Members:
22127008 – Đặng Châu Anh
22127014 – Nguyễn Kim Anh
22127147 – Đỗ Minh Huy
22127170 – Trần Dịu Huyền

PROPOSAL

I. Project topic

1. Project objectives:

Spotify, a leading music streaming platform, has gained significant popularity in Vietnam. Its charts offer real-time insights into the evolving music preferences of Vietnamese listeners. This project involves some actionable steps:

- Gathering Spotify streaming data specific to Vietnam.
- Analyzing trends, identifying popular songs, albums, and artists.
- Presenting the findings in a way that provides insights into local music preferences.

2. Project purpose:

The Vietnamese music industry has significant growth in 2024, with a surge in both quantity and quality of music products. This project uses Spotify data to better understand the changing tastes of Vietnamese listeners. By identifying popular songs, albums, and new trends, the project will provide useful insights to help industry professionals, artists, and fans stay updated on current music preferences.

3. Data Sources Details:

About:

- Data are scraped from two websites that provide weekly top 200 music charts in Viet Nam: https://charts.spotify.com/charts/overview/vn
- The data spans from January 1, 2024 – October 31, 2024
- The dataset consists of 9 columns and 2000 rows (approximately)

Attributes:

Attribute	Description	Data type (Estimated)
rank	The song's rank in top 200	int
uri	A unique identifier - URI for the song on Spotify	string
artist_names	The names of the artists involved in the song, including singers and producers	string
track_name	The song's title	string
source	Source company that release the track	string
peak_rank	The highest position the track has reached on the chart.	int
previous_rank	The track's rank on the previous chart.	int
days_on_chart	The number of days the track has remained on the chart.	int
streams	The total number of streams the track has received.	int

II. Planning

1. Method:

Phase	Tasks	Timeline	Deliverables
Business Understanding	- Define Objectives: Analyze genre trends, artist rankings, and optimal release times. - Assess Situation: Identify data sources, constraints, and resources. - Data Goals: Set specific data goals (e.g., tracking genre popularity).	Week 1	Project plan with objectives, timeline, and resources.
Data Understanding	- Data Collection: Gather data from streaming platforms, social media, and charts.	Weeks 2 – 3	Data summary and

	- Explore Data: Analyze patterns, distributions, and correlations. - Data Quality Check: Verify completeness and consistency.		initial EDA report.
Data Preparation	- Data Cleaning: Remove duplicates, handle missing values. - Feature Engineering: Create new features (e.g., monthly streaming trends). - Data Transformation: Encode, normalize, and integrate data.	Weeks 4 – 5	Cleaned, prepared dataset
Modeling	- Model Selection: Choose time-series models for trend analysis and clustering for artist segmentation. - Build Models: Analyze genre trends, artist rankings, and peak release periods. - Model Assessment: Validate model performance and interpret results.	Weeks 6 – 8	Model performance report and preliminary insights.
Evaluation	- Review Results: Ensure alignment with business objectives. - Refine Models: Adjust models as needed for accuracy.	Weeks 9	Evaluation report and key insights summary.
Deployment	- Dashboard Development: Create an interactive dashboard for trend and artist insights. - Final Report: Summarize findings with actionable recommendations. - Stakeholder Presentation: Walk through key insights and dashboard.	Weeks 10 – 12	Interactive dashboard, final report, presentation.

2. Technologies:

- Programming Language: Python
  - Data Manipulation and Analysis: Pandas, NumPy.
  - Data Visualization: Matplotlib, Seaborn, Plotly
  - Machine Learning and Statistical Analysis: Scikit-learn
  - Others: Requests (for handling HTTP requests if accessing APIs, JSON)
  - Data Analysis and Development Environment: Jupyter Notebook
- Version Control: Git and GitHub

III. Expected result

The expected results from this project are expected to provide a detailed overview of the Vietnamese music landscape in 2024, offering actionable insights for stakeholders in the industry. By identifying the populary trends, we determine the most popular music genres in Viet Nam and reveal which genres dominate the charts over time. Futhermore, we will find out the top artists in that year. By that way, the project will pinpoint the key players and potential growth areas within the Vienamese music industry. This analysis will provide valueable guidance for music participants. Additionally, we will find the mean of number of tracks released per month to identify peak periods for new music, examining the exact time to release a track will motivate the streaming engagement.

IV. References

Slides from lecturer