

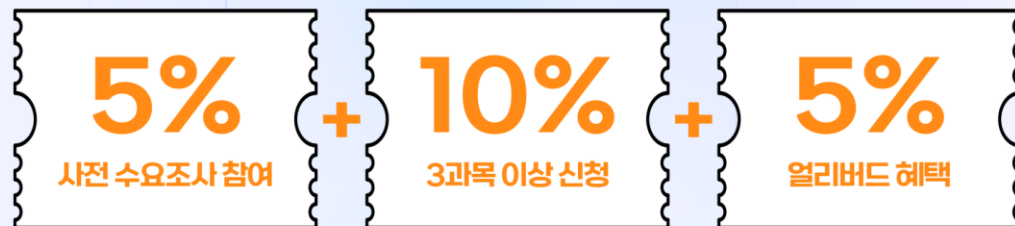


내 분야의 텍스트로 논문쓰기

텍스트마이닝 학습하기

와이즈인컴퍼니 2025년 오프라인 교육 수강생 모집

어려운 통계강의를 **기본원리부터 차근차근**, 초보자도 알기 쉽게
개념강의와 데이터 실습을 통합한 강의 커리큘럼!



최대 할인 혜택 20%

6월 20일 까지 등록 시 최대 20% 할인 혜택(중복할인 가능)

과정 소개

10시~17시(점심시간 1시간)

SPSS 과정

1. SPSS 기초통계분석	강사 : 이창선	07.05(토)	200,000원
2. SPSS 중급통계분석	강사 : 이창선	07.12(토)	200,000원
3. SPSS 회귀분석	강사 : 이창선	07.19(토)	200,000원
4. SPSS 다변량분석	강사 : 김원표	07.26(토)	200,000원

AMOS 과정

1. AMOS 기본	강사 : 김원표	08.02(토)	250,000원
2. AMOS 고급	강사 : 김원표	08.09(토)	250,000원

Excel 과정

1. Excel 기본	강사 : 사경환	08.23(토)	150,000원
2. Excel 고급	강사 : 사경환	08.30(토)	150,000원

과정당 5명 이하 신청 시 폐강

강사 소개

김원표(와이즈인컴퍼니 대표)

한양대학교 겸임교수
데이터분석 경력 22년 / 솔루션개발 기획 총괄

사경환(밸류애드 대표)

사회조사분석사회 전문위원
사회조사분석사 1급 및 2급 자격보유

이창선(와이즈인컴퍼니 연구팀 부장)

데이터분석 경력 15년
통계학 석사 전공

장소

삼경교육센터(한강대로 391 센트럴프라자) - 서울역 인근

할인 혜택



최대 할인 혜택 20%

6월 20일 까지 등록 시 최대 20% 할인 혜택(중복할인 가능)

2과목 이상 신청 시 할인쿠폰 발급을 위해
아래 연락처(메일 가능)로 연락 부탁드립니다

교육상담 문의

교육관련 문의는 이메일 및 전화로 상담 가능합니다

담당자 이호성 교육팀 부장

메일 hosung@wiseinc.co.kr

전화 02-558-5144(내선 : 203)



김 원 표

現 (주)와이즈인컴퍼니 대표
한양대 겸임교수

【 주요 경력 】

- 20년간 [2,000건](#) 이상의 통계분석/ 빅데이터 프로젝트 수행
- [20권](#) 이상 통계분석/ 빅데이터 관련 서적 출간
- 연간 [2만명](#) 이상의 수강생이 검증한 전문 강사
- AI 통계, 데이터분석 솔루션 " [데이터인\(DataIN\)](#) " 개발 기획 총괄
- 국책연구소 인공지능 (딥러닝) 프로젝트 다수 수행

CONTENTS



텍스트마이닝 학습하기

1. 텍스트마이닝의 핵심지식 담기
2. 데이터 불러오기와 워드 클라우드
3. 감성분석
4. 토픽모델링과 LDA시각화

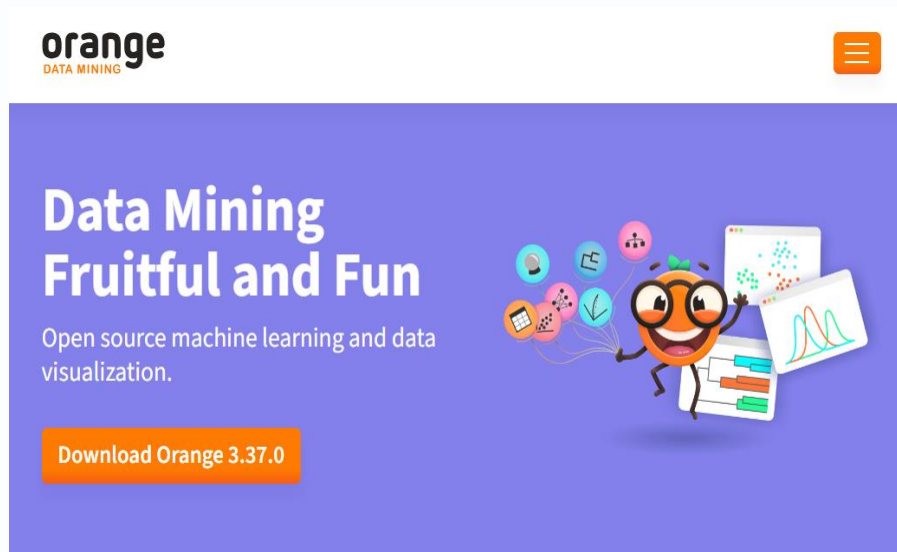
Part 01

텍스트마이닝의 핵심지식 담기



1. 무료 머신러닝 솔루션 ORANGE3 활용

ORANGE3

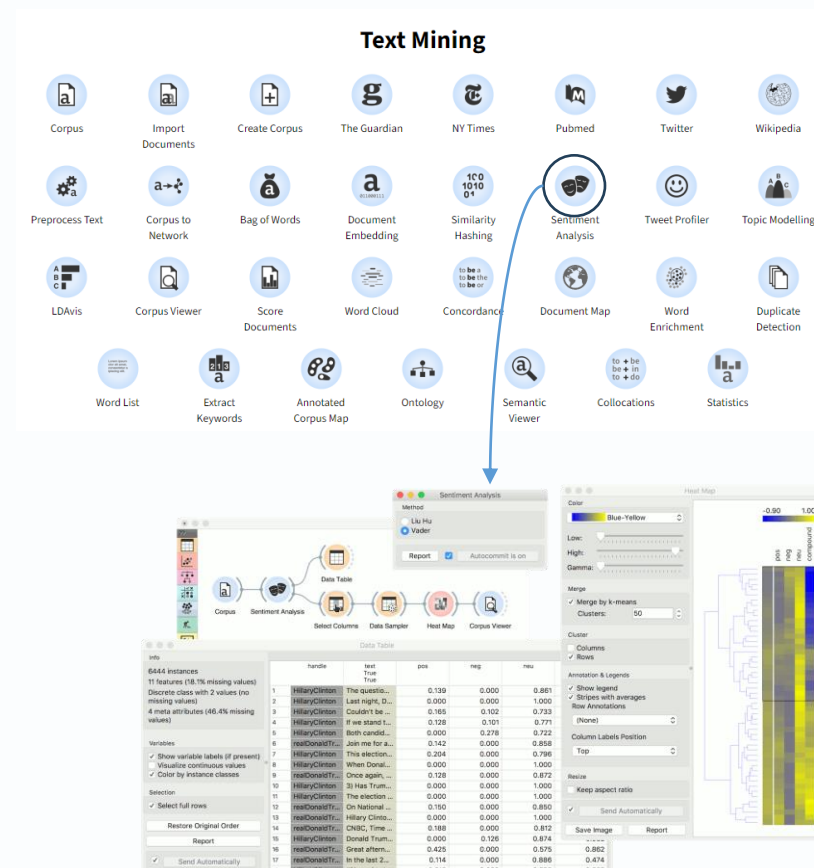


다운로드

<https://orangedatamining.com/download/>

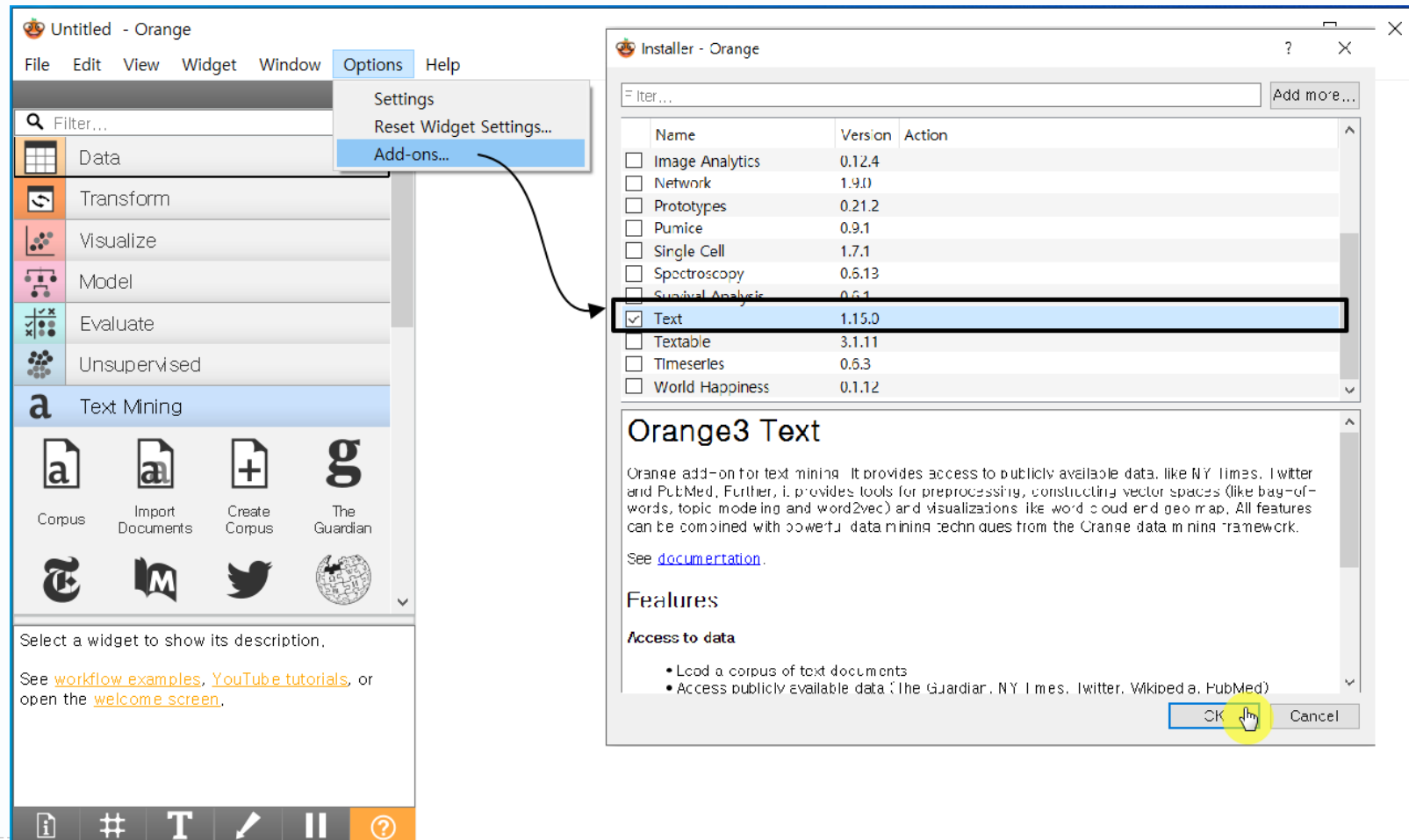
텍스트마이닝 상세 설명법

<https://orangedatamining.com/widget-catalog/>



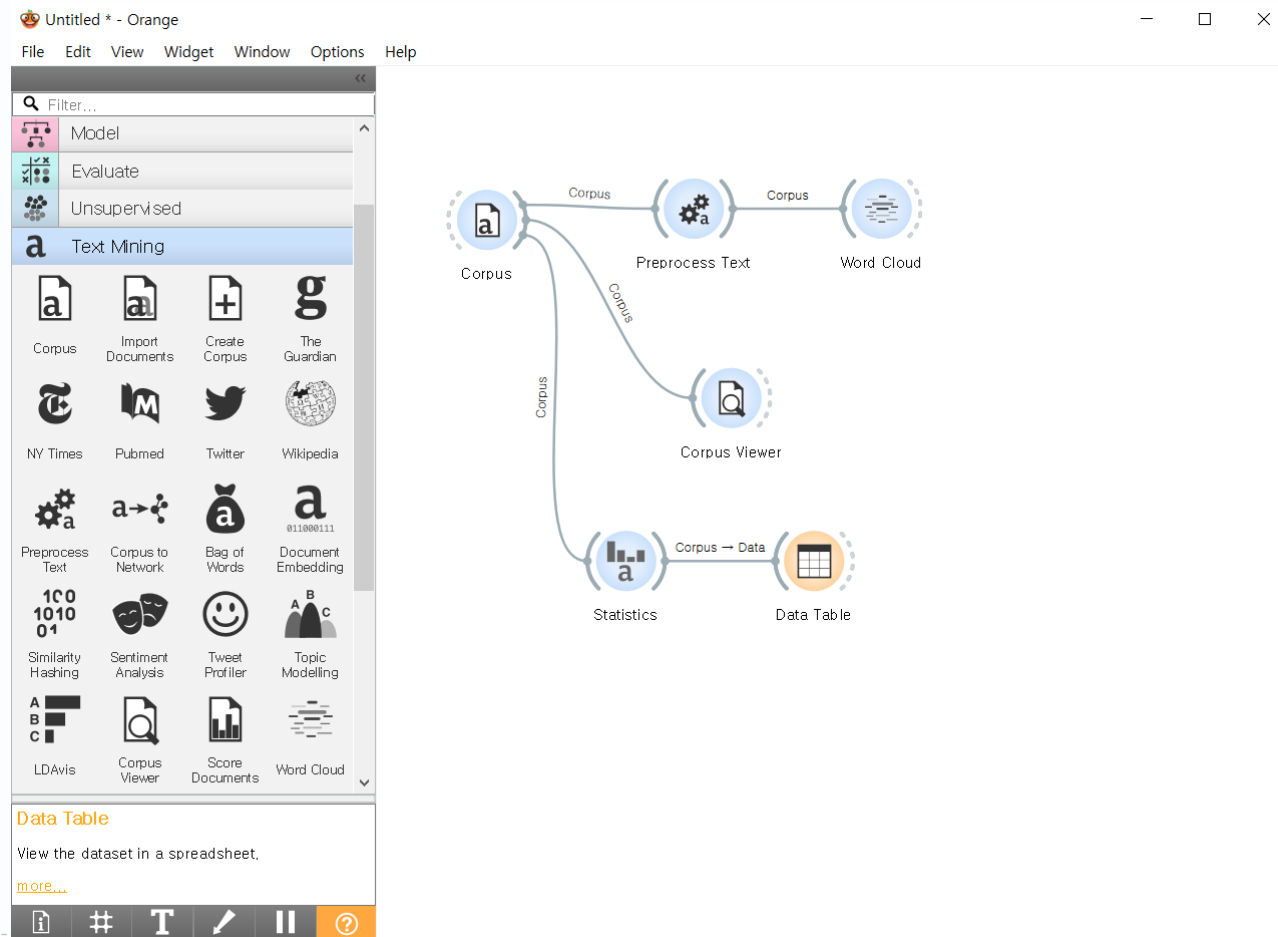
1. 무료 머신러닝 솔루션 ORANGE3 활용

- 텍스트마이닝 위젯 추가방법: Options > Add-ons 클릭 후, Text 선택하여 'OK'로 추가 설치 (꽤 많은 시간이 소요됨)



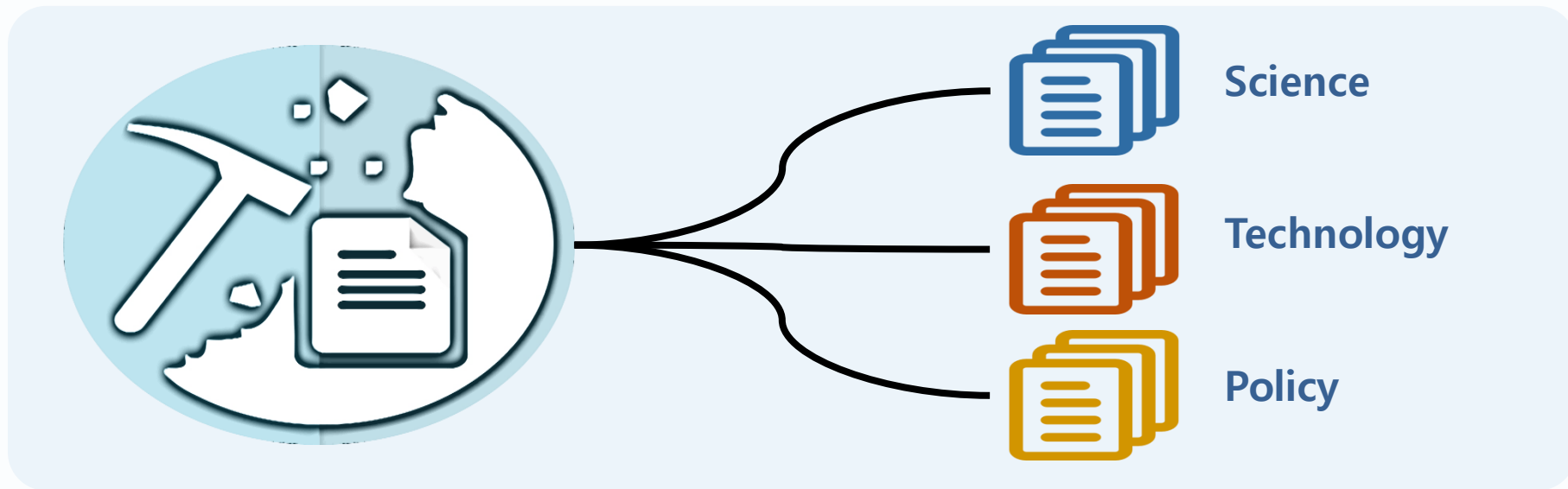
1. 무료 머신러닝 솔루션 ORANGE3 활용

- Text Mining 기능이 추가 설치됨. 왼쪽의 위젯을 오른쪽 그림으로 프로세스에 따라 설정하여 분석 수행



2. 텍스트마이닝 기본 개념

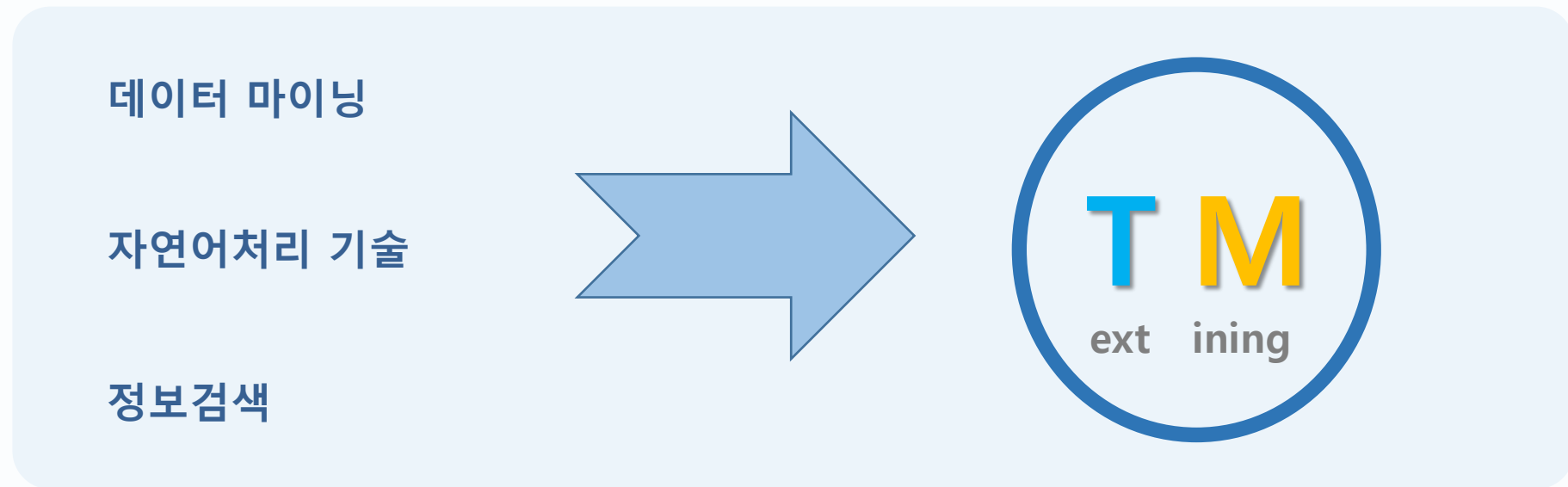
- 텍스트 형태로 이루어진 정형/비정형 데이터들을 자연어 처리 방식(Natural Language Processing)과 문서처리 방법을 적용하여 **유용한 정보를 추출**하고 가공하는 기술



- 텍스트란 일반 문서나 도서뿐만 아니라 웹페이지, 블로그, 전자저널, 이메일 등 전자문서도 포괄한 자료원천으로 우리 일상에서 가장 사람과 가까운 형태의 정보

2. 텍스트마이닝 기본 개념

- 텍스트 마이닝은 데이터로부터 유용한 인사이트를 발굴하는 **데이터마이닝(Data Mining)**, 언어를 정보로 변환하기 위한 **자연어처리(Natural Language Process)**, **정보검색** 등 다양한 분야가 접목되어 발전한 학문, 기술

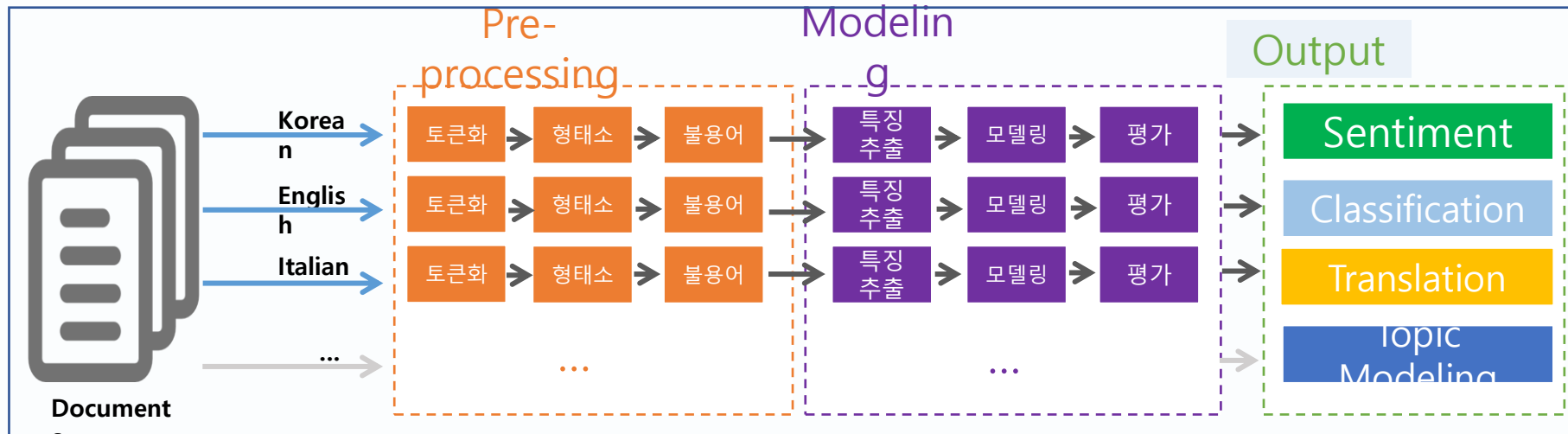
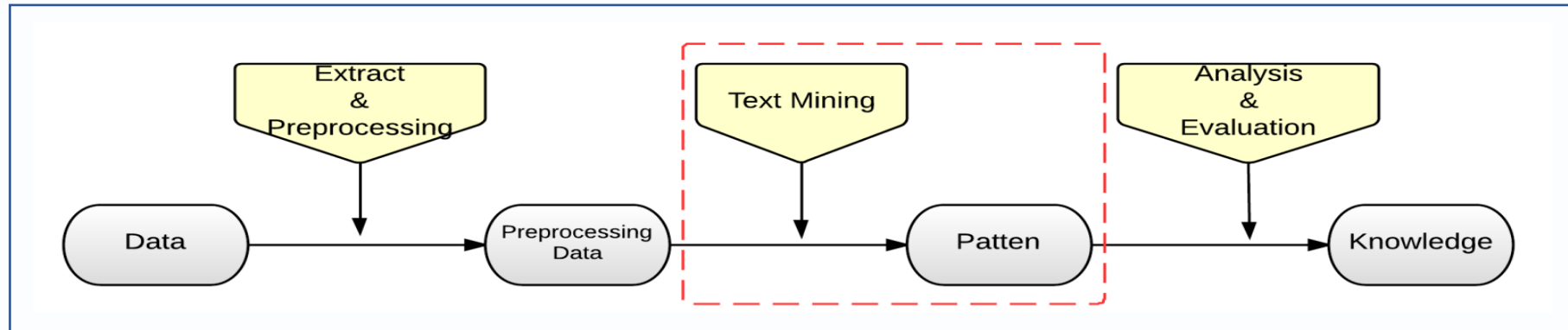


- 즉, 텍스트마이닝이란 텍스트 데이터로부터 새로운 고급 정보를 이끌어 내는 과정이라 할 수 있음

2. 텍스트마이닝 기본 개념

- **문서 요약** (Summarization) : 논문/신문/보고서 요약
- **문서 분류** (Classification) : 자동 범주화 (예: 뉴스 기사 분석 -> 사회/경제/금융 등으로 분류)
- **문서 군집** (Clustering) : 유사 단어 또는 유사 문서 간의 군집 분석
- **특징 추출** (Feature Extraction) : 주요 키워드 추출

2. 텍스트마이닝 기본 개념



2. 텍스트마이닝 기본 개념

가. Sentiment Analysis

- 텍스트 마이닝의 한 분야로 **사용자의 의견, 감정, 태도** 등을 텍스트로부터 분석하는 방법
- 텍스트 내 키워드와 관련된 감성 어휘의 빈도수를 분석하여 **긍정/부정/중립**으로 분류하고 사용자의 의견을 분석함
(예 :상품평, 후기, 선호도 등)

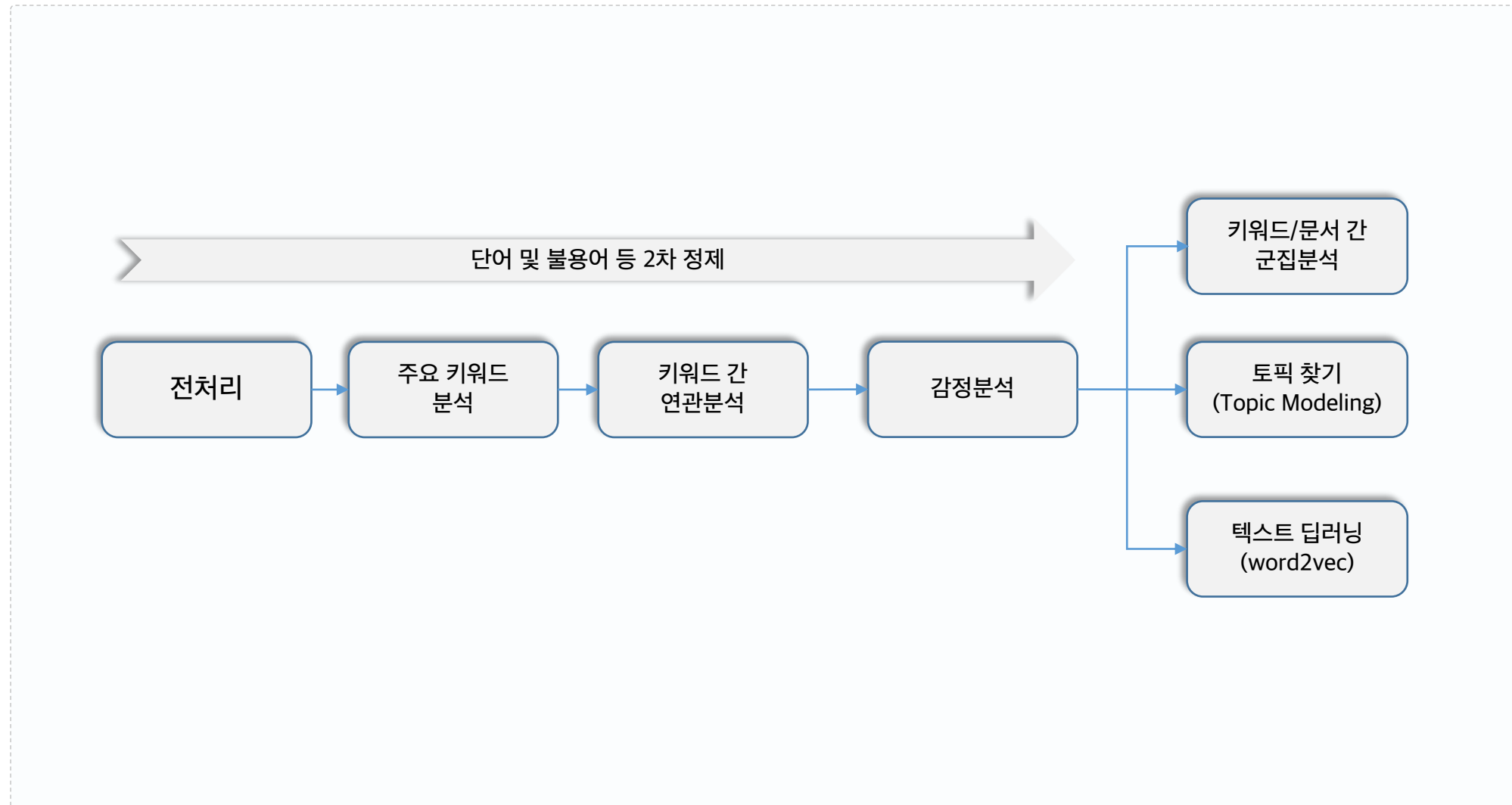
나. Topic Modeling

- 문서를 이루고 있는 키워드들을 바탕으로 문서에서 **주제(Topic)**를 도출하기 위해 사용되는 **통계적 분석방법**
- **비정형 텍스트 분석**에 많이 사용되고 있으며 다양한 종류의 데이터에도 적용이 가능

다. Word2Vec/Doc2Vec

- 단어를 벡터 공간으로 표현하는 방법으로 인공지능 기반의 임베딩 기법
- 단어 사이의 **분포 관계**를 바탕으로 단어를 일정한 의미를 갖는 **벡터로 변환**하는 기술
- Word2Vec 알고리즘의 확장으로 **문장, 단락 또는 전체 문서**와 같이 더 큰 텍스트를 표현

3. 텍스트마이닝 프로세스



Part 02

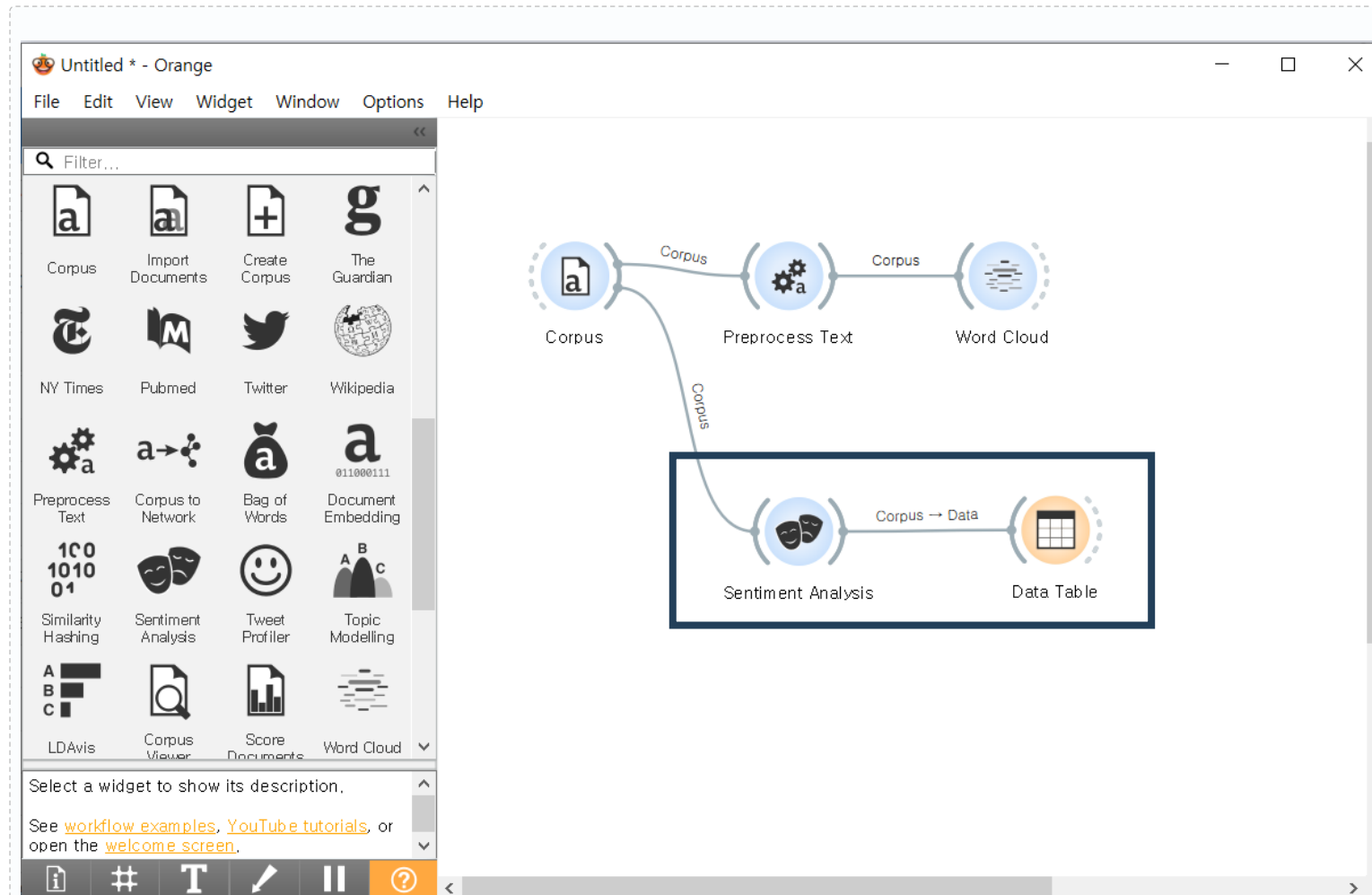
데이터 불러오기와 워드 클라우딩

1. ORANGE 구성

Part 03

감성분석

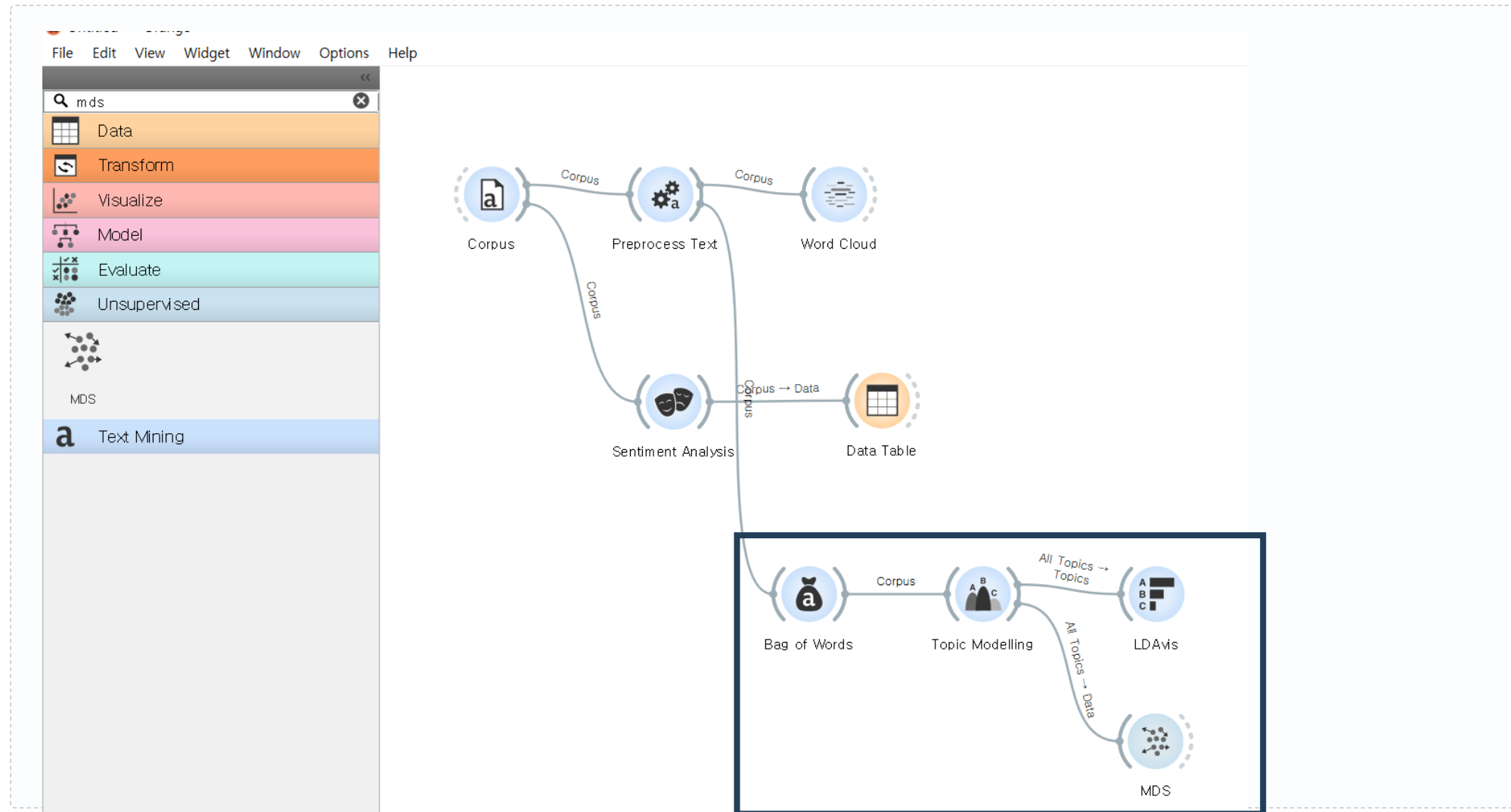
1. ORANGE 구성



Part 04

토픽모델링과 LDA시각화

1. ORANGE 구성





THANK YOU

(주)와이즈인컴퍼니 / 서울시 강남구 언주로 309, 기성빌딩 3층 / T 02.558.5144 / F 02.558.5146