**Частина 1: Вибір теми та опис проєкту**

**Приклади:**

***Тема:*** *Аналіз особистої історії прослуховування в Spotify*

***Опис проєкту:***

*Мій PET-проєкт буде спрямований на перевірку знань із пошуку, обробки та аналізу інформації з різних джерел та форматів.*

*Обрала особисті дані, бо цікаво дізнатися щось нове про себе. Обрала саме спотіфай, бо дані з Гуглу про мою активність містять багато пустих значень і є неінформативними для аналізу.*

*Під час аналізу буде добре видно, як змінювалися мої музичні вподобання протягом 2023-2024 років (з квітня 2023, коли я встановила додаток, і по квітень 2024).*

**Частина 2: Вибір технологій та стеку**

* **Excel:** використовувала для очистки від викидів **(**музика для сну зайняла перше місце в рейтингу, але це не те, що я слухаю свідомо, просто щоб заснути, а вона потім грає цілу ніч, накручуючи прослуховування**).**
* **Python**: Використовувала Python для перетворення у потрібний формат даних, які я аналізувала в **Tableau.**
* **Tableau**: Використовувала для аналізу та створення візуалізацій та дашбордів для кращого візуального представлення результатів.

**Частина 3: Пошук та вибір датасетів**

Датасет формувала самостійно із даних, які мені видала на запит платформа *Spotify. В мене їх вийшло два – 2023 і 2024 (дані прослуховування за 2023 і 2024 роки)*

**Частина 4: Розробка плану та структури проєкту**

План роботи:

1. Отримання даних (формування запиту в акаунті *Spotify, завантаження архіву з даними на ПК);*
2. Вивчення даних (визначення формату та способів обробки для приведення у формат, зручний для аналізу в**.**
3. Аналіз в **Tableau.**
4. Представлення результатів (дашборд, презентація).

**Частина 5: Розробка та аналіз даних**

Під час цього етапу мною було визначено, що початковий формат файлів, які я отримала - .json – не показує потрібної інформації, у зручному для аналізу форматі, коли його зчитати у Python саме як json, тому я перетворила його за допомогою бібліотеки asposecells у формат csv.

**import** json

**from** pprint **import** pprint

pip install aspose**-**cells

**import** jpype

**import** asposecells

jpype.**startJVM**()

**from** asposecells.api **import** Workbook

workbook **=** Workbook("StreamingHistory\_music\_1.json")

workbook.**save**("StreamingHistory\_music\_1.csv")

**Частина 6: Візуалізація та дашборд**

На мою думку, найінформативнішими в моєму випадку будуть стовпчикові діаграми, бо я аналізуватиму кількість та тривалість прослуховування певних пісень та виконавців, а також загальний час прослуховування музики протягом дня, місяця, року.

Візуальне оформлення висновків – Сanva.

Посилання на проект в Tableau - <https://public.tableau.com/views/SpotifyData_17146347047270/2023-2024?:language=en-US&:sid=&:display_count=n&:origin=viz_share_link>

**Частина 7: Презентація проєкту**

Посилання на презентацію на гугл-диску - <https://docs.google.com/presentation/d/1IcXYdtRtLVMr60BEV8Fh-ZZ8mqJY8_uT/edit?usp=sharing&ouid=110392508222704849095&rtpof=true&sd=true>

**Частина 8: GitHub репозиторій**

Посилання на репозиторій в **GitHub** - <https://github.com/valentun4uk/02052024>