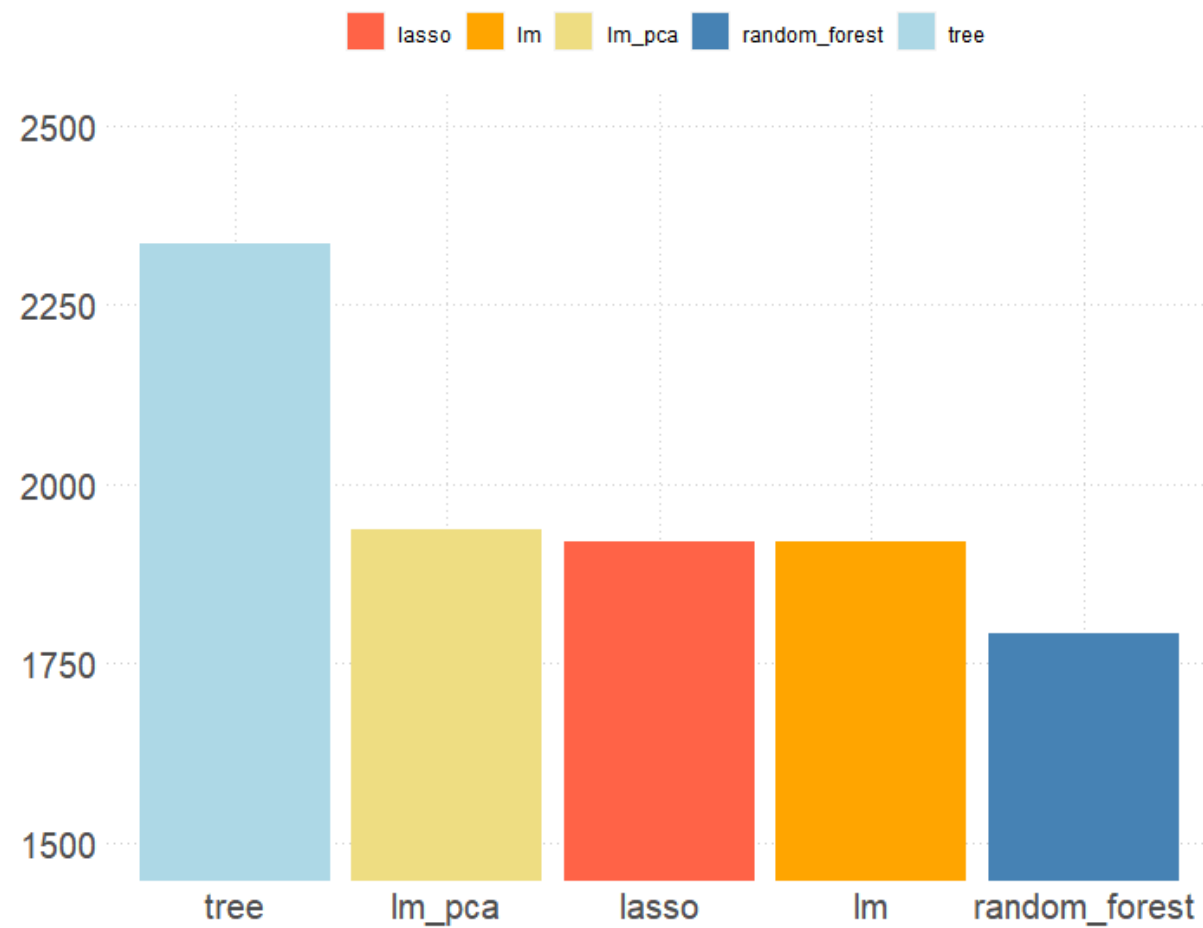


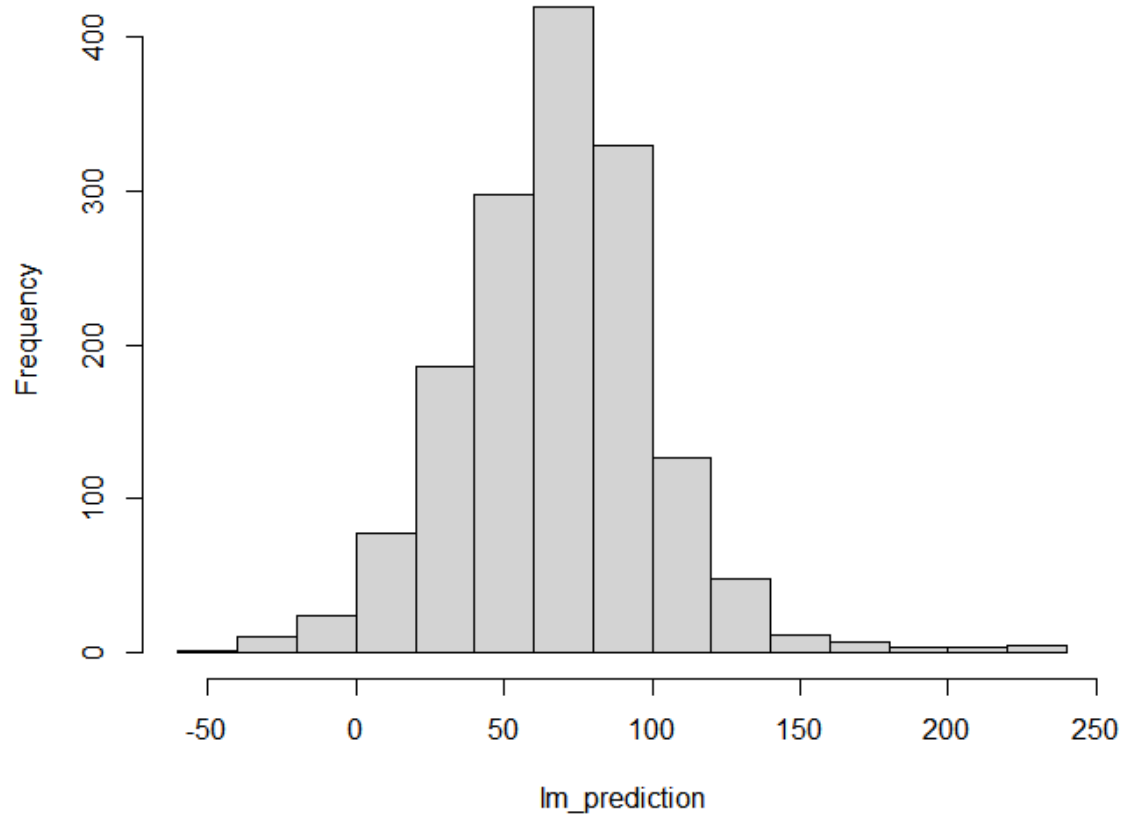
1. 선형회귀모형
2. PCA를 이용한 선형회귀모형
3. LASSO
4. Tree 모형
5. Random Forest 모형

10-fold CV error를 기준으로 모형 선택

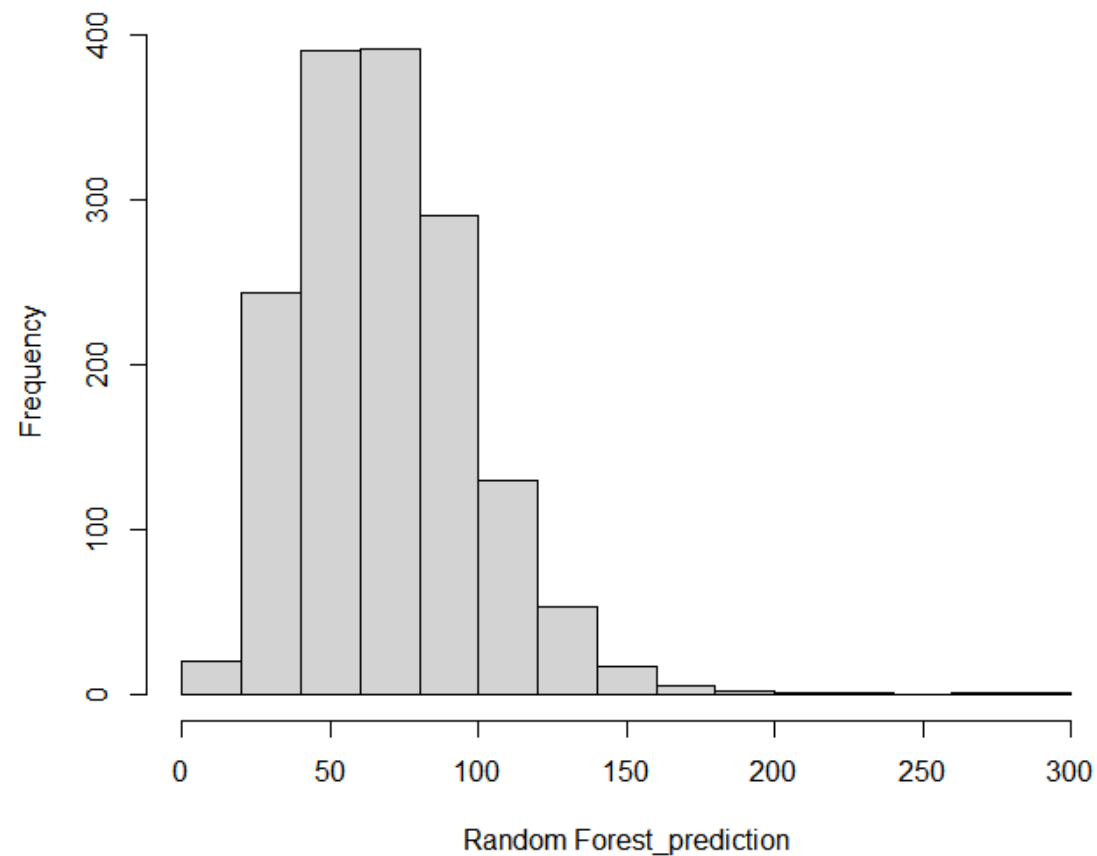


# 선형회귀모형 이용 시 문제점과 Random Forest 모형

중회귀모형 예측값 분포

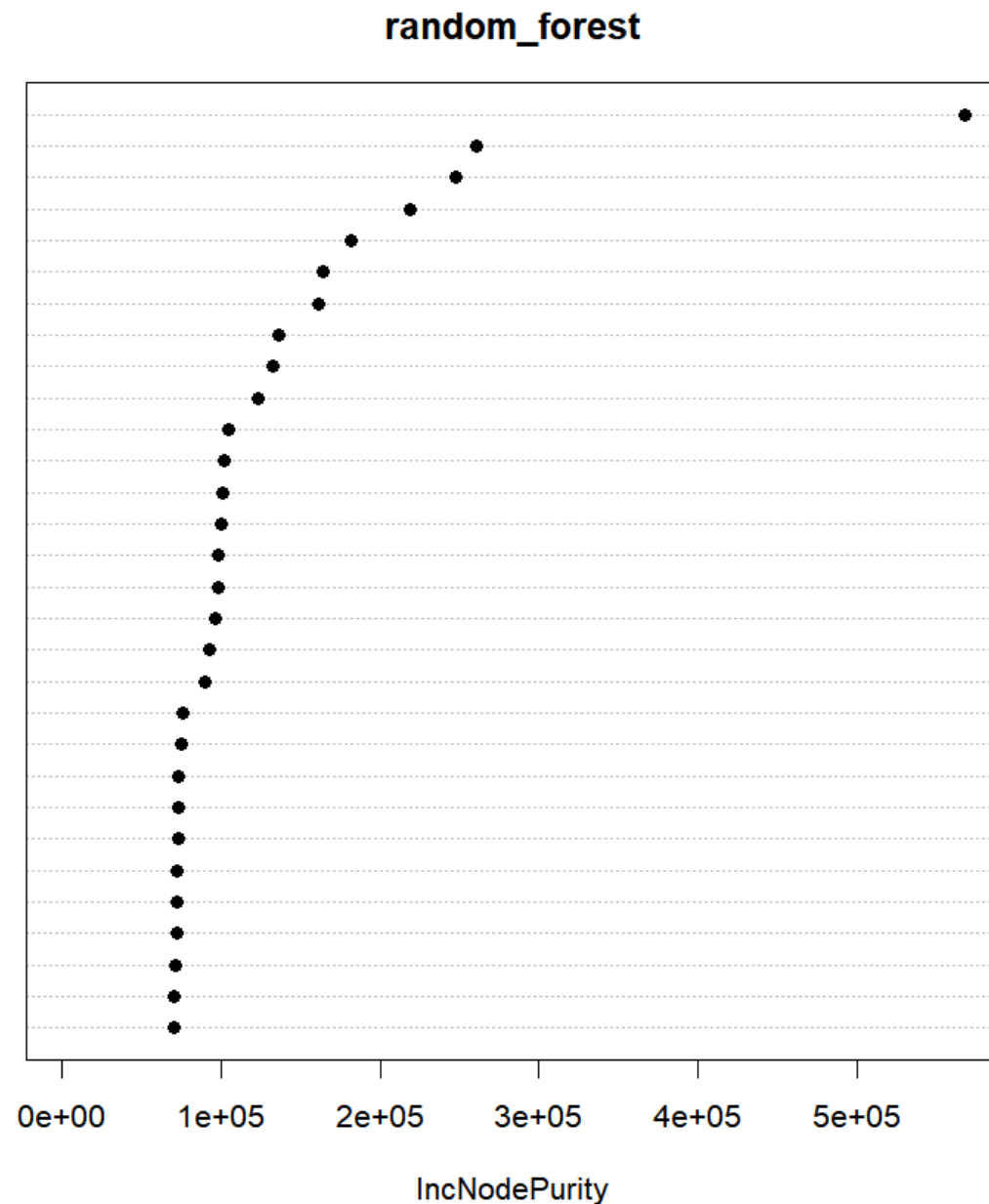


Random Forest 예측값 분포

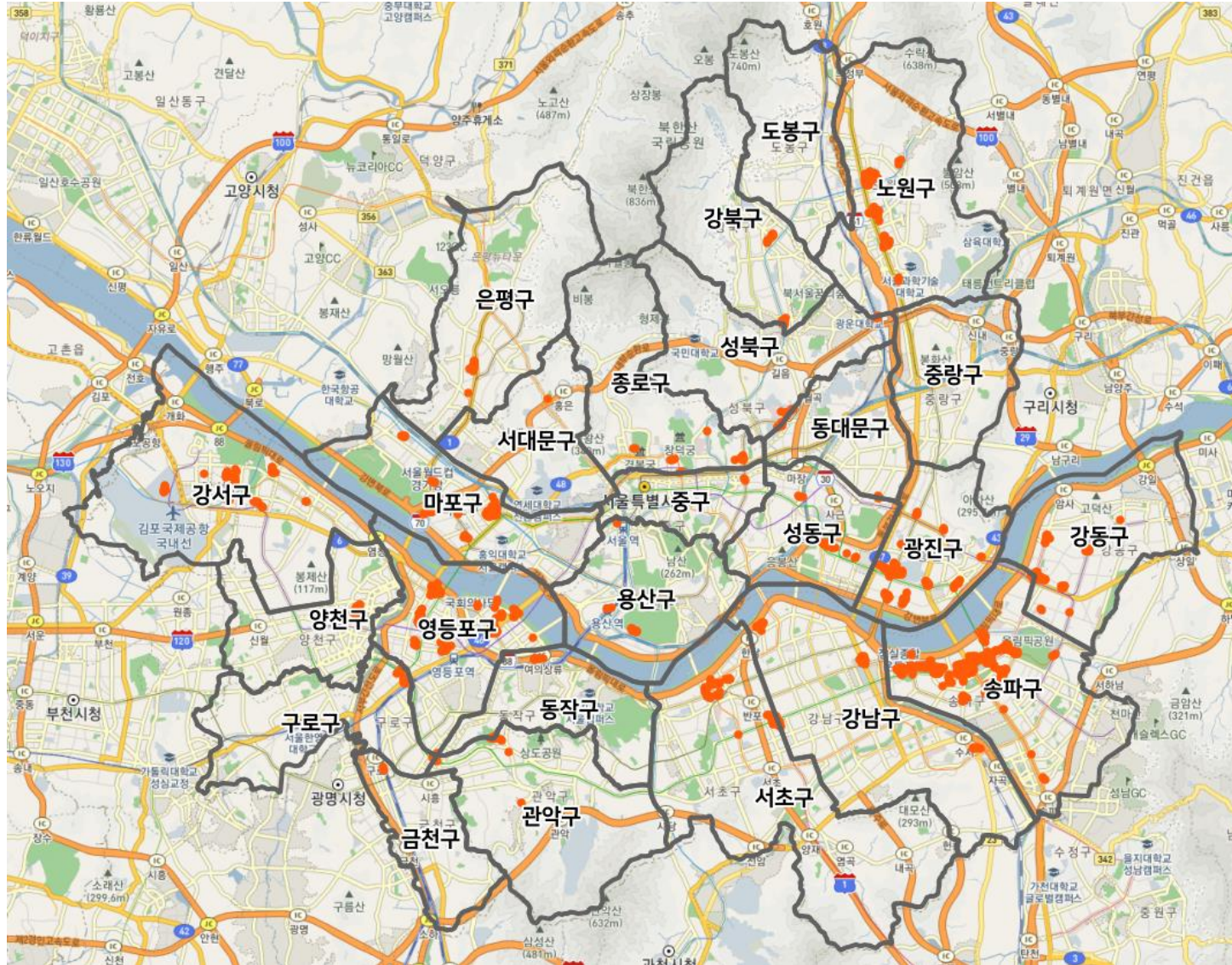


# Random Forest 모형 이용 시, 변수별 중요도

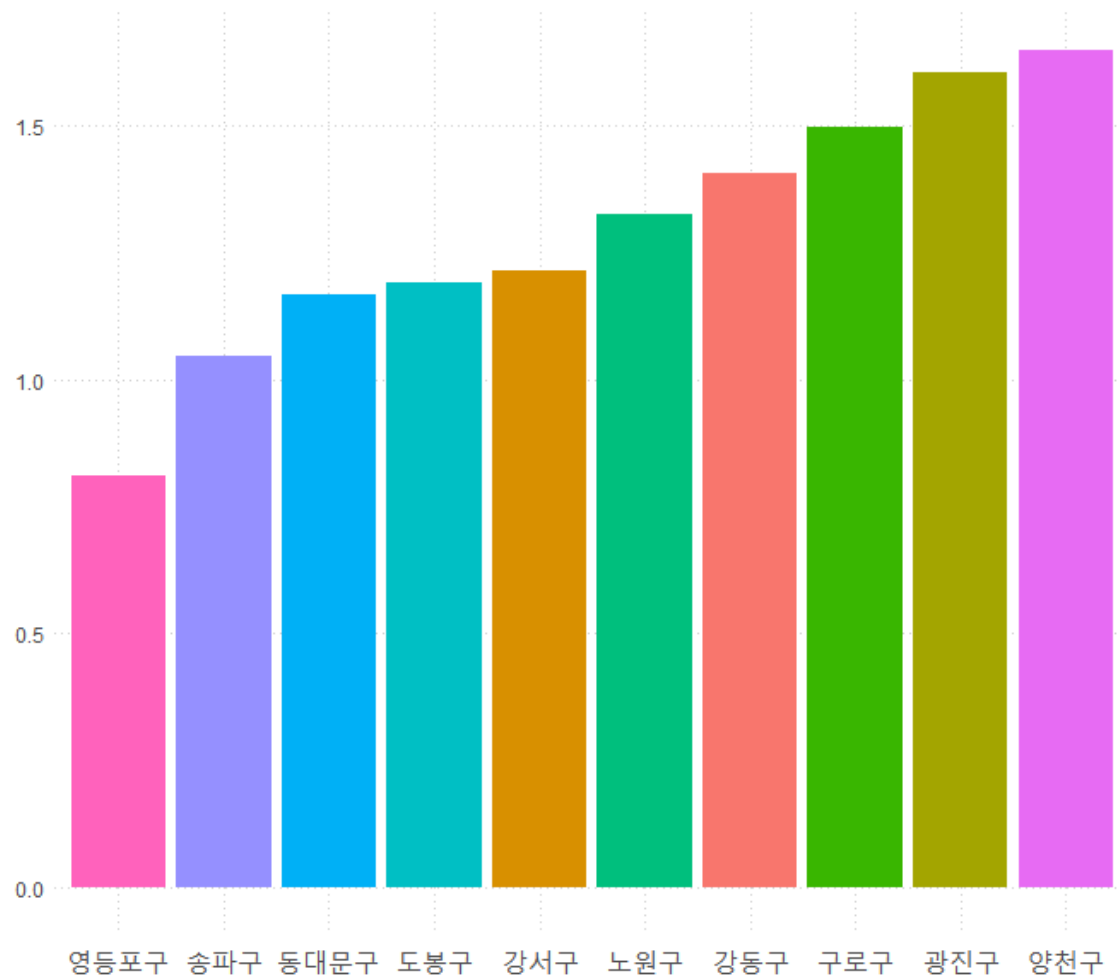
평균\_경사  
 지하철\_승  
 유동20대  
 거리\_지하  
 거리\_관광  
 버스\_승객  
 유동30대  
 거리\_하천  
 거리\_자전  
 유동10대  
 거리\_특화  
 거리\_공원  
 거리\_문화  
 거리\_영화  
 주거\_50대  
 거리\_상업  
 주거\_40대  
 거리\_대학  
 유동40대  
 거리\_체육  
 유동10대미  
 거리\_공공  
 거리\_교통  
 거리\_의료  
 유동70대이  
 주거\_20대  
 유동60대  
 유동50대  
 거리\_주차  
 near



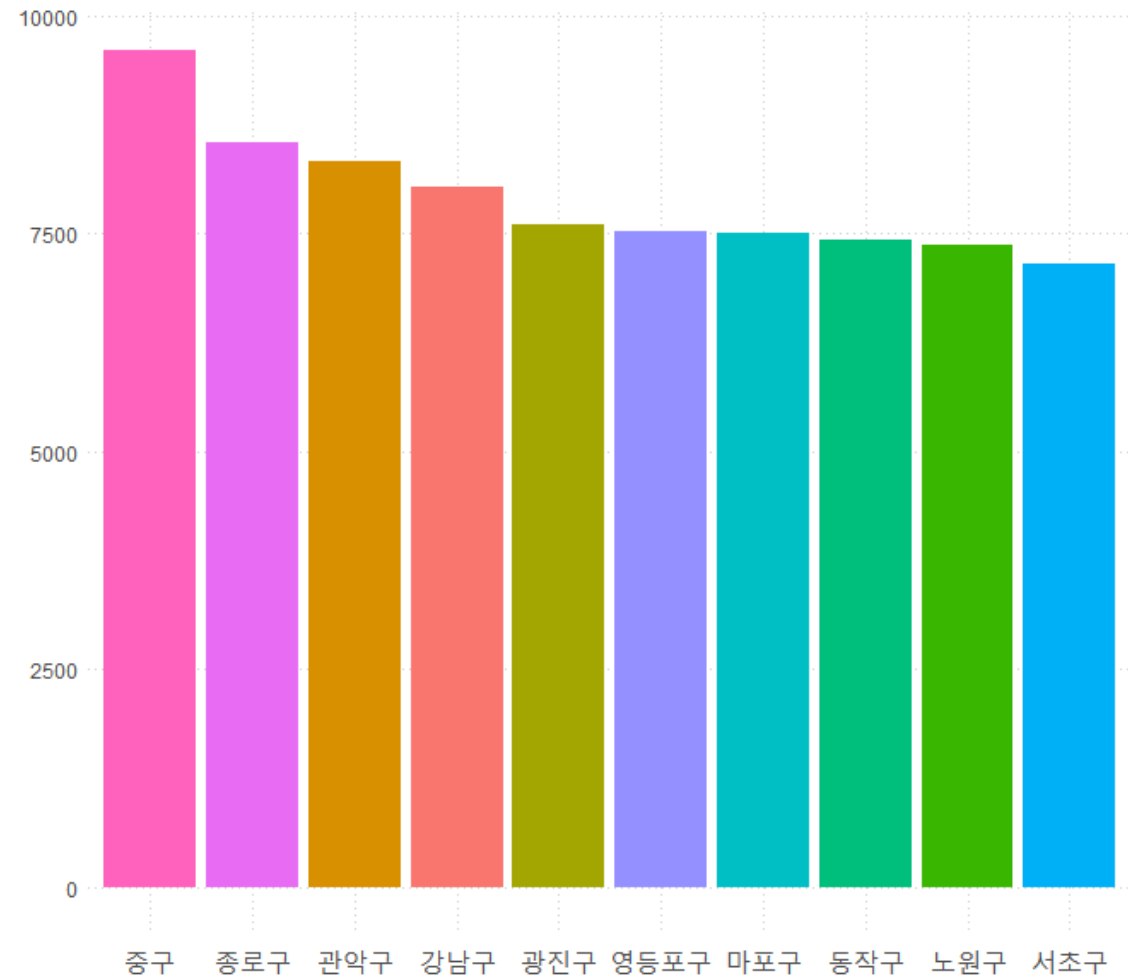
## 총 이용량 예측값 중 상위 5% (122.67명 이상)



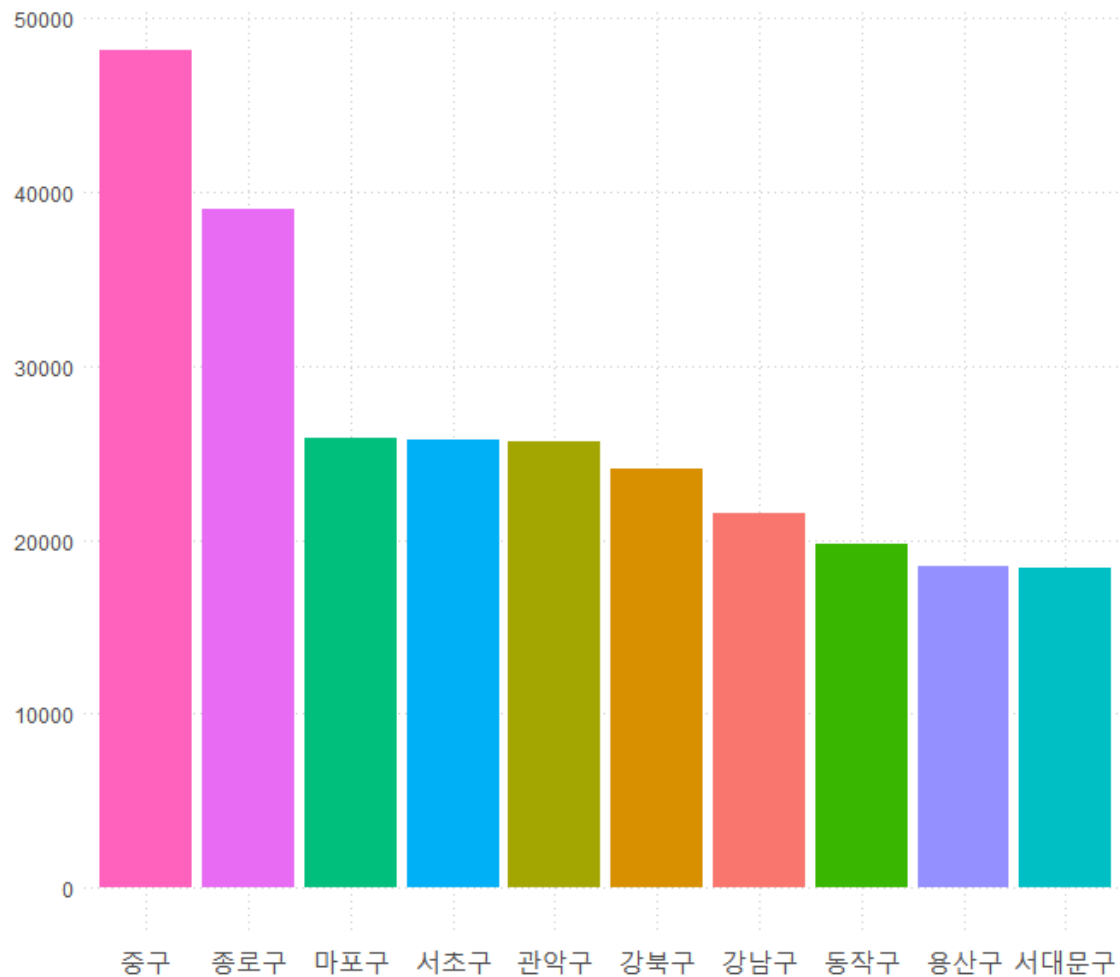
일정 반경 내의 평균 경사도



일정 반경 내의 30대 이하 유동인구 수



일정 반경 내의 평균 대중교통 이용자 수



이전에 진행했던 기초통계분석과  
직관에 어느정도 들어맞는 결과가  
도출되었다고 할 수 있다.