

### 3.3.1 수치형 데이터 : 상관계수 - 실무예제



다음은 2013년 전국 주요 지점별 유동 인구 현황의 일부이다. 남자 20대 vs. 여자 20대, 남자 10대 vs. 여자 50대의 상관계수를 구하여 비교하고, 중복 속성으로 판단할 수 있을지 검토해 보시오.

◎ 데이터 파일 : ch3-1(유동인구수).csv

◎ 원본 튜플 수 : 23,221개

조사일자	시간대	X좌표	Y좌표	행정구역명	남자 10대	남자 20대	남자 30대	남자 40대	남자 50대	여자 10대	여자 20대	여자 30대	여자 40대	여자 50대
2010-06-21	12시~13시까지	343099	417482	대전광역시 서구 월평동	2	24	68	50	31	4	37	64	44	26
2010-06-21	19시~20시까지	343099	417482	대전광역시 서구 월평동	19	44	28	33	21	14	56	49	43	18
2010-06-20	12시~13시까지	343099	417482	대전광역시 서구 월평동	13	33	34	61	55	13	32	29	28	12
2010-06-20	19시~20시까지	343099	417482	대전광역시 서구 월평동	23	33	32	547	129	12	39	13	46	4
2010-06-21	12시~13시까지	343121	417343	대전광역시 서구 월평동	0	9	27	21	6	5	24	20	10	6

출처: 공공데이터포털([www.data.go.kr](http://www.data.go.kr))

☞ 해답은 [ch3-1.ipynb](#) 참고