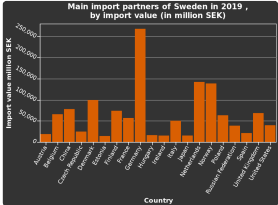


Chih-Hui Ho Giovanna Marinho Felipe Viana
Varad Pimpalkhute* Rodolfo Tonoli Andre Von Zuben
 Articul8.ai

Question in five languages

Scientific image



- English: What was the main import partner of Sweden in 2019?
- Hindi: 2019 में स्वीडन का मुख्य आयात साझेदार कौन था?
- Japanese: 2019年にスウェーデンの主要な輸入パートナーは誰でしたか？
- Korean: 2019년 스웨덴의 주요 수입 파트너는 누구였습니까?
- Portuguese: Quem foi o principal parceiro de importação da Suécia em 2019?

Table image

Top ten keyword searches with highest search:study ratio from ICPSR's website.

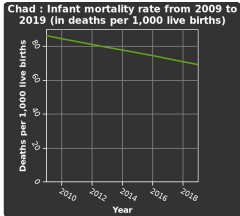
Search Phrase	# Exact Phrases	# Searches Containing Phrase	# ICPSR Studies	Search:Study Ratio	% Search Exa-5	% Search Exa-5
social media	336	812	20	40.6	26.19	26.6
NCAA	136	323	11	29.4	35.29	28.17
LGBT	216	658	25	26.3	24.07	22.8
justice	114	156	7	22.3	43.86	39.1
2012 election	118	396	18	22	17.8	15.4
trafficking	362	585	23	20.2	30.39	30.5
generation	255	291	15	19.4	76.51	44.75
immigrant	147	481	10	12.9	16.37	11.96
assault and risk	88	149	14	10.6	36.36	27.52
demoralization	523	323	30	10.4	96.59	95.59

- English: What is the search:study ratio for the keyword search "demoralization"?
- Hindi: दमन शोध की शोध:शोध दर क्या है शब्द शोध \दमन\ के लिए?
- Japanese: キーワード 検索 "デモラライゼーション" の検索 : 学習比は何ですか？
- Korean: 다음은 검색어 \\'더모랄리제이션\'에 대한 검색:학습 비율을 말하고 있습니다.
- Portuguese: Qual é a razão de pesquisa: estudo para a pesquisa de palavras-chave "demoralização"?

Figure 2: Additional question-image pairs in the proposed DocQIR dataset. Top: Example from the scientific image subset. Bottom: Example from the table image subset.

Question in five languages

Scientific image



- English: How many deaths per 1,000 live births were there in 2009?
- Hindi: 2009 में परत हिजार जीवित प्रवाह में कतिने मृत्यु हो गए थे?
- Japanese: 2009年の1,000人の新生児の死亡率は何でしたか？
- Korean: 2009년에 1,000명의 생존아기당 사망자 수는 몇명인가요?
- Portuguese: Quantas mortes por cada 1.000 nascimentos vivos houve em 2009?

Table image

Table 1: Comparison of Baseline Demographic and Characteristics Between African American and Caucasian Subjects

Characteristic	African American (N=1,000)	Caucasian (N=1,000)	p-value
Age (years)	38.0 (SD=10.0)	38.0 (SD=10.0)	1.000
Sex (male)	500 (50.0%)	500 (50.0%)	1.000
Body Mass Index (BMI)	26.5 (SD=4.5)	26.5 (SD=4.5)	1.000
Education (years)	12.5 (SD=2.0)	12.5 (SD=2.0)	1.000
Income (USD)	15,000 (SD=10,000)	15,000 (SD=10,000)	1.000
Marital Status (married)	500 (50.0%)	500 (50.0%)	1.000
Health Status (healthy)	500 (50.0%)	500 (50.0%)	1.000
Smoking Status (smoker)	250 (25.0%)	250 (25.0%)	1.000
Alcohol Consumption (alcohol)	250 (25.0%)	250 (25.0%)	1.000
Exercise Frequency (exercise)	250 (25.0%)	250 (25.0%)	1.000
Stress Level (stress)	250 (25.0%)	250 (25.0%)	1.000
Sleep Quality (sleep)	250 (25.0%)	250 (25.0%)	1.000
Overall Health (health)	250 (25.0%)	250 (25.0%)	1.000

- English: What is the difference in the percentage of African American and Caucasian subjects who are male?
- Hindi: अफ्रीकन अमेरिकन और कैस्काइन वषियों के परतशित का अंतर क्या है?
- Japanese: アフリカ系アメリカ人とカソニック系アメリカ人の男 性の割合の違いは何ですか？
- Korean: 아프리카계 미국인과 백인의 성별에 따른 비율의 차이가 무엇인가요?
- Portuguese: Qual é a diferença no percentual de sujeitos africanos americanos e caucasianos que são do sexo masculino?

Figure 3: Additional question-image pairs in the proposed DocQIR dataset. Top: Example from the scientific image subset. Bottom: Example from the table image subset.

Question in five languages

Scientific image

The world's leading container ship operators as of October 12, 2020, based on TEU capacity of own ships

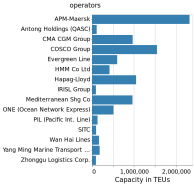


Table image

Table 3. Adjusted model of delay discounting rate including cognitive measures			
	β	T-value	p
BRAD	0.956	1.315	0.191
MDD	-0.192	-0.387	0.699
SCZSCAD	1.979	2.710	0.008*
Age	0.006	0.496	0.621
Sex	0.090	0.259	0.796
Antipsychotic Tx	-1.263	-1.997	0.048*
Mood stabilizer Tx	0.346	0.695	0.488
Antidepressant Tx	0.170	0.393	0.695
Smoking status	-0.728	-1.282	0.202
SWM	0.029	2.976	0.004*
AST	-0.004	-0.221	0.826
PAL	0.001	0.166	0.869

English: What is the capacity of APM-Maersk in TEUs?

Hindi: क्या APM-Maersk की टीईयू इंडरिफि क्पमता है?

Japanese: APM-MaerskのTEUsの容量は何ですか？

Korean: APM-Maersk의 TEUs 용량은 몇 개인가요?

Portuguese: Qual é a capacidade do APM-Maersk em TEUs?

English: What is the adjusted model of delay discounting rate including cognitive measures?

Hindi: नरियात मॉडल के लरि क्या है दलिया डसिकाउंरि रेट का अनुसंधान श्रेणी के साथ?

Japanese: 調整された遅延割引率のモデルは何か、認知尺度を含むですか？

Korean: 지연 감소를에 대한 조정 모델은 어떤 것인가요?

Portuguese: Qual é o modelo ajustado de taxa de desconto de atraso incluindo medidas cognitivas?

Figure 4: Additional question-image pairs in the proposed DocQIR dataset. Top: Example from the scientific image subset. Bottom: Example from the table image subset.