

## 목차

- 개요
- 코로나 전후의 소비 패턴 분석 배경
- 데이터 정의 및 처리
- 데이터 분석
- 결과 평가 및 해석
- 기대 효과

## 개요

- 코로나 시기를 거쳐 충청남도 시민들의 소비패턴을 확인하여 소득 대비 소비 사용량으로 경제 추이를 확인한다.

## 코로나 전후의 소비 패턴 분석 배경:

1. 팬데믹으로 인한 경제 불확실성
2. 디지털 전환 및 비대면 소비 증가
3. 재난 지원금의 경제 영향 확인
4. 산업 및 시장 변화 확인

## 데이터 정의 및 처리:

1. 데이터 정의: 데이터 연도별 월소득, 카드 사용 금액 평균, 텔레세일즈 사용 금액 평균, **cs**선수금 사용금액 평균 등 관심 데이터 정의.
2. 데이터 처리: **csv** 파일의 데이터 프레임 병합 및 결측치 처리 후 그래프로 출력하는 시각화 처리과정을 거침.
3. **CSV**파일 코드 설명

라이브러리 импорт 및 설정:

`pandas`, `matplotlib`, `seaborn`, `numpy` 등 필요한 라이브러리를 импорт합니다.

한글 폰트를 설정합니다. 운영 체제에 따라 적절한 폰트 경로를 지정합니다.

데이터 불러오기:

2018년, 2020년, 2023년의 데이터를 각각 `CSV` 파일에서 불러옵니다.

각 데이터프레임에 `'YEAR'` 열을 추가하여 해당 연도를 표시합니다.

데이터 병합:

`pd.concat()`을 사용하여 세 개의 데이터프레임을 하나로 병합합니다.

관심 변수 선택:

카드 사용액 열을 자동으로 찾습니다 (`'CARD'`와 `'GRAMT'`가 포함된 열 이름).

`'YEAR'`, `'MM_INCM'`, 카드 사용액, `'CR_NTSL_USE_GRAMT_AVG'`, `'CSADVC_USE_GRAMT_AVG'` 열을 선택합니다.

변수명 한글화:

선택된 변수들의 이름을 한글로 변경합니다.

데이터 분석 준비:

선택된 변수들만 포함하는 새로운 데이터프레임(`df_analysis`)을 생성합니다.

시각화:

a. 선 그래프:

연도별로 각 변수의 평균값 추이를 보여줍니다.

b. 막대 그래프:

연도별로 각 변수의 평균값을 비교합니다.

c. 박스 플롯:

이상치를 제외한 데이터의 분포를 연도별로 보여줍니다.

d. 히트맵:

변수 간의 상관관계를 색상으로 표현합니다.

e. 산점도:

월 소득과 카드 사용액의 관계를 연도별로 표시합니다.

f. 파이 차트:

각 연도별로 지출 항목의 비중을 원형 그래프로 보여줍니다.

데이터 처리 및 분석 과정:

`groupby()`를 사용하여 연도별 평균을 계산합니다.

`melt()` 함수를 사용하여 데이터를 재구조화합니다 (박스 플롯용).

`corr()` 메서드로 변수 간 상관관계를 계산합니다.

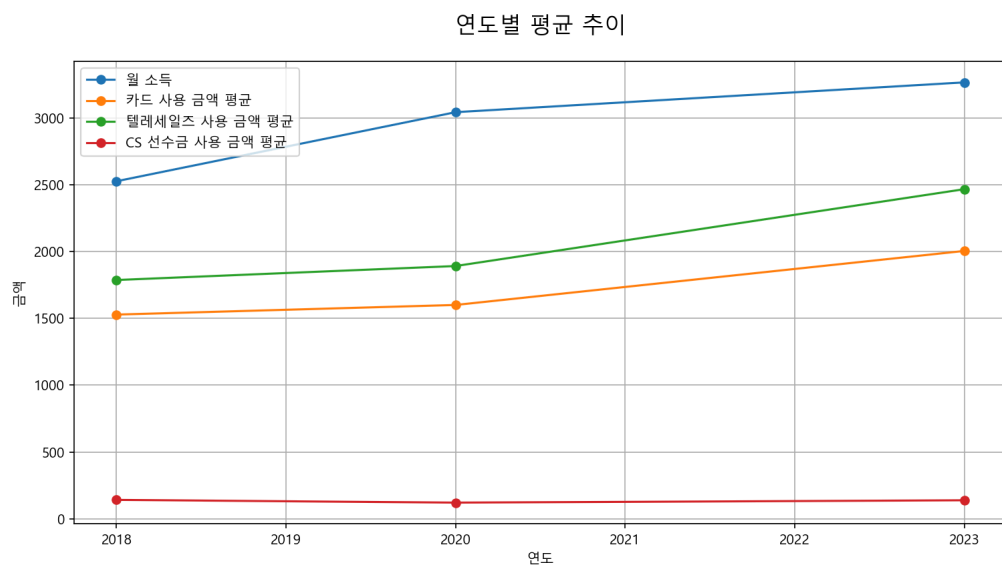
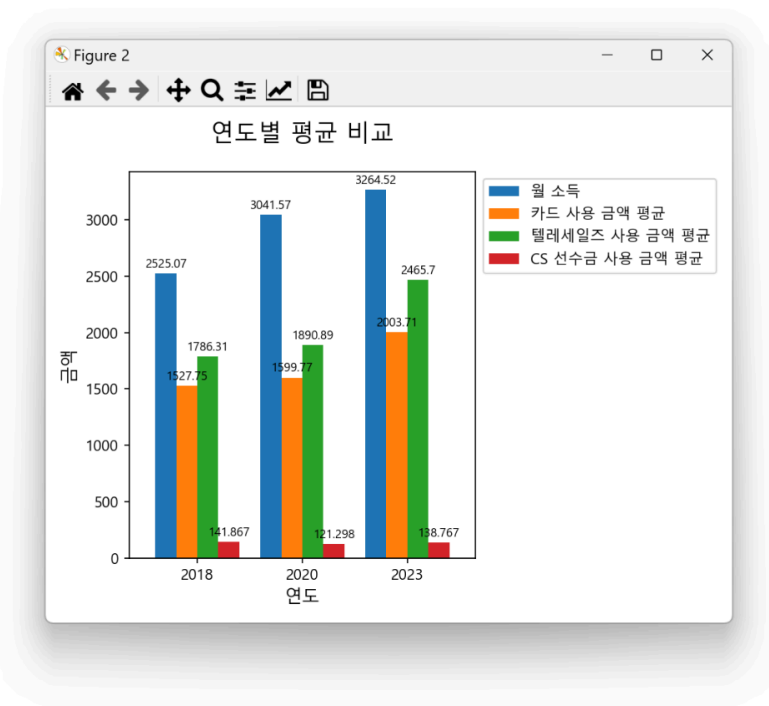
시각화 세부 설정:

그래프의 크기, 제목, 레이블, 범례 등을 세부적으로 조정합니다.

일부 그래프에서는 데이터 레이블을 추가하여 가독성을 높입니다.

## 데이터 분석:

그래프 결과(순서대로 평균 비교 막대그래프, 평균 추이 선그래프, 연도별 분포 박스 그래프, 연도와 소득 및 소비 히트맵, 월소득과 카드 사용액의 산점도, 연도별 지출 비중의 원그래프)



연도별 분포 비교 (이상치 제외)

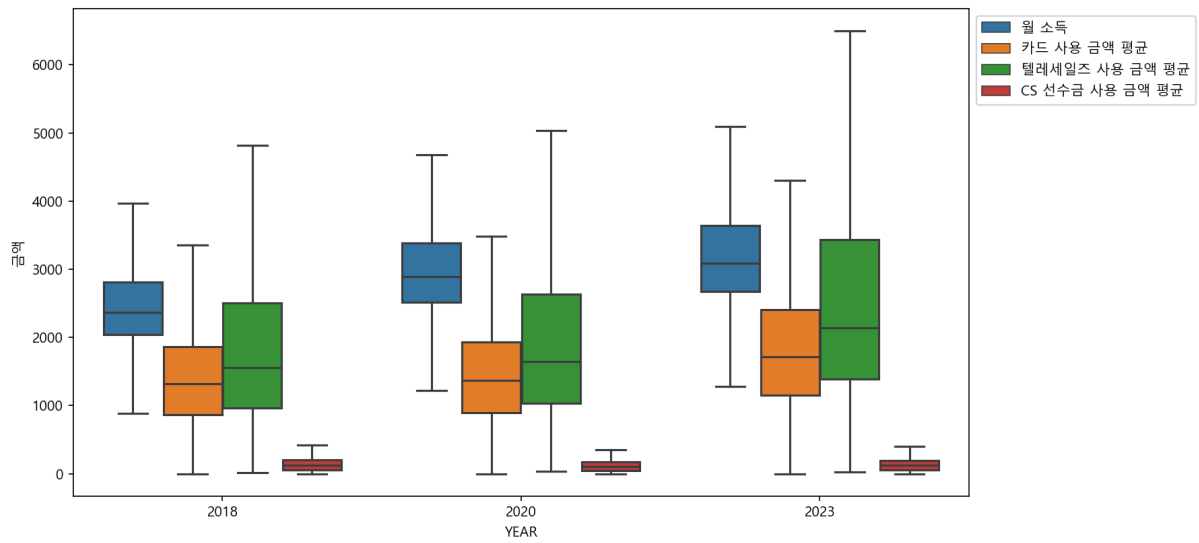
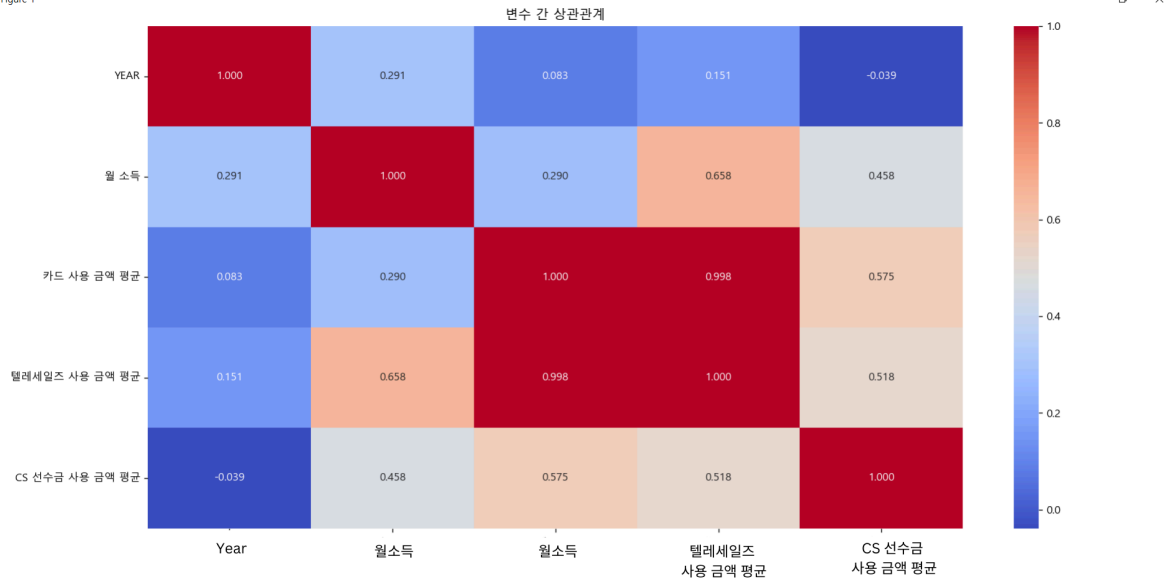
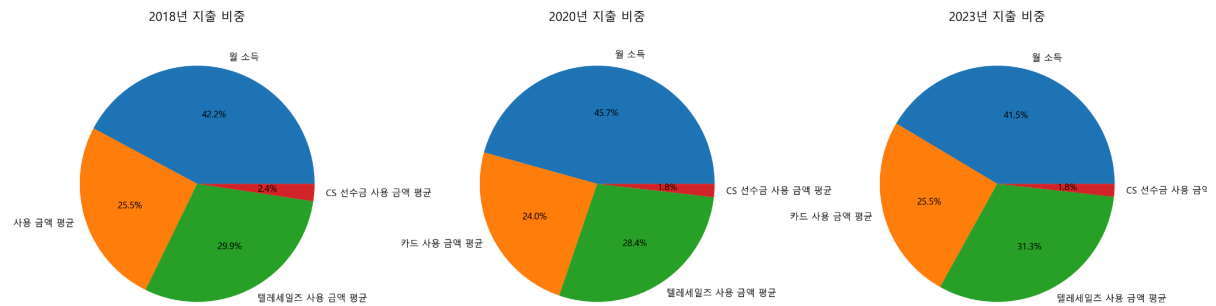
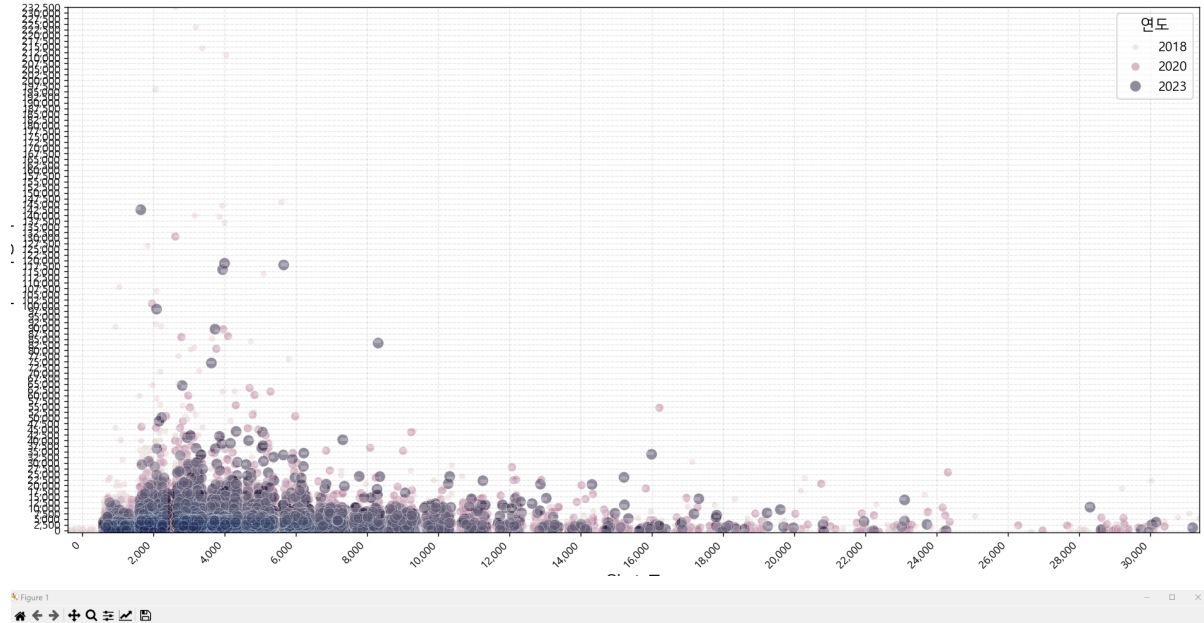


Figure 1





## 결과 평가 및 해석:

1. 연도별 평균 비교 막대그래프: 모든 막대그래프는 연도별로 비교하였을 때 시간에 따라 상승하는 형태를 가진다. 이를 통하여 **2018년도부터 2023년도**까지의 충청남도 경제는 소비와 소득이 상승하고 있는, 팬데믹으로 인한 경제 하락 사태를 보이지 않았다는 사실을 알 수 있다.
2. 연도별 평균 추이 선그래프: 월 소득은 **2018년에서 2020년도로 넘어가는 동안** 연평균 **50만원**의 소득이 증가하였지만, **2020년도 코로나 이후 2023년까지의**

소득 증가는 연평균 10만원 정도로 코로나로 인한 소득 증가량이 줄어들었다는 것을 확인할 수 있다. 또한 카드 및 텔레세일즈(통화 서비스 판매) 사용금액은 코로나를 기준으로 전의 연평균 증가량은 약 5만원이었지만, 코로나 이후의 연평균 증가량은 약 20만원으로 소비 증가량은 코로나로 인해 늘어났다는 것을 확인할 수 있다.

3. 연도별 분포 비교 박스 그래프(이상치 제외): 각 변수의 평균의 분포를 박스 플롯으로 나타낸것이다. 월 소득은 2018년도부터 2023년도 까지 비교적 안정적이며 큰 변동은 없었다. 카드 사용 금액 평균은 일정 수준을 유지하고있지만, 소폭의 감소가 관찰된다. 비교적으로, 텔레세일즈 사용 금액 평균은 변동이 크고 시간에 지남에 따라 증가하는 추세가 보인다. CS 선수금 사용 금액 평균은 매우 낮고 변동이 적으며 이 그래프를 통해 각 항목의 경향과 변이를 분석하여 경제 상황이나 소비 패턴의 변화를 확인할 수 있다.
4. 연도와 소비 및 소득의 히트맵: 우측으로 하강하는 1.0은 히트맵에서 동일항목을 나타내기 때문에 이를 제외하고 히트맵 출력값을 해석하였다. 히트맵에서 양수는 연관이 높다는 것을 나타내고, 음수는 이와 반대의 연관을 가지고 있다는 것을 나타낸다. 이를 통해  $\pm 0.5$ 를 기준으로, 연도에 따른 소득과 소비 관계는 거의 없으며, 각 소비에서의 연관성은 높다는 사실을 알 수 있다. 특히 카드 사용과 텔레세일즈 사용금액에서의 큰 연관성이 있다는 것을 확인할 수 있다.
5. 월소득과 카드 사용액 간 산점도: 세로축은 카드 사용액을, 가로축은 월 소득을 나타내며, 각 축의 단위는 숫자 뒷자리의 1000이 생략된 단위임을 기술한다. 2018년도의 경우, 2023년도에 비하여 월 소득이 20만원에서 50만원 이내임에도 150만원 이상의 소비 (이하 과소비)를 하는 경우가 여럿 있다는 것을 통하여 2020년도를 기준으로 경제적 위험에 대한 경각심이 증가하였다는 사실을 확인할 수 있다.

6. 연도별 지출 비중: 3년도 모두 소득보다 지출이 많은 양상으로, 충청남도의 경제활동이 활발한 편에 속한다는 것을 알 수 있다. 20년도의 경우 소비활동의 저하로 인하여 소득 수준이 가장 높게 나타났으며, CS 선수금 지원 금액의 경우(사전 서비스 지원금) 코로나를 기준으로 줄어들게 되어 서비스에 대한 지출이 줄어드는 사회양상을 확인할 수 있다. 또한 카드 사용 금액의 경우, 2020년도는 비대면으로 인한 경제활동의 저하현상으로 인해 단독으로 24%를 기록하였다는 것을 알 수 있다.

## 기대 효과

1. 정책 수립에 기여
  - 분석 결과를 바탕으로 정책 수립 및 충청남도의 경제 회복 전략을 마련하는 것을 기대할 수 있음.
2. 시장 예측력 향상
  - 소비 패턴 분석을 통해 시장 예측력을 향상하는 현상을 기대할 수 있음.
3. 비즈니스 전략 수립
  - 충남을 대상으로 하는 비즈니스 전략 수립 시 유용한 데이터가 될 수 있음.