 모두에게 공개되고 모두가 활용할 수 있는 데이터 활동은 어떻게 가능한지 묻는 데에서 데이터 액티비즘은 출발합니다. ‘모두가 참여할 수 있고, 모두에게 이로운 공익데이터’의 세계로 여러분을 초대합니다!

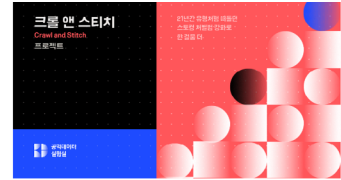
빠디는 시민이 주도하는 열린 데이터와 시민 기술 생태계를 만듭니다.



시민사회 공익데이터활동



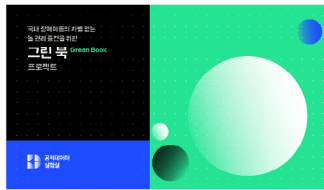
시민에게 도움이 되는 데이터, We are full when we are full



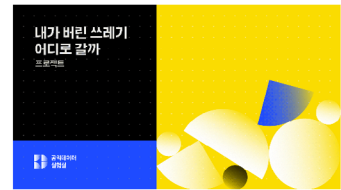
끝날 때까지 끝난 게 아니다, 크롤&스티치



앞으로도 계속 될 서로넷의 노동인권복지 이야기



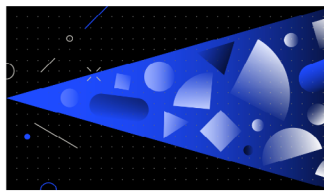
한뼘 한뼘 현장조사로 만들어낸 데이터가 세상을 바꿀 때까지



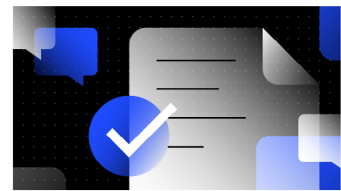
쓰레기를 따라 걷는 여행, 그 막다른 종착역



시민중심 도시데이터 생태계 구현



공익데이터실험실 가을스프린트 회고



설문조사! 우리가 직접 하고 싶을 때 '꼭' 알아야 할 것들: 사회조사 기초 강의

빠디는 시민과 함께 공익데이터를 수집하고,
시빅 해커들의 커뮤니티를 지원하며,
공익데이터를 구축하는 열린 플랫폼을 만듭니다.

모두의 참여와 기여로 공익데이터를 구축하고,
시민들이 데이터에 대한 권리를 갖고 함께 사회문제를 해결합니다.

데이터 교육 데이터로 문제를 해결하기 위한 리터러시 교육과 데이터 액티비즘 활동을 기획하고 실행하는 과정을 제공합니다.

데이터트러스트 시민 주도로 공익 활동에 필요한 데이터를 모으는 공익데이터 플랫폼을 활성화합니다.

목차

Intro.		06
가이드북 활용방법		

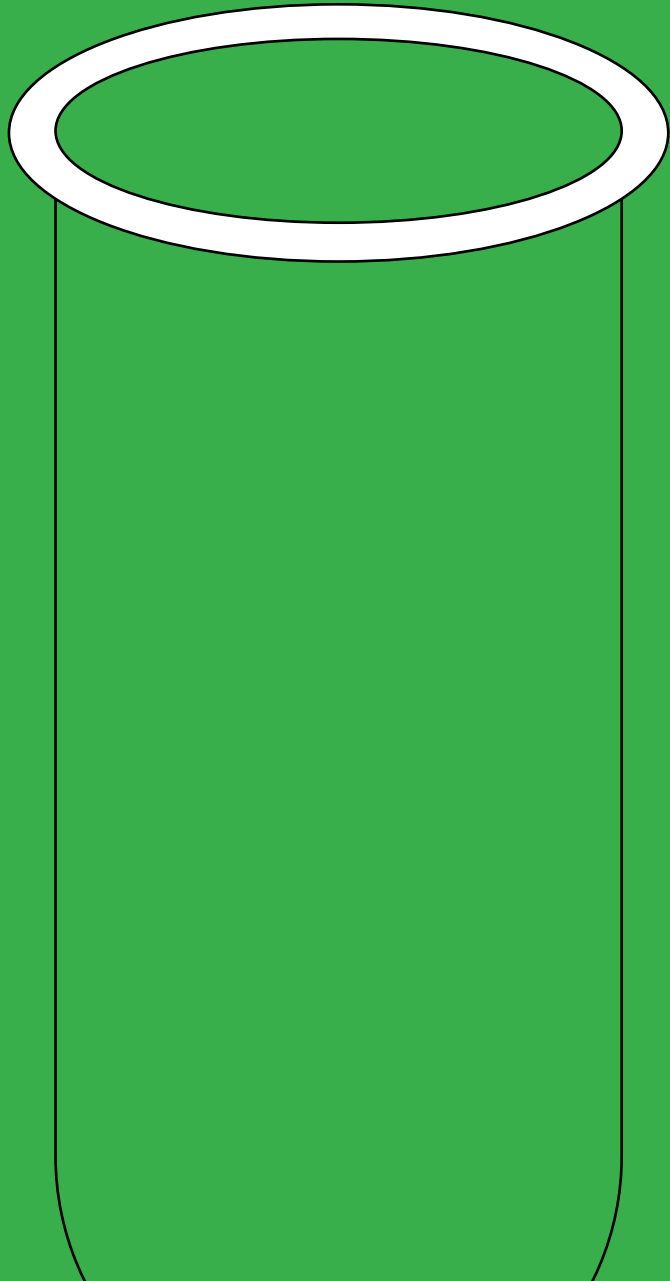
Chapter1.	01. 문제 탐색하기	08
액션 플랜 세우기	02. 문제 정의하기	09
	03. 데이터 액션 플랜	11

Chapter2.	01. 데이터 수집	14
데이터 활동 시작하기	02. 데이터 정제	17
	03. 데이터 분석	19
	04. 데이터 시각화	19
	05. 데이터 액션	21

Outro. 부록	문제 해결을 위한 데이터 액션 캔버스	23
: 데이터 액티비즘 툴킷 (Tool kit)	더 나은 협업을 위한 오픈소스 프로젝트 향해 가이드	24

Intro.

가이드북 활용방법



누구나 시작하고 시도할 수 있는 공익데이터 에티비즐 가이드북

공익데이터 액티비즘 가이드북은

누구나 쉽게 데이터 액티비즘을 시작할 수 있도록 돕는 툴킷입니다.

이럴 때 활용하세요	— 문제 정의가 어렵거나 시나리오가 그려지지 않을 때 — 데이터 기반 문제해결 활동을 시작하고 싶을 때 — 프로젝트 안에서 방향을 점검하고 싶을 때
-----------------------	--

이런 경험을 만들 수 있어요	— 내가 관심있고 공감하는 문제를 구체적으로 정리할 수 있어요. — 사용자 경험 중심의 디지털/데이터 문제해결 프로젝트를 시도할 수 있어요. — 데이터 기반의 활동 툴킷으로, 필요한 데이터나 기술과의 연계가 가능해요.
--------------------------------	---

챕터1은 데이터 활동을 위한 가이드를 만드는 단계입니다.

문제(관심 주제)를 설정하기 위한 밑작업으로서 다양한 이슈와 관련 데이터들을 탐색하고, 이를 바탕으로 핵심문제를 정의하며, 구체적인 데이터 수집계획을 세워보는 순서로 구성됩니다.

챕터2는 본격적인 데이터 액션 플랜을 도출하는 단계입니다.

빠디에서 진행한 다양한 공익데이터 활동은 모두 이 프로세스를 따라 진행됩니다. ‘데이터 수집 → 정제 → 분석 → 시각화 → 솔루션(액션)’의 단계별 방법론과 팁들을 제공합니다.

더 구체적인 공익데이터의 활동사례가 궁금하다면!

사회문제를 발견하는 데이터 액티비즘

→ [팔공프로젝트 : 대구시 어린이 교통사고 및 교통안전 시설물 분석](#)

사회문제 해결을 촉구하는 데이터 액티비즘 : 강서구 장애아동

→ [그린북프로젝트 : 강서구 장애아동친화놀이터 데이터 수집](#)

흩어진 데이터를 모아 아카이빙하는 데이터 액티비즘

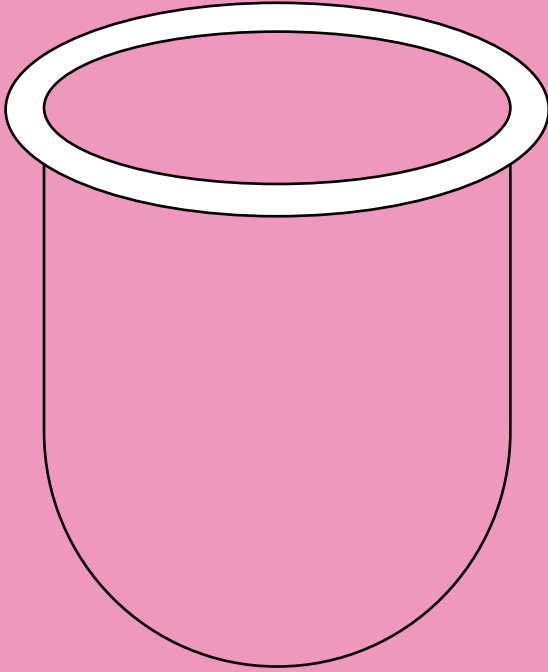
→ [내가 버린 쓰레기 어디로 갈까?](#)

없는 데이터를 만들어 쓰기 좋게 공유하는 데이터 액티비즘

→ [일상관찰가 프로젝트 : 춘천 우리동네리어나카 데이터](#)

Chapter1.

데이터 액션 플랜 세우기



**01. 문제
탐색하기**

어떤 문제를 다룰 것인가? 내 관심 분야에는 어떤 데이터들이 모여 있을까?

처음부터 핵심 문제를 정확하게 정의하기는 어렵겠죠.

문제 탐색은 본격적인 데이터 활동계획을 세우기에 앞서, 문제의식을 구체화하고 가설을 세워보기 위한 밑작업에 해당합니다.

어떤 데이터들이 있는지 자유롭게 찾아보고 아이디어션해봅니다.

명확히 정리되지 않는다면 메모해둔다는 마음으로, 다양한 데이터를 자유롭게 찾아보세요.

어디에 들어가서 뭘 찾아야할지 모르겠다면, 아래 자료를 참고하세요!

자료 검색 데이터, 어디서 찾을 수 있을까?

직접 데이터를 수집해요

■ 정보공개청구

<https://www.opengirok.or.kr> 에서 시민으로서 정보공개 청구를 진행하고, 이의제기를 해보세요.

■ 크라우드 소싱 Crowd Sourcing

사람들이 공동으로 데이터를 생산하고 수집하는 활동입니다. 구글 설문, 플랫폼을 활용해 데이터를 만들어보세요.

■ 웹파싱, 크롤링

웹페이지 데이터 등 데이터를 모을 수 있어요.
구글 스프레드시트, 엑셀로도 가능해요.

■ 논문 데이터 활용

RISS, KISS, 구글학술, 국회도서관, 네이버 전문 정보 사이트에서 키워드검색해 살펴보세요.
원데이터 출처를 찾아보세요.

■ 기구, 기관, 단체 데이터 공개 요청

국제기구, 공공기관, 비영리단체 공개된 데이터를 홈페이지에서 찾고, 이메일로 공개 요청해보세요.

데이터 포털, 검색엔진 - 키워드

■ 공공데이터포털 data.go.kr

■ 서울시 열린데이터광장 data.seoul.go.kr

■ 국가통계포털 kosis.kr

■ 뉴스타파데이터포털 data.newstapa.org

■ 데이터스토어 datastore.or.kr

■ 한국데이터거래소 kdx.kr

■ 데이터안심 구역 dsz.kdata.or.kr

■ 구글 google.com

■ 깃헙 github.com

■ 캐글 kaggle.com

기관, 지역, 주제, 분야에 따라

다양한 데이터 포털과 검색엔진을 활용해요.

— 키워드, 키워드 + 데이터로 검색해보세요.

— 원데이터인지 아닌지도 항상 체크해보세요.

— 믿을 수 있는 출처에 대한 고민을 해보세요.

**02. 문제
정의하기**

어떤 문제를 다루고 싶나요?

데이터를 기반으로 보여주고 싶은 사회문제나, 드러내고 싶은 문제 상황을 정리해봅시다.
데이터는 그 자체로는 의미를 가지지 않기 때문에, 어떤 문제와 어떤 가설을 세우는가에 따라
수집하고 다듬을 데이터의 방향이 정해질 수 있어요.

문제 상황 : 어떤 문제 상황에 놓여 있나요?

내 관심 이슈를 바탕으로, 데이터를 통해 접근하고 싶은 문제를 그려볼까요?

문제의 시기, 장소, 이해관계자, 관련 제품이나 서비스, 제도 등을 생각나는대로 적어보세요.
문제가 발생하는, 혹은 문제를 겪은 상황을 자세히 묘사할수록 문제는 명확해집니다.

문제 정의 : 핵심 문제가 무엇인가요?

위 문제 상황을 요약해, 핵심 문제를 하나의 문장 혹은 키워드로 정리해보세요.

당사자 : 문제 상황의 당사자는 누구인가요?

아래 양식을 참고해 문제의 당사자 퍼소나(persona) 프로필을 작성해보세요.

당사자의 특성이나 생각, 행동, 처한 상황을 구체적으로 디자인한다면 해결하고 싶은 문제와 데이터 수집의 방향도 더욱 분명해질 수 있어요.

<p>사진(그림)</p>	<p>라이프스타일 #키워드1 #키워드2 #키워드3 — — —</p>
<p>이름 : 나이 : 직업 : 가족관계 : 사는 지역 : 캐릭터 — 성향, 특성, MBTI..</p>	<p>문제 상황에서의 불편사항(불만) — — 문제 상황에서의 니즈 — —</p>

사용자 경험 : 이 데이터 액션(활동)은 당사자에게 어떤 경험을 줄 수 있나요?

문제정의, 문제해결, 사회변화 등 이 데이터를 수집하고 활동을 발전시킬 때 어떤 경험을 줄 수 있나요?

문제의 당사자와 나를 포함한 활동참여자, 활동 과정에서 도움을 줄 수 있는 기여자의 경험이 각각 다를 수 있겠죠?

<ul style="list-style-type: none"> — 문제의 당사자 — 이해관계자 (기획자, 참여자 등) — 기여자 (ex. 데이터 제공자, 분석가, 디자이너 등)

03. 데이터 액션 플랜

다뤄보고 싶은 문제 혹은 주제 분야가 조금 더 분명해졌나요? 이제 위에서 정의한 문제 해결에 기여하거나, 문제를 발전시키는 데 도움이 되는 데이터를 생각해봅시다.

‘문제 탐색’ 단계에서 모은 데이터들 가운데 관련이 있을 것 같은 데이터, 필요하다고 생각하는 데이터 목록을 나열해보는 것도 좋은 방법이에요. 확실하지 않아도 괜찮아요. 앞으로 꾸준히 활동 방향과 데이터를 다듬어나갈 수 있으니까요.

수집하고 싶은 데이터 : 무엇을 수집할 건가요?

* 여기서 데이터란 수식이나 지표로 된 엑셀 데이터 외에 이미지나 텍스트 등의 비정형 데이터를 모두 포함합니다.

데이터 수집계획 : 궁금한 데이터를 모으고 정리하기 위한 구체적인 계획을 적어보세요.

내가 설정한 문제와 관련해 공개된 데이터가 있나요? 안내한 링크를 참고해 공공데이터, 민간데이터를 열심히 찾아봅시다. 원하는 데이터가 없을 수도 있어요. 데이터가 없거나 새롭게 만들어야 하는 경우, 아래 내용을 참고해보세요.

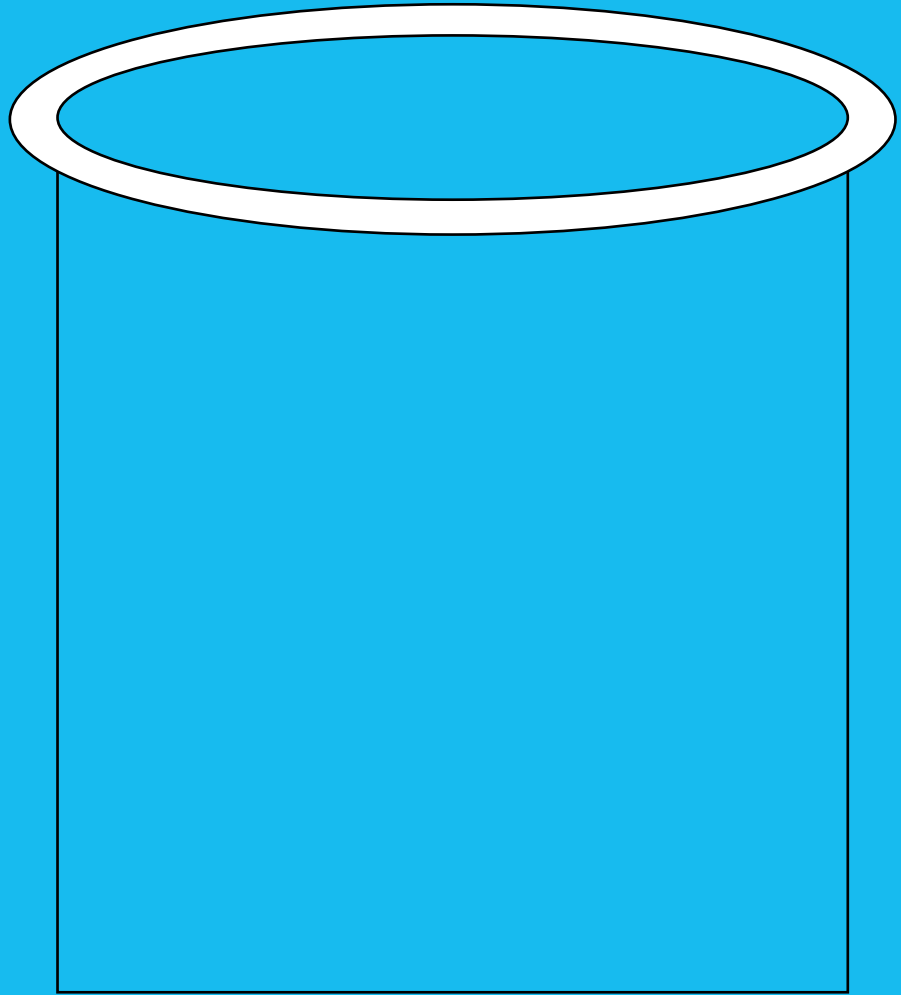
- 혼자서 수집하기 어렵다면, 누구와 얼마나 함께 해야 하나요?
- 어떤 데이터를 주로 다룰까요? (형식, 수집방법, 기간 등)
- 수집한 데이터를 어떻게 입력할 건가요?

수집한 데이터를 어떻게 정리하면 이 문제를 효과적으로 보여줄 수 있나요? 결과물로 보여주기 위한 솔루션, 데이터를 통해 실천하고 싶은 액션을 구체적으로 적어보세요.

(ex. 차트나 그래프로 시각화한다, 완성된 스프레드시트에서 문제 이슈를 콘텐츠로 정리한다)

Chapter2.

데이터 활동 시작하기



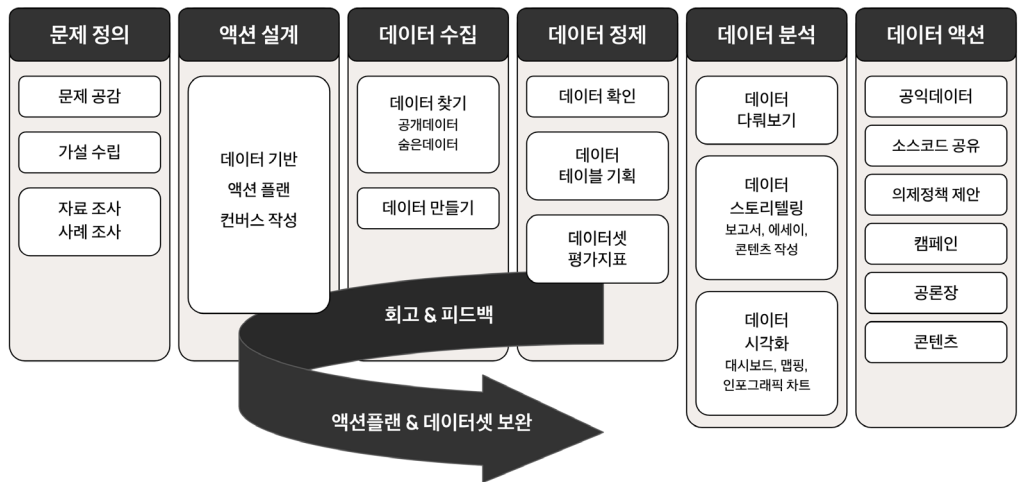
이제 데이터의 바다를 향해가기 위한 준비 작업은 끝났어요.
 정의한 문제와 수집계획에 따라 ‘나만의 방식’으로 프로젝트를 진행해봅니다.

데이터 액티비즘의 활동 실험은,
 데이터를 기반으로 협력 프로젝트를 진행하는 것이 핵심이라고 할 수 있어요.

실험을 시작하기위한 몇 가지 약속

- 어려운 문제와 궁금한 질문을 지속적으로 탐구해요.
- 작은 행동과 기여로 활동을 진행해보아요.
- 협력을 중요하게 생각하고, 함께 기록하고 공유해요.
- 회고와 성찰을 병행하며, 점진적으로 개선해나가요.
- 오픈소스로 결과물을 공개해요.

데이터 액티비즘 순서도



데이터 액티비즘이 이루어지는 전체 프로세스는 위와 같습니다. 단순히 수집하고 그래프나 차트를 만들기만 하는 것이 아니라, 정의한 문제를 바탕으로 맥락과 의미를 갖고 스토리텔링을 진행하는 과정이 데이터 액티비즘의 핵심입니다.



나의 데이터 활동에 적절한 방식은 무엇일까요?

본격적인 '수집 > 정제 > 분석/시각화 > 액션(실천활동)'의 과정을 단계별로 알아봅니다!

01. 데이터 수집

Find Data

데이터를 수집하는 방법은 굉장히 다양해요.

적절한 방법을 선택하기 위해서는 우선 내가 수집하고 싶은 데이터의 특성을 잘 알아야 합니다.

내가 보고 싶은 데이터는 어떤 데이터인가요?

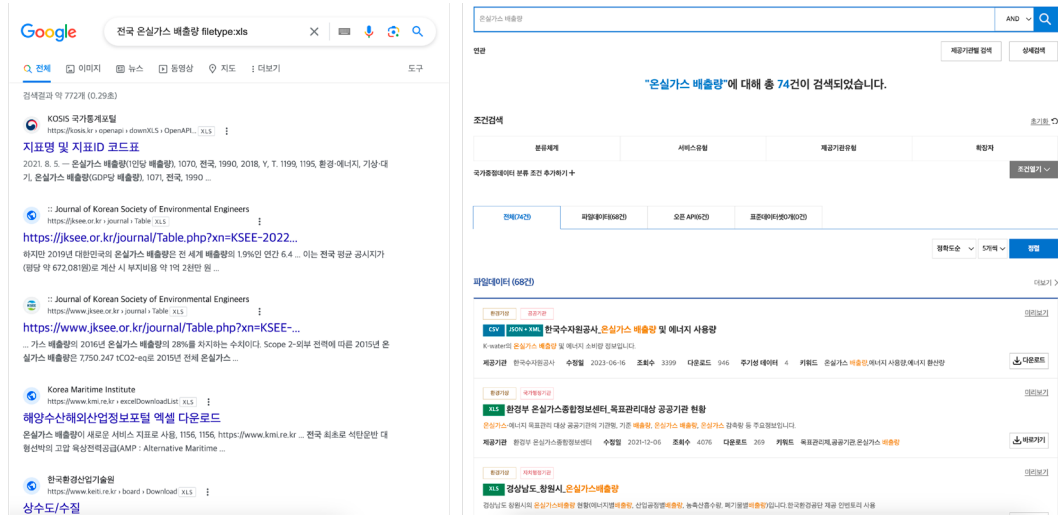
- 공개된 데이터가 있나요? 공공데이터, 민간데이터를 찾아보세요.
- 숨어있는 데이터가 있나요? 정보공개청구 또는 데이터 제공신청 등을 이용해보세요.
- 데이터를 직접 만들어야 할까요? 크라우드 소싱으로 데이터를 만드는 작업을 설계해보세요.
 - 얼마나 많은 사람들과 함께 해야 할까요?
 - 어떤 데이터를 주로 다룰까요?
 - 어떤 식으로 데이터를 입력할 수 있나요?

위 체크리스트에 따라, 아래와 같은 수집방법을 활용할 수 있어요.

- [] 공개되어 있는 데이터인지 : 자료 검색
- [] 숨어있는 데이터인지 : 정보공개청구, 공공데이터제공신청, 민간데이터 제공협조요청 등
- [] 새로 정의하여 만들어야 하는 데이터인지 : 크라우드소싱, 현장조사, 전화 인터뷰

공개되어 있는 데이터라면 : 자료 검색

자료 검색은 가장 기본적인 데이터 수집의 수단이죠. 검색을 ‘잘’ 하는 것만으로도 대부분의 데이터를 찾을 수 있어요. 공공기관 문서나 데이터는 주로 구글검색(구글링)으로도 쉽게 접근할 수 있고, 기사의 경우 국내포털 검색이 효과적일 수 있습니다.



위와 같이 구글에 원하는 파일형식(pdf, xls)이나 연도별로도 원하는 데이터를 효과적으로 검색할 수 있어요. 공공정보의 경우 주요 키워드의 검색만으로도 관련된 관공서 홈페이지나 공공데이터포털, 혹은 이슈별로 체계적인 데이터를 아카이브하고 관리하는 포털사이트로 연결될 수 있습니다.

보다 자세한 자료 검색처는 **챕터1의 ‘문제 탐색’** 파트를 다시 한번 참고합니다. 충분한 검색만으로도 관련 데이터에 쉽게 접근하고 이해도를 높이는 것이 가능할 거예요.

숨어있는 데이터라면 : 정보공개청구, 데이터 제공신청

데이터 검색으로 원하는 데이터를 얻지 못한다면? 충분히 받아볼 수 있는 공공데이터임에도 내가 원하는 조건(지역, 연령 등)의 데이터는 개방되어 있지 않다면? 정보공개청구나 공공데이터 제공신청을 활용할 수 있습니다.

정보공개청구는 목적을 묻지 않는 것이 원칙이지만, 공공데이터 제공신청의 경우 활용목적과 원하는 존재하는 데이터의 형태인지 등을 분명히 명시해야 한다는 차이가 있으니, 이를 구분해 활용하는 것이 중요합니다.

쉽고 간단하게 배우는 정보공개청구 출처 : 투명사회를 위한 정보공개센터

정보공개 청구, 누가 할 수 있나?

대한민국 국민이라면 누구나 정보공개청구가 가능합니다. 청소년도, 고도스에 있는 수형자도, 외국에 체류 중이라도 모두 가능합니다. 또한 외국인의 경우라도 ①국내에 일정한 주소를 두고 거주하거나 학술·연구를 위해 일시적으로 체류하는 사람이나 ②국내에 사무소를 두고 있는 법인 또는 단체라면 정보공개청구가 가능합니다.

☞ 공공기관의 정보공개에 관한 법률(이하 정보공개법) 제5조와 정보공개법 시행령 제3조

정보공개 청구, 어디에 할 수 있나?

공공기관이라면 어디든 정보공개청구가 가능합니다.

정보공개법에 따르면 공공기관이란 국회, 법원, 헌법재판소, 중앙선거관리위원회와 중앙행정기관 및 관련 법률에 따른 위원회 등의 국가기관과 지방자치단체, 그리고 관련 법률에 따른 공공기관과 대통령령으로 정하는 기관(예, 대학교, 중고등학교, 지방공사 등)을 말합니다.

자세한 내용은 어렵지 않으니 근거 법률을 국가법령정보센터에서 검색해서 찾아보셔도 좋습니다. 그래도 확인이 어려우시면 정보공개센터로 연락 주세요.

☞ 정보공개법 제2조와 공공기관의 운영에 관한 법률 제2조, 정보공개법 시행령 제2조

정보공개 청구, 어떻게 할 수 있나?

정보공개청구는 '정보공개청구서'라는 양식에 해당하는 내용을 작성하여 우편·팩스·직접방문·인터넷 등 다양한 방법으로 청구할 수 있습니다. 정보공개포털 홈페이지에서 온라인으로 정보공개 청구를 하면 매우 쉽고 간단하게 청구가 가능합니다. 또한 직접 청구서 작성이 어려우신 경우는 정보공개청구를 하실 해당 공공기관에 가서서 정보공개 담당자에게 안내를 받아 '정보공개 구술 청구'를 하실 수도 있습니다.

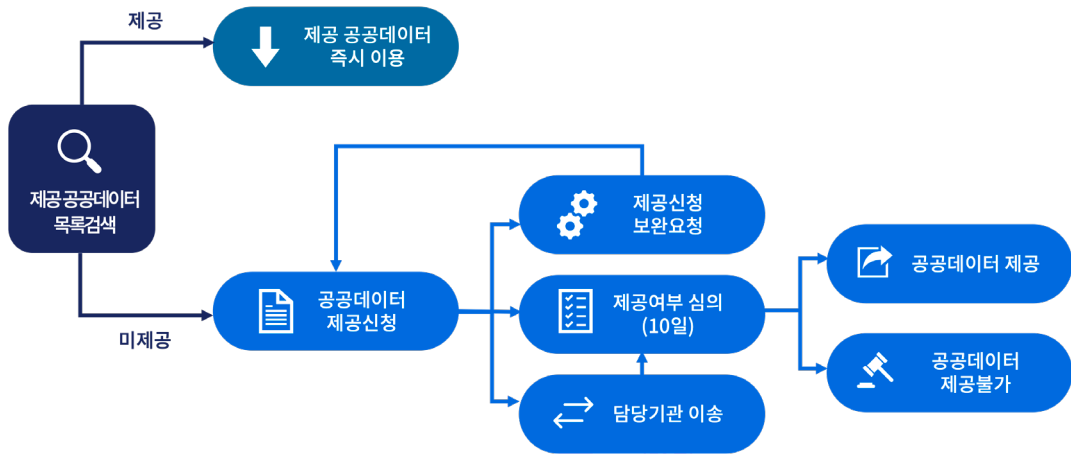
온라인으로 정보공개청구하기

자, 그럼 가장 간단하고 편리한 온라인 정보공개청구를 본격적으로 함께 해보겠습니다. 온라인을 통해 정보공개청구를 하기 위해서는 가장 먼저 정보공개포털 홈페이지에 접속해야 합니다. 대한민국 대부분의 공공기관에 대한 정보공개청구는 '대한민국 정보공개포털' (<https://www.open.go.kr/>)을 통해서 할 수 있습니다.

정보공개포털은 정보공개청구 이외에도 원문공개, 정보목록 검색, 사전공표정보 게시, 정보공개제도 안내, 각종 관련 서식 제공, 수수료 안내 등 정보공개에 관한 통합 서비스를 제공하는 홈페이지입니다.



공공데이터포털을 활용한 데이터 제공신청 출처 : 공공데이터포털



데이터를 직접 만들어야 한다면 : 클라우드소싱

내가 원하는 데이터가 검색이나 제공신청을 통해 얻을 수 없다면? 문제의식은 분명한데 이를 종합해 보여줄 단일 데이터셋이 아직 세상에 존재하지 않는다면? 이를 처음으로 만들어내는 작업이 필요하겠죠.

데이터 클라우드소싱이란 다른 사람들과 함께 데이터셋을 구축하는 작업으로, 시민데이터 활동의 중요한 요소이기도 합니다. 한 가지 사례로 2020년 강서구 장애아동친화놀이터 관련 데이터 문제를 다룬 '스몰빅'팀은 클라우드소싱을 통해 장애아동친화지표로 검증하기 위한 **자체적인 데이터셋**을 만들기도 했어요.

#	구분	놀이시설명	주소	전화번호	위도	경도	면적	홈페이지	점검일	점검자	놀이에 활제어, 유모차 접근가능 여부	신체장애아 등이 함께 놀 수 있는 놀이기구	다양한 감각을 발달시킬 수 있는 놀이기구	휠체어, 유모차가 접근 가능한 화장실	기타 교편 실자 확인
1	공원	강서환경공원 어린이놀이터	서울특별시 강서구 방화2동	02-3780-0621	37.5859954	126.8171323		hangeans	2020. 10. 15	손한진	예	없음	모래놀이터	남자, 여자 화장실	여자
2	근린공원	공원내부근린공원	서울특별시 강서구 가양동	02-2600-4186	37.57063459	126.849086	25,680		2020. 10. 16	전희진	예	없음	없음	화장실 없음	없음
3	근린공원	구암근린공원	서울특별시 강서구 허문포동	02-2600-4186	37.57599867	126.8537291	29,844		2020. 10. 27	김홍주	예	안전벨트가 있는	없음	남자, 여자 화장실	여자
4	근린공원	경신근린공원	서울특별시 강서구 가양동	02-2600-4186	37.57425363	126.8413319	135,002.7		2020. 10. 30	이혜진	예	안전벨트가 있는	모래놀이터	화장실 없음	없음
5	근린공원	까치산근린공원	서울특별시 강서구 화곡동	02-2600-4112	37.53693781	126.8439364	1,805		2020. 10. 12	이은주	아니오	없음	맛줄을 잡고 오르	화장실 없음	없음
6	근린공원	등촌근린공원	서울특별시 강서구 경시동	02-2600-4112	37.56368324	126.8413901	10,052		2020. 10. 16	전희진	예	없음	없음	여자 화장실	여자
7	근린공원	미국나무근린공원	서울특별시 강서구 미국동	02-2600-4112	37.566183	126.813091	17,283		2020. 10. 15	손한진	예	없음	잔디 운동장	화장실 없음	없음
8	근린공원	미국하늬근린공원	서울특별시 강서구 서울동	02-2600-4112	37.556699	126.812659	24,597		2020. 9. 26	문다정	예	없음	없음	남자, 여자 화장실	남자
9	근린공원	별산근린공원	서울특별시 강서구 미국동	02-2600-4112	37.55453963	126.8258374	17,826		2020. 9. 26	이지영	예	없음	없음	여자 화장실	없음
10	근린공원	병화산근린공원	서울특별시 강서구 금남동	02-2600-4186	37.580907	126.8141197	72,132		2020. 10. 15	손한진	예	없음	모래놀이터	남자, 여자 화장실	없음
11	근린공원	병재산근린공원	서울특별시 강서구 화곡4동	02-2600-4184	37.539906	126.84975	917,190		2020. 10. 15	민형근	예	없음	없음	화장실 없음	없음
12	근린공원	서남성근린공원	서울특별시 강서구 등촌동	02-2600-4112	37.55886513	126.849984	10,249		2020. 10. 16	전희진	예	없음	없음	여자 화장실	여자
13	근린공원	서울식물원근린공원	서울특별시 강서구 미국동	02-2600-4112	37.569783	126.830068	503,875		2020. 10. 16	전희진	예	없음	없음	여자 화장실	여자
14	근린공원	수령산근린공원	서울특별시 강서구 경시동	02-2600-4112	37.55066787	126.8298924	83,807		2020. 9. 26	김민영	예	없음	없음	화장실 없음	없음
15	근린공원	용기굴근린공원	서울특별시 강서구 양정동	02-2600-4186	37.57356247	126.8216952	9,724		2020. 10. 13	김은경	예	없음	없음	남자, 여자 화장실	남자
16	근린공원	우정근린공원	서울특별시 강서구 우정산	02-2600-4185	37.55282044	126.843808	359,435		2020. 10. 14	민형근	아니오	충심물 잡기	아니오	화장실 없음	없음
17	근린공원	원당근린공원	서울특별시 강서구 경서동	02-2600-4112	37.559782	126.8425661	10,096		2020. 9. 26	김유현	예	없음	없음	남자, 여자 화장실	없음
18	어린이공원	가곡어린이공원	서울특별시 강서구 내방산	02-2600-4181	37.559964	126.837765	2,748		2020. 9. 26	김유현	아니오	없음	없음	화장실 없음	없음
19	어린이공원	가양 근동어린이공원	서울특별시 강서구 가양동	02-2600-4183	37.562405	126.857527	1,612		2020. 10. 16	전희진	예	없음	없음	여자 화장실	여자
20	어린이공원	가양 무지개어린이공원	서울특별시 강서구 가양동	02-2600-4183	37.565086	126.851657	1,500		2020. 9. 26	조성훈	예	없음	없음	화장실 없음	없음
21	어린이공원	가양 비둘기어린이공원	서울특별시 강서구 가양동	02-2600-4183	37.559504	126.862944	1,623		2020. 10. 16	전희진	예	없음	없음	여자 화장실	여자
22	어린이공원	가양어린이공원	서울특별시 강서구 가양동	02-2600-4183	37.569758	126.844789	1,500		2020. 10. 15	손한진	예	없음	없음	화장실 없음	없음
23	어린이공원	개화어린이공원	서울특별시 강서구 방화동	02-2600-4183	37.580248	126.812534	1,512		2020. 10. 15	손한진	아니오	없음	없음	화장실 없음	없음
24	어린이공원	굴방어린이공원	서울특별시 강서구 화곡동	02-2600-4181	37.53331316	126.8560495	978		2020. 10. 11	이혜진	아니오	없음	없음	화장실 없음	없음
25	어린이공원	광명어린이공원	서울특별시 강서구 외방산	02-2600-4181	37.546857	126.822944	1,739		2020. 10. 14	박시현	예	없음	다양한 시각, 촉각	화장실 없음	없음

클라우드소싱의 대표적인 방식으로 설문조사가 있습니다. 데이터 스토리텔링을 위해, 스토리의 큰 부분을 제공하는 개인들로부터 데이터를 수집하기 위한 효과적인 도구입니다.

또 한 가지 방법은, 데이터셋 스프레드시트 자체를 함께 공유하는 것입니다. 굉장히 간단하고 당연한 방법이지만, 이것이 데이터 활동 영역에서 활발히 활용된 것은 불과 몇 년 사이의 일이지요. 2017년 1월 미국에서 여성 행진 예상 군중과 관련한 **데이터셋**을 만들었을 때, 간단한 작성가이드와 함께 만들어진 데이터셋이 소셜미디어를 통해 공유되고, 전 세계 모든 사람들이 사이트에 데이터를 추가했던 사례는, 클라우드소싱의 강점과 가능성을 잘 보여줍니다.

02. 데이터 정제 Clean Data

이제 기존의 문제의식과 가설을 바탕으로 수집한 데이터를 잘 정제해봅시다. 사실 수집을 시작하기 전 초기 데이터셋을 구성하는 작업과 수집 이후 이를 정제하는 작업은 하나로 연결되어 있습니다.

데이터 스토리텔링을 위해 목적에 맞게 원하는 형태의 ‘행’과 ‘열’을 만들며, 복잡하고 지저분하게 수집된 데이터를 일관성 있게 정돈(tidy data)하고 가공하는 과정이기도 합니다.

데이터 수집 & 정제 체크리스트

어떤 데이터 형태인가요?

- 수치형 데이터, 테이블 형식으로 시작해보세요.
- 가공한 데이터인지, 원데이터인지 확인해보세요.

열과 행은 어떻게 구성되었고, 관계형 데이터인가요?

- 어떤 데이터 형태인가요?
- 데이터 수가 몇 개인가요? 열과 행의 개수를 세어주세요.
- 열과 행의 세부사항을 적어주세요.
- 데이터 간 관계를 만들어볼 수 있는지 살펴보세요.

데이터 오류를 가지고 있나요?

- 모든 데이터에 하나의 값(key값 포함)이 정확히 들어갔는지 확인해보세요.
- 데이터 중 비어있거나 중복되는 건 없는지 확인해보세요.
- 잘못된 데이터가 적혀있는지 확인해보세요.

출처가 믿을 수 있는 데이터인가요?

- 어떻게 수집했는지 수집방법을 살펴보세요.
- 데이터가 가공되어 있는지를 확인해보세요.
- 투명하게 공개되어 있는 데이터인지 확인해보세요.

데이터를 활용하는 목적은 무엇일까요?

- 사람들이 이해할 수 있는지 검토해보세요.

데이터셋 명세서

- 활용 목적
- 주제 분야
- 데이터 형태 (파일 형식)
- 데이터 수집 방식
- 조사 대상
- 조사 기간
- 데이터 항목 (특성)
- 연도 (업데이트 추가)
- 이용 시 유의사항

‘문제 정의 > 데이터 수집 > 데이터 정제’의 과정은 단순히 순차적으로 진행된 후 마무리되는 것은 아닙니다. 수집 및 정제 과정에서 문제의식이 수정되거나 확장될 수 있고, 이 경우 데이터셋의 구성 또한 바뀔 수 있기 때문이죠.

따라서 ‘데이터 정제 > 피드백 및 회고 > 문제 재정의(구체화)’ 과정이 순환적으로 이루어짐으로써 데이터 활동의 짜임새를 더 하는 것은 액티비즘 프로세스 안에서 중요한 과정이 됩니다.

03. 데이터 분석 Analyze Data

내가 하고 싶은 스토리텔링을 기획하고, 지금까지 만들고 모아온 데이터를 바탕으로 맥락과 의미를 찾고 만드는 과정입니다.

일반적인 데이터 분석의 기법을 활용하지 않더라도, 공익데이터 액티비즘의 방향 안에서 데이터 수집과 솔루션 도출을 위한 의미를 부여하는 단계로 이해하면 좋을 거예요. 이를 위해 우리는,

- 데이터에서 의미 찾기 : 목적과 부합하는 정보가 들어있는지 파악하고 해석해봐요.
- 데이터의 품질 파악하기 : 부족한 데이터, 데이터의 특징을 알아봐요.

데이터 분석을 통해 맥락과 의미 만들기

기술적(descriptive) 분석

- 주어진 데이터를 있는 그대로 보여주거나, 설명하는 분석 방법
- “어떻게 보여줄 것인가?” 혹은 “어떻게 정리할 것인가?”라는 질문과 연결돼요.

탐색적(exploratory) 분석

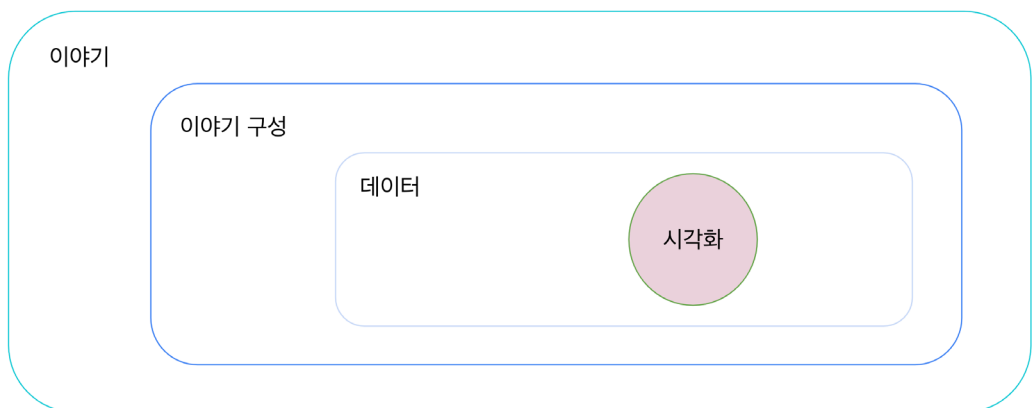
- 주어진 데이터를 탐색하여, 변수 간의 상관관계나 경향성을 찾아보는 분석 방법

확증적(confirmatory) 분석

- 주어진 데이터를 통계적으로 추론/분석해, 변수간의 인과관계 등 가설을 검증하는 분석 방법. 상관관계의 설명력을 높이고, 인과관계를 도출하기 위함.

04. 데이터 시각화 Visualize Data

“어떻게 하면 나의 데이터 스토리를 효과적으로 보여줄 수 있을까?”

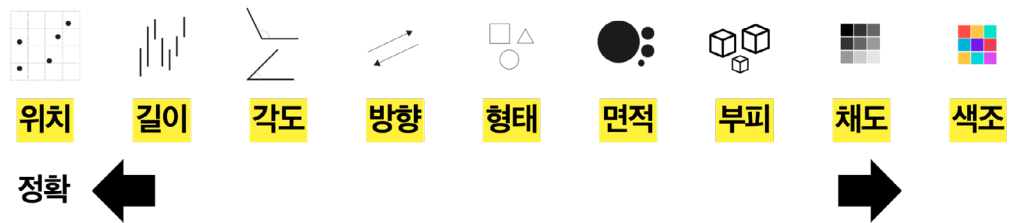


많은 데이터를 시각적 요소로 요약해서 데이터의 의미를 한눈에 파악할 수 있도록 전달하는 단계입니다. 데이터 시각화는 활용대상과 데이터를 파악하고, 핵심지표를 선정해 인사이트를 도출하는 과정으로 진행합니다.

데이터 형식에 따라 시각화 목적과 유형을 다르게 설정할 수 있어요. 비교, 추이, 비중, 관계, 지도 등의 분류 기준을 둘 수 있습니다.

데이터 형태	시각화 목적	
시계열 데이터 (시간변화에 따른) 범주형 데이터 (텍스트, 지역, 날짜, 숫자) 공간 데이터 (위치, 좌표)	비교 추이 비중 비율 관계 분포 지도	버블 차트 막대 차트 선 차트 타임라인 차트 파이 차트 트리맵 와플 차트 산점도 네트워크 시각화 맵핑

- 시각적 단서들을 사람들이 얼마나 정확하게 해석하는 지 통계학자들이 실험
- 사람들이 시각화를 어떻게 해석하고 정보를 인지하는 지 가늠할 수 있는 가이드



참고문헌 출처: 네이션 아우 - '데이터 포인트'
참고문헌 출처: Graphical Perception and Graphical Methods for Analyzing Scientific Data

데이터 시각화의 유용한 도구들을 함께 소개합니다! 코딩을 모르는 시민들도 어렵지 않게 작업할 수 있는 좋은 툴들이 다양하게 존재한답니다.

코딩이 필요없는 툴

- 스프레드시트, 구글맵,
- Data Wrapper, Flourish

어떨 때 활용할까?

- 코딩 없이 간단하게 시각화
- 쉽고 빠르게 시각화
- 막대차트, 선, 그래프 등 기본 유형을 어렵지 않게 활용

코딩 툴 혹은 코딩이 필요한 툴

- R의 ggplot2 등
- Python의 matplotlib, seaborn, folium 등
- Javascript의 chart.js, d3.js 등

어떨 때 활용할까?

- 반복적으로 차트를 제작해야 할 때
- 용량이 큰 데이터를 시각화할 때

05. 데이터 액션 Data Action

어떤 형태로 결과물(솔루션)을 도출할 것인지, 도출한 데이터를 바탕으로 구체적인 액션을 실천하는 단계입니다. 데이터 스토리텔링을 기반으로 처음 설정한 문제의식과 가설을 검증하고, 이를 확산하거나 직접적인 문제 해결에 기여할 수 있는 액션의 방식을 고민해보세요.

데이터셋/소스코드 공유하기

- 프로젝트를 통해 도출한 데이터셋 자체가 하나의 결과물이 될 수 있어요. 스프레드시트 데이터셋을 잘 정제해 누구나 쓸 수 있도록 오픈소스 링크로 공유합니다.
- 누구나 자유롭게 활용할 수 있는 자동화된 소스코드를 공유할 수도 있습니다.

의제 및 정책 제안하기

- 문제 해결에 기여하거나 문제 해결을 직접적으로 촉구하는 정책 및 의제 제안활동을 할 수 있습니다.

캠페인으로 확산하기

- 온오프라인 캠페인을 열어, 다양한 시민들에게 확산하거나 의견을 촉구할 수 있어요.
- 더 많은 캠페인의 방법과 사례가 궁금하다면 → [캠페인즈](#)

공론장으로 이슈 나누기

- 데이터를 통해 검증한 문제의식과 이슈들을 더 많은 시민들과 나누기 위한 공론장을 기획해 진행합니다.
- 2023년 시민 공약데이터 실험실에서는 [그린워싱 데이터 공론장](#)과 [세입자주거권 데이터 공론장](#)을 진행하기도 했어요.

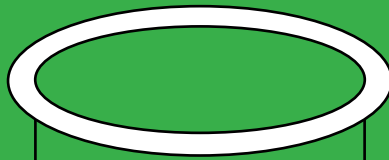
정당한 문제를 실천활동으로 보여주는 데이터 액션의 결과물은 6페이지의 공약데이터 활동사례에서 자세히 확인할 수 있어요.

Outro. 부록 : 데이터 액티비즘 툴킷 (Tool kit)

* 챗터1과 챗터2의 가이드를 툴킷 템플릿으로 제공합니다.

* 출처 : 빠띠 디지털 사회혁신 프로젝트 툴킷 v2021

누구나 시작하고 시도할 수 있는 공익데이터 액티비즘 가이드북



**문제해결을
위한 데이터
액션 캔버스**

프로젝트명 :	
문제 어떤 사회문제를 다룰 것인가요?	목표 어떤 변화가 있었으면 하나요?
대상 누구를 위한 프로젝트인가요?	기여자 누구와 함께 하는 프로젝트 인가요?
사용자 경험 프로젝트로 사람들은 어떤 경험을 할 수 있을까요?	
데이터 액션 플랜	
데이터 수집 단계 궁금한 데이터를 찾고 알아가는 과정으로 수집계획을 적어보세요.	데이터 정제 단계 원하는 형태의 행과 열로 데이터셋을 구성하고 다루는 과정으로 정제 계획을 적어보세요.
데이터 분석 단계 데이터를 살펴보고 스토리텔링을 하는 과정으로 질문, 이야기거리를 적어보세요.	데이터 시각화 단계 효과적으로 메시지를 전달하는 과정으로 시각화 목적과 시나리오를 작성해보세요.
데이터 액션 어떤 형태와 결과물로 데이터기반 활동을 할지 결정해보세요. []공익데이터 []소스코드 공유 []의제정책 제안 []캠페인 []공론장 []콘텐츠 []아카이브	
걸림돌 법률 이슈(개인정보, 저작권) 등 문제가 될 수 있는 부분을 적어보고, 주기적으로 대안을 고민해봐요.	마일스톤 프로젝트 진행상황을 체크리스트처럼 적어보고, 주기적으로 점검해봐요.
공익데이터셋명 :	
데이터셋 개요 데이터셋이 어떻게 구성되어있고, 만들어졌는지 설명해주세요. 링크 데이터셋 링크도 공유해주세요.	

더 나은 협업을 위한 오픈소스 프로젝트 향해 가이드

데이터 액티비즘 활동을 프로젝트로 진행해보고 싶다면,
더 나은 협력과 협업을 위한 향해 가이드를 참고해보세요.
프로젝트 과정과 회고를 기록해서 공유하고, 자료로 아카이브해볼 수 있어요.

향해일지 작성방법

- 구글문서, 노션, hack md, 깃헙페이지에 향해일지를 쓰기 위한 문서를 엽니다.
- 프로젝트 진행사항과 회고를 한 문단 씩입니다.

프로젝트명

- 프로젝트 활동가 이름 : (@아무개 식으로 적어주세요.)
- 프로젝트 소모임 일정 : (가안으로 적어주세요.)
- 우리 프로젝트에는 아래와 같이 강의자문 또는 활동가 기여가 필요해요.
(구체적으로 작성해주세요)
- 향해일지 링크 :

날짜/제목

기억에 남았던 활동

데이터활동 과정 속에 있었던 좋았던 점, 인상깊었던 일, 작업들을 적어주세요.

개선해야할 점

프로젝트에서 개선해야할 점이나 일감을 적어주세요.

필요한 점

자문, 교육이 필요한 사항을 적어주세요.

프로젝트 로그

프로젝트에서 결정한 사항, 진행사항을 회의록처럼 간단하게 적어주세요. 데이터, 자료 출처 url을 함께 적어주세요. 모든 걸 아카이브하고 싶은 마음에 사진도 남겨주세요!

- 일시 :
- 참여자 :
- 데이터, 자료 출처 url을 꼭 함께 적어주세요.
- 모든 걸 아카이브하고 싶은 마음에 사진도 남겨주세요.

누구나 시작하고 시도할 수 있는 공익데이터 액티비즘 가이드북

발행일 | 2023년 12월

발행처 | 사회적협동조합 빠띠

주소 | 서울특별시 서대문구 수색로 43 2층

홈페이지 | parti.coop

이메일 | contact@parti.coop

대표전화 | 070-7610-6417

기획·편집 | 사회적협동조합 빠띠

디자인 | 청년협동조합 몽땅

지원 | 아름다운재단

* 이 가이드북은 아름다운재단 2023년 변화의시나리오 지원사업으로 진행한 ‘시민 공익데이터 실험실’의 일환으로 제작되었습니다.

* 이 저작물은 **Creative Commons Attribution 4.0 Korea License**에 따라 이용할 수 있습니다.

뭔가 단단히 잘못된 것 같긴 한데, 나만 이런 건지 고민했던 적 혹시 있으신가요?
디지털 시대에 더 많은, 더 나은 민주주의가 필요한지도 모릅니다.
'믿을 수 있는 데이터를 투명하게 공개한다면'이라는 질문에서
빠뜨리는 그 실마리를 찾았습니다.

누구나 어렵지 않게 시도할 수 있는
'공익데이터'의 세계에 여러분을 초대합니다!

누구나 시작하고
시도할 수 있는
공익데이터 액티비즘
가이드북