DATA ANALYST DUO (POWERBI PROJECT WORK) _ COHORT 2
Name: Vishnupriya.K

PROBLEM STATEMENT: Using the provided dateset from a Domino's Pizza outlet in Mumbai, create a comprehensive power bi report to uncover actionable insights to help the outlet increase sales & improve operational efficiency. Your report should identify trends in pizza orders, customer preferences & sales performance & recommended strategies to optimize the menu, pricing & promotional activities.

1. Objective:

To identify the actionable insights to help outlet increase sales, changes in menu, customer buying choices and improve operational efficiency.

2. Data Cleaning and preparation:

The given data set was clean with no blanks/empty rows and can be considered for further analysis.

- The data in excel file need to load into power bi, then check the data type of each columns in a tables and change to respective data type.
- Then transform the data using power query.
- In the given datasets, we have 4 tables contains order details, order_id, pizzas and pizza id with transactions details of 2022 year.
- **All these tables have common columns which gives relationship b/w all these tables.**
- Merge all these columns into one common table using merge query command for better analysis.
- Merge order_details, order id tables & based on pizza_type_id merge pizza name, category & price from pizza's and pizza's id tables.
- Now all required columns are grouped in one table named as **(updated_orders)** for further analysis.

3. Data Analysis:

The Analysis was divided into following categories:

1. Sales highlights:

The following are the dax (Data Analysis Expressions) functions used to calculate sales highlights:

Column/Measure name	Dax formula used
Revenue	Revenue = Updated_orders[Price] * Updated_orders[quantity]
Total orders	Total orders = DISTINCTCOUNT(Updated_orders[order_id])

DATA ANALYST DUO (POWERBI PROJECT WORK) _ COHORT 2
Name: Vishnupriya.K

Total Items ordered	Total Items ordered = COUNT(Updated_orders[order_details_id])
Avg_ordervalue	Avg_ordervalue = SUM(Updated_orders[Revenue])/[Total orders]
Max sales	Max sales = MAX(Updated_orders[Revenue])
# of items	# of items = DISTINCTCOUNT(pizza_types[pizza_type_id])
# of categories	# of categories = DISTINCTCOUNT(pizza_types[Category])

Based on above formulas, the following highlights on dashboard was created.
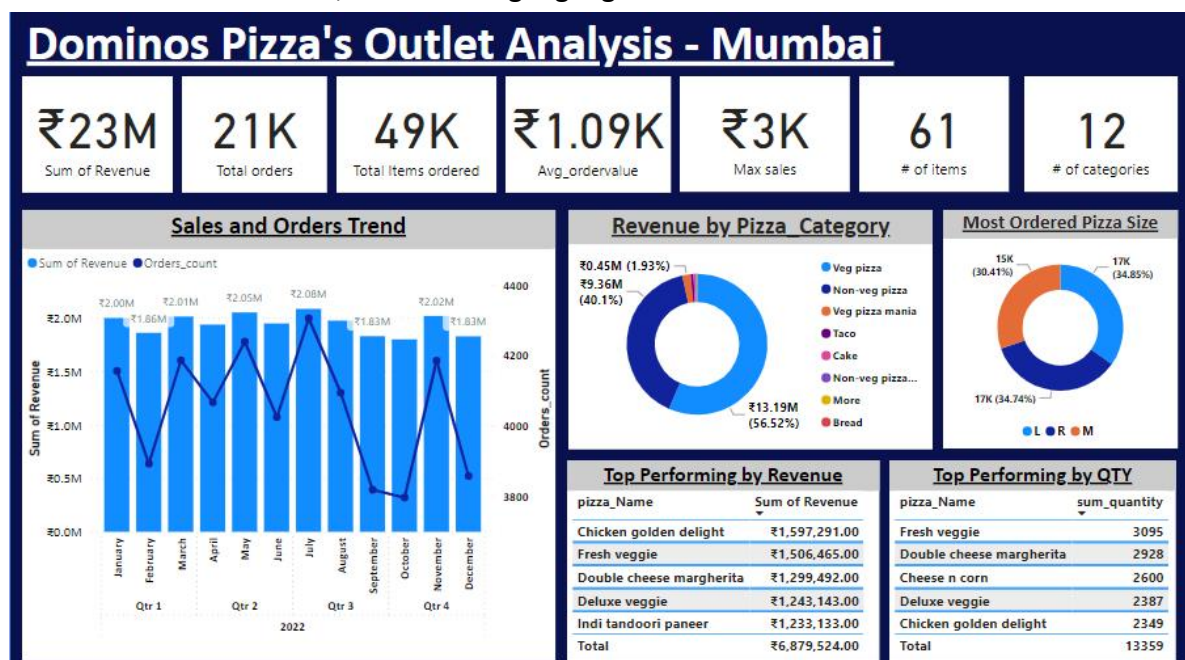
a) **Sales & Orders trend:**

- Total revenue generated from the outlet is 23 millions, if we see the orders count & sales, as orders increases sales are increasing which shows a positive correlation b/w both of them.

- In sales on monthly and quarterly basis, **highest sales & orders was recorded in July of 2.08 M(4303 orders) & least sales was occurred on 1.8 M (3797 orders) in October.**

DATA ANALYST DUO (POWERBI PROJECT WORK) _ COHORT 2
Name: Vishnupriya.K

- In Qtr1, Feb shows much decline in orders & sales and again at Qtr3 in Sept followed to Qtr 4 in Oct, Dec. The reasons can be due to local factors of particular location such as festivals, health issues/holidays/ outlet renovation etc.

b) Revenue by Pizza Category & Size:

- There are total of 12 categories are available out of that only 8 categories have generated revenue and remaining 4 categories **(DIP, Parcel,Pasta veg & pasta non-veg)** doesn't observed any single transaction in that year.

- Out of all categories **veg pizza holds 56%** of sales generated followed by **non_veg pizza (40%), veg pizza mania (2%) and others (< 2%),** which tells that veg & non_veg category pizza's have high demand compared to others.

- Most people preferring to buy **large(~35%) & regular (~35%) size** pizzas than medium (30%) size.

C) **Top Performing & least performing Pizza's in terms of revenue & Qty:**

- Chicken golden delight generates more revenue followed by fresh veggie, double cheese, deluxe veggie, indi tandoori.

- In terms of qty most ordered item was fresh veggie followed by double cheese, cheese n corn, Chicken golden delight.

- If we see, in non veg pizza's, top performed was chicken golden delight was generated more revenue but ordered less, it is possible due to high price of non_veg pizza than veg one's.

pizza_Name	Sum of Revenue
Taco mexicana-veg(single)	₹32,311.00
Onion	₹27,876.00
Crinkle fries	₹27,324.00
Non veg loaded	₹13,761.00
Garlic breadsticks	₹3,052.00
Total	₹104,324.00

pizza_Name	sum_quantity
Red velvet lava cake	289
Crinkle fries	276
Veg loaded	276
Non veg loaded	99
Garlic breadsticks	28
Total	968

- Above picture represents bottom 5 performing pizza's which can summarized using filters. Garlic bread-sticks was seen least ordered followed by non veg pizza mania, crinkle fries in all menu items.

   2. Weekly & Hourly peak time analysis:

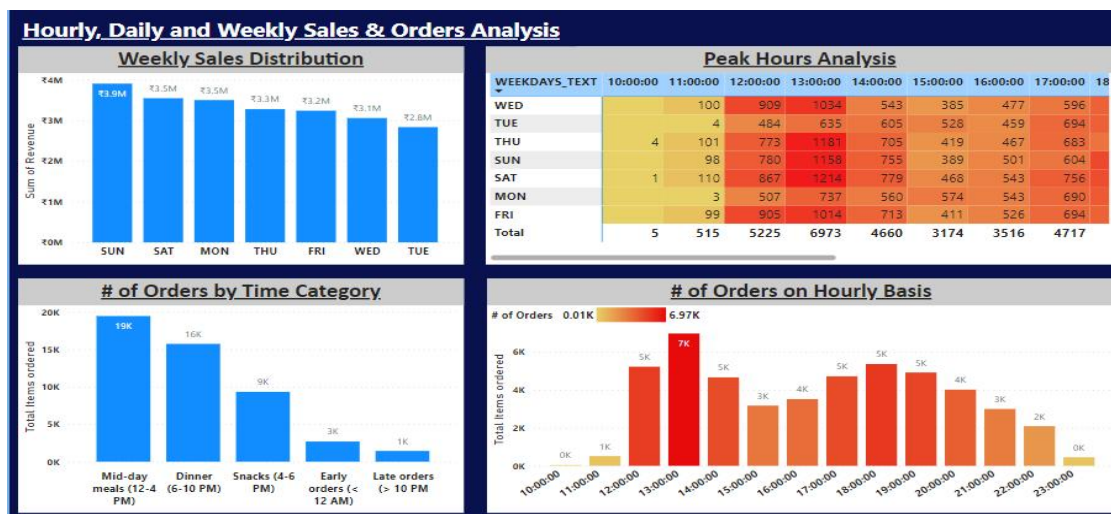# DATA ANALYST DUO (POWERBI PROJECT WORK) _ COHORT 2
Name: Vishnupriya.K

In the given data ordered time & date was available for each order_id. Based on the these information, sales was calculated on weekend & weekdays to identify sales pattern and distribution of time in hourly basis to identify peak hours.

The following are the dax (Data Analysis Expressions) functions used to calculate distribution of time in hourly basis to identify peak hours.

Column/Measure name	Dax formula used
Weekday	Weekday = WEEKDAY(Updated_orders[date].[Date])
Weekday_Text	WEEKDAYS_TEXT = SWITCH(TRUE,Updated_orders[Weekday] = 1, "SUN", Updated_orders[Weekday] = 7, "SAT", Updated_orders[Weekday] = 2, "MON", Updated_orders[Weekday] = 3, "TUE", Updated_orders[Weekday] = 4, "WED", Updated_orders[Weekday] = 5, "THU", "FRI")
time_category	time_category = SWITCH(TRUE, Updated_orders[time] < TIME(12, 0, 0), "Early orders (< 12 AM)", Updated_orders[time] < TIME(16,0,0), "Mid-day meals (12-4 PM)", Updated_orders[time] < TIME(18,0,0), "Snacks (4-6 PM)", Updated_orders[time] < TIME(22,0,0), "Dinner (6-10 PM)", "Late orders (> 10 PM")
Custom_time (to round off time in hourly basis)	Custom_time=Number.Round(Number.From([time])*(60*24/60))/(60*24/60)

- The following was the dashboard created for weekly & hourly sales analysis.

DATA ANALYST DUO (POWERBI PROJECT WORK) _ COHORT 2
Name: Vishnupriya.K

- From the dashboard, we can see that Sunday, Saturday has highest sales followed by Monday and Tuesday, Wednesday have least sales.

- Overall orders are high at lunch (especially 1 pm) & dinner times (especially 6 pm).

Peak Hours Analysis

WEEKDAYS_TEXT	10:00:00	11:00:00	12:00:00	13:00:00	14:00:00	15:00:00	16:00:00	17:00:00	18:00:00	19:00:00	20:00:00	21:00:00	22:00:00	23:00:00	Total
WED		100	909	1034	543	385	477	596	730	593	476	288	214	28	6373
TUE		4	484	635	605	528	459	694	743	639	483	371	248	24	5917
THU	4	101	773	1181	705	419	467	683	619	665	536	342	238	21	6754
SUN		98	780	1158	755	389	501	604	897	878	702	653	527	164	8106
SAT	1	110	867	1214	779	468	543	756	865	673	522	318	177	30	7323
MON		3	507	737	560	574	543	690	785	810	839	691	452	164	7355
FRI		99	905	1014	713	411	526	694	726	661	456	334	234	23	6796
Total	5	515	5225	6973	4660	3174	3516	4717	5365	4919	4014	2997	2090	454	48624

- Above is the Peak hours analysis map, which helps in identifying the most busiest hours in a day & week.

- In a day, most orders are happening at 12 to 1 pm, followed by 5 to 7 pm which are lunch & dinner timings. In terms of week, expect Tuesday & Monday remaining all days have highest orders floating at lunch time (12 - 1 pm). For dinner time (5-7 pm) almost all weeks have high orders floating.

- This gives us clear picture of managing the staff based on peak hours & helps to give week off to employees on less order days like Tuesday & Monday and more staff & preparation is needed at lunch & dinner timings also help in early preparation of inventory levels according to orders.

   3. **Best pizza combinations & menu optimization using Market basket Analysis :**

Market Basket Analysis:

The Basket Analysis tells us the products bought together by consumers. In example of grocery store, this analysis might show that whipped cream and berries are frequently bought together. Based on this result, you can:
– place the products close to each other (if in a physical store) to encourage cross-selling.
– suggest the second product when the consumer places the first product in the cart (if in an online store).
– sell the products together.
– offer discounts for buying the products together.

DATA ANALYST DUO (POWERBI PROJECT WORK) _ COHORT 2
Name: Vishnupriya.K

This analysis is not limited to products. The basket analysis is also useful to explore the visited pages of a website, the questions raised in support tickets, and even the dishes ordered in a restaurant.

The Basket Analysis Network and a Basket Analysis Chart are the two things will be build in Basket analysis.

Three key concepts are necessary to understand for basket analysis in Power BI: Support, Confidence, and Lift.

For example, I have a data with 10 transactions of a grocery store, among that need to find the whipped cream and berries are frequently bought together or not by using 3 factors.

1. ***Support:*** It indicates the frequency of products being bought together.

Support of basket = (No of transactions including both products/Total no of transactions)

EX: Suppose we have No of transactions including both products = 3 & Total no of transactions = 10 then support for basket = (3/10) * 100 = 30%, it mean 30% of transactions have berry & whipped cream together.

2. *Confidence:* It indicates the % of transactions that contains the two products together, out of transactions containing one of the two products.

Confidence is calculated separately for each product in basket.

Confidence of product one = (Support of basket/support of product one)

Confidence of product two = (Support of basket/support of product two)

Confidence indicates the direction of cross selling but it doesn't provide the strength of relation b/w 2 products.

3. *Lift:* It indicates the strength of relationship for a basket of two products.

Lift = support of basket /(support of product one* support of product two)

- Lift ~ 1 : No important relationship

- Lift > 1 : Two products bought together frequently

- Lift < 1: Two products bought together less frequently

DATA ANALYST DUO (POWERBI PROJECT WORK) _ COHORT 2
Name: Vishnupriya.K

Based on above three factors, the best both together combos are identified for our pizza sales data. For this, the combinations need to make from the each of the given items in data set with remaining all items.

Ex: The data set has 61 distinct items each item has combination with remaining 60 items. Through this the combinations will be huge, to reduce the complexity, I have used the top 6 product combinations with remaining 56 items by using cross join which gives me total 78 distinct combinations (by eliminating self combinations & repeated one's).

Following are the dax used to calculate the above formulas.

Column name	Dax formula used	Application
MBA (New column in a table)	MBA = FILTER(CROSSJOIN(SELECTCOLUMNS(VALUES('Top list'[Top selling items]), "product1", 'Top list'[Top selling items]), SELECTCOLUMNS(VALUES(Updated_orders[pizza_type_id]), "product2", Updated_orders[pizza_type_id])), [product1] > [product2])	If we divide the dax into 2 parts then: 1. Cross join - helps to make combinations b/w top 6 selling items with remaining all 61 items. 2. Filter helps to eliminating self combinations & repeated one's to get distinct combinations.
Basket	Basket = 'MBA'[product1]& " - "&'MBA'[product2]	Helps to identify the combinations
Support Basket	Support Basket = VAR prod1 = MBA[product1] VAR prod2 = 'MBA'[product2] var prod1Transactions = SELECTCOLUMNS(filter(Updated_orders, Updated_orders[pizza_type_id] = prod1), "Transaction ID", Updated_orders[order_id])	Based on support formula, If we divide the dax into 2 parts then: 1. Identifying no of transactions which includes individual transactions 2. By doing intersection, identifying common transactions.

|		

	``
	var prod2Transactions =
	    SELECTCOLUMNS(
	            filter(Updated_orders, Updated_orders[pizza_type_id] = prod2),
	        "Transaction ID", Updated_orders[order_id])
	
	var bothprodtranscations = INTERSECT(prod1Transactions,prod2Transactions)
	RETURN
	COUNTROWS(bothprodtranscations)/[Total orders]
	``	

Confidence of prod1	``` Confidence of prod1 = var prod1 = MBA[product1] var supportprod1 = COUNTROWS(FILTER(Updated_orders,Updated_orders[pizza_type_id] = prod1)) / [Total orders] RETURN 'MBA'[Support Basket] / supportprod1 ```	Calculating individual support & then dividing with support basket.
Confidence of prod2	``` Confidence of prod2 = var prod2 = MBA[product2] var supportprod2 = COUNTROWS(FILTER(Updated_orders,Updated_orders[pizza_type_id] = prod2)) / [Total orders] RETURN 'MBA'[Support Basket] / supportprod2 ```	Calculating individual support & then dividing with support basket.
Lift	``` lift = var prod1 = 'MBA'[product1] var prod2 = MBA[product2] var prod1support = COUNTROWS(FILTER(Updated_orders,Updated_orders[pizza_type_id] = prod1)) / [Total orders] var prod2support = ```	Calculating individual support & then applying in lift formula.

| | COUNTROWS(FILTER(Updated_orders,Updated_orders[pizza_type_id] = prod2)) / [Total orders]

RETURN

MBA[Support Basket] / (prod1support * prod2support) | |

After Executing of all these dax formulae, Market basket table will be created with all above columns. Visualize the data based on following graphs.

1. **Basket network:** Add ZOOMCHARTS NETWORK CHARTS into visualizations bar through download by google or by add more visuals in visualizations bar which connects directly to chrome.
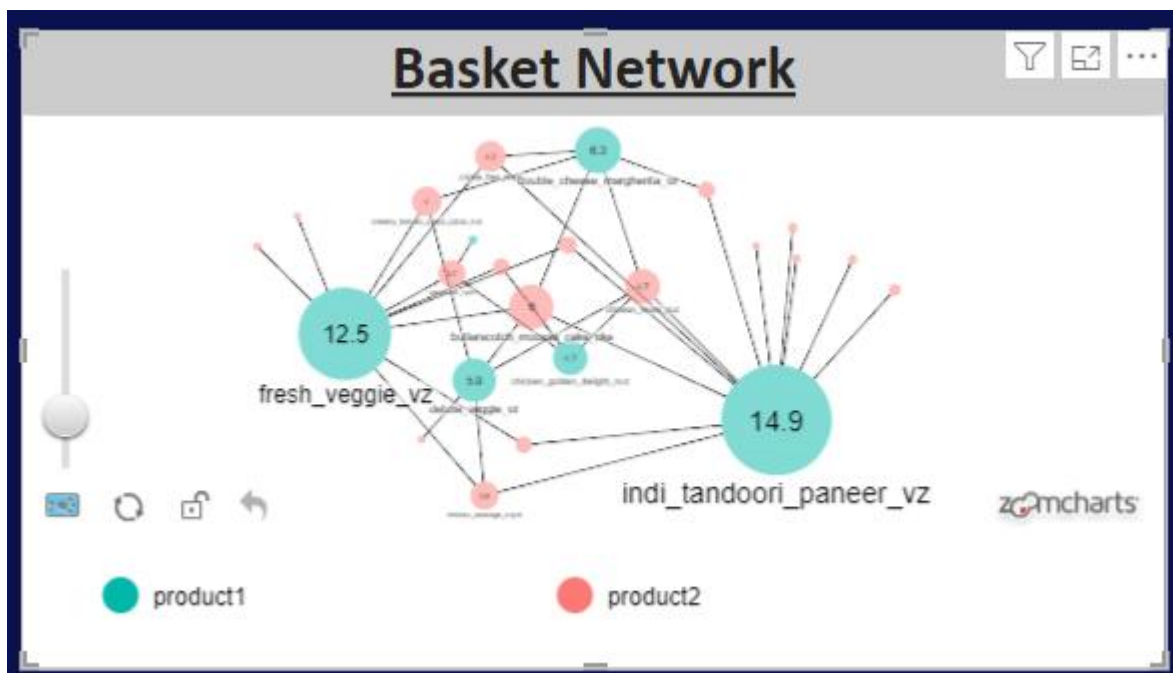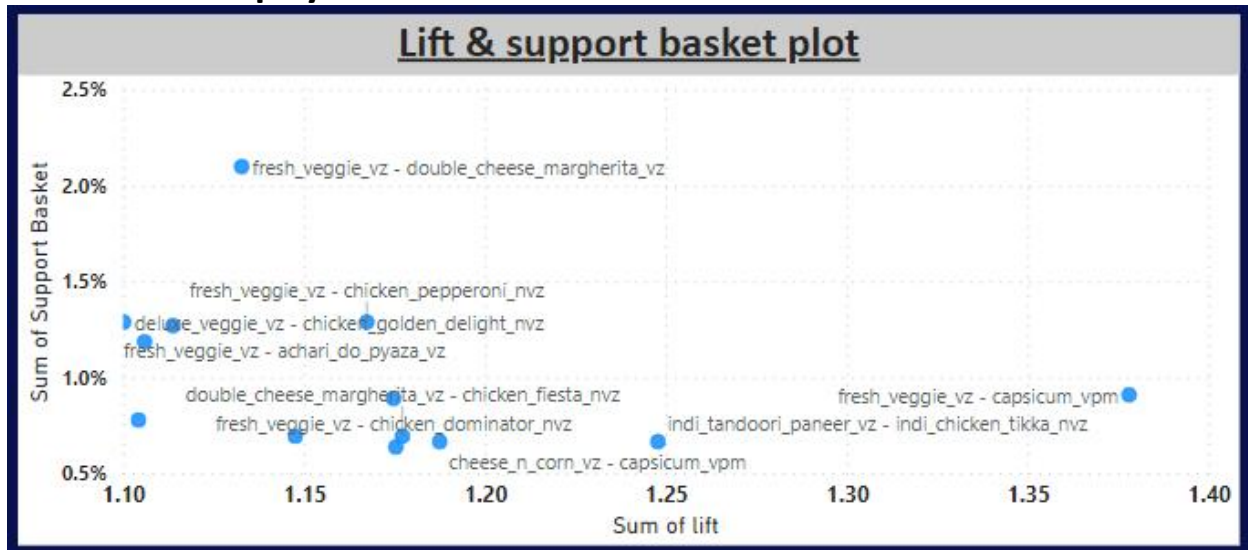

The above is the basket network map which can be create using ZOOMCHARTS NETWORK CHARTS visual by adding columns & lift value in fields.

2. Scatter plot: Explains the best combinations based on support & lift values.
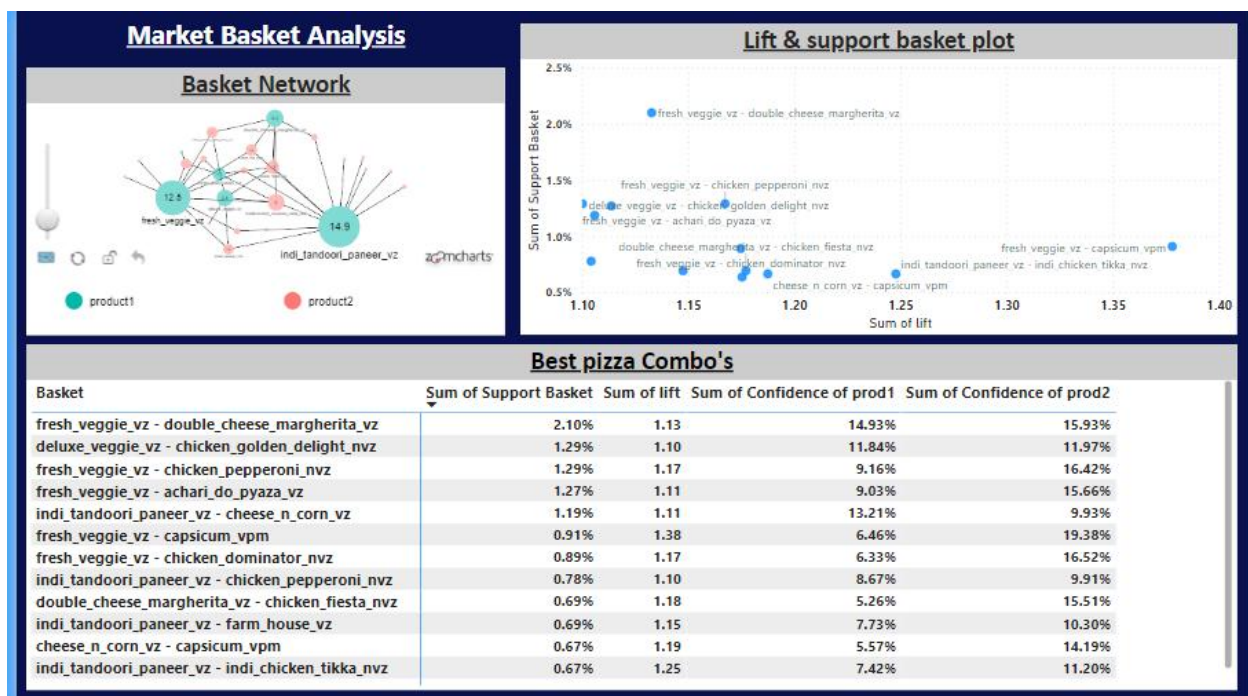
3. Best combo's Matrix:

The pizza names along with lift, support & confidence values are visualized, & overall analysis will look like below:

Basket	Sum of Support Basket	Sum of lift	Sum of Confidence of prod1	Sum of Confidence of prod2
fresh_veggie_vz - double_cheese_margherita_vz	2.10%	1.13	14.93%	15.93%
deluxe_veggie_vz - chicken_golden_delight_nvz	1.29%	1.10	11.84%	11.97%
fresh_veggie_vz - chicken_pepperoni_nvz	1.29%	1.17	9.16%	16.42%
fresh_veggie_vz - achari_do_pyaza_vz	1.27%	1.11	9.03%	15.66%
indi_tandoori_paneer_vz - cheese_n_corn_vz	1.19%	1.11	13.21%	9.93%
fresh_veggie_vz - capsicum_vpm	0.91%	1.38	6.46%	19.38%
fresh_veggie_vz - chicken_dominator_nvz	0.89%	1.17	6.33%	16.52%
indi_tandoori_paneer_vz - chicken_pepperoni_nvz	0.78%	1.10	8.67%	9.91%
double_cheese_margherita_vz - chicken_fiesta_nvz	0.69%	1.18	5.26%	15.51%
indi_tandoori_paneer_vz - farm_house_vz	0.69%	1.15	7.73%	10.30%
cheese_n_corn_vz - capsicum_vpm	0.67%	1.19	5.57%	14.19%
indi_tandoori_paneer_vz - indi_chicken_tikka_nvz	0.67%	1.25	7.42%	11.20%

In all the above visuals, **lift value was filtered as greater than 1.1 & support basket was taken as greater than 0.6 to get best combinations.**

These filters may vary based on values obtained according to business needs.

DATA ANALYST DUO (POWERBI PROJECT WORK) _ COHORT 2
Name: Vishnupriya.K

Total 13 combinations were identified which are frequently bought & have strong relationship b/w two products.

4. <u>Insights & Recommendations:</u>

- The **sales was high in June** & **orders count was reduced more in Sept & Oct** , need to find what are the potential reasons for less orders based on the geographic location details of the outlet, any festival seasons or any health issues occur (like covid cases) etc.

- Among all products, **veg & non-veg pizza's have high demand with overall 96.6% share,** so this out let is good for pizza's rather than other product categories.

- In terms of size, Large & medium holds ~35% share each & 30% by regular pizza's. it tells us people prefer the pizza as their complete lunch or dinner rather than snacks.

- In terms of revenue & qty sold, chicken_golden delight, fresh veggie, double cheese margarita, indi tandoori, deluxe veggie, cheese n corn are top 6 performers.

- In terms of qty sold, majority are fresh veggie, double cheese margarita, cheese n corn which are less price compared to remaining veg pizza's. Similarly in non veg pizza, chicken golden delight was in medium price compared to other non veg pizza category. it can be reason for people preferring to purchase qty due to cost effective.

- Among 12 categories only 8 categories have generated revenue and 4 categories **(DIP, Parcel,Pasta veg & pasta non-veg)** doesn't observed any single transaction in that year. There are total 9 items under this categories. It means people not preferring to have dips with pizza's it might be due to additional cost, people are more likely to buy pasta pizza rather than original pasta items.

- Needs Reducing the inventory levels of these items and to sale the existing stock, free dips need to provide based on pizza orders count and can cross sell these items with top selling items with discounts may attract the customers.

- In a day, most orders are happening at **12 to 1 pm, followed by 5 to 7 pm** which are lunch & dinner timings. In terms of week, expect Tuesday & Monday remaining all days have highest orders floating at lunch time (12 - 1 pm). For dinner time (5-7 pm) almost all weeks have high orders floating.

- This gives us clear picture of managing the staff based on peak hours & helps to give week off to employees on less order days like Tuesday & Monday and more staff is needed at

lunch & dinner timings also help in early preparation of inventory levels according to orders which helps to improve overall operational efficiency.

- In terms of best combinations, **fresh veggie - double cheese margarita were frequently bought together with (support 2.1% & lift of 1.3)** followed by delux_veggie - chicken_golden delight, fresh veggie - chicken pepperoni etc. Total 13 combinations were identified which are frequently bought & have strong relationship b/w two products (picture attached above).

- With these, new combo's can be added in menu card with best combinations combo pack's to choose by customers (" TRY OUR BEST COMBO PIZZA'S") with offering discounts will improve sales efficiency.