|  |  |  |  |
| --- | --- | --- | --- |
| **3주차 Homework** | | | |
| 학번 | 32164420 | 이름 | 조정민 |

※ 문제를 해결하여 제출하시오. 필요시 실행 결과를 캡쳐하여 붙여넣는다.

|  |
| --- |
| **[Python 복습] 주피터를 활용해 결과를 구하시오.** |

1. 문자열 ‘Big data processing’를 변수에 저장하고 그것을 한 문자씩 출력하는 프로그램을 작성하시오(1점).

Answer:

|  |
| --- |
| font: consolas / Font size: 10  s = ‘Big data processing’  for element in s:  print(element) |

2. 1부터 50까지 자연수의 합을 구하고 결과를 출력하는 프로그램을 작성하시오(2점).

Answer:

|  |
| --- |
| Font: consolas / Font size: 10  sum = 0  for i in range(1, 51):  sum += i  print(sum) |

3. 피보나치 수열을 구하는 함수를 정의하고, 그 함수를 이용해 수열의 40번째 값을 구하시오. 피보나치 수열은 첫째 항과 둘째 항이 1이고 세번째 항부터는 바로 앞의 두 항의 합인 수열이다(1, 1, 2, 3, 5, 8, 13, …). (3점)

Answer:

|  |
| --- |
| Font: consolas / Font size: 10  def fibonacci(num):  a, b = 0, 1  for i in range(num):  a, b = b, a+b  return a  print(fibonacci(40)) |

|  |
| --- |
| **[3주차 확인 학습]** |

1. 하둡과 하둡 에코시스템에 대해 중학생 동생에게 설명한다고 가정하고, 자신의 언어로 설명해 보시오. (4점)

Answer:

|  |
| --- |
| Font: 맑은 고딕(한글) / Font size: 10  시간이 흐르면서 모든 분야에 대량의 데이터를 사용하게 되었고 이를 빅데이터라고 한다.  이 빅데이터를 쉽게 분석하고 처리할 수 있었던 것은 하둡이 개발되면서 부터이다.  하둡은 대용량의 데이트를 저비용으로 더 빠르게 분석활 수 있는 소프트웨어이며, 빅데이터의 처리와 분석을 위한 플랫폼 중 사실상의 표준으로 자리잡고 있다.  하둡을 쉽게 예를 들어 설명하자면, 한 명의 교사가 1만 명 학생의 시험을 관리한다고 생각해보자. 해당 교사는 혼자서는 1만 명 학생의 시험지를 관리하기 매우 힘들다. 때문에, 알바생을 채용하여 시험지를 복사하고 답안지도 작성하고, 채점도 하는 등 여러 역할을 분산시켜 알바생들에게 할 일을 줄 것이다. 이러한 부분들을 다 모아서 결과를 보고하도록 하는 시스템이 하둡이다.  이 시스템은 여러 대의 컴퓨터로 데이터를 분석하고 저장하는 방식으로 분석에 필요했던 많은 비용과 시간을 단축할 수 있게 된다. 국내 뿐만이 아니라 해외에서도 빅데이터를 다루는 개발자는 하둡 생태계를 통해 빅데이터를 보다 원활하고 효율적으로 분석할 수 있게 된다. 하둡은 데이터의 분산 저장(HDFS), 분산처리(MapReduce), 2가지를 모두 지원한다. 하둡의 기능은 크게 수집, 저장, 처리, 관리 4가지가 있다.  하둡의 코어 프로젝트는 위에서 말한 분산저장(HDFS), 분산처리(MapReduce)이지만 그 외에도 데이터를 수집, 저장, 처리, 관리해주는 다양한 서브 프로젝트들이 많다. 하둡 에코시스템이란 것은 그 프레임워크를 이루고 있는 다양한 서브 프로젝트들의 모임이다. 쉽게 말해서, 하둡과 관련된 프로젝트(프레임워크)들을 하둡 에코시스템이라고 생각하면 된다. |