PROJECT | PYTHON FOUNDATIONS

DATA-DRIVEN STRATEGIES FOR FOODHUB

SHAISHAV MERCHANT

AGENDA

- EXECUTIVE SUMMARY
- BUSINESS PROBLEM OVERVIEW AND SOLUTION APPROACH
- DATA OVERVIEW
- EDA - UNIVARIATE ANALYSIS
- EDA - MULTIVARIATE ANALYSIS
- CONCLUSION & RECOMMENDATIONS

EXECUTIVE SUMMARY

This comprehensive analysis delves into FoodHub's operations within the competitive New York City food delivery market, examining customer demand patterns, operational efficiency, and satisfaction metrics across 1,898 orders.

EXECUTIVE SUMMARY

- FOODHUB IS A KEY PLAYER IN NYC'S COMPETITIVE ONLINE FOOD DELIVERY MARKET, CONNECTING CUSTOMERS WITH RESTAURANTS.
- THE ANALYSIS COVERS 1,898 ORDERS, FOCUSING ON CUSTOMER DEMAND, OPERATIONAL EFFICIENCY, AND SATISFACTION METRICS.
- AMERICAN CUISINE IS THE MOST POPULAR, WITH 31% OF ORDERS, FOLLOWED BY JAPANESE, ITALIAN, AND CHINESE.
- THE MEDIAN ORDER COST IS $14.14, WITH NO OUTLIERS, INDICATING CONSISTENT PRICING ACROSS ORDERS.
- WEEKENDS ARE BUSIER WITH 71% OF ORDERS, SUGGESTING A SHIFT IN CONSUMER BEHAVIOR DURING THESE DAYS.
- RATINGS ARE MISSING FOR 39% OF ORDERS, BUT AMONG RATED ORDERS, THE MAJORITY ARE RATED '5'.
- FOOD PREPARATION AND DELIVERY TIMES ARE COSISTENT, AVERAGING AROUND 27 AND 24 MINUTES, RESPECTIVELY.
- SHAKE SHACK LEADS IN REVENUE, WITH THE TOP 14 RESTAURANTS CONTRIBUTING TO OVER 54% OF TOTAL REVENUE.
- A SLIGHT NEGATIVE CORRELATION EXISTS BETWEEN FOOD PREPARATION TIME AND RATINGS, INDICATING EFFICIENCY MAY BOOST SATISFACTION.
- NO STRONG CORRELATION IS FOUND BETWEEN ORDER COST, FOOD PREPARATION TIME, AND DELIVERY TIME, SUGGESTING OTHER FACTORS MAY INFLUENCE THESE VARIABLES.

BUSINESS PROBLEM OVERVIEW & SOLUTION APPROACH

This section defines the challenge or opportunity that the business seeks to address through data analysis. It also outlines the methodologies and analytical techniques that will be employed to explore the data, extract insights, and develop data-driven recommendations to solve the identified business problem.

PROBLEM OVERVIEW

PROBLEM DEFINITION:

CONTEXT

THE BURGEONING RESTAURANT LANDSCAPE IN NEW YORK CITY IS NOT ONLY COMPETITIVE BUT ALSO A CRITICAL PART OF THE DAILY LIFE OF ITS INHABITANTS. WITH THE ADVENT OF ONLINE FOOD DELIVERY PLATFORMS, THERE HAS BEEN A SIGNIFICANT SHIFT IN CONSUMER BEHAVIOR. FOODHUB, A LEADING FOOD AGGREGATOR, OFFERS A COMPREHENSIVE SOLUTION, LINKING CUSTOMERS TO RESTAURANTS THROUGH A STREAMLINED DIGITAL INTERFACE.

OBJECTIVE

THE PRIMARY OBJECTIVE IS TO CONDUCT AN EXHAUSTIVE DATA-DRIVEN ANALYSIS TO UNCOVER ACTIONABLE INSIGHTS THAT WILL ENHANCE CUSTOMER EXPERIENCE AND OPERATIONAL EFFICIENCY. LEVERAGING A DATASET THAT ENCAPSULATES VARIOUS ASPECTS OF CUSTOMER ORDERS—RANGING FROM COST, RATINGS, AND FOOD PREPARATION TIME TO DELIVERY TIME—THE ANALYSIS SEEKS TO ADDRESS THE FOLLOWING:

1. IDENTIFY KEY DRIVERS INFLUENCING CUSTOMER DEMAND FOR VARIOUS RESTAURANTS.
2. ASSESS THE EFFICACY AND TIME-EFFICIENCY OF THE FOOD PREPARATION AND DELIVERY PROCESS.
3. EVALUATE CUSTOMER SATISFACTION THROUGH RATINGS AND CORRELATE IT WITH OTHER VARIABLES LIKE COST AND DELIVERY TIME.

SOLUTION APPROACH

SOLUTION APPROACH:

DATA PREPROCESSING:

- UTILIZE PANDAS FOR DATA CLEANING, HANDLING MISSING VALUES, AND STRUCTURING THE DATASET FOR ANALYSIS.

EXPLORATORY DATA ANALYSIS (EDA):

- **UNIVARIATE ANALYSIS:** APPLY PANDAS AND SEABORN TO ANALYZE SINGLE VARIABLES, SUCH AS COSTS AND RATINGS.
- **MULTIVARIATE ANALYSIS:** LEVERAGE SEABORN TO EXPLORE RELATIONSHIPS BETWEEN VARIABLES LIKE DAY-OF-THE-WEEK, RATINGS TO FOOD PREP. AND DELIVERY TIME.

DATA VISUALIZATION:

- UTILIZE SEABORN TO GENERATE INSIGHTFUL CHARTS AND PLOTS THAT CAN HELP BUSINESS DECISIONS.

THE END GOAL IS TO PRODUCE A COMPREHENSIVE BUSINESS PRESENTATION THAT OFFERS ACTIONABLE INSIGHTS AND RECOMMENDATIONS. THIS PRESENTATION AIMS TO DEFINE DATA-BACKED BUSINESS STRATEGIES.

DATA OVERVIEW

Data Overview encapsulates the structure and composition of FoodHub dataset. The data dictionary provides a clear definition of each attribute, while the examination reveals the dataset's shape and inherent characteristics. Analysis also offers a summary of key statistics, laying the groundwork for a deeper analytical exploration.

DATA DICTIONARY

COLUMN	TYPE	REMARKS
order_id	int64	Unique ID of the order
customer_id	int64	ID of the customer who ordered the food
restaurant_name	object	Name of the restaurant
cuisine_type	object	Cuisine ordered by the customer
cost_of_the_order	float64	Cost of the order
day_of_the_week	object	Indicates whether the order is placed on a weekday (Mon-Fri) or weekend (Sat-Sun)
rating	object	Rating given by the customer out of 5
food_preparation_time	int64	Food preparation time is the difference in minutes between order confirmation and pick-up confirmation.
delivery_time	int64	Delivery time is the minute difference between pick-up and drop-off timestamps.

RECORDS	COLUMNS	DATA TYPE
1,898	9	3 (int64, float64, object)

Dataset Information:

- The data is sourced from FoodHub, a food aggregator.
- It encompasses transactional details for a total of 1,898 orders.
- Each row records a unique customer order, detailing restaurant, cuisine, cost, day of order, rating, and prep and delivery times.
- Content Insights:
  - Customer Details: Insights into customers preferences.
  - Restaurants: Which restaurants are most engaged.
  - Cuisine Preferences: Types of cuisines ordered by customers.
  - Cost Metrics: Breakdown of the monetary value of each order.
  - Feedback Mechanism: Each order comes with a customer rating, providing a quality measure.
  - Time Analysis: The dataset captures the duration for food preparation and the subsequent delivery time.
- Time Frame: While exact order dates aren't provided, the dataset categorizes orders based on when they were placed: during the Weekday or over the Weekend.

DATA OVERVIEW

Answers 1 to 5

Answer 1:
Foodhub Order dataset (foodhum_orders.csv) has 1,898 rows and 9 column.

Answer 2:
The dataset has 3 data types:

- float64: cost_of_the_order
- int64: order_id, customer_id, food_preparation_time and delivery_time
- object: restaurant_name, cousine_type, day_of_the_week and rating

Answer 3:
There are no missing value (NaN) in the data. However, there are 736 (39%) orders where rating is 'Not given'.

Answer 4:
Food Preparation Time:

- Minimum: 20 mins,
- Average: 27.37 mins and
- Maximum: 35 mins.

Answer 5:
There are 736 orders where ratings are 'not given'.

UNIVARIATE ANALYSIS

Univariate analysis is a crucial step for statistical examination, concentrating on the singular examination of each variable within the FoodHub dataset. It aims to reveal individual patterns, identify anomalies, and summarize the distinct properties of each variable, paving the way for subsequent, more intricate analysis.

UNIVARIATE ANALYSIS

Answer 6

i. **Order IDs** are unique through the dataset.

- There are 1,898 orders.
- Each row in the dataset represents one order.

ii. FoodHub serves **1,200 customers**.

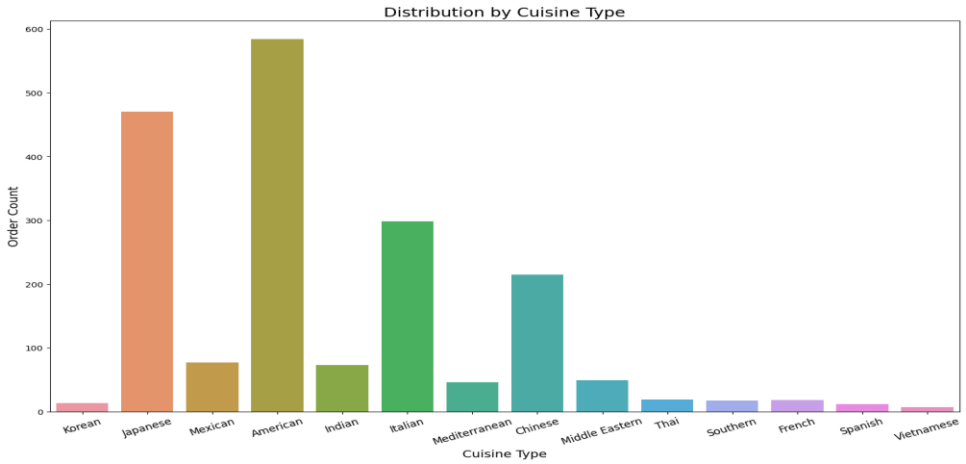
- Most (65%) customers ordered only once, 22% ordered twice.
- The data indicates potential customer retention issues at FoodHub.
- Consider strategies to boost repeat orders, like loyalty programs or personalized marketing.

iii. FoodHub works with **178 restaurants**.

- Top 14 contribute to 54% of business.
- Shake Shack dominates with 21% of orders, underscoring a stark disparity among the top 14 restaurants.
- Give incentives to top performers and encourage others.

iv. FoodHub offers **14 different types of cuisines** through the partner restaurants.

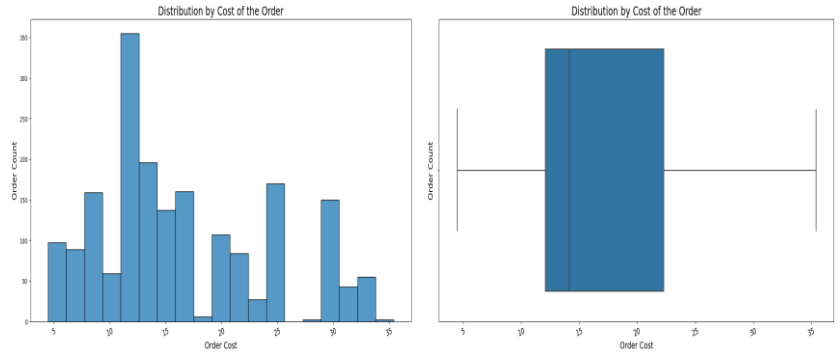
- American cuisine is most popular with 584 orders (31%)
- Followed by Japanese cuisine with 470 orders (25%)
- Italian 298 orders (16%) &
- Chinese 215 orders (11%).
- Vietnamese cuisine is ordered the least with 7 orders.



UNIVARIATE ANALYSIS

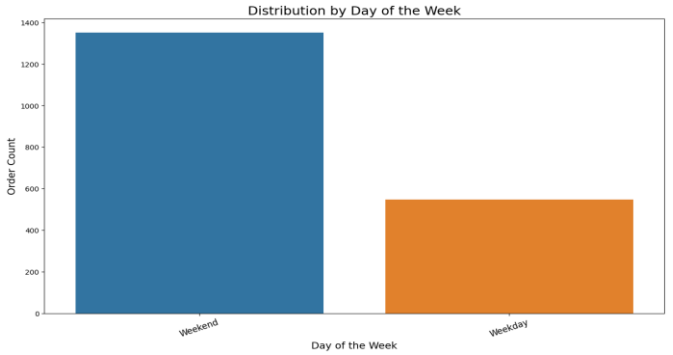
Answer 6

vi. **Cost of the Order**



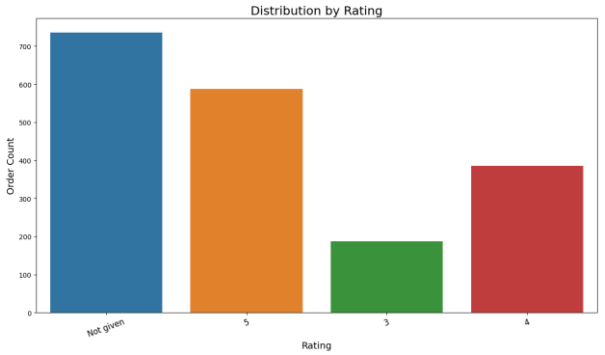
- It is positively skewed with 50% of orders between $11 and $23.
- There are no outliers.
- Median order cost is $14.14, Minimum of $4.57 and Maximum $35.41
- The data suggest there is wide variation of order cost.

vii. **Day of the week**



- It has only two values 'Weekend' and 'Weekday'.
- Analysis shows that majority, 1,351 (71%) orders were placed during weekend,
- Whereas only 547 (29%) of orders were placed during Weekday.

viii. **Rating**

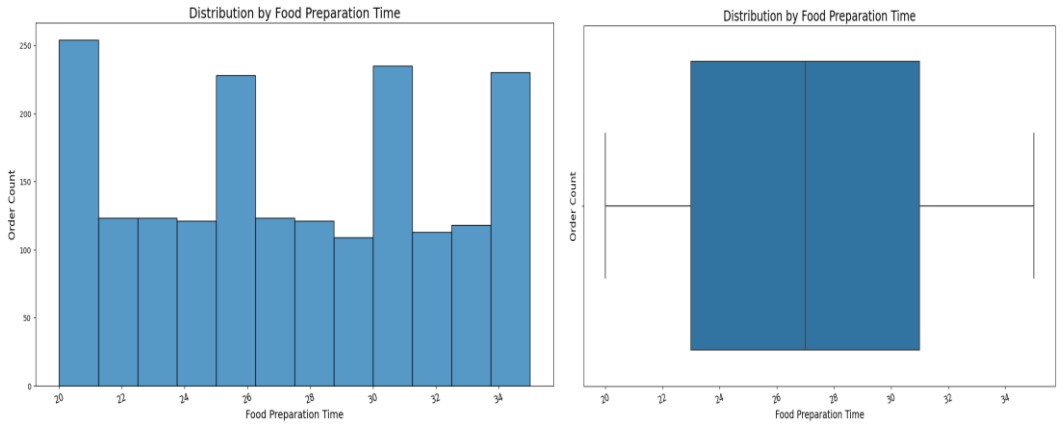


- It has 4 values, 'Not given', '5', '4' and '3'.
- Majority of orders, 736 (39%), do not have rating ('Not given')
- From the rated orders, Maximum orders 588 (31%) receive rating of '5'. 386 has rating of '4' and remaining 188 were rated '3.

UNIVARIATE ANALYSIS

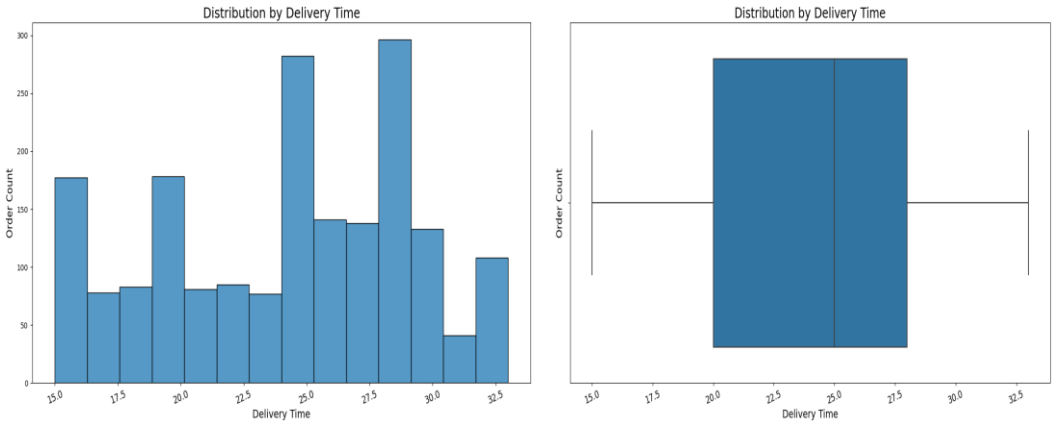
Answer 6

ix. **Food preparation time**



- It has quad-modal distribution on Histplot.
- With mean and median around 27 minutes, it shows that the data is symmetrically distributed around median.
- Boxplot also suggests 50% of preparation time is between 23 to 31 minutes and there are no outliers.

x. **Delivery time**



- Histplot show left skewed distribution.
- Boxplot shows 50% of orders take between 20 to 28 minutes.
- There are not outliers.

UNIVARIATE ANALYSIS

Answers 7 to 11

Answer 7:

Top 5 restaurants by order count are as follows:

Rank	Restaurants	Orders
1	Shake Shack	219 (21%)
2	The Meatball Shop	132 (7%)
3	Blue Ribbon Sushi	119 (6%)
4	Blue Ribbon Fried Chicken	96 (5%)
5	Parm	68 (4%)

Answer 8:

Most popular cuisine on Weekend is **American cuisine** with **415 orders, 31%** out of all orders booked during Weekend.

Answer 9:

29.24% (555) of orders cost over 20 dollars.

Answer 10:

Mean order delivery time of orders is **24.16 minutes**.

Answer 11:

Top 3 customers by number of orders placed are;

Customer ID	Order count
52832	13
47440	10
83287	9

We've noted that a majority, 65%, of customers have placed only a single order, while 23% have ordered twice. A minimal fraction of customers have made more than two orders.

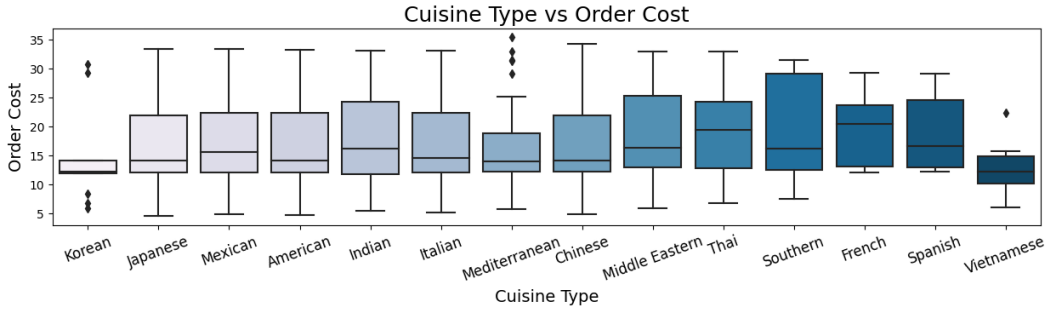
MULTIVARIATE ANALYSIS

Multivariate analysis delves into the intricate interplay between multiple variables of the FoodHub dataset. This analytical approach seeks to discern correlations, explore causal relationships, and understand the collective impact of these variables on consumer behavior and operational efficiency, providing a holistic view of the dataset's dynamics.

MULTIVARIATE ANALYSIS

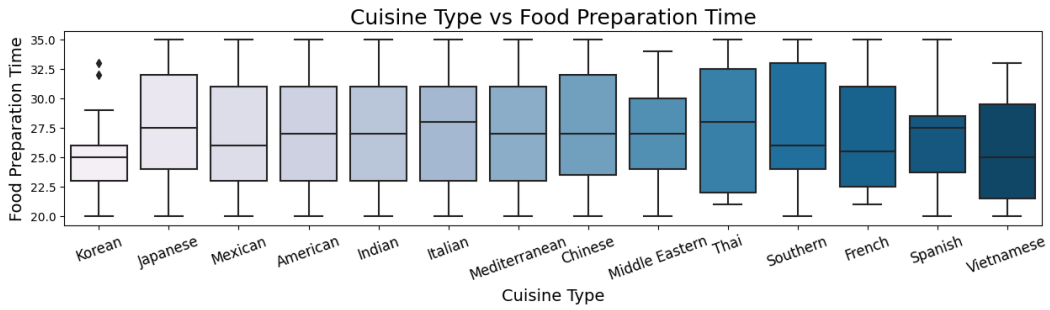
ANSWER 12

i. **Cuisine vs Cost of the Order**



- Korean, Mediterranean, and Vietnamese cuisines show extreme cost outliers. Other cuisine show uniform pricing.
- Most cuisines fall within a $14-$16 cost bracket; French cuisine has a higher median of $20, showing a negative skew.
- Southern cuisine has the widest cost range, with an IQR from $12 to $29, suggesting high variability.
- French cuisine ranks as the most expensive, and Vietnamese as the most affordable in terms of median cost.

ii. **Cuisine vs Food Preparation Time**

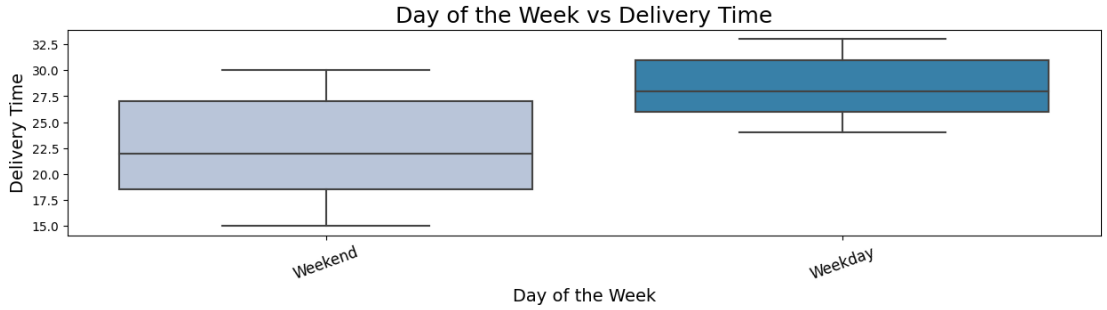


- Korean cuisine shows preparation time outliers, mainly 32-33 minutes, above the median of 25 minutes.
- Most cuisines prepare food in 26-28 minutes. Thai cuisine's preparation time varies most, with a 10-minute IQR from 22 to 32 minutes. Korean cuisine's 3-minute IQR from 23 to 26 minutes indicates low variability, similar to Spanish and Middle Eastern cuisines with 4 and 5-minute IQRs.
- Prep times generally span 20-35 minutes, but Korean cuisine ranges from 20-29 minutes, with outliers increasing this range.

MULTIVARIATE ANALYSIS

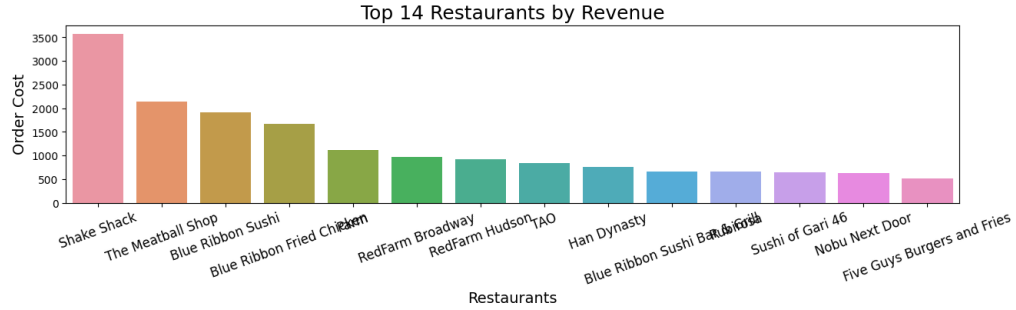
Answers 12

iii. **Day of the Week vs Delivery Time**



- No delivery time outliers detected for Weekdays or Weekends.
- Weekday deliveries typically take longer, with a median time of 27.5 minutes versus 22 minutes on Weekends.
- Weekend delivery times are more varied, with a 9-minute IQR (18-27 minutes), while Weekdays show a tighter 5-minute IQR (26-31 minutes).
- Delivery times span 15-30 minutes on Weekends and 24-33 minutes on Weekdays, with both showing a slight positive skew towards longer times.

iv. **Top 14 Restaurants by Revenue**

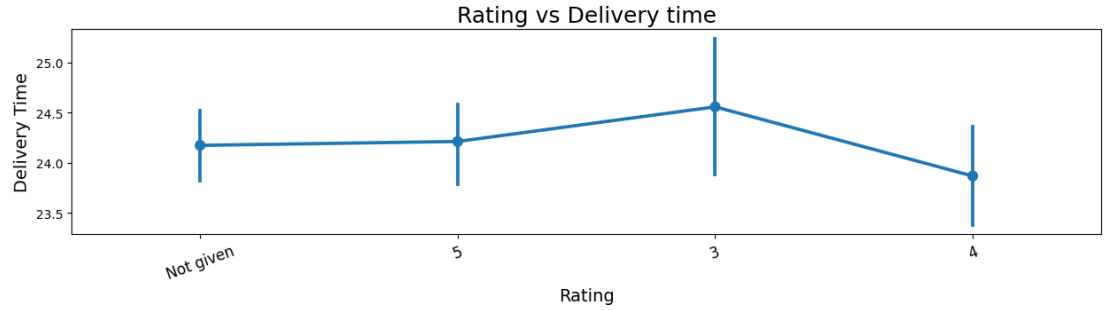


- The top 14 restaurants account for 54% of FoodHub's total orders and revenue.
- Shake Shack leads with $3,579.53 in revenue; The Meatball Shop and Blue Ribbon Sushi are next, with $2,145.21 and $1,903.95, respectively.
- Revenue for these top restaurant ranges from $506.47 to $3,579.53.
- There's a notable disparity in revenue among the top 14, with Five Guys earning the least at $506.47.

MULTIVARIATE ANALYSIS

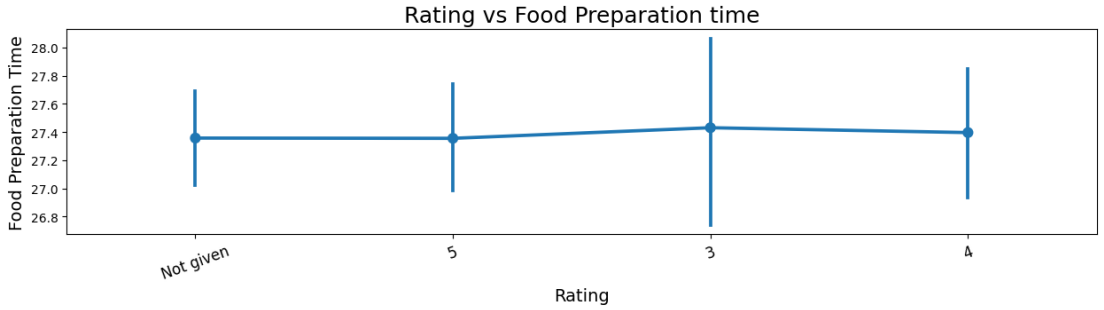
Answers 12

v. **Rating vs Delivery Time**



- About 39% of orders lack a rating.
- 3-star ratings reflect longer, variable delivery times averaging 24.50 minutes, while 4-stars have the shortest at 24 minutes, and 5-stars the most consistent, within a 23.75 to 24.65-minute interval.
- The data indicates trends between delivery times and ratings, but no clear correlation is evident.

vi. **Rating vs Food Preparation Time**

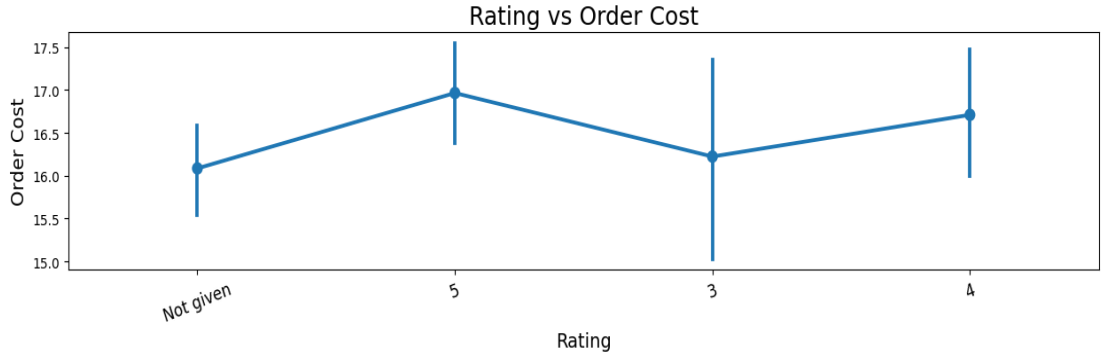


- Quicker, consistent food preparation correlates with higher ratings, as 5-star ratings match a narrow prep time range of 27 to 27.8 minutes,
- While 3-star ratings associate with longer, more variable prep times.
- Data indicates that efficiency in prep time may enhance ratings.

MULTIVARIATE ANALYSIS

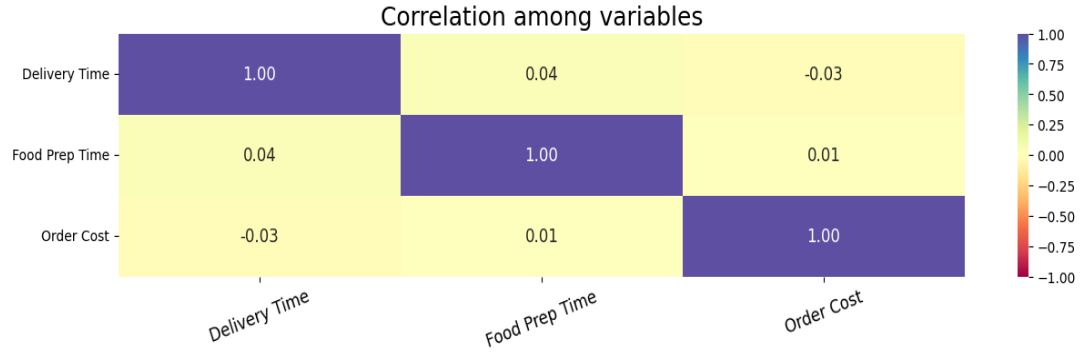
Answers 12

vii. **Rating vs Cost of the Order**



- The Point plot reveals higher-cost orders usually garner higher ratings.
- A rating of 3 often coincides with the lowest average order cost, alongside a broad confidence interval, reflecting cost variability.
- Conversely, a rating of 5 aligns with a tight confidence interval in order costs, indicating consistent pricing among top-rated orders.
- The trend suggests a general association between higher order costs and increased customer satisfaction.

viii. **Correlation among variables**



- Variables like order cost, food prep time, and delivery time lack strong correlation.
- Order cost has a minor positive correlation with prep time, hinting pricier orders take more prep time.
- There's a small negative correlation between order cost and delivery time, suggesting faster delivery for higher-priced orders.
- Food prep and delivery time have a weak positive correlation, indicating longer prep could extend delivery times.

MULTIVARIATE ANALYSIS

Answers 13 to 16

Answer 13:
Top restaurants fulfilling criteria of having 50 or more orders and average rating of 4 are as follows:

Rank	Restaurants	Rating
1	The Meatball Shop	4.51
2	Blue Ribbon Fried Chicken	4.33
3	Shake Shack	4.28
4	Blue Ribbon Sushi	4.22

Answer 14:
The net revenue earned by FoodHub is $6,166.30.

Answer 15:
10.54% (200) of orders take more than 60 minutes to deliver from the time order was place.

Answer 16:

- Average delivery time for orders on Weekdays is 28 minutes.
- Average delivery time for orders on Weekends is 22 minutes.

Weekdays have 6 minutes longer delivery time than weekends, which may imply that restaurants might be in busy business district

CONCLUSION & RECOMMENDATIONS

This section synthesizes key findings into actionable insights and strategic guidance for FoodHub. It not only encapsulates the essence of the data-driven journey but also charts a course for informed decision-making and targeted improvements, ensuring that the insights gleaned from the analysis translate into tangible business outcomes.

CONCLUSION - ANSWER 17

- FOODHUB EFFECTIVELY MEETS DIVERSE CULINARY PREFERENCES, WITH A STRONG LEANING TOWARDS AMERICAN AND JAPANESE CUISINES.
- PRICING STRATEGIES ARE CONSISTENT, WITH A MEDIAN ORDER VALUE THAT ALIGNS WITH CUSTOMER EXPECTATIONS.
- OPERATIONAL PROCESSES ARE STREAMLINED, WITH FOOD PREPARATION AND DELIVERY TIMES KEPT WITHIN REASONABLE LIMITS.
- CUSTOMER SATISFACTION, AS MEASURED BY RATINGS, IS GENERALLY HIGH, BUT THERE'S ROOM FOR IMPROVEMENT GIVEN THE MISSING RATINGS.
- THE WEEKEND SURGE IN ORDERS INDICATES A POTENTIAL AREA FOR TARGETED MARKETING AND STAFFING STRATEGIES.
- SHAKE SHACK'S REVENUE DOMINANCE HIGHLIGHTS THE IMPORTANCE OF BRAND STRENGTH AND CUSTOMER LOYALTY.
- THE LACK OF CORRELATION BETWEEN COST AND DELIVERY TIMES SUGGESTS THAT CUSTOMERS DO NOT PRIORITIZE SPEED OVER QUALITY.
- THE ANALYSIS UNDERSCORES THE NEED FOR A MORE ROBUST FEEDBACK MECHANISM TO BETTER UNDERSTAND CUSTOMER SATISFACTION DRIVERS.

RECOMMENDATIONS - ANSWER 17

- ENHANCE CUSTOMER LOYALTY PROGRAMS, FOCUSING ON TOP SPENDERS TO INCREASE REVENUE.
- IMPLEMENT TARGETED PROMOTIONS DURING WEEKENDS TO CAPITALIZE ON HIGHER ORDER VOLUMES.
- ENCOURAGE CUSTOMERS TO RATE WITH INCENTIVES TO IMPROVE DATA ON CUSTOMER SATISFACTION.
- EXPLORE PARTNERSHIPS WITH TOP-PERFORMING RESTAURANTS TO EXPAND OFFERINGS AND MEET DEMAND.
- OPTIMIZE FOOD PREPARATION PROCESSES TO MAINTAIN OR IMPROVE EFFICIENCY WITHOUT COMPROMISING QUALITY.
- DEVELOP A MORE GRANULAR TRACKING SYSTEM FOR FOOD PREPARATION AND DELIVERY TO IDENTIFY BOTTLENECKS.
- CONDUCT FURTHER ANALYSIS ON THE IMPACT OF ORDER COST ON CUSTOMER SATISFACTION AND REPEAT BUSINESS.
- CONSIDER DYNAMIC STAFFING SOLUTIONS TO ADDRESS THE INCREASED DEMAND ON WEEKENDS AND IMPROVE DELIVERY TIMES.
- LEVERAGE DATA INSIGHTS TO PERSONALIZE MARKETING STRATEGIES, TAILORING OFFERS TO CUSTOMER PREFERENCES AND ORDERING PATTERNS.
- REGULARLY REVIEW AND ADJUST PRICING STRATEGIES TO ENSURE COMPETITIVENESS WHILE MAINTAINING PROFITABILITY.

THANK YOU

SHAISHAV MERCHANT

PROJECT - FOODHUB | PYTHON FOUNDATIONS