

코퍼스언어학

기말과제

홍정하

제출 기한 및 형식

- ▶ 배점: 20점
- ▶ 제출 기한: 12월24일(금) 23시59분
- ▶ 제출형식
 - 과제 지침 및 평가 기준을 참고하여 데이터 가공 및 분석 과정에서 사용된 R 코드, 출력결과, 데이터 분석 의미를 레포트 형식(서론, 본론, 결론)을 갖추어 아래아한글 또는 MS Word로 작성하여 제출(pdf 가능)
 - R 스크립트 파일 또는 ppt 파일은 제출 불가
- ▶ 분량: A4 10장 이상
 - 편집용지 글꼴, 문단크기는 상식적 수준에서 설정할 것. 만약 과도한 편집용지 설정, 문단크기 및 글꼴 크기를 설정할 경우 분량에 따라 감점 부과
 - 분량 미달: A4 용지 -1장당 -2점

과제 지침

▶ 텍스트 선택

- 수업에서 다루지 않은 텍스트를 자유롭게 선택
- 단, 파일은 반드시 2개 이상으로 구성되어야 함. ⇒ 파일 수 부족 시 하나당 -2점 감점

▶ 필수 포함 데이터 분석 기법

- 빈도분석, n-gram, 연어, 키워드분석, 탐색적 분석(계층적 군집 분석, 주성분분석, 대응분석 중 택1) ← 서로 다른 4개 분석 기법 사용, 즉, 서로 다른 데이터 분석 기법 4개가 최소 각각 한 번 이상 사용되어야 함.
- 탐색적 분석은 계층적 군집분석, 주성분분석, 대응분석 중 한 가지만 사용해도 되며, 만약 분석의 완결성을 위해 두 개 이상 서로 다른 기법을 사용 가능
- 필수 포함 데이터 분석 기법 누락 시 기법 하나당 -2점 감점

- ▶ 필수 포함 코드(총 7개 코드)
 - for 제어문
 - if 제어문
 - 함수 paste
 - 함수 substr 또는 함수 substring 중 하나
 - 집합론 함수 중 하나
 - 함수 grep, grepl, gsub, strsplit 중 하나 이상 사용하여 정규 표현식을 반드시 한 번 이상 사용
 - 함수 match 또는 %in% 중 하나
 - 필수 포함 코드 누락 시 하나당 -2점 감점
- ▶ 수업에서 다루지 않은 코드 및 데이터 분석 기법은 사용할 수 없음. ⇒ 위반 시 하나당 -2점 감점

- ▶ 데이터 가공 및 분석 과정에서 사용된 R 코드 및 출력결과를 모두 제시해야 하며, 데이터 분석 결과의 의미 해석에 대한 타당한 근거를 제시해야 함.
- ▶ 레포트 형식으로 제출
 - 서론, 본론, 결론의 구성
 - 관찰 데이터를 선택한 이유 제시
 - 데이터 가공 과정에 사용한 R 코드 및 가공 과정에서 도출된 결과 제시해야 하며, 가공 과정의 필요성에 대해 기술
 - 데이터 분석에 사용한 R 코드 및 분석 결과를 제시하는 한편 데이터 분석 결과를 어떤 의미로 해석할 수 있는지를 근거와 함께 기술

평가 기준 (감점제)

▶ 형식 평가

- A4 용지 10장 이상 ⇒ 분량미달: 1장당 -2점
- 수업에서 다루지 않은 코드 및 데이터 분석 기법 사용 시 하나 당 -2점
- 텍스트 선택 기준 위반 시 하나당 -2점
- 필수 포함 데이터 분석 기법 누락 시 하나당 -2점
- 필수 포함 코드 누락 시 하나당 -2점
- R 이외의 다른 도구 및 프로그래밍 언어 사용 시 0점 처리
- 서론, 본론, 결론 구성으로 과제가 작성되지 않았다면 -2점
- 관찰 데이터 선정 이유 미제시 -2점
- 데이터 가공 과정 및 필요성 누락 건별 -2점
- 데이터 분석 과정 및 해석 누락 건별 -2점

▶ 내용 평가

- 코드 사용 및 통계분석 기법 적용의 적절성
 - 데이터 가공 및 분석에 사용된 코드 사용의 적절성 ⇒ 부적절 항목 하나당 -2점 감점
 - 데이터 가공 방식 및 데이터 분석 기법 사용의 적절성 ⇒ 부적절 항목 하나당 -2점 감점
- 이유와 해석 설명의 적절성 및 타당성
 - 관찰 데이터 선정 이유/가공 과정 필요성/데이터 분석 결과의 의미 해석의 적절성 및 타당성
 - 건별 -2점 감점

▶ 기타 감점 내역 발견 시 사안에 따라 추가 감점