# Question 1

|  |  |
| --- | --- |
| a | print(8 \*\*4) |
| b | inputString = "Split this string"  print(inputString.split()) |
| c | planet = "Earth"  diameter = 12742  print("The diameter of {} is {} kilometers.".format(planet, diameter)) |
| d | the\_list = [1,2,[3,4],[5,[100,200,['target']],23,11],1,7]  print(the\_list[3][1][2]) |
| e | the\_dic = {'k1':[1,2,3,{'tricky':['oh','man','inception',{'target':[1,2,3,'hello']}]}]}  print(the\_dic['k1'][3]['tricky'][3]['target'][3]) |
| f | def checkWord(inputString):      return 'elephant' in inputString.lower().split()  results = checkWord('where is the elephant ?')  print(results) |
| g | def countElephantWord(inputString):      count = 0      for word in inputString.lower().split():          if word == 'elephant':              count += 1      return count  inputString = "This elephant is heavier than the other elephant dude!"  print(inputString.lower().split())  num = countElephantWord(inputString)  print(num) |
| h | def checkSpeed(speed, birthday):    isBirthday = False    results = ""    if birthday:      speed = speed - 5    print(speed)    if speed <= 60:      results = "Low speed"    if speed >= 61 and speed <= 80:      results = "Medium speed"      if speed >= 81:      results = "High speed"    return results  print(checkSpeed(61,True)) |

# Question 2

|  |  |
| --- | --- |
| a | schema = ["employee\_name","department","state","salary","age","bonus"]  df = spark.createDataFrame(data=data, schema = my\_schema ) |
| b | df.show()  df.printSchema() |
| c | df.groupBy("department").min("salary").show()  df.groupBy("department").max("salary").show()  df.groupBy("department").avg("salary").show() |
| d | df.groupBy("country").min("salary").show()  df.groupBy("country").max("salary").show()  df.groupBy("country").avg("salary").show() |