

金融大数据作业1

施宇 191250119

请解释数据和信息的区别和关系。

联系：数据是人们用于记录事物情况的物理符号，是信息的符号表示或称载体。信息是数据的内涵，是对数据的定义的解释。不同知识、经验的人，对于同一数据的理解，可得到不同信息。

区别：数据是表示信息的，但并非任何数据都能表示信息，信息只是加工处理后的数据，是数据所表达的内容。另一方面信息不随表示它的数据形式而改变，它是反应客观现实世界的知识，而数据则具有任意性，用不同的数据形式可以表示同样的信息，信息只是加工处理后的数据，是数据所表达的内容。数据的格式往往随着载体它的物理设备的形式而改变，具有多种多样的形式，也可以加以转换，但其中包含的信息内容不会改变，即信息不随载体的物理设备形式的改变而改变。信息可以离开信息系统而独立存在，也可以离开信息系统的各个组成和阶段而独立存在。

何为大数据？

大数据指一个超大的、难以用现有常规的数据库管理技术和工具处理的数据集。大数据技术描述了一种新一代技术和构架，用于以很经济的方式、以高速的捕获、发现和分析技术，从各种超大规模的数据中提取价值。大数据是需要新处理模式才能具有很强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。大数据就是“未来的新石油”。

大数据有几个特征？

大体量（Volume）：即可从数百TB到数十数百PB，甚至EB的规模。

多样性（Variety）：即大数据包括各种格式和形态的数据。数据来源多，数据类型多，关联性强。

时效性（Velocity）：即很多大数据需要在一定的时间限度下得到及时处理

准确性（Veracity）：即处理的结果要保证一定的准确性。大数据隐含着更准确的事实。

大价值（Value）：即大数据包含很多深度的价值，大数据分析挖掘和利用将带来巨大的商业价值。价值密度低，是大数据的一个典型特征。

请简述金融数据的特征。

一般具有“流数据”特征，需要在短时间内快速处理。具有逻辑关系紧密，处理实时性要求高，可展示性需求强等特征。