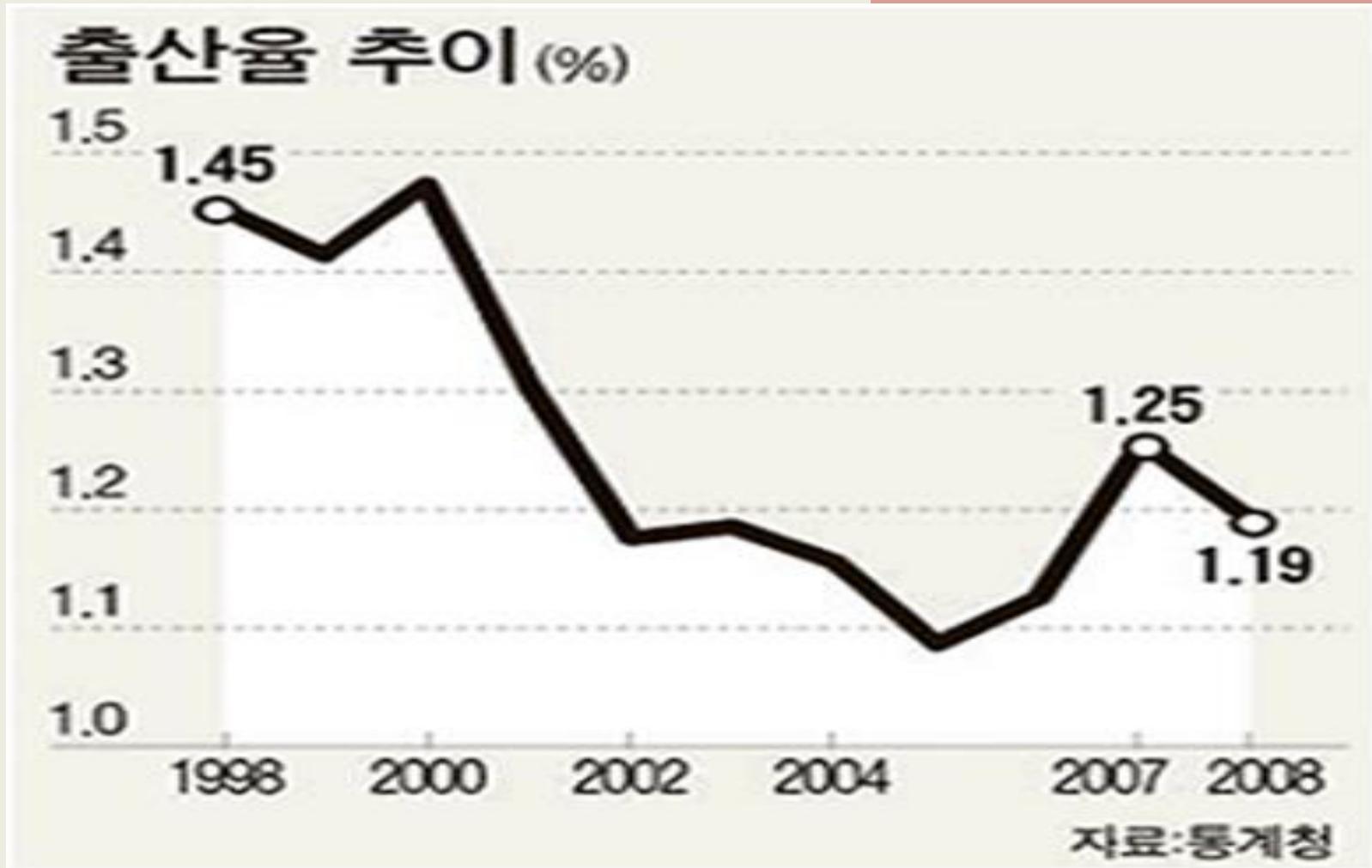
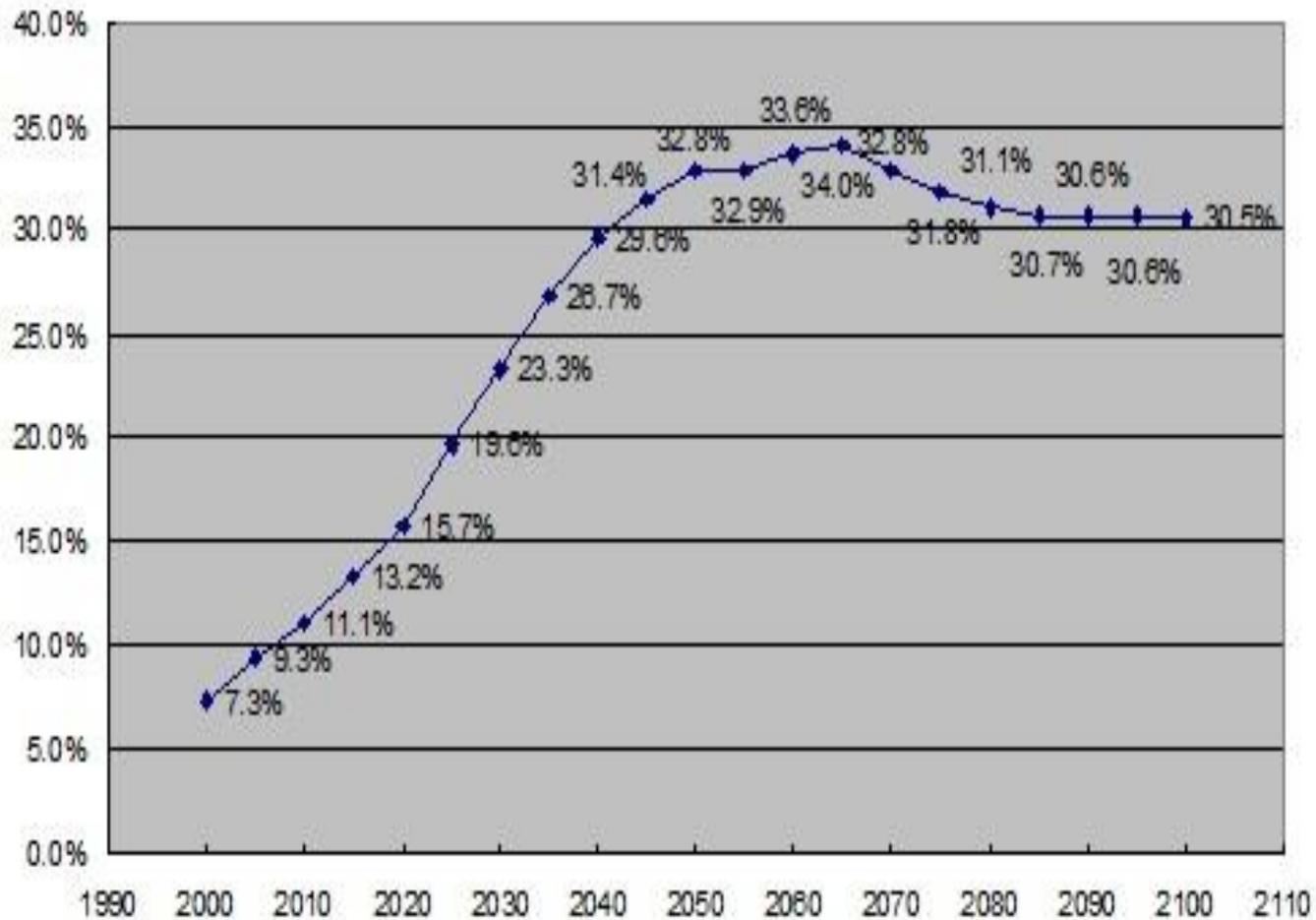


10장: 상관관계분석 (Correlation Analysis)



출산율 <==> 국가발전간 상관관계
는?

65세 이상 노인인구비중(%)



노인인구 <==> 국가발전간 상관관계는?

1. 상관관계분석의 변수

- 두 변수 모두 등간변수(혹은 비율변수)
- 두 변수간에 인과관계는 고려하지 않는다.
- 즉 어느 변수가 독립변수인가, 어느 변수가 종속변수를 따지지 않고,
다만 두 변수간 “관계의 유무” 및 “관계의 정도”
만 조사

2. 상관관계계수(r)

- 두 변수간 관계의 정도를 "**r**"로 나타낸다.
- "**r**"은 두 변수간 표준화된 **공분산**(공통된 부분)을 나타낸다.
- **공분산**이 클수록 "**r**"이 커지는데, 역으로 공분이 작을수록 "**r**"이 작아진다.
(우리는 인생을 살아가면서 사람들간에 공통관심사에 관심이 많다.
 - 우리 연구의 쫓점도 변수(현상)들의 **공분산**에 관심이 많다.
 - 왜냐하면, 변수 변화의 범위와 공통변화여부를 알 수 있기 때문
- 양의 관계는 "**+**"로 나타내고, 음의 관계는 "**-**"로 나타낸다.
- r의 범위는 **-1 --> 0 --> +1**.
- 절대값 1에 가까울수록 강한 관계이며,
0에 가까울수록 약한 관계

3. 해석요령

- 교차분석과 동일
- "+, -"부호와 관계없이

- ① 0 - 0.3까지 : 약
- ② 0.3 - 0.6(0.7) : 중
- ③ 0.6(0.7) - 1 : 강

4. 산점도(산포도 Scatter Plot)

- 두 변수간 관계를 그래프로 나타냄

● SPSS 적용

- ▲ 메뉴 --> 그래프 --> 레거시 대화상자
 - > 산점도/점도표 --> 「단순산점도」 선택
 - > 「정의」 클릭
 - > X축과 Y축에 변수 삽입 --> 확인

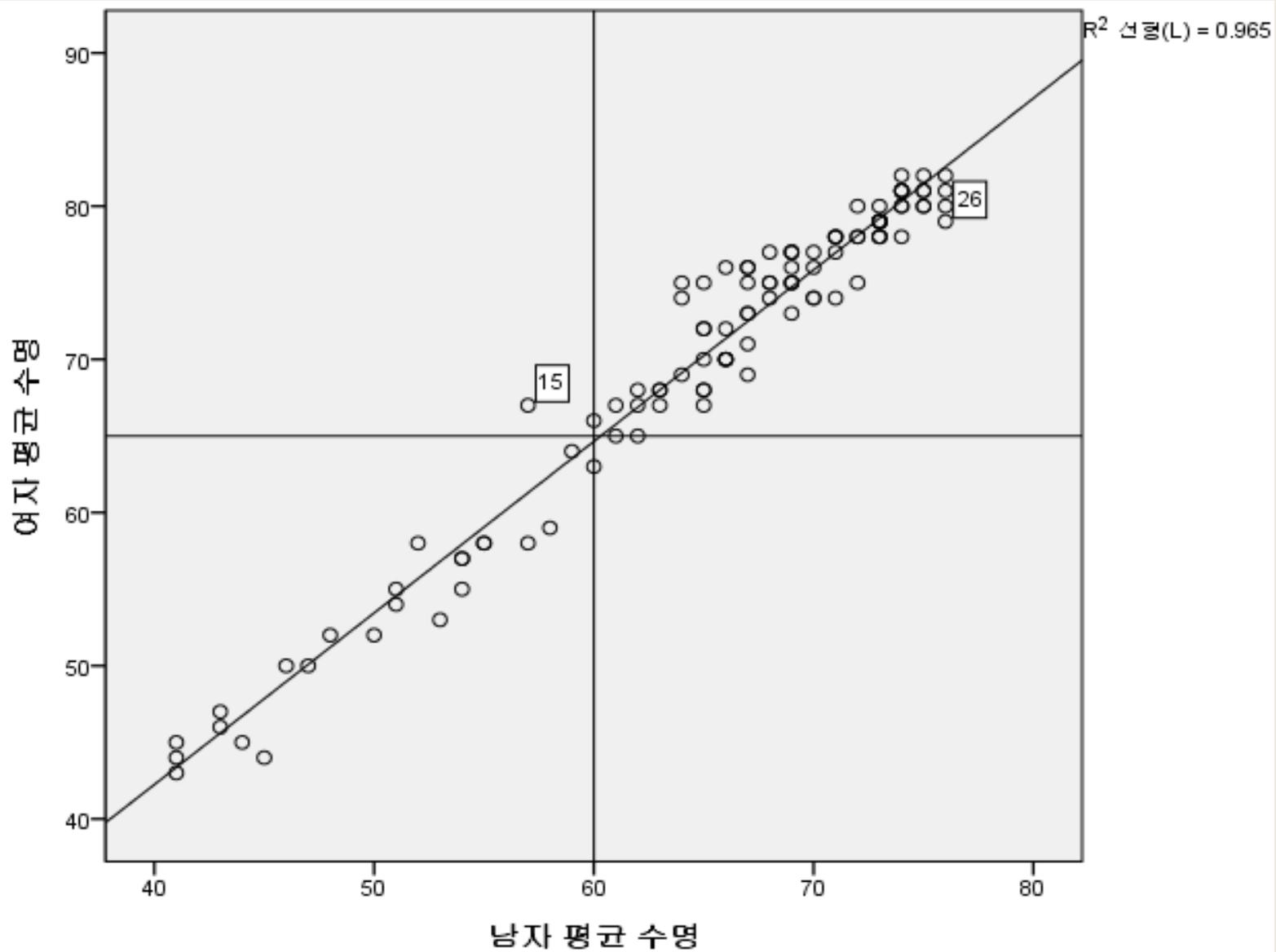
▲ 「World 95」 데이터 이용

▲ 메뉴 --> 그래프 --> 레거시 대화상자 --> 산점도/점도표
 → 「단순산점도」 선택 --> 「정의」 클릭

--> X축에 **남자평균수명**, Y축에 **여자평균수명** 삽입 --> 확인
 → 그래프가 나오면, 그래프 위에 커서를 올려서 2번 클릭

→ 「도표편집기」의 「옵션」에서 남자평균수명의 평균값을
 표시하기 위해 「X축참조선」클릭,
 --> 여자평균수명은 「Y축참조선」클릭

--> 「요소」에서 「데이터설명모드」클릭
 --> 표시를 그래프 상에 관심 있는 데이터위에 올려서 클릭
 --> 「전체적합선」클릭하여 **적합선**과 **R^2** (=설명력)구한다



- 위 그래프는 1995년 전 세계 남자와 여자간 평균수명의 상관관계를 제시하고 있다.
- 두 변수간 상관관계는 정의 방향이며 그 관계가 **매우 강**하다.
- 두 변수간 상호 설명능력을 나타내는 $R^2 = 0.97$ 로 거의 "1"이다.
- **해석**: 전 세계 각국에서 "남자의 평균수명이 높으면 여자의 평균수명도 높다"는 뜻이며 or 거꾸로 "여자의 평균수명이 높으면 남자의 평균수명도 높다"고 할 수도 있다.
- **or 해석**: 남자평균수명과 여자평균수명은 상관관계가 매우 높다!

⇒인제 다른 내용을 잠시 살펴보면...

⇒관심국가로는 먼저 15번 브라질이다.

(남: 57세; 여자: 67세

→ 남녀간 평균수명이 **큰 차이**)

⇒다른 국가들은 남녀간 수명차이가 5살 정도인데 브라질은 관한 연구는 좀 더 이루어져야 할 것이다.

⇒다음 관심국가는 26번 코스타리카인데,

(남: 76세; 여: 79세

→남녀 모두 평균수명이 매우 높다.

아마도 민주주의가 잘 정착되어 있고
기후가 좋기 때문이 아닌가 판단됨.)

5. SPSS 상관관계분석

▲ 「World 95」 데이터 이용

▲ 메뉴 --> 분석 --> 상관분석
--> **이변량상관계수**

→ 왼쪽 변수창에서 "1인당 GDP" 변수와 "AIDS환자수"
변수를 오른쪽 창으로 이동

→ 확인

		1인당 GDP	AIDS 환자 수
1인당 GDP	Pearson 상관계수	1	.258**
	유의확률 (양쪽)		.008
	N	109	106
AIDS 환자수	Pearson 상관계수	.258**	1
	유의확률 (양쪽)	.008	
	N	106	106

★ 예제 1

- 「World 95」 데이터를 이용한다.
- 「여자수명», 「남자수명», 「영아사망», 「GDP», 「AIDS환자」 5개 변수를 활용하여 상관분석을 하고, 표를 만들어 → 설명을 붙인다!

★ 예제 2

- 「World 95」 데이터를 이용한다.
- X변수에 「GDP», Y변수에 「영아사망」을 이용하여 산점도를 그린다!