

이름

데이터테크 전공 20173204 곽명빈

데이터 시각화 기말프로젝트

(출산률과 고용률, 실업률)



seok830621



1. 조사 목적
2. 조사 방법
3. 시각화
4. 결론



곽명빈

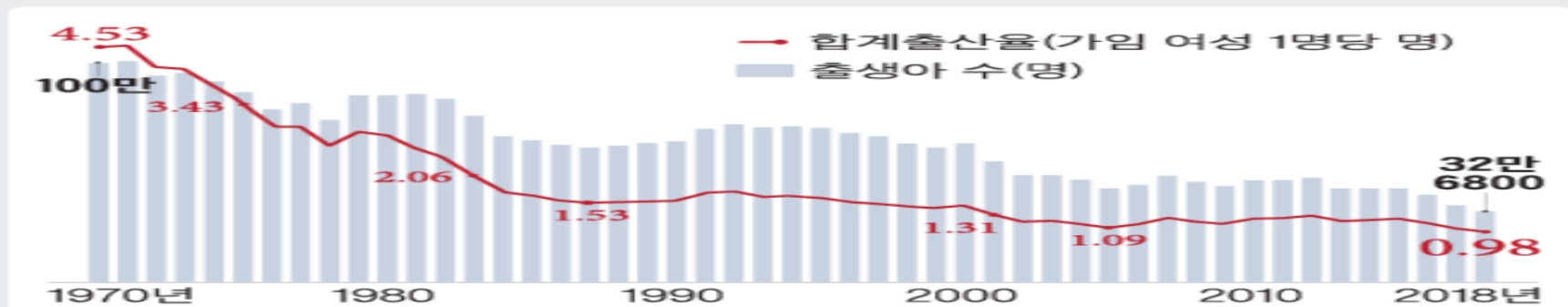


10년간 100조 퍼붓고도 출산율 0.98명...지구 꼴찌 한국

152조 쏟고도...세계 첫 '0명대 출산율'

출산율 0.98명 '인구소멸 국가' 진입...한국 첫 0명대 국가 됐다

합계출산율 통계 작성 이래 최저치



자료: 통계청

② 중앙일보

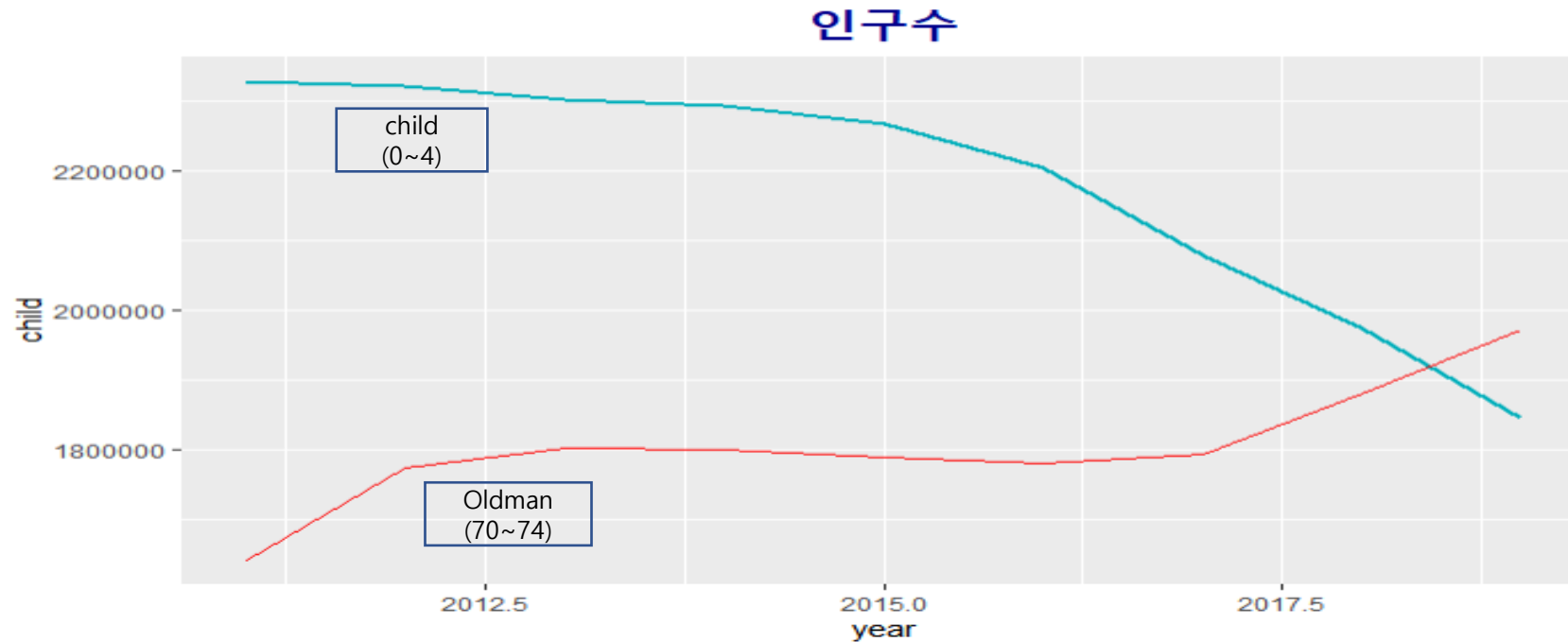
한국의 출산율이 날이 갈수록 떨어져 가고 있는 가운데
출산율에 영향을 미치는 요인은 무엇인가?에 대해 알아보기 위하여 조사해보았다.



곽명빈



```
co <- ggplot(data = man,
             aes(x = year, y = child))+
  geom_line(color = "#00AFBB", size = 1, group = 1)+
  geom_line(aes(x= year , y = oldman), color="red")+
  ggtitle("인구수") +
  theme(plot.title = element_text(family = "serif", face = "bold", hjust = 0.5, size = 20, color =
"darkblue"))
```



아이는 하향세를 보이고있고 노인은 상향그래프를 그리고있다 = 고령화 진행중



곽명빈



◆ 김창순> 그래서 저희들이 저출생 원인을 분석할 때는 저출생 원인이 상당히 다양하고 복합적인 요소로 구성돼 있습니다. 왜 아이를 안 낳느냐는 원인을 따져보면 크게 보면 경제적 측면이 있죠. 거시적인 측면으로 아이를 낳기 위해서는 생활이 안정돼야 하는데 일자리가 없다든지, 고용문제. 또는 소득이 적다든지 또는 교육에 관한 부담이 너무 많다든지 하는 그런 경제적 문제가 큰 것으로 나타나고요. 또 하나는 사회문화적 요인입니다. 사회문화적으로 아까 잠깐 말씀드렸습니다만, 우리나라 사회가 가부장적이고 가사분담이나 육아분담이 너무 여성에게 편중돼있다는 거죠. 여성의 경우 상당히 부담을 느끼고 직장과 가정을 다 돌보기엔 상당히 어렵다, 지금 현재. 그래서 이러한 사회문화적 요인이 저출생을 야기하는 요인으로 작용하고 있고요. 세 번째는 가치관이죠. 가정이나 자녀라든지, 이런 것들에 대한 가치관적인 요인들이 옛날과 달라졌다, 그래서 이러한 부분들이 종합적으로 작용해서 저출생의 원인이 되고 있는데, 사유리 씨 같은 경우 가족 간, 가족을 이루는 그런 관념도, 개념도 상당히 다양화되고 있다. 결혼과 출산에 대한 인식이 많이 변해가고 있는데, 저희가 2020년도 사회조사결과를 보니까 13세 이상은 남녀 중에서 59.7%가 결혼을 하지 않아도 함께 살 수 있다, 이렇게 응답했어요. 결혼하지 않아도 동거할 수 있다. 그 다음에 31.7%는 결혼 안 해도 자녀는 가질 수 있다. 이렇게 답변했거든요. 그래서 가족간, 가족에 대한 인식에 대한 역할이나 생각이 특히 청년 세대에서 많이 변하고 있다. 그래서 이러한 변화를 우리 사회가 어떻게 수용할 수 있느냐 하는 것도 앞으로 우리가 해야 할 큰 정책적 과제가 아닌가, 이런 생각이 듭니다.



곽명빈



◆ 김창순> 그래서 저희들이 저출생 원인을 분석할 때는 저출생 원인이 상당히 다양하고 복합적인 요소로 구성돼 있습니다. 왜 아이를 안 낳느냐는 원인을 따져보면 크게 보면 경제적 측면이 있죠. 거시적인 측면으로 아이를 낳기 위해서는 생활의 안정돼야 하는데 일자리가 없다든지, 고용문제. 또는 소득이 적다든지 또는 교육에 관한 부담이 너무 많다든지 하는 그런 경제적 문제가 큰 것으로 나타나고요. 또 하나는 사회문화적 요인입니다. 사회문화적으로 이가 잠깐 말씀드렸습니다만은, 우리나라 사회가 가부장적이고 가사분담이나 육아분담이 너무 여성에게 편중돼있다는 거죠. 여성의 경우 상당히 부담을 느끼고 직장과 가정을 다 돌보려한 상황이 어렵다, 지금 현재. 그래서 이러한 사회문화적 요인이 저출생을 야기하는 요인으로 작용하고 있고요. 세 번째는 가치관입니다. 가정이나 자녀라든지, 이런 것들에 대한 가치관적인 요인들이 옛날과 달라졌다, 그래서 이러한 부분들이 종합적으로 작용해서 저출생의 원인이 되고 있는데 사유각 씨 같은 경우 가족 간, 가족을 이루는 그런 관념도, 개념도 상당히 다양화되고 있다. 결혼과 출산에 대한 인식이 많이 변해가고 있는데, 저희가 2020년도 사회조사결과를 보니까 13세 이상은 남녀 중에서 59.7%가 결혼을 하지 않아도 함께 살 수 있다, 이렇게 응답했어요. 결혼하지 않아도 동거할 수 있다. 그 다음에 31.7%는 결혼 안 해도 자녀는 가질 수 있다. 이렇게 답변했거든요. 그래서 가족간, 가족에 대한 인식에 대한 역할이나 생각이 특히 청년 세대에서 많이 변하고 있다. 그래서 이러한 변화를 우리 사회가 어떻게 수용할 수 있느냐 하는 것도 앞으로 우리가 해야 할 큰 정책적 과제가 아닌가, 이런 생각이 듭니다.

저출산의 원인을 3가지로 분류할 수 있다.

1. 경제적 요인

2. 사회문화적 요인

3. 가치관



곽명빈



출산물

Year	Mean
2000	1.467
2001	1.301
2002	1.164
2003	1.175
2004	1.144
2005	1.075
2006	1.117
2007	1.251
2008	1.184
2009	1.151
2010	1.226
2011	1.251
2012	1.403
2013	1.286
2014	1.299
2015	1.364
2016	1.297
2017	1.171
2018	1.088
2019	1.026

실업률

Year	Mean
2000	3.94
2001	3.52
2002	2.82
2003	3.06
2004	3.13
2005	3.14
2006	2.88
2007	2.76
2008	2.71
2009	3.02
2010	3.03
2011	2.77
2012	2.61
2013	2.6
2014	3.01
2015	2.99
2016	3.15
2017	3.21
2018	3.49
2019	3.5

고용률

Year	Mean
2000	58.51
2001	59.11
2002	60.3
2003	59.68
2004	60.14
2005	59.92
2006	60.02
2007	60.21
2008	59.91
2009	59.18
2010	58.99
2011	59.36
2012	59.75
2013	59.88
2014	60.47
2015	60.72
2016	60.78
2017	61.09
2018	60.98
2019	61.31

각 데이터가 시/군/구 별로 구분되어 있어 각 시의 평균으로 데이터 정리



곽명빈



```
a <- ggplot(data = baby,
  aes(x=Year, y=Mean))+
  geom_line(color = "#00AFBB", size = 1, group = 1)+
  ggtitle("출산률") +
  theme(plot.title = element_text(family = "serif", face = "bold", hjust = 0.5, size = 20, color =
"darkblue"))

b <- ggplot(data = lost,
  aes(x=Year, y=Mean))+
  geom_line(color = "#00AFBB", size = 1, group = 1)+
  ggtitle("실업률") +
  theme(plot.title = element_text(family = "serif", face = "bold", hjust = 0.5, size = 20, color =
"darkblue"))

c <- ggplot(data = job,
  aes(x=Year, y=Mean))+
  geom_line(color = "#00AFBB", size = 1, group = 1)+
  ggtitle("고용률") +
  theme(plot.title = element_text(family = "serif", face = "bold", hjust = 0.5, size = 20, color =
"darkblue"))

grid.arrange(a,b,c, ncol = 2, nrow = 2)
## 꺾은선 그래프를 통해 출산률과 실업률, 고용률을 표현 |
```

각 데이터의 연도별 변화를 확인하기 위해 line chart 그리기



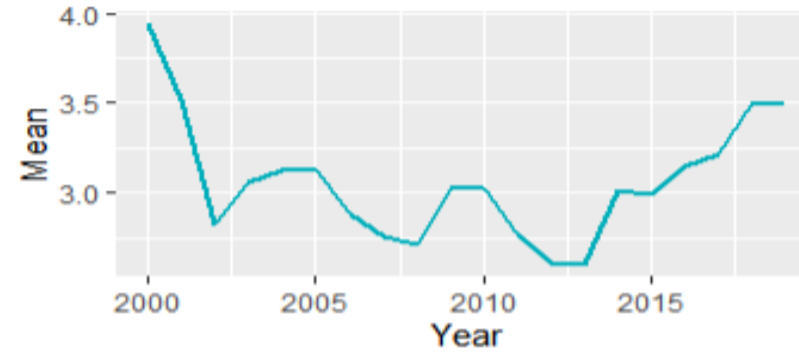
곽명빈



출산물



실업률



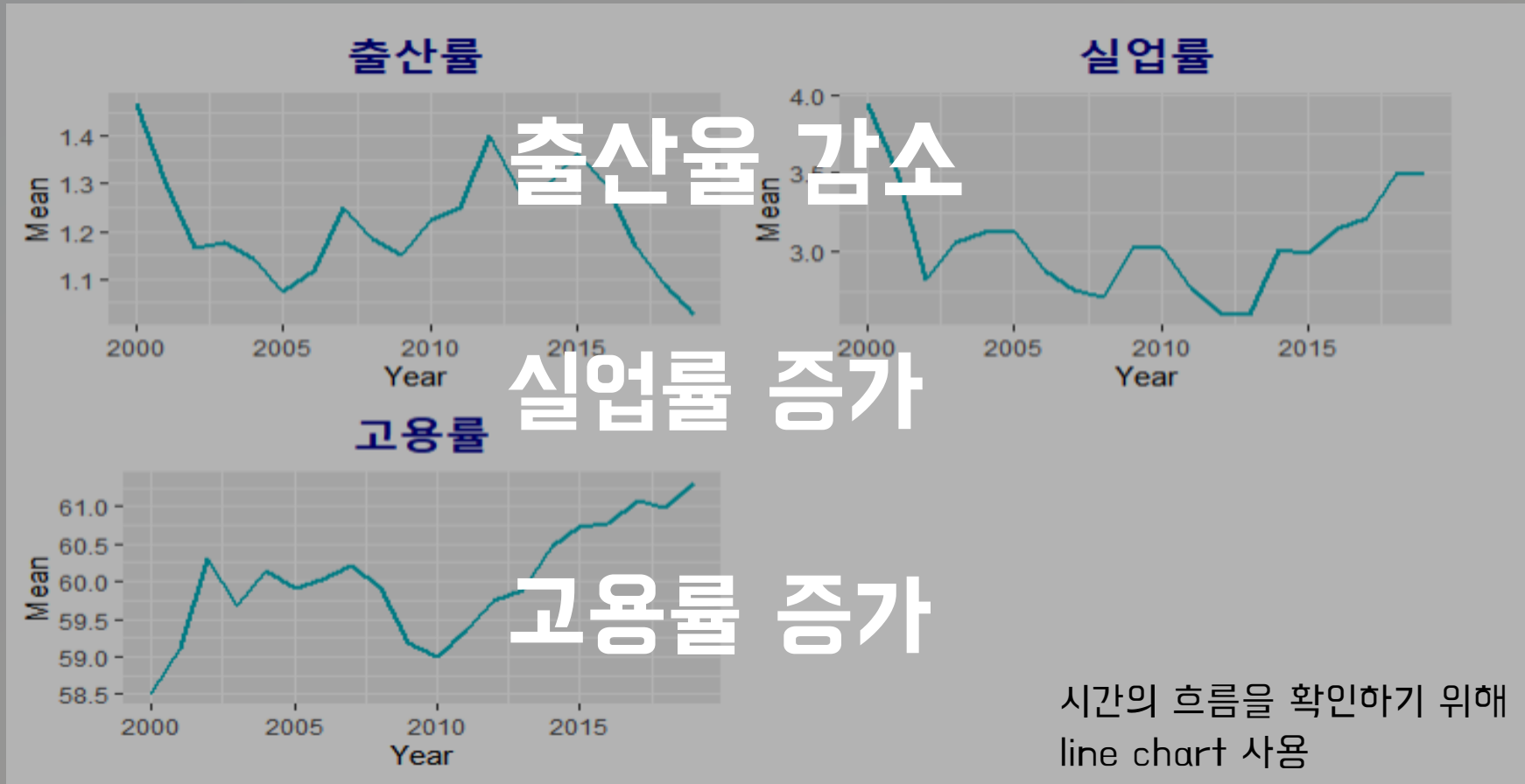
고용률



시간의 흐름을 확인하기 위해
line chart 사용



곽명빈





곽명빈



```
mean1 <- cbind(baby$Mean, lost$Mean, job$Mean)
```

상관관계 확인을 위하여 데이터정리

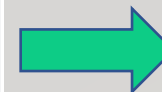
```
> mean1
      [,1] [,2] [,3]
[1,] 1.467 3.94 58.51
[2,] 1.301 3.52 59.11
[3,] 1.164 2.82 60.30
[4,] 1.175 3.06 59.68
[5,] 1.144 3.13 60.14
[6,] 1.075 3.14 59.92
[7,] 1.117 2.88 60.02
[8,] 1.251 2.76 60.21
[9,] 1.184 2.71 59.91
[10,] 1.151 3.02 59.18
[11,] 1.226 3.03 58.99
[12,] 1.251 2.77 59.36
[13,] 1.403 2.61 59.75
[14,] 1.286 2.60 59.88
[15,] 1.299 3.01 60.47
[16,] 1.364 2.99 60.72
[17,] 1.297 3.15 60.78
[18,] 1.171 3.21 61.09
[19,] 1.088 3.49 60.98
[20,] 1.026 3.50 61.31
```



```
colnames(mean1) <- c("baby1", "lost1", "job1")
mean1 <- data.frame(mean1)
```

데이터 프레임형태로 전환

	baby1	lost1	job1
1	1.467	3.94	58.51
2	1.301	3.52	59.11
3	1.164	2.82	60.30
4	1.175	3.06	59.68
5	1.144	3.13	60.14
6	1.075	3.14	59.92
7	1.117	2.88	60.02
8	1.251	2.76	60.21
9	1.184	2.71	59.91
10	1.151	3.02	59.18
11	1.226	3.03	58.99
12	1.251	2.77	59.36
13	1.403	2.61	59.75
14	1.286	2.60	59.88
15	1.299	3.01	60.47
16	1.364	2.99	60.72
17	1.297	3.15	60.78
18	1.171	3.21	61.09
19	1.088	3.49	60.98
20	1.026	3.50	61.31



```
```{r}
cor(mean1$baby1, mean1$lost1)
cor(mean1$baby1, mean1$job1)
cor(mean1$lost1, mean1$job1)
```
```

```
[1] -0.004047275
[1] -0.4278203
[1] -0.05192147
```

상관계수



곽명빈



```
```{r}
d <- ggplot(data = mean1,
 aes(x = baby1, y = lost1))+
 geom_point(size = 2, color = "skyblue")+
 ggtitle("Scatter plot : 출산률과 실업률")+
 stat_smooth(method = 'lm', se=F)+
 theme(plot.title = element_text(size = 15))

e <- ggplot(data = mean1,
 aes(x = baby1, y = job1))+
 geom_point(size = 2, color = "skyblue")+
 ggtitle("Scatter plot : 출산률과 고용률")+
 stat_smooth(method = 'lm', se=F)+
 theme(plot.title = element_text(size = 15))

f <- ggplot(data = mean1,
 aes(x = lost1, y = job1))+
 geom_point(size = 2, color = "skyblue")+
 ggtitle("Scatter plot : 실업률과 고용률")+
 stat_smooth(method = 'lm', se=F)+
 theme(plot.title = element_text(size = 15))

grid.arrange(d,e,f, ncol = 2, nrow = 2)
```

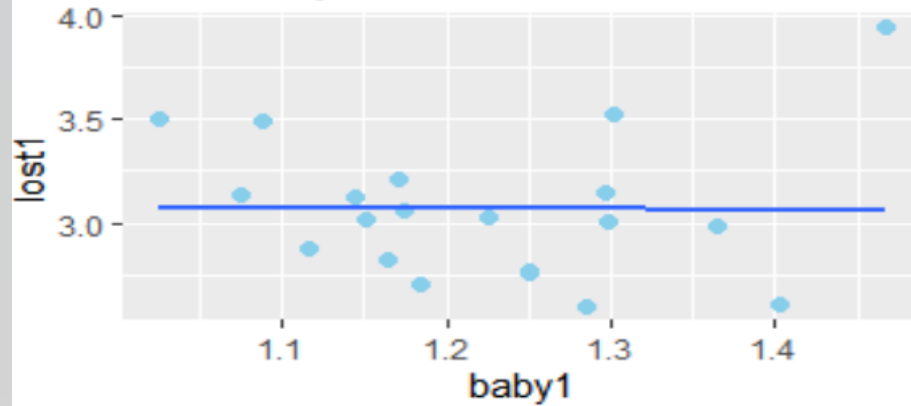
각 변수별로 관계를 시각화 하기 위하여 geom\_point 사용



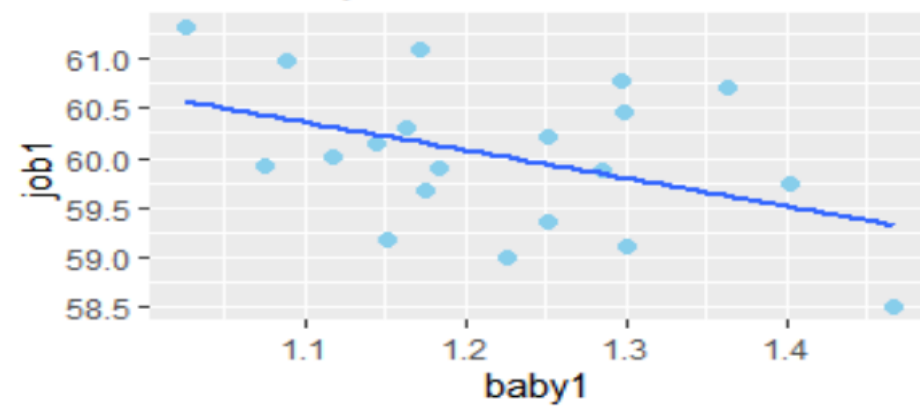
곽명빈



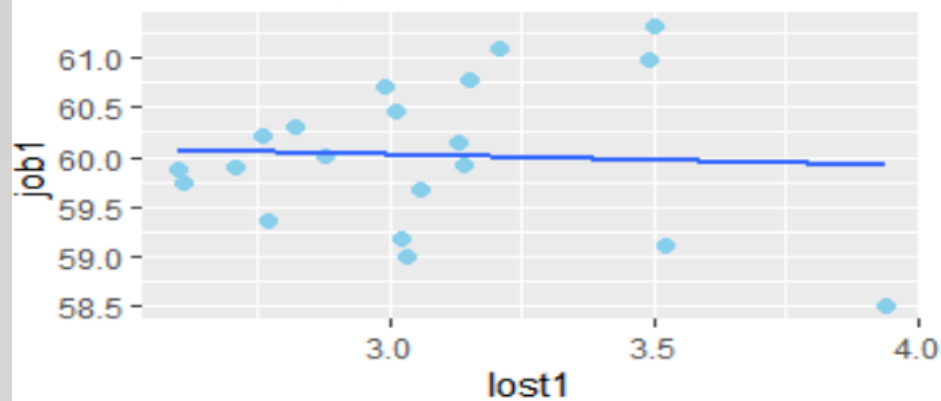
Scatter plot : 출산률과 실업률



Scatter plot : 출산률과 고용률



Scatter plot : 실업률과 고용률



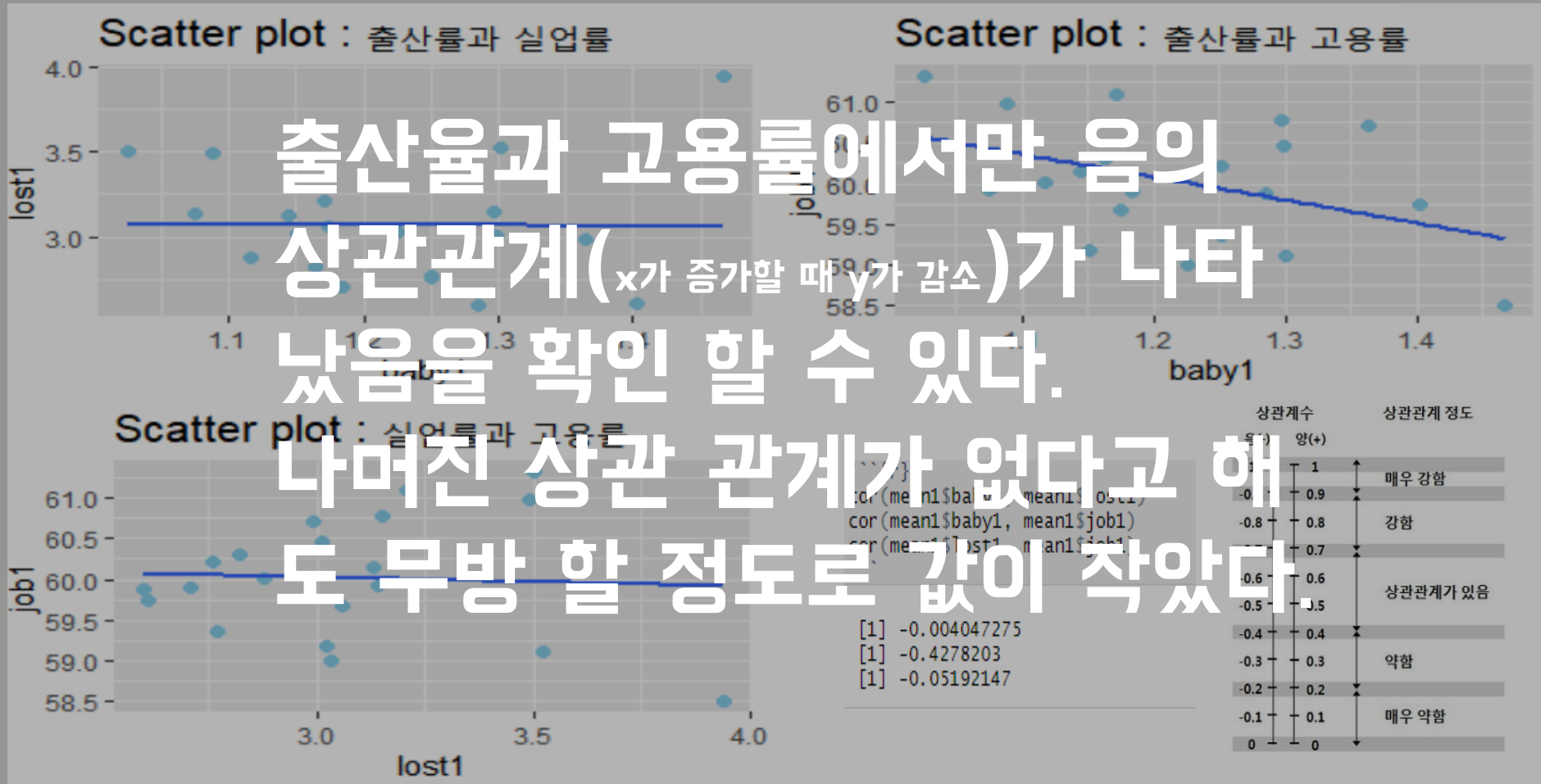
```
```{r}
cor(mean1$baby1, mean1$lost1)
cor(mean1$baby1, mean1$job1)
cor(mean1$lost1, mean1$job1)
```
```

```
[1] -0.004047275
[1] -0.4278203
[1] -0.05192147
```

| 상관계수 |      | 상관관계 정도  |
|------|------|----------|
| 음(-) | 양(+) |          |
| -1   | 1    | 매우 강함    |
| -0.9 | 0.9  |          |
| -0.8 | 0.8  | 강함       |
| -0.7 | 0.7  |          |
| -0.6 | 0.6  | 상관관계가 있음 |
| -0.5 | 0.5  |          |
| -0.4 | 0.4  |          |
| -0.3 | 0.3  | 약함       |
| -0.2 | 0.2  |          |
| -0.1 | 0.1  | 매우 약함    |
| 0    | 0    |          |



곽명빈

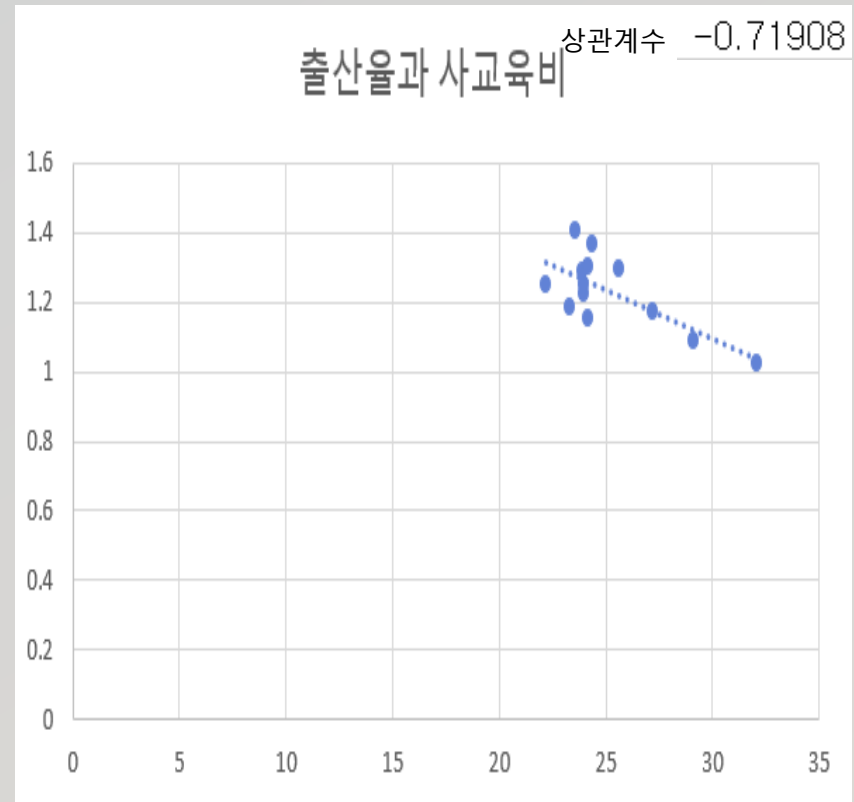




곽명빈



| Year | Mean(단위: 만) | Mean  |
|------|-------------|-------|
| 2007 | 22.2        | 1.251 |
| 2008 | 23.3        | 1.184 |
| 2009 | 24.2        | 1.151 |
| 2010 | 24          | 1.226 |
| 2011 | 24          | 1.251 |
| 2012 | 23.6        | 1.403 |
| 2013 | 23.9        | 1.286 |
| 2014 | 24.2        | 1.299 |
| 2015 | 24.4        | 1.364 |
| 2016 | 25.6        | 1.297 |
| 2017 | 27.2        | 1.171 |
| 2018 | 29.1        | 1.088 |
| 2019 | 32.1        | 1.026 |



사교육비와 출산율을 산점도로 표현해보았다.

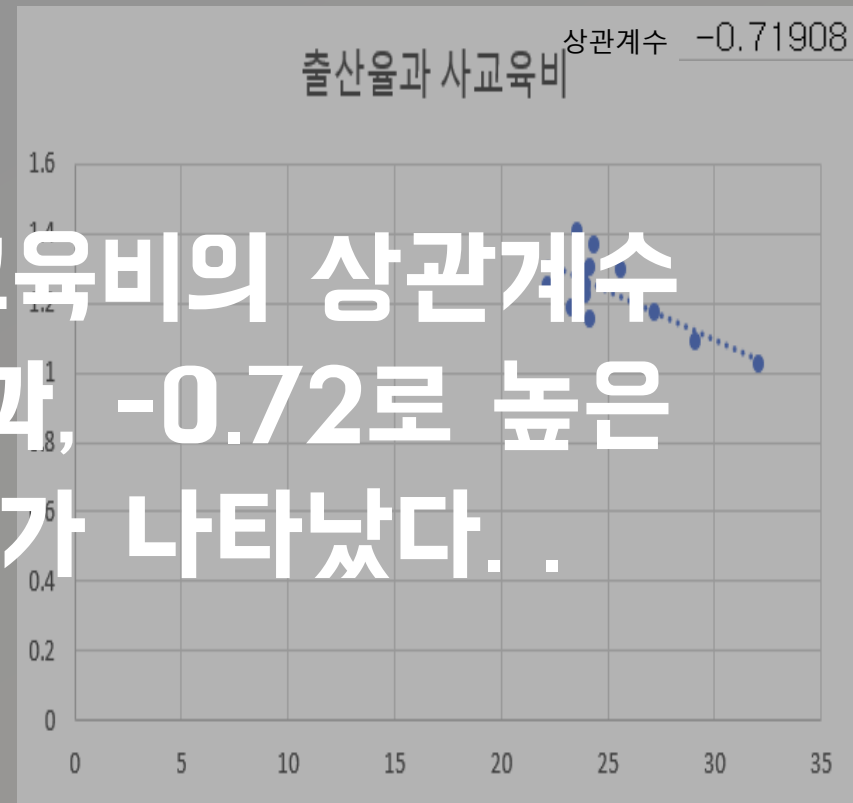


곽명빈



| Year | Mean(단위: 만) | Mean  |
|------|-------------|-------|
| 2007 | 22.2        | 1.251 |
| 2008 | 23.3        | 1.184 |
| 2009 | 24.2        | 1.151 |
| 2010 | 24.6        | 1.216 |
| 2011 | 24.6        | 1.251 |
| 2012 | 23.6        | 1.405 |
| 2013 | 23.9        | 1.286 |
| 2014 | 24.2        | 1.332 |
| 2015 | 24.4        | 1.364 |
| 2016 | 25.6        | 1.297 |
| 2017 | 27.2        | 1.171 |
| 2018 | 29.1        | 1.088 |
| 2019 | 32.1        | 1.026 |

출산율과 사교육비의 상관계수를  
구해본 결과, -0.72로 높은  
음의상관관계가 나타났다.

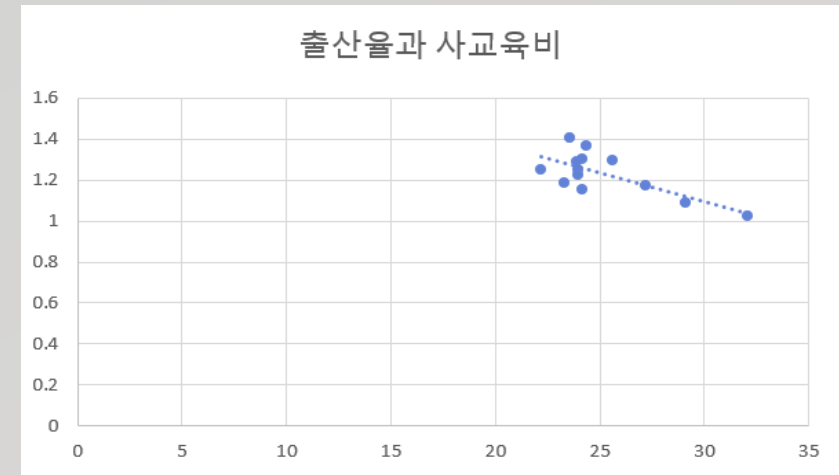
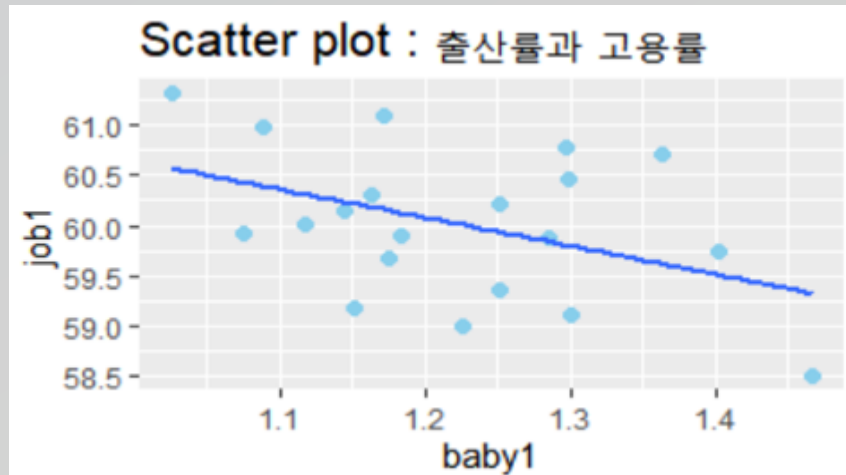


사교육비와 출산율을 산점도로 표현해보았다.





곽명빈



이번 조사를 통해 출산율에 관계가 있는 요소는 고용률과 사교육비라는 사실을 알 수 있었다. 실업률에는 매우 낮은 음의 상관관계라 무시해도 될 수준 이였고, 특히 사교육비가 높아질수록 출산율이 낮아진다는 결과를 산점도를 통해 알 수 있었다. 이를 통해 출산율에는 경제적 요인(사교육비)가 영향이 있을 수 있다는 사실을 알 수 있다.

고용률과 출산율이 왜 음의 상관관계를 가지는지 생각해 보았는데, 한정된 일자리양에 노동 인구가 줄어들어 나타나는 현상이라고 생각했다.

