해외 관세청 주요 고시정보 제공 서비스

KOTRA 빅데이터팀 남환우 PM

<멘토 소개>

▶ 성명: 남환우

▶ 학력: 연세대학교 경제학과 졸업 (학사)

▶ 경력

KOTRA Invest KOREA (2005 ~2011) KOTRA 운영지원실 / 수출지원실 (2012~2017) KOTRA 빅데이터팀 (2018 ~ 현재)

▶ 담당업무

KOTRA 빅데이터 플랫폼(트라이빅) 기획 및 운영 데이터분석모델개발 (Analytics & AI Engineer) 데이터아키텍트 (DA)

▶ 자격증

데이터분석전문가(ADP) 빅데이터분석기사 데이터아키텍트전문가(DAP) 경영지도사(CMC)

▶ 분석모델

국가별 산업별 상관계수 기반 신시장유망기업 추천모델 (통계모델) 기업의 품목별 국가별 수출금액 기반 미진출 해외시장추천모델 (추천모델) 수출기업 중 KOTRA 회원이 아닌 기업의 회원화 확률모델 (ML - 분류) 한국 수출금액 예측 모델 (DL - 회귀) 텍스트 데이터 기반 해외시장동향 탐색 (Text Mining)

▶ 분석언어: Python, R, SQL

<프로젝트 개요>

1. 프로젝트명: 해외관세청 고시 정보 분석 서비스

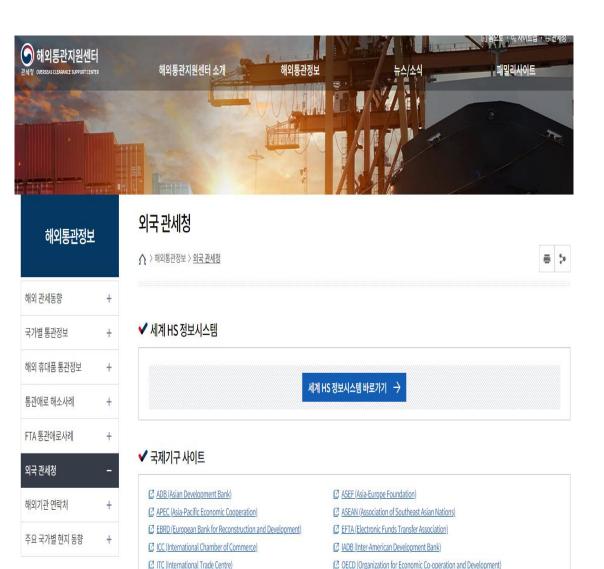
2. 주요내용

- ▶ 한국의 주요 교역대상국 관세청 고시 정보 수집 (5개 이내)
- ▶ 대상국의 수출입 규제정보 알림 서비스 (사건 발생시)
- 고시정보 텍스트 분석을 통한 이상동향 탐지

3. 추진방안

- ▶ 대상 교역국 결정 및 해당국 관세청 사이트 분석 (크롤링)
- ▶ 수출입규제 고시 특징 분석 및 알림서비스 (사건 발생시)
- ▶ 고시 및 뉴스 텍스트 마이닝을 통한 동향파악 (상시/주기적)

<관세청 해외통관지원센터: 각 국 관세청 사이트 연계>



UN (United Nations)

[2] WIPO (World Intellectual Property Organization)

[2] TWB (The World Bank)

∠ WCO (World Customs Organization)

customs.go.kr/foreign/cm/cntnts/cntntsView.d	o?mi=3178&cntntsId=932		B
		∠ UN (United Nations)	
	WCO (World Customs Organization)	WIPO (World Intellectual Property Organization)	
	WTO (World Trade Organization)		

✔ 아시아 관세청 사이트

[2 <u>바레인 (Bahrain)</u>	② <u>홍콩 (Hong kong, China)</u>
[2 <u>일본 (Japan)</u>	② 몰디브 (Maldives)
[2 <u>싱가포르 (Singapore)</u>	[건 <u>아랍에미리트연합국 (United Arab Emirates)</u>
[2] <u>방글라데시 (Bangladesh(</u>	[² 인도 (India)
[2 요르단(Jordan)	[2 미얀마 (Myanmar)
[건 <u>스리랑카 (Sri Lanka)</u>	[건 브루나이 (Brunei Darussalam)
② 인도네시아 (Indonesia)	[2 쿠웨이트 (Kuwait)
[2 <u>파키스탄 (Pakistan)</u>	[2 대만 (Taiwan)
[베트남 (Vietnam)	② 캄보디아 (Cambodia)
[2 이란 (Iran (Islamic Republic Of))	[즈 필리핀 (Philippines)
☑ <u>태국 (Thailand)</u>	[2 <u>예엔 (Yemen)</u>
[<u>중국 (China)</u>	[2 이스라엘 (Israel)
[<u>말레이시아 (Malaysia)</u>	② 몽골 (Mongolia)

✔ 아메리카 관세청 사이트

[2 <u>가이아나 (Guyana)</u>	[2 과테말라 (Guatemala)
② 니카라과 (Nicaragua)	[2 도미니카 공화국 (Dominican Republic)
[<u>멕시코 (Mexico)</u>	[2 <u>미국 (United States of America)</u>
② 베네수엘라 (Venezuela)	[2] 볼리비아 (Bolivia)
[2 <u>브라질 (Brazil)</u>	[2 아르헨티나 (Argentina)
② <u>우루과이 (Uruguay)</u>	② <u>자메이카 (Jamaica)</u>
[2 <u>칠레 (Chile)</u>	② 캐 <u>나</u> 吐(Canada)

<중국관세청 사이트: 해관총서>





<21년 중국 관세청 고시 제81호: 요소수 사태 촉발원인>



附件

少须头他位验的进口口简单日求明整衣

序号	海关商品编号	商品名称	当前海关 监管条件	调整后的 海关监管 条件	
1	2827101000	肥料用氯化铵		В	
2	3102100010	尿素(配额内,不论是否水溶液)	A	A/B	
3	3102100090	尿素(配额外,不论是否水溶液)	A	A/B	
4	3102300000	硝酸铵(不论是否水溶液)	A	A/B	
5	3102400000	硝酸铵与碳酸钙等的混合物(包括硝酸铵与其他无效肥及无机物的混合物)		В	
6	3102600000	硝酸钙和硝酸铵的复盐及混合物		В	
7	3102800000	尿素及硝酸铵混合物的水溶液(包括氨水溶液)		В	
8	3102909000	其他矿物氮肥及化学氮肥(包括上述编号未列名的混合物)		В	
9	3103111000	重过磷酸钙[按重量计五氧化二磷(P205)含量≥35%]	A	A/B	
10	3103119000	其他按重量计五氧化二磷 (P205) 含量≥35%的过磷酸钙	A	A/B	
11	3103190000	其他过磷酸钙	A	A/B	
12	3103900000	其他矿物磷肥或化学磷肥		В	
13	3104202000	纯氯化钾(按重量计氯化钾含量≥99.5%)		В	
14	3104209000	其他氯化钾	A	A/B	
15	3104300000	硫酸钾	A	A/B	
16	3104901000	光卤石, 钾盐及其他天然粗钾盐		В	
17	3104909000	其他矿物钾肥及化学钾肥		В	
18	3105100010	制成片状及类似形状或零售包装的硝酸铵(零售包装每包毛重不超过10千克)		В	
19	3105100090	制成片状及类似形状或零售包装的三十一章其他货品(零售包装 每包毛重不超过10千克)		В	
20	3105200010	化学肥料或矿物肥料(配额内,含氮、磷、钾三种肥效元素)	A	A/B	
21	3105200090	化学肥料或矿物肥料(配额外,含氮、磷、钾三种肥效元素)	A	A/B	
22	3105300010	磷酸氢二铵(配额内)	A	A/B	
23	3105300090	磷酸氢二铵(配额外)	A	A/B	
24	3105400000	磷酸二氢铵(包括磷酸二氢铵与磷酸氢二铵的混合物)	A	A/B	
25	3105510000	含有硝酸盐及磷酸盐的肥料(包括矿物肥料或化学肥料)	A	A/B	
26	3105590000	其他含氮、磷两种元素肥料(包括矿物肥料或化学肥料)	A	A/B	
27	3105600000	含磷、钾两种元素的肥料(包括矿物肥料或化学肥料)	A	A/B	
28	3105901000	有机-无机复混肥料	A	A/B	
29	3105909000	其他肥料	A	A/B	

KOTRA 해외시장뉴스 텍스트 데이터 분석

데이터 출처

KOTRA 해외시장뉴스 (기간: 2019.4.5 ~ 2022.7.7)

<제공데이터: 총 50건>

연	밴	보도일자	기사제목	기사본문	출처
0	1	2022-05-16	중국, 스마트 제조업의 미래는 어디	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스
1	2	2020-10-16	베트남의 스마트팜산업 육성 및 투자정책	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스
2	3	2020-06-23	터키, 스마트 시티 구축을 위한 움직임	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스
3	4	2020-05-21	[동남아지역] 포스트 코로나, 동남아 시장의 유망분야는	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스
4	5	2020-02-27	獨 스마트 전기계량기 설치, 2020년 전면 시행	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스
5	6	2019-11-18	[기고] 中 충칭시 스마트산업분야의 발전 방향 및 정책	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스
6	7	2019-05-13	헝가리 스마트 시티, 절러에게르세그(Zalaegerszeg)	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스
7	8	2022-07-06	물가상승과 주변국 수출규제 속 태국의 경제 전망	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스

<기사내용 보완: 크롤링>

	기사본문	출처	키워드	기사내용
0	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	스마트	₩n₩n₩n₩n중국, 스마트 제조업의 미래는 어디₩n₩n₩n₩n투자진출₩n중국₩n₩
1	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	스마트	₩n₩n₩n₩n베트남의 스마트팜산업 육성 및 투자정책₩n₩n₩n₩n투자진출₩n베트남
2	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	스마트	₩n₩n₩n₩n터키, 스마트 시티 구축을 위한 움직임₩n₩n₩n₩n현장·인터뷰₩n터
3	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	스마트	₩n₩n₩n₩n[동남아지역] 포스트 코로나, 동남아 시장의 유망분야는₩n₩n₩n
4	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	스마트	₩n₩n₩n₩n獨 스마트 전기계량기 설치, 2020년 전면 시행 ₩n₩n₩n₩n통상
5	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	스마트	₩n₩n₩n₩n[기고] 中 충칭시 스마트산업분야의 발전 방향 및 정책₩n₩n₩n
6	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	스마트	₩n₩n₩n₩n헝가리 스마트 시티, 절러에게르세그(Zalaegerszeg)₩n₩n₩
7	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	규제	₩n₩n₩n₩n물가상승과 주변국 수출규제 속 태국의 경제 전망₩n₩n₩n₩n경제·무
8	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	규제	₩n₩n₩n₩n아르헨티나, 주요 산업군에 외환규제 완화₩n₩n₩n₩n통상·규제₩n아
9	https://dream.kotra.or.kr/kotranews/cms/news/a	해외시장뉴스	규제	₩n₩n₩n₩n이스라엘 표준개혁안 발효로 수입규제 대폭 완화₩n₩n₩n₩n통상·규제

<바이그램 유사도: 2개의 인접단어(명사)별 분석>

	가가 물가	가격 가 량	가격 가 장	가격 가지	가격 개선	가격 갤런	가격 경쟁 력	가격 계 속	가격 고려	가격 공공재	 희망 다낭	희망 신규	희망 전 자상거래	희망 프 로그램	희생 의지	희소식 독일	희소식 에 너지	히비츠 발라	히터 보일러	히트 추가
가가 물가	1.0	0.000000	0.000000	0.0	0.0	1.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
가격 가량	0.0	1.000000	0.000000	0.0	0.0	0.0	0.688724	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000000	0.0	0.0	0.0
가격 가장	0.0	0.000000	1.000000	0.0	0.0	0.0	0.000000	0.678907	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
가격 가지	0.0	0.000000	0.000000	1.0	0.0	0.0	0.000000	0.000000	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
가격 개선	0.0	0.000000	0.000000	0.0	1.0	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
가격 갤런	1.0	0.000000	0.000000	0.0	0.0	1.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
가격 경쟁력	0.0	0.688724	0.000000	0.0	0.0	0.0	1.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.688724	0.0	0.0	0.0
가격 계속	0.0	0.000000	0.678907	0.0	0.0	0.0	0.000000	1.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
가격 고려	0.0	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.000000	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.000000	0.0	0.0	0.0
가격 공공재	0.0	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.000000	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0
10 rows >	34138	columns																		

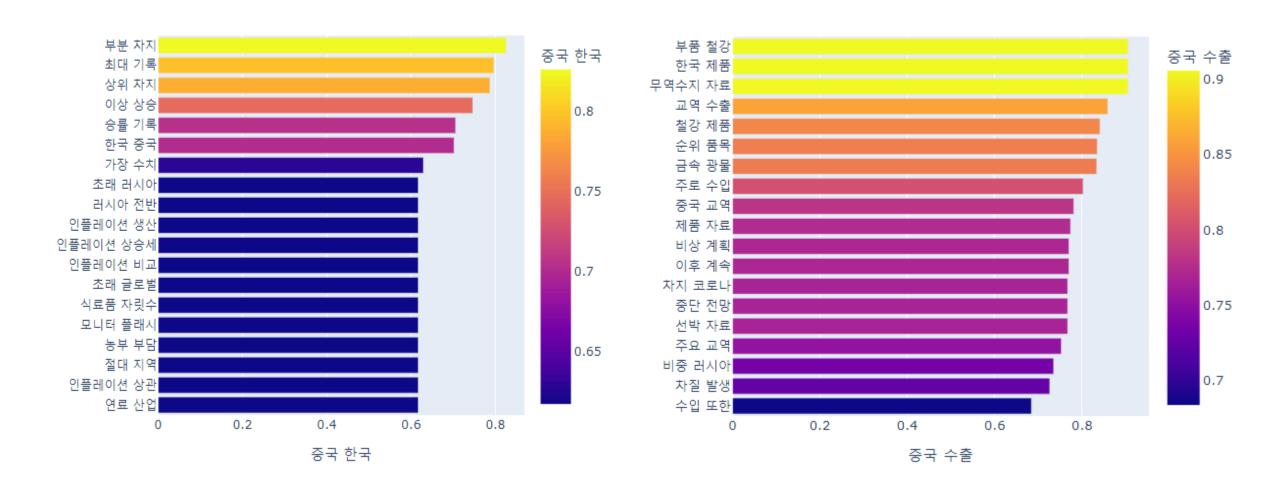
34138 * 34138 (8.7GB)

<Key Word: 중국 한국>

<Key Word: 중국 수출>

중국 한국 연관단어 top 20

중국 수출 연관단어 top 20

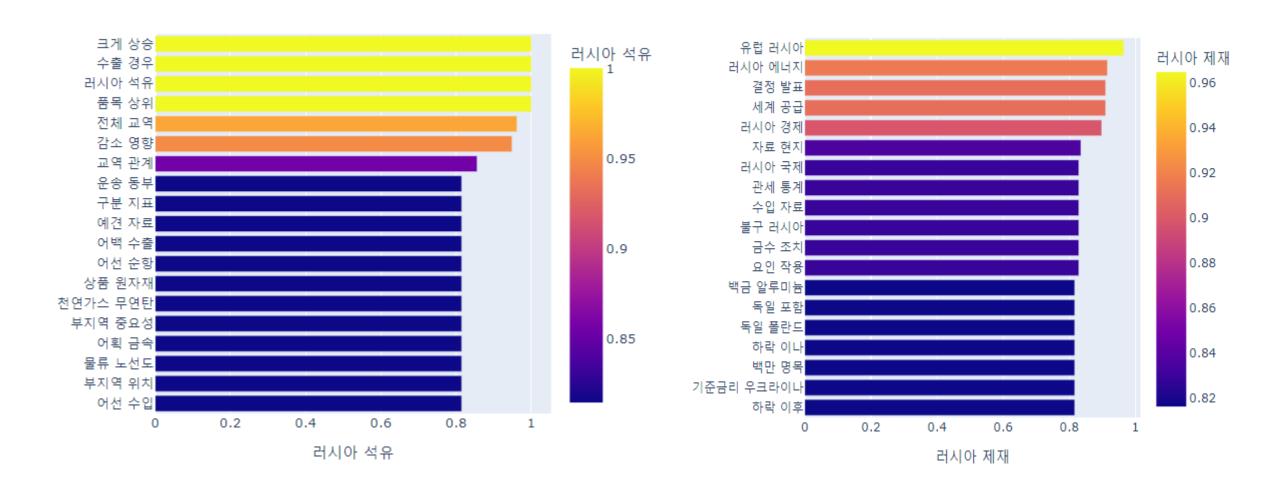


<Key Word: 러시아 석유>

<Key Word: 러시아 제재>

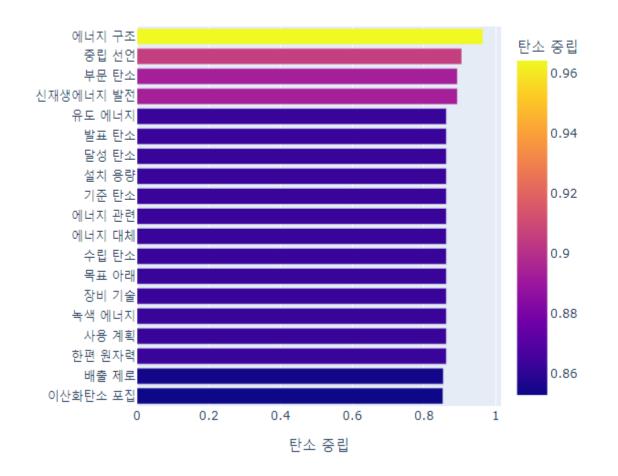
러시아 석유 연관단어 top 20

러시아 제재 연관단어 top 20



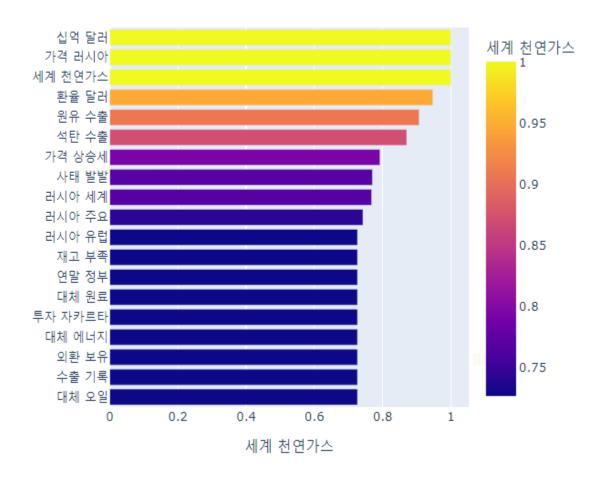
<Key Word: 탄소 중립>

탄소 중립 연관단어 top 20



<Key Word: 세계 천연가스>

세계 천연가스 연관단어 top 20



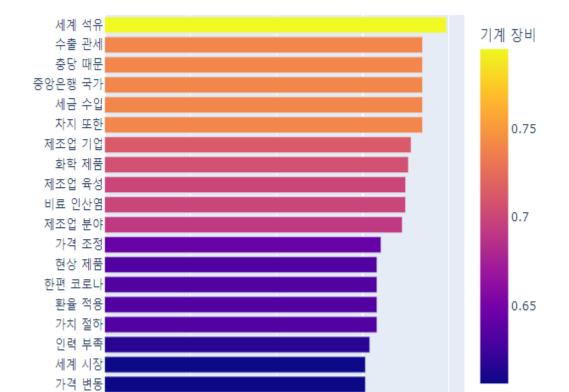
<Key Word: 기계 장비>

0.2

0

<Key Word: 중고 기계>

기계 장비 연관단어 top 20



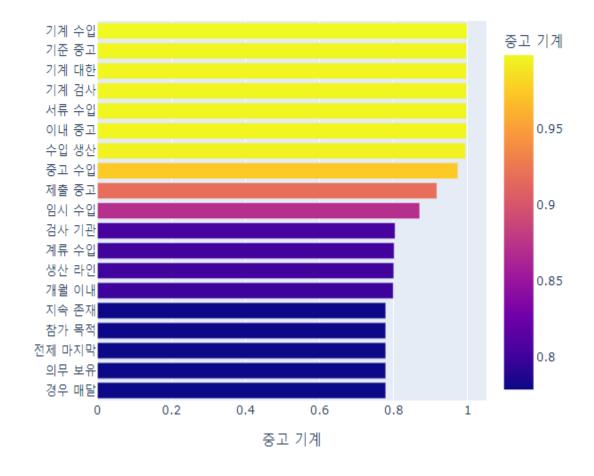
0.4

기계 장비

0.6

0.8

중고 기계 연관단어 top 20



<결론>

- 1. 데이터 원천의 중요성
 - > 텍스트 데이터 원천에 따라 분석결과가 크게 달라짐
 - > 필요 데이터수: 최소 50개 이상의 데이터 필요
 - * 시계열 분석의 경우 단위 기간별 최소 30~50개 이상 必

2. 활용방안

- 1) 유니그램(1단어) 보다 바이그램(2단어) 분석이 효과적
 - * 단, 바이그램 분석을 위해서 많은 컴퓨팅 자원 필요
- 2) 유니그램의 경우 트렌트 파악을 위해 많은 배경지식 필요
- 3) 사이트 크롤링을 위해서 체계적인 데이터 소스 관리 필요

감사합니다.