



项目规格书

数据工程的顶点

写上去

标准	符合规格
计划范围	撰写的内容包括项目中所采取步骤的概述。 最终数据模型的目的是明确的。
解决其他情况	本文在以下情况下描述了此项目的逻辑方法： <ul style="list-style-type: none">• 数据增加了100倍。• 管道每天每天早上7点运行。• 该数据库需要100多人访问。
保卫决定	工具，技术和数据模型的选择是合理的。

执行

标准	符合规格

标准	符合规格
项目代码是干净且模块化的。	所有编码脚本均具有直观，易于遵循的结构，并将代码分为逻辑功能。变量和函数的命名遵循PEP8样式准则。该代码应运行没有错误。
质量检查	该项目至少包括两次数据质量检查。
资料模型	<ul style="list-style-type: none">ETL过程产生了写入中概述的数据模型。包含最终数据模型的数据字典。数据模型适合于已确定的目的。
数据集	<p>该项目包括：</p> <ul style="list-style-type: none">至少2个数据源超过一百万行数据。至少两种数据源/格式 (csv, api, json)

让您的项目脱颖而出的建议！

为了使您的项目脱颖而出：

- 处理大量数据。
- 合并难以合并的数据集。
- 丰富来自多个不同来源的数据。
- 包括有关如何使用数据提出见解的建议。
- 撰写有关您的项目的博客文章并链接到它。