[평생] 멀티미디어 일반형 양식

[원고] 빅데이터 수집 시스템 개발

14회차: 시계열 데이터 다루기

내용전문가	백현숫	교수설계(한기대)	홍길동
협력업체	업체명	교수설계(협력업체)	홍길동

NCS 분류 정보	20-01-02-09	정보통신 - 정보기술 - 정보기술개발 - 빅데이터플랫폼구축
능력단위 정보	03	빅데이터 수집시스템 개발
	2001020903_17v1.1	빅데이터 수집시스템 설계하기
느겁다니 이 시 거 버		
능력단위 요소 정보		

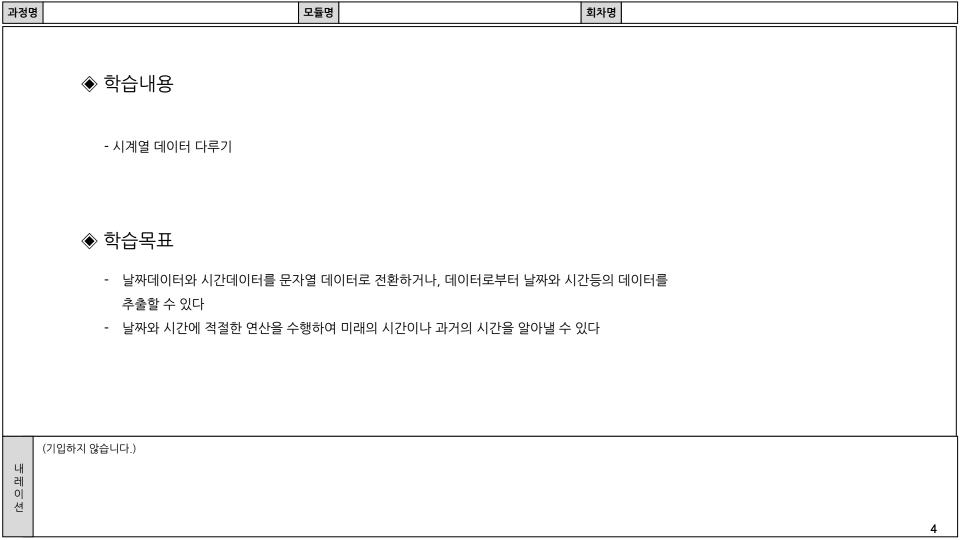
업무	작성자	버전	작성일	특이사항
원고 작성	백현숙	v.1.0	2019/09/11	2회차 원고
한기대 피드백		v.1.1		
원고 보완		v.2.0		
자문 진행		v.2.1		
원고 보완				

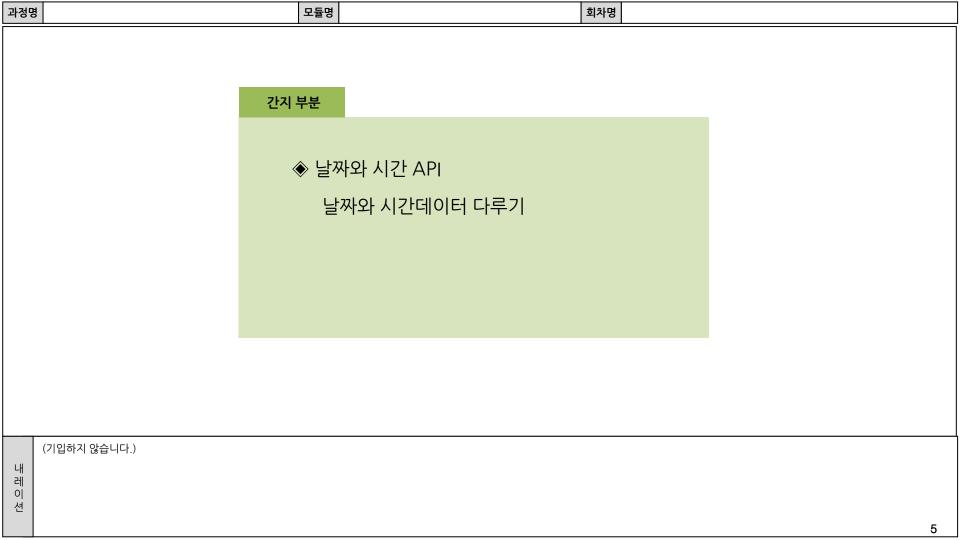


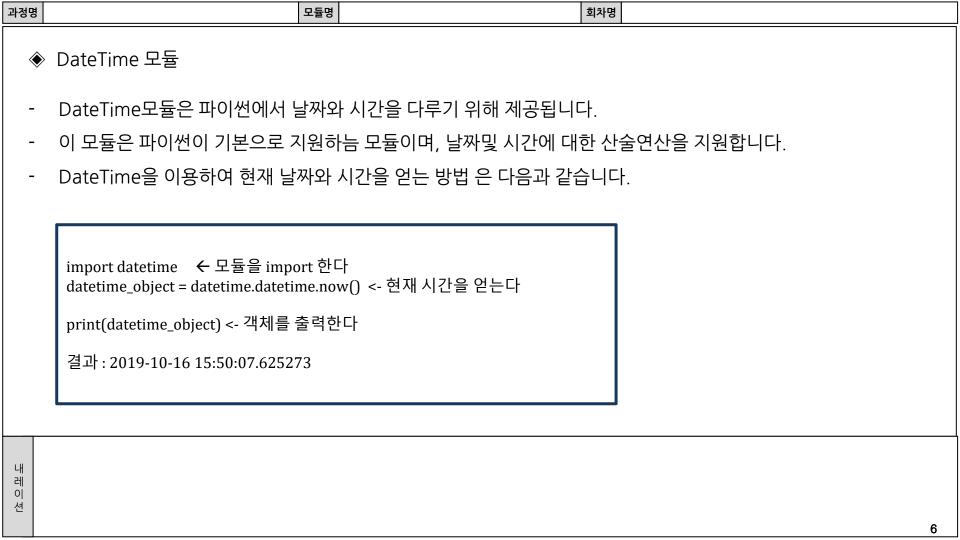
◆ 학습 목차 - 학습 목차를 작성해 주세요. (NCS 비적용 과정일 경우에는 NCS 및 능력단위 정보를 기입하지 않아도 됩니다.)

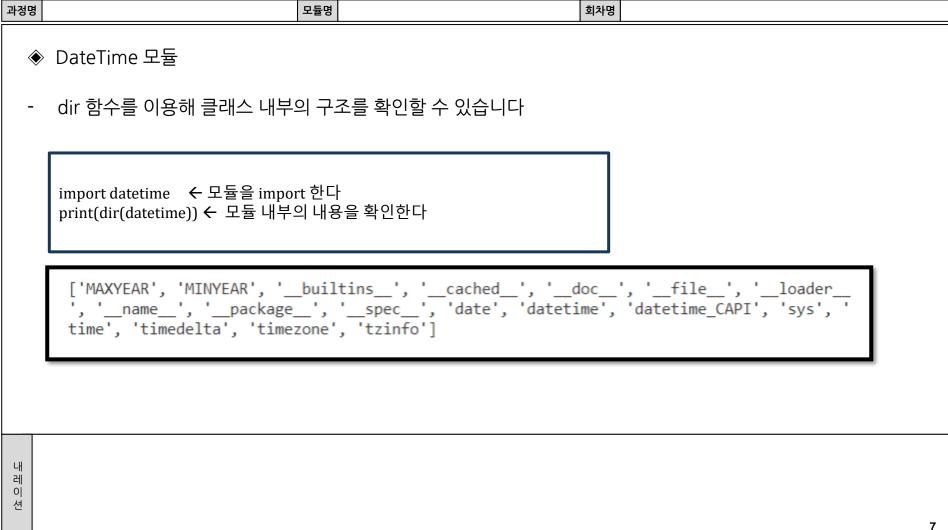
NCS 분류 정보	20-01-02-09	정보통신 - 정보기술 - 정보기술개발 - 빅데이터플랫폼구축
능력단위 정보	능력단위코드 기입	빅데이터 수집시스템 개발

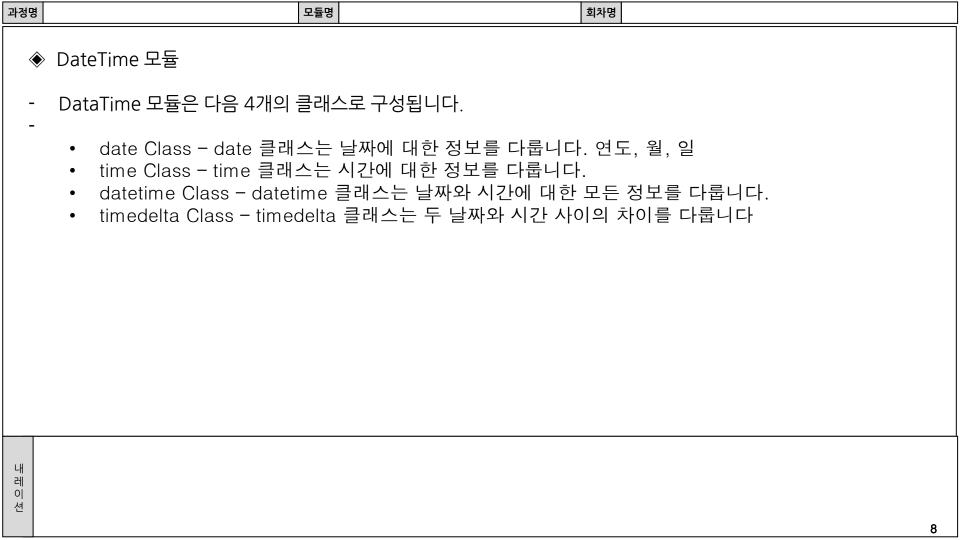
모듈명	회차명		능력단위 요소		수행준거
1. 모듈명 기입	01. 회차명 기입	1	능력단위요소명 기입	1.1	수행준거 기입
	02. 회차명 기입	1	능력단위요소명 기입	1.2	수행준거 기입
	03. 회차명 기입	1	능력단위요소명 기입	1.3 1.4	수행준거 기입 수행준거 기입
	04. 회차명 기입	1	능력단위요소명 기입	1.4	수행준거 기입
2. 모듈명 기입	05. 회차명 기입	2	능력단위요소명 기입	2.1 2.2	수행준거 기입 수행준거 기입
					2

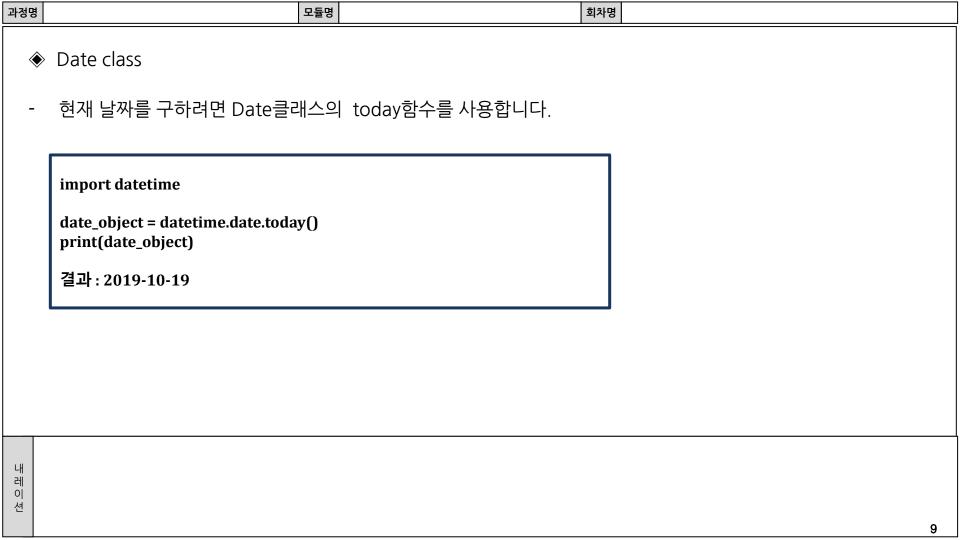


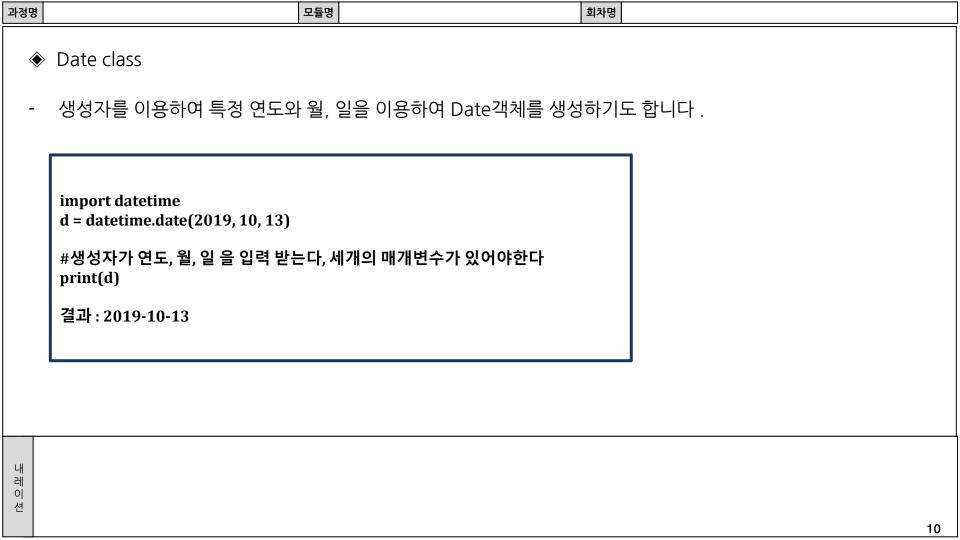


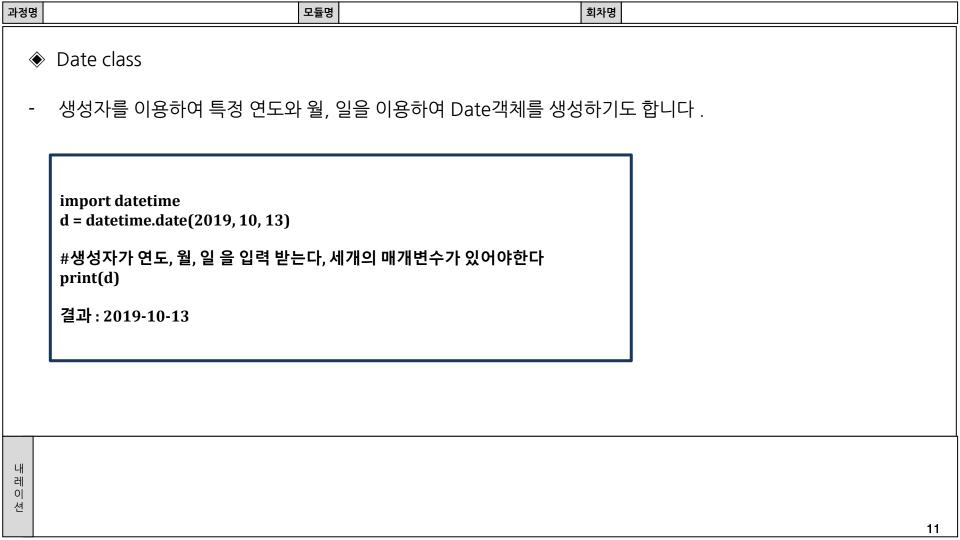


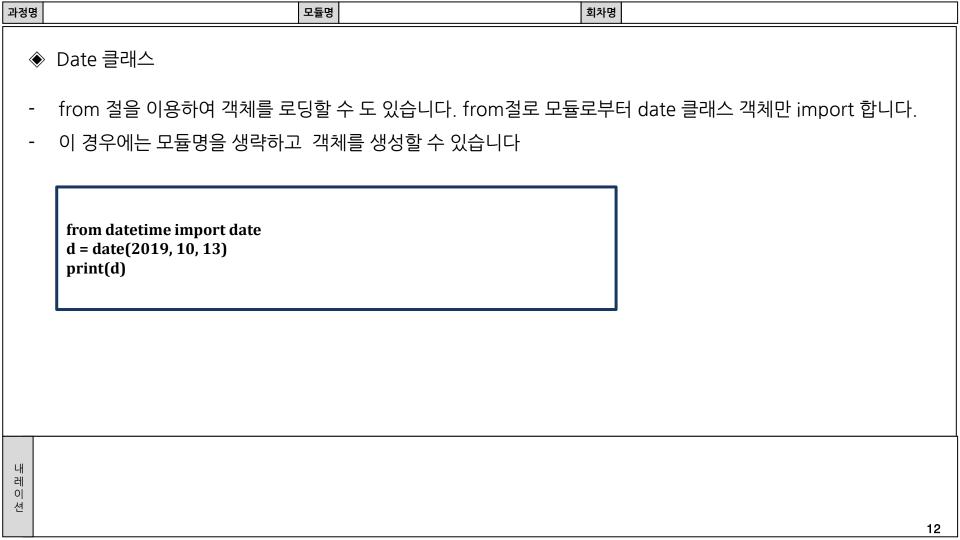




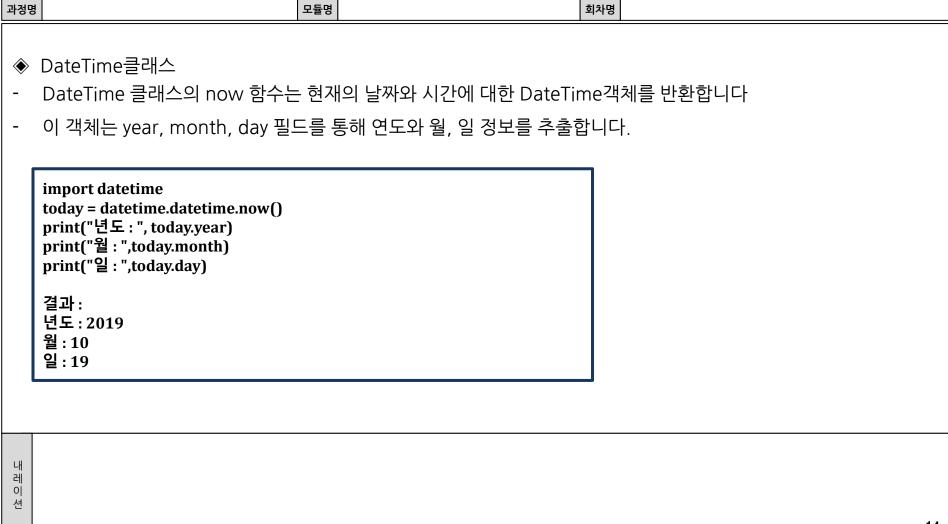


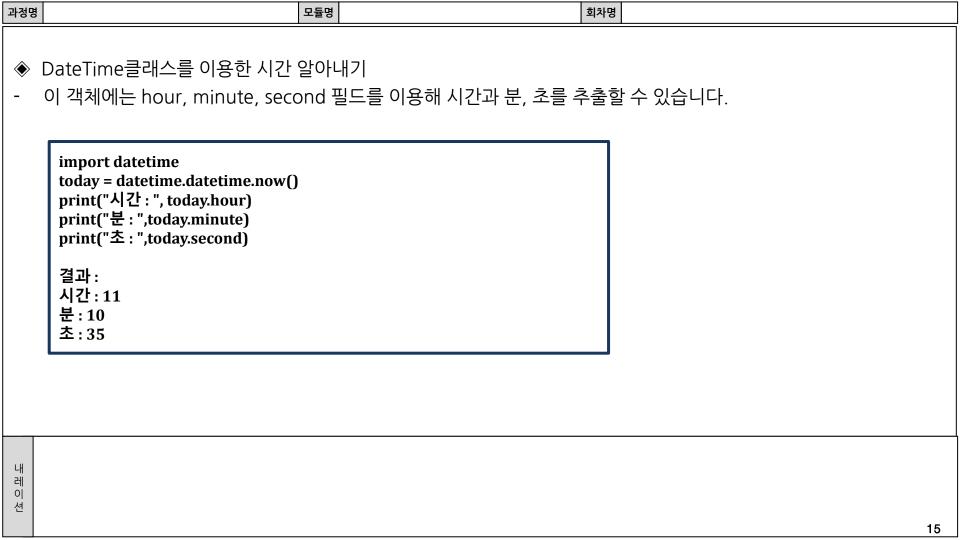






과정명	모듈명	회차명	
-	timestamp로 부터 날짜 정보 얻기 timestamp는 유닉스/리눅스 계열에서 사용하는 시간을 나타내는 정		
-	1970년1월1일 0시0분0초부터 몇초가 지났는 지를 나타내는 정수값이 타임스탬프는 지구자전축의 움직임을 고려하는 윤초(Leap Secon 숫자가 아니라 단순히 하루를 86400초로 계산하여 순차적으로 증가	nd)까	지 반영하여 아주 엄격하게 정의한
	from datetime import date value = 1567345678 timestamp = date.fromtimestamp(1567345678) print("timestamp =", 1567345678) print("Date =", timestamp)		
	결과 : timestamp = 1567345678 Date = 2019-09-01		
내 레 이 션			13





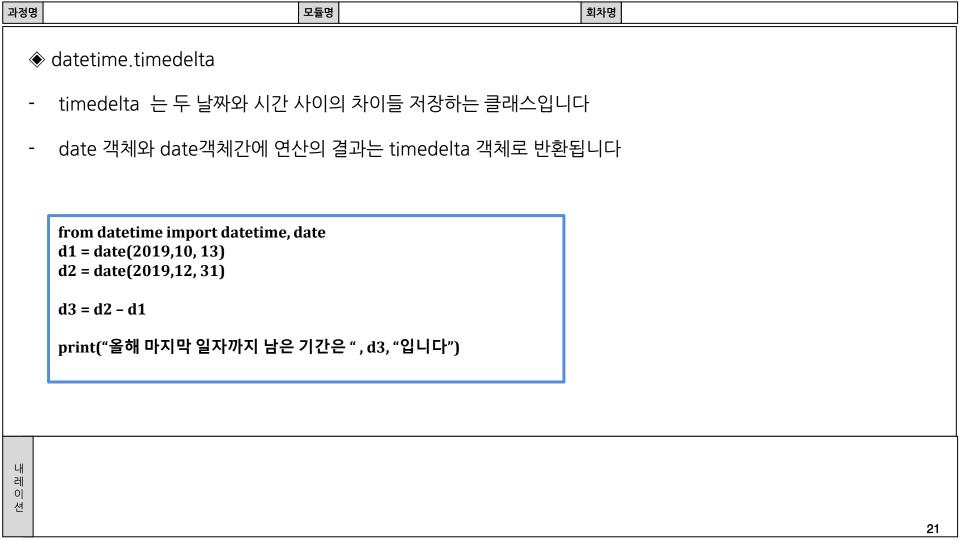
과정명		모듈명	회차명
-	DateTime클래스를 이용한 요일 DateTime클래스는 weekday라 weekday함수는 요일을 숫자로 t	는 함수를 제공합니다 -	
-	월요일:0, 화요일:1, 수요일:2, 목	요일:3, 금요일:4, 토요일:5 일요일:6	
	import datetime today = datetime.datetime.now() print("년도:", today.year) print("월:",today.month) print("일:",today.day) 결과: 년도:2019 월:10 일:19		
내 레 이 션			16

과정명	모듈명 회차명	
-	DateTime클래스를 이용한 시간 알아내기 DateTime객체에는 weekday()함수를 제공하는데 이 함수는 요일에 대한 정보를 반환합니다 월요일:0, 화요일:1, 수요일:2, 목요일:3, 금요일:4, 토요일:5 일요일:6	
	weekday = ["월요일", "화요일", "수요일", "목요일", "금요일", "토요일", "일요일"] today = datetime.datetime.now() print("요일 : ",today.weekday()) day = today.weekday() print("{0}-{1}-{2} 는 {3} 입니다".format(today.year, today.month, today.day, weekday[day]))	
내 레 이 션		17

과정	8	모듈명	회차명
	◈ datetime 모듈 예제 파일명 : exam14_1.py		
į	mport datetime		
1	orint("현재날짜와 시간 알아보기") coday = datetime.datetime.now() orint(today)		
l	#내부구조 확인-클래스의 내부구조를 print() print(dir(datetime)) print()	확인할 수 있다	
l	#현재날짜 orint() date_object = datetime.date.today() orint(date_object)		
내레			
이 선			
			18

과정당	3	모듈명	회차명	
(#연도, 월, 일 을 입력하여 date객체 상 d = datetime.date(2019, 10, 13) print(d)	!성하기		
f a	#from 절을 이용하여 모듈을 import 한 #이럴경우에는 모듈명을 사용하지 않고 rom datetime import date a = date(2019, 10, 13) print(a)	ŀ다 ፲ 클래스명만으로 객체 생성이 가능하다		
t r	#timestamp로 부터 날짜 추출하기 value = 1567345678 imestamp = date.fromtimestamp(15 print("timestamp =", 1567345678) print("Date =", timestamp)	567345678)		
내 레 이 션				
			19	ر

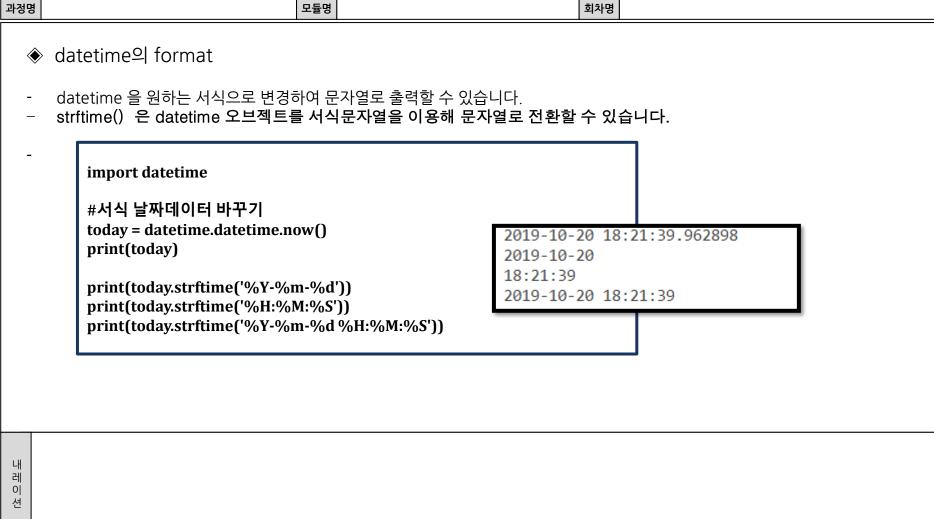
과정명	모듈명	회차명
#각 필드를 이용해 날짜와 시간 추출히 today = datetime.datetime.now() print(type(today)) print("년도:", today.year) print("월:",today.month) print("일:",today.day)	·기	
today = datetime.datetime.now() print("시간:", today.hour) print("분:",today.minute) print("초:",today.second)		
weekday = ["월요일", "화요일", "수요 today = datetime.datetime.now() print("요일:",today.weekday()) day = today.weekday()	목요일:3, 금요일:4, 토요일:5 일요일:6 .일", "목요일", "금요일", "토요일", "일요 rmat(today.year, today.month, today	
내 레 이 션		20

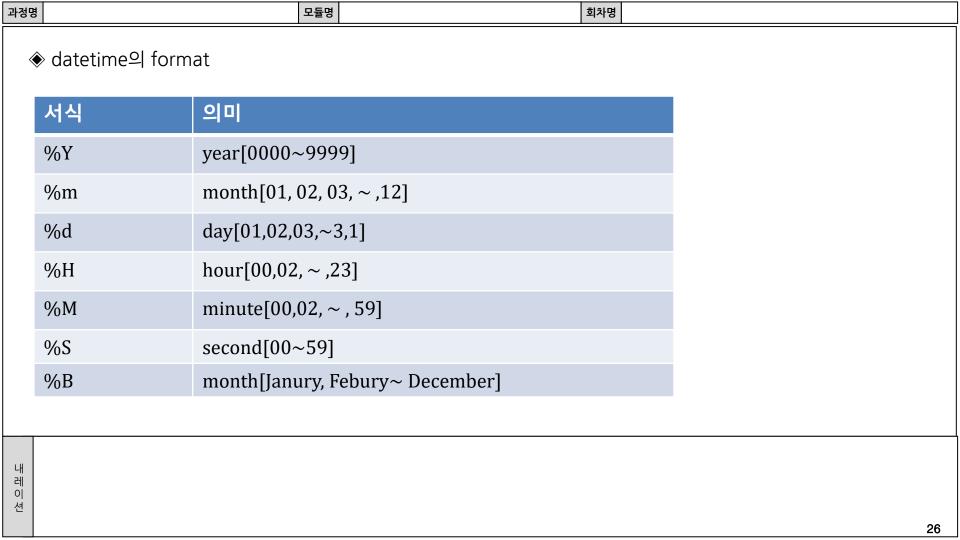


과정명		모듈명	회차명
	timedelta 의 인자값 1주: datetime.timedelta(weeks=1) 1일: datetime.timedelta(days=1) 1시간: datetime.timedelta(hours=1 1분: datetime.timedelta(minutes=1 1초: datetime.timedelta(seconds=1 1밀리초: datetime.timedelta(millise))) 1) econds=1)	회차명
내 레 이 션			22

과정명		모듈명		회차명	
1 *	⇒ datetime 모듈 예제 ト일명 : exam14_2.py				
d d: # d: p:	om datetime import datetime, date 1 = date(2019,10, 13) 2 = date(2019,12, 31) 날짜연산 3 = d2 - d1 rint("올해 남은날", d3) rint("type of d3 =", type(d3))	, time	type of d3 = t3 = 2:08:20	<cl></cl>	ays, 0:00:00 ass 'datetime.timedelta'> ass 'datetime.timedelta'>
t2 t3	= datetime(year = 2019, month = 2 = datetime(year = 2019, month = 3 = t1 - t2 rint("t3 =", t3) rint("type of t3 =", type(t3))				
내 레 이 션					

과정	경	모듈명	회차명
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	#timedelta 사이의 연산 t1 = timedelta(weeks = 2, days = 5, t2 = timedelta(days = 4, hours = 11, t3 = t1 - t2 print("t3 =", t3) t1 = timedelta(seconds = 33) #음수 t2 = timedelta(seconds = 54) t3 = t1 - t2 print("t3 = ", t3) print("t3 = ", t3) print("t3 = ", abs(t3)) #기간을 초로 나타내기 t = timedelta(days = 1, hours = 2, seconds("total seconds =", t.total_seconds") #지금 부터 3시간 뒤 current = datetime.today() after = current + timedelta(hours=3) print(current) print(after)	minutes = 4, seconds = 54) 로 나타날 수 도 있다 econds = 33, microseconds = 233)	t3 = 14 days, 13:55:39 t3 = -1 day, 23:59:39 t3 = 0:00:21 total seconds = 93633.000233 2019-10-20 01:10:24.044908 2019-10-20 04:10:24.044908
내 레 이 션			24

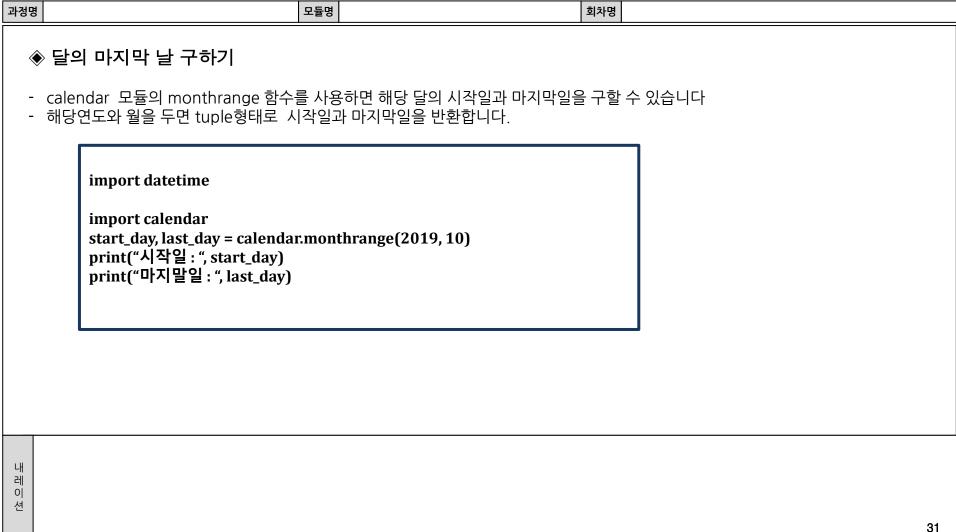




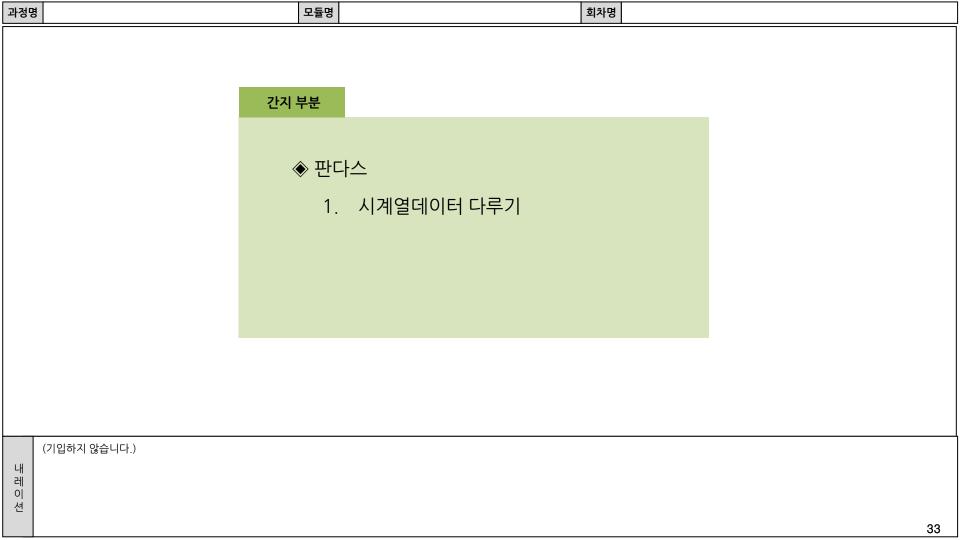
과정명		모듈명		회차명	
4	★ 서시과 타임존 예제				
#	파일명 :exam14_3.py				
ir	mport datetime				
t	:서식 날짜데이터 바꾸기 oday = datetime.datetime.now() rint(today)				
p	orint(today.strftime('%Y-%m-%d')) orint(today.strftime('%H:%M:%S')) orint(today.strftime('%Y-%m-%d %F	H:%M	:%S'))		
s #	orint(today.strftime('%Y-%m-%d %E Date = '2019-10-10 14:22:34' str형태의 날짜를 date 타입으로 전환 late1 = datetime.datetime.strptime(s print(type(date1)) print(date1)	하기	e, '%Y-%m-%d %H:% M :%S')		
내					
레 이 션					
					2

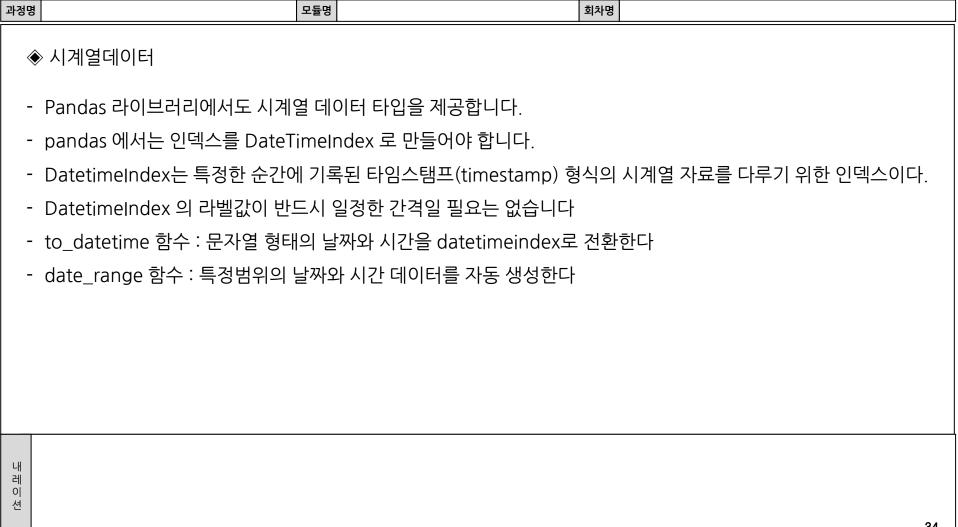
과정명	8	모듈명	회차명
f iii f r t t c c	#타임존 from datetime import datetime import pytz format = "%Y-%m-%d, %H:%M:%S" local = datetime.now() print("현재지역 :", local.strftime(formitz_NY = pytz.timezone('America/Newdatetime_NY = datetime.now(tz_NY) print("뉴욕시간 :", datetime_NY.strftintz_London = pytz.timezone('Europe/Ldatetime_London = datetime.now(tz_print("런던시간 :", datetime_London.strint("런던시간 :", datetime_London.strint("	nat)) v_York') me(format)) _ondon') _London)	
내 레 이 션			29

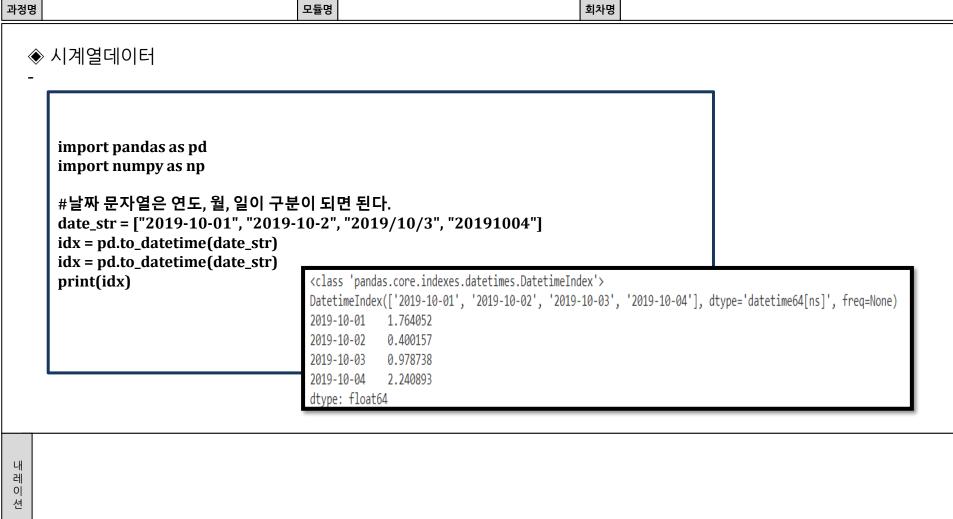
과정명		모듈명	회차명	
froin fcolo lo tz da po tz da	타임존 nport pytz prmat = "%Y-%m-%d, %H:%M:%Sizeal = datetime.now() rint("현재지역 :", local.strftime(forme.ny = pytz.timezone('America/Newatetime_NY = datetime.now(tz_NY) rint("뉴욕시간 :", datetime_NY.strftire_London = pytz.timezone('Europe/Latetime_London = datetime_London.strint("런던시간 :", datetime_London.strint("datetime_London.strint	at)) '_York') me(format)) _ondon') _London)		
내 레 이 션				30



과정명		모듈명	회차명
11	≫ 달의 마지막 날 구하기 예제		
Ī	다일명 : exam14_4.py		
	날짜 연산하기 nport datetime		
s	nport calendar tart_day, last_day = calendar.month rint(start_day) rint(last_day)	nrange(2019, 10)	
1	2019년 한해의 말일들을 출력합니다 or i in range(1,13): start_day, last_day = calendar.mo print(i,"월의 마지막날은 ", last_da		
내 레 이			
션			32

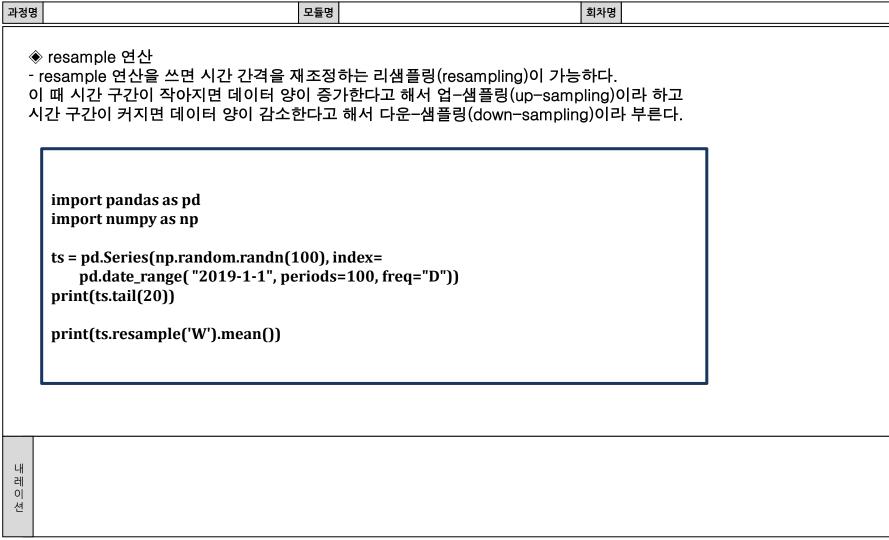




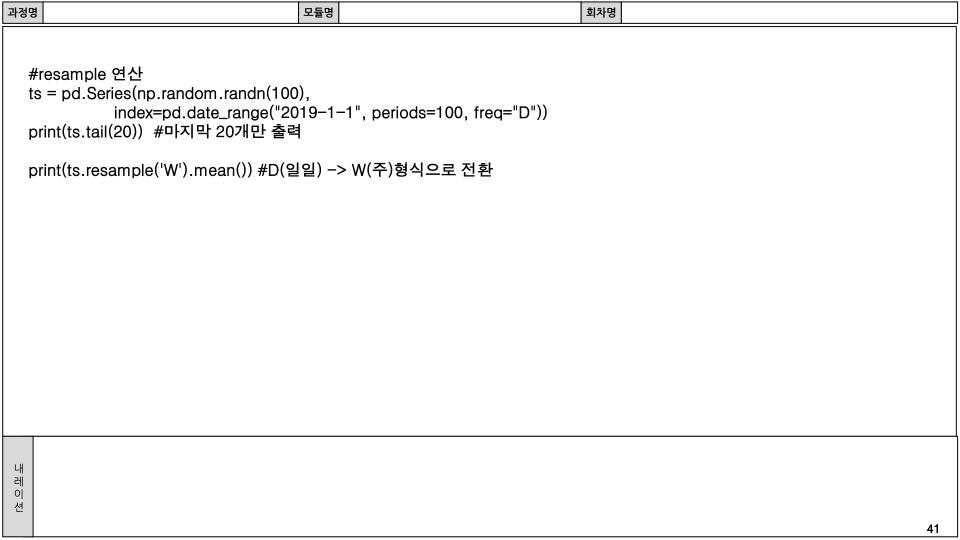


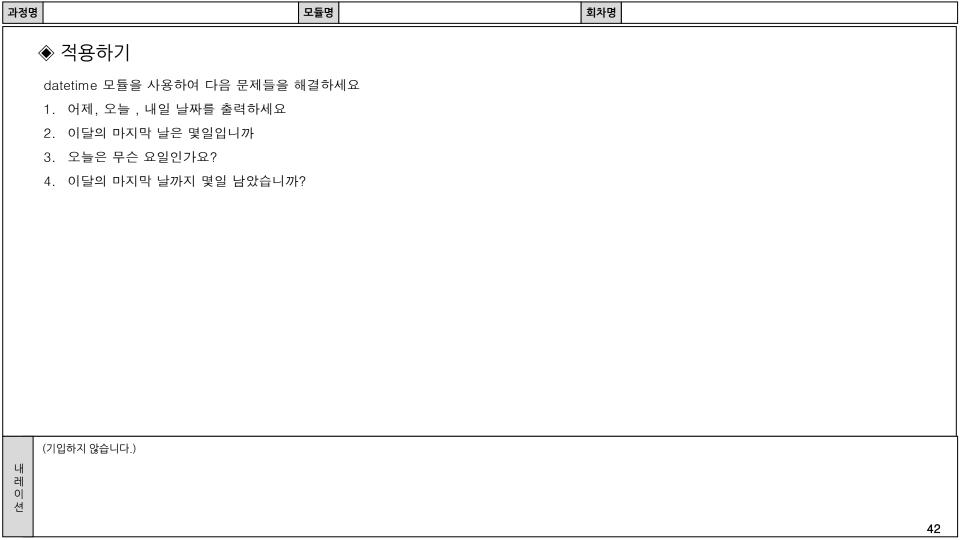
과정명 모듈명 회차명 ◆ to datetime 예제 파일명: exam14_5.py import pandas as pd import numpy as np #날짜 문자열은 연도-월-일이 구분이 되면 된다. date_str = ["2019-10-01", "2019-10-2", "2019/10/3", "20191004"] idx = pd.to_datetime(date_str) print(type(idx)) print(idx) <class 'pandas.core.indexes.datetimes.DatetimeIndex'> DatetimeIndex(['2019-10-01', '2019-10-02', '2019-10-03', '2019-10-04'], dtype='datetime64[ns]', freq=None) np.random.seed(0) 0.400157 0.978738 s = pd.Series(np.random.randn(4), index=idx) 2019-10-04 2.240893 dtype: float64 print(s) 1.764052345967664 0.9787379841057392 #인덱스를 이용하여 출력할 수 있다 2.240893199201458 print(s['2019-10-01']) print(s['2019-10-02'])



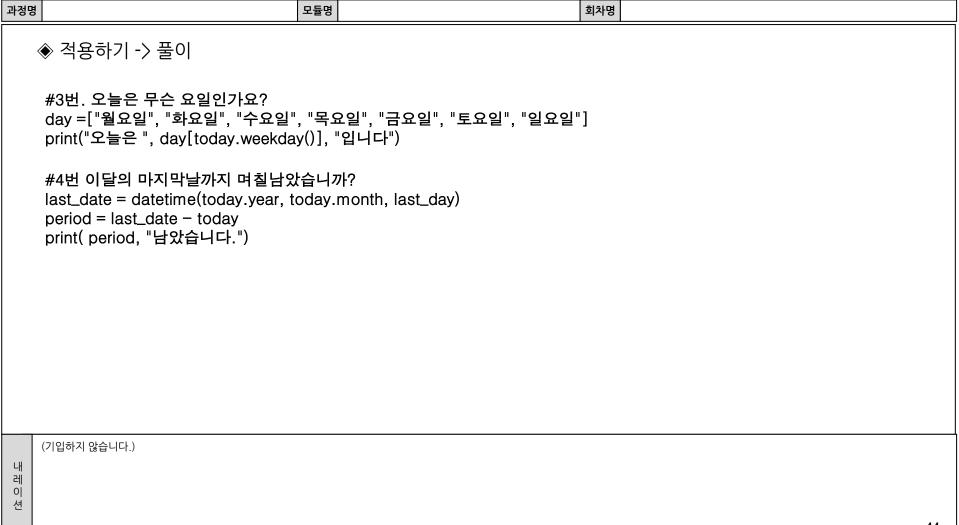


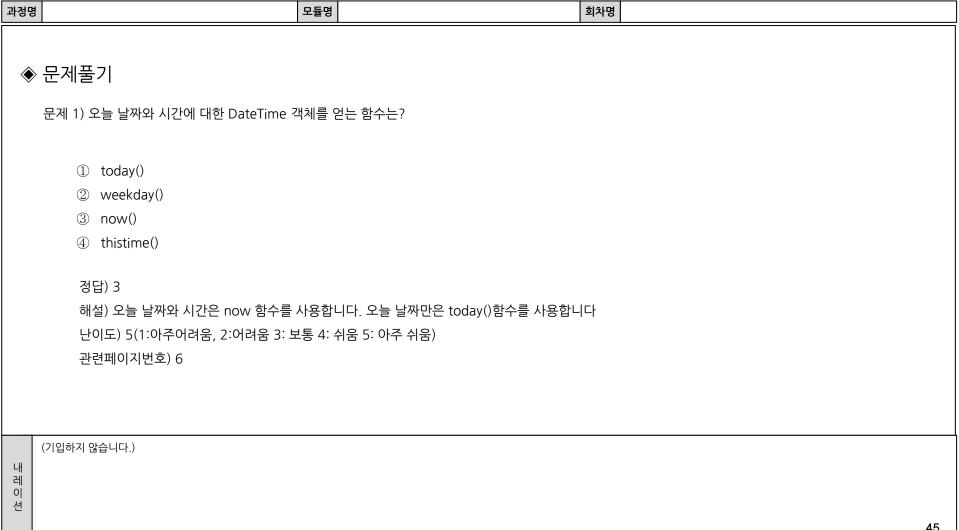
	과정명		모듈명	ā	회차명	
		◈ date_range 함수 예제 파일명 : exam14_6.py				
		import pandas as pd import numpy as np				
		#2019년 9월 1일부터 2019년 9월 30일까지 생 d = pd.date_range(start="2019-9-1", end="20print(d)")				
		d = pd.date_range(start="2019-1-1", periods print(d)	=60)			
		#2019년 1월중 일요일만 d = pd.date_range(start="2019-10-1", end=' print(d)	'2019-10-30", freq='W')			
		#2019년 각 달의 마지막날만 데이터 생성 np.random.seed(0) ts = pd.Series(np.random.randn(12), index=pd.date_range(start="2019-01-01", print(ts)	periods=12, freq="M"))			
	내레					
	네 이 션					
١					Λ	Λ





과정명		모듈명	회차명	
	◈ 적용하기- 풀이			
	from datetime import datetime, time import calendar	edelta		
	today = datetime.now() yesterday = today - timedelta(days tomorrow = today + timedelta(days			
	#1. 오늘 어제 내일 print("오늘 : ", today) print("어제 : ", yesterday) print("내일 : ", tomorrow)			
	#2번.이달의 마지막날 start_day, last_day = calendar.mont print("이달의 마지막 날은 ", last_day			
내 레 이 션	(기입하지 않습니다.)			43

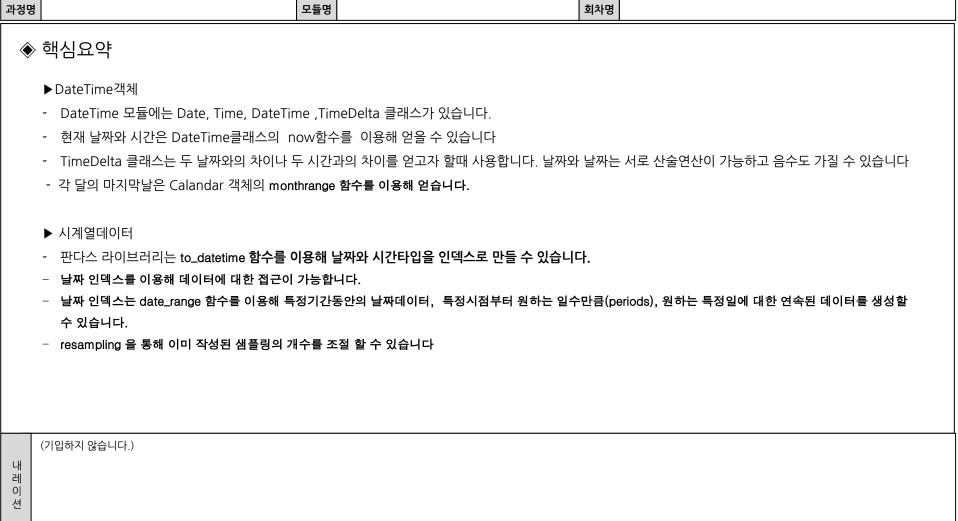




번호	문제	정답	난이도	해설	관련학습보기
2	오늘이 무슨 요일인지 숫자로 요일을 반환하는 함수는? ① now ② today ③ weekday ④ day	3	4	weekday 함수는 해당 요일을 수치자료로 반환합니다. 월요일을 0으로 보고 차례대로 숫자를 부여합니다	16
3	지금으로부터 6시간 뒤의 시간을 알아내려고 합니다. 맞는 코드를 작 성하세요	해설참조	1	current = datetime.today() after = current + timedelta(hours=6)	24
4	2019년 2월달의 마지막일을 할려고 합니다. 올바른 코드를 삭성하세 요	해설참조	4	start_day, last_day = calendar.monthra nge(2019, 2)	32

번호	문제	정답	난이 도	해설	관련학습 보기
5	다음 궁 타입스탬프에 대한 설명을 잘못된 것은 ? ① timestamp는 유닉스/리눅스 계열에서 사용하는 시간을 나타내는 정수입니다. ② 1970년1월1일 0시0분0초부터 몇초가 지났는 지를 나타내는 정수 값입니다 ③ 이 타임스탬프는 지구자전축의 움직임을 고려하는 윤초(Leap Second)까지 반영하여 아주 엄격하게 정의한 숫자입니다 ④ 단순히 하루를 86400초로 계산하여 순차적으로 증가시킨 것입니다	3	4	타입스탬프는 단순히 하루를 86400초로 계산하여 순차적으로 증가신것입니다	13
6	strftime함수의 서식문자와 내용을 잘못 연결한것은? ① %Y - 연도 ② %m - 월 ③ %M - 월 ④ %d - 일	3	3	월은 %m을 사용합니다. 대문자는 분을 나타냅니다	22
7	strftime함수를 사용하여 오늘 날짜를 문자열의 형태로 다음 처럼 출력하 려면 어떤 서식을 써야 하는지 바른것을 고르시오 결과: 2019-10-20 October ① %Y-%M-%D ② %Y-%M-%D %B ③ %Y-%m-%d %B ④ %Y-%m-%d %h	3	2	연도 월 일 월이 영어로 나왔으므로 %Y-%m-%d %B를 사용 한다	26

번호	문제	정답	난이 도	해설	관련학습보기
8	판다스 라이브러리의 to_datetime 을 이용해 DateTime 인덱스를 생성하려고 합니다. 문자열로 지정한 날짜 입력 형식이 잘못 된것은? ① "2018, 9, 1" ② "2019-9-1" ③ "2019/90/04" ④ "20191004"	1	4	날짜는 2019-10-01 이나 2019/10/02 또는 201091005 형 태로 주어야 합니다.	35
9	일정 기간동안 datetimeindex를 생성하려고 합니다. 2019.9.1일부터 30일간 연속된 날짜데이터를 생성하는 코드를 작성하시오 import pandas as pd	해설참조	1	pd.date_range(start="2019-9-1", periods=30)	37
10	1년동안 각달의 마지막 날만 pandas 라이브러리를 이용해 작성하려고 합 니다. 올바른 코드를 작성하십시오	해설참조	1	pd.date_range(start="2019-01-01", periods=12, freq="M"))	38



과정명	,	모듈명	회차명					
	↑ 됩奏마요							
♦ 학습 맺음								
	- 이번 회차에서는 파이썬에서 날짜와 시간을 다루는 부분을 학습해보았습니다 수고많으셨습니다							
	◈ 참고자료							
	- https://docs.anaconda.com/anacond	da/						
	- https://code.visualstudio.com/docs/g	getstarted/themes						
1	- https://pandas.pydata.org/pandas-	-docs/stable/getting_started/tutorials.html						
1	- https://www.programiz.com/python	<u>ı-programming/datetime</u>						
1	- https://allenjeon.tistory.com/235							
1	- https://www.programiz.com/python	<u>ı-programming/datetime</u>						
il	 https://datascienceschool.net/view-note 	ebook/8959673a97214e8fafdb159f254185e9/						
1	- https://pandas.pydata.org/pandas-docs/version/0.24/reference/api/pandas.DatetimeIndex.html							
<u> </u>								
	(기입하지 않습니다.)							
내레								
이 션								
			50					



- 주요학습내용 : 해당 회차 또는 레슨의 주요학습내용을 자세히 기입해 주세요.
- 검색 키워드: 학습자가 검색창에 어떤 검색어를 입력하면 본 회차 또는 본 레슨이 검색될 수 있을지 검색 키워드를 5개 기입해 주세요.

제목	주요학습내용	검색 키워드1	검색 키워드2	검색 키워드3	검색 키워드4	검색 키워드5
14. 시계열데이터다루기	DateTime모듈, Pandas library	DateTime	timedelta	timezone	to_datetime	DatetimeInde x