

Submitted by:
Riddhi Singh

Take Home Assignment - Spoonshot

Before we start with the analysis, some cleaning and rearranging of the excel file provided must be done.

We clean the excel file by first combining all three sheets into one and removing irrelevant columns such as dish description and restaurant links.

We also remove the dollar sign from the "Dish Price" column in the sheet, so that all values processed can be seen as numbers when the analysis is done in python. Further, it is noted that there are some empty fields in the file which result in a #NAME? Or #VALUE! Error. These are removed later on during the EDA process.

We start the analysis by reading the file, saved as a CSV.

We read the column names in the now edited file. Following were the result:

```
RestaurantName_Type  
Restaurant_Type  
Restaurant_Menu  
Dish_Price  
Dish_Price_NoDollar  
Dish_Type  
Cuisine  
Dish_Course  
Advertised_ingredients  
Adv_Dishes
```

Then, I wanted to check which type of restaurant is most there in the city of New York. The reason for this is simple - Supply and Demand. One thing must exist in large quantities if the demand for the same from the public is quite large. What we get from this is that PUBS are the most popular type of restaurant. Therefore, if the restaurant 'Rays Gourmet Market' wishes to increase its popularity, it must have a bar, at the very least, inside the restaurant, so as to serve the crowd that visits pubs as well and redirect them there, which will implore them to not only try the experience there but also focus on the food as well, thereby increasing the ROI, the popularity and increase profit as well.

```
In [4]: #most common restaurant type
df.Restaurant_Type.mode()
```

```
Out[4]: 0    Pub
dtype: object
```

```
In [5]: #finding how many restaurant types are there
df['Restaurant_Type'].value_counts()
```

```
Out[5]: Pub                61
Pizzeria                 36
Café                    27
Barbecue restaurant      4
Brasserie and bistro      4
Steakhouse Restaurants    1
Name: Restaurant_Type, dtype: int64
```

Now, we cross match all the restaurants that have the same names as the Pubs names produced and remove the columns of restaurant names, types and the restaurant name menu column as well.

We get this output for further analysis.

	Dish_Price_NoDollar	Dish_Type	Cuisine	Dish_Course	Advertised_ingredients	Adv_Dishes
0	9.95	Dumpling	Sichuan	Starters	potato salad	Katz's Classic Luncheon Cold Cut Platter
1	7.95	Dumpling	Chinese	Starters	corned beef	Katz's Classic Luncheon Cold Sandwiches
2	8.95	Milk Based Drinks	Sichuan	Drinks	water	Bottled Water
3	8.95	Milk Based Drinks	Chinese	Drinks	sour pickles	Full Sour Pickles - 1 Qt
4	9.95	Dumpling	Chinese	Starters	potato salad	Small Potato Salad

Now, we remove the errors - #NAME? And #VALUE! Form the dish price column.

Due to the removal of dollar sign, certain values have errors in excel sheet. We must remove these errors before proceeding forward.

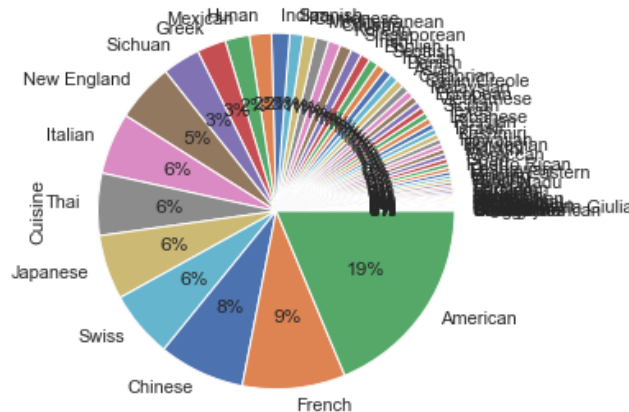
```
In [14]: df["Dish_Price_NoDollar"].replace({"#NAME?": "0"}, inplace=True)
print(df)
```

```
df["Dish_Price_NoDollar"].replace({"#VALUE!": "0"}, inplace=True)
print(df)
```

Now, I wanted to see the values presented in the Cuisine column and understand which one is most famous and available in the city.

```
df.Cuisine.value_counts().sort_values().plot(kind = 'pie', autopct = '%1.0f%%')
```

```
Out[19]: <AxesSubplot:ylabel='Cuisine'>
```



From this, we see that American cuisine is most available. However, for this particular restaurant to optimise itself, it must also focus on French and Chinese Cuisines as well to attract that crowd too.

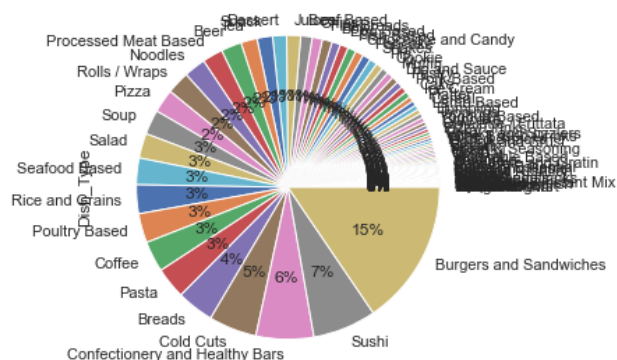
Similarly, we see the dish type available. Most widely sold are burgers and sandwiches, therefore Rays Gourmet Market must have this, apart from this, it must also sell Sushi, Breads, Coffee, along with poultry based products too. Others are optional.

```
In [46]: count_dish = count_of_dish_Course['Dish_Type'].value_counts()
count_dish
```

```
Out[46]: Burgers and Sandwiches      9389
Pasta                                2095
Cold Cuts                           2077
Poultry Based                       1901
Rice and Grains                     1789
Seafood Based                       1692
Rolls / Wraps                       1554
Noodles                             1316
Processed Meat Based                1313
Pizza                               1238
Sushi                               1238
Beef Based                          786
Flat Breads                         640
Fish Based                          571
Egg Based                           571
Bowl                                500
Pork Based                          428
Pancake                             358
Breads                              348
Lamb Based                          312
Salad                               292
Cereal                              272
Pie                                  205
Curry                              199
Omelette / Frittata                 198
Steaks and Sizzlers                 180
Soup                                168
Biriyani                            155
```

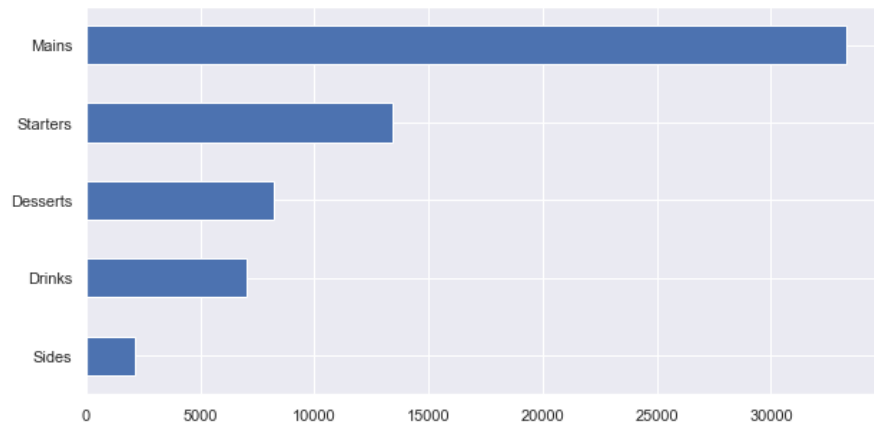
```
In [22]: df.Dish_Type.value_counts().sort_values().plot(kind = 'pie', autopct = '%1.0f%%')
```

```
Out[22]: <AxesSubplot:ylabel='Dish_Type'>
```



```
In [25]: df.Dish_Course.value_counts().sort_values().plot(kind = 'barh')
```

```
Out[25]: <AxesSubplot:>
```



Mains course is most sold in the city. Rays Gourmet Restaurant must have mains dishes.

Which dish type is famous under dish course?

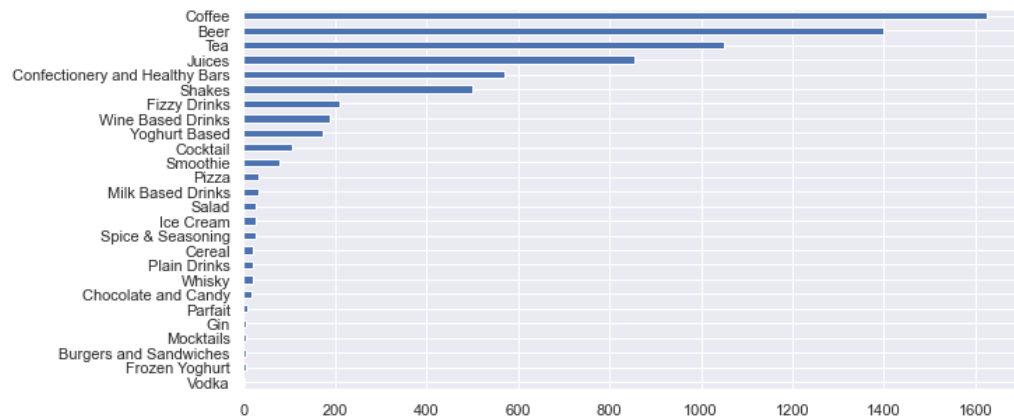
```
In [26]: a= df.loc[df['Dish_Course'] == 'Mains', 'Dish_Type']
a
```

```
Out[26]: 6 Meal
7 Meal
8 Yoghurt Based
14 Skewer
18 Meal
...
63822 Confectionery and Healthy Bars
63825 Confectionery and Healthy Bars
63994 Confectionery and Healthy Bars
64031 Confectionery and Healthy Bars
64135 Confectionery and Healthy Bars
Name: Dish_Type, Length: 33324, dtype: object
```

Next, we identify what all drinks must be on the menu at the restaurant:

```
In [30]: count_of_drinks_df.Dish_Type.value_counts().sort_values().plot(kind = 'barh')
```

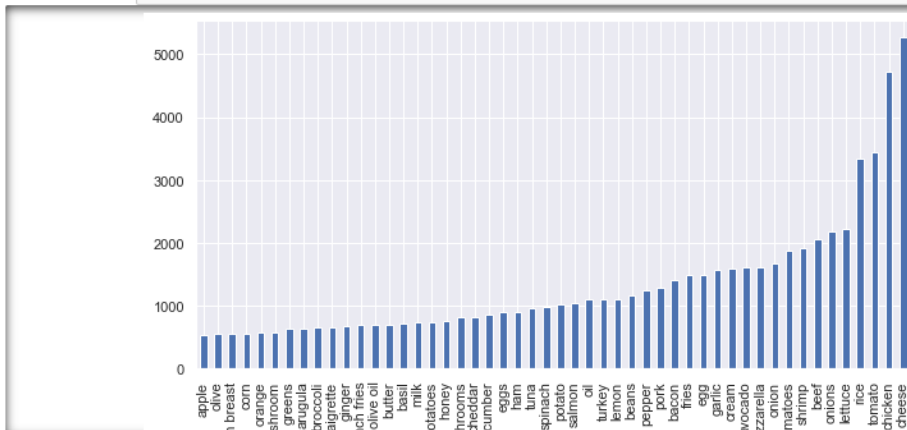
```
Out[30]: <AxesSubplot:>
```



While coffee, beer, tea, juices, shakes, wine based drinks, cocktails are present in large quantities, due to the pub/rest augmentation, they must also include options such as whiskey, gin, mocktails, vodka which are far less available in other pubs which can easily attract customers.

Lastly, we look at the advertising and costing section.

```
In [34]: adver_ingre.Advertised_ingredients.sort_values().plot(kind = 'bar')
plt.rcParams['figure.figsize'] = [2,8]
```



We find, that the most advertised products are cheese and chicken followed by tomato, rice and lettuce and more. Rays Gourmet Market must advertise these products which are likely to gain customers and retain the current ones as well.

```
In [43]: comp_dish_price_type.plot('Cuisine', 'Dish_Price_NoDollar', kind='bar')
plt.rcParams["figure.figsize"] = (12,42)
```

Lastly, I wanted to compare how much a person is willing to pay for the cuisines available. The most expensive cuisine just so happens to be Bengali with a maximum price of 31Dollars.

```
In [37]: comp_dish_price_type.loc[comp_dish_price_type['Dish_Price_NoDollar'].idxmax()]
```

```
Out[37]: Cuisine          Bengali
Dish_Price_NoDollar    31.916216
Name: 6, dtype: object
```

Finding the average price a person will pay for food.

The graph too, indicates the same(notebook)

```
In [38]: comp_dish_price_type.mean(axis=0)
```

```
C:\Users\riddh\AppData\Local\Temp\ipykernel_16256,
uctions (with 'numeric_only=None') is deprecated;
ore calling the reduction.
comp_dish_price_type.mean(axis=0)
```

```
Out[38]: Dish_Price_NoDollar    12.701099
dtype: float64
```

The average price a person is willing to pay at any restaurant is 12 Dollars. Therefore, Rays Gourmet must have prices around this range to increase any profit in the sales they generate.

Ultimately, from this analysis, these are a few of the things that the restaurant must look into to perform better:

- 1) Add a bar to their restaurant
- 2) The bar must have of course all the popular drinks but should include whiskey, gin, mocktails, vodka as well.
- 3) American cuisine is a must. Focus must also be shed on French and Chinese cuisines.
- 4) Mains as a dish type must be there along with starters. Additionally, deserts and sides must also be there for the full experience.
- 5) While mains dishes must be there, sushi, burgers and sandwiches, breads, pasta must also be available.
- 6) Coffee is the most popular drink which should be on the menu with tea, juices, shakes etc.
- 7) Bengali cuisine is most expensive.
- 8) Average spending a person is willing to do is 12 dollars. Prices must be based around this.