

In [1]:

```
import cx_Oracle
from selenium import webdriver
from selenium.common.exceptions import NoSuchElementException
```

In [2]:

```
con = cx_Oracle.connect('hr/password@localhost:1521/xe')
cur = con.cursor()
```

In [3]:

```
front= "http://210.95.199.40/hrdp/ti/ptiao/PTIA00100L.do?
startDate=20170101&endDate=20180717&pageRow=100&pageIndex="
```

In [4]:

```
con = cx_Oracle.connect('hr/password@localhost:1521/xe')
cur = con.cursor()
```

In []:

```
driver = webdriver.Firefox()

for k in range(1,50253):
    k = str(k)
    url = front+k
    #print(url)
    #print("-----")
    driver.get(front+k)
    driver.get(url)
    unicodeError = []
    for i in range(1,101):
        try:
            i = str(i)
            #고유번호
            no = no +1

            #강좌명 /html/body/div[3]/div[3]/div[3]/div/form/ul/li[2]/div[1]/p/a
            path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/p[1]/a"
            courseName = driver.find_element_by_xpath(path)
            courseName = courseName.text
            courseName = courseName.replace(".", "")
            courseName = courseName.replace(' ', '')

            #학원이름
            path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/dl/dt"
            academy = driver.find_element_by_xpath(path)
            academy = academy.text

            #주소 & 번호
            path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/dl/dd"
            addressTel = driver.find_element_by_xpath(path)
            addressTel = addressTel.text

            #기간
            path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[1]/li[1]/span"
            term = driver.find_element_by_xpath(path)
            term = term.text

            #훈련기관직종별취업률
            path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[1]/li[2]"
            employRate = driver.find_element_by_xpath(path)
            employRate = employRate.text
```

#취업처 임금평균

```
path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[1]/li[3]"
employWage = driver.find_element_by_xpath(path)
employWage = employWage.text
```

#훈련비

```
path="/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[1]/li[4]"
trainPrice = driver.find_element_by_xpath(path)
trainPrice = trainPrice.text
```

#모집정원

```
path="/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[1]/li[5]"
recruitLimit = driver.find_element_by_xpath(path)
recruitLimit = recruitLimit.text
```

#주말여부

```
path="/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[1]/li[6]"
weekendAble = driver.find_element_by_xpath(path)
weekendAble = weekendAble.text
```

#훈련시간

```
path="/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[2]/li[1]"
trainPeriod = driver.find_element_by_xpath(path)
trainPeriod = trainPeriod.text
```

#과정만족도

```
path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[2]/li[2]"
corSatisfaction = driver.find_element_by_xpath(path)
corSatisfaction = corSatisfaction.text
```

#평균연령대

```
path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[2]/li[3]"
meanAge = driver.find_element_by_xpath(path)
meanAge = meanAge.text
```

#NCS적용여부

```
path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[2]/li[5]"
ncs = driver.find_element_by_xpath(path)
ncs = ncs.text
```

#직종 (직무수준)

```
path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[2]/li[4]"
occupation = driver.find_element_by_xpath(path)
occupation = occupation.text
```

#관련자격증

```
path = "/html/body/div[3]/div[3]/div[3]/div/form/ul/li["+i+"]/div[1]/ul[2]/li[6]"
releatedCertificate = driver.find_element_by_xpath(path)
releatedCertificate = releatedCertificate.text
```

```
#print("-----")
```

```
#print(no,courseName,academy, addressTel, term,employRate,employWage,trainPrice,
      recruitLimit, weekendAble, trainPeriod,corSatisfaction, meanAge,occupation,ncs,re
leatedCertificate);
```

```
sql_insert= 'INSERT INTO RAW_HRD_NET
(NO,COURSENAME,ACADEMY,ADDRESSTEL,TERM,EMPLOYRATE,EMPLOYWAGE,TRAINPRICE,RECRUITLIMIT,WEEKENDABLE,TE
RIOD,CORSATISFACTION,MEANAGE,OCCUPATION,NCS,RELEATEDCERTIFICATE) values (:1, :2, :3, :4, :5, :6,
:7, :8, :9, :10, :11, :12, :13, :14, :15, :16)'
```

```
cur.execute(sql_insert,(no,courseName,academy, addressTel, term,employRate,employWage,
                        trainPrice,recruitLimit, weekendAble,
trainPeriod,corSatisfaction,
                        meanAge,occupation,ncs,releatedCertificate))

con.commit()
```

```
except NoSuchElementException :
```

```
#print("-----")
```

```
#print(no,courseName,academy, addressTel, term,employRate,employWage,trainPrice,
      recruitLimit, weekendAble, trainPeriod,corSatisfaction, meanAge,ncs);
```

```
sql_insert= 'INSERT INTO RAW_HRD_NET
(NO,COURSENAME,ACADEMY,ADDRESSTEL,TERM,EMPLOYRATE,EMPLOYWAGE,TRAINPRICE,RECRUITLIMIT,WEEKENDABLE,TE
ERIOD,CORSATISFACTION,MEANAGE) values (:1, :2, :3, :4, :5, :6, :7, :8, :9, :10, :11, :12, :13) '

cur.execute(sql_insert,(no,courseName,academy, addressTel, term,employRate,employWage,t
rainPrice,
                                recruitLimit, weekendAble, trainPeriod,corSatisfaction, meanAge
)

con.commit()
```