Team Name: HDMI Cord   **Group #120**
Shaun Heffernan, Landon Burchill, Nick Farr

Project Name: VinoWise

Problem: What problem are we trying to solve? [0.25 point]

The variety of wine becomes overwhelming for many people, people struggle to find
wine similar to the ones they enjoy. We are trying to assist people in selecting a wine similar to
ones they enjoy. Similar to music, people have distinct and unique tastes and opinions that can be
further enjoyed with optimization of Algorithmic driven recommendations. Similar to Spotify,
VinoWise will provide helpful and enjoyable recommendations, that both make the user try
something new and something that historically they like.

Motivation: Why is this a problem? [0.25 point]
Human Sommelier's are not always accurate when recommending a wine, and oftentimes people
find themselves disappointed when they spend a lot of money on a bottle of wine they do not
enjoy. This project has the potential to fix these issues and provide wine enjoyers a better
recommendation based on algorithmic and data driven decision making

Features: When do we know that we have solved the problem? [0.25 point]

We know we have solved the problem when we have a positive impact on the customers'
experience buying a new wine. Sorting and filtering options that allow users to customize their
search based on attributes such as price, variety, rating, and region. In addition, Accurate wine
recommendations based on user preferences and historical data will be another key indicator of
this project's success.

Data: (Public data set we will be using and the link to the public data set) or (Schema of randomly generated data - i.e. what are the different columns in our dataset and the respective data types) [0.25 point]

We will use this Public Kaggle Wine Reviews dataset:
https://www.kaggle.com/datasets/zynicide/wine-reviews/data

The dataset contains over 130,000 wine reviews with the following key attributes:

- Variety (str): Type of grape used.
- Country (str): Country of origin.
- Province (str): Region of the country.
- Winery (str): The producer of the wine.

- Points (int): Expert rating (e.g., 85–100 scale).
- Price (float): Cost of the bottle in USD.
- Description (str): Review text that describes taste and characteristics.

Tools: Programming languages or any tools/frameworks we will be using [0.25 point]

Python: Data cleaning, algorithm implementation, and backend processing.

- Flask: Web framework for front-end integration.
- Pandas & NumPy: Data manipulation and processing.
- Matplotlib & Seaborn: Visualization of wine trends and analytics.

Visuals: *Note: Wireframes and Concept Photos at the end of the document.*

- User input page where users can enter their favorite wine or preferences (e.g., price range, rating, variety).
- Results page displaying sorted recommendations with filtering options.
- Option to provide feedback to refine future recommendations.

Strategy: Preliminary _algorithms_ or _data structures_ you may want to implement and how would you represent the data [0.25 points]

- Sorting Algorithms:
  - Merge Sort: Stable sorting method for handling large datasets efficiently.
  - Quick Sort: Used for performance optimization when sorting by a single attribute.
  - Hybrid Approach: Combining sorting algorithms to handle different query types dynamically.
- Data Representation:
  - Store wines in a dictionary-based structure for efficient lookup. Another possibility is using a tree based on numeric attributes wine scores}
  - Implement a similarity scoring system that ranks wines based on multiple attributes (e.g., price, points, variety, and description analysis).

    Distribution of Responsibility and Roles: Who is responsible for what?
    Shaun Heffernan:
    - Responsible for designing, implementing, and testing various sorting algorithms Optimize the performance and time complexity of algorithms.
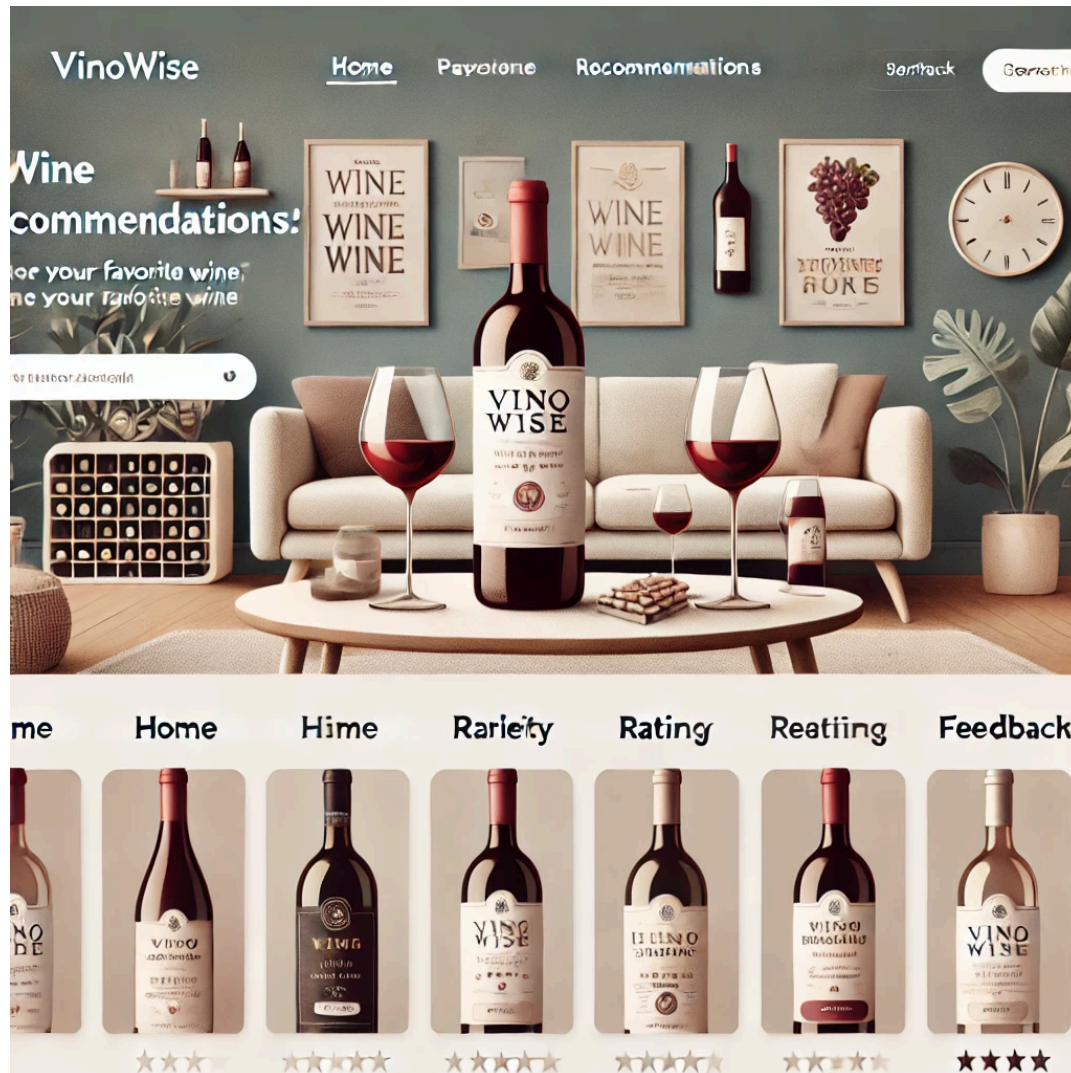   **Landon Burchill:**
    - Responsible for cleaning the data and converting it into a usable format effective for sorting algorithms. Building an efficient dictionary.

    Nicholas Farr: Responsible for designing the UI, and assisting both Shaun and Landon.

Extras:

Wireframes and Structure of the Application:

Concepts generated by DALLE by OpenAI:

~ Website will have one central search bar where you enter your favorite wine/ wine you want a similar wine recommendation. Then will have price filters and other filters to sort by. Once an input is entered a similar wine by a variety of metrics will then appear with other similar results

References

https://www.kaggle.com/datasets/zynicide/wine-reviews/data

https://github.com/zackthoutt/wine-deep-learning