

Canvas #06

Details of datasets

Designed by: _____

Date: _____

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