

**[양식2] 근태기록부**

## 2019 공공 빅데이터 청년 인턴십

일일근태기록부

(2019년 11월 )

☐ 참여기관: 화성시    ☐ 수련생: 김동혁    ☐ 기관담당자: 이 지 호

일자	요일	성명	출근시간	퇴근시간	근태사항					확인 (수련생)	담당자 확인
					결근	지각	조퇴	외근	휴가		
1	금	김동혁	9:00	18:00							임지훈
2	토	—	—	—							
3	일	—	—	—							
4	월	김동혁	9:00	18:00							임지훈
5	화	김동혁	9:00	18:00							임지훈
6	수	김동혁	9:00	18:00							임지훈
7	목	김동혁	9:00	18:00							임지훈
8	금	김동혁	9:00	18:00							임지훈
9	토	—	—	—							
10	일	—	—	—							
11	월	김동혁	9:00	18:00							임지훈
12	화	김동혁	9:00	18:00							임지훈
13	수	김동혁	9:00	18:00							임지훈
14	목	김동혁	10:00 (수능)	18:00							임지훈
15	금	김동혁							0	—	—

일자	요일	성명	출근시간	퇴근시간	근태사항					확인 (수련생)	담당자 확인
					결근	지각	조퇴	외근	휴가		
16	토	—	—	—							
17	일	—	—	—							
18	월	김동혁	9:00	18:00						gmm	임지호
19	화	김동혁	9:00	18:00						gmm	임지호
20	수	김동혁	9:00	18:00						gmm	임지호
21	목	김동혁	9:00	18:00						gmm	임지호
22	금	김동혁	9:00	18:00						gmm	임지호
23	토	—	—	—							
24	일	—	—	—							
25	월	김동혁	9:00	18:00						gmm	임지호
26	화	김동혁	9:00	18:00						gmm	임지호
27	수	김동혁	9:00	18:00						gmm	임지호
28	목	김동혁	9:00	18:00						gmm	임지호
29	금	김동혁	9:00	18:00						gmm	임지호
30	토	—	—	—							
31											
합계									1	gmm	임지호

2019 공공 빅데이터 청년 인턴십  
인턴십 프로젝트 주간 활동보고서  
(2019년 11월)

☐ 참여기관: 화성시    ☐ 수련생 성명: 김동혁

활동 일시	월-주차	기간	참여기관 담당자 확인
	11월 1주차	2019.11.1.(금)	
활동 내용	1일차 (11.1)	금일에는 CCTV 초상보고서를 80% 작성하였다. 초상조사순위를 코드북과 귀성사인으로 검색한 뒤 정리하면 될것이다. 이후에는 통계조사상행위 에가서 필요한 아과와 사무용품, 조사용품등을 전달해주었다. 복거배 ADSP를 준비하였고 피해한 정보를 안내받을 통해 수집하였다. 각 CCTV분석이 끝나면 데이터베이스 관리 및 SQL 언어를 공부하여 DB 관리연습을 하려고 한다.	이지훈

## 2019 공공 빅데이터 청년 인턴십

### 인턴십 프로젝트 주간 활동보고서

### (2019년 11월)

☐ 참여기관: 화성시    ☐ 수련생 성명: 김동혁

활동 일시	월-주차	기간	참여기관 담당자 확인
	11월 2주차	2019.11.4.(월)~2019.11.8.(금)	
활동 내용	1일차 (11.4)	음일은 데이터 분석준비가 공부를 하였다. 통계분석파트에서 기본통계 개념부터 측정방법, 분포 추정 및 가설검정, 회귀분석, 주성분 분석, 시계열 데이터 다루는 방법에 대해 학습하였다. 시계열 데이터를 한번도 다루지 않아서 개념확립이 어려웠지만 응용문제를 풀어봄으로써 어느 정도 내용을 머릿속에 확립하였다. 데이터분석사직전에 있어서 차분차소, 변수선택법이 매우 중요한 요소인데, 여러가지 방법이 있다는것을 알았다.	김우영
	2일차 (11.5)	음일에는 각 지역 통계상황실에 필요 물품 전달 및 조사가 시작되어 여러 가지가 상황을 관리하기 위해서 줄장을 다녀왔다. 또한 남부에서 취지한 여러번에 걸쳐 작성된 통계제자를 전달하였다. 또한 복귀하여 APSP 학습을 하였다. 시계열 데이터중, 정상성과 비정상성 시계열의 차이점과 정규 분산에 따른 차분, 변환방법에 대해 공부하였다. 유동과 같이 시계열성을 띄는 데이터에서 활용하여 인위성을 해결도 줄어갈다.	임지훈
	3일차 (11.6)	음일은 정형데이터마이닝에 대해 공부하였다. 지도학습과 비지도학습의 분석 방법종류와 각 분석의 개념을 학습하였다. CCTV 분석을 했을때 설명변수들을 활용하여 군집분석을 진행해왔는데 어떤 방식으로 분석이 진행되는지, 변수의 특성별로 사용해야 하는 방법이 차이가 있다는 것을 알았다. 이후 연관규칙과 측정지표 (치지도, 신뢰도, 향상도 등) 계산 방법을 공부하여 해당 선출문제를 풀어 정리를 하였다.	임지훈
	4일차 (11.7)	음일은 비정형 데이터 마이닝에 대해 공부하였다. 기존에 숙지되어있던 데이터들은 대부분 정형 데이터이며 시간이 지날수록 비정형 데이터의 양이 방대해지고, 처리와 군집이 매우 복잡하여 쉽게 접할 수 <del>없었다</del> 없었다. 텍스트 마이닝을 알아보는데, 빈도분석, 도메인, 군집분석과 비슷한 군집방법들을 학습했다 이전에 분자량이 분명한 화합물 분해분석데이터로써 빈도분석정도를 해보며 기초를 다졌다.	임지훈
	5일차 (11.8)	음일도 비정형데이터 마이닝을 학습하였다. 군집분석을 좀더 구체적으로 공부하였고, 어떤 거리적도를 사용하는지 여러가지 군집알고리즘 중 무엇을 사용할 것인지 최적의 군집수란 뭘까가 적당할 것인지에 대해 알았다. 통계학전공을 하면서 분산과 프로그래밍 수업을 주로 들었고 선형대수학, 행렬을 접한 경험이 드문데 기본적으 데이터를 볼 때 두가지를 사용하여 변수가 어떤 관계를 가지는지 알아야 할 것이라고 생각했다.	임지훈

# 2019 공공 빅데이터 청년 인턴십

## 인턴십 프로젝트 주간 활동보고서

### (2019년 11월)

□ 참여기관: 화성시    □ 수련생 성명: 김동혁

활동 일시	월-주차	기간	참여기관 담당자 확인
	11월 3주차	2019.11.11.(월)~2019.11.15(금)	
활동 내용	1일차 (11.11)	금일 오전에는 제헌절에 이어서 선형대수학과 행렬에 대해 공부하였다. 덧셈과 곱셈의 연산후에 나타나는 변화라 군조에 초점을 두는 것인데, 주어진 데이터(배치)를 선형사상(일차원)에서 표현하는 방식이라고 한다. 흔히 고딩에서 말하는 선형과 유사한데, 가령 평면을 지나는 직선에 대한 대칭변환(rotation), 수직선상에 가해지는 이동을 나타내고 있다. 이후에는 통계상황설로 줄장을 다녔고, 내일 이어서 확률론 예정이다.	임지호
	2일차 (11.12)	금일에도 선형대수를 이어서 학습했다. 3차원에서의 기저벡터를 표현하면서 차원축소를 위해 정규화하는 과정을 학습했다. 기저벡터끼리 내적해서 크기가 0이 되는 것이 바로 좌표계인데, 여태 GIS를 사용하면서 공부했던 좌표계와는 다른 개념이었다. GIS에서 사용하는 polynomial 도 폴리까지를 만족한다면 벡터라고 한다. 두개의 함수를 사용해도 표현이 가능하다. 논점이 신기했다. 이후에는 분기값이 연속인 데이터와 분기데이터를 공부하였다.	임지호
	3일차 (11.13)	금일 오전에는 통계변수 작성을 도왔다. 사면체 지표를 지명별, 연평별, 성별 등으로 구분되어있는 자료를 피벗테이블을 활용하여 기존데이터에 맞춰 줄장을 하였다. 이후에는 전날 분기값이 정해진 데이터를 활용하여 분기보고서를 수정하였다. 그중, 행정등급데이터 및 토지이용데이터 같은 GIS에서 새로 받은 데이터가 있었고, 이를 어떻게 사용해야 할지 EPA 작업의 필요성을 느꼈다.	임지호
	4일차 (11.14)	금일 오전에는 분기보고서를 수정하기 앞서 새로받은 데이터를 병합하여 EPA를 하는 과정을 가졌다. 각 지자체별 5대 병합이 병합된 데이터를 병합하였다. 이어 토지이용목적에 따른 지자체별 명제를 다시 병합하고, 이전에 가지고 있지 않았던 상가건물 중, 병원·유치원·공원·은행·ATM·학교 등 모두 좌표에 적어보며 지자체별 개수를 병합하였다. 어떤 지리적, 행정적인 요소가 병합 방식에 영향을 미치는지 알아본 뒤, 정리된 자료를 가칭으로 부여하였다.	임지호
	5일차 (11.15)	휴 가	임지호



# 2019 공공 빅데이터 청년 인턴십 인턴십 프로젝트 주간 활동보고서 (2019년 11월)

□ 참여기관: 화성시 □ 수련생 성명: 김동혁

활동 일시	월-주차	기간	참여기관 담당자 확인
	11월 4주차	2019.11.18(월)~2019.11.22.(금)	
활동 내용	1일차 (11.18)	금일 오전에는 지난주 분석했던 자료를 다시 정리하였고, 기존에 작성했던 보고서에는 다르게 분석방법이 정해져있는 양식과 기초데이터를 다시 좌표에 찍어 그림을 교체해야 하는 작업이 꽤 오래걸렸다. 특히 중간에 범죄등급을 분석하는데 격자의 어떤 특성도 범죄발생등급과는 상관성이 적어서, 추가로 할 분석작업은 없을 것 같다. 또한 불포준분석모델과는 다르게 범죄데이터를 사용함으로써 근거리에서 정보를 좀더 찾아주어야 할 것 같다.	임지훈
	2일차 (11.19)	금일은 오전에 모두 출장을 다녀왔다. 오전에는 대리와 사무용품들을 구입하여, 각 통계 상황실에 들렀다. 오후에는 목요일에 진행하는 용문회 대상 교육을 하는데 필요한 용품을 구입하였다. 필요 개수가 있어서 서지 없이 소외되었다. 사무실로다 보러 작성을 마쳐버렸다. 현재 도매로 주매의 설치업체를 방문할 방법이 떠오르지가 않아서 찾는 중이다. 각 격자별 도매 구매가 아니라 하이퍼 리스크를 정리할 수 있을 것 같다.	임지훈
	3일차 (11.20)	금일은 정과표와 위험도를 주소와 좌표로 변환할 수 있는 지오코딩 프로그램을 찾아보았고, 굉장히 복잡한 절차로 이루어져있어 학습하는데 시간이 걸렸다. GIS 상에서 점좌표를 격자의 중심점으로 찍었고 그림을 GIS가 자동으로 인식하여 위험도 좌표를 찍어주는 프로그램을 사용하였다. 이후 해당 분석프로그램을 사용하여 위험도를 주소로 변경하여 리스크를 만드는 데 큰 도움이 되었다. 내일날로 보고서가 완성될 것 같다.	임지훈
	4일차 (11.21)	금일 오전에는 하이퍼 리스크를 선출한 뒤, 상의 밑에서 코드를 다 시도 이이작을 정리하여 보고서를 최종작성하였다. 분석방법의 피드백을 받았고, 수정할 부분은 없지만 조금 보충설명이 필요한 부분을 물어주었다. 오후에는 빅데이터 교육준비를 끝내고 들리었다. 국가직업안전원에서 경찰청이 직접 교육을 하였고, 빅데이터에 대한 호기심을 불러일깨우는 유익한 듣기였다.	김동혁
	5일차 (11.22)	금일 오전에는 어제 링장님이 요청받은 화성시 공무사업 시작화를 드렸다. 이전에 자료를 보낸 이후에 공무사업 건수의 개수에 변동이 있었고, 그날도 자료준비 분석명이 바뀌어서 편법이 필요했다. 추가로 3번 부서별 공무사업현황을 금일 주기로 원하시서 선택사항을 제출하였다. 오후에 4시부터는 내일준비회의가 진행되어서 공무를 하였다. 통계분석부분은 괜찮지만 분석기력과 거버넌스 PART는 아직 보충이 필요하다.	임지훈

# 2019 공공 빅데이터 청년 인턴십

## 인턴십 프로젝트 주간 활동보고서

### (2019년 11월)

□ 참여기관: 화성시 □ 수련생 성명: 김동혁

활동 일시	월-주차	기간	참여기관 담당자 확인
	11월 5주차	2019.11.25.(월)~2019.11.29.(금)	
활동 내용	1일차 (11.25)	금일에는 분석보서를 최종적으로 피드백을 받았다. 분석가님이 직접 코딩하여 추출한 화성시 top 100개소와 내가 뽑은 top 100개소의 위치가 대부분이 있어서 데이터를 통합하고, 보서를 위한 수정해야 하는 부분이 생겼다. 휴먼에올릴 보고서로는 경찰청에서 제공하는 범죄비율 데이터를 아스킹 수정을 해야한다. 가한은 28일 목요일까지로 정했으며 분석가님에게 사용한 데이터를 받음과 동시에 GIS, 보서를 마무리하면 될 것 같다.	임지호
	2일차 (11.26)	금일은 휴먼에올릴 보고서 수정을 하였다. 결과보고용과 휴먼용을 따로 작성하였으며, GIS상에 보이는 범죄비율을 아스킹하는 작업이었다. 어느정도까지 유틸이 생기게 되는지 몰라서 제일 큰 변수인 동부서 서부에서의 범죄발생 건수 정도도 아스킹 하였다. 이후 답변을 검토가 들어가지 않지만 범죄가 0~10등분으로 일어나 발생했는지 빈도 히스토그램 정도도 만들어야겠다.	임지호
	3일차 (11.27)	금일은 K-ICT 데이터 그랜드 컨퍼런스에 다녀왔다. 담당자들과 함께 참석하였으며, 여러기업에서 데이터 분석업무를 설명해주는 연사들이 있었고 해당기업에서 실시한 프로젝트에 대해 강의하였다. 하지만 보낼바라지 개인과성, 적합성, 정확성등의 기준을 설명해주었이 크게 와닿았다. 주제는 지수(선자수, 모바일수등)를에 파악이 않는 의미를 두는 것 같았다. 데이터를 많이 분석하는일에 큰 자부심을 느껴야 판교 생겼다.	임지호
	4일차 (11.28)	금일도 보고서 작성(수정)을 계속했다. 보고서 상에서 데이터분석 시각화 부분은 보는데이 이해할 수 있는 것이라는 판단하에 삭제하고, 쉽게 알아볼 수 있는 지도 데이터로 수정하였다. 제1계급지명부터 제3계급지명등 나눠있는 도시 용도를 사용하여 계급지명 명칭변 C4V까지 개소등 기존에 설치되어있는 C4V를 용도 세부적으로 분석하여 다양성을 높였다.	임지호
	5일차 (11.29)	금일 보서를확 다작성하였다. 결과발급대로 휴먼 시스템 표준분석모델에 올라갈 것 이라고 하였다. 실시간으로 바뀌는 데이터와 새로 제공받은 데이터 때문에 계속 해서 바뀌었지만, 오늘은 끝으로 최종마무리 하였다. 이후 운영계획서 상으로 인턴십, 혹은 태블로 시각화가 있는데 이를 과제로 받기전에 미리 학습하는 것이 좋을 것 같다. Java와 C 쪽을 학습해볼 계획이다.	임지호