

과제 설명서

(2021-2학기 데이터 분석 및 시각화 프로젝트 “현금 없는 사회”)

산업시스템공학과 2019112470 김민서, 2019112466 문지윤, 2019112476 하지수

1. 주제

전자금융의 발달로 전자지급수단의 이용비중이 크게 증가하면서 “현금 없는 사회(Cashless Society)”로의 진입이 전세계적으로 진행되고 있다. 현금 없는 사회란 동전과 지폐를 사용하지 않고 신용카드, 모바일 지갑 등 비현금 지급수단을 주로 사용하는 사회를 의미한다.

그에 따라 “현금 없는 사회”라는 주제에 대해 각 나라별 특성을 분석 후, 우리나라에 반영하여 인사이트를 도출하고자 한다.

2. EDA 목표

먼저 실제 데이터와 뉴스 기사를 참고하여 현금 없는 사회 발달에 더딘 나라들과 빠른 나라들을 아래와 같이 선별하였다.¹

- 그룹 1. 현금 없는 사회 발달이 더딘 나라 : 나이지리아, 타이, 일본, 사우디아라비아, 베트남, 필리핀

- 그룹 2. 현금 없는 사회 발달이 빠른 나라 : 스웨덴, 영국, 네덜란드, 미국, 노르웨이, 덴마크

총 12개 국가와 우리나라의 데이터를 통해 분석할 목표와 세부 질문을 아래와 같이 선정하였다.

1) 경제 지표를 분석하여 현금 없는 사회를 나타낼 수 있는 대체 대표 변수를 탐색

=> 현금 없는 사회에서 계좌 보유 비중이 높은가?

=> 현금 없는 사회에서 ATM 기기의 수가 점점 줄어든 것인가?

=> 현금 없는 사회에서 상업 은행 수가 점점 줄어든 것인가?

2) 경제 외 지표를 분석하여 현금 없는 사회에서 나타나는 경제 외 요인을 탐색

=> 현금 없는 사회에서 65세 이상 인구 수 비율이 높을 것인가?

=> 현금 없는 사회에서 개인 인터넷 사용률이 높을 것인가?

=> 현금 없는 사회에서 Co2 배출량이 많을 것인가?

=> 현금 없는 사회에서 재생 에너지 사용률이 높을 것인가?

¹ 참고 : Share of cash in offline payment transactions at points-of-sale - such as retail or restaurants in 41 countries worldwide in 2020 / 한국은행, 최근 현금 없는 사회 전환 국가들의 주요 이슈와 시사점_최종 (보도자료, 2020)

3) 1과 2의 분석 결과에 우리나라의 특징을 결합시켜 인사이트 도출

=> 우리나라는 현금 없는 사회가 발달한 나라인가, 덜 발달한 나라인가?

3. 데이터 설명

1) 사용 데이터셋

① 경제 지표

- Account ownership at a financial institution or with a mobile-money-service provider (% of population ages 15+) : 15세 이상 계좌를 가지거나 모바일 서비스를 이용하는 비율
- GDP per capita (current US\$) : US 달러로 나타낸 1인당 GDP 데이터
- Automated teller machines (ATMs) (per 100,000 adults) : 성인 100,000명 당 자동화기기 (ATM)수
- Commercial bank branches (per 100,000 adults) : 성인 100,000명 당 상업 은행 수

② 경제 외 지표

- Population ages 65 and above (% of total population) : 65세 이상 인구 수 비율
- Individuals using the Internet (% of population) : 개인 인터넷 사용 비율
- CO2 emissions (metric tons per capita) : 1인당 이산화탄소 배출량
- Renewable energy consumption (% of total final energy consumption) : 총 에너지 소비량 대비 재생 에너지 소비량 비율

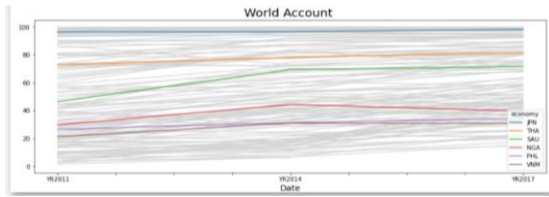
2) 데이터 전처리

- 몇몇 변수들에서 일정 년도 이후의 데이터만 존재하고, 각 나라별로 데이터셋 중간에 NaN 값이 존재하는 것을 확인하였다. 따라서 전체 행이 결측치인 경우에만 해당 행을 삭제하였다. 그 후 그룹별, 변수별로 데이터프레임을 생성하여 사용하였다.

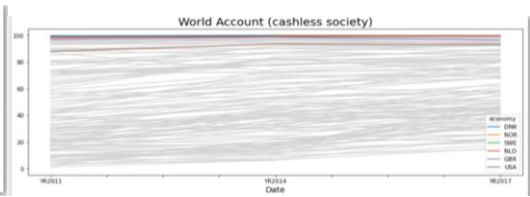
4. EDA 결과

1) 경제 지표 분석

- 은행 계정 소유권 (Account) : 그룹 1보다 2의 계정 소유가 더 높은 것을 알 수 있다. 해당 변수는 경제 성장률, 은행 수 등 다른 높은 상관관계를 띄는 요인이 존재하거나 은행 계정을 만드는 목적이 다양할 것이라 생각되어 적절한 대체 변수로 보기는 어렵다고 판단하였다.

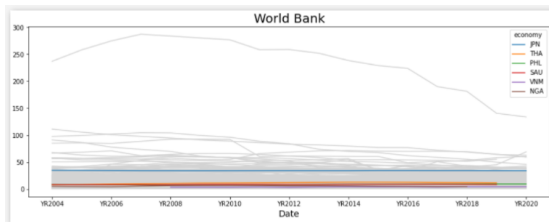


(좌) 그룹 1 = 현금 없는 사회 덜 발달

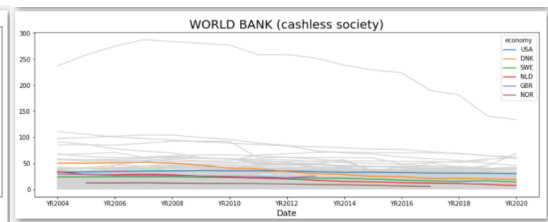


(우) 그룹 2 = 현금 없는 사회 발달

- 상업 은행 수 (Bank) : 그룹 1은 증가, 그룹 2는 감소하는 추세이지만 결과가 미미하여 추가 시각화를 진행한 결과 그룹 1은 증가, 그룹 2는 감소하는 형태를 보였다. 현금 없는 사회에서 은행 지점 수가 줄어든 것이라는 가정에 맞았고, 대체 변수로서 연관이 있다고 판단하였다.



(좌) 그룹 1 = 현금 없는 사회 덜 발달



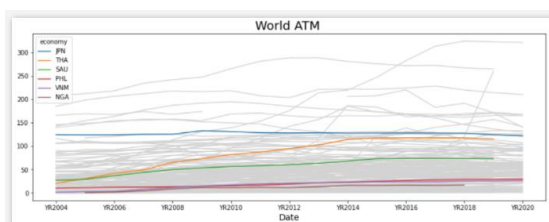
(우) 그룹 2 = 현금 없는 사회 발달

economy	NSA	THA	JPN	SAU	VNM	PHL	GBR	GBR	NLD	USA	NOR	DNK
YR2004	4.700000	7.750000	34.610000	8.130000	nan	8.210000	23.570000	29.030000	33.610000	32.690000	nan	50.080000
YR2005	4.180000	8.270000	34.880000	7.860000	nan	8.000000	23.520000	28.820000	28.070000	33.280000	12.080000	50.200000
YR2006	3.780000	9.080000	34.120000	7.950000	nan	7.800000	23.780000	26.430000	27.700000	33.940000	12.010000	50.520000
YR2007	5.210000	8.690000	33.900000	8.010000	nan	7.530000	26.340000	28.460000	24.770000	34.770000	12.310000	51.840000
YR2008	6.270000	10.310000	33.650000	8.020000	3.260000	7.570000	23.890000	26.030000	27.570000	35.160000	11.700000	50.450000
YR2009	6.480000	10.800000	33.810000	8.300000	3.280000	7.490000	22.960000	25.410000	25.230000	35.910000	11.020000	48.800000
YR2010	6.560000	11.000000	33.820000	8.360000	3.210000	7.510000	22.500000	24.700000	23.020000	35.420000	10.960000	40.860000
YR2011	6.410000	11.340000	33.900000	8.250000	3.540000	7.640000	21.700000	24.060000	21.310000	35.150000	10.730000	38.730000
YR2012	5.820000	11.680000	33.950000	8.170000	3.100000	7.930000	21.810000	22.080000	19.650000	34.900000	9.930000	34.080000
YR2013	5.900000	12.090000	33.100000	8.170000	3.240000	8.240000	21.840000	25.140000	17.300000	33.400000	8.940000	33.010000
YR2014	5.630000	12.520000	33.600000	8.510000	3.830000	8.620000	21.330000	25.130000	14.790000	32.420000	8.510000	27.790000
YR2015	4.980000	12.530000	34.140000	8.550000	3.750000	8.760000	19.320000	23.000000	13.990000	32.700000	7.730000	24.690000
YR2016	4.740000	12.350000	34.100000	8.530000	3.830000	8.840000	17.840000	21.460000	12.460000	32.100000	6.160000	24.050000
YR2017	4.440000	11.860000	34.000000	8.500000	3.450000	8.990000	16.170000	21.920000	11.920000	31.210000	5.500000	20.710000
YR2018	4.300000	11.700000	34.010000	8.360000	3.920000	9.020000	14.910000	21.000000	10.940000	30.880000	nan	20.850000
YR2019	nan	11.240000	33.920000	8.160000	3.980000	9.200000	16.420000	20.000000	9.240000	30.460000	nan	20.000000
YR2020	nan	nan	33.920000	nan	4.010000	9.170000	13.820000	nan	6.990000	29.690000	nan	18.660000

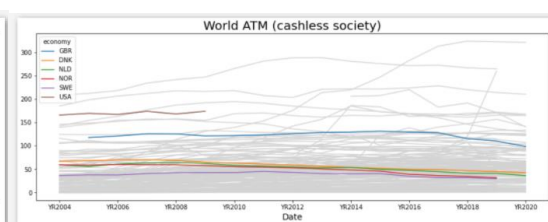
(좌) 그룹 1 = 현금 없는 사회 덜 발달

(우) 그룹 2 = 현금 없는 사회 발달

- ATM 수 : 일본을 제외한 그룹 1은 증가, 그룹 2는 감소하는 추세를 보였지만 더 자세한 판단을 위해 추가 시각화를 진행하였다. 그룹 1은 증가, 그룹 2는 감소하는 형태를 보여 ATM 기기의 수가 줄어든 것이라는 가정을 확인하였고 대체 변수로서 연관이 있다고 판단하였다.



(좌) 그룹 1 = 현금 없는 사회 덜 발달



(우) 그룹 2 = 현금 없는 사회 발달

economy	NSA	THA	JPN	GBR	VNM	PHL	GBR	GBR	NLD	USA	NOR	DNK
YR2004	nan	20.855230	124.869047	27.168022	1.409589	10.311727	46.469533	117.048614	59.270231	165.783891	59.138061	27.023411
YR2005	0.680126	30.653426	122.900401	29.132537	2.912677	11.432779	37.731229	117.048614	59.270231	165.783891	59.138061	27.023411
YR2006	1.790070	42.198578	124.118344	37.028182	3.487822	12.314818	40.791878	117.048614	59.270231	165.783891	59.138061	27.023411
YR2007	4.475741	49.237019	125.130850	44.087129	7.572170	12.925114	40.487856	117.048614	59.270231	165.783891	59.138061	27.023411
YR2008	8.641506	35.340906	125.445950	49.851293	11.825384	13.172395	42.108873	117.048614	59.270231	165.783891	59.138061	27.023411
YR2009	11.493693	73.356627	132.807932	52.688766	14.725051	14.008426	42.795309	117.048614	59.270231	165.783891	59.138061	27.023411
YR2010	11.258850	81.936410	130.022266	55.683388	17.175098	15.106225	42.795309	117.048614	59.270231	165.783891	59.138061	27.023411
YR2011	11.983614	90.912884	128.792117	58.982119	19.483448	16.804617	42.202288	117.048614	59.270231	165.783891	59.138061	27.023411
YR2012	11.939028	94.788477	127.911026	60.423021	20.671119	16.851103	43.042637	117.048614	59.270231	165.783891	59.138061	27.023411
YR2013	13.311999	102.350231	128.324077	63.822811	21.887815	17.925518	40.545933	117.048614	59.270231	165.783891	59.138061	27.023411
YR2014	16.183267	114.213988	127.818367	68.199297	22.686490	23.343138	40.186892	117.048614	59.270231	165.783891	59.138061	27.023411
YR2015	16.211488	117.231610	127.648821	70.700488	23.742326	25.040999	40.545933	117.048614	59.270231	165.783891	59.138061	27.023411
YR2016	16.758416	117.624930	127.765540	74.030365	24.243536	27.041891	34.766163	117.048614	59.270231	165.783891	59.138061	27.023411
YR2017	16.328233	117.791645	127.868080	74.108822	24.144839	28.135137	31.995328	117.048614	59.270231	165.783891	59.138061	27.023411
YR2018	16.930262	117.554120	127.883708	75.999132	25.321024	28.888888	31.861754	117.048614	59.270231	165.783891	59.138061	27.023411
YR2019	nan	118.088888	124.118344	78.888888	25.904020	28.970231	29.620241	117.048614	59.270231	165.783891	59.138061	27.023411
YR2020	nan	nan	121.705089	nan	26.264738	29.718833	nan	nan	36.627483	36.027683	nan	42.564959

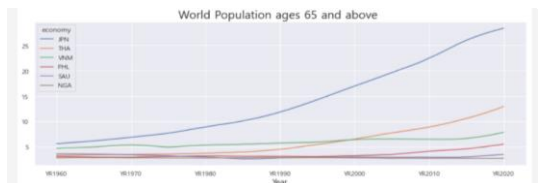
(좌) 그룹 1 = 현금 없는 사회 덜 발달

(우) 그룹 2 = 현금 없는 사회 발달

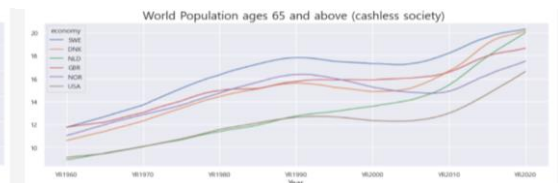
=> ATM, 상업 은행 수는 현금 없는 사회의 발달 정도와 비례하며, 대표 변수로 볼 수 있다.
은행 계정 소유권은 GDP와 관련이 있으며 GDP가 높을수록 은행 계정 수 비율이 높다.

2) 경제 외 지표 분석

- 65세 이상 인구 수 : 일본을 제외한 그룹 1의 65세 이상 인구 수가 그룹 2보다 작게 나타났다. 일본의 결과에는 다른 요인이 작용할 것이라고 생각하며, 현금 없는 사회가 발달된 나라일수록 65세 이상 인구 수가 많을 것이라는 가정을 확인하였다.

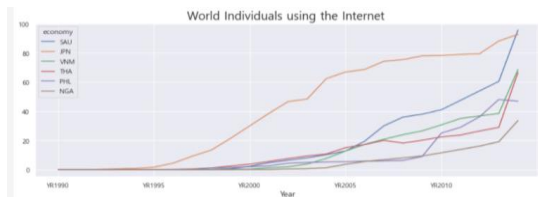


좌) 그룹 1 = 현금 없는 사회 덜 발달

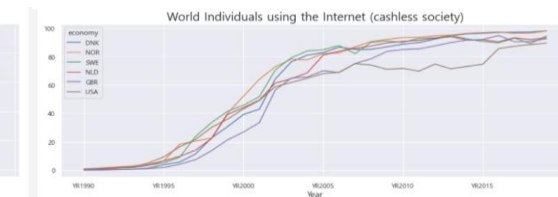


(우) 그룹 2 = 현금 없는 사회 발달

- 개인 인터넷 사용률 : 그룹 1과 2 모두 인터넷 사용률이 증가하고 있는 추세이지만 그룹 2에서 더 높은 것을 알 수 있다. 현금 없는 사회가 발달된 나라일수록 개인 인터넷 사용률이 많을 것이라는 가정을 확인하였다.

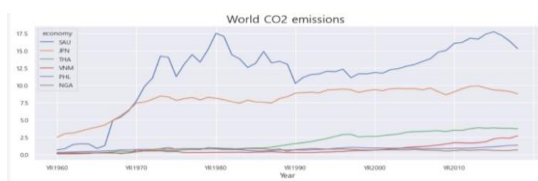


좌) 그룹 1 = 현금 없는 사회 덜 발달

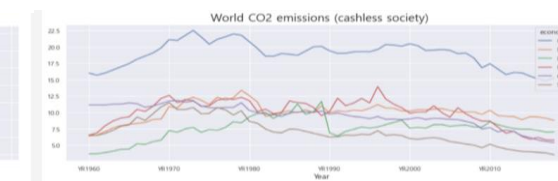


(우) 그룹 2 = 현금 없는 사회 발달

- CO2 배출량 : 그룹 1보다 그룹 2의 CO2 배출량이 많은 것을 확인 할 수 있지만 뚜렷한 차이가 보이지 않아 현금 없는 사회와 연관된 변수로 사용할 수 없다고 판단하였다.

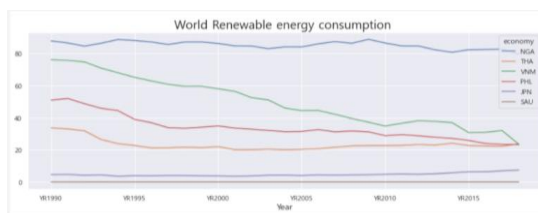


좌) 그룹 1 = 현금 없는 사회 덜 발달

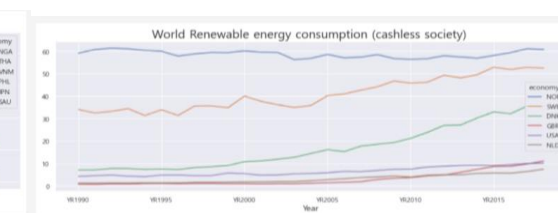


(우) 그룹 2 = 현금 없는 사회 발달

- 재생 에너지 소비량 : 그룹별 차이가 없어 관련 변수로 사용하기 어렵다고 판단하였다.



좌) 그룹 1 = 현금 없는 사회 덜 발달

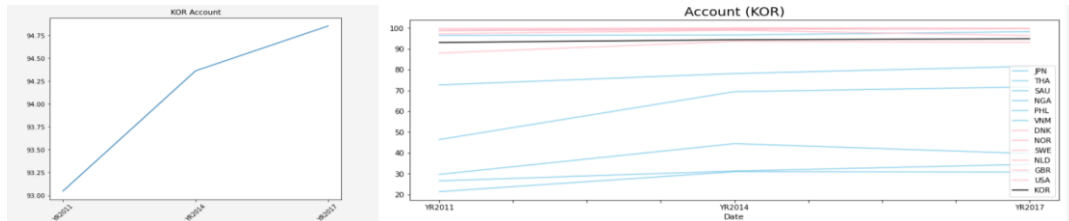


(우) 그룹 2 = 현금 없는 사회 발달

=> 65세 이상 인구 수와 개인 인터넷 사용률은 현금 없는 사회가 발달될수록 더 높은 값을 가졌으며, 현금 없는 사회와 관련된 변수로 사용할 수 있다. CO2 배출량과 재생 에너지 소비량은 현금 없는 사회로의 발달 정도와 무관한 값을 보였다.

3) 우리나라 특징 분석

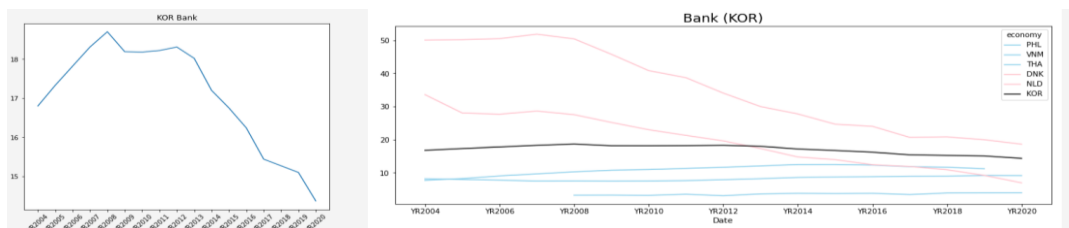
- 은행 계정 소유권 (Account) : 증가하는 추세이며 그룹 2와 비슷하게 높은 편에 속한다.



(좌) 우리나라 은행 계정 소유권

(우) 그룹 1(파랑) = 현금 없는 사회 덜 발달 / 그룹 2(분홍) = 현금 없는 사회 발달

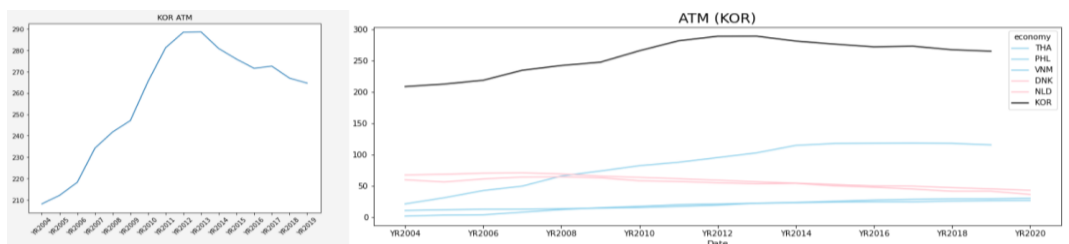
- 상업 은행 수 (Bank) : 상업 은행 수가 감소하는 그룹 2와 비슷한 양상을 보인다.



(좌) 우리나라 상업 은행 수

(우) 그룹 1(파랑) = 현금 없는 사회 덜 발달 / 그룹 2(분홍) = 현금 없는 사회 발달

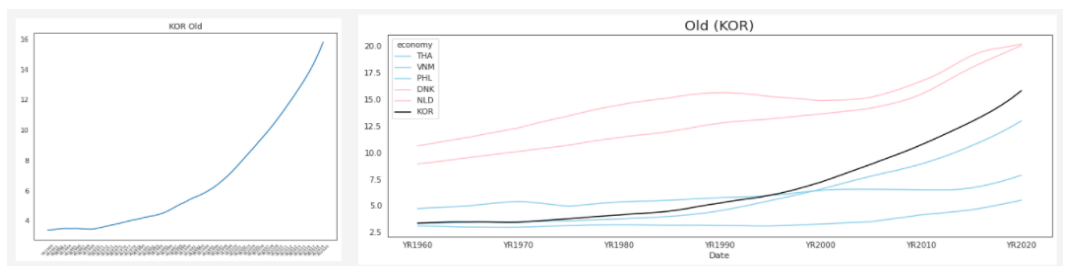
- ATM 수 : ATM 수가 감소하는 그룹 2와 비슷한 양상을 보인다.



(좌) 우리나라 ATM 수

(우) 그룹 1(파랑) = 현금 없는 사회 덜 발달 / 그룹 2(분홍) = 현금 없는 사회 발달

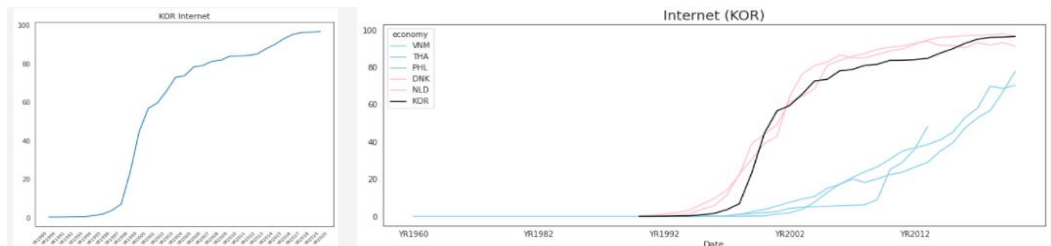
- 65세 이상 인구 수 : 그룹 1과 그룹 2의 중간 정도의 양상을 보인다.



(좌) 우리나라 65세 이상 인구 수

(우) 그룹 1(파랑) = 현금 없는 사회 덜 발달 / 그룹 2(분홍) = 현금 없는 사회 발달

- 개인 인터넷 사용률 : 개인 인터넷 사용률이 더 높은 그룹 2와 비슷한 양상을 보인다.



(좌) 우리나라 개인 인터넷 사용률

(우) 그룹 1(파랑) = 현금 없는 사회 덜 발달 / 그룹 2(분홍) = 현금 없는 사회 발달

=> 우리나라의 은행 계정 소유권은 비교 국가들 중 높은 편에, GDP는 중간 정도에 속한다. 경제 지표 분석에서 ATM과 상업은행 수가 현금 없는 사회가 발달한 나라들과 비슷한 양상을 띄었다. 경제 외 지표 분석에서 65세 이상 인구 수는 중간 정도의 양상을 띄며, 개인 인터넷 사용률은 현금 없는 사회가 발달한 나라들과 비슷한 양상을 띈다.

5. 결론 및 추가사항

EDA 분석을 통한 경제, 경제 외 지표와의 비교분석 결과 우리나라는 현금 없는 사회가 발달한 나라로 판단할 수 있다.

앞선 세부 질문들에 대한 결론을 아래와 같이 정리하였다.

1) 경제 요인

- 현금 없는 사회에서 계좌 보유 비중이 높은가? => 상관 없는 변수이다.
- 현금 없는 사회에서 ATM 기기의 수가 점점 줄어든 것인가? => 줄어든다.
- 현금 없는 사회에서 상업 은행 수가 점점 줄어든 것인가? => 줄어든다.

2) 경제 외 요인

- 현금 없는 사회에서 65세 이상 인구 수 비율이 높을 것인가? => 높다.
- 현금 없는 사회에서 개인 인터넷 사용률이 높을 것인가? => 높다.
- 현금 없는 사회에서 Co2 배출량이 많을 것인가? => 상관 없는 변수이다.
- 현금 없는 사회에서 재생 에너지 사용량이 높을 것인가? => 상관 없는 변수이다.

처음의 가정과 다르게 3개의 변수가 현금 없는 사회의 발달 정도와 관련 없는 변수로 나왔고 나머지 4개의 가정은 맞는 것을 확인할 수 있었다. 현금 없는 사회는 최근 떠오르는 이슈로 년 단위로 보기보다는 월 단위로 보는 것이 현재의 상황이나 변화를 알 수 있음에 바람직하지만 년 단위의 데이터라 원하는 흐름을 캐치하기 어려웠다는 한계점을 가지며 프로젝트를 마친다.