Meal Price Optimization

½ Consulting, Inc.

Group 1:
Qiuxia Chen (Katalia)
Xuechen He (Celia)
Mei Lu (May)
Junkai Zhong (Caesar)

Structure

- Background
- Data Analysis
- Modeling
- Conclusion

Background

Meal Delivery Company

Multiple Regions and Cities

Various Fulfillment Centers

Dispatching Meal Order

14 Meal Categories, 4 Cuisines – 51 Meal IDs

Industry Overview

Physical Footprint

15,000	Distribution centers	*Texas, California, and Florida are leading states*
288 million	Square feet of distribution space	*Equivalent to 5,000 football fields*
8.7 billion	Cases delivered	*Nearly 24 million cases per day*
3.2 billion	Vehicle miles per year	*Around the Earth 330,000 times, or 1,700 roundtrips to the sun*
820 million	Gallons of fuel per year	*More than 1,240 Olympic-size swimming pools*
131,000	Drivers	*3.75 percent of all truck drivers in the U.S.*

Economic Footprint

$280 billion	Industry annual sales	*Approximately the GDP of Louisiana*
350,000	Employed by industry	*5.25 percent of all wholesale jobs*
700,000	Ancillary jobs	*Industry + ancillary jobs = population of Delaware*
$51 billion	U.S. GDP	*0.25 percent of the total U.S. economy*
$14 billion	Federal, state, and local tax revenues	*$7.2 billion in federal taxes, and $6.9 billion in state and local taxes*

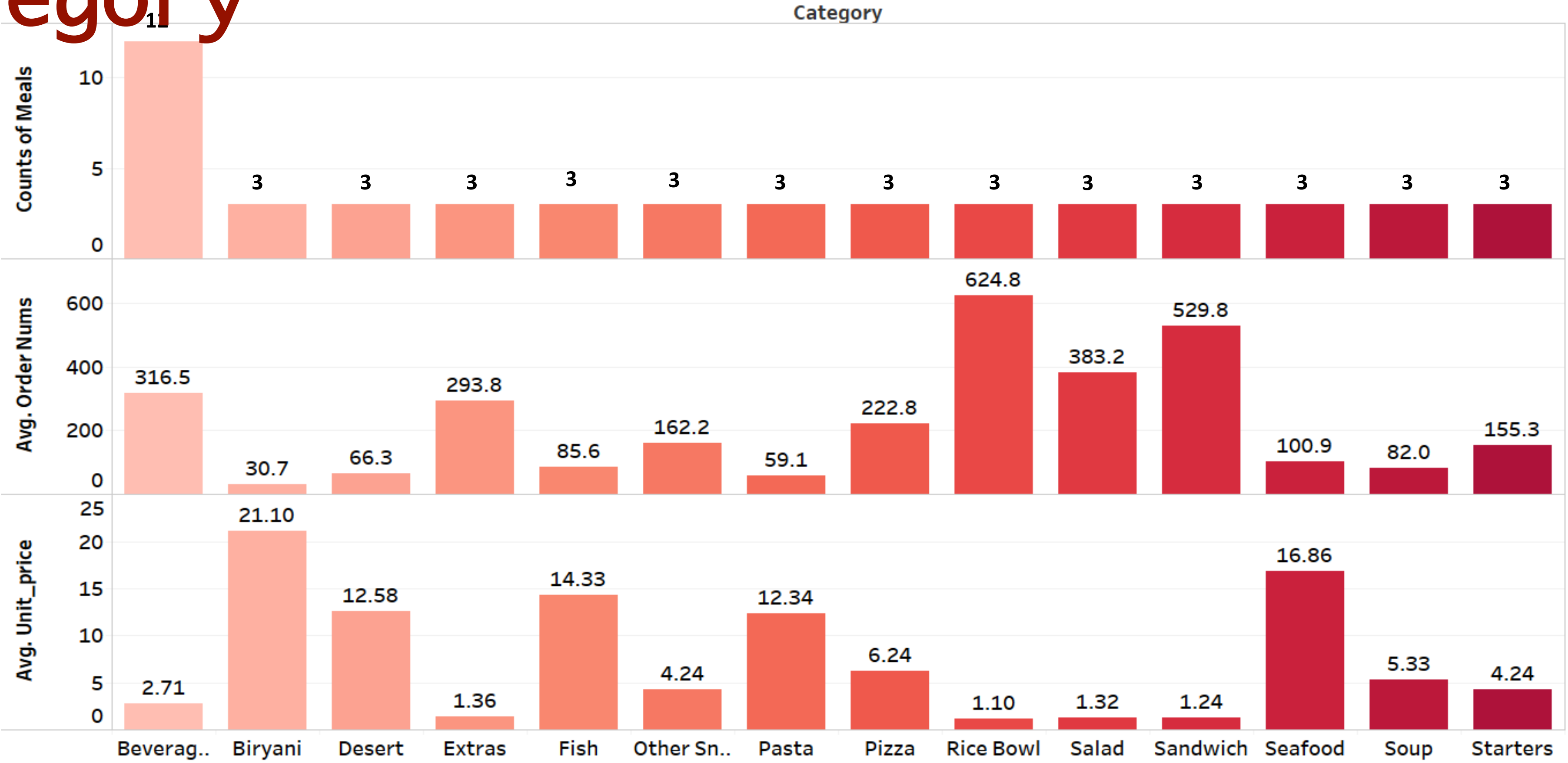
Data Overview

Variable	Definition
id	Unique ID
week	Week No
center_id	Unique ID for fulfillment center
meal_id	Unique ID for the meal
checkout_price	Final price including discount, taxes & delivery charges
base_price	Base price of the meal
emailer_for_promotion	Email sent for promotion of meal
homepage_featured	Meal featured at homepage
num_orders	(Target) Orders Count

Variable	Definition
center_id	Unique ID for fulfillment center
city_code	Unique code for city
region_code	Unique code for region
center_type	Anonymized center type
op_area	Area of operation (in km^2)

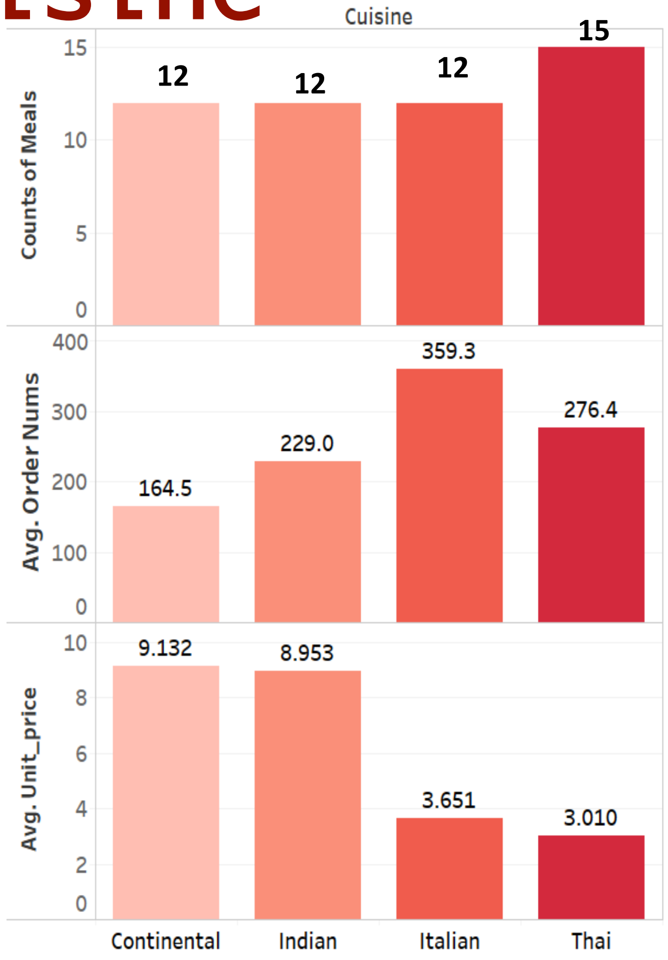
Variable	Definition
meal_id	Unique ID for the meal
category	Type of meal (beverages/snacks/soups…)
cuisine	Meal cuisine (Indian/Italian/…)

Exploratory Data Analysis: by Category



- 14 categories, 51 meal products
- Most popular meal: Rice Bowl, Sandwich and Salad
- Most expensive meal: Biryani, Seafood and Fish
- The order numbers seem to be negatively correlated with the unit price

Exploratory Data Analysis: by Cuisine



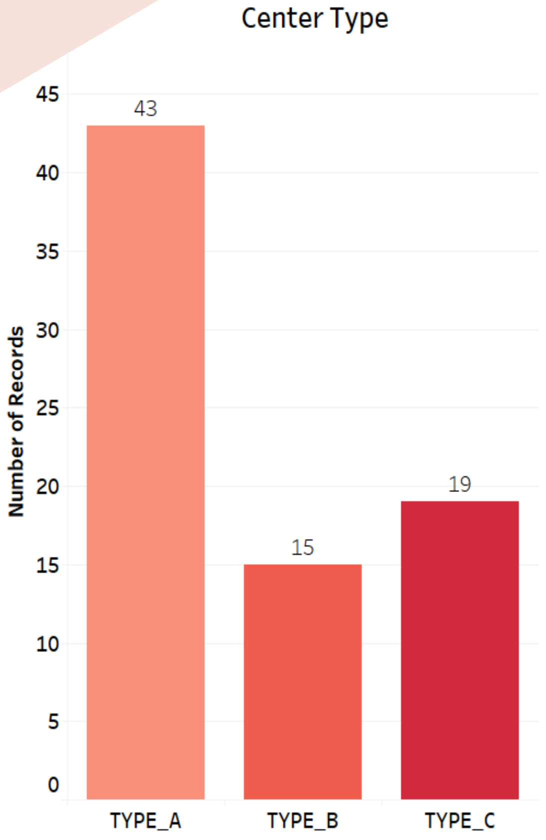
- 4 cuisines
- Order numbers: Italian > Thai > Indian > Continental
- Unit price: Continental > Indian > Italian > Thai

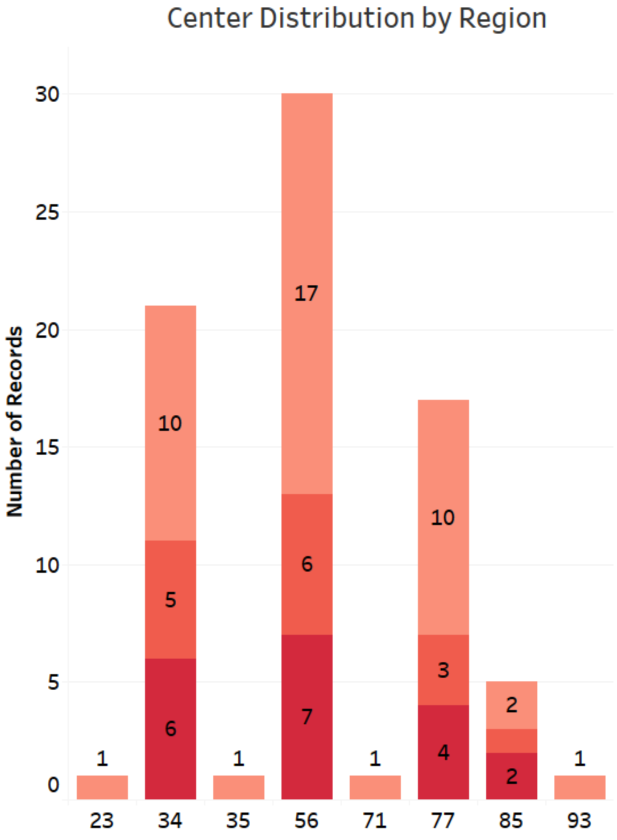
Category- Cuisine combination

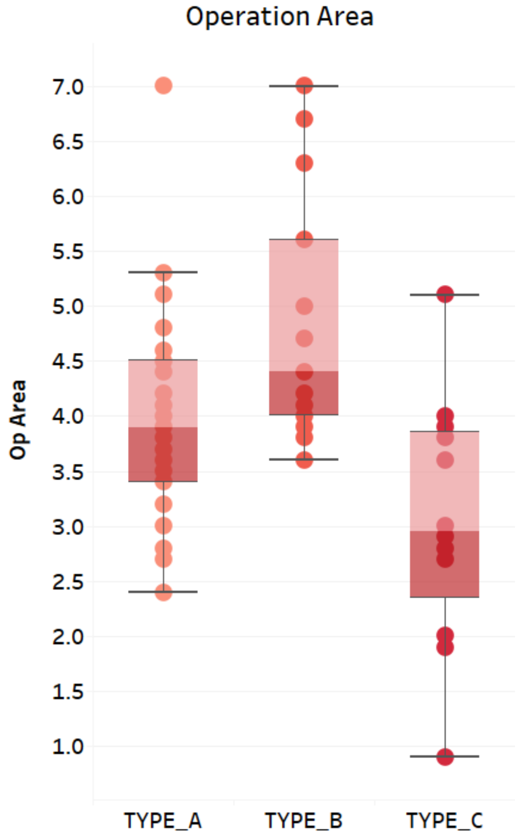
Category	Continental	Cuisine Indian	Italian	Thai
Beverages	3	3	3	3
Biryani		3		
Desert		3		
Extras				3
Fish	3			
Other Snacks				3
Pasta			3	
Pizza	3			
Rice Bowl		3		
Salad			3	
Sandwich			3	
Seafood	3			
Soup				3
Starters				3

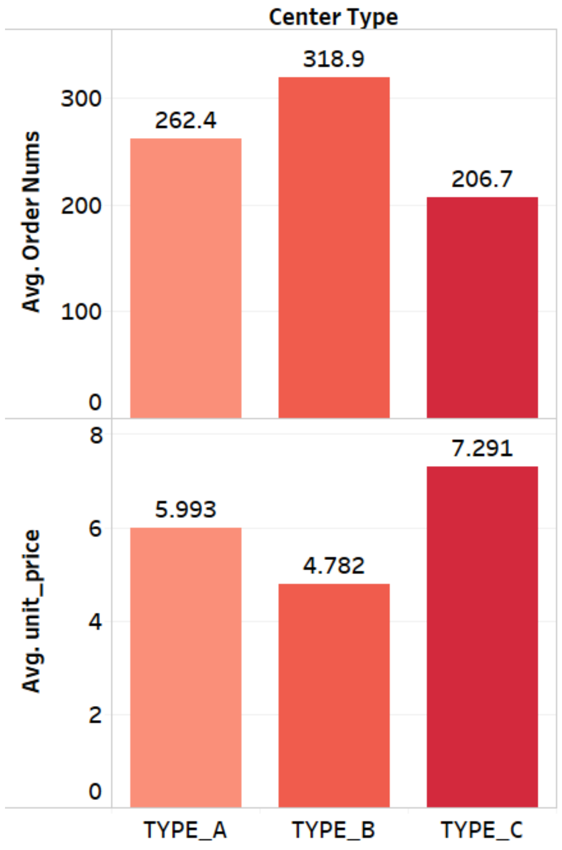
- Beverage 4 favors, other categories belong to one specific cuisine
- Each cuisine contain 4~5 different kinds of meals

Exploratory Data Analysis: by Center









- 3 types, 77 centers
- number: type A > type C > type B

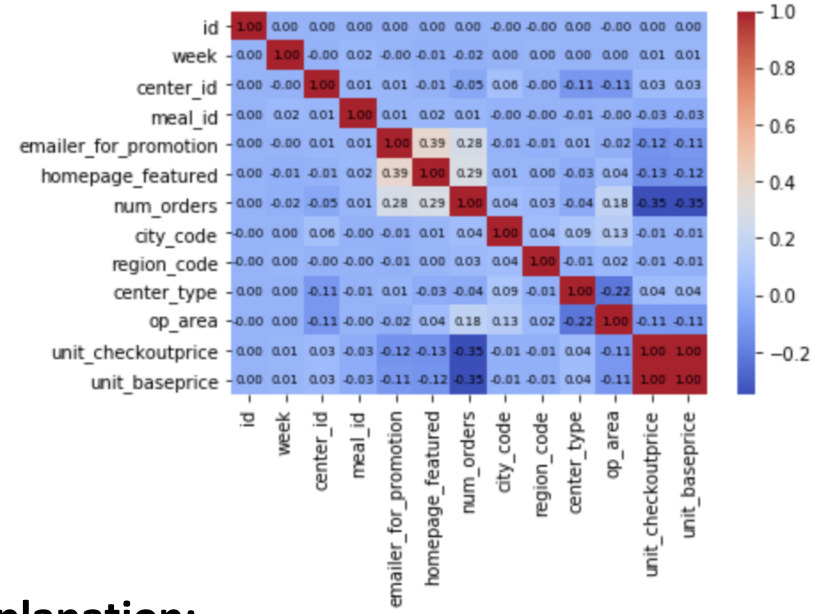
- 8 regions
- Type A center in every region

- Average area: B > A > C

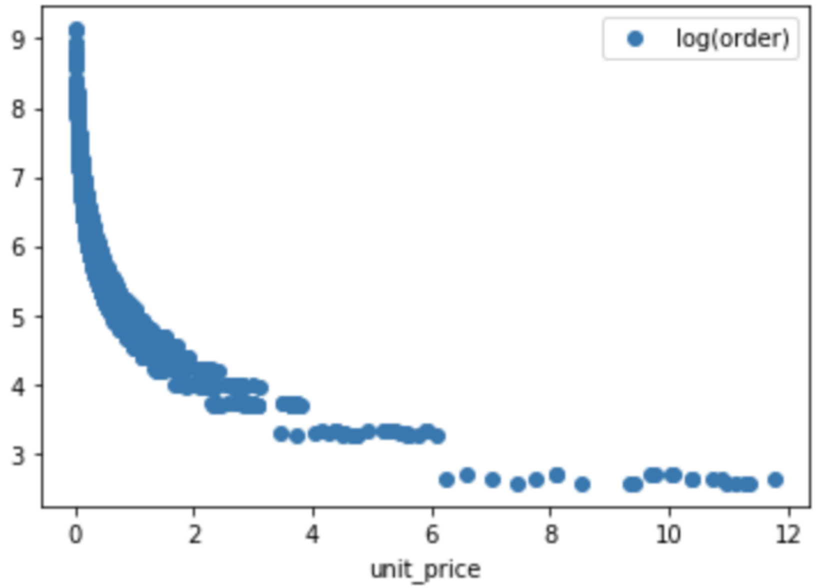
- Order number: B > A > C
- Unit price: C > A > B

How to build Demand Model

Correlation Matrix



log linear model



Data explanation:

It is shown that the number of orders is highly correlated with unit price, emailer promotion, homepage featured, and op_ a rea.

The relationship between unit checkout price and number of orders, we will run a log linear model

Demand Model

- center_type A, B, C as dummy variables **type_A, type_B and type_C**
- Emailer_for_promotion, homepage_featured are binary variables
- Log(num_orders) is dependent variable

$$\text{Log(D)} = \beta_0 + \beta_1 \text{ unit\_checkout\_price} + \beta_2 \text{ promotion} + \beta_3 \text{ featured} + \beta_4 \text{ area} + \beta_5 \text{ type\_A} + \beta_6 \text{ type\_B} + \beta_7 \text{ type\_C}$$

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.6674	0.004	974.947	0.000	3.660	3.675
unit_checkoutprice	-0.0965	0.000	-778.615	0.000	-0.097	-0.096
emailer_for_promotion	0.4284	0.004	96.660	0.000	0.420	0.437
homepage_featured	0.4370	0.004	112.536	0.000	0.429	0.445
op_area	0.1183	0.001	102.187	0.000	0.116	0.121
center_type_TYPE_A	1.2224	0.002	652.833	0.000	1.219	1.226
center_type_TYPE_B	1.2610	0.003	441.239	0.000	1.255	1.267
center_type_TYPE_C	1.1840	0.002	600.397	0.000	1.180	1.188

Model for Cuisines

Cuisines	Model	Price Coff.	R-square
Thai	Log(D) = 3.8 - 0.19*unitprice + 0.16*promotion + 0.61*featured + 0.11*area + 1.26*type_A + 1.42*type_B + 1.11*type_C	-0.1857	0.694
Italian	Log(D) = 3.8 - 0.13*unitprice + 0.42*promotion + 0.41*featured + 0.17*area + 1.25*type_A + 1.19*type_B + 1.35*type_C	-0.1327	0.667
Indian	Log(D) = 3.66 - 0.1*unitprice + 0.85*promotion + 0.19*featured + 0.09*area + 1.2*type_A + 1.25*type_B + 1.21*type_C	-0.0952	0.642
Continental	Log(D) =3.56 - 0.06*unitprice + 0.39*promotion + 0.49*featured + 0.08*area + 1.19*type_A + 1.09*type_B + 1.28*type_C	-0.0641	0.768

Data explanation:

The coefficient on unit price for Continental, Indian, Italian, Thai food are 0.06 < 0.1< 0.13 < 0.19 respectively, we can conclude that the Thai food is more price sensitivity than others.

Price Sensitivity for Categories

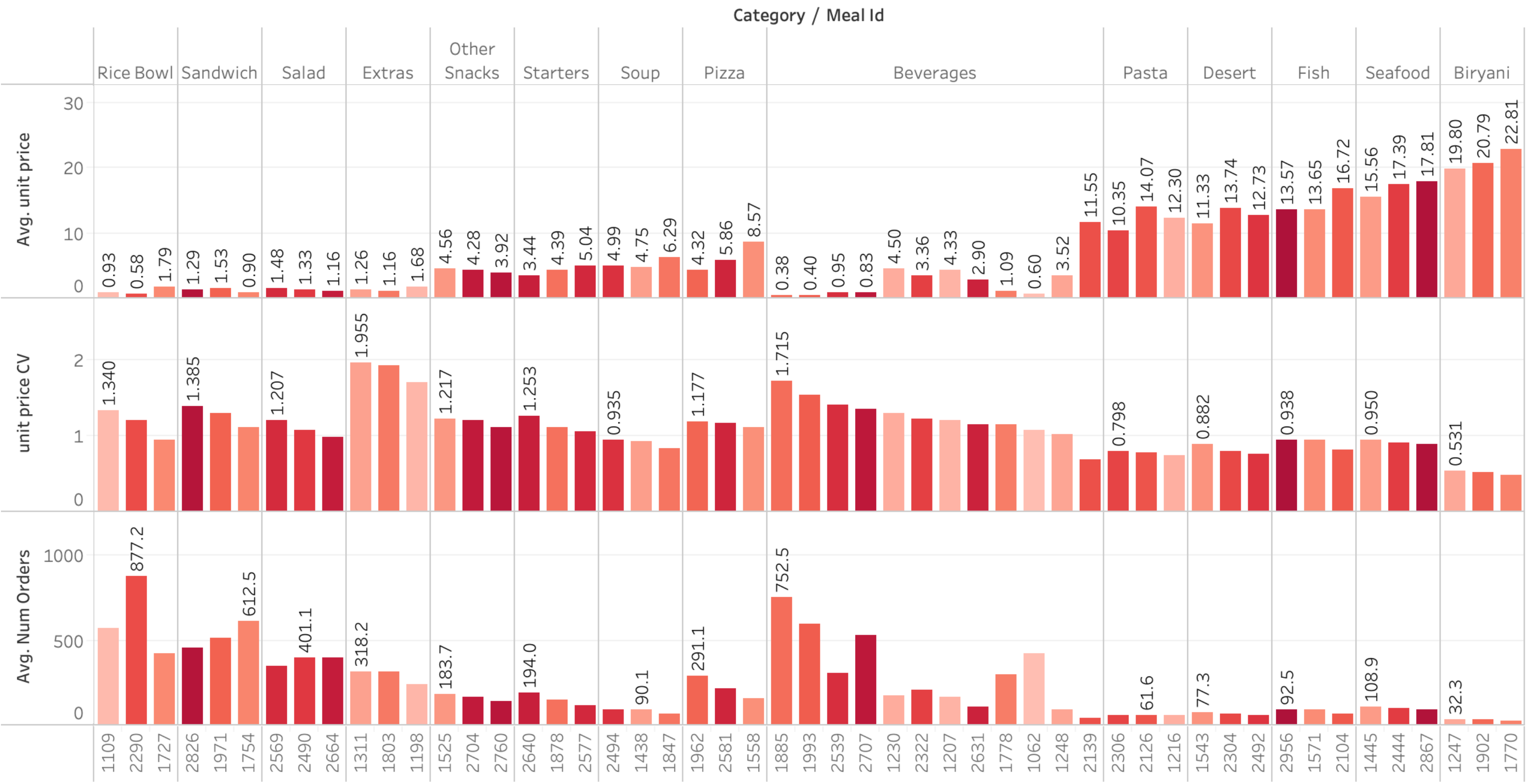
Coefficient of Unit Checkout Price

Beverage	**Desert**	Fish	Pasta	Rice Bowl	Sandwich	Soup
-0.17	**-0.67**	**-0.06**	**-0.07**	**-0.34**	**-0.3**	**-0.15**
Biryani	Extras	Other Snacks	Pizza	Salad	**Seafood**	Starters
-0.05	**-0.19**	**-0.14**	**-0.08**	**-0.32**	**-0.04**	**-0.14**

Data explanation:

- Desert has the largest price sensitivity, price increase by $1, the demand decrease by 67%
- Seafood has the smallest price sensitivity, demand change by 4% with the price goes up $1.

Item Selection



Average of unit price, unit price CV and average of Num Orders for each Meal Id broken down by Category. Color shows details about Meal Id. The view is filtered on Category, which keeps 14 of 14 members.

We choose three items with the highest price coefficient of variation in different unit level and categories
Beverage : Meal id 1885 ($0.17), Starters: Meal id 2640($3.44), Desert: Meal id 1543($11.33)

Item Based Model

Desert - 1543

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.4151	0.012	295.277	0.000	3.392	3.438
unit_checkoutprice	-0.0681	0.000	-205.273	0.000	-0.069	-0.067
emailer_for_promotion	0.2524	0.013	19.491	0.000	0.227	0.278
homepage_featured	0.0524	0.009	5.571	0.000	0.034	0.071
op_area	0.0647	0.003	19.138	0.000	0.058	0.071
center_type_TYPE_A	1.0713	0.006	192.223	0.000	1.060	1.082
center_type_TYPE_B	1.2456	0.008	151.650	0.000	1.230	1.262
center_type_TYPE_C	1.0981	0.006	194.089	0.000	1.087	1.109

R-square: 0.842

Starters - 2640

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.6843	0.014	260.614	0.000	3.657	3.712
unit_checkoutprice	-0.1411	0.001	-126.930	0.000	-0.143	-0.139
emailer_for_promotion	0.5626	0.021	27.413	0.000	0.522	0.603
homepage_featured	0.2094	0.015	13.971	0.000	0.180	0.239
op_area	0.1086	0.004	25.996	0.000	0.100	0.117
center_type_TYPE_A	1.2088	0.006	189.022	0.000	1.196	1.221
center_type_TYPE_B	1.6047	0.010	162.661	0.000	1.585	1.624
center_type_TYPE_C	0.8709	0.009	98.682	0.000	0.854	0.888

R-square:0.822

Beverage - 1885

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.9922	0.018	219.539	0.000	3.957	4.028
unit_checkoutprice	-0.6497	0.009	-73.492	0.000	-0.667	-0.632
emailer_for_promotion	0.1757	0.027	6.487	0.000	0.123	0.229
homepage_featured	0.3940	0.012	34.056	0.000	0.371	0.417
op_area	0.2628	0.006	47.546	0.000	0.252	0.274
center_type_TYPE_A	1.4580	0.009	170.611	0.000	1.441	1.475
center_type_TYPE_B	1.4135	0.013	107.596	0.000	1.388	1.439
center_type_TYPE_C	1.1206	0.009	118.112	0.000	1.102	1.139

R-square: 0.662

Revenue Optimization

Assuming that we apply model

1) With only Center Type A, which has an average operation area of 4.08
2) No promotion notice
3) No homepage feature

The number of orders would be
Beverage-1885: Log(D) = 3.99 – 0.65*unit_price + 0.26*4.08 + 1.46*1
Soup-1438: Log(D) = 3.50 – 0.16*unit_price + 0.06*4.08 + 1.17*1
Desert-1543: Log(D) = 3.42 – 0.07*unit_price + 0.06*4.08 + 1.07*1

Meal id	Optimal Price	demand	Optimal Revenue	Current revenue	Percentage of change
1885	1.54	247.35	380.54	204.83	+85.78%
1438	7.09	75.75	536.83	412.96	+29.99%
1543	14.68	42.56	624.83	778.64	-19.75%

Model Evaluation

Beverage – 1885
Interaction Effect

Dep. Variable:	log(order)	R-squared:	0.66
Model:	OLS	Adj. R-squared:	0.66
Method:	Least Squares	F-statistic:	3125
Date:	Wed, 29 Apr 2020	Prob (F-statistic):	0.0
Time:	22:59:37	Log-Likelihood:	-8554.
No. Observations:	11092	AIC:	1.713e+0
Df Residuals:	11084	BIC:	1.718e+0
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025
const	3.9942	0.018	220.089	0.000	3.959
unit_checkoutprice	-0.6473	0.009	-73.312	0.000	-0.665
emailer_for_promotion	0.3097	0.033	9.327	0.000	0.245
homepage_featured	0.3912	0.012	33.865	0.000	0.369
op_area	0.2621	0.006	47.519	0.000	0.251
center_type_TYPE_A	1.4577	0.009	170.946	0.000	1.441
center_type_TYPE_B	1.4134	0.013	107.813	0.000	1.388
center_type_TYPE_C	1.1230	0.009	118.538	0.000	1.104
price_pro	-0.7324	0.105	-6.945	0.000	-0.939

Dep. Variable:	log(order)	R-squared:	0.667
Model:	OLS	Adj. R-squared:	0.667
Method:	Least Squares	F-statistic:	3170.
Date:	Wed, 29 Apr 2020	Prob (F-statistic):	0.00
Time:	23:03:03	Log-Likelihood:	-8501.2
No. Observations:	11092	AIC:	1.702e+04
Df Residuals:	11084	BIC:	1.708e+04
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	4.0029	0.018	221.410	0.000	3.967	4.038
unit_checkoutprice	-0.6308	0.009	-70.791	0.000	-0.648	-0.613
emailer_for_promotion	0.1629	0.027	6.053	0.000	0.110	0.216
homepage_featured	0.4837	0.014	35.693	0.000	0.457	0.510
op_area	0.2571	0.006	46.675	0.000	0.246	0.268
center_type_TYPE_A	1.4612	0.008	172.098	0.000	1.445	1.478
center_type_TYPE_B	1.4196	0.013	108.732	0.000	1.394	1.445
center_type_TYPE_C	1.1222	0.009	119.086	0.000	1.104	1.141
price_feat	-0.4453	0.036	-12.483	0.000	-0.515	-0.375

Price_pro = unit_price * Emailer_for_promotion
Price_feat = unit _price * homepage_featured

Model	NO interaction	Add unit_price * emailer_for_promotion	Add unit_price * homepage_featured
Significant at 1% level	/	yes	yes
Adj.R^2	0.662	0.663	0.667
MSE	0.2740	0.2726	0.269

Business Insights

- ❖ Item specific cost-efficient business investments
  - Emailer promotion
  - Homepage feature
  - Center type
- ❖ Future implication due to Covid-19
- ❖ Possible partnership with other operators from the industry

Limitation

- Revenue optimization for the most part
- Unit Checkout Price vs. Unit Base Price
- Center Information, Possible Multicollinearity
- Omitted Variable Bias

Thank You For Your Attention