**R기반의 데이터 분석 프로젝트 기획안**

|  |  |
| --- | --- |
| **팀 명** | 5조 |
| **팀 원** | 박민국, 윤정인 |
| **프로젝트 명** | 로또 번호 분석 |
| **프로젝트 주제 및 내용**  해결하고자 하는 문제 최종 산출물의 청사진 | 과거 로또번호 데이터를 이용하여 어떤 번호가 얼마나 많이 추첨되었는지를 분석하고 추첨될 가능성이 높은 번호를 추려냅니다. 또한 당첨금액과 당첨된 사람들의 수를 비교 분석하여 평균적인 당첨확률과 당첨 금액을 파악합니다. 추가적으로 홀수와 짝수의 비율을 분석하여 홀수와 짝수 중에서는 어떤 수가 더 확률이 높을지 파악해보고자 합니다. |
| **주제 선정 배경**  주제 선정의 정당성  산출될 결과의 유용성 | 요즘들어 개인정보 노출로 인해 광고메세지가 늘어나고 있습니다. 그런 광고들 중에는 로또 번호를 알려준다고 하는 광고도 간간히 보입니다. 하지만 이런 광고가 신뢰할 수 있는 정보를 주는지는 알 수 없습니다.  저희는 이런 사설 업체에서 뽑아주는 로또 번호가 정말 가능성이 있는지에 대한 의구심이 들었고 실제 데이터를 기반한 정보가 맞는지 확인해보고 싶어졌습니다.  따라서, 로또 공식홈페이지에 있는 API에 기반하여 평균적인 당점확률을 계산해보고 직접 각 번호의 확률성을 구해 확률에 기반한 번호를 직접 뽑아보고자 했습니다. |
| **데이터 수집 계획**  데이터 수집 사이트  정보와 데이터 수집  내용 | 로또 공식홈페이지 |
| **프로젝트 수행 일정**  수행 일정 | 아이디어 도출 : 09.03~09.10  자료 수집 및 전처리 : 09.11~09.16  데이터 분석및 시각화 : 09.17~09.22  결론 도출 및 보고서 작성 : 09.23 |