

Общая информация

Для выполнения задания необходимо завести аккаунт на [GitHub](#)
Для получения доступа к заданию заполните форму по [Ссылке](#)
Описание задачи, правила проведения соревнования и данные размещены в GitHub-репозитории, доступном после заполнения формы выше.
Просим оставить обратную связь по [Ссылке](#) после решения задания.

Желаем удачи!

Помогите Лютику разобрать интервью

Мир **Неверленда** переживает экономический бум. Изобретение колесных зверей, работающих на зачарованном масле, ускорило общий прогресс. Появились полноценные предприятия по масштабному производству и распределению этого масла. Для пополнения мощностей колесных зверей стали открываться отдельные точки - Зачарованные маслостанции. Появилась большая потребность в рабочих совершенно разного рода, в повышении качества условий работы для привлечения лучших соискателей.



Йеннифер и Трисс работают в компании **"Цинтрийские Масла Феникса"**. Для анализа рабочей среды они решили провести дополнительные внутренние интервью с работниками компании. В ходе этих интервью они стремились понять истинное

состояние дел в компании, выявить скрытые проблемы и улучшить рабочие процессы, получить инсайты для привлечения новых сотрудников. Каждое интервью велось в формате открытого диалога, а самопишущее цидарийское перо фиксировало ключевые моменты и важные детали для дальнейшего анализа.

Судьба распорядилась так, что именно *Лютик*, недавно устроившийся в **ЦМФ** в качестве начинающего аналитика, оказался перед задачей анализа полученной статистики. Лютик, полный энтузиазма и амбиций, был взволнован своей первой серьезной задачей. Он знал, что от его работы зависит не только дальнейший успех подразделений, но и его репутация в компании.



Вы найдете приложенный файл, который предварительно оцифровал и обработал *Лютик*.

Задание

1. Предложите решение по выделению подмножества тематик и их описания из общего корпуса цитат.
2. Предположим, Йеннифер и Трисс поставят эту активность как регулярную, проводя интервью раз в месяц по разным доменам предприятия. Предложите варианты, как Лютику сделать для них легковесную систему по обработке новых цитат из опросов с учетом уже накопленной истории извлеченных тематик и тегов.

Вспомогательные материалы

1. Witcher Wiki: https://witcher.fandom.com/wiki/Witcher_Wiki
2. Git Book: <https://git-scm.com/book/en/v2>
3. Git LFS: <https://git-lfs.com>
4. Connect with SSH: <https://docs.github.com/en/authentication/connecting-to-github-with-ssh>
5. PIP: <https://pip.pypa.io/en/stable>
6. Conda: <https://docs.conda.io/en/latest>
7. Jupyter Docker Stacks: <https://jupyter-docker-stacks.readthedocs.io/en/latest>
8. Data science with JupyterLab: <https://docs.docker.com/guides/use-case/jupyter>
9. PEP: <https://peps.python.org>