Canvas #06

Details of datasets

Designed by: Mark Spencer, VP Marketing Gym Sports

Date: \_\_\_\_\_May 15\_\_\_\_

BONUS POINTS	Explanations	Dataset 1:	Dataset 2:	Dataset 3:
1 to 5 (1 = hard, 5 = easy)		_Fitness machines_	_Body scan	Google Fit / Apple Health
Machine readable?	if the data is in a .docx or pdf file, software can't read it. A database or even a csv file is better.	5	5	5
Structured or not?	if the dataset is "Excel like' then it is quite structured. Free text, web pages or pictures are typically very unstructured.	5	5	5
Follows universal categories or is it firm specific?	a dataset following INSEE or Eurostat categories is quite universal.	3	4	4
Time series?	is the data collected several times across months or years?	5	4	4
Personal and sensitive data?	Personal data comes with more constraints. Sensitive data even more.	2	1	1
Complete?	No missing records, years, values, and no errors.	5	4	4
Sum of points per dataset	Add up the points to get a total. A higher total shows a more favorable dataset	25	23	23

Copyright © 2017-2019, emlyon business school. DDBM is for you to use without restriction in modeling your own or other people's businesses. If you wish to use DDBM in original or adapted to sell it as a tool, you must contact the copyright holder for permission.