

## VUE DÉTAILLÉE SUR LES JEUX DE DONNÉES

POINTS BONUS 1 = Difficile 5 = Facile	Explications	Dataset 1 : _____	Dataset : 2 _____	Dataset : 3 _____
Format lisible par un programme informatique?	<i>Si la donnée est sous format .docx ou pdf, un programme de code peut difficilement la lire. Une base d donnée ou même un fichier csv sont plus faciles</i>			
Structuré ou non?	<i>Si la donnée peut tenir dans Excel, elle est probablement structurée. Le texte libre, une page web, ou des images sont des formats non structurés.</i>			
Suis des catégories universelles ou spécifiques?	<i>Des données suivant des codifications INSEE ou Eurostat sont a priori assez universelles</i>			
Séries temporelles?	<i>La donnée est-elle collectée à travers le temps? La fréquence de collecte est-elle suffisante?</i>			
Données personnelles?	<i>La donnée personnelle impose des contraintes</i>			
Données complètes?	<i>Y a-t-il des valeurs manquantes, des valeurs erronées, des dates manquantes...</i>			
<b>Total: somme des points par dataset</b>	<i>Faites la somme des points. Un total élevé indique un dataset relativement plus facile à utiliser.</i>			