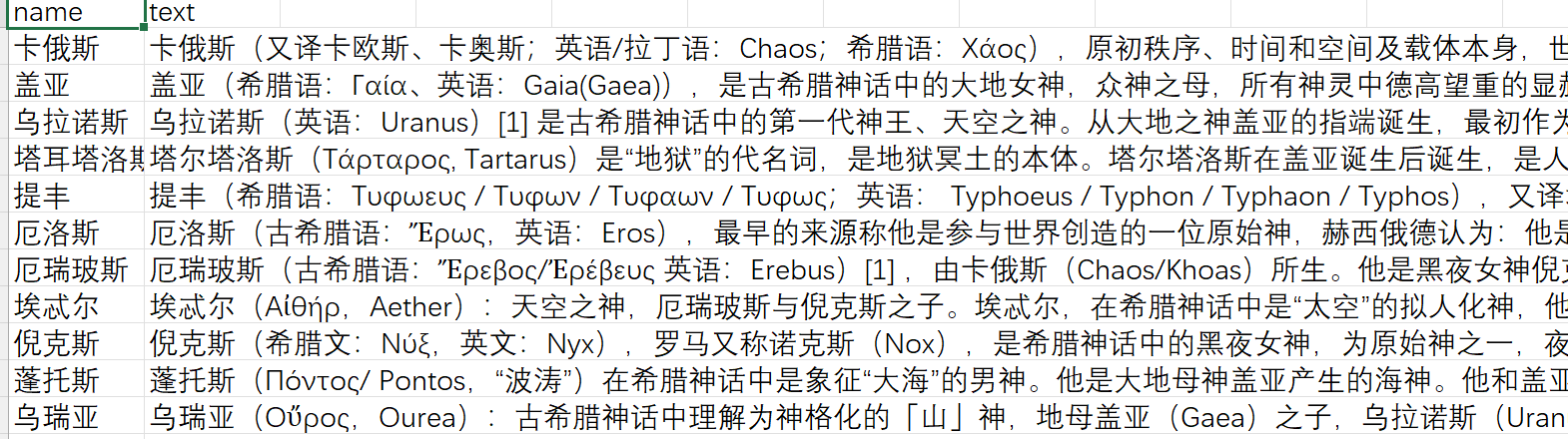
我们关注百度百科词条——希腊神话（西方神话体系之一），为他们在主要神祗、次要神祗、著名的人类和半神三个分目下罗列的人物进行人物关系知识图谱的构建。

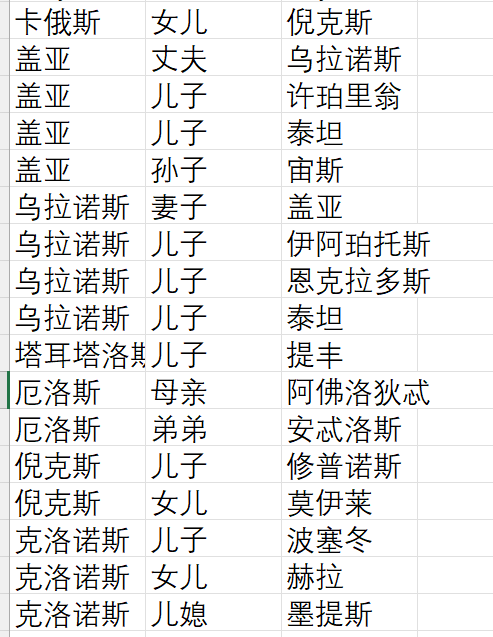
Step1：爬取以上三个分目下的人物以及其对应的链接，得到raw\_url.csv

Step2：按照链接读取每个百科词条的摘要部分，其中涵盖百科所介绍的人物关系，得到raw\_text.csv（如下图所示）



注意：读取的url会由于同名导致text缺失，我们手动补全了这部分缺失

Step3：通过爬取百度百科的关系定义，以及对文本的分析，我们提取了不同人物间的三元组，得到关系文件 relation1.txt



Step4：除了人物间的关系外，我们希望找到人物间通过事件的关联。因此，我们查询了一本书《古希腊神话》（https://pan.baidu.com/s/1sltzETR?\_at\_=1637829256318）见文件story.txt，将每个小故事作为一个事件（新节点），连接所有与该节点有关联的人物节点。将所有故事关联节点关系存储到all\_intersection.txt中。

Step5：从古希腊神话故事[乌拉诺斯]提取人物相关的故事，制成人物-故事关系文件

Step6：整理实体和关系文件制成csv文件，其中person\_entity包含人物，story\_entity包含故事的标题，person\_intersection表示人物之间有交集，person\_relation包含更加具体的人物之间关系的信息，story\_relation包含人物与故事的关系，entity中包含所有实体，relations包含所有关系。

Step7：graph.py将关系文件部署到neo4j，并且可视化

