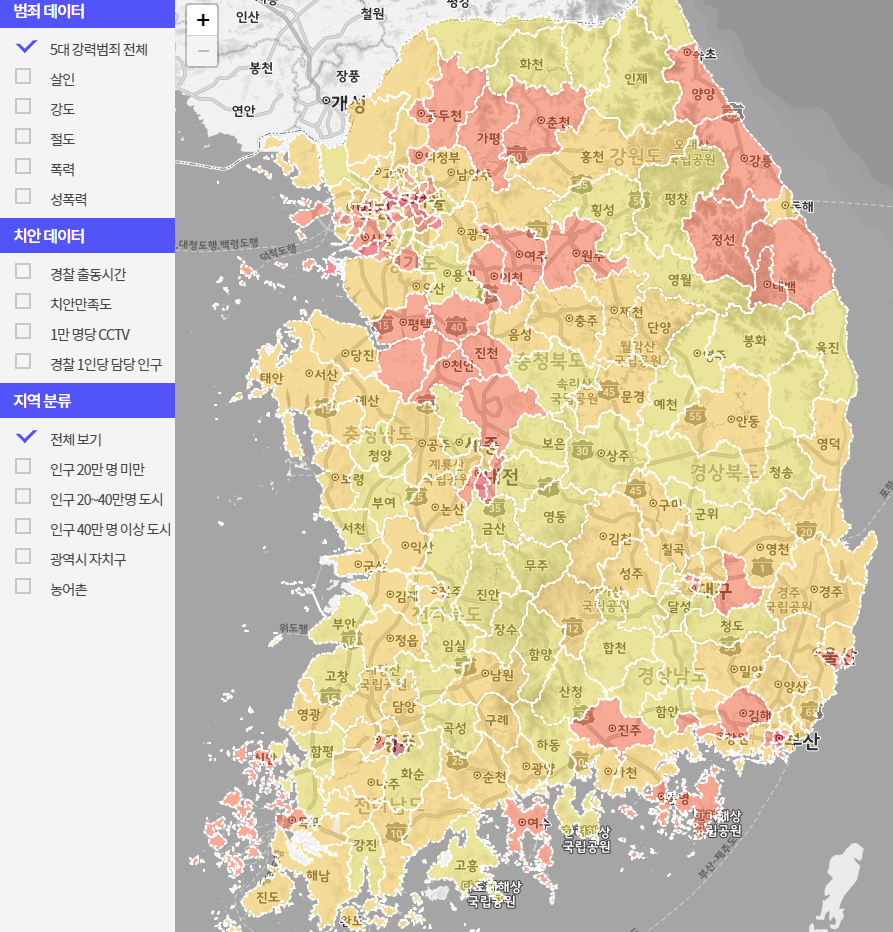
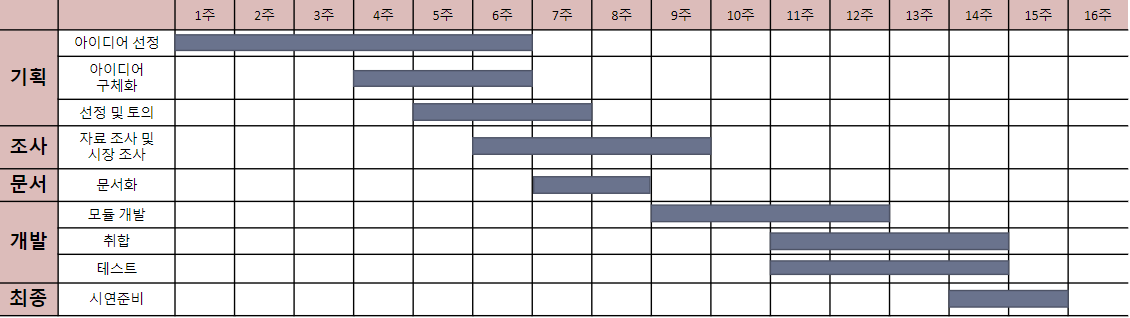
1. **위기탈출 헬조선**
   1. **주제 :**공공데이터 통계 분석을 통하여 내 주변의 위험도 확률 분석 및 지역 루머 비교 검증
   2. **개요 및 개발 동기**옛 말 중에 먼저 가는데 순서 없다는 말이 있습니다. 우리는 언제 어디서 어떤 사건 사고가 발생할지 알 수 없습니다. 어제까지만 해도 건강하기만 하던 사람들이 갑작스런 사고를 당해 다치기도, 불행하게는 생을 마감하기도 합니다. 이러한 불확실한 미래는 우리들을 불안에 떨게 하며, 각종 소문이나 미신들을 믿게 만듭니다.  이렇듯 미래의 일은 단 0.1초 이후도 먼저 알 수 없습니다. 하지만 역사를 통해 우리가 미래를 대비 하듯이, 모든 일은 과거의 일들을 통해서 어느정도 예측이 가능합니다.   
       그래서 생각하였습니다. 각종 지역들에서 일어난 과거 사건 사고들에 대한 공공 데이터들과, 자연재해, 날씨 등 사건사고에 관련될 만한 모든 데이터들을 수집해 통계를 낸다면, 당장 오늘이나 내일에 조금이라도 안전하게 다닐수 있지 않을까? 이렇게 시작한 아이디어는, 나의 알 수 없는 미래를 조금이나마 확률을 통해 알려준다는 재미를 선사함과 동시에 전체 지역적인 통계를 다시한번 가공하여서 떠돌고있는 지역적 루머들이 (ex. ‘마계인천’, ‘대구 = 고담시티’ 등) 실제로 있음직한 말이라는 것을 검증해내는 아이디어로 나아가게 되었습니다.
   3. **목표**  
       어플리케이션의 푸시알람을 통해 오늘의 운세처럼 매일 오늘은 내가 얼마나 안전한가에 대한 정보를 알림받습니다. 어플리케이션을 통해 오늘 내가 사고가 날 확률, 좋지 않은 일을 당할 확률 등 여러가지 확률을 확인 할 수 있을 뿐만 아니라, 지역별 위험도 더보기 항목을 통하여 지역별 위험도가 어느정도인지 시각적으로 바로 확인 할 수 있습니다. 각종 사건사고들이 발생한 데이터들을 종합하여 통계를 내 신뢰도 있는 지역별 위험도를 제공하고, 그 통계를 바탕으로 낸 각종 사건들의 확률을 통하여 사용자들에게 재미를 선사하는 것이 최종 목표입니다.
   4. **사전조사**  
      4-1.시장 조사  
       전국의 범죄와 치안에 관련된 시장 조사를 하던 결과, 프로젝트와 비슷한, 흥미로운 사이트 하나를 발견하게 되었습니다.  SBS데이터저널리즘팀의 <마부작침>에서 만든 범죄여지도라는 사이트 입니다. 사이트를 들어가게되면, 지도를 볼 수 있으며, 각각의 지역마다의 5대강력범죄 데이터가 얼마나 있는지 확인이 가능합니다. 또한 전국의 범죄 데이터들을 모두 통계를 내어, 어느지역이 가장 위험한지 나타내어 줍니다. 아래는 사이트의 사진입니다

****  
  
(그림 1) SBS데이터저널리즘팀 <마부작침>의 2016 범죄여지도  
  
  
**** (그림 2) SBS데이터저널리즘팀 <마부작침>의 지역별 범죄 순위

4-2. 차별점  
 이 사이트와 현재 이 프로젝트와의 차별점은 다음과 같습니다.  
 1. 범죄 데이터만이 아닌 다양한 데이터(교통사고, 화제 등)의 사용으로 더 정확한 지역별 위험도의 통계 산출 가능.  
 2.  지역 지도를 보여주는것 만이 아닌, 각 지역별 사건 사고 발생 데이터들을 종합하여 통계를 내어 각종 확률들을 제공해주어 흥미 유발.  
 3. 가고자하는 지역을 입력할 경우, 그 지역의 위험도와 각종 사건 사고발생 확률을 제공하여 흥미유발.

* 1. **사용 시나리오**  
       
       
       
       
       
     1) 공공데이터 포털에서 데이터들을 가져옵니다.  
     2) 모아온 데이터들을 바탕으로 가공 및 통계분석에 사용합니다.  
     3) 사용자 어플리케이션에 가공된 통계데이터들을 나타내 줍니다.
  2. **개발 진행 방향**1) 공공데이터 포털을 이용하여 데이터들을 수집- 사고발생 현황(연간) 데이터셋  
      - 교통사고정보 (무단횡단사고 다발지, 자전거사고 다발지, 교통사망사고 정보) 데이터셋  
      - 각종 범죄 통계 (범죄 발생시간 및 요일, 범죄발생장소, 범죄발생지역) 데이터셋  
      - 종관 기상 관측 데이터셋  
      - 지역별 연간 화재 발생 데이터셋  
       
     2) 필요 데이터 확보 전략   
      사용할 데이터는 모두 공공데이터 포탈에 저장되어있습니다. 이 데이터들의 다운로드 URL을 미리 저장을 해놓고, 데이터셋이 업데이트 될 때마다 (연간, 월간, 주간데이터들을 최신화 하기 위함.) 재다운로드 받아 통계 분석을 위한 최신데이터셋을 유지합니다.  
        
     3) 데이터 가공 및 학습   
      - 수집한 데이터를 AWS RDS를 사용하여 저장 및 분석  
      - 범죄 데이터중 절도, 추행 등 가장 많이 발생한 범죄종류에 대해서 높은 가중치 부여  
      - 범죄 데이터중 살인, 교통사고사망 등 적게 발생한 사건 사고 종류에 대해서 낮은 가중치 부여  
      - 각각 가중치가 적용된 데이터들을 AWS 기계학습 서비스를 이용해 학습  
      - 데이터를 처리해 나온 지역별 위험한 정도를 전체적으로 통계를 냄  
      - 전체적인 통계를 바탕으로 지역별 루머에 대한 실제 검증  
     4)  예상 출력 데이터셋  
      - 오늘 생존률  
      - 교통사고 확률
  3. **사용 효과 및 기대**당장 어디론가 갔을 때, 사건 사고가 발생할 확률을 제공하는 전례없는 서비스를 제공함으로써 사용자에게 재미를 제공합니다. 단순히 무작위의 확률을 제공하는게 아닌 실제 신뢰할 수 있는 공공데이터를 가지고 통계를 내 상당히 믿을만 하다고 느낄 수 있습니다.  
       뿐만아니라 지역별 통계를 다시 전체적으로 데이터를 가공하면, 어느지역이 사건사고가 발생할 확률이 얼마나 더 높은지 알 수 있습니다. 그를 통해 어떤 지역이 상대적으로 더 안전한지 위험한 지 알 수 있으며, 더 나아가서는 소문으로 떠돌아다니는 지역별 루머들을 검증해 볼 수 있습니다.  
       데이터들을 가공해서 나온 확률들이 어느정도의 신뢰도가 나온나면, 보험사와의 협업을 생각 해 볼 수도 있겠습니다. 어디에서 어떤일이 일어날지 알 수 없는 미지의 확률에서, 조금이나 몇 퍼센트에 해당하는 지역별 위험도에 대한 정보를 갖게된다면, 그에따른 상품의 변화가 충분히 일어날 수 있다고 생각합니다.
  4. **개발 애로사항**
* 공공데이터 포털에서 데이터를 가져오므로 데이터 확보에서는 큰 어려움이 따르지 않을 것으로 보이나, 이 데이터들을 가공하는 알고리즘 개발의 신뢰도 및 정확도가 가장 커다란 문제로 사료됩니다.
* 어플리케이션을 통하여 서비를 제공 할 때, 어떠한 입력데이터를 받을지에 대한 선별이 필요하며, 이 입력 데이터에 대해 어떻게 처리할 것인지에 대한 구상이 어려울 것으로 생각 됩니다.
* 현실적으로 바라보았을 때, 프로젝트의 성격상 잠깐의 일회성인 흥미를 끌다가 잊혀질 가능성이 굉장히 높습니다. 그러므로 프로젝트의 재사용성을 위하여 지속적인 흥미를 유발할 수 있는 좋은 컨텐츠 개발이 요구됩니다.
  1. **개발 일정  
     **

**10. 참고자료**- 공공데이터 포털 : https://www.data.go.kr/  
- SBS데이터저널리즘팀 <마부작침>의 2016 범죄여지도 : http://mabu.newscloud.sbs.co.kr/20170308crimemap/web/index.html  
- 이미지 출처 : <https://www.pinterest.co.kr/>

|  |  |  |  |
| --- | --- | --- | --- |
| **사용 데이터 셋** | **세부 데이터셋** | **링크** | **기타** |
| 사고발생총괄현황 | 사고발생 현황(연간) | <https://www.data.go.kr/dataset/15014225/fileData.do> | 매년 발간되는 재난연감 통계자료를 기초로 한 사고발생현황자료 |
| 교통사고정보 | 무단횡단사고다발지, 자전거사고다발지, 교통사망사고정보 | <https://www.data.go.kr/dataset/15003493/fileData.do> | 교통사고정보를 위치데이터 기반 제공 교통사고 항목 별 정보제공 |
| [범죄통계] 발생시간·요일 | 범죄 발생시간 및 요일 | <https://www.data.go.kr/dataset/3074459/fileData.do> |  |
| [범죄통계] 발생 및 검거 현황 (지방경찰청별) | [범죄통계] 발생 및 검거 현황 (지방경찰청별),[범죄통계] 2015년 발생시간·요일 | <https://www.data.go.kr/dataset/3074451/fileData.do> | 전국 경찰관서에 고소, 고발, 인지 등으로 형사입건된 사건의 발생, 검거, 피의자에 대한 죄종별 분석 현황 |
| [범죄통계] 범죄발생장소 (장소별) |  | <https://www.data.go.kr/dataset/3074463/fileData.do> |  |
| [범죄통계] 범죄발생지 (지역별) |  | <https://www.data.go.kr/dataset/3074462/fileData.do> |  |
| 종관기상관측 |  | <https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36> | 종관기상관측장비(ASOS)로 부터 관측된 자료 |