

- 환경융합 빅데이터 플랫폼 -데이터셋 분석하기

2020.07

- 목차 -
- I. 환경데이터 포털의 구성
- Ⅱ. 데이터 얻기
- Ⅲ. 타블로 사용하여 데이터 분석하기
- IV. R-FLOW 사용하여 데이터 분석하기

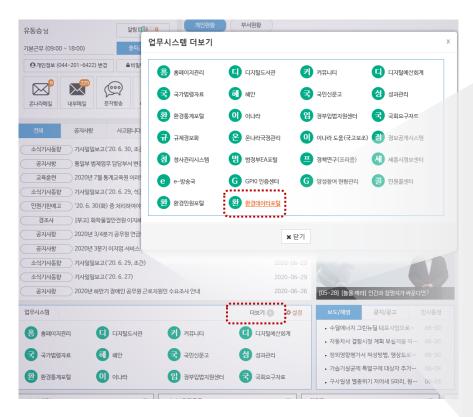




1. 환경데이터 포털의 구성

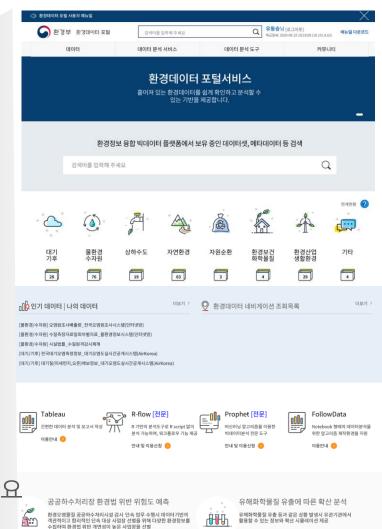
I.1 환경데이터 포털서비스를 아시나요?





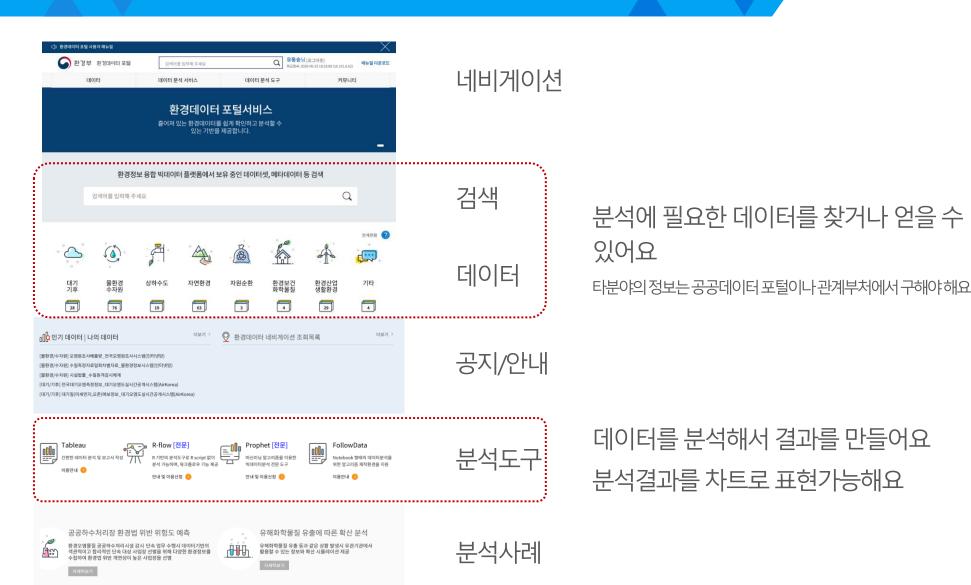


이지샘터 중·하단 업무시스템에서 더보기를 클릭하세요 팝업창이 나오면 하단에 있는 환경데이터포털을 선택하세요



I.2 환경데이터 포털서비스 화면





I.3 환경데이터 포털서비스 메뉴

자원순환

3

28

76

19

63

환경보건 화학물질

4

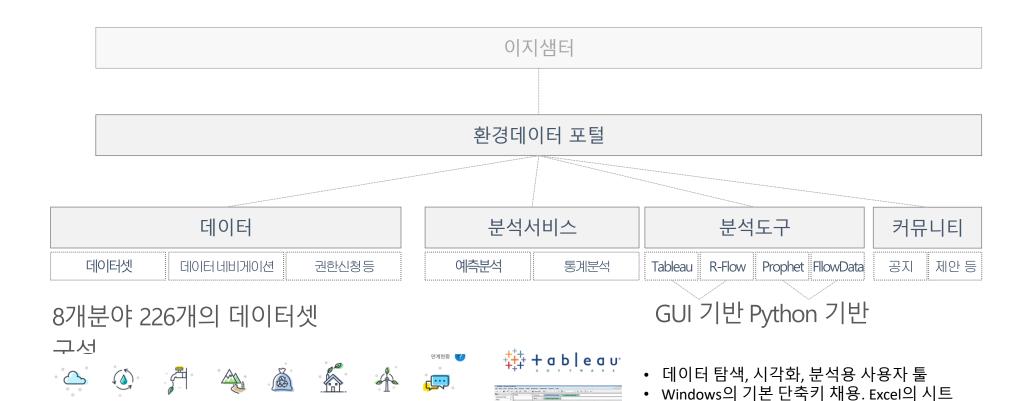
환경산업 생활환경

29

기타

4





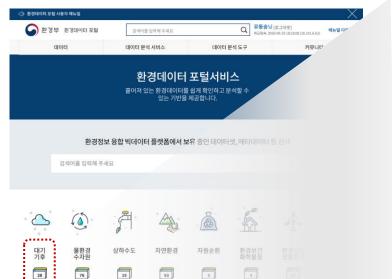
개념 채용.

직관적인 UI 구성

• IE의 화면 구성을 채용하여 인터넷방식의

Ⅱ. 데이터 얻기













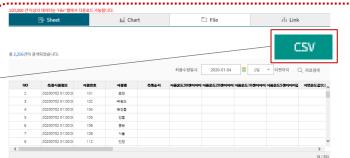




이지샘터 중·하단 업무시스템에서 더보기를 클릭하 팝업창이 나오면 하단에 있는 환경데이터포털을 선택하여 데이터셋의 내용을 확인합니다

데이터 개요

CSV 파일을 다운로드 할 수 있습니다.



테이터 형태

Ⅱ. 데이터 얻기(데이터 네비게이션)



노드 달힘		
3 노드 열림		■ 탄소포인트제 운영시스템
데이터셋 있음		☑ 통합기후변화홍보포털
		₫ 냉매정보관리시스템 ₫ 탄소종합관리
		○ 단도등합단다 ○ 온실가스배출량산정시스템
	온실가스감축 🗗	☑ 기후변화모델링시스템
	일바대기질 🚹	
대기/기후 🖸	자동차배출가스 🕒	국가온실가스종함관리시스템
	사업장대기질 🔼	₫ 배출권등록부시스템
		₫ 상쇄등록부시스템
물환경/수자원 🕒		온실가스검증심사원 문제은행시스템
223,712		하수도종합정보관리시스템
17 L S	하수도 💽	□ 하수도정보 □ 하수도정보
환경부 🚭	상수도 🕒	□ 하 가 보 하고 기 계측시스템 □ 한강수계 하수관망 계측시스템
상하수도 🗗		
자연환경 🕒	토양지하수 🚨	
자원순환 🔒		
	4	
환경보건/화학물질 ① 환경산업/생활환경 ②		

데이터 네비게이션 기능을 활용하여 원하는 분석자료를 얻을 수 있습니다.

총 1건이 검색되었습니다.

엑셀다운로드

순번	정보 대분류	정보 소분류	시스템	데이터셋여부	데이터베이스	테이블
1	상하수도	하수도	한강수계 하수관망 계측시스템	N	-	-



III. 타블로 사용하여 데이터 분석하기(실습 1)

Ⅲ-0. 타블로 온라인으로 데이터 분석 준비



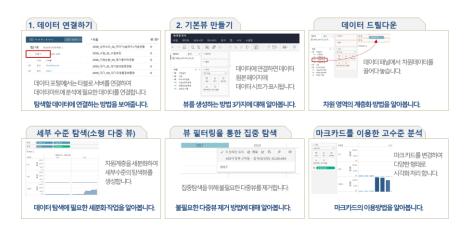
타블로 무료체험판을 얻습니다.



Ⅲ. 타블로 사용하여 데이터 분석하기(실습 1)



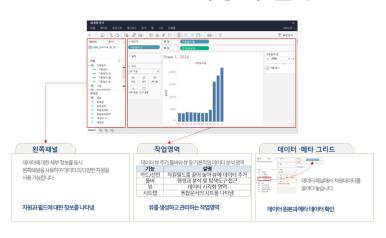
Ⅲ-1. 데이터를 탐색할 기본뷰



Ⅲ-3. 뷰에 고급분석 적용하기



III-2. Tableau 사용자 환경



Ⅲ-4. 선반 및 카드 참조



Ⅲ -1. 데이터를 탐색할 기본뷰



타블로에서 데이터를 분석하는 기본적인 방법을 통해 기본기능을 살펴봅니다.

데이터 연결하기



탐색할 데이터에 연결하는 방법을 보여줍니다.

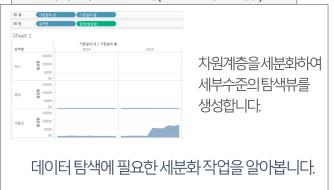
기본뷰 만들기



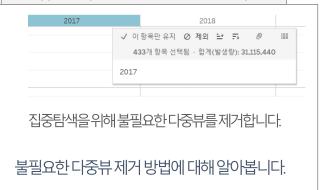
데이터 드릴다운



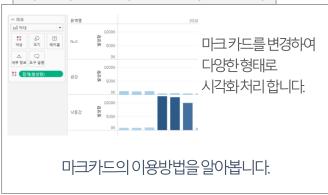
세부 수준 탐색(소형 다중 뷰)



뷰 필터링을 통한 집중 탐색



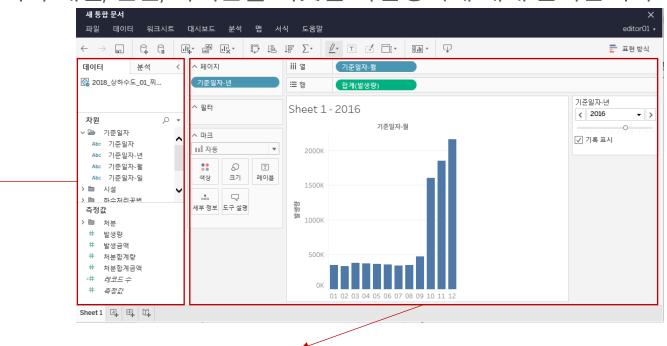
마크카드를 이용한 고수준 분석`



III-2. Tableau 사용자 환경



타블로 페이지와 패널, 선반, 아이콘을 비롯한 작업영역에 대해 알아봅니다.



왼쪽패널

데이터에 대한 세부정보를 표시 왼쪽패널을 사용하여 데이터의 다양한 차원을 사용 가능합니다.

차원과 필드에 대한 정보를 나타냄

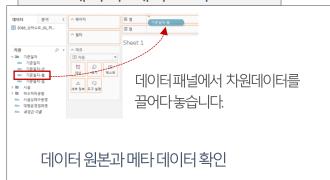
작업영역

데이터 뷰추가 툴바와 뷰 등 기본적인 데이터 분석 영역

기능		설명	
	카드/선반	차원필드를 끌어 놓아 뷰에 데이터 추가	
	툴바	명령과 분석 및 탐색도구 접근	
	뷰	데이터 시각화 영역	
	시트탭	통합문서의 시트를 나타냄	

뷰를 생성하고 관리하는 작업영역

데이터 '메타 그리드

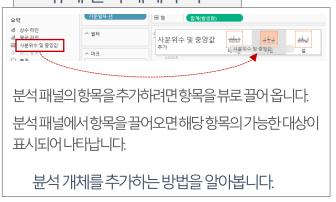


Ⅲ-3. 뷰에 고급분석 적용하기



고급분석을 사용하기 위해 작업영역 왼쪽에 나타나는 분석 패널에서 참조선, 박스 플룻, 추세선 예측 및 기타 항목을 뷰로 끌어올수 있습니다.

뷰에 분석 개체 추가



분석 개체 삭제하기



뷰의 분석개체 편집



분석개체 정의

- 상수라인
- 평균라인

- 사분위수 및 중앙값
- 박스플롯
- 총계

- 95% CI 평균
- 95% CI 중앙값

- 추세선
- 예측

Ⅲ-4. 선반 및 카드 참조



선반과 카드를 활용한 다양한 표현방법을 알아봅니다.

뷰를 시작하는 옵션

Case	설명			
끌어놓기	패널에서 필드를 끌어 놓기(Drag&Drop)			
더블클릭	패널에서 필드 두번클릭			
표현형식	필드선택후 표현방식 사용하기			
여기에 놓기	여기에 필드 놓기 격자에 끌어놓기			

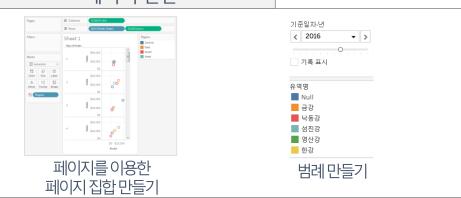
필드를 배치하는 방법에 대해 알아봅니다.



마크카드



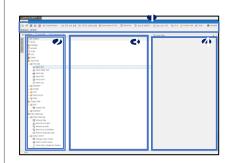
페이지 선반



IV. R-FLOW 사용하여 데이터 분석하기(실습 2)



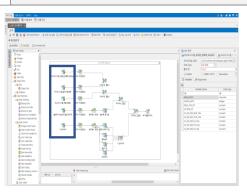
1. 데이터 연결하기



텍스트기반의R분석도구를사용자가 편하게사용할수있도록사용자 친화적인 그래픽 환경으로 제공합니다.

R-How의 분석 화면을 소개합니다.

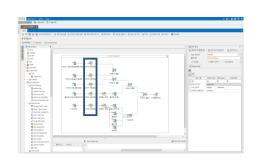
2. 데이터 세트



데이터 분석에 필요한 데이터셋을 불러와작업공간에 배치합니다.

데이터 분석에 필요한 데이터셋을 불러오는 방법을 알아봅니다.

3. 데이터 전처리



데이터전처리에필요한타입변경, 파생변수추가, 데이터셋결합, 라샘플링등다양한방법을 알아봅니다.

본격적인 데이터 분석에 앞서 필요한 전처리 과정을 알아봅니다.

뷰 필터링을 통한 집중 탐색



수학적함수를 사용하여데이터를 분석하고분석한결과를 그래프로 출력합니다.

다양한 수학적 함수 사용법과 그래프 표현법을 알아봅니다.

V. follow Data 사용하여 데이터 분석하기(실습 3)



분석환경

데이터 수집 (엑셀처리)

전처리 (수학 함수) 통계분석 (예측) 통계분석 (예측)

환경 데이터 포털



다차원 배열 등 수학 관련

작업 패키지

NumPy



회귀분석과 시계열분석 등 전통적인 통계분석 패키지



회귀분석과 통계분석 관련 시계열분석 등 다양한 **시각화** 분석데이터

검증데이터



- · 시점: 2018년 4분기 데이터
- · 지역:경남양산북부동 (PM10미세먼지)
- · 데이터명:전국대기오염측정정보
- · 시점 : 2019년 1분기 데이터
- 지역:경남양산북부동 (PM10미세먼지)
- 데이터명:전국대기오염측정정보

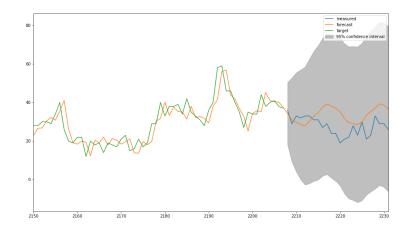
엑셀파일 처리 패키지



Follw Data



예측분석



감사합니다.