

경영데이터분석기초

- SPSS, Excel 활용한 데이터 분석 -

유 진 호

jhyoo@smu.ac.kr

Intro

- 인사
- 강의 내용 소개
- 교재, 참고자료
- 평가방법

연구실 : (M관)월해관 (M316)

연락처 : 02-781-7574 | 세부전공 : 경영정보, 정보보호, 개인정보보호, 빅데이터분석

● 경력

- 회장, 한국정보통신보안윤리학회, (2023.01~)
- 편집위원장, <국방과 보안> 학술지 , 국방보안연구소, (2019.06~)
- LINC+사업단장, 상명대학교, (2019.03~2022.05)
- 산학연구처장, 상명대학교, (2019.03~2021.09)
- 산학협력단장, 상명대학교, (2019.03~2021.09)
- ISMS/ISMS-P(정보보호관리체계/개인정보보호관리체계) 인증심사위원회 위원, 한국인터넷진흥원, (2017.09~)
- 부회장, 한국신용카드학회, (2014.03~)
- 부회장, 한국정보처리학회, (2014.03~)
- 이사, 한국정보보호학회, (2013.03~)
- 단장, 한국인터넷진흥원, (2004.10~2013.02)
- Managing Consultant, IBM, (2000.01~2004.09)
- 연구원, 한국전자통신연구원, (1993.11~1999.12)

학습목표 및 내용

- 데이터 분석에 대해 이해하고
- 데이터 분석과 실습을 통해 데이터 분석 역량을 제고한다.
- *수강생들이 실습할 수 있도록 도와주고 실습과제를 수행하도록 지원하는 수업

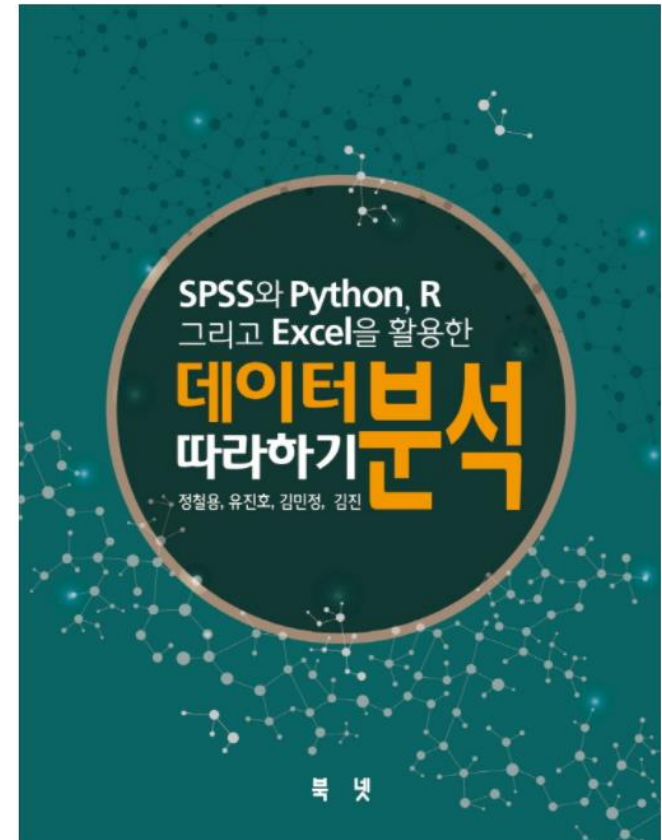
- 데이터 분석을 위해 Excel과 SPSS 통계소프트웨어를 가지고 분석한다.
- Excel 데이터를 SPSS로 불러서 다양한 통계 분석을 실시하고,
- SPSS script를 활용하여 통계분석을 하는 방법도 습득한다

- 현장에서 사용하는 데이터를 가지고
- 실습중심, 사례분석 중심으로 수업을 할 예정

노트북 활용 권장!

교재 및 참고자료

- **데이터분석 따라하기(SPSS와 Python, R 그리고 Excel을 활용한)**
 - 저자: 정철용, 유진호, 김민정, 김진
 - 출판사: 북넷



SPSS, Excel 중심으로 수업 예정

Part 1. SPSS를 활용한 데이터분석

1. SPSS 시작하기
2. 데이터 불러와서 분석하기
3. 기술통계량 분석
4. 상관분석
5. 교차분석(카이제곱 검정)
6. 평균차이 검정
7. 회귀분석
8. 로지스틱 회귀분석
9. 의사결정나무 분석
10. SPSS Syntax 사용하기

Part 2. 파이썬(Python)을 활용한 데이터분석

1. 파이썬 설치하기
2. 파이썬 기본문법
3. 데이터 불러오기
4. 기초통계 분석하기
5. plot 하기
6. 교차분석하기
7. 상관분석하기
8. 평균차이 검정하기
9. 회귀분석하기
10. 로지스틱 회귀분석하기
11. 의사결정나무 분석하기
12. 데이터 Set Merge하기

Part 3. Excel을 활용한 데이터 분석

1. 조건부 서식을 활용하여 데이터 이해하기
2. 상대참조, 절대참조 이용하여 새로운 값 만들기
3. 통계함수 이용하기
4. 논리함수(IF) 이용하기
5. 날짜함수 이용하기
6. 텍스트함수 이용하기
7. VLOOKUP, HLOOKUP
8. 수학함수 이용하기
9. 재무함수 이용하기
10. 차트활용하기
11. 데이터정렬 및 필터링
12. 피벗

외향



수리



글쓰기

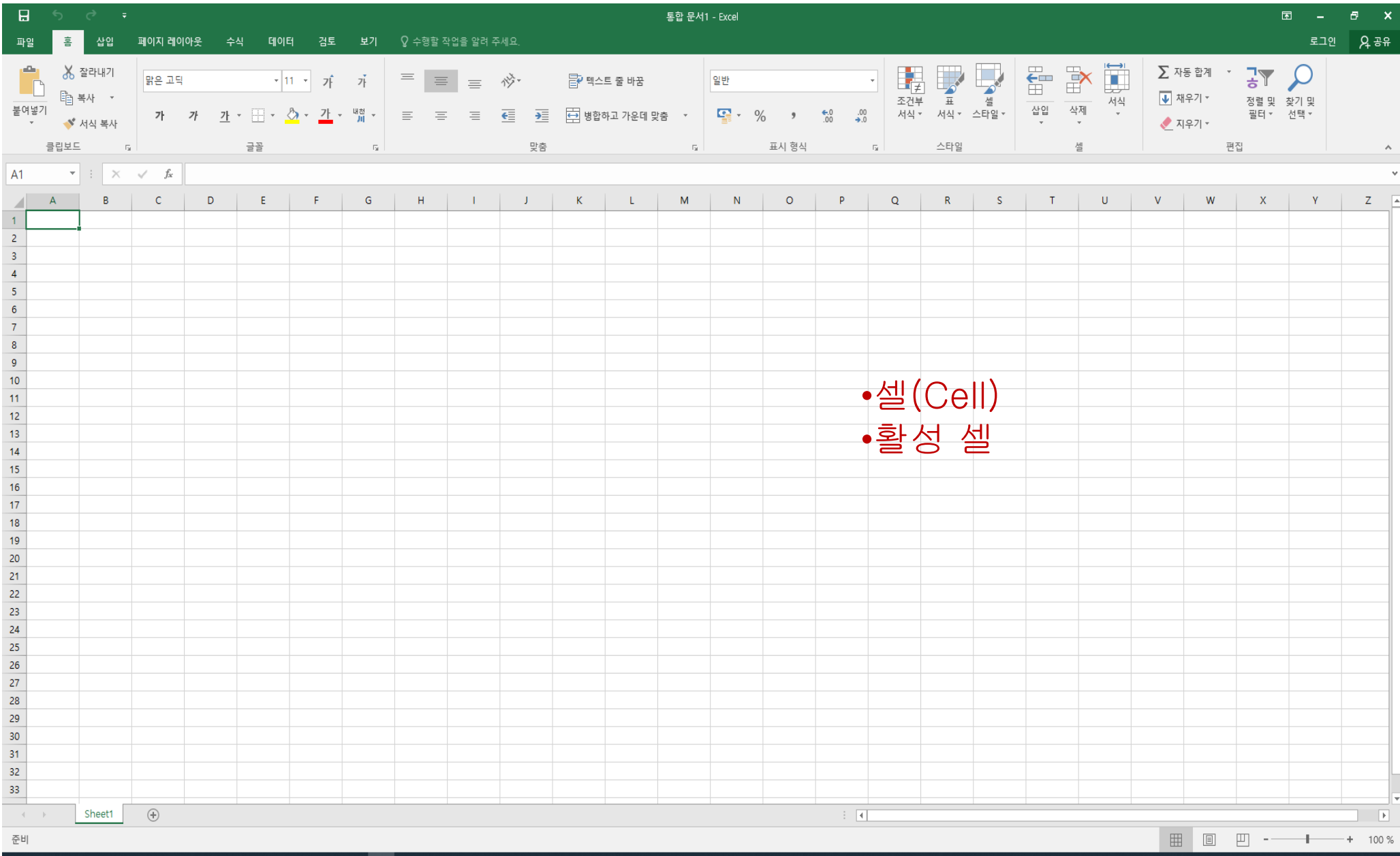
내향

평가방법

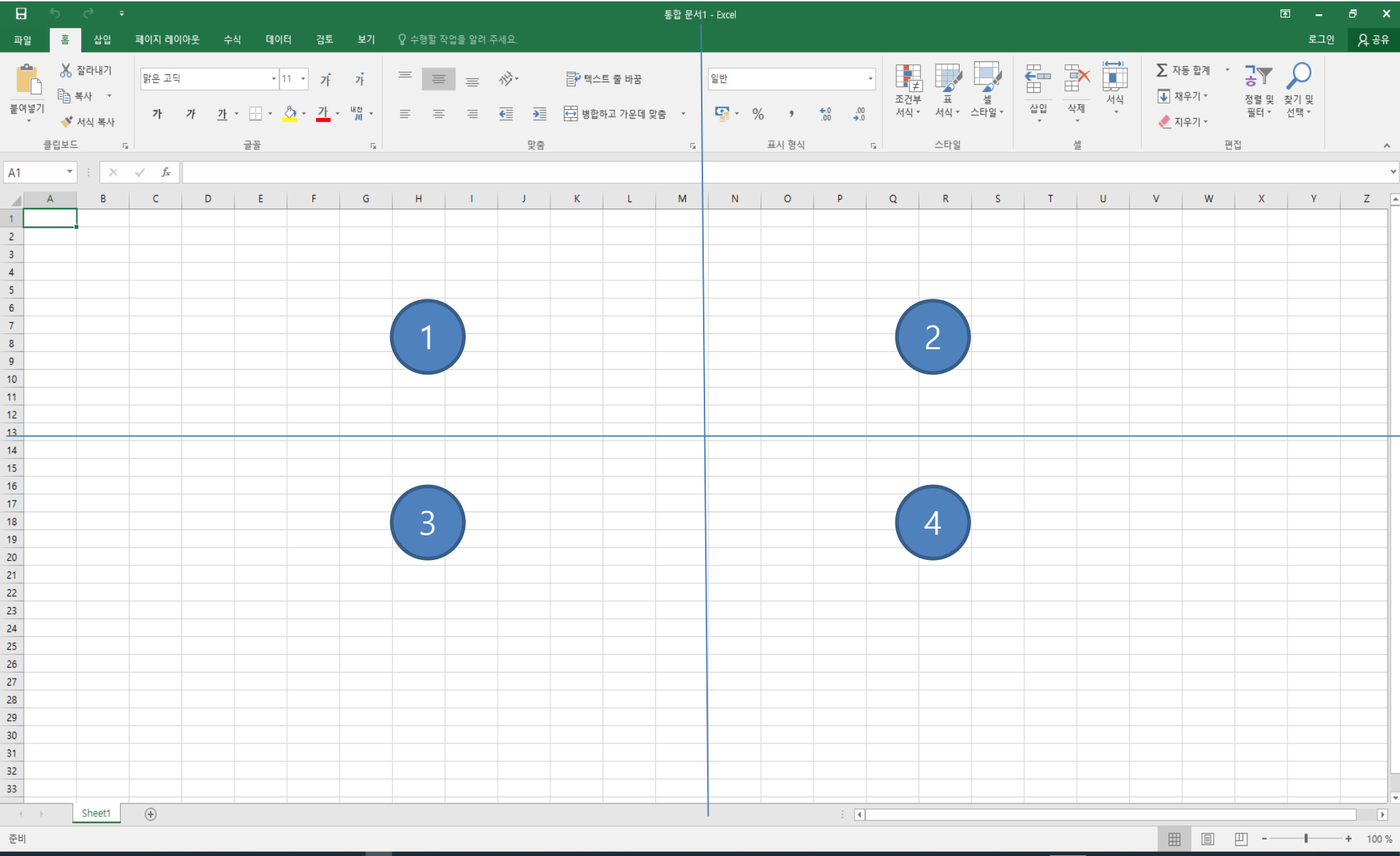
- **과제(미션): 100%**
 - 수시과제: 30%
 - 중간과제: 30%
 - 기말과제: 40%
- **출석: 학교지침에 따라 $\frac{1}{4}$ 이상 결석시 F**

Q&A

jhyoo@smu.ac.kr



- 셀(Cell)
- 활성 셀





파일

홈

삽입

페이지 레이아웃

수식

데이터

검토

보기

💡 수행할 작업을 알려 주세요.



붙여넣기



잘라내기



복사



서식 복사

클립보드

맑은 고딕

11

가

가

가

가

가

가

가

가

가

내천

글꼴



텍스트 줄 바꿈



병합하고 가운데 맞춤

맞춤

A1



fx

A

B

C

D

E

F

G

H

I

J

K

L

M

1

2

3

4

5

6

7

8

9

10

11

12

13

1

일반



%

,

←.0
.00

.00
→.0

조건부
서식 ▼

표
서식 ▾



셀 스타일



삼입

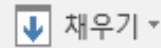


삭제


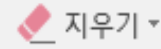


서식

Σ 자동 합계 ▼



↓ 채우기 ▾

 지우기 ▾정렬 및
필터 >

찾기 및 선택 ▼

^

Z

2

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

3

4



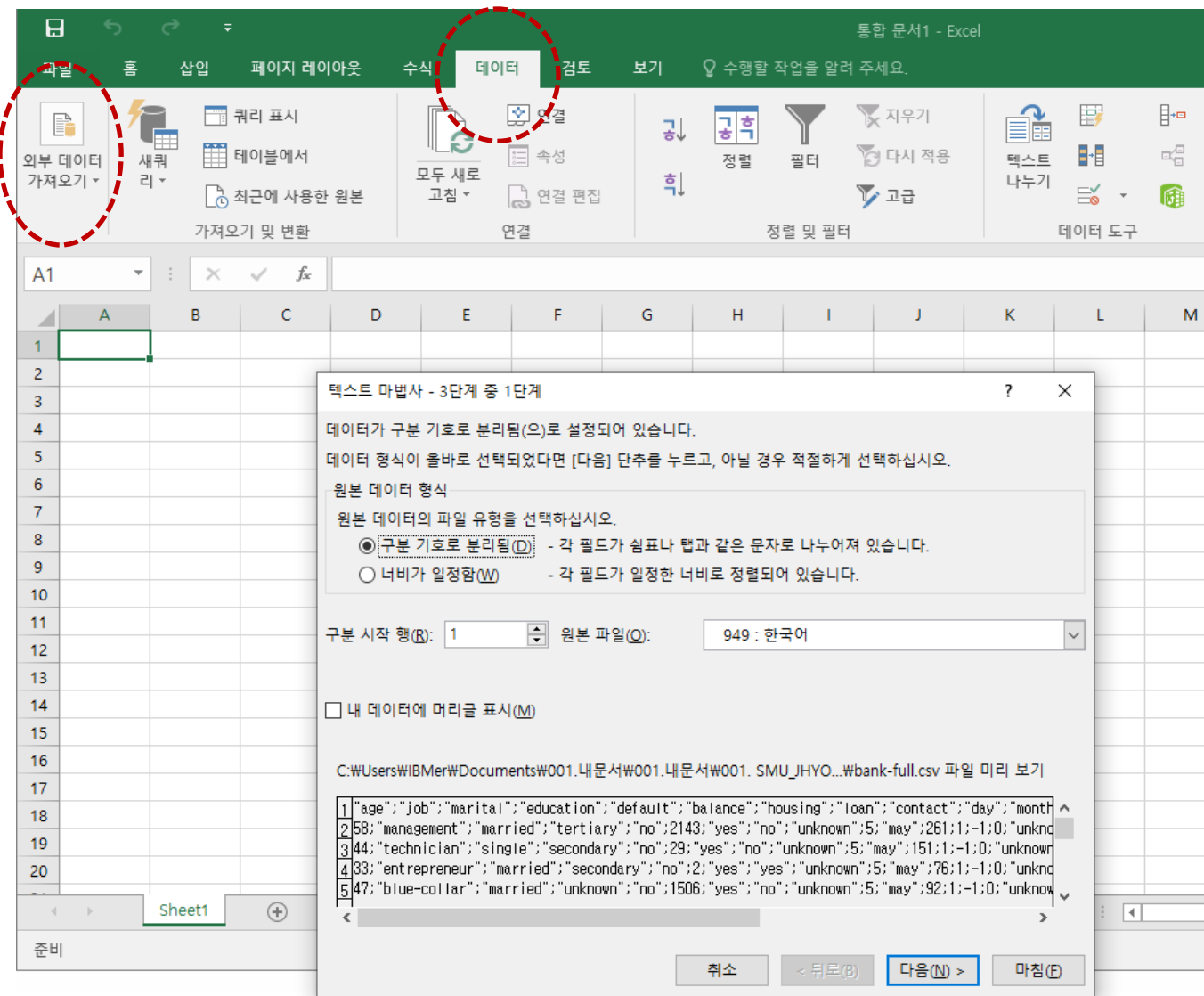
100 %

데이터 파일 미리보기

bank.csv - Windows 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

```
"age";"job";"marital";"education";"default";"balance";"housing";"loan";"contact";"day";"month";"duration";"campaign";"pdays";"previous";"poutcome";"y"
30;"unemployed";"married";"primary";"no";1787;"no";"no";"cellular";19;"oct";79;1;-1;0;"unknown";"no"
33;"services";"married";"secondary";"no";4789;"yes";"yes";"cellular";11;"may";220;1;339;4;"failure";"no"
35;"management";"single";"tertiary";"no";1350;"yes";"no";"cellular";16;"apr";185;1;330;1;"failure";"no"
30;"management";"married";"tertiary";"no";1476;"yes";"yes";"unknown";3;"jun";199;4;-1;0;"unknown";"no"
59;"blue-collar";"married";"secondary";"no";0;"yes";"no";"unknown";5;"may";226;1;-1;0;"unknown";"no"
35;"management";"single";"tertiary";"no";747;"no";"no";"cellular";23;"feb";141;2;176;3;"failure";"no"
36;"self-employed";"married";"tertiary";"no";307;"yes";"no";"cellular";14;"may";341;1;330;2;"other";"no"
39;"technician";"married";"secondary";"no";147;"yes";"no";"cellular";6;"may";151;2;-1;0;"unknown";"no"
41;"entrepreneur";"married";"tertiary";"no";221;"yes";"no";"unknown";14;"may";57;2;-1;0;"unknown";"no"
43;"services";"married";"primary";"no";-88;"yes";"yes";"cellular";17;"apr";313;1;147;2;"failure";"no"
39;"services";"married";"secondary";"no";9374;"yes";"no";"unknown";20;"may";273;1;-1;0;"unknown";"no"
43;"admin."; "married";"secondary";"no";264;"yes";"no";"cellular";17;"apr";113;2;-1;0;"unknown";"no"
36;"technician";"married";"tertiary";"no";1109;"no";"no";"cellular";13;"aug";328;2;-1;0;"unknown";"no"
20;"student";"single";"secondary";"no";502;"no";"no";"cellular";30;"apr";261;1;-1;0;"unknown";"yes"
31;"blue-collar";"married";"secondary";"no";360;"yes";"yes";"cellular";29;"jan";89;1;241;1;"failure";"no"
40;"management";"married";"tertiary";"no";194;"no";"yes";"cellular";29;"aug";189;2;-1;0;"unknown";"no"
56;"technician";"married";"secondary";"no";4073;"no";"no";"cellular";27;"aug";239;5;-1;0;"unknown";"no"
37;"admin."; "single";"tertiary";"no";2317;"yes";"no";"cellular";20;"apr";114;1;152;2;"failure";"no"
25;"blue-collar";"single";"primary";"no";-221;"yes";"no";"unknown";23;"may";250;1;-1;0;"unknown";"no"
31;"services";"married";"secondary";"no";132;"no";"no";"cellular";7;"jul";148;1;152;1;"other";"no"
38;"management";"divorced";"unknown";"no";0;"yes";"no";"cellular";18;"nov";96;2;-1;0;"unknown";"no"
42;"management";"divorced";"tertiary";"no";16;"no";"no";"cellular";19;"nov";140;3;-1;0;"unknown";"no"
44;"services";"single";"secondary";"no";106;"no";"no";"unknown";12;"jun";109;2;-1;0;"unknown";"no"
```



- CSV 파일을 엑셀에서 불러오기
- CSV[comma separated value]
 - 쉼표(comma)를 기준으로 항목을 구분하여 저장한 데이터,
 - 수많은 애플리케이션에서 취급하는 범용 형식,
 - CSV 형식의 파일은 텍스트 파일로 보존하여 문서 처리기나 편집기에서 열람·편집할 수 있음.

T16																	
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	outcome	y
2	30	unemployed	married	primary	no	1787	no	no	cellular	19	oct	79	1	-1	0	unknown	no
3	33	services	married	secondary	no	4789	yes	yes	cellular	11	may	220	1	339	4	failure	no
4	35	management	single	tertiary	no	1350	yes	no	cellular	16	apr	185	1	330	1	failure	no
5	30	management	married	tertiary	no	1476	yes	yes	unknown	3	jun	199	4	-1	0	unknown	no
6	59	blue-collar	married	secondary	no	0	yes	no	unknown	5	may	226	1	-1	0	unknown	no
7	35	management	single	tertiary	no	747	no	no	cellular	23	feb	141	2	176	3	failure	no
8	36	self-employed	married	tertiary	no	307	yes	no	cellular	14	may	341	1	330	2	other	no
9	39	technician	married	secondary	no	147	yes	no	cellular	6	may	151	2	-1	0	unknown	no
10	41	entrepreneur	married	tertiary	no	221	yes	no	unknown	14	may	57	2	-1	0	unknown	no
11	43	services	married	primary	no	-88	yes	yes	cellular	17	apr	313	1	147	2	failure	no
12	39	services	married	secondary	no	9374	yes	no	unknown	20	may	273	1	-1	0	unknown	no
13	43	admin.	married	secondary	no	264	yes	no	cellular	17	apr	113	2	-1	0	unknown	no
14	36	technician	married	tertiary	no	1109	no	no	cellular	13	aug	328	2	-1	0	unknown	no
15	20	student	single	secondary	no	502	no	no	cellular	30	apr	261	1	-1	0	unknown	yes
16	31	blue-collar	married	secondary	no	360	yes	yes	cellular	29	jan	89	1	241	1	failure	no
17	40	management	married	tertiary	no	194	no	yes	cellular	29	aug	189	2	-1	0	unknown	no
18	56	technician	married	secondary	no	4073	no	no	cellular	27	aug	239	5	-1	0	unknown	no
19	37	admin.	single	tertiary	no	2317	yes	no	cellular	20	apr	114	1	152	2	failure	no
20	25	blue-collar	single	primary	no	-221	yes	no	unknown	23	may	250	1	-1	0	unknown	no
21	31	services	married	secondary	no	132	no	no	cellular	7	jul	148	1	152	1	other	no
22	38	management	divorced	unknown	no	0	yes	no	cellular	18	nov	96	2	-1	0	unknown	no
23	42	management	divorced	tertiary	no	16	no	no	cellular	19	nov	140	3	-1	0	unknown	no
24	44	services	single	secondary	no	106	no	no	unknown	12	jun	109	2	-1	0	unknown	no

데이터 불러오면 행의 개수는? 4,522 (데이터는 4,521)

엑셀파일로 저장하기 (bank.xlsx)

데이터 도구

[illegible]



정보

저장

다른 이름으로
저장

인쇄

공유

내보내기

게시

닫기

다른 이름으로 저장



최근 항목

SangMyung University



OneDrive - SangMyung...
100587@sangmyung.kr



사이트 - SangMyung U...
100587@sangmyung.kr

기타 위치



이 PC



위치 추가



찾아보기



C: > Users > lenovo > OneDrive - SangMyung U...

bank

Excel 통합 문서 (*.xlsx)

[기타 옵션...](#)

새 폴더

여기에 표시할 항목을 찾을 수 없습니다.

Input variables:

bank client data:

1 – age (numeric)

2 – job : type of job (categorical: "admin.", "unknown", "unemployed", "management", "housemaid", "entrepreneur", "student", "blue-collar", "self-employed", "retired", "technician", "services")

3 – marital : marital status (categorical: "married", "divorced", "single"; note: "divorced" means divorced or widowed)

4 – education (categorical: "unknown", "primary", "secondary", "tertiary")

5 – default: has credit in default? (binary: "yes", "no"), 채무불이행

6 – balance: average yearly balance, in euros (numeric), 연간 평균 잔액

7 – housing: has housing loan? (binary: "yes", "no"), 주택담보대출

8 – loan: has personal loan? (binary: "yes", "no"), 개인신용대출

related with the last contact of the current campaign:

9 – contact: contact communication type (categorical: "unknown", "telephone", "cellular")

10 – day: last contact day of the month (numeric)

11 – month: last contact month of year (categorical: "jan", "feb", "mar", ..., "nov", "dec")

12 – duration: last contact duration, in seconds (numeric), 통화지속시간

other attributes:

13 – campaign: number of contacts performed **during this campaign** and for this client (numeric, includes last contact)

14 – pdays: number of days that **passed by** after the client was last contacted from a previous campaign (numeric, -1 means client was not previously contacted)

15 – previous: number of contacts performed **before this campaign** and for this client (numeric)

16 – poutcome: outcome of the previous marketing campaign (categorical: "unknown", "other", "failure", "success")

Output variable (desired target):

17 – y – has the client **subscribed a term deposit**? (binary: "yes", "no"), 정기에금에 가입하였나?

과제1

엑셀에서

- **Age, balance, duration 변수에 대한**
 - 각각의 총합(sum), 평균(Average), 최대값(Max), 최소값(Min)을 구하기
 - **교재(P.188~195) 참고하기**

=sum(A2:A4521)

Q&A

jhyoo@smu.ac.kr