**Case Study 1**

Task 1:

import xlrd

import xlwt

import datetime

def combine\_sheets(excel\_path,sv\_path):

1)

# read excel

workbook = xlrd.open\_workbook(excel\_path) //new excel

new\_workbook = xlwt.Workbook()

new\_sheet = new\_workbook.add\_sheet(u'Combine', cell\_overwrite\_ok=True)

style = xlwt.XFStyle()// convert integer to date

# count quarters

quarter\_count = {'quarter1':0,'quarter2':0,'quarter3':0,'quarter4':0}

# num of sheets

sheet\_names = workbook.sheet\_names()

start\_row = 0

for sheet\_name in sheet\_names:

sheet = workbook.sheet\_by\_name(sheet\_name)

# num of rows

n\_rows = sheet.nrows

for j in range(n\_rows):

row = sheet.row\_values(j)

# add one more col to header

if sheet\_name == 'Compliants\_Details':

if j==0:

row.append('Quarter')

else:

quarter = season\_filter(row[2])

quarter\_count[quarter]+=1

row.append(quarter)

print(row)

#columns K

for k in range(len(row)):

if sheet\_name == 'Compliants\_Details' and (k==2 or k==11):

style.num\_format\_str = 'YY/M/D'

new\_sheet.write(start\_row+j, k, row[k],style)

else:

new\_sheet.write(start\_row + j, k, row[k])

start\_row+=n\_rows+3

# add a new table counter\_count

for n,quarter in enumerate(quarter\_count.keys()):

new\_sheet.write(start\_row + n, 0, quarter)

new\_sheet.write(start\_row+n,1,quarter\_count[quarter])

new\_workbook.save(sv\_path)

# "Quarter"

# nov'17 to jan'18 - quarter1;

# feb'18 to apr'18 - quarter2;

# may'18 to july'18 - quarter3;

# aug'18 to oct'18 - quarter4

def int2date(dates):

delta=datetime.timedelta(days=dates)

today=datetime.datetime.strptime('1899-12-30','%Y-%m-%d')+delta

return today.date()

def season\_filter(num):

date = int2date(num)

year = date.year

# determine time for the season

if date < datetime.date(year,1,18):

return 'quarter1'

elif date > datetime.date(year,11,17):

return 'quarter1'

elif datetime.date(year,2,18) < date < datetime.date(year,4,18):

return 'quarter2'

elif datetime.date(year, 5, 17) < date < datetime.date(year, 7, 18):

return 'quarter3'

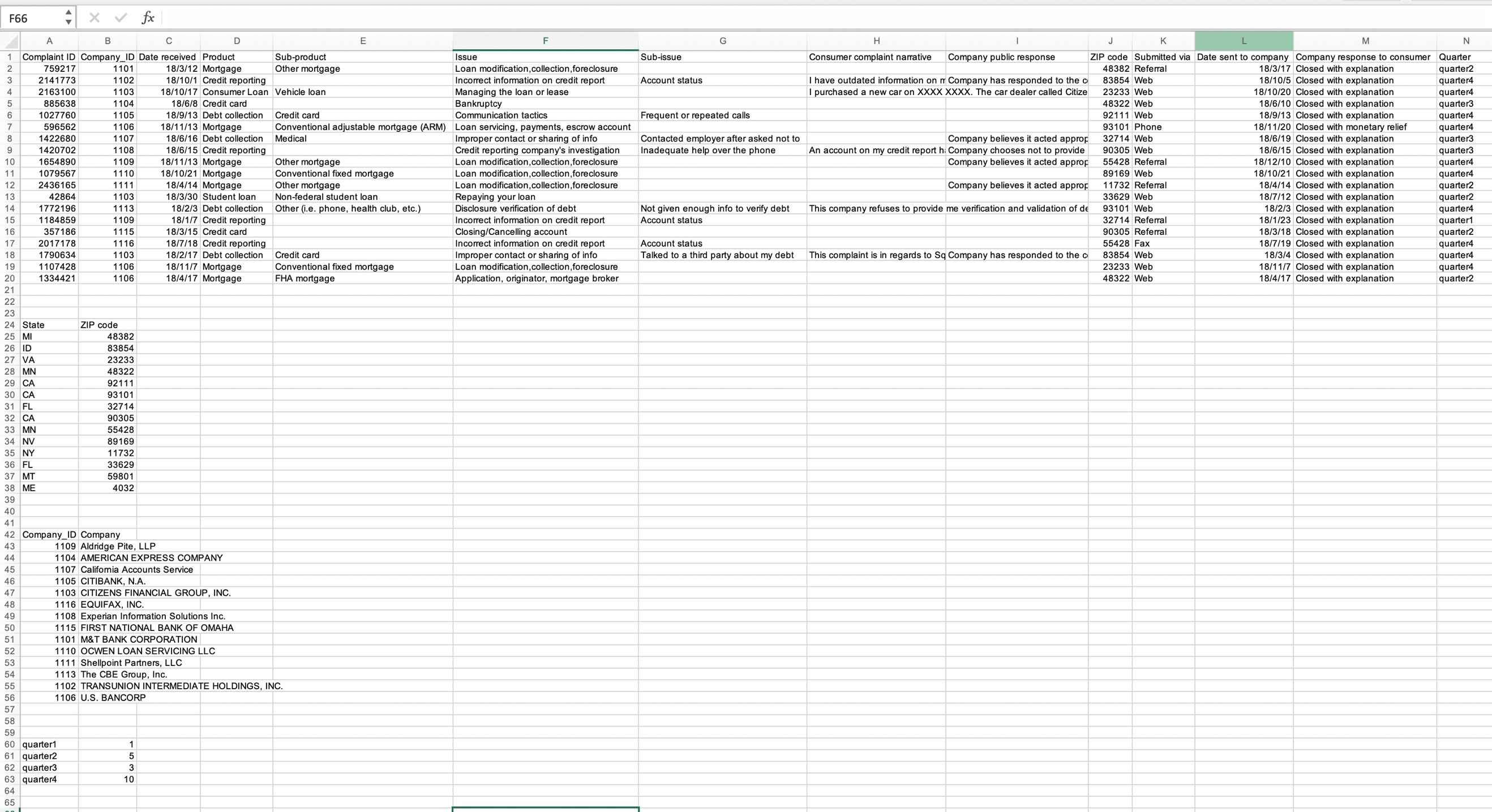
else:

return 'quarter4'

excel\_path = '/Users/welcomemymaster/Desktop/tmp/Dataset 1.xlsx'

sv\_path = '/Users/ welcomemymaster /Desktop/tmp/combine.xls'

combine\_sheets(excel\_path,sv\_path)



Task 2:

1)

CREATE Tab\_4 AS (

SELECT \*

FROM Compliants\_Details

LEFT JOIN Locations\_Details

ON Compliants\_Details.ZIP code = Locations\_Details. ZIP code

LEFT JOIN Company\_Details

ON Compliants\_Details .Company\_ID = Company\_Details. Company\_ID);

2)

SELECT

Date received AS DATE\_RECEIVED,

SUM(Complaint ID) AS COMPLAINTS\_NO

FROM

Tab\_4

WHERE year(DATE\_RECEIVED) = 2018

GROUP BY

month(DATE\_RECEIVED);

Task3: