

Chapter. 00

강의소개

I 미디어 파트 개요 및 소개

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 조코딩

Chapter. 00

미디어 파트 개요 및 소개

I 유튜브에서 누구나 배울 수 있는 쉬운 코딩 채널을 만들어가는 조코딩입니다.



I 유튜브 데이터를 분석하여 인기 채널 순위 시각화 및 인기 키워드를 분석합니다.



1. 인기 채널 순위 시각화
2. 인기 영상 타이틀 분석

Chapter. 02

인기 채널 순위 시각화

| 환경 설정 (라이브러리, | 폰트, 데이터)

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 조코딩

Chapter. 01

환경 설정 (라이브러리, 폰트, 데이터)

I 데이터 분석을 위한 환경 설정을 진행합니다.

1. 한글 폰트 설치 및 적용
2. Warning 출력 무시
3. 라이브러리 import
4. 데이터 가져오기

Chapter. 02

인기 채널 순위 시각화

I 데이터 구조 파악 및 정리

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 조코딩

Chapter. 02

데이터 구조 파악 및 정리

I 데이터를 불러오고 구조를 파악 후 분석하기 위해 정리합니다.

1. 데이터 불러오기
2. 데이터 구조 파악
3. 데이터 정리하기

Chapter. 03

인기 채널 순위 시각화

| seaborn을 활용한 데이터 시각화

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 조코딩

Chapter. 03

seaborn을 활용한 데이터 시각화

ITOP 100 채널의 순위를 시각화 해봅니다.

1. 조회수 합계 기준 TOP 100 채널 뽑기
2. Seaborn barplot을 활용한 데이터 시각화

Chapter. 01

인기 영상 타이틀 분석

I 제목 추출 및 정제하기

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 조코딩

Chapter. 01

제목 추출 및 정제하기

I 인기 영상의 제목을 분석하여 인기 영상에서 자주 사용되는 키워드를 찾아보겠습니다

1. 데이터 정리하기

2. 정규표현식을 이용한 한글 제목 추출하기

Chapter. 02

인기 영상 타이틀 분석

|soynlp 개념 설명 및 한글 | 단어 추출의 원리

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 조코딩

Chapter. 02

soynlp 개념 설명 및 한글 단어 추출의 원리

lsoynlp란?

KoNLPy, soynlp 한글 데이터 가공 라이브러리

KoNLPy는 미등록단어 문제로 새로운 단어를 인식 불가 (ex.

“너무너무너무”)

soynlp는 학습 기반으로 새로운 단어도 인식 가능

Isosynlp란?

WordExtractor?

- 통계 기반 단어 추출기 (문서집합에서 통계 기반 자주 등장하는 단어열 추출)

Tokenizer?

- WordExtractor의 결과를 이용하여 단어의 경계를 따라 문장을
(Tokenizer - 띄어쓰기 O, MaxScoreTokenizer - 띄어쓰기 X)

단어열로 분해

I 한글 단어 추출 과정

1. soynlp 설치
2. 전체 제목 데이터를 이용해 WordExtractor로 단어 추출
3. LTokenizer로 단어의 경계를 따라 L-R 분리
4. L만 남겨서 의미를 지닌 단어 추출

I 한국어에서 의미를 지니는 단어 (명사/동사/형용사/부사)는 어절 왼쪽에 있음 (L-R)

짜장면/명사(L) + 을/조사(R)

먹/동사(L) + 었어/어미(R)

Chapter. 03

인기 영상 타이틀 분석

| 제목 단어 추출 및 빈도 | 수 구하기

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 조코딩

I 제목 단어 추출 및 빈도 수 구하는 과정

1. soynlp 설치

2. 전체 제목 데이터를 이용해 WordExtractor로 단어 추출

3. LTokenizer로 단어의 경계를 따라 L-R 분리

4. L만 남겨서 의미를 지닌 단어 추출

5. 인기 영상 제목 단어 빈도수 Count

Chapter. 04

인기 영상 타이틀 분석

I워드 클라우드 그리기

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 조코딩



I 진행 순서

1. WordCloud 그려보기

2. StopWord 설정 및 사이즈 조정 후 다시 그려보기

I StopWord란?



