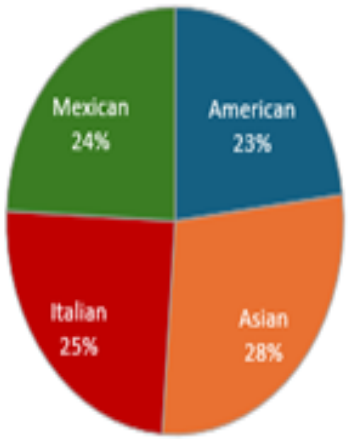
RESTAURANT ORDER ANALYSIS

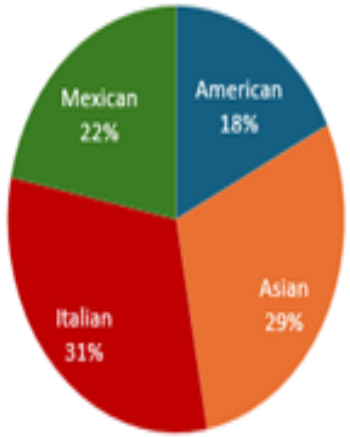
Total Orders: **5,343**

Total Revenue: **$159,218**

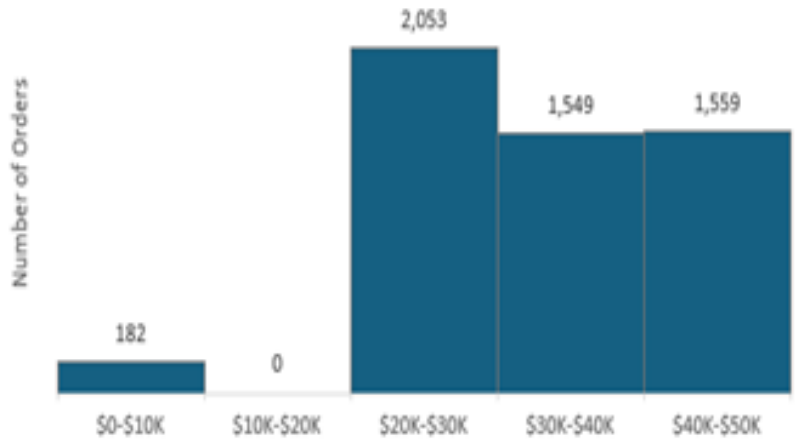
Order Breakdown



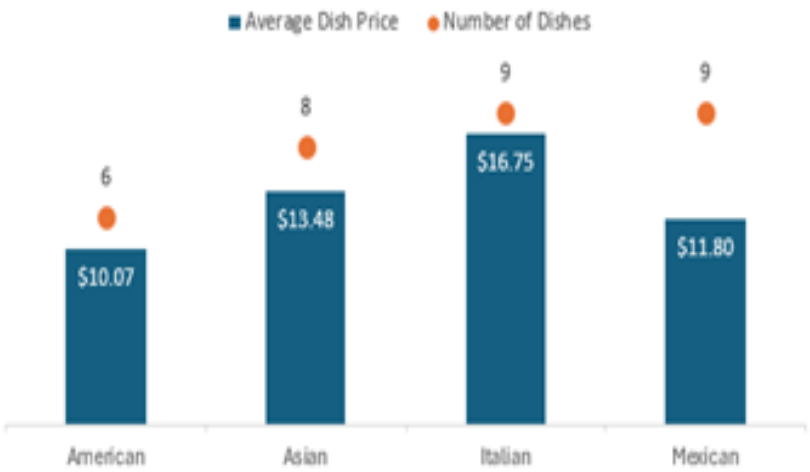
Revenue Breakdown



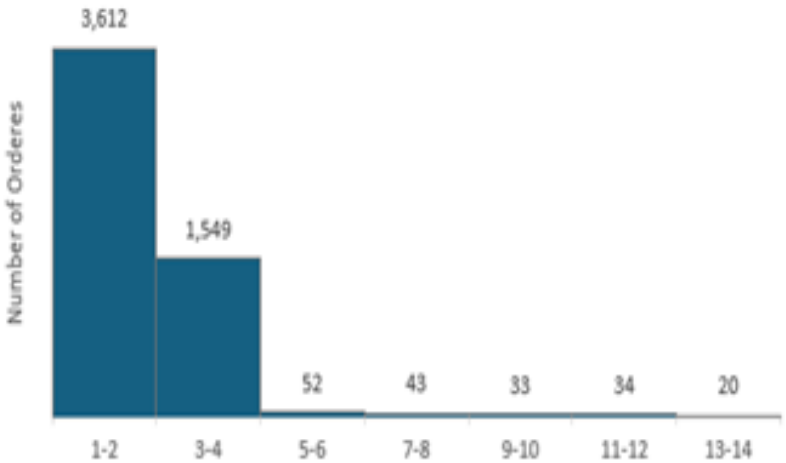
Revenue Distribution



Dish Statistics



Dish Distribution



Least Expensive Dish:	**Edamame**
Least Expensive Dish Category:	**Asian**
Least Expensive Dish Price:	**$5.00**

Most Expensive Dish:	**Shrimp Scampi**
Most Expensive Dish Category:	**Italian**
Most Expensive Dish Price:	**$19.95**

Least Ordered Dish:	**Chicken Tacos**
Least Ordered Dish Category:	**Mexican**
Number of Orders:	**123**
Total Revenue:	**$1,470**

Most Ordered Dish:	**Hamburger**
Most Ordered Dish Category:	**American**
Number of Orders:	**622**
Total Revenue:	**$8,055**

Objective

- I am a newly hired data analyst for Taste of the World Cafe.
- The restaurant has diverse menu offerings and serves generous portions.
- The restaurant debuted a new menu at the start of the year.
- The restaurant wants to investigate customer data to determine which dishes are doing well and which dishes customers like best.

Approach

- I performed a product sales analysis of the customer data.
- My analysis aims to answer the following questions:
  - What was the share of the number of orders and total revenue for each dish category?
  - What was the average dish price and number of dishes for each dish category?
  - What was the least and most expensive dish on the menu?
  - What was the least and most frequently ordered dish on the menu?
  - How was the number of dishes and total revenue distributed per number of orders?
- I wrote SQL queries to answer the questions for my analysis.
- I created an Excel dashboard that summarizes the answers to my questions.
- I uploaded all the files for this project onto my GitHub.

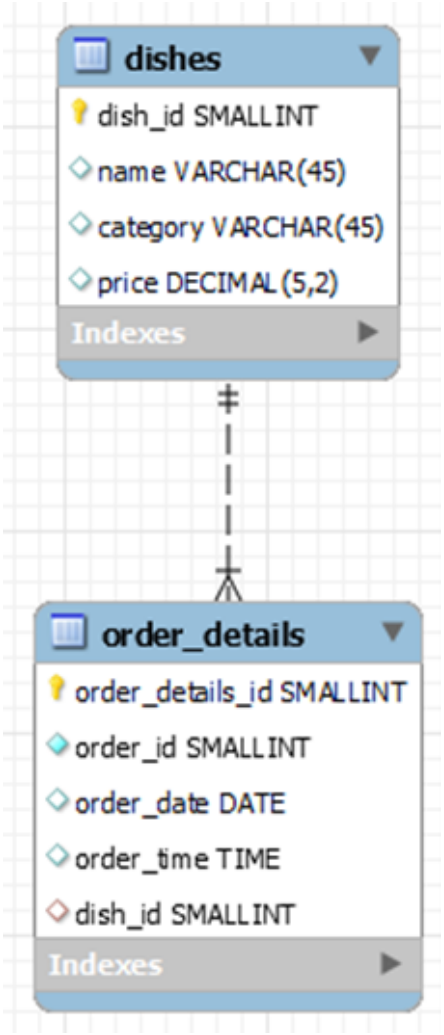
About the Customer Data

- I created a MySQL database and data model.
  - The database consists of the following tables:
    - A “menu_items” dimension table containing the names of all dishes.
    - An “order_details” fact table containing 12,266 records of transaction data.
- The customer data was mostly clean and structured.
  - Therefore, my data cleaning strategy was not very involved.
  - I renamed some of the columns and tables to simplify the data.
    - I renamed the “menu_item” table to “dishes.”
    - I renamed the “menu_item_name” column to “name”
    - I renamed the “menu_item_id” column to “dish_id”.

Assumptions

- There were orders with no purchased dishes. (NULL dish IDs).
  - I assumed that these represented canceled orders or orders not processed properly.
- I decided not to include these orders in my analysis.
  - This is because these orders will not generate revenue for the business.
  - The scope of my analysis mainly focused on revenue maximization and product demand.
    - It did not include any factors affecting why orders were canceled or not processed properly.

Data Model


dishes
dish_id SMALLINT
name VARCHAR(45)
category VARCHAR(45)
price DECIMAL(5,2)
Indexes
order_details
order_details_id SMALLINT
order_id SMALLINT
order_date DATE
order_time TIME
dish_id SMALLINT
Indexes

SQL Queries

- Dish Statistics
  1. number of dishes and average dish price per dish category
  2. least and most expensive dishes
  3. least and most ordered dishes
- Order Statistics
  1. number of orders and total revenue per dish category
  2. number of orders and total revenue for all dishes
  3. order distribution

Dish Statistics

```sql
-- Find the number dishes and average dish price per dish category
SELECT
    category,
    COUNT(*) AS num_dishes,
    ROUND(AVG(price), 2) AS avg_dish_price
FROM dishes
GROUP BY category
ORDER BY category;
```

category	num_dishes	avg_dish_price
American	6	10.07
Asian	8	13.48
Italian	9	16.75
Mexican	9	11.80

Back to SQL Queries

Dish Statistics

```sql
-- Find the least and most expensive dishes
WITH cte AS
(
    SELECT
        name,
        category,
        price,
        RANK() OVER(ORDER BY price DESC) AS desc_rank,
        RANK() OVER(ORDER BY price) AS asc_rank
    FROM dishes
)
SELECT
    name,
    category,
    price
FROM cte
WHERE desc_rank = 1 OR
      asc_rank = 1;
```

name	category	price
Edamame	Asian	5.00
Shrimp Scampi	Italian	19.95

Back to SQL Queries

Dish Statistics

```sql
-- Find the least and most ordered dishes
WITH cte AS
(
    SELECT
        d.name,
        d.category,
        COUNT(*) AS num_orders,
        SUM(d.price) AS total_revenue,
        RANK() OVER(ORDER BY COUNT(*) DESC) AS desc_rank_num,
        RANK() OVER(ORDER BY COUNT(*)) AS asc_rank_num
    FROM order_details o
    LEFT JOIN dishes d ON o.dish_id = d.dish_id
    WHERE d.dish_id IS NOT NULL
    GROUP BY d.dish_id
)
SELECT
    name,
    category,
    num_orders,
    total_revenue
FROM cte
WHERE desc_rank_num = 1 OR
      asc_rank_num = 1;
```

name	category	num_orders	total_revenue
Chicken Tacos	Mexican	123	1469.85
Hamburger	American	622	8054.90

Back to SQL Queries

Order Statistics

```sql
-- Find the number of orders and total revenue per dish category
SELECT
    d.category,
    COUNT(DISTINCT o.order_id) AS num_orders,
    SUM(d.price) AS total_revenue
FROM order_details o
LEFT JOIN dishes d ON d.dish_id = o.dish_id
WHERE d.dish_id IS NOT NULL
GROUP BY d.category
ORDER BY d.category;
```

category	num_orders	total_revenue
American	2152	28237.75
Asian	2635	46720.65
Italian	2292	49462.70
Mexican	2266	34796.80

Back to SQL Queries

Order Statistics

```sql
-- Find the number of orders made and total revenue for all dishes
SELECT
    COUNT(DISTINCT o.order_id) AS num_orders,
    SUM(d.price) AS total_revenue
FROM order_details o
LEFT JOIN dishes d ON o.dish_id = d.dish_id
WHERE d.dish_id IS NOT NULL;
```

num_orders	total_revenue
5343	159217.90

Back to SQL Queries

Order Statistics

```sql
-- Find the distribution of the number of dishes and total revenue per number of orders
WITH cte AS
(
    SELECT
        o.order_id,
        COUNT(*) AS num_dishes,
        SUM(d.price) AS total_revenue
    FROM order_details o
    LEFT JOIN dishes d ON o.dish_id = d.dish_id
    WHERE d.dish_id IS NOT NULL
    GROUP BY o.order_id
)
SELECT
    COUNT(*) AS num_orders,
    num_dishes,
    SUM(total_revenue) AS total_revenue
FROM cte
GROUP BY num_dishes
ORDER BY num_dishes;
```

num_orders	num_dishes	total_revenue
2053	1	27243.05
1559	2	40792.25
847	3	33645.45
702	4	36765.65
32	5	2109.65
20	6	1536.15
21	7	2013.60
22	8	2286.90
17	9	2106.40
16	10	2073.10
14	11	2037.30
20	12	3071.25
13	13	2251.30
7	14	1285.85

Back to SQL Queries

Dish Insights

- Asian and Italian dishes were the most expensive.
- These were the least and most expensive dishes:
  - Edamame (Asian Dish, $5.00)
  - Shrimp Scampi (Italian Dish, $19.95)

Order Insights

- There was a total of 5,343 orders generating $159,218 in revenue.
  - Most orders contained Asian and Italian dishes (28% and 25% respectively).
  - Most revenue came from Asian and Italian dishes also (29% and 31% respectively).
- Most orders generated $20,000 or more in revenue and had between one to four dishes.
- These were the least and most frequently ordered dishes:
  - Chicken Tacos: (123 orders, $1,470 in revenue).
  - Hamburger: (622 orders, $8,055 in revenue).

Conclusion

- Customers seem to like all the dishes on the menu.
  - The total number of orders for each dish category was almost equally distributed.
- Customers are willing to spend a lot of money.
  - The revenue distribution is left-skewed, with orders generating less than $20,000 being insignificant.
- Customers do not prefer to purchase many dishes.
  - The dish distribution is heavily right-skewed, with orders having more than four dishes being insignificant.

Recommendations

- Prioritize marketing campaigns on Asian and Italian dishes.
  - Most of the revenue came from Asian and Italian dishes.
  - Therefore, this strategy will likely increase revenue by attracting more customers to the top revenue-driving dish category.
- Reduce the price of the Chicken Tacos.
  - The demand for this dish is low because it was the least frequently ordered dish.
  - Therefore, this strategy will likely make more customers purchase this dish.
- Upsell on Hamburgers
  - The demand for this dish is high because it was the most frequently ordered dish.
  - Therefore, this strategy will likely make more customers purchase this dish.