

1. 다음 데이터로 cluster dendrogram 을 그리고 20과 같은 군집인 숫자는 ? (10점) (거리 계산=euclidean, 군집분석 = complete)  
40 ,55, 60, 95, 80, 90, 105, 10, 20, 30

40과 같은 군집의 숫자는 (                      )

2. 다음 데이터로 cluster dendrogram 을 그리고 20과 같은 군집인 숫자는 ? (10점) (거리 계산=euclidean, 군집분석 = single)  
40 ,55, 60, 95, 80, 90, 105, 10, 20, 30

40과 같은 군집의 숫자는 ( (                      ), (                      ), (                      ) )

3. 판별분석의 'default' 데이터로부터 다음을 구하라 (50점)

3.1 파산 확률과 income 사이에 관계는 ? (10점) p-value = (                      ), 관계가 ( ① 있다, ② 없다)

3.2 Income이 25,000인 사람의 파산확률은 ? (10점) (                      ) %

3.3 파산 확률과 balance + student 사이에 관계는 ? (30점)

balance p-value = (                      ), income과 파산 확률은 관계가 ( ① 있다, ② 없다) (10점)

student p-value = (                      ), income과 파산 확률은 관계가 ( ① 있다, ② 없다) (10점)

balance=2,000, 학생(student) 일 때 파산 확률은 ? (                      ) % (10점)

4. 머신러닝 중 하나인 의사결정나무 분석 결과 다음과 같이 나타났다. (15점)

Predicted	Reference	
	0	1
0	300	100
1	100	550

4.1 Accuracy ? (                      ) %

4.2 Sensitivity ? (                      ) %

4.3 Specificity ? (                      ) %

5. 'dataTransactions.tab' 데이터를 활용하여 연관규칙탐사를 수행하고 표 안을 채워라 (15점)

조건 : ① 일반식품, 화장품 코너 제외, ② 15개 이상 품목 구입만 고려, ③ 지지도 20% 이상, 신뢰도 80% 이상  
④ 스포츠 코너 중심 분석 (향상도 1.4 이상)

	lhs	rhs	support	confidence
[1]		영캐주얼		