MESS DATA ANALYSIS & PREDICTIONS

Ayush Agarwal (ME21B1076)
Raghvendra Mishra (ME21B1075)

INDEX

01.

DATA COLLECTION

DATA COLLECTION

Date	Day	Month	Year	Day_Of_Week	day_index	Breakfast	Breakfast_Wastage	Lunch	Lunch_Wastage	Snacks	Snacks_Wastage	Dinner	Dinner_Wastage	Total_Daily_Comsumers	Total_Daily_Wastage
01/08/2023	1	8	2023	Tue	1	241	19.36	522	31.11	291	13.47	456	24.58	1510	88.52
02/08/2023	2	8	2023	Wed	2	428	27.05	623	35.73	319	12.31	674	34.74	2044	109.83
03/08/2023	3	8	2023	Thu	3	289	20.83	747	42.8	559	21.79	632	32.83	2227	118.25
04/08/2023	4	8	2023	Fri	4	269	18.11	585	36.35	458	14.14	550	26.38	1862	94.98
05/08/2023	5	8	2023	Sat	5	392	26.67	427	25.04	272	11.45	816	40	1907	103.16
06/08/2023	6	8	2023	Sun	6	315	19.09	651	40.59	289	13.05	607	28.91	1862	101.64
07/08/2023	7	8	2023	Mon	7	285	20	502	29.74	427	13.75	517	25.39	1731	88.88
08/08/2023	8	8	2023	Tue	8	393	25.81	573	35.11	378	17.07	507	26.57	1851	104.56
09/08/2023	9	8	2023	Wed	9	353	24.51	709	42.2	421	16.41	640	27.82	2123	110.94
10/08/2023	10	8	2023	Thu	10	274	17.91	778	45.76	558	19.68	559	27.68	2169	111.03
11/08/2023	11	8	2023	Fri	11	342	21.61	651	38.81	313	11.59	577	30.94	1883	102.95
12/08/2023	12	8	2023	Sat	12	340	26.35	533	31.55	439	16.94	466	26.18	1778	101.02
13/08/2023	13	8	2023	Sun	13	273	21.93	676	42.76	506	17.85	764	37.67	2219	120.21
14/08/2023	14	8	2023	Mon	0	455	27.96	500	29.54	560	18.04	660	34.65	2175	110.19
15/08/2023	15	8	2023	Tue	1	341	27.39	664	39.56	259	11.99	495	26.68	1759	105.62
16/08/2023	16	8	2023	Wed	2	398	25.15	563	32.29	292	11.27	640	32.99	1893	101.7
17/08/2023	17	8	2023	Thu	3	354	25.51	594	34.04	629	24.52	487	25.3	2064	109.37
18/08/2023	18	8	2023	Fri	4	263	17.71	603	37.47	422	13.03	671	32.18	1959	100.39
19/08/2023	19	8	2023	Sat	5	297	20.21	444	26.03	364	15.32	559	27.4	1664	88.96
20/08/2023	20	8	2023	Sun	6	300	18.18	683	42.59	346	15.63	630	30	1959	106.4
21/08/2023	21	8	2023	Mon	7	360	25.26	789	46.75	552	17.78	532	26.13	2233	115.92
22/08/2023	22	8	2023	Tue	8	399	26.21	503	30.82	400	18.07	459	24.05	1761	99.15
23/08/2023	23	8	2023	Wed	9	251	17.43	622	37.02	294	11.46	608	26.43	1775	92.34
24/08/2023	24	8	2023	Thu	10	358	23.4	619	36.41	665	23.46	727	35.99	2369	119.26
25/08/2023	25	8	2023	Fri	11	359	22.69	553	32.97	488	18.07	663	35.55	2063	109.28
26/08/2023	26	8	2023	Sat	12	263	20.39	410	24.27	432	16.67	459	25.79	1564	87.12
27/08/2023	27	8	2023	Sun	13	292	23.45	726	45.92	522	18.41	776	38.27	2316	126.05
28/08/2023	28	8	2023	Mon	0	314	19.29	531	31.37	532	17.13	717	37.64	2094	105.43
29/08/2023	29	8	2023	Tue	1	283	22.73	691	41.18	254	11.76	396	21.35	1624	97.02
30/08/2023	30	8	2023	Wed	2	343	21.67	782	44.85	450	17.36	502	25.88	2077	109.76
31/08/2023	31	8	2023	Thu	3	340	24.51	807	46.24	481	18.75	605	31.43	2233	120.93
01/09/2023	1	9	2023	Fri	4	333	22.43	477	29.64	598	18.46	652	31.27	2060	101.8

DATA COLLECTION

92x15
Mess dataset dimensions

CLEAN DATA
Data didn't have impurities

Date	Day	Month	Year	Day_Of_Week	day_index	Breakfast	Breakfast_Wastage	Lunch	Lunch_Wastage	Snacks	Snacks_Wastage	Dinner	Dinner_Wastage	Total_Daily_Comsumers	Total_Daily_Wastage
01/08/2023	1	8	2023	Tue	1	241	19.36	522	31.11	291	13.47	456	24.58	1510	88.52
02/08/2023	2	8	2023	Wed	2	428	27.05	623	35.73	319	12.31	674	34.74	2044	109.83
03/08/2023	3	8	2023	Thu	3	289	20.83	747	42.8	559	21.79	632	32.83	2227	118.25
04/08/2023	4	8	2023	Fri	4	269	18.11	585	36.35	458	14.14	550	26.38	1862	94.98
05/08/2023	5	8	2023	Sat	5	392	26.67	427	25.04	272	11.45	816	40	1907	103.16
06/08/2023	6	8	2023	Sun	6	315	19.09	651	40.59	289	13.05	607	28.91	1862	101.64
07/08/2023	7	8	2023	Mon	7	285	20	502	29.74	427	13.75	517	25.39	1731	88.88
08/08/2023	8	8	2023	Tue	8	393	25.81	573	35.11	378	17.07	507	26.57	1851	104.56
09/08/2023	9	8	2023	Wed	9	353	24.51	709	42.2	421	16.41	640	27.82	2123	110.94
10/08/2023	10	8	2023	Thu	10	274	17.91	778	45.76	558	19.68	559	27.68	2169	111.03
11/08/2023	11	8	2023	Fri	11	342	21.61	651	38.81	313	11.59	577	30.94	1883	102.95
12/08/2023	12	8	2023	Sat	12	340	26.35	533	31.55	439	16.94	466	26.18	1778	101.02
13/08/2023	13	8	2023	Sun	13	273	21.93	676	42.76	506	17.85	764	37.67	2219	120.21
14/08/2023	14	8	2023	Mon	0	455	27.96	500	29.54	560	18.04	660	34.65	2175	110.19
15/08/2023	15	8	2023	Tue	1	341	27.39	664	39.56	259	11.99	495	26.68	1759	105.62
16/08/2023	16	8	2023	Wed	2	398	25.15	563	32.29	292	11.27	640	32.99	1893	101.7
17/08/2023	17	8	2023	Thu	3	354	25.51	594	34.04	629	24.52	487	25.3	2064	109.37
18/08/2023	18	8	2023	Fri	4	263	17.71	603	37.47	422	13.03	671	32.18	1959	100.39
19/08/2023	19	8	2023	Sat	5	297	20.21	444	26.03	364	15.32	559	27.4	1664	88.96
20/08/2023	20	8	2023	Sun	6	300	18.18	683	42.59	346	15.63	630	30	1959	106.4
21/08/2023	21	8	2023	Mon	7	360	25.26	789	46.75	552	17.78	532	26.13	2233	115.92
22/08/2023	22	8	2023	Tue	8	399	26.21	503	30.82	400	18.07	459	24.05	1761	99.15
23/08/2023	23	8	2023	Wed	9	251	17.43	622	37.02	294	11.46	608	26.43	1775	92.34
24/08/2023	24	8	2023	Thu	10	358	23.4	619	36.41	665	23.46	727	35.99	2369	119.26
25/08/2023	25	8	2023	Fri	11	359	22.69	553	32.97	488	18.07	663	35.55	2063	109.28
26/08/2023	26	8	2023	Sat	12	263	20.39	410	24.27	432	16.67	459	25.79	1564	87.12
27/08/2023	27	8	2023	Sun	13	292	23.45	726	45.92	522	18.41	776	38.27	2316	126.05
28/08/2023	28	8	2023	Mon	0	314	19.29	531	31.37	532	17.13	717	37.64	2094	105.43
29/08/2023	29	8	2023	Tue	1	283	22.73	691	41.18	254	11.76	396	21.35	1624	97.02
30/08/2023	30	8	2023	Wed	2	343	21.67	782	44.85	450	17.36	502	25.88	2077	109.76
31/08/2023	31	8	2023	Thu	3	340	24.51	807	46.24	481	18.75	605	31.43	2233	120.93
01/09/2023	1	9	2023	Fri	4	333	22.43	477	29.64	598	18.46	652	31.27	2060	101.8

02.

EDA & PRE-PROCESSING

```
LunchMenu
Poori DumAloo BeansCarrotPoriyal Rice Vathakolambu                          7
Phulka Panchratandal OnionPakoda PeruguPachadi Rice BrinjalMochaiGravy      7
Phulka RajmaDal AlooGobiMatarDry Rice Sambhar                               7
Phulka PumpkinKaalaChanna Rice BeetrootChanaPoriyal                         7
Phulka CornPeasMasala SproutedDal Rice                                      7
Phulka AlooGobiMatar PaneerBiryani                                          7
Phulka PeasMasala AlooMasalaWedges Rice Sambhar                             7
Phulka VegKurma BeetrootChannaPoriyal Rice Vathakolambu                     7
Phulka MixedDal BrinjalPoriyal MixVegKarakozhambu                           6
Poori PunjabiAlooMatar KovakaiFry Rice AndhraTomatoDal                      6
Phulka MixVegMatar LaukiChanadal Gobi65 Rice                                6
Phulka MixVegCurry ChilliSoyaBeanDry PeruguPachadi Rice Sambhar             6
Phulka PaneerButterMasala VegBiryani OnionRaitha                            6
Phulka DalMakhani BhindiFry Rice Sambhar                                    6
Name: count, dtype: int64
```

```
[51]
# Checking Different Meals in the menu

y =[]

x = [data.LunchMenu[x].split() for x in range(15)]
for i in x:
    for j in i:
        y.append(j)

x = [data.BFMenu[x].split() for x in range(15)]
for i in x:
    for j in i:
        y.append(j)

x = [data.SnkMenu[x].split() for x in range(15)]
for i in x:
    for j in i:
        y.append(j)

x = [data.DinMenu[x].split() for x in range(15)]
for i in x:
    for j in i:
        y.append(j)

y = set(y)
len(y)

...  87
```

FIXING MESS MENU ITEMS

Exploring and analyzing how to work with mess menu items as to draw inferences or possibly use for model training.

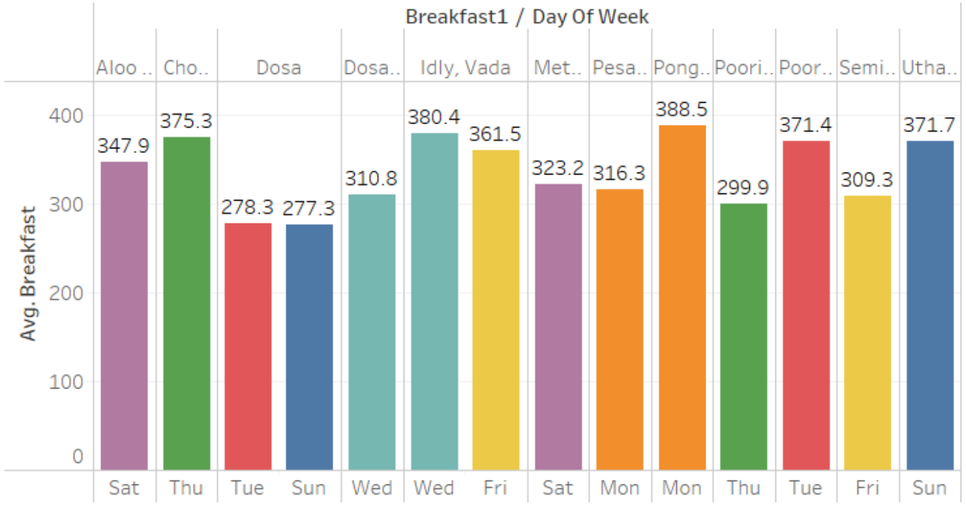
EDA

03.

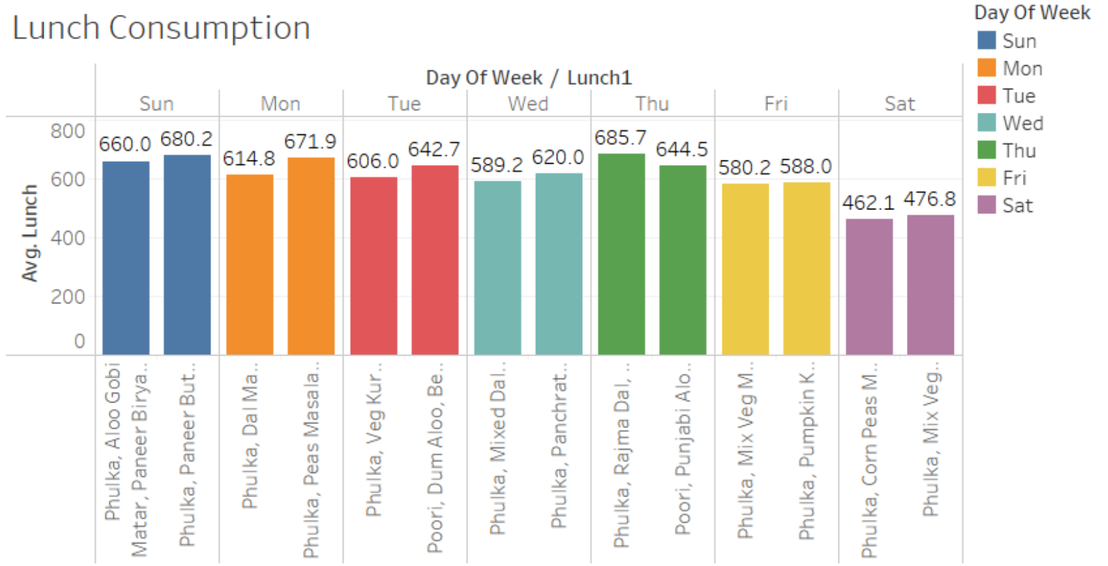
VISUALIZATIONS

VISUALIZATION

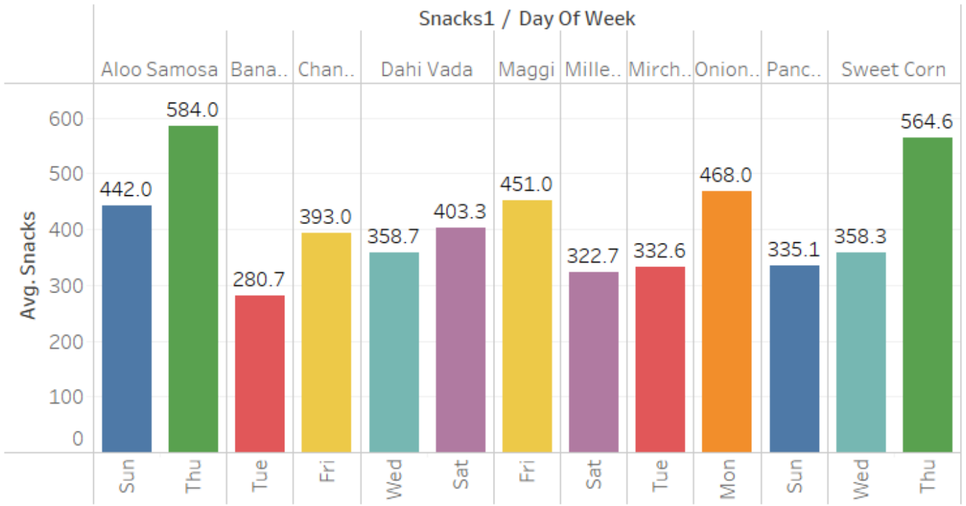
BreakFast Consumption



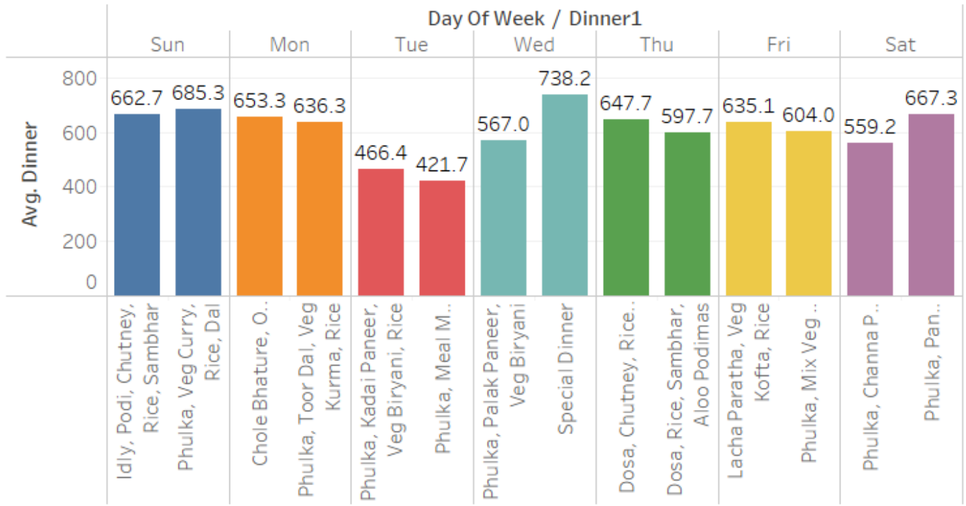
Lunch Consumption



Snacks Consumption



Dinner Consumption

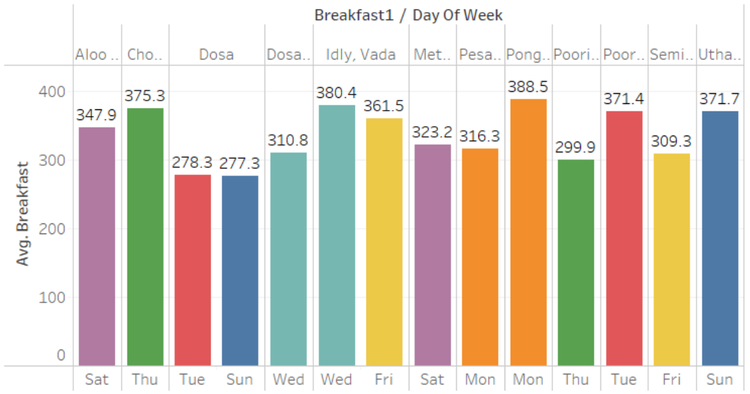


INFERENCES

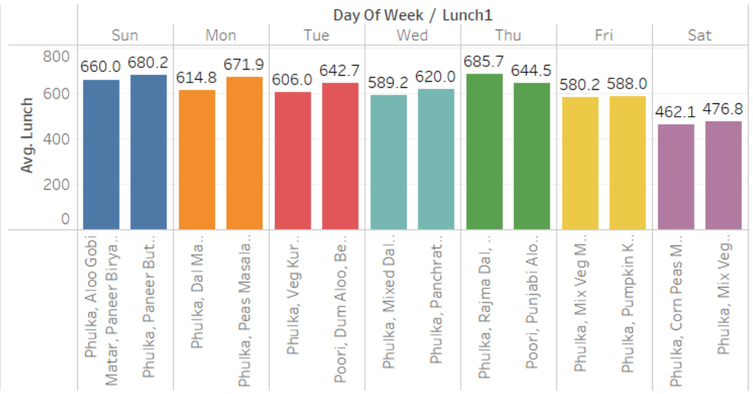
FAVORED FOOD

These graphs inform us about the mess days that the students generally prefer the menu and other days the menu is not to taste. Such as odd-week Tuesday generally sees a decrease in consumption.

BreakFast Consumption

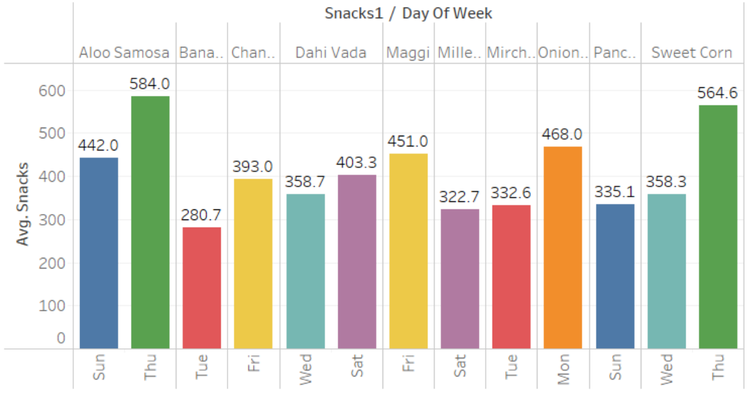


Lunch Consumption

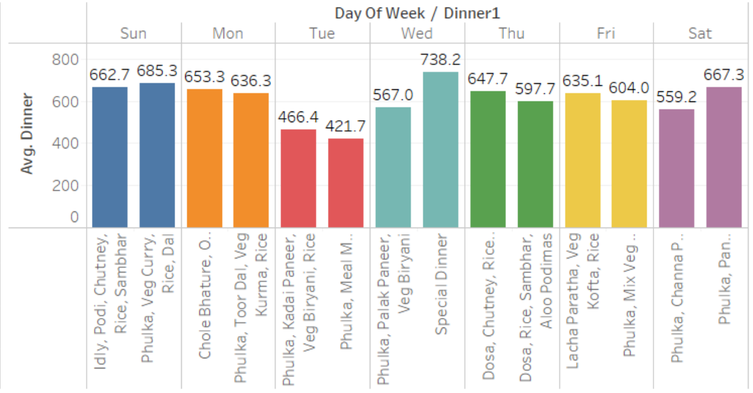


Day Of Week
Sun
Mon
Tue
Wed
Thu
Fri
Sat

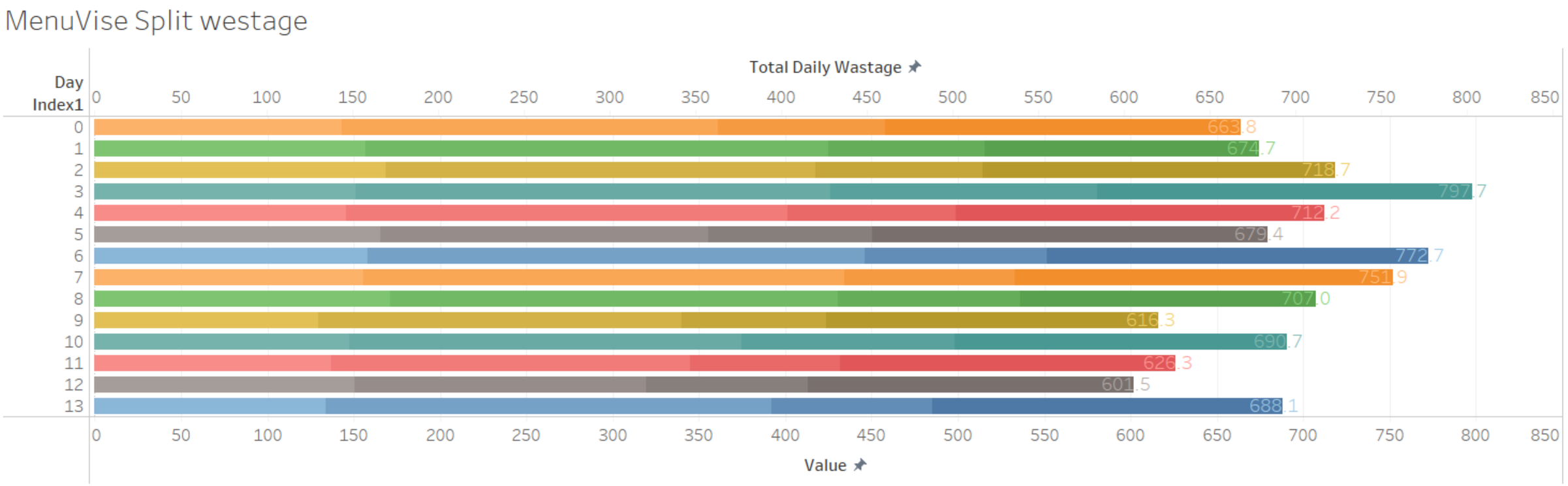
Snacks Consumption



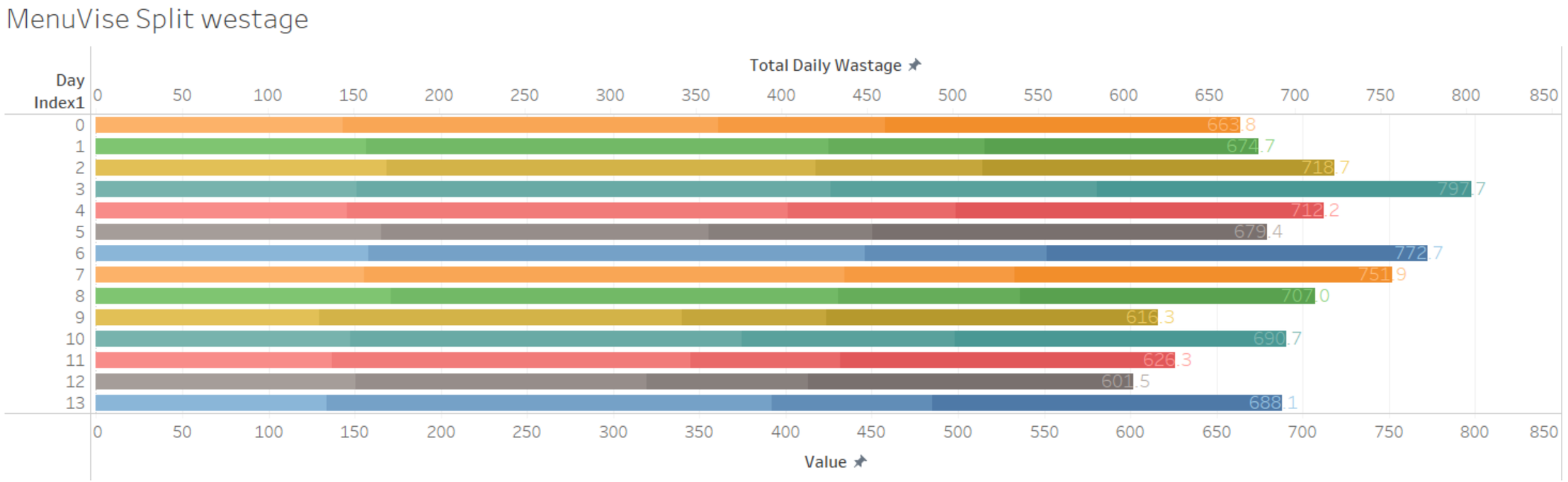
Dinner Consumption



VISUALIZATION



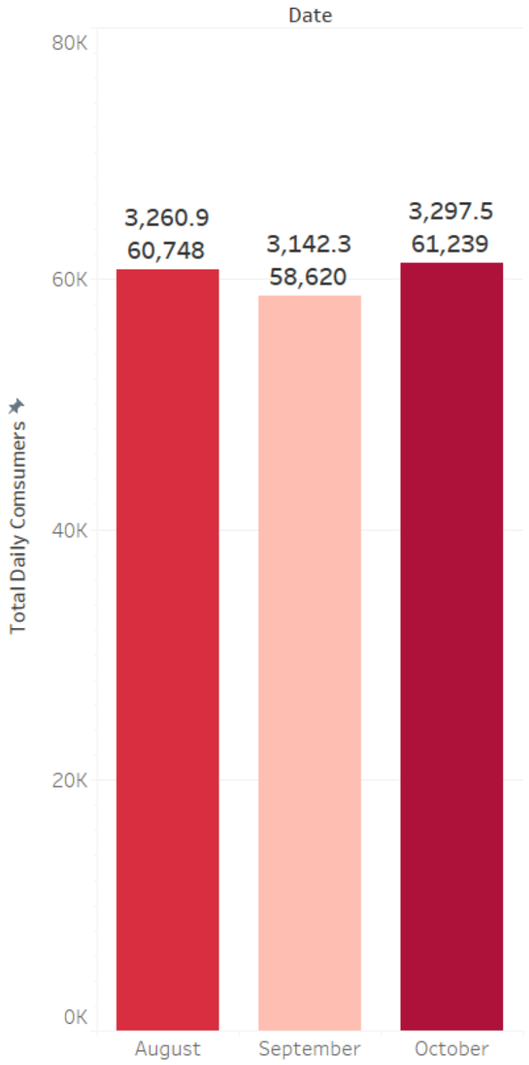
INFERENCES

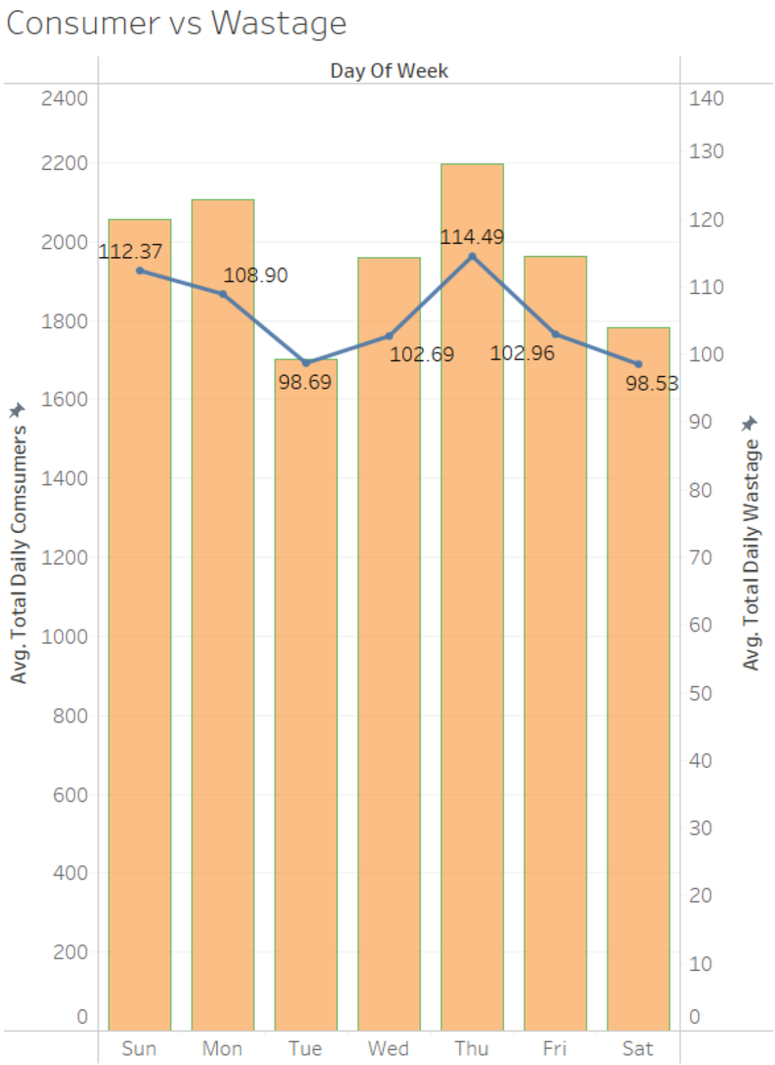


SURPRISE ATTACK

In this we can see that most wastage is seen on Thursday of the odd-week(No. 3) and Sunday (No. 6). Since they don't have the most amount of consumers as seen from previous graphs, this infers that the menu items have a more probable cause of being bad as unforeseen by the consumers.

VISUALIZATION & INFERENCE





04.

MODEL TRAINING/TESTING

MODEL SELECTION

To conclude the best model that can be used for our dataset, we take into consideration quite a few of them:

LINEAR REGRESSION

DECISION TREE REGRESSOR

RANDOM FOREST REGRESSOR

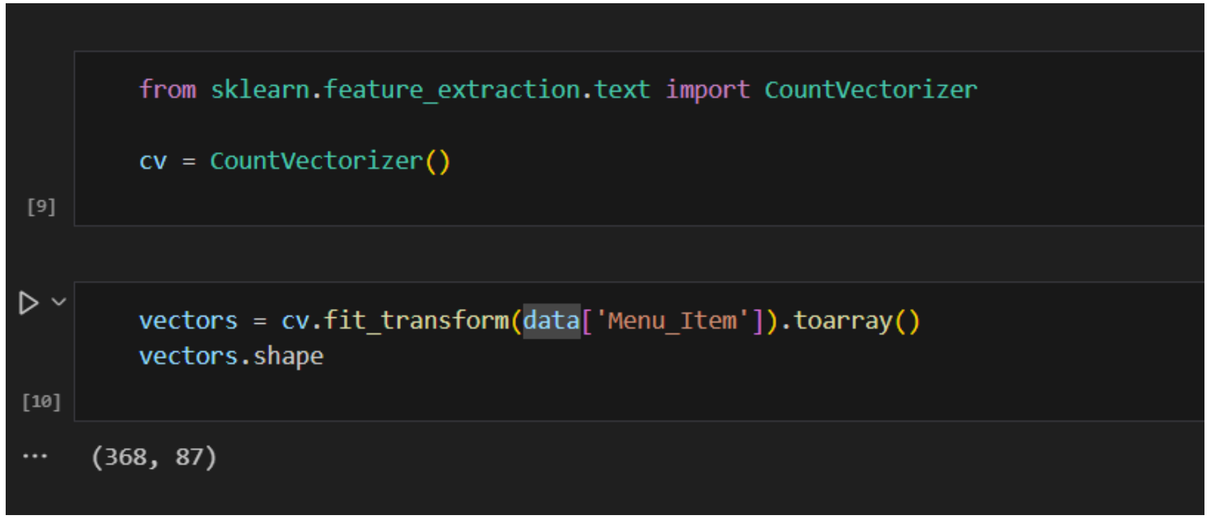
GRADIENT BOOSTING REGRESSOR

SUPPORT VECTOR REGRESSOR

MPL REGRESSOR

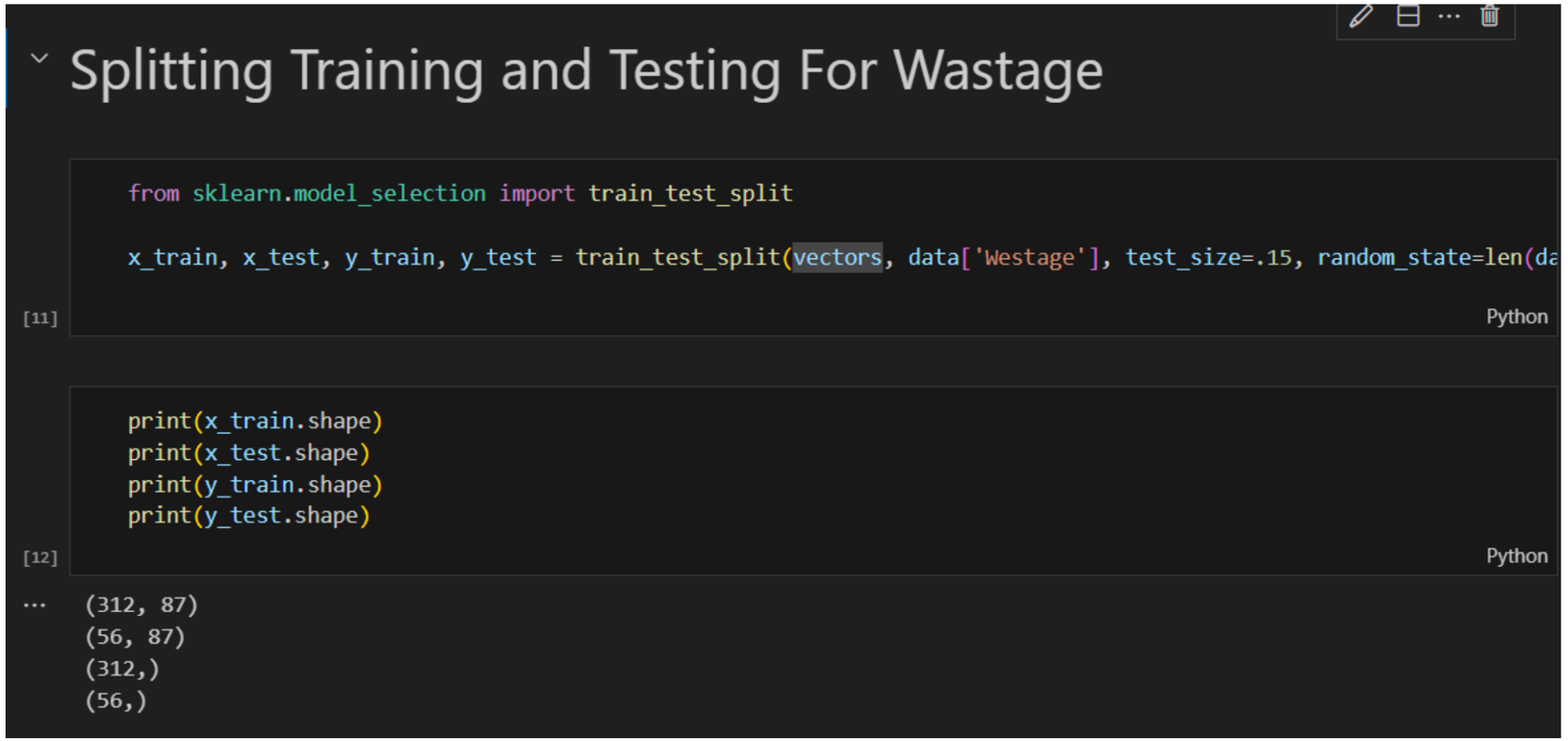
VECTORIZATION

Before train/testing of data we vectorize the menu items as it is the only non-numeric data in the data-set and we need the it to train the data models.



TEST/TRAIN SPLIT

We use 15% of the data as testing data and the rest of it to train the data models



RESULTS

Results

The Results and accuracy scores of the models as summary

Wastage

Model	Train/Test	RMSE	R2 Score
Linear Regression	Train	5.387168726471041	0.8885819065769296
Linear Regression	Test	6.165576225981842	0.8598599395694546
Decison Tree Regressor	Train	3.822037457053275	0.8947229929393822
Decison Tree Regressor	Test	6.399925922950648	0.8671618994634029
Random Forest Regressor	Train	3.8300151552084847	0.8945032493616568
Random Forest Regressor	Test	6.279218965718302	0.8696673164187455
Gradient Boosting Regressor	Train	5.889639200602035	0.837771451830743
Gradient Boosting Regressor	Train	8.77539541398132	0.8178561951677401
Support Vector Regressor	Train	7.273659611899595	0.7996489770383194
Support Vector Regressor	Test	10.488283623342653	0.7823031561322215
MPL Regressor	Train	5.521357362319088	0.847915677292914
MPL Regressor	Train	7.453949223457471	0.8452843879349432

From the above Table we can say that Random Forest Regressor is the best Regression model for our Task of calculating wastage