**第8章 文本数据可视化**

**【练习题】**

1. **文本数据的类型与可视化目的**
2. 列举五种常见的文本数据类型，并说明它们的典型应用场景。  
   b) 为什么我们需要对文本数据进行可视化？请列出文本数据可视化的三个主要功能。
3. **文本可视化技术**
4. 什么是词云（Word Cloud）？它是如何通过字体大小和颜色来传递文本信息的？  
   b) 与传统词云相比，Wordle 有哪些改进？它是如何增强文本数据的视觉表达的？
5. **文本可视化分析流程**
6. 简述文本可视化分析的基本流程，包括数据预处理、特征提取与可视化设计。  
   b) 在文本可视化中，交互设计的重要性是什么？请举例说明如何通过交互设计增强用户的探索体验。
7. **文本分析技术**
8. 解释分词（Tokenization）在文本分析中的作用，并讨论它在不同语言（如英文和中文）中的处理差异。  
   b) 请解释 TF-IDF 的核心原理，并说明它在文本分类和信息检索中的应用。
9. **可视化技术的应用**
10. 主题河（ThemeRiver）如何用于文本的时间序列分析？请描述其工作原理及应用场景。  
    b) 什么是 LDA 主题模型？如何利用它进行文档的主题分析？
11. **文本关系可视化**
12. 什么是短语网络（PhraseNet），它如何展现文本中词汇间的关系？  
    b) 与单词树（WordTree）相比，句子树（SentenTree）有哪些独特优势？请结合实际案例说明其应用价值。