현장실습 활동보고

(주)PDXen 현장실습생 김아영

<목 차 >

- ◆ 현장실습 일시&기관
- ◆ 업무 보고
 - 1) Web 개발 폐암 검사일지 관리시스템
 Donor 정보 관리시스템
 - 2) 통계 분석 대장암 데이터 심혈관 데이터 치매 데이터

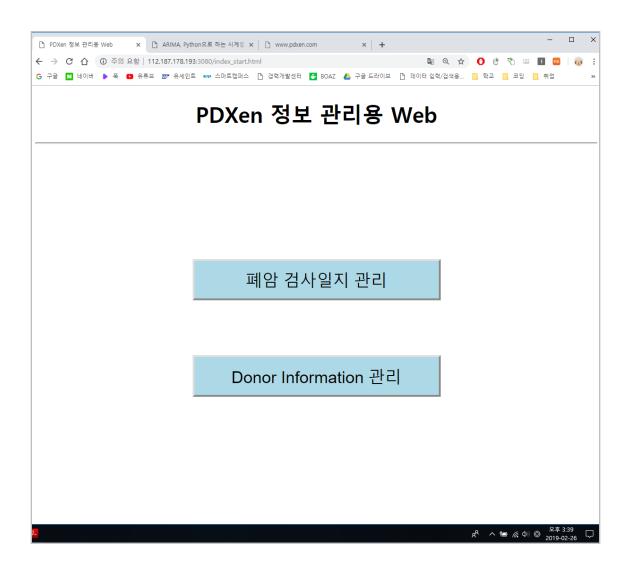
◆ 현장실습 일시&기관

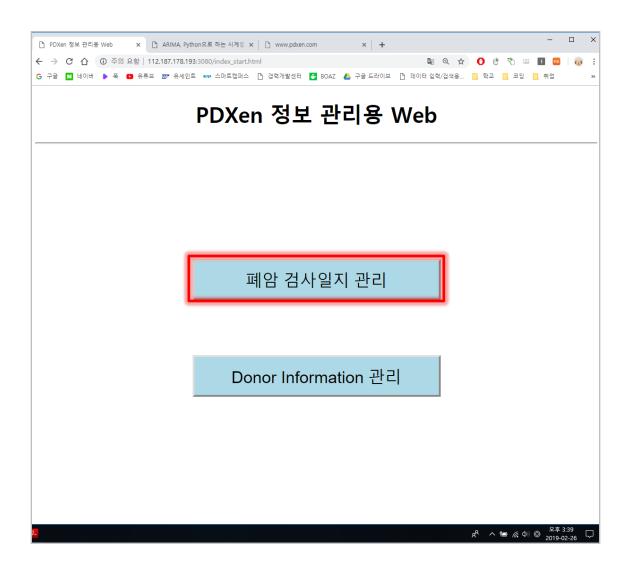
· 일시: 1/2 ~ 2/28 (2달간)

· 기관 : (주)PDXen

경기도 성남시 수정구 대왕판교로 815 판교 제2테크노밸리 기업지원허브 771호

◆ 업무 보고 - 1) Web 개발





〈기존 제안서〉

데이터 생성



데이터 확인/검색



데이터 생성



데이터 확인/검색

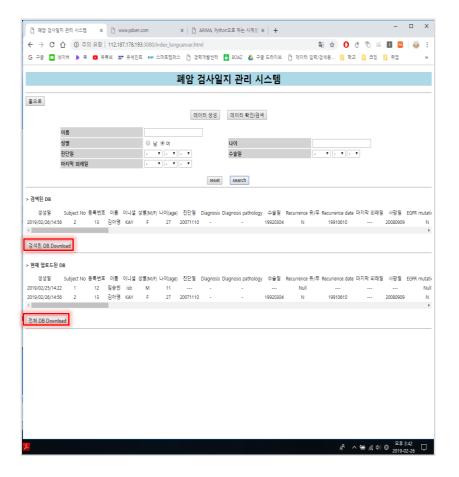
id	Field	Туре	Null	Key	Default	Extra
Initial		int(11)	NO	PRI	NULL	auto_increment
Initial	created	varchar(20)	YES		I NULL	
Initial		varchar(20)	YES			
sex	I name	varchar(10)	YES I			
diagnosis_date		varchar(10)	YES:			
diagnosis_date varchar(20) YES NULL va		varchar(5)	YES			
diagnosis_pathology		L varchar(20)	YES			
Surfery_date		varchar(20)	YES:			
recurrence_date	surgery_date	varchar(20)	YES		NULL	
lastcare_date varchar(20)	recurrence	varchar(5)	YES			
eafr_mutation		varchar(20)	YES I			
rosl_mutation	death_date	varchar(20)	YES			
alk_mutation	egfr_mutation ros1 mutation		AE2			
	ll alk mutation	varchar(5)	YES		NULL	
pathology_differentiation	k_ras_mutation	varchar(5)	YES			
pathology_differentiation		varchar(50)	YES			
pathology_differentiation	tumor_stage_8th	varchar(20)	YES		NULL	
pathology_differentiation	pathology_location	varchar(20)	YES			
pathology_differentiation	pathology_size_mm pathology_visceral_pleural	varchar(20)	YES		NULL	
pathology_differentiation	pathology_resection_margion_fullterm	varchar(20)	YES			
tobacco_smoking_history_indicator	pathology_lymph_node	varchar(20)	YES:			
reg men	pathology_differentiation	varchar(20)	YES		NULL	
reg men	tobacco_smoking_history_indicator	varchar(5)	YES			
reg men	year_ot_tobacco_smoking_onset year_of_tobacco_smoking_cessation	L varchar(20)	YES			
last_date	regimen	varchar(20)	YES I		NULL	
cycle_no	start_date	varchar(20)	YES			
recurrence_notate_ recurrence_date_ recurrenc	rast_date	varchar(5)	YES			
rt	recurrence_progressive_disease	L varchar(5)	YES			
rt	recurrence_date_ regimen	varchar(20) varchar(20)	VES I		I NULL I NULL	
Cycle_no_ wbc warchar(s) YES NULL warchar(10) YES NULL wa	rt	varchar(5)	YES		NULL	
wbc		varchar(20)				
ha	cycle_no_ wbc	varchar(10)	YES I			
Jymphocyte yarchar(10) YES NJLL yarcha	hg	varchar(10)	YES			
Jymphocyte yarchar(10) YES NJLL yarcha		varchar(IU)	AE2			
Jymphocyte yarchar(10) YES NJLL yarcha	seg_neutrophil	varchar(10)	YES I		NULL	i
sodium potassium chloride comolality_serum calcium phosphorus bun creatinine glucose g		varchar(10)	YES I			
potassium chloride chloride comolality_serum comolality_s	l sodium	varchar(10)	YES			
osmolality_serum calcium phosphorus bun creatinine glucose uric_acid cholesterol protein_total albumin bilirubin_total alkaline_phosphatase ast alt egfr_mdrd egfr_ckd_epi Idh	potassium	l varchar(10)	IYES I			
calcium phosphorus bun creatinine glucose uric_acid cholesterol protein_total al laurinine_altase altal ine_phosphatase altal ine_phosphatase altal ine_acid cefr_mdrd eefr_mdrd eefr_mdrd eefr_mdrd eefr_mdrd eefr_dd_epi ldh varchar(10) varchar(10			YES I			
bun varchar(10) VES NULL varch		varchar(10)	YES İ			i i
creatinine glucose uric_acid uric_acid varchar(10) YES varchar(10) VES varchar			YES I			
alucose varchar(10) VES NULL v			YES			
cholesterol roterin total albumin total albu	glucose	l varchar(10)	YES İ		İ NULL	
al la line_phosphatase alt leefr_mdrd eefr_ckd_epi ldh ldh ldh ldh ldh leefr_mdrd eefr_ckd_epi ldh ldh ldh ldh ldh ldh ldh ldh ldh ldh	uric_acid cholesterol	varchar(10) varchar(10)	YES I			
Database: pdxendb Database: pdxendb Table: mytable	protein_total	varchar(10)	YES			
Database: pdxendb ast alt esfr_ndrd esfr_ckd_epi ldh Database: pdxendb varchar(b)	albumin	varchar(10)	YES I		NULL	
eafr_mdrd eafr_ckd_epi	alkaline_phosphatase	Mala	LYES _		وام ما ٠	ما ام مر م
eafr_mdrd eafr_ckd_epi	l ast	Datar	120	Φ.		X E N A N
eafr_ckd_epi varchaTable :\mvtable			حيرح		Pu	CITAD
varchar able NMVTable	egfr_ckd_epi	varcha=(10)	YES		I NULL	
70 rows in set (0.00 sec)	Idh	varchar 1🤦	nIΔ		' my	Tahla
	70 rows in set (0.00 sec)	Ia	\mathcal{L}		· 111 y	LUDIC

데이터 확인/검색

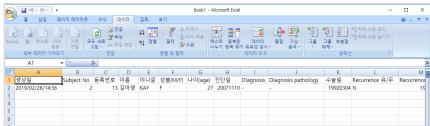




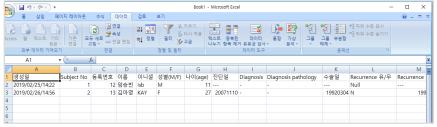
csv 파일 확인

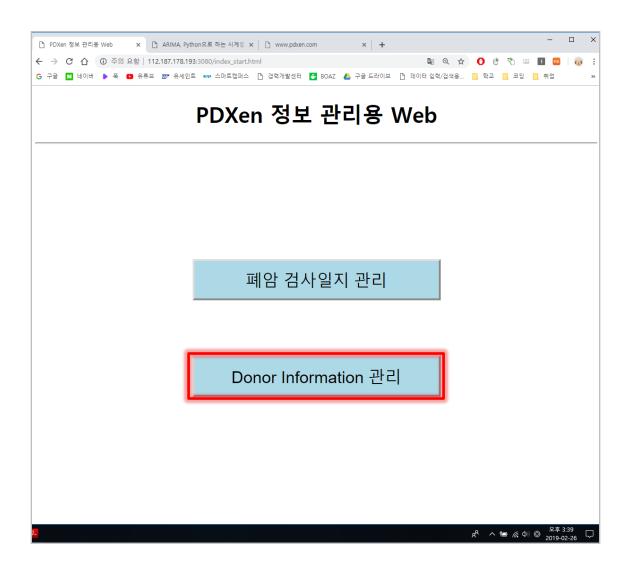


-> my_searched_export.csv

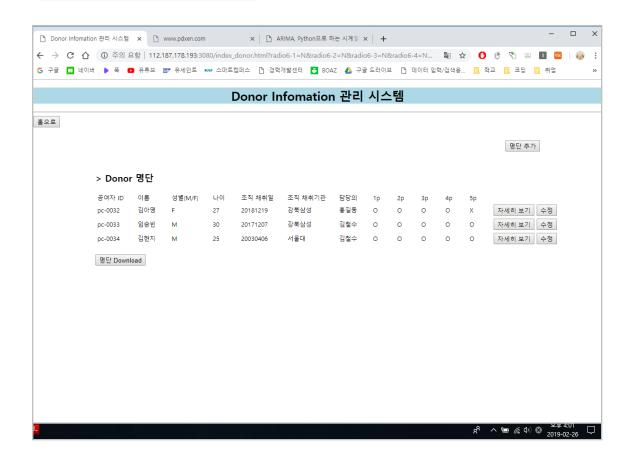


-> my_export.csv

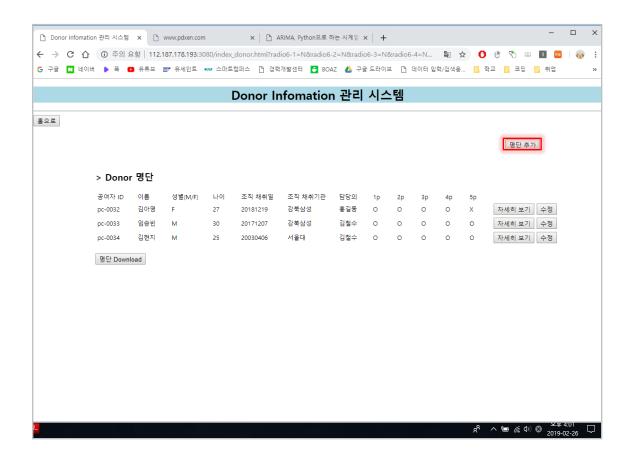




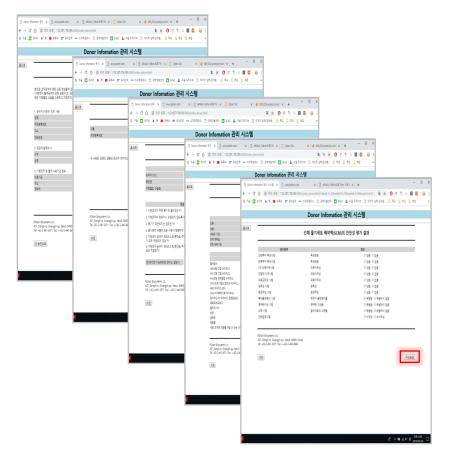
명단 확인



명단 확인

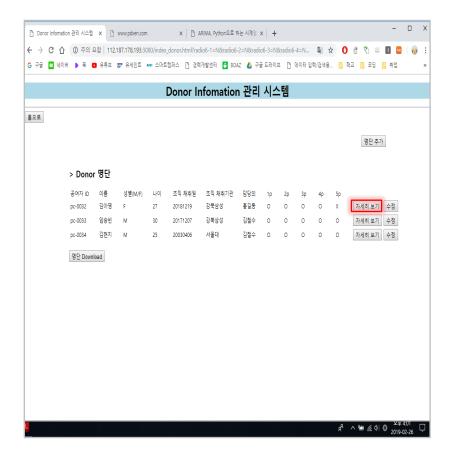


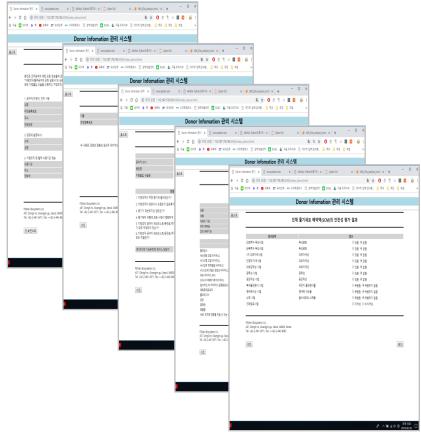
명단 추가



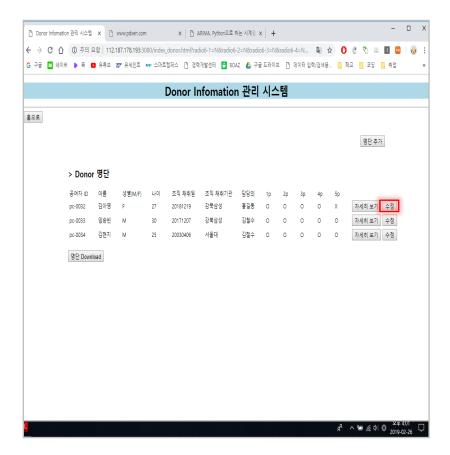
Field	Туре	Null	Key	Default	Extra
created identification name race sex age resident_no address phone cell_origin doctor_department doctor_department doctor_name surgery_date amount charge_pdxen charge_pdxen charge_pdxen charge_tdate exect_l_note check_1 note check_2 note check_2 note check_3 note check_3 check_3 check_3 note check_5 check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 note check_5 check_5 check_5 check_5 check_5 check_5 check_5 check_5 check_5 check_5 check_1 blood_test_1 blood_test_2 blood_test_3 blood_test_1 blood_test_1 blood_test_1 blood_test_1 safety_test_1 safety_test_2 safety_test_5 safety_test_8 safety_test_8 safety_test_8 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1 safety_test_1	varchar(20) varchar(20) varchar(20) varchar(5) varchar(20) varchar(20) varchar(20) varchar(20) varchar(10) varchar(20) v		PRI		
safety_test_11 page1 page2 page3 page4	varchar(5)			dxen	
page5 page5 	Varchar Ta	ble	: n	nytab	le

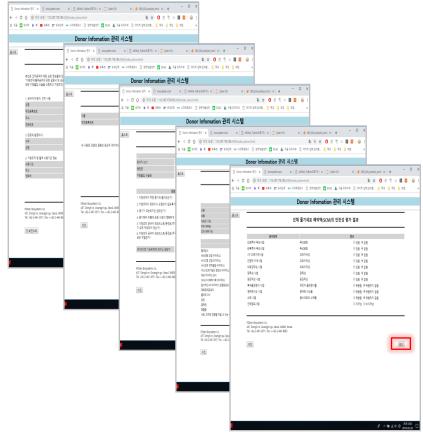
자세히 보기



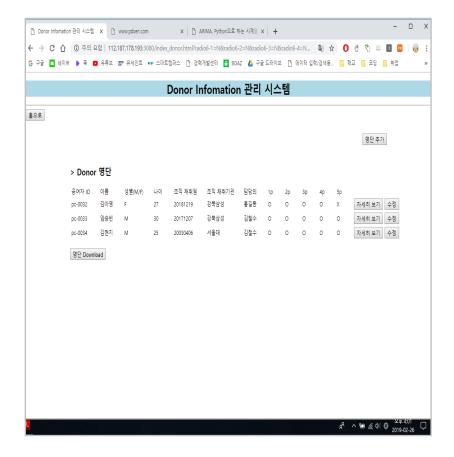


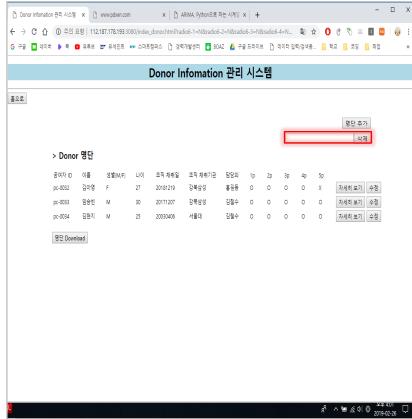
수정



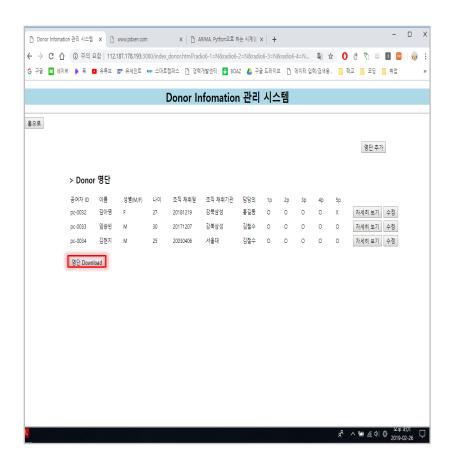


삭제

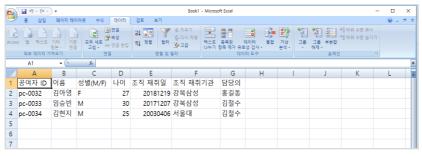




csv 파일 확인



-> my_export.csv



◆ Web 개발 - 개발 언어/툴

ㆍ개발 언어

ㆍ개발 툴

[프론트엔드] - HTML5 CSS JavaScript





[백엔드] - Node.js MySQL

ATOM Editor Chrome Browser

◆ 업무 보고 - 2) 통계 분석

features	속성	전체(n=1318)	Na	0(n=558)	Na	1(n=760)	Na	정규성		분산성	분석방법	p-value	ROC Curve	AUC
Age	연속형	62.54 ± 12.28		58.31 ± 13.10		65.64 ± 10.62		Name of residency and the part of the part	х	х	Mann-Whitney U test	0.0000	Aug. 1860 10 10 10 10 10 10 10 10 10 1	0.680
Sex	범주형	0 - 776(58.9%)		0 - 282(50.5%)		0 - 494(65.0%)					Chi-squared	0.0000		
		1 - 542		1 - 276		1 - 266					test			I

1. 데이터 및 결측치 파악

2. 자료형 & 분포 확인

· 범주형: 가장 빈도가 높은 레벨(%)

ㆍ연속형:평균 ± 표준편차

3. 그룹별 분석

· 범주형: Chi-squared test

· 연속형: 정규성, 등분산성 확인 후

(히스토그램, QQ plot 이용)

① Independent T test

② Welchs T test

3 Mann-Whitney U test

4. ROC curve 추가(연속형 변수)

◆ 통계 분석 - 대장암 데이터

총 데이터 : 70336 개

정상 (acrn_yn=0) : 69376 개

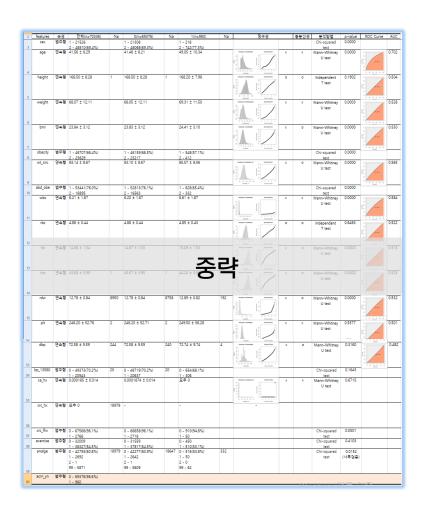
환자 (acrn_yn=1) : 960 개

총 변수 : 59 가지(acrn_yn 제외)

범주형 : 22 가지

연속형 : 36 가지

의미 없는 변수 : 1 가지



◆ 통계 분석 - 심혈관 데이터

총 데이터 : 1318 개

정상 (DX=0) : 558 개

환자 (DX=1) : 760 개

총 변수 : 61 가지 (DX 제외)

범주형 : 7 가지

연속형 : 46 가지

의미 없는 변수 : 8 가지



◆ 통계 분석 - 치매 데이터

총 데이터 : 7844 개

정상 (DX=0, 1) : 7043 개

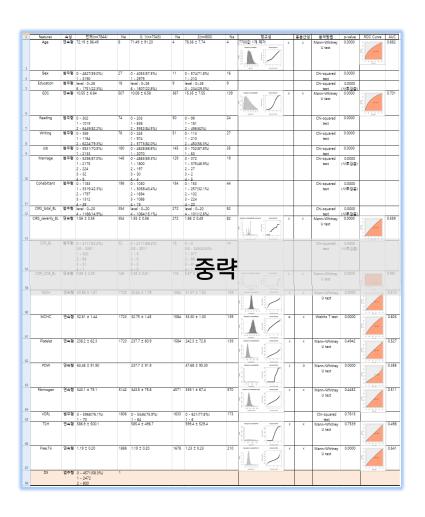
환자 (DX=2) : 800 개

총 변수 : 66 가지 (DX 제외)

범주형 : 21 가지

연속형 : 44 가지

의미 없는 변수 : 1 가지



◆ 통계 분석 - 분석 언어/툴

・분석 언어

・분석 툴

R



R studio

감사합니다:)