

```
In [46]: import pandas as pd
import numpy as np
import seaborn as sns
hotpot = pd.read_csv("hotpot.csv")
hotpot.head()
```

Out[46]:

	bundleID	soupbase	maindishl	maindishll	dipping	snack	avg_rating
0	1	Tomato Soup Base	Tofu Pudding	Smashed Shrimp Paste	Fragrant	Deep-Fried Crispy Pork	5.47
1	2	Tomato Soup Base	Tofu Pudding	Smashed Shrimp Paste	Fragrant	Dough Stick With Fennel	7.88
2	3	Tomato Soup Base	Tofu Pudding	Smashed Shrimp Paste	Spicy	Deep-Fried Crispy Pork	6.20
3	4	Tomato Soup Base	Tofu Pudding	Smashed Shrimp Paste	Spicy	Dough Stick With Fennel	8.00
4	5	Tomato Soup Base	Tofu Pudding	Smashed Shrimp Paste	Seafood	Deep-Fried Crispy Pork	8.24

```
In [4]: hotpot.isnull().values.any()
```

Out[4]: False

As we can see above that there are no null value in the dataset.

```
In [10]: hotpot.describe()
```

Out[10]:

	bundleID	avg_rating
count	384.000000	384.000000
mean	192.500000	6.758594
std	110.995495	1.670385
min	1.000000	1.690000
25%	96.750000	5.570000
50%	192.500000	6.810000
75%	288.250000	7.972500
max	384.000000	10.160000

The describe function gives the descriptive statistics of the variables which summarises the distribution of the variables in the dataset.

```
In [6]: hotpot_new = pd.get_dummies(hotpot, drop_first=True, columns=['soupbase',
    , 'maindishI',
    'maindishII', 'dipping', 'snack'])
```

The `get_dummies` function is used above to convert the categorical value into the dummy values.

```
In [7]: hotpot_new.columns
```

```
Out[7]: Index(['bundleID', 'avg_rating', 'soupbase_Mushroom Soup Base',
    'soupbase_Three Delicacies Soup Base', 'soupbase_Tomato Soup Bas
    e',
    'maindishI_Special Mutton', 'maindishI_Spicy Marinated Beef',
    'maindishI_Tofu Pudding', 'maindishII_Fish Cakes',
    'maindishII_Lobster Tails', 'maindishII_Smashed Shrimp Paste',
    'dipping_Seafood', 'dipping_Spicy', 'snack_Dough Stick With Fenn
    el'],
    dtype='object')
```

Determining the coefficient values of your model inputs.

```
In [8]: X = hotpot_new[['soupbase_Mushroom Soup Base',
    'soupbase_Three Delicacies Soup Base', 'soupbase_Tomato Soup Bas
    e',
    'maindishI_Special Mutton', 'maindishI_Spicy Marinated Beef',
    'maindishI_Tofu Pudding', 'maindishII_Fish Cakes',
    'maindishII_Lobster Tails', 'maindishII_Smashed Shrimp Paste',
    'dipping_Seafood', 'dipping_Spicy', 'snack_Dough Stick With Fenne
    l']]

y = hotpot_new['avg_rating']
```

Building the linear model using the outcome variable Average rating

```
In [11]: from sklearn.linear_model import LinearRegression
    from sklearn import metrics
    lmodel = LinearRegression()
    lmodel.fit(X,y)
```

```
Out[11]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normaliz
    e=False)
```

```
In [12]: lmodel.intercept_
```

```
Out[12]: 6.161718750000002
```

```
In [13]: coef_hotpot = pd.DataFrame(lmodel.coef_,X.columns, columns=['Coefficient',
coef_hotpot
```

Out[13]:

	Coefficient
<b>soupbase_Mushroom Soup Base</b>	0.058750
<b>soupbase_Three Delicacies Soup Base</b>	-0.123229
<b>soupbase_Tomato Soup Base</b>	0.100521
<b>maindishI_Special Mutton</b>	0.805521
<b>maindishI_Spicy Marinated Beef</b>	0.831979
<b>maindishI_Tofu Pudding</b>	0.223542
<b>maindishII_Fish Cakes</b>	0.194375
<b>maindishII_Lobster Tails</b>	-0.038438
<b>maindishII_Smashed Shrimp Paste</b>	0.354687
<b>dipping_Seafood</b>	-0.117344
<b>dipping_Spicy</b>	-0.119844
<b>snack_Dough Stick With Fennel</b>	0.148021

From the model we can see that Special mutton and the spicy marinated beef has the highest positive coefficient. This infers that the customers like these dish the most and would have a positive effect on the revenue generate. This might be because the carnival attendees prefer spicy food during the winter season. Also the fact even though maine is famous all over the world for its fresh lobster, the customers prefer meat over the lobster during winter may be because the reason that the lobster fishing is readuced drastically and the customers are skeptical about the freshness of the lobster. This might be the reason why lobster tail has a negative coefficient. Surprisingly the customers prefer Fish cake and sprimp paste main dish and has a positive coefficient. The spicy dip, seafood dip and the three delicacies soup base has the most negative coefficient. The lobsterland management should focus on including the dishes that have positive coefficient during the memo selection.

```
In [14]: hotpot_avg_rating=hotpot.sort_values(by="avg_rating",ascending=False)
```

In [15]: hotpot\_avg\_rating

Out[15]:

	<b>bundleID</b>	<b>soupbase</b>	<b>maindishI</b>	<b>maindishII</b>	<b>dipping</b>	<b>snack</b>	<b>avg_rating</b>
<b>221</b>	222	Mushroom Soup Base	Spicy Marinated Beef	Smashed Shrimp Paste	Seafood	Dough Stick With Fennel	10.16
<b>47</b>	48	Tomato Soup Base	Spicy Marinated Beef	Fish Cakes	Seafood	Dough Stick With Fennel	9.97
<b>223</b>	224	Mushroom Soup Base	Spicy Marinated Beef	Crab Sticks	Fragrant	Dough Stick With Fennel	9.96
<b>75</b>	76	Tomato Soup Base	Special Mutton	Smashed Shrimp Paste	Spicy	Dough Stick With Fennel	9.93
<b>40</b>	41	Tomato Soup Base	Spicy Marinated Beef	Lobster Tails	Seafood	Deep-Fried Crispy Pork	9.93
...	...	...	...	...	...	...	...
<b>293</b>	294	Three Delicacies Soup Base	Tofu Pudding	Smashed Shrimp Paste	Seafood	Dough Stick With Fennel	3.24
<b>353</b>	354	Three Delicacies Soup Base	Beef Tripe	Lobster Tails	Seafood	Dough Stick With Fennel	3.20
<b>93</b>	94	Tomato Soup Base	Special Mutton	Fish Cakes	Spicy	Dough Stick With Fennel	3.12
<b>76</b>	77	Tomato Soup Base	Special Mutton	Smashed Shrimp Paste	Seafood	Deep-Fried Crispy Pork	2.63
<b>64</b>	65	Tomato Soup Base	Beef Tripe	Lobster Tails	Seafood	Deep-Fried Crispy Pork	1.69

384 rows × 7 columns

The average customer rating of the food bundle with Mushroom soup base , spicy marinated beef, smashed shrimp paste, seafood dipping and dough stick with fennel snack is highest.

In [28]: vendor\_cost= pd.read\_csv( 'vendor\_costs.csv' )

In [29]: `vendor_cost`

Out[29]:

	<b>Item</b>	<b>Item Category</b>	<b>Cost Per Serving (Dollars)</b>
<b>0</b>	Tomato Soup Base	Soup Base	3.40
<b>1</b>	Classic Spicy Soup Base	Soup Base	3.50
<b>2</b>	Mushroom Soup Base	Soup Base	2.90
<b>3</b>	Three Delicacies Soup Base	Soup Base	4.60
<b>4</b>	Tofu Pudding	Main Dish I	5.70
<b>5</b>	Spicy Marinated Beef	Main Dish I	6.00
<b>6</b>	Beef Tripe	Main Dish I	5.30
<b>7</b>	Special Mutton	Main Dish I	6.10
<b>8</b>	Smashed Shrimp Paste	Main Dish II	5.30
<b>9</b>	Crab Sticks	Main Dish II	4.70
<b>10</b>	Lobster Tails	Main Dish II	6.00
<b>11</b>	Fish Cakes	Main Dish II	4.70
<b>12</b>	Fragrant	Dipping Sauce	0.20
<b>13</b>	Spicy	Dipping Sauce	0.10
<b>14</b>	Seafood	Dipping Sauce	0.15
<b>15</b>	Deep-Fried Crispy Pork	Snack	0.60
<b>16</b>	Dough Stick With Fennel	Snack	0.70

```
In [30]: vendor_cost= vendor_cost[['Item', 'Cost Per Serving (Dollars) ']]
         vendor_cost
```

Out[30]:

	<b>Item</b>	<b>Cost Per Serving (Dollars)</b>
0	Tomato Soup Base	3.40
1	Classic Spicy Soup Base	3.50
2	Mushroom Soup Base	2.90
3	Three Delicacies Soup Base	4.60
4	Tofu Pudding	5.70
5	Spicy Marinated Beef	6.00
6	Beef Tripe	5.30
7	Special Mutton	6.10
8	Smashed Shrimp Paste	5.30
9	Crab Sticks	4.70
10	Lobster Tails	6.00
11	Fish Cakes	4.70
12	Fragrant	0.20
13	Spicy	0.10
14	Seafood	0.15
15	Deep-Fried Crispy Pork	0.60
16	Dough Stick With Fennel	0.70

```
In [38]: all_items=[ 'Tomato Soup Base',
                    'Classic Spicy Soup Base',
                    'Mushroom Soup Base',
                    'Three Delicacies Soup Base',
                    'Tofu Pudding',
                    'Spicy Marinated Beef',
                    'Beef Tripe',
                    'Special Mutton',
                    'Smashed Shrimp Paste',
                    'Crab Sticks',
                    'Lobster Tails',
                    'Fish Cakes',
                    'Fragrant',
                    'Spicy',
                    'Seafood',
                    'Deep Fried Crispy Pork',
                    'Dough Stick with Fennel']

all_cost=[3.40 ,3.50 ,2.90 ,4.60 ,5.70 ,6.00 ,5.30 ,6.10 ,5.30 ,4.70 ,6.
00 ,4.70 ,0.20 ,0.10 ,0.15 ,0.60 ,0.70]
```

```
In [39]: df2 = pd.DataFrame({'all_items':all_items, 'all_cost':all_cost})
```

```
In [40]: bundle_222= ['Mushroom Soup Base', 'Spicy Marinated Beef', 'Smashed Shrimp Paste', 'Seafood', 'Dough Stick with Fennel']
bundle_48=['Tomato Soup Base', 'Spicy Marinated Beef', 'Fish Cakes', 'Seafood', 'Dough Stick with Fennel']
bundle_223=['Mushroom Soup Base', 'Spicy Marinated Beef', 'Crab Sticks', 'Fragrant', 'Dough Stick with Fennel']
bundle_75=['Tomato Soup Base', 'Special Mutton', 'Smashed Shrimp Paste', 'Spicy', 'Dough Stick with Fennel']
```

```
In [41]: df1 = pd.DataFrame({'bundle_222':bundle_222, 'bundle_48':bundle_48, 'bundle_223':bundle_223, 'bundle_75':bundle_75})
```

```
In [43]: rename_dict = df2.set_index('all_items').to_dict()['all_cost']
```

```
In [44]: df1 = df1.replace(rename_dict)
df1
```

Out[44]:

	<b>bundle_222</b>	<b>bundle_48</b>	<b>bundle_223</b>	<b>bundle_75</b>
<b>0</b>	2.90	3.40	2.9	3.4
<b>1</b>	6.00	6.00	6.0	6.1
<b>2</b>	5.30	4.70	4.7	5.3
<b>3</b>	0.15	0.15	0.2	0.1
<b>4</b>	0.70	0.70	0.7	0.7

```
In [45]: df1.sum()
```

```
Out[45]: bundle_222    15.05
bundle_48    14.95
bundle_223    14.50
bundle_75    15.60
dtype: float64
```

Since the price of the hotpot is set as standard 15 dollars for all the customers, the bundle with Tomato Soup Base, Spicy Marinated Beef, Fish Cakes, Seafood and Dough Stick With Fennel costs under 15 dollars. The Lobsterland management should consider this menu selection in order to provide the customer of the hotpot a very satisfying experience and to create an opportunity to generate higher revenue from the Hotpot.

## ***One Page Memo***

*Soup: Tomato Soup Base*

*Main Dish 1: Spicy Marinated Beef*

*Main Dish 2: Fish Cake*

*Dip: Seafood*

*Snack: Dough Stick With Fennel*