



🦓 제쥬대학교 I. Excel 데이터 정렬

1. 데이터 정렬

- 항목별로 일정한 순서에 따라 데이터를 나열해 주는 기능을 정렬이라고 함
- 데이터 정렬은 다음의 순서로 함
 - · 정렬할 기준 셀로 셀 포인터를 이동
 - ㆍ[홈]-[편집] 그룹에서 [정렬 및 필터]를 클릭하여 [숫자 내림차순 정리]를 선택

2. 실습

- 강기춘 홈페이지에서 b3-ch1-3.xlsx 파일을 다운로드

	Α	В	С	D	E	F	G	Н
1	name	prin	micro	macro	sum	mean	PF	grade
2	김기훈	98	82	95	275	91.66667	pass	Α
3	박수동	100	92	80	272	90.66667	pass	Α
4	원선희	50	45	75	170	56.66667	fail	F
5	위계영	50	100	100	250	83.33333	pass	В
6	최동팔	80	95	95	270	90	pass	Α
7	최종열	90	60	60	210	70	pass	С
8	최종수	30	30	30	90	30	fail	F
9	김기팔	80	25	25	130	43.33333	fail	F
10	이상수	65	70	70	205	68.33333	pass	D
11	강창수	95	90	90	275	91.66667	pass	Α

- 예 1 : 데이터를 평균을 중심으로 내림차순으로 정렬하려면 다음의 순서로 함
 - · A2부터 H11까지 셀을 선택한 후 [데이터]-[정렬 및 필터] 그룹에서 [정렬]을 클릭하면 정렬 대화상자 가 나타남
 - ·정렬 대화상자에서 열-정렬 기준의 🛛를 클릭하면 나타나는 선택 항목 중 mean(평균)을 선택
 - · 정렬의 💵를 클릭하면 나타나는 선택 항목 중 내림차순을 선택하고 확인을 누름

정렬		? ×
*ू↓ 기준 추가(<u>A</u>) 🗙 기준 삭제(<u>D</u>)	🖹 기준 복사(C) 🔺 🔻 옵션(Q) · 내 데이터에 머리글 표시(<u>H</u>)
열	정렬 기준	정렬
정렬 기준 mean 🗸	값 ~	내림차순 🗸
L		확인 취소

	Α	В	С	D	E	F	G	Н
1	name	prin	micro	macro	sum	mean	PF	grade
2	김기훈	98	82	95	275	91.66667	pass	Α
3	강창수	95	90	90	275	91.66667	pass	Α
4	박수동	100	92	80	272	90.66667	pass	Α
5	최동팔	80	95	95	270	90	pass	Α
6	위계영	50	100	100	250	83.33333	pass	В
7	최종열	90	60	60	210	70	pass	С
8	이상수	65	70	70	205	68.33333	pass	D
9	원선희	50	45	75	170	56.66667	fail	F
10	김기팔	80	25	25	130	43.33333	fail	F
11	최종수	30	30	30	90	30	fail	F



- 예 2 : 데이터를 prin(경제원론)을 중심으로 오름차순으로 정렬하려면 다음의 순서로 함
 - · A2부터 H11까지 셀을 선택한 후 [데이터]-[정렬 및 필터] 그룹에서 [정렬]을 클릭하면 정렬 대화상자 가 나타남
 - · 정렬 대화상자에서 열-정렬 기준의 ☑를 클릭하면 나타나는 선택 항목 중 prin(경제원론)을 선택
 - · 정렬의 ☑를 클릭하면 나타나는 선택 항목 중 오름차순을 선택하고 확인을 누름

정렬					?	×
• 기준 추가(<u>A</u>)	🗙 기준 삭제(<u>D</u>)	🗈 기준 복사(<u>C</u>)	▲ ▼ 옵션() ☑ 내 데이터	에 머리글 표	시(<u>H</u>)
열		정렬 기준		정렬		
정렬 기준 prin	~	값	\sim	오름차순		\sim
				확인	취소	٤

	А	В	С	D	E	F	G	н
1	name	prin	micro	macro	sum	mean	PF	grade
2	최종수	30	30	30	90	30	fail	F
3	원선희	50	45	75	170	56.66667	fail	F
4	위계영	50	100	100	250	83.33333	pass	В
5	이상수	65	70	70	205	68.33333	pass	D
6	최동팔	80	95	95	270	90	pass	Α
7	김기팔	80	25	25	130	43.33333	fail	F
8	최종열	90	60	60	210	70	pass	С
9	강창수	95	90	90	275	91.66667	pass	Α
10	김기훈	98	82	95	275	91.66667	pass	Α
11	박수동	100	92	80	272	90.66667	pass	Α

🔏 제조대학교 표. R 데이터 정렬

1. dplyr

- dplyr 패키지는 데이터 전처리(preprocessing)를 위한 패키지
- 데이터 전처리란 원하는 형태로 데이터를 가공하는 것을 말함
- dplyr 패키지를 install 하기 위해서 Console 창에서 install.packages("dplyr")을 입력
- dplyr에서 주로 사용하는 함수는 다음과 같음
 - · select() : 특정 열(변수)을 추출
 - · arrange() : 데이터를 오름차순이나 내림차순으로 정렬
 - · filter() : 조건을 주어 행을 추출
 - · mutate() : 기존의 변수(열)를 이용하여 새로운 변수(열) 생성
 - · summarise()(with group_by) : 집단별로 통계량 계산
- %>%는 '무엇인가를 전송한다'는 의미의 파이프 연산자(pipe operator)로서 동일한 자료 집합에서 많 은 작업을 할 때 유용
 - · 왼쪽에서 계산된 결과 값을 파이프 오른쪽의 함수 첫 번째 입력 값으로 넘겨주는 역할을 함

(1) select()

- 데이터의 특정 열(변수)을 선택
- 예 1 : prin 및 micro를 선택
- 예 2 : prin부터 macro까지 한꺼번에 선택
- 예 3 : micro만 제외하고 선택
- 예 4 : 열이름(변수명)이 cro로 끝나는 열만 선택
- 예 5 : 변수명을 변경

b3-ch2-1.R				>	select prin	(df, prin, m micro	icro)
library(dplyr)				1 2	98 100	82 92	
│ │ name<-c("김기훈","박수동","원선희","위계영","최동팔","최종열	","최종수","?	김기팔","이상수",'	'강창수")	3 4	50 50	45 100	
prin<-c(98,100,50,50,80,90,30,80,65,95)				5 6	80 90	95 60	
micro<-c(82,92,45,100,95,60,30,25,70,90)		7 8	30 80	30 25			
macro<-c(95.80.75.100.95.60.30.25.70.90)		9 10	65 95	70 90			
df<-data frame(name prin micro macro)		> :	select	(df, prin:m	iacro)		
dfsum<- df srin+ df smicro+ df smacro				1	98	82 95	i i
$df_{man}^{2} - df_{nm}^{2}$				2	100 50	92 80 45 75	5
$df \Phi DE (df \Phi man N = 60 name fail)$				4	50 80	100 100) 5
$dt\phi$ reads (it less ($dt\phi$ reason $\lambda = 00$, μ as s, fair)	"D" ifalaa("O" ifala	6	90	60 60))
disgrade<-ileise(dismean >= 90, A, ileise(dismean >= 80,	, B, Heise(di\$mean >= 70,	C, Heis	8	30 80	25 25) \$
e(df\$mean >= 60, "D", "F"))))				9 10	65 95	70 70 90 90)
				> se	lect(df, ends_wi	th("cro"))
select(df, prin, micro) 1 82 95							
select(df, prin:macro)				2	92 45	80 75	
select(df, -micro)				4	100	100	
select(df, ends_with("cro"))				6	95 60	95 60	
rename(df,이름=name,경제원론=prin, 미시경제=micro,거시경	제=macro,君	통점=sum,평균=m	nean,이수	7	30	30	
=PF,학점=grade)				9	70	70	
				10	90	90	
<pre>> select(df, -micro)</pre>	01=	경제의로 미지?	테기시경			펴그 이스 친	하지
1 김기훈 98 95 275 91.66667 pass A	1 김기훈	- 8세권는 비시경 98	82	^⊪ =⊂⊟ 95	275	91.66667	pass A
2 박수동 100 80 272 90.66667 pass A	2 박수동	100	92	80	272	90.66667	pass A
3 원선희 50 75 170 56.66667 tail F 4 이계여 50 100 250 92 22222 page P	3 원선희	50	45	75	170	56.66667	fail F
4 위계8 50 100 250 85.55555 pass B 5 최동팔 80 95 270 90.00000 pass A	4 위계 3 5 최동팔	80	95	95	230	90,00000	pass b pass A
6 최종열 90 60 210 70.00000 pass C	6 최종열	90	60	60	210	70.00000	pass C
7 최종수 30 30 90 30.00000 fail F	7 최종수	30	30	30	90	30.00000	fail F
8 김기팔 80 25 130 43.33333 fail F	8 김기팔	80	25	25	130	43.33333	fail F
9 미상수 65 70 205 68.33333 pass D 10 강창스 95 90 275 91 66667 pass A	9 이상수 10 강찬스	65	70	70	205	68.33333 01.66667	pass D
10 강경우 95 90 275 91.66667 pass A	10 상상수	95	90	90	275	91.66667	pass A

(2) arrange()

- 특정변수를 기준으로 오름차순 또는 내림차순으로 정렬
- 예1: 평균을 기준으로 내림차순으로 정렬
- 예 2 : 경제원론을 기준으로 오름차순으로 정렬

b3-ch2-1 B		>	df %>%	arran	ge(des	c(mean))
	-		name	prin	micro	macro sum mean PF grade
library(dplyr)		1	11년 1년 가차스	98	82	95 275 91.66667 pass A
name<-c("김기훈","박수동","원선희","위계영","최동팔","최종열","최종		3	* 박수동	100	92	80 272 90.66667 pass A
수","김기팔","이상수","강창수")		4	최동팔	80	95	95 270 90.00000 pass A
prin<-c(98,100,50,50,80,90,30,80,65,95)		6	취계 8 최종열	90	60	60 210 70.00000 pass C
micro = -c(82, 02, 45, 100, 05, 60, 30, 25, 70, 00)		7	이상수	65	70	70 205 68.33333 pass D
THICION=C(02,92,43,100,93,00,30,23,70,90)		8	원선희	50	45	75 170 56.66667 fail F
macro<-c(95,80,75,100,95,60,30,25,70,90)		9	· 취족스	20	25	25 130 43.33333 Tall F
df<-data.frame(name,prin,micro,macro)		11	-10-	50	50	50 50 50.0000 Fall P
df\$sum<-df\$prin+df\$micro+df\$macro						
df\$mean<-df\$sum/3		>	df %>%	arran	ae(pri	n) .
df\$PF<-ifelse(df\$mean >= 60, "pass","fail")		Ľ	name	prin	micro	macro sum mean PF grade
df\$grade<-ifelse(df\$mean >= 90, "A", ifelse(df\$mean >= 80, "B", if		1	죄송수 원성히	30 50	30 45	30 90 30.00000 fail F 75 170 56 66667 fail F
alse(df mean > 70 "C" if $alse(df$ mean > 60 "D" "F"))))		3	휘계영	50	100	100 250 83.33333 pass B
		4	이상수	65	70	70 205 68.33333 pass D
		5	최동팔	80	95	95 270 90.00000 pass A
df[order(df\$mean,decreasing=T),]		6	김기딸 치조여	80	25	25 130 43.33333 fail F
df %>% arrange(desc(mean))		8	의 등 물 강창수	90	90	90 275 91,66667 pass A
		9	김기훈	98	82	95 275 91.66667 pass A
df[order(df\$prin),]		1) 박수통	100	92	80 272 90.66667 pass A
df %>% arrange(prin)						





1. 데이터 입력

- 데이터베이스(DB)는 데이터를 구성하고, 관리하며, 검색하는 도구임
- 대표적인 DB는 통계청의 국가통계포털(KOSIS)임
- 다음의 데이터를 입력하면 엑셀은 자동적으로 데이터베이스로 취급
- 강기춘 홈페이지에서 dbase.xlsx 파일을 다운로드

	A	В	С	D	E	F	G
1	학번	이름	생년월일	성별	본적	소득	
2	6789	김상욱	1970-02-10	남	부산	1000	
3	2346	김숙희	1971-03-29	여	대구	2000	
4	3456	박윤정	1965-09-07	여	수원	3500	
5	3455	윤석호	1974-12-07	남	청주	2500	
6	1234	니경규	1968-01-01	남	광주	1500	
7	4546	이동균	1969-05-23	남	민천	5000	
8	2341	조동민	1967-08-30	남	경기	8000	
9	6543	최정민	1972-04-01	여	제주	4000	
10	2345	홍길동	1973-10-22	남	서울	3800	
11							

1 학번 ▼ 이름 ▼ 생년월일 ▼ 성별 ▼ 본적 ▼ 소득 2 6789 김상욱 1970-02-10 남 부산 100 5 3455 윤석호 1974-12-07 남 청주 250	A
26789 김상욱1970-02-10 남부산10053455 윤석호1974-12-07 남청주250	1 학번 🔽
5 3455 윤석호 1974-12-07 남 청주 250	2 6789
	5 3455
6 1234 니경규 1968-01-01 남 광주 150	3 1234
_7 4546 이동균 1969-05-23 남 인천 500	7 4546
<u>8</u> 2341 조동민 1967-08-30 남 경기 800	3 2341
10 2345 홍길동 1973-10-22 남 서울 380	0 2345

- · 성별 풀다운 목록을 클릭하면 나타나는 대화 창에서 남자만 선택하고 확인을 누름
- 예 1 : 남학생만 검색

	A	B	C	D	E	F	G
1	학번 🔽	이름 🔽	생년월일 💌	성별 🔽	본적 💌	소득 🖃	
2	6789	김상욱	1970-02-10	남	부산	1000	
3	2346	김숙희	1971-03-29	여	대구	2000	
4	3456	박윤정	1965-09-07	여	수원	3500	
5	3455	윤석호	1974-12-07	남	청주	2500	
6	1234	니경규	1968-01-01	남	광주	1500	
7	4546	이동균	1969-05-23	남	인천	5000	
8	2341	조동민	1967-08-30	남	경기	8000	
9	6543	최정민	1972-04-01	여	제주	4000	
10	2345	홍길동	1973-10-22	남	서울	3800	
11							
4.0							

- 화살표를 클릭하면 그 열에 있는 고유한 항목이 풀다운 목록에 모두 나타남
- [데이터]-[정렬 및 필터] 그룹에서 [필터]를 클릭하면 첫 번째 열의 각 행 바로 옆에 화살표가 나타남
- 데이터를 검색하기 위해서는 데이터베이스의 임의의 셀로 셀 포인터를 이동

2. 데이터베이스 검색



- 예 2 : 남학생이면서 본적이 서울
 - 먼저 성별 풀다운 목록을 클릭하여 남자만 선택하여 검색
 - · 다음으로 본적 풀다운 목록을 클릭하여 서울을 선택하고 확인을 누름



- 예 3 : 조건이 사전 목록으로 정의되어 있지 않은 경우 사용자 정의를 이용
 - · 예를 들어, 소득이 2000보다 많고 4000미만인 학생들을 검색하고자 할 경우
 - · 소득 풀다운 목록을 누르고 숫자 필터-사용자 지정 필터를 선택하면 나타나는 사용자 지정 자동필터 대화상자에 찾을 조건을 설정해 주고 확인을 누름

사용자 지정 자동 필터	?	×
찾을 조건: 소득		
> 4000		\sim
● 그리고(A) ○ 또는(D)		
< 2000		~
? 기호를 사용하여 한 문자를 나타낼 수 있습니다. * 기호를 사용하여 여러 문자를 나타낼 수 있습니다.		
확인	취	소

	A	В	С	D	E	F	G
1	학번 💌	이름 💌	생년월일 💌	성별 ▼	본적 💌	소득 🖅	
4	3456	: 박윤정	1965-09-07	여	수원	3500	
5	3455	: 윤석호	1974-12-07	남	청주	2500	
10	2345	홍길동	1973-10-22	남	서울	3800	
11							

제주대학교 표. R 데이터베이스

1. filter()

- 특정 조건을 만족시키는 데이터를 검색(추출)
- 예 1 : 남학생만 검색
- 예 2 : 남학생이면서 본적이 서울
- 예 3 : 소득이 2000보다 많고 4000미만인 학생 검색

제주 4000 8 6543 최정민 26390 여 9 2345 홍길동 26959 남 서울 3800 b3-ch2-2.R > data %>% filter(data\$성별 library(openxlsx) "남") == 생년월일 성별 본적 소득 학버 이름 library(dplyr) 남 부산 1000 1 6789 김상욱 25609 남 남 청주 2500 광주 1500 2 3455 윤석호 27370 data<-read.xlsx("http://kanggc.iptime.org/book/data/dbase.xlsx") 1234 니경규 24838 남 인천 5000 data -4 4546 이동균 25346 남 경기 8000 5 2341 조동민 24714 encoding = "UTF-8"남 서울 3800 6 2345 홍길동 26959 data %>% filter(data\$성별 == "남") data %>% filter(data\$성별 == "남" & data\$본적 == "서울") > data %>% filter(data\$성별 == "남" & data\$본적 == "서울") data %>% filter(data\$소득 > 2000 & data\$소득 < 4000) * 이름 생년윌일 성별 본적 소득 학번 남 서울 3800 1 2345 홍길동 26959 (참고) 논리연산자 &:그리고(and) > data %>% filter(data\$소득 > 2000 & data\$소득 < 4000) 이름 생년월일 성별 본적 소득 학번 여 수원 3500 남 청주 2500 |: 또는(or) 3456 박윤정 23992 2 3455 윤석호 27370 !:부정 남 서울 3800 3 2345 홍길동 26959 >, <, <=, >= : 관계연산자

> data

학변

1 6789 김상욱

2 2346 김숙희

3 3456 박윤정

5 1234 니경규

6 4546 이동균

7 2341 조동민

3455 윤석호

이름

생년월일 성별

25609

26021

23992

27370

24838

25346

24714

본적 소득

남

여

여

남남

남남

부산 1000

대구 2000

수원 3500

청주 2500

광주 1500

인천 5000

경기 8000

2. 실습(simulation 가상 데이터)	<pre>> mydata</pre>
dplyr-simul.R	3 1 15 43 75
library(dplyr)	5 1 20 60 70
set.seed(1234)	6 1 14 67 70 7 2 30 53 75
mydata<-data.frame(class = sample(3,10, replace=T),	8 2 24 68 78 9 3 30 63 56
id = sample(30,10, replace=T),	10 2 4 63 52
mid = as.integer(rnorm(10, mean=60, sd=10)),	<pre>> filter(mydata, class==1)</pre>
final = as.integer(rnorm(10, mean=70, sd=8)))	1 1 15 43 75
mydata	2 1 20 60 70
	<pre>> filter(mydata_class==3_id==30)</pre>
filter(mydata, class==1)	class id mid final
filter(mydata, class==3, id==30)	1 3 30 63 56
filter(mydata, class==1 class ==2)	<pre>> filter(mydata, class==1 class ==2)</pre>
filter(mydata, class %in% c(1,2))	class id mid final 1 2266870
filter(mydata, class %in% c(1,2), final==70)	2 2 6 73 69
filter(mydata, ((class==1) (class==2)) & final==70)	4 1 20 60 70
	5 1 14 67 70 6 2 30 53 75
<pre>> filter(mydata, class %in% c(1,2), f/nal==>0) class id mid final</pre>	7 2 24 68 78
1 2 26 68 70	8 2 4 63 52
2 1 20 60 70 3 1 14 67 70	<pre>> filter(mydata_class %in% s(1 2))</pre>
	class id mid final
<pre>> filter(mydata, ((class==1) (class==2)) & final==70 class_id_mid_final</pre>)) 3 1 15 43 75
1 2 26 68 70	4 1 20 60 70 5 1 14 67 70
2 1 20 60 70 3 1 14 67 70	6 2 30 53 75 7 2 34 68 78
	8 2 4 63 52





1. 부분합 계산

- 부분합 계산이란 특정한 필드를 기준으로 데이터를 그룹별로 분류
- 각 그룹별로 합계뿐만 아니라 평균, 최댓값, 최솟값, 표본표준편차, 표준편차, 표본분산, 분산, 숫자개수
 등 부분적인 계산을 자동으로 해 주는 기능을 말함

2. 실습

- 강기춘 홈페이지에서 subtotal.xlsx 파일을 다운로드
- 경영, 경제, 회계, 무역 등 4개학과
- 각 학과별 1-4학년 학생 20명

	Å	В	С	D	E	F	G
1	이름	출석	중간고사	기말고사	총점	학년	학과
2	강리라	9	44	18	42.8	1	경영
3	이동심	7	15	12	24.8	1	경영
4	강희영	9	75	76	78.4	1	경영
5	김민찬	7	62	76	69.2	1	경영
6	박자영	10	52	60	64.8	1	경영
7	이은영	6	16	20	26.4	2	경영
8	김은주	7	83	72	76	2	경영
9	김일원	9	16	78	55.6	2	경영
10	방인성	8	64	55	63.6	2	경영
11	문성철	5	42	67	53.6	2	경영
12	강지천	0	27	13	16	3	경영
13	박나한	9	73	65	73.2	3	경영
14	심은숙	5	40	70	54	3	경영
15	강민숙	7	30	64	51.6	3	경영
16	김영두	10	74	65	75.6	3	경영
17	강진혜	1	45	24	29.6	- 4	경영
18	강춘호	7	69	83	74.8	- 4	경영
19	이승원	10	35	58	57.2	- 4	경영
20	박주원	10	42	50	56.8	- 4	경영
21	김양옥	6	63	81	69.6	- 4	경영

- 예 1 : 학생 전체의 중간고사, 기말고사 및 총점의 평균을 계산 · =average(c2:c81)로 중간고사 평균을 계산하고 나머지도 동일한 방법으로 계산 82 50.3375 48.225 54.2

- 예 2 : 학년별 중간고사, 기말고사 및 총점의 평균을 계산
- · 학년 순으로 그룹을 묶을 경우 학년 필드에 셀 포인터를 위치
- · [홈]-[편집] 그룹에서 [정렬 및 필터]를 클릭하여 [숫자 오름차순 정리]를 선택
- · 학년 순으로 정렬되면 [데이터]-[윤곽선] 그룹에서 [부분합]을 클릭
- · 부분합 대화상자에 아래 왼쪽 그림과 같이 입력
- · 학년별 중간고사, 기말고사 및 총점의 평균을 계산

부분합	?	\times
그룹화할 항목(<u>A</u>):		
학년		\sim
사용할 함수(<u>U</u>):		
평균		\sim
부분합 계산 항목(<u>D</u>):		
☐ 출석 ☑ 중간고사		^
I☑ 기발고사 ☑ 총점		
다. 학년 학과		~
✓ 새로운 값으로 대치(C)		
□ 그룹 사이에서 페이지 나누기(P)	
☑ 데이터 아래에 요약 표시(<u>S</u>)		
모두 제거(<u>R</u>) 확인	취소	2

123		A	B	C	D	E	F G
	1	이름	출석	중간고사	기말고사	총점	학년 학과
[[·]]	2	강리라	9	44	18	42.8	1 경영
·	3	이동심	7	15	12	24.8	1 경영
·	4	강희영	9	75	76	78.4	1 경영
· · ·	5	김민찬	7	62	76	69.2	1 경영
·	6	박자영	10	52	60	64.8	1 경영
· ·	7	문봉기	10	51	91	76.8	1 경제
·	8	고이영	7	64	22	48.4	1 경제
· · ·	9	김철수	9	41	31	46.8	1 경제
· ·	10	오시창	8	61	43	57.6	1 경제
·	11	강라세	8	52	50	56.8	1 무역
· ·	12	김건엽	7	73	91	79.6	1 무역
·	13	백은희	7	54	47	54.4	1 무역
$ \cdot $	14	신진호	8	47	71	63.2	1 무역
·	15	김상대	10	60	80	76	1 회계
·	16	민양수	7	38	45	47.2	1 회계
· ·	17	고희선	5	37	34	38.4	1 회계
· · ·	18	박신정	6	49	49	51.2	1 회계
Ē	19			51,4706	52.7059	57.44	1 평균
Γ·]	20	이은영	6	16	20	26.4	2 경영
1.1.	21	김은주	7	83	72	76	2 경영
· ·	22	김일원	9	16	78	55.6	2 경영
· ·	23	방인성	8	64	55	63.6	2 경영
·	24	문성철	5	42	67	53.6	2 경영
· ·	25	강행준	7	57	71	65.2	2 경제
· ·	26	고영준	8	71	85	78.4	2 경제
·	27	정열순	9	40	44	51.6	2 경제
·	28	박호민	7	47	40	48.8	2 경제
·	29	박규성	3	29	57	40.4	2 경제
·	30	강경은	9	25	29	39.6	2 무역
· '	31	박신애	8	55	79	69.6	2 무역
· ·	32	이광림	9	72	68	74	2 무역
	33	명한숙	6	86	46	64.8	2 무역
· '	34	조재진	4	39	73	52.8	2 무역
· ·	35	원옥순	10	48	52	60	2 회계
1.	36	김숙희	9	41	40	50.4	2 회계
· ·	37	김자안	8	36	34	44	2 회계
· '	38	공병호	6	44	32	42.4	2 회계
	39	김성옥	5	48	31	41.6	2 회계
L È I	40			47.95	53.65	54.94	2 평균
I T · I	41	강지천	0	27	13	16	3 경영
	42	박나한	9	73	65	73.2	3 경영
	43	신은숫	5	40	70	54	3 2 9



- 예 3 : 학과별 중간고사, 기말고사 및 총점의 평균을 파악
 - · 학과 순으로 그룹을 묶을 경우 학과 필드에 셀 포인터를 위치
 - · [홈]-[편집] 그룹에서 [정렬 및 필터]를 클릭하여 [텍스트 오름차순 정리]를 선택
 - · 학과 순으로 정렬되면 [데이터]-[윤곽선] 그룹에서 [부분합]을 클릭
 - · 부분합 대화상자에 아래 왼쪽 그림과 같이 입력
 - · 학과별 중간고사, 기말고사 및 총점의 평균을 계산

부분합 ?	×
그룹화할 항목(<u>A</u>):	
학과	\sim
사용할 함수(<u>U</u>):	
평균	\sim
부분합 계산 항목(<u>D</u>):	
□ 이름 □ 축석	^
· · · · · · · · · · · · · · · · · · ·	
☑ 기월고자 ☑ 총점	
□ 학년	*
☑ 새로운 값으로 대치(<u>C</u>)	
🗌 그룹 사이에서 페이지 나누기(P)	
☑ 데이터 아래에 요약 표시(<u>S</u>)	
모두 제거(<u>R</u>) 확인	취소

123		A	В	С	D	Е	F	G	Н
	1	이름	출석	중간고사	기말고사	총점	학년	학과	
ГГС	2	강리라	9	44	18	42.8	1	경영	
	3	이동심	7	15	12	24.8	1	경영	
1 .	4	강희영	9	75	76	78.4	1	경영	
	5	김민찬	7	62	76	69.2	1	경영	
	6	박자영	10	52	60	64.8	1	경영	
1 .	7	이운영	6	16	20	26.4	2	경영	
	8	김은주	7	83	72	76	2	경영	
	ğ	긴억원	9	16	78	55.6	2	경영	
	10	바이서	8	64	55	63.6	2	겨여	
	11	무선첩	5	42	67	53.6	2	겨여	
	12	고 o 글 가 지 처	0	27	13	16	3	<u>ਕ</u> ਕ	
	13	바니하	ä	73	65	73.2	3	<u> </u>	
	14	지요스	5	40	70	54	3	<u> </u>	
	14	200 소	7	20	64	51 G	2	'0 0 74 04	
	10	기여도	10	74	65	75.6	2	'00 2104	
	17	11 이 ㅜ 가지 쉐	10	14	00	20.6	3	'0'0' 74'04	
	10	응안에 가운호	7	40	24	29.0	4	70-70 74 04	
	10	상준오	10	25	00	[4.0 [7.0	4	200	
	19	비궁권	10	30	00	51.2	4	73 73 74 04	
1.	20	막우원	10	42	50	50.0	4	73 73	
14.	21	김장족	6	40.00	01	69.6	4	33	547
匚 .	22		10	48.35	55.35	55.68		23	영판
1.	23	군종기	10	51	91	/0.0	1	경제	
11.	24	고미영	1	64	22	48.4	1	경제	
1 .	25	김철수	9	41	31	46.8	1	경제	
· · ·	26	오시장.	8	61	43	57.6	1	경제	
1 .	21	강행운	1	5/	/1	65.2	2	경제	
11.	28	고영운	8	/1	85	78.4	2	경제	
11.	29	성열순	9	40	44	51.6	2	경제	
· · ·	30	막호민	1	47	40	48.8	2	경제	
· · ·	31	막귀성	3	29	57	40.4	2	경제	
1 .	32	김수형	9	64	40	59.6	3	경제	
1 · ·	33	양은순	9	33	30	43.2	3	경제	
11.	34	김미수		46	20	40.4	3	경제	
1 .	35	김영수	9	37	34	46.4	3	경제	
1 .	36	고유경	9	/1	42	63.2	3	경제	
11.	37	우진호	6	60	45	54	3	경제	
1 ·	38	이금숙	10	66	50	66.4	4	경제	
1 .	39	이준용	10	72	55	70.8	4	경제	
1 ·	40	장호철	8	80	64	73.6	4	경제	
1 ·	41	강정희	10	63	39	60.8	4	경제	
IL.	42	고일동	1	53	27	34	4	경제	
	43			55.3	46.5	56.32		경제	평균
1 [• .	44	강라세	8	52	50	56.8	1	무역	
1 ·	45	김건엽	7	73	91	79.6	1	무역	
1 ·	46	백은희	7	54	47	54.4	1	무역	
1 ·	47	신진호	8	47	71	63.2	1	무역	
	48	강경은	9	25	29	39.6	2	무역	
	49	박신애	8	55	79	69.6	2	무역	

제조대학교 표. R 부분합 계산

1. summarise

- 학생 전체의 기초통계량을 계산
- 예 1 : 중간고사, 기말고사, 총점의 평균을 계산
- 예 2 : 중간고사, 기말고사, 총점의 표준편차를 계산

b3-ch2-3.R
library(openxlsx)
library(dplyr)
df<-read.xlsx("http://kanggc.iptime.org/book/data/subtotal-e.xlsx")
df
summary(df)
summarise(df,mean_mid = mean(mid), mean_final = mean(final), mean_total = mean(total))
summarise(df,sd_mid = sd(mid), sd_final = sd(final), sd_total = sd(total))

> summary(df)

name	attend	mid	final	total	class	dept
Length:80	Min. : 0.000	Min. :15.00	Min. : 4.00	Min. : 9.6	Min. :1.000	Length:80
Class :character	1st Qu.: 6.750	1st Qu.:39.00	1st Qu.:32.00	1st Qu.:45.7	1st Qu.:2.000	Class :character
Mode :character	Median : 8.000	Median :51.00	Median :46.00	Median :54.0	Median :3.000	Mode :character
	Mean : 7.388	Mean :50.34	Mean :48.23	Mean :54.2	Mean :2.575	
	3rd Qu.: 9.000	3rd Qu.:63.00	3rd Qu.:65.50	3rd Qu.:64.8	3rd Qu.:3.250	
	Max. :10.000	Max. :86.00	Max. :91.00	Max. :79.6	Max. :4.000	
<pre>> summarise(df,me mean_mid mean_f 1 50.3375 48</pre>	an_mid = mean(mid inal mean_total .225 54.2	d), mean_final =	mean(final), me	an_total = mear	n(total))	
<pre>> summarise(df,sd, sd_mid sd_fin;</pre>	_mid = sd(mid), s al sd_total	d_final = sd(fir	nal), sd_total =	sd(total))		
1 17.53907 21.216	77 15.14808					



2. summarise()(with group_by)

- 학년별, 학과별 등 그룹의 기초통계량을 계산
- 예 1 : 학년별 중간고사, 기말고사, 총점의 평균을 계산
- 예 2 : 학과별 중간고사, 기말고사, 총점의 평균을 계산



· 새로운 변수(열)을 생성 · 실습(simulation 데이터)			class "	id	mid	÷ final	\$	
- 세도군 친구(일)를 영영 - 신슈(simulation 데이터)		1	2	26	68	70		
· 실습(simulation 데이터)		2	2	6	73	69		
		3	1	15	43	75		
· 예 1 : 중간고사와 기말고사를 합한 총점 변수	만들기	4	3	14	53	66		
에 2, 초거의 이용법이 편구 법수 마트기		- 5	1	20	60	70		
· 에 2 : 공심을 이용하여 평균 면수 만들기		6	1	14	67	70		
dplyr-simul.R		7	2	30	53	75		
		8	2	24	68	78		
at seed(1234)		9	3	30	63	56		
$vdata \leq -data frame(class = sample(3.10 replace=T)$		10	2	4	63	52		
id = comple(20, 10, contace = T)		-	class 🗦	id [‡]	mid [‡]	final	† total	mean
nd = a integer(merm(10, mean=60, ad=10))		1	2	26	68	70	138	69.0
(10 - as integer(110111(10, 11ea11-60, sd - 10))),		2	2	6	73	69	142	71.0
tinal = as.integer(rnorm(10, mean=70, sd=8)))		3	1	15	43	75	118	59.0
ydata		4	3	14	53	66	119	59.5
utate(mydata, total=mid+final, mean=total/2)		5	1	20	60	70	130	65.0
ydata<-mutate(mydata, total=mid+final, mean=total/2)		0	2	30	53	70	128	64.0
		8	2	24	68	78	146	73.0
ansmute(mydata, total=mid+final, mean=total/2)		9	3	30	63	56	119	59.5
vdata trans $-$ transmute(mvdata total=mid+final maan=total/2)		10	2	4	63	52	115	57.5







제조대학교 함수를 이용한 그림그리기

1. 선 그래프







Consumption of Korea(2000-2016)







Scatter plot of GDP and Consumption





4. 상자그래프

b1-ch3-4-plot.R

library(openxlsx)

sample1<-read.xlsx("http://kanggc.iptime.org/book/data/stat-1.xlsx")</pre>

mid<-sample1\$mid

final<-sample1\$final

total<-sample1\$total

grade<-sample1\$grade

summary(sample1)

par(mfrow=c(1,3))

boxplot(mid, main="Box plot of mid")

boxplot(final, main="Box plot of final")

boxplot(total, main="Box plot of total")





Pie Chart of Total Score



6. 막대그래프	b1-ch3-6-plot.R
	library(openxlsx)
	sample1<-read.xlsx("http://kanggc.iptime.org/book/data/stat-1.xlsx")
	mid<-sample1\$mid
	final<-sample1\$final
	total<-sample1\$total
	grade<-sample1\$grade
	total
	grade
	counts<-table(total, grade)
	barplot(counts, main="Bar Chart of Total Score", xlab="Grade")



Bar Chart of Total Score

7. 잎-줄기 그래프

The decimal point is 1 digit(s) to the right of the







ggplot을 이용한 그림그리기



0. ggplot2 레이어 구조





- ggplot2는 데이터를 시각화하는 패키지로 보통 3단계로 구성되어 있음
 - · 1단계 : 배경 설정으로 데이터 축을 설정
 - · 2단계 : 그래프 추가(점, 막대, 선 등)
 - · 3단계 : 세부 설정 추가(축 범위, 색, 표식 등)
- ggplot2의 함수 구조의 예를 들면 다음과 같음 ggplot(data=data1, aes(x=var1, y=var2))+geom_point()+xlim(3,6) (1단계) (2단계) (3단계) ·1단계에서 data는 사용할 데이터, aes의 괄호 안은 x축 변수, y축 변수 · 2단계는 그래프의 종류를 나타내는데 주로 사용하는 종류는 다음과 같음 geom_point() : 산포도 geom_smooth() : 평활그래프 geom_bar() : 막대그래프(빈도 막대그래프로 X축만 설정) geom_col() : 막대그래프(집단간 차이를 나타냄) geom_boxplot() : 상자그래프 geom_histogram(): 히스토그램 geom line(): 선그래프 · 3단계에서 xlim의 괄호 안은 x축에 그릴 데이터의 범위를 나타냄

1. 선 그래프

```
b1-ch3-1-gaplot.R
library(openxlsx)
library(ggplot2)
library(gridExtra)
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/sample1-
n.xlsx")
year<-sample1$year
gdp<-sample1$gdp
consumption <- sample 1 $ consumption
gdp
consumption
qp1<-qplot(year, gdp, group=1, geom="line")</pre>
qp2<-qplot(year, consumption, group=1, geom="line")</pre>
marrangeGrob(grobs=list(qp1, qp2), nrow=2, ncol=1)
plot1<-ggplot(data=sample1, aes(x=year, y=gdp, group=1)) + geom
_line() + ggtitle("GDP of Korea(2000-2016)")+ theme(plot.title = ele
ment_text(hjust = 0.5))
plot2<-ggplot(data=sample1, aes(x=year, y=consumption, group=1)
) + geom_line() + ggtitle("Consumption of Korea(2000-2016)") + the
me(plot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot1, plot2), nrow=2, ncol=1)
```



2. 히스토그램

b1-ch3-2-ggplot.R library(openxlsx) library(ggplot2) library(gridExtra) sample1<-read.xlsx("http://kanggc.iptime.org/book/data/sa mple1-n.xlsx") year<-sample1\$year gdp<-sample1\$gdp consumption <- sample 1 \$ consumption gdp consumption qp1<-qplot(gdp, bins=8, geom="auto")</pre> qp2<-qplot(consumption, bins=8, geom="auto")</pre> marrangeGrob(grobs=list(qp1, qp2), nrow=2, ncol=1) plot1<-ggplot(data=sample1, aes(x=gdp)) + geom_histogra m(fill="white",bins=8) + ggtitle("Histogram of GDP")+ theme (plot.title = element_text(hjust = 0.5)) plot2<-ggplot(data=sample1, aes(x=consumption)) + geom _histogram(fill="red",bins=8) + ggtitle("Histogram of Consu mption") + theme(plot.title = element_text(hjust = 0.5)) marrangeGrob(grobs=list(plot1, plot2), nrow=2, ncol=1)



3. 산포도

b1-ch3-3-ggplot.R	1
library(openxlsx)	ç
library(ggplot2)	mptio
library(gridExtra)	consu
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/sample1-	6
n.xlsx")	
year<-sample1\$year	4
gdp<-sample1\$gdp	
consumption<-sample1\$consumption	
gdp	1
consumption	
plot1<-qplot(gdp, consumption, geom="auto")	ption
plot2<-ggplot(data=sample1, aes(x=gdp, y=consumption)) + geom_	unsuu
point(colour="red", size=2) + ggtitle("Scatter plot of GDP and Consu	86
mption")+ theme(plot.title = element_text(hjust = 0.5))	
marrangeGrob(grobs=list(plot1, plot2), nrow=2, ncol=1)	4



4. 상자그래프

b1-ch3-4-ggplot.R
library(openxlsx)
library(ggplot2)
library(gridExtra)
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/stat-1.xlsx")
mid<-sample1\$mid
final<-sample1\$final
total<-sample1\$total
grade<-sample1\$grade
qp1<-qplot(x=1, y=mid, geom="boxplot")
qp2<-qplot(x=1, y=final, geom="boxplot")
qp3<-qplot(x=1, y=total, geom="boxplot")
marrangeGrob(grobs=list(qp1, qp2, qp3), nrow=1, ncol=3)
plot1<-ggplot(data=sample1, aes(x=1, y=mid)) + geom_boxplot() + ggtitle("
Boxplot of Mid")+ theme(plot.title = element_text(hjust = 0.5))
plot2<-ggplot(data=sample1, aes(x=1, y=final)) + geom_boxplot() + ggtitle("
Boxplot of Final")+ theme(plot.title = element_text(hjust = 0.5))
plot3<-ggplot(data=sample1, aes(x=1, y=total)) + geom_boxplot() + ggtitle("
Boxplot of Total")+ theme(plot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot1, plot2, plot3), nrow=1, ncol=3)



5. 원그래프

b1-ch3-5-ggplot.R						
library(openxlsx)						
library(ggplot2)						
library(gridExtra)						
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/stat-1.xlsx")						
mid<-sample1\$mid						
final<-sample1\$final						
total<-sample1\$total						
grade<-sample1\$grade						
t.grade<-data.frame(table(sample1\$grade))						
ggplot(t.grade, aes(x="", y=Freq, fill=Var1)) + geom_bar(width=1, stat=						
"identity") + coord_polar(theta="y") + ggtitle("Pie Chart of Total Score"						
)+ theme(plot.title = element_text(hjust = 0.5))						



6. 막대그래프-1

b1-ch3-6-ggplot.R							
library(openxlsx)							
library(ggplot2)							
library(gridExtra)							
sample1<-read.xlsx("http://kanggc.iptime.org/book/data/stat-1.xl							
sx")							
mid<-sample1\$mid							
final<-sample1\$final							
total<-sample1\$total							
grade<-sample1\$grade							
ggplot(data=sample1, aes(x=grade)) + geom_bar() + ggtitle("Bar C							
hart of Grade")+ theme(plot.title = element_text(hjust = 0.5))							



7. 막대그래프-2

b1-ch3-7-ggplot.R library(openxlsx)

library(dplyr) library(ggplot2)

library(gridExtra)

df<-read.xlsx("http://kanggc.iptime.org/book/data/subtotal-e.xlsx") df

```
dept_name_1 <- df %>%
group_by(dept, class) %>%
summarise(mean_total = mean(total))
dept_name_1
```

plot1<-ggplot(data=dept_name_1, aes(x=dept, y=mean_total, fill=class
)) +</pre>

```
geom_col(position="dodge2") + ggtitle("Bar Chart of Total Mean by d
ept & class")+ theme(plot.title = element_text(hjust = 0.5))
plot2<-ggplot(data=dept_name_1, aes(x=class, y=mean_total, fill=dept
)) +</pre>
```

```
geom_col(position="dodge2") + ggtitle("Bar Chart of Total Mean by cl
ass & dept")+ theme(plot.title = element_text(hjust = 0.5))
marrangeGrob(grobs=list(plot1, plot2), nrow=2, ncol=1)
```







1. 피벗 테이블

- 부분합의 확장형이라고 할 수 있으며 다량의 데이터 목록을 요약해서 새로운 테이블을 구성
- X축과 Y축을 기준으로 복잡한 데이터를 쉽게 재배치함으로써 데이터의 식별과 분석이 용이
- 행 영역, 열 영역, 데이터 영역, 페이지 영역에 데이터베이스 필드를 위치시켜 작성

2. 실습

- 강기춘 홈페이지에서 subtotal.xlsx 파일을 다운로드
- 경영, 경제, 회계, 무역 등 4개학과
- 각 학과별 1-4학년 학생 20명

	A	В	C	D	E	F	G	
1	이름	출석	중간고사	기말고사	총점	학년	학과	
2	강리라	9	44	18	42.8	1	경영	
3	이동심	7	15	12	24.8	1	경영	
4	강희영	9	75	76	78.4	1	경영	
5	김민찬	7	62	76	69.2	1	경영	
6	박자영	10	52	60	64.8	1	경영	
7	이은영	6	16	20	26.4	2	경영	
8	김은주	7	83	72	76	2	경영	
9	김일원	9	16	78	55.6	2	경영	
10	방인성	8	64	55	63.6	2	경영	
11	문성철	5	42	67	53.6	2	경영	
12	강지천	0	27	13	16	3	경영	
13	박나한	9	73	65	73.2	3	경영	
14	심은숙	5	40	70	54	3	경영	
15	강민숙	7	30	64	51.6	3	경영	
16	김영두	10	74	65	75.6	3	경영	
17	강진혜	1	45	24	29.6	- 4	경영	
18	강춘호	7	69	83	74.8	- 4	경영	
19	이승원	10	35	58	57.2	- 4	경영	
20	박주원	10	42	50	56.8	- 4	경영	
21	김양옥	6	63	81	69.6	- 4	경영	

피벗 테이블 만들	7	?	×					
분석할 데이터를 선택하십시오.								
● 표 또는 범위 선택(S)								
표/범위①:	Sheet1!\$A\$1:\$G\$81		1					
○ 외부 데이터 원분	본 사용(<u>U</u>)							
연결 선택(C)							
연결 이름:			-					
피벗 테이블 보고서를	을넣을 위치를 선택하십시오.							
 새 워크시트(N) 								
○ 기존 워크시트(E))							
위치(<u>L</u>):			.					
여러 테이블을 분석할 것인지 선택								
□ 데이터 모델에 이 데이터 추가(<u>M</u>)								
	확인	취	소 					

- · 피벗 테이블 만들기 대화상자가 나타남
- · 학과에 셀 포인터를 두고 [삽입]-[표] 그룹에서 피벗 테이블을 선택
- · 학과 순으로 그룹을 묶어야 되는데 이미 정렬이 되어 있음
- 예 1 : 학과별-학년별 총점의 평균을 계산

_

· 확인을 클릭하면 피벗 테이블 만들기 화면이 나타남

· 피벗 테이블 필드 목록에서 보고서에 추가할 필드를 페이지 필드, 행 필드, 열 필드, 데이터 필드로 드래 그 하면 피벗 테이블이 만들어 짐



- · 학과를 행 필드, 학년을 열 필드, 총점을 데이터 필드(Σ 값)로 드래그
- ·데이터 필드의 합계:총점 오른 쪽에 있는 ▼를 클릭하고 하단에 나오는 값 필드 설정을 선택하면 값 필드 설정 대화상자가 나타남
- 선택한 필드의 데이터 항목에서 평균을 선택하고 확인을 누르면 피벗 테이블이 만들어 짐

값 필드 설정	? ×	-						
원본 이름: 총점			A	В	С	D	Е	F
사용자 지정 이름(<u>C</u>): 합계 : 총점		1						
값 요약 기준 값 표시 형식		3	평균 : 총점	열 레이블 💌	-			1 -1
값 필드 요약 기준(<u>S</u>)		4	행 레이블 ▼ 계여	1	55.04	54.09	4 ·	종합계
요약에 사용할 계산 유형을 선택하십시오. 선택한 필드의 데이터		6	[3 8] 경제 모여	50 57.4	56.88	51.13333333	61.12	56.32
합계 ^		8	 회계	53.2	47.68	50,00000007	59.30 51.04	50.62
개수 평균 치대가		9	총합계	57.43529412	54.94	48.48695652	57.28	54.2
최 4 값 곱		10						
표시 형식(<u>N</u>) 확인	취소							

- 예 2 : 학과별-학년별 총점의 표본표준편차를 계산
- · 학과 순으로 그룹을 묶어야 되는데 이미 정렬이 되어 있음
- · 학과에 셀 포인터를 두고 [삽입]-[표] 그룹에서 피벗 테이블을 선택하면 피벗 테이블 만들기 대화상자 가 나타나고 확인을 클릭하면 피벗 테이블 만들기 화면이 나타남
- · 피벗 테이블 필드 목록에서 학과를 행 필드, 학년을 열 필드, 총점을 데이터 필드(Σ 값)로 드래그
- ·데이터 필드의 합계:총점 오른 쪽에 있는 ▼를 클릭하고 하단에 나오는 값 필드 설정을 선택하면 값 필드 설정 대화상자가 나타남
- · 선택한 필드의 데이터 항목에서 표본표준편차를 선택하고 확인을 누르면 피벗 테이블이 만들어 짐



2. 피벗 차트

- 예 1: 학과별-학년별 총점의 평균을 그림

· 학과 순으로 그룹을 묶어야 되는데 이미 정렬이 되어 있음

· 학과에 셀 포인터를 두고 [삽입]-[차트] 그룹에서 피벗 차트를 선택하면 피벗 차트 만들기 대화상자가 나타나고 확인을 클릭하면 피벗 차트 만들기 화면이 나타남





- · 피벗 차트 필드 목록에서 보고서에 추가할 필드를 선택하는데 학과를 축(범주), 학년을 범례(계열), 총 점을 Σ 값 필드로 드래그
- · Σ 값 필드의 합계:총점 오른 쪽에 있는 ▼를 클릭하고 하단에 나오는 값 필드 설정을 선택하면 값 필드 설정 대화상자가 나타남
- · 선택한 필드의 데이터 항목에서 평균을 선택하고 확인을 누르면 피벗 차트가 만들어 짐



- 예 2 : 학과별-학년별 총점의 표본표준편차를 그림
- · 학과 순으로 그룹을 묶어야 되는데 이미 정렬이 되어 있음
- · 학과에 셀 포인터를 두고 [삽입]-[차트] 그룹에서 피벗 차트를 선택하면 피벗 차트 만들기 대화상자가 나타나고 확인을 클릭하면 피벗 차트 만들기 화면이 나타남
- · 피벗 테이블 필드 목록에서 보고서에 추가할 필드를 선택하는데 학과를 축(범주), 학년을 범례(계열), 총점을 Σ 값 필드로 드래그
- · Σ 값 필드의 합계:총점 오른 쪽에 있는 ▼를 클릭하고 하단에 나오는 값 필드 설정을 선택하면 값 필드 설정 대화상자가 나타남
- · 선택한 필드의 데이터 항목에서 표본표준편차를 선택하고 확인을 누르면 피벗 차트가 만들어 짐



J 제조대학교 Ⅱ. R 피벗 테이블 및 피벗 차트

1. 피벗 테이블

b3-ch2-5-rev.R library(openxlsx) library(dplyr) library(ggplot2) library(gridExtra) df<-read.xlsx("http://kanggc.iptime.org/book/data/subtotal-e.xlsx") df dept_name_1 <- df %>% group_by(dept, class) %>% summarise(mean_total = mean(total)) dept_name_1 dept_name_2 <- df %>% group_by(dept, class) %>% summarise(sd_total = sd(total)) dept_name_2 plot1<-ggplot(data=dept_name_1, aes(x=dept, y=mean_total, fill=class)) + geom_col(position="dodge2") plot2<-ggplot(data=dept_name_2, aes(x=dept, y=sd_total, fill=class) + geom_col(position="dodge2")

marrangeGrob(grobs=list(plot1, plot2), nrow=2, ncol=1)

	> (#) # (1 2	dept_nam A tibble Groups: dept c <i><chr> <</chr></i> 경영 경영	e_1 : 16 x dept lass mo db7> 1 2	3 [4] ean_total <i><db1></db1></i> 56 55.0
	3 4 5 7 8 9 10)경영에제제제제역 역	3 4 1 2 3 4 1 2	54.1 57.6 57.4 56.9 51.1 61.1 63.5 60.2
	11 12 13 14 15 16	-무역 모역계 회계 회계 회계	3 4 1 2 3 4	38.7 59.4 53.2 47.7 51. 51.0
	> (# / # (dept_name A tibble Groups: dept c 경영	e_2 : 16 x dept lass so db7> 1	3 [4] 4_total <i><db1></db1></i> 21.8
	2 3 4 5 6 7	경영 경영 경제 경제 경제	2 3 4 1 2 3	18.3 23.9 17.5 13.8 15.0 9.23
)	8 9 10 11 12 13 14	경제역역역역 무무역역 회계	4 1 2 3 4 1 2	15.9 11.4 14.0 21.0 10.2 16.1 7.71
	15	히계	2	11 /

16 회계

5.87

4

제주대학교 JEJU MATIONAL UNIVERSITY

2. 피벗 차트

b3-ch2-5-rev.R library(openxlsx) library(dplyr) library(ggplot2) 60 library(gridExtra) mean_total df<-read.xlsx("http://kanggc.iptime.org/book/data/subtotal-e.xlsx") Df 20 dept_name_1 <- df %>% 0 group_by(dept, class) %>% 경영 summarise(mean_total = mean(total)) 25 dept_name_1 20 dept_name_2 <- df %>% 15 10group_by(dept, class) %>% summarise(sd_total = sd(total)) dept_name_2 5-0 plot1<-ggplot(data=dept_name_1, aes(x=dept, y=mean_total, fill=class 경영)) + geom_col(position="dodge2") plot2<-ggplot(data=dept_name_2, aes(x=dept, y=sd_total, fill=class)) + geom_col(position="dodge2") marrangeGrob(grobs=list(plot1, plot2), nrow=2, ncol=1)

