UCLA Extension - Introduction to Data Science
Ali El-Annan

!! Please read carefully !!

Let's write a story...well, a sort of story. We're an investigative journalist and we're investigating possible corporate corruption. The esteemed (or perhaps not so esteemed head of the department) has a budget of $1,500,000. Interestingly enough he hired two of his kids (whom seem to be missing qualifications and vetting) and we suspect are part of the budget leak. The department has a strict anti-nepotism policy (which is poorly and badly enforced). The department is budgeted for 10 employees and we have a list of salaries related to roles. 4 are accountants, 4 are data analysts, and 2 are meant for contractors. However it appears the employee budget spend is $1,500,000! We need to isolate is there a budget overspend? Where is the possible overspend coming from?

- Accountants Budgeted Salaries: $100,000
- Data Analyst Budgeted Salaries: $100,000
- Contractor Budgeted Salaries: $80,000
- Total number of employees: 10 (4 accountants, 4 data anaysts, 2 contractors)
- Total employee budget: 4*100,000 + 4*100,000 + 2*80,000
- Actual spend...?

Luckily, we're not alone --- we have our trusty

Python (and Python program)

at our side, it will help us track/tabulate/and de-bug..fix the code!

(1) Well, well, well...let's take some notes...

Our trusty pet Python (and accompanying computer program...) will help us, let us fire up our friend Spyder, where we will run all our commands, we'll need to store some of this information...and for that, we'll need variables

```
#global variables, we will access throughout the program
accountantSalaryVar = 100000
dataAnalystSalaryVar = 100000
contractorSalaryVar = 80000
employeeCap = 10
totalAllowedBudget = '$960,000'
actualBudget = '$1,500,000'
```

Let's actually do some investigate journalism and request a list of the employee names! And most importantly, let's fine out whose related to the Head of Department...we only know who in the office is related to one another

```
#more global variables, we will access throughout the program
listOfEmployees = ['Ronald', 'Bilal', 'Tariq', 'Cinderlla', 'Luan', 'Viet', 'Sameer', 'Reina', 'Perla', 'Ronalada']

dictOfEmployeeSiblings =  {
      "Ronald": "Sibling",
      "Bilal": "No relative",
      "Tariq": "No relative",
      "Cinderlla": "No relative",
      "Luan": "No relative",
      "Viet": "No relative",
      "Sameer": "No relative",
      "Reina": "No relative",
      "Perla": "No relative",
      "Ronalada": "Sibling" }
```

(2) We need to do some initial forensics...

So, we know the max is 10 employees, we know that the budget HAS to be $960,000 but how much over is it? Do we have more than 10 employees? Are any of them even related? Luckily we can create some functions to help us --- we'll eventually have to hand these over to the compliance department, we can than permanently track this

```
#lets' test to see if the list we were given passes the number of employees check
totalAllowedBudget = totalAllowedBudget.strip('$')
totalAllowedBudget = totalAllowedBudget.replace(',','')
actualBudget = actualBudget.strip('$')
actualBudget = actualBudget.replace(',','')

overBudget = abs(int(totalAllowedBudget) - int(actualBudget))
print("We are overBudget " + str(overBudget))

def checkEmployeeCap(numberOfEmployeesCap, listOfEmployees):
    #local variable, accessible only within a function
    numberOfExistingEmployee = len(listOfEmployees)
```

```
    if (numberOfEmployeesCap < numberOfExistingEmployee):
        print("We have more than the employee cap...")
    else:
        print("We have 10 or less employees...")

#lets' test to see if the list we were given passes the number of employees check
checkEmployeeCap(employeeCap, listOfEmployees)

def checkEmployeeSiblings(dictOfEmployeeSiblings):
  for name, relation in dictOfEmployeeSiblings.items():
    if relation == "Sibling":
      print(name)

#lets' test to see if the list we were given passes the number of employees check
checkEmployeeSiblings(dictOfEmployeeSiblings)
```

So, we know the max is 10 employees, we know that the budget HAS to be $960,000 but how much over is it? ... 540000!!! Do we have more than 10 employees? ... No Are any of them even related? ... Yes (Ronalada, Ronald) Things aren't lookin good...(for corruption existing!)

(3) We have to speak the truth (with numbers)!

So looks like we know Ronald, Ronalada are related, we know the department budget is over $540,000 and only 10 employees...let us send a message to shareholders, compliance...to the law (hopefully they haven't been bought out...

```
print("Dear Management, Shareholders and the Law\n\
    We have found $540,000 overspend in budget with 2 relatives, Ronald and Ronalada. Good news is we now have some functions to run throug
      Sincerely,\
        Trying to catch corruption...\
  ")
```

```
Additional Exercises: Go back and see what functions you can add? Maybe something that calculates what Accountants + Data Analysts + Contracto

Things to explore/understand:
Casting, and common functions:
int(), str(), functions are common, we work in a world of numbers/and strings
https://docs.python.org/3/library/stdtypes.html?highlight=casting%20int

Datastrucutres:
dict, lists, very common datastrucutures, you DON'T have to memorize each function, but after a while you'll find you recall them faster, if r
https://docs.python.org/3/tutorial/datastructures.html

Most things in the natural world are lists (i.e. a list of data points)? or key value pairs (i.e. an address is a location/key of a persons ho

Notes: You can use Spyder OR Jupytr for this, I would recommend Spyder IDE from anaconda
```