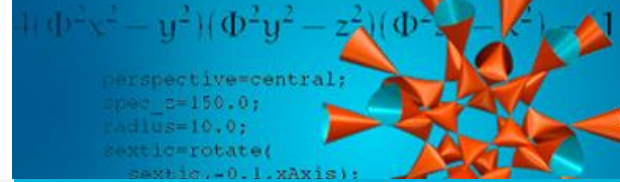


Dataintegration Project

Niewiedzący



Integration Pipeline :

- Problem : Country Entity fehlt uns.
- Lösung : Aus den anderen Datensätzen einzigartige Ländernamen extrahieren.
- Problem : Die Arbeitslosenraten waren auf verschiedene Spalten aufgeteilt.
- Lösung : Datensatz manipulieren.

- Problem : Country codes waren nicht in allen Datensätzen vorhanden.
- Lösung : Library country code verwenden.
- Verwendete Algorithmen : Apriori Approach für ucc, naive Kalkulation für fds und Validation mit plis Intersection.
- String similarities haben uns viel geholfen, um Ähnlichkeiten zu finden.

	Country	Country_code
70	EGYPT, ARAB REP.	EGY
71	EL SALVADOR	SLV
72	EQUATORIAL GUINEA	GNQ
73	ERITREA	ERI
74	ESTONIA	EST
75	ESWATINI	SWZ
76	ETHIOPIA	ETH
77	EURO AREA	NA
78	EUROPE & CENTRAL ASIA	NA
79	EUROPE & CENTRAL ASIA (EXCLUDING HIGH INCOME)	NA
80	EUROPE & CENTRAL ASIA (IDA & IBRD COUNTRIES)	NA
81	EUROPE AND CENTRAL ASIA	NA

	Country	Country_code	year	under_30
70	India (Urban)		2015	99.319190
71	Indonesia	IDN	2015	98.820730
72	Indonesia (Rural)		2015	99.616060
73	Indonesia (Urban)		2015	98.124250
74	Iran	IRN	2015	91.199320
75	Iraq	IRQ	2015	99.782860
76	Ireland	IRL	2015	30.852390
77	Israel	ISR	2015	52.350380
78	Italy	ITA	2015	40.188130
79	Jamaica	JAM	2015	94.090200
80	Japan	JPN	2015	30.482990
81	Jordan	JOR	2015	96.134430
82	Kazakhstan	KAZ	2015	98.128240

year	Country	Happiness_Rank	Happiness_Score	GDP	HLExpectancy	Freedom	Generosity	Corruption	Country_code
2015	Sweden	8	7.364	1.33171	0.91087	0.65980	0.36262	0.43844	SWE
2015	New Zealand	9	7.286	1.25018	0.90837	0.63938	0.47501	0.42922	NZL
2015	Australia	10	7.284	1.33358	0.93156	0.65124	0.43562	0.35637	AUS
2015	Israel	11	7.278	1.22857	0.91387	0.41319	0.33172	0.07785	ISR
2015	Costa Rica	12	7.226	0.95578	0.86027	0.63376	0.25497	0.10583	CRI
2015	Austria	13	7.200	1.33723	0.89042	0.62433	0.33088	0.18676	AUT
2015	Mexico	14	7.187	1.02054	0.81444	0.48181	0.14074	0.21312	MEX

Country	Country_code	Unemployment_rate	year
Africa Eastern and Southern	AFE	6.49	2015
Afghanistan	AFG	11.13	2015
Africa Western and Central	AFW	4.63	2015
Angola	AGO	7.39	2015
Albania	ALB	17.19	2015
Arab World	ARB	10.97	2015
United Arab Emirates	ARE	1.77	2015
Argentina	ARG	7.52	2015
Armenia	ARM	18.26	2015
Australia	AUS	6.05	2015

ER-Diagram :

