



Drink Genie

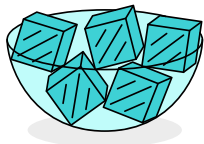
Axel Lopez

Introduction

Business Case

- The objective of this business case, is to develop a personalized drink recommendation system using Natural Language Processing (NLP) techniques.
- This system aims to provide tailored drink suggestions to users based on their preferred drink choices.

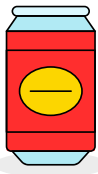


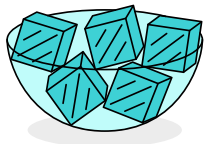


Data Collection



idDrink	strDrink	strTags	strCategory	strIBA	strAlcoholic
11000	Mojito	Iba	Cocktail	Contemporary Classics	Alcoholic
14366	Lemon Drop	New Era	Cocktail	None	Alcoholic

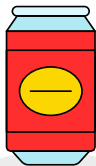




Data Collection



strGlass	strInstructions	strIngredients	...	strMeasure	...
Highball glass	Muddle mint leaves...	Light Rum	...	2	
Cocktail glass	Shake and strain into...	Vodka	Cocktail	2.25	



Data Wrangling

Techniques

The data wrangling process consisted of:

- Dropping unnecessary columns
- String methods to normalize the values
- Dict Mapping
- Filling Values
- Casting



Exploratory Data Analysis (EDA)

Describe:

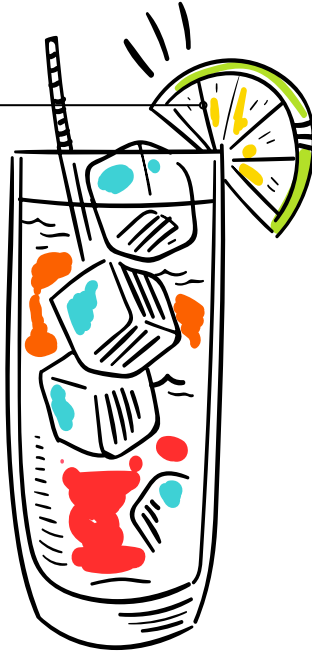
Object (count, unique, top, freq)

Numerical (mean, min, stdc, max)

Summary Statistics

	idDrink	strDrink	strTags	strCategory	strIBA	strAlcoholic	strGlass	strInstructions	strIngredient1
count	426	426	101	426	61	426	426	426	426
unique	426	426	44	11	3	3	31	391	137
top	17222	A1	IBA, Contemporary Classic	Ordinary Drink	Unforgettables	Alcoholic	cocktail glass	Shake all ingredients with ice, strain into a ...	Gin
freq	1	1	21	192	25	380	101	12	59

4 rows × 23 columns



Exploratory Data Analysis (EDA)

Describe:

Object (count, unique, top, freq)

Numerical (mean, min, stdc, max)

Summary Statistics

	strMeasure1	strMeasure2	strMeasure3	strMeasure4	strMeasure5	strMeasure6	strMeasure7
count	406.000000	406.000000	271.000000	171.000000	97.000000	35.000000	7.000000
mean	3.470121	2.557697	1.986986	2.642813	3.493361	2.940361	0.719976
std	6.312767	7.815599	5.341855	11.104402	14.560903	7.157672	0.680170
min	0.000000	0.000000	0.005000	0.021000	0.014000	0.003381	0.001691
25%	1.000000	0.500000	0.338000	0.250000	0.210000	0.200000	0.269070
50%	1.500000	1.000000	0.750000	0.625000	1.000000	1.000000	0.500000
75%	2.000000	2.000000	1.500000	1.500000	1.000000	1.000000	1.000000
max	50.000000	128.000000	67.000000	128.000000	128.000000	33.000000	2.000000



Feature Engineering

Process

For the purposes of the model and for one of the visualizations, feature engineering was employed.

- To create a Word Cloud in Tableau there was some usage of creating Dataframes, using melt and loc to modify this newly created dataframe
- For the model, there was a need to create a new column to be able to properly use the vectorizer



Model

Steps

To create the NLP model, the following steps had to be performed:

1. Preprocessing
2. Vectorizer
3. Training
4. Deployment



Integration & Development



Usage

- The recommendation system was put on a Streamlit website, and it will theoretically enhance user experience, increase sales and customer satisfaction.
- It will allow whichever company decides to utilize it to keep track on their Return on Investment (ROI) and other ratios they deem important.

Conclusion

