



온라인 쇼핑몰 매출관계 분석

- 국가통계포털 <https://kosis.kr/index/index.do>
- 온라인쇼핑몰 판매매체별 / 상품군별거래액
- 합계등을 제외하고 선택(전기전자통신 2020년전후로 분리되어서 미리 합침)
- 행렬을 전환

 No description has been provided for this image

 No description has been provided for this image

```
In [1]: import matplotlib
import matplotlib.pyplot as plt
import matplotlib.font_manager as fm
# import matplotlib.dates as mdates

import pandas as pd
from datetime import datetime
```

```
In [2]: # font 설정 D2coding
d2_path = "./ref/D2Coding-Ver1.3.2-20180524.ttf"
fm.fontManager.addfont(d2_path)
matplotlib.rcParams["font.family"] = "D2Coding"
```

```
In [5]: file_path = "./ref/온라인쇼핑몰_판매매체별_상품군별거래액.csv"
ori = pd.read_csv(file_path, encoding="EUC-KR") # utf-8
ori.columns = ["category", "date", "internet", "mobile"]
ori.tail()
```

```
Out[5]:
```

	category	date	internet	mobile
2203	기타	2024.08	61304	120200
2204	기타	2024.09	63863	126583
2205	기타	2024.10	68403	128216
2206	기타	2024.11 p)	73246	135527
2207	기타	2024.12 p)	75613	144134

```
In [11]: print("수정전")
print(ori['date'].tail())
ori['date'] = ori['date'].map(lambda x: x.rstrip(' p'))
print("수정후")
print(ori['date'].tail())
```

수정전

```
2203      2024.08
2204      2024.09
2205      2024.10
2206      2024.11 p)
2207      2024.12 p)
Name: date, dtype: object
```

수정후

```
2203      2024.08
2204      2024.09
2205      2024.10
2206      2024.11
2207      2024.12
Name: date, dtype: object
```

```
In [22]: ori['date'] = pd.to_datetime(data["date"], format='%Y.%m')
ori.dtypes
ori.tail()
```

```
Out[22]:
```

	category	date	internet	mobile
2203	기타	2024-08-01	61304	120200
2204	기타	2024-09-01	63863	126583
2205	기타	2024-10-01	68403	128216
2206	기타	2024-11-01	73246	135527
2207	기타	2024-12-01	75613	144134

```
In [38]: ori[130:170]
```

Out[38]:

	category	date	internet	mobile
130	가전.전자.통신기기	2019-11-01	568012	789341
131	가전.전자.통신기기	2019-12-01	518944	694448
132	가전.전자.통신기기	2020-01-01	400922	662108
133	가전.전자.통신기기	2020-02-01	494994	755756
134	가전.전자.통신기기	2020-03-01	518520	780014
135	가전.전자.통신기기	2020-04-01	507840	738910
136	가전.전자.통신기기	2020-05-01	588332	927906
137	가전.전자.통신기기	2020-06-01	553380	905941
138	가전.전자.통신기기	2020-07-01	502863	816872
139	가전.전자.통신기기	2020-08-01	548239	987774
140	가전.전자.통신기기	2020-09-01	462687	849047
141	가전.전자.통신기기	2020-10-01	547418	987672
142	가전.전자.통신기기	2020-11-01	663471	1161320
143	가전.전자.통신기기	2020-12-01	611410	1105244
144	가전.전자.통신기기	2021-01-01	610054	1137541
145	가전.전자.통신기기	2021-02-01	512864	981223
146	가전.전자.통신기기	2021-03-01	612790	1052288
147	가전.전자.통신기기	2021-04-01	546939	976497
148	가전.전자.통신기기	2021-05-01	575594	1118449
149	가전.전자.통신기기	2021-06-01	563485	1072052
150	가전.전자.통신기기	2021-07-01	658089	1256457
151	가전.전자.통신기기	2021-08-01	576436	1088677
152	가전.전자.통신기기	2021-09-01	514785	1056165
153	가전.전자.통신기기	2021-10-01	661197	1363174
154	가전.전자.통신기기	2021-11-01	645257	1354614
155	가전.전자.통신기기	2021-12-01	637081	1122314
156	가전.전자.통신기기	2022-01-01	539706	1107412
157	가전.전자.통신기기	2022-02-01	599163	1202958
158	가전.전자.통신기기	2022-03-01	603523	1058191
159	가전.전자.통신기기	2022-04-01	556310	998166

	category	date	internet	mobile
160	가전.전자.통신기기	2022-05-01	579325	1137206
161	가전.전자.통신기기	2022-06-01	548771	1135013
162	가전.전자.통신기기	2022-07-01	567002	1207117
163	가전.전자.통신기기	2022-08-01	587503	1203957
164	가전.전자.통신기기	2022-09-01	459133	1160551
165	가전.전자.통신기기	2022-10-01	653866	1420255
166	가전.전자.통신기기	2022-11-01	587368	1275010
167	가전.전자.통신기기	2022-12-01	572088	1175938
168	가전.전자.통신기기	2023-01-01	506317	1166020
169	가전.전자.통신기기	2023-02-01	551798	1207891

```
In [64]: ex1_data = ori.copy()
# ex1_data[ex1_data['category']== '음식서비스' ]#
obj = ['컴퓨터 및 주변기기', '농축수산물', '의복', '신발', '가방', '음식서비스', '음·식료품',
ex1_data = ex1_data[ex1_data['category'].isin(obj)]
ex1_data.sample(5)
```

```
Out[64]:
```

	category	date	internet	mobile
622	가방	2020-11-01	49657	117055
542	신발	2022-03-01	98582	228639
78	컴퓨터 및 주변기기	2023-07-01	338837	337786
512	신발	2019-09-01	78129	122879
1723	여행 및 교통서비스	2024-08-01	705579	1626290

```
In [108... fig , ax = plt.subplots(1,2,figsize=(15,5))

category = obj # ['컴퓨터 및 주변기기', '농축수산물', '의복', '신발', '가방', '음식서비스',
for cate in category:
    category_data = ex1_data[ex1_data['category'] == cate]
    x_data = category_data['internet']
    y_data = category_data['mobile']
    ax[0].scatter( x_data , y_data, s = 20, label = cate )

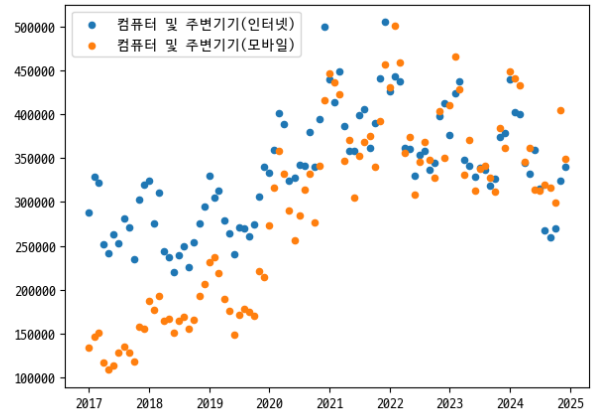
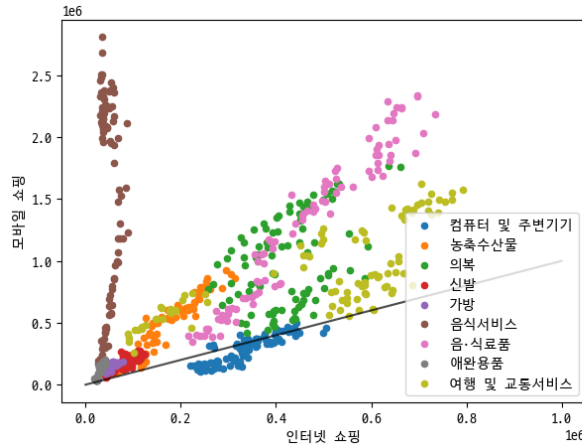
ax[0].set_xlabel("인터넷 쇼핑")
ax[0].set_ylabel("모바일 쇼핑")
ax[0].legend(loc="best")
ax[0].plot([0,1000000],[0,1000000], 'k-', alpha=0.7)

tdata = category_data['date']
second = ex1_data[ex1_data['category'] ==obj[0]]
```

```
idata =second['internet']
mdata =second['mobile']

ax[1].scatter( tdata , idata, s = 20, label = obj[0]+'(인터넷)')
ax[1].scatter( tdata , mdata, s = 20, label = obj[0]+'(모바일)')
ax[1].legend()

plt.show()
```



선색

- 'k': 검은색 (black)을 의미합니다. Matplotlib에서는 여러 가지 색을 한 글자로 나타낼 수 있습니다.
- 'b': 파란색 (blue)
- 'g': 녹색 (green)
- 'r': 빨간색 (red)
- 'c': 청록색 (cyan)
- 'm': 자홍색 (magenta)
- 'y': 노란색 (yellow)
- 'k': 검은색 (black)
- 'w': 흰색 (white)

선종류

- '-': 실선 (solid line)을 의미합니다. 선 스타일에는 다양한 종류가 있습니다.
- '-': 실선 (solid line)
- '--': 파선 (dashed line)
- '-.': 점선-파선 혼합 (dash-dot line)
- '...': 점선 (dotted line)

인사이트

인터넷 vs 모바일

- 전체적으로 인터넷과 모바일 쇼핑 거래액 모두 증가하는 추세
- 2020~2022년 정점
- 코로나19 팬데믹으로 인해 온라인 쇼핑 증가
- 비대면 소비 확산, 원격 근무 증가 등으로 관련 제품(컴퓨터, 주변기기 등) 수요 증가
- 모바일 쇼핑 비중 증가
- 초반에는 인터넷(PC 기반) 쇼핑이 더 컸지만, 최근에는 모바일 쇼핑이 상당한 수준으로 성장
- 모바일 쇼핑의 접근성, 결제의 편리함 등이 기여했을 가능성

결론

- 인터넷 쇼핑이 전반적으로 성장하며, 특정 상품군(음식서비스, 여행 등)에서 강한 상관 관계를 보임
- 컴퓨터 및 주변기기 쇼핑의 경우 인터넷(PC)에서 시작했지만, 모바일 쇼핑이 빠르게 성장
- 향후 모바일 쇼핑이 더욱 강세를 보일 가능성이 높고, 온라인 중심의 유통 구조가 더욱 자리 잡을 전망