1. Write a python code to Implement a bubble-sort on a list. Explain your program.
2. Write a python program to implement an insertion sort on a list. Explain your code.
3. Looking at the below code, write down the final values of A0, A1, ...An. Please explain the answers.

A0 = dict(zip(('a','b','c','d','e'),(1,2,3,4,5)))

A1 = range(10)

A2 = sorted([i for i in A1 if i in A0])

A3 = sorted([A0[s] for s in A0])

A4 = [i for i in A1 if i in A3]

A5 = {i:i\*i for i in A1}

A6 = [[i,i\*i] for i in A1]

1. You work for a startup that offers loans. You have an event log that records every time a user inserted a verification image into a page. (One user can insert multiple images.) The event\_log SQL table looks like this:

|  |  |
| --- | --- |
| **user\_id** | **event\_date\_time** |
| 7494212 | 1535308430 |
| 7494212 | 1535308433 |
| 1475185 | 1535308444 |
| 6946725 | 1535308475 |
| 6946725 | 1535308476 |
| 6946725 | 1535308477 |
| … | … |

…and it has over one billion rows.  
*Note: If the*event\_date\_time*column’s format doesn’t look familiar, google “epoch timestamp”!*

**Write an SQL query to find out how many users inserted more than 1000 but less than 2000 images in their applications!**

1. You have two files in hdfs one having date range with two columns start date and end date and another having two columns with date and visitors field. You have to write a spark code which gives date range having maximum no. of visitors using both tables.
2. Find What is Wrong in this Query?

SELECT subject\_code, AVG (marks) FROM students WHERE AVG(marks) > 75

GROUP BY subject\_code;

1. Draw what you’d perceive to be an architecture where data goes from its rawest form and transforms into model ready data used by Decision Science teams. Please also recommend hardware/software that can be used to facilitate this architecture.