




# Capstone Project for The Cupping Room

---

Predicting Product Preferences for The Cupping Room  
Diane Zhu





# Table Of Contents

---

**01**

## **Introduction**

Background

**02**

## **Statistics**

Data Analysis

**03**

## **Conclusion**

Recommendations



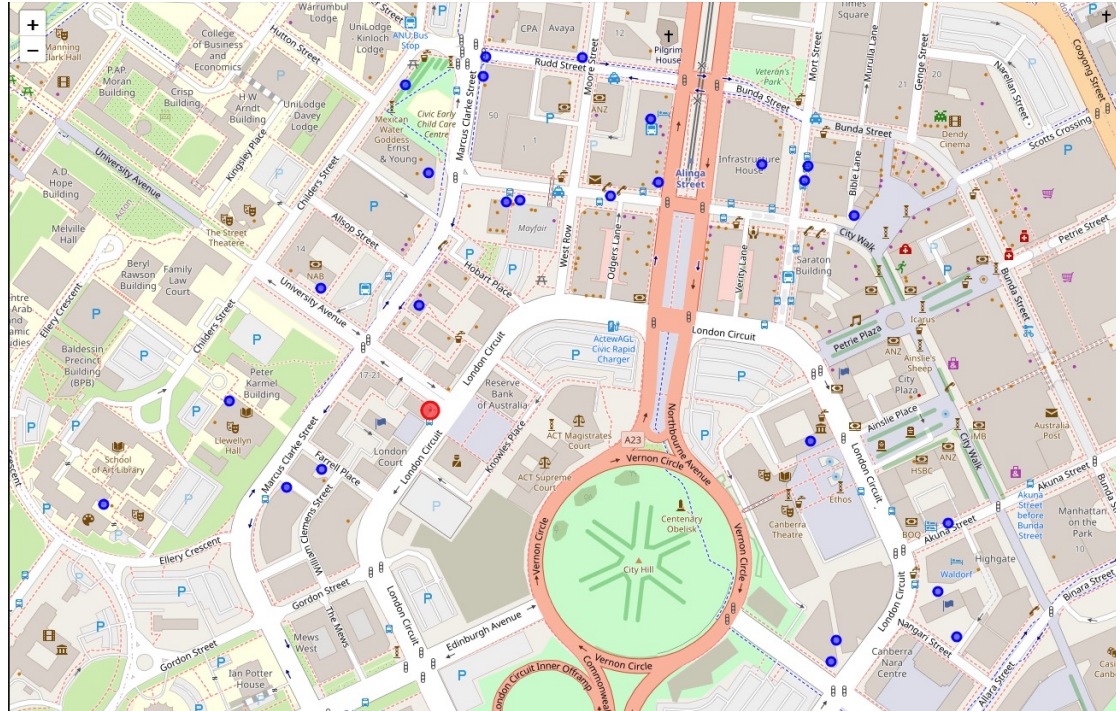


# The Cupping Room



The Cupping Room is a modern café located in Canberra, Australia, providing high-quality coffee experiences to customers for the past 10 years. The Cupping Room is designed by the team from Ona Coffee, which is Canberra's largest specialty coffee roaster, to share passion and knowledge about coffee with customers. The Cupping Room positions itself as a diverse café community and tries to open new café in Canberra to serve more customers.

# Existing Café in Canberra



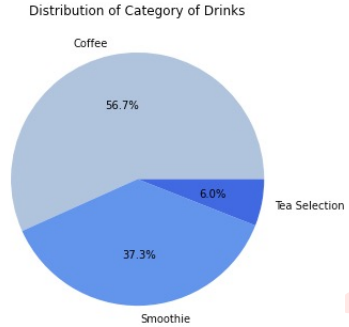
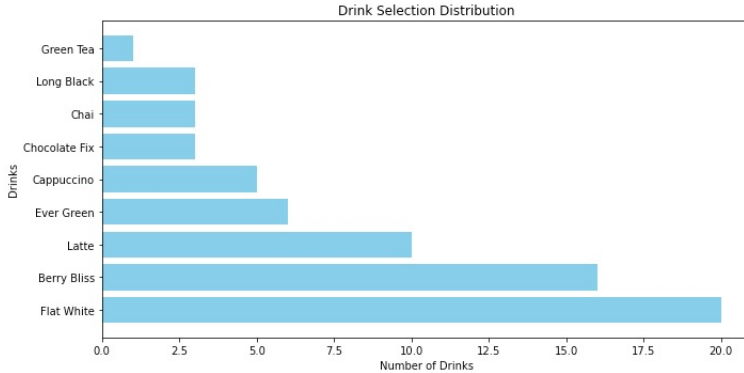
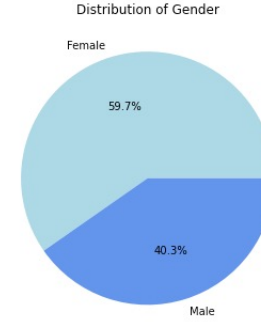
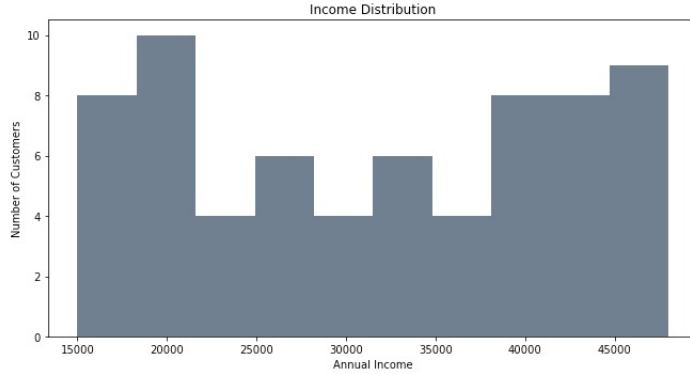
# Data Preparation

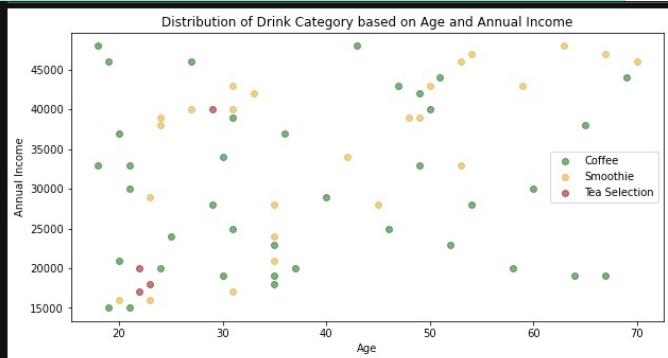
```
[42]: # encode categorical variables
# gender
le = preprocessing.LabelEncoder()
gender = coffee['Gender']
le.fit(gender)
gender_array = le.transform(gender)
coffee['gender_new'] = gender_array
display(coffee)
```

	drink	category	time	Gender	Age	Annual_Income	gender_new
0	Latte	Coffee	2019-08-15 07:36:47	Male	19.0	15000.0	1
1	Flat White	Coffee	2019-08-15 07:36:20	Male	21.0	15000.0	1
2	Ever Green	Smoothie	2019-08-15 07:33:56	Female	20.0	16000.0	0
3	Berry Bliss	Smoothie	2019-08-15 07:06:30	Female	23.0	16000.0	0
4	Chocolate Fix	Smoothie	2019-08-15 06:50:58	Female	31.0	17000.0	0
...	...	...	...	...	...	...	...
62	Berry Bliss	Smoothie	2019-07-18 07:14:38	Female	67.0	47000.0	0
63	Berry Bliss	Smoothie	2019-07-18 07:01:46	Female	54.0	47000.0	0
64	Berry Bliss	Smoothie	2019-07-18 07:00:34	Male	63.0	48000.0	1
65	Cappuccino	Coffee	2019-07-18 06:58:55	Male	18.0	48000.0	1
66	Latte	Coffee	2019-07-17 13:20:46	Female	43.0	48000.0	0

67 rows x 7 columns


# Exploratory Data Analysis






	drink	category	time	Gender	Age	Annual_Income	gender_new	category_index
0	Latte	Coffee	2019-08-15 07:36:47	Male	19.0	15000.0	1	2
1	Flat White	Coffee	2019-08-15 07:36:20	Male	21.0	15000.0	1	2
2	Ever Green	Smoothie	2019-08-15 07:33:56	Female	20.0	16000.0	0	1
3	Berry Bliss	Smoothie	2019-08-15 07:06:30	Female	23.0	16000.0	0	1
4	Chocolate Fix	Smoothie	2019-08-15 06:50:58	Female	31.0	17000.0	0	1
...	...	...	...	...	...	...	...	...
62	Berry Bliss	Smoothie	2019-07-18 07:14:38	Female	67.0	47000.0	0	1
63	Berry Bliss	Smoothie	2019-07-18 07:01:46	Female	54.0	47000.0	0	1
64	Berry Bliss	Smoothie	2019-07-18 07:00:34	Male	63.0	48000.0	1	1
65	Cappuccino	Coffee	2019-07-18 06:58:55	Male	18.0	48000.0	1	2
66	Latte	Coffee	2019-07-17 13:20:46	Female	43.0	48000.0	0	2

## Original Data Distribution



# Model Evaluation



The Confusion Matrix for KNN model is:

		PREDICTION			
		Tea	Smoothie	Coffee	
OBSERVATION	Tea	0	0	1	1
	Smoothie	0	2	3	5
	Coffee	0	5	10	15
		0	7	14	21

Classification Report for KNN Model

	precision	recall	f1-score	support
0	0.00	0.00	0.00	1
1	0.29	0.40	0.33	5
2	0.71	0.67	0.69	15
accuracy			0.57	21
macro avg	0.33	0.36	0.34	21
weighted avg	0.58	0.57	0.57	21



# Recommendations

---

Although the fictional dataset has its limitations when implementing the KNN model, it is still possible to find out the pattern of customer behaviors for The Cupping Room. First, it is highly recommended to develop more coffee products for the new café, and tea products may not be as favorable as coffee or smoothie. Second, customers with a higher annual income may prefer coffee over smoothie and tea. Therefore, when selecting new location for the café, The Cupping Room should take the surrounding offices or business areas into consideration.





# Thanks

---

Diane Zhu