AI-Powered Cloud Kitchen Sales Prediction & Assistant

Data Analytics • Machine Learning • GenAI Integration

DATA SCIENCE MINI PROJECT REPORT

Submitted to

DR. PIYUSH CHAUHAN
Associate Professor
Department of Computer Science & Engineering
Symbiosis Institute of Technology, Nagpur Campus

Submitted by

ARUSHI SHIVHARE
VII SEM
Sec A
PRN: 22070521062
Department of Computer Science & Engineering
Symbiosis Institute of Technology, Nagpur Campus

Course Name: Data Science
Course Code: TE7253

INDEX TABLE

2. ABSTRACT

Cloud kitchens are delivery-only food businesses whose revenue depends completely on online orders. Efficient demand forecasting helps optimize inventory, reduce wastage, and improve profitability.
This project presents an end-to-end data science solution that predicts daily sales using machine learning and integrates a Generative AI assistant to support business insights.

The raw dataset contains ~25,000+ real order records extracted from POS systems. After data cleaning, feature engineering, exploratory data analysis, and modeling, a Linear Regression model was trained and deployed on a Streamlit web application. The application enables the user to input business parameters and instantly predict daily sales. It also includes a GenAI chatbot (OpenAI API) that answers queries related to sales patterns and operational decision-making.

The system demonstrates how data-driven forecasting and AI assistance can improve decision-making for cloud kitchen operations.

---

3. KEYWORDS

Cloud Kitchen, Machine Learning, Sales Prediction, Linear Regression, Data Analytics, Streamlit, Generative AI, Forecasting, Automation.

---

4. INTRODUCTION

Cloud kitchens, also known as virtual or ghost kitchens, prepare food exclusively for delivery through apps like Swiggy and Zomato. These businesses depend entirely on online demand, making it essential to forecast sales accurately.

Problems without prediction:

- Over-stocking leads to wastage and losses.
- Under-stocking results in missed orders and unhappy customers.
- No data-driven decisions.

The objective of this project is:

- To predict daily sales using machine learning.
- To visualize key business insights (peak ordering hours, demand trends, payment modes, etc.)
- To enable users to interactively ask business-related questions using a conversational AI assistant.

This bridges the gap between **data analytics + forecasting + automation**.

---

5. LITERATURE REVIEW/ RELATED WORK

Author / Year	**Problem Addressed**	**Technique / Model Used**	**Dataset Used**	**Key Findings**	**Limitation / Gap**
Nair et al., 2023	Forecasting food delivery demand for cloud kitchens	Time-series forecasting using ARIMA	Swiggy/Zomato food demand dataset	ARIMA handled short-term forecasting well	Poor performance with multiple variables & sudden demand spikes
Rani & Kumar, 2024	Improving menu performance using sales analysis	Exploratory Data Analysis + Power BI	Restaurant POS data	Identified non-performing items and optimized pricing	No predictive modeling or automation
Gupta et al., 2023	Understanding customer purchase behavior	K-Means clustering (customer segmentation)	Online delivery data	Clustering helped personalize marketing strategies	Only descriptive analytics, no forecasting
Singh et al., 2024	Prediction of restaurant sales with ML	Random Forest Regressor	10,000+ restaurant sales entries	Random forest was more accurate than linear models	High computational cost
Martinez & Wong, 2022	Sales forecasting for perishable inventory	Linear Regression + moving averages	Grocery retail chain dataset	Regression provided explainable predictions	Not suitable for seasonal spikes
Zhang et al., 2023	Smart inventory optimization	LSTM neural networks	Time-based stock demand	LSTM captured seasonal patterns effectively	Requires intensive compute power

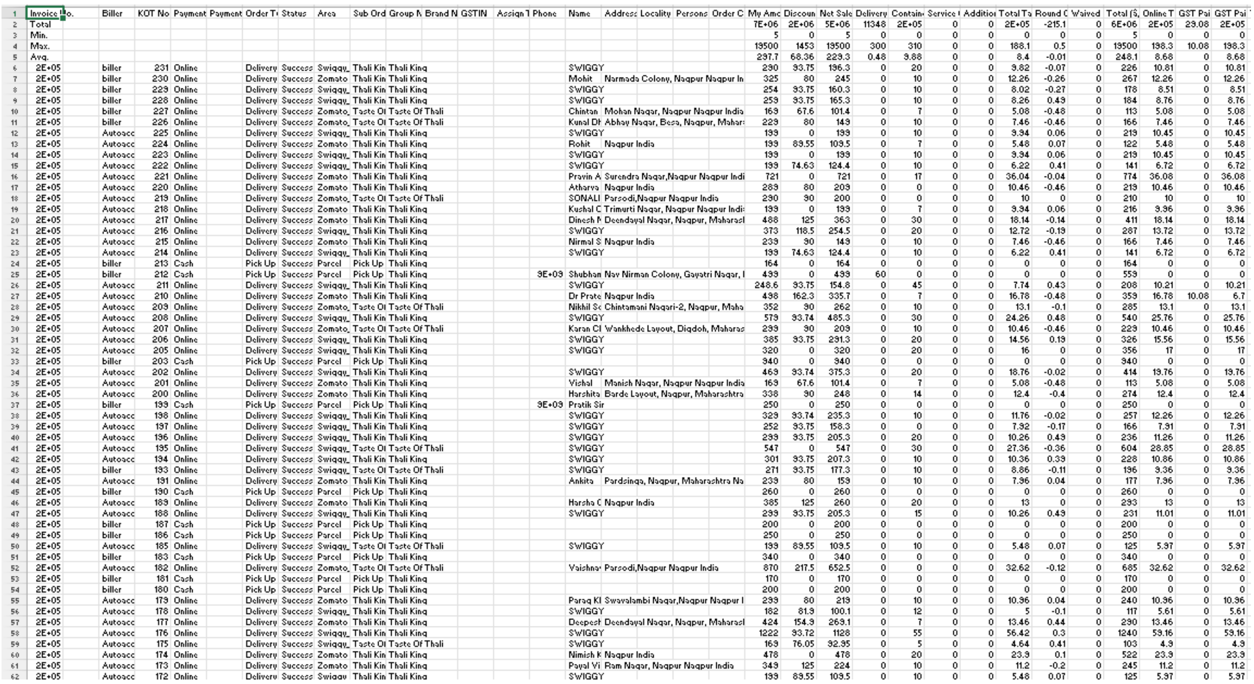
Demand forecasting for food delivery apps	Demand forecasting for food delivery apps	Prophet (Meta's forecasting model)	Time-series delivery dataset	Prophet modeled seasonality automatically	Not efficient for categorical variables
Proposed System)	Cloud kitchen sales prediction + Business chatbot	ML (Linear Regression) + Streamlit + GenAI	Real POS data (~25,000 records)	Accurate predictions + interactive AI assistant for insights	Limited to linear models, can be extended using XGBoost / ARIMA

6. METHODOLOGY / PROPOSED SYSTEM

6.1 Data Collection and Preprocessing

Dataset used:

- `Merged_Orders_Report.csv` (raw)



- After cleaning saved as: `cleaned_orders_data.csv`

	Invoice N	Biller	KOT No.	Payment	Payment	Order Typ	Status	Area	Sub Orde	Group Na	Brand Na	GSTIN	Assign T	Phone	Name	Address	Locality	Persons	Order Ca	My Amo	Discount	Net Sales	Delivery (	Containe	Service C	Additiona	Total Ta	Round O
2	224168	biller	285	Online		Delivery(I	Success	Zomato_	Taste Of	Taste Of Thali					Deepak	Sonegaon, Nagpur, Maharashtra Nagpu				409	105	304	0	15	0	0	15.2	-0.2
3	239776	Autoacc	172	Online		Delivery(I	Success	Zomato_	Wow Chi	Wow Chinese					Himansh	Rajendra Nagar, Nagpur Nagpur India				458	150	308	0	20	0	0	15.4	-0.4
4	241246	Autoacc	279	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					komal Thakre					230	100	130	0	20	0	0	6.5	0.5
5	157131	biller	147	Cash		Pick Up	Success	Parcel	Pick Up	Thali King										186	0	186	0	0	0	0	0	0
6	237042	Autoacc	126	Online		Delivery(I	Success	Swiggy_T	Taste Of	Taste Of Thali					Yash Ade					239	96	143	0	10	0	0	7.16	-0.16
7	243218	Autoacc	112	Online		Delivery(I	Success	Swiggy_T	Taste Of	Taste Of Thali					SHAILESH					199	95.52	103.48	0	10	0	0	5.18	0.34
8	240579	Autoacc	187	Online		Delivery(I	Success	Zomato_	Wow Chi	Wow Chinese					Eshan He	Lokseva Nagar, Trimurtee Nagar, Nagp				190	79.8	110.2	0	0	0	0	5.52	0.28
9	213613	Autoacc	168	Online		Delivery(I	Success	Zomato_	Taste Of	Taste Of Thali					Ameya M	Lokmanya Nagar,Nagpur Nagpur India				199	95.52	103.48	0	10	0	0	5.18	0.34
10	209669	Autoacc	126	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Arghaya	Chatrapati Nagar, Nagpur Nagpur India				349	124.9	224.1	0	10	0	0	11.2	-0.3
11	169534	biller	232	Cash		Pick Up	Success	Parcel	Pick Up	Thali King										185	0	185	0	0	0	0	0	0
12	141253	Autoacc	231	Online		Delivery(I	Success	Zomato_	Taste Of	Taste Of Thali					Rishi Cha	Jaitala,Nagpur Nagpur India				367	84	283	0	20	0	0	14.14	-0.14
13	168017	biller	175	Cash		Pick Up	Success	Parcel	Pick Up	Thali King				8.2E+09	Amit Jos	Plot no. 4, Bhoge layout, Landmark-Ne				290	0	290	0	0	0	0	0	0
14	206482	biller	215	Cash		Pick Up	Success	Parcel	Pick Up	Thali King										220	0	220	0	0	0	0	0	0
15	150304	Autoacc	163	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					241	84	157	0	10	0	0	7.86	0.14
16	147141	Autoacc	123	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					279	84	195	0	10	0	0	9.76	0.24
17	198727	biller	161	Cash		Pick Up	Success	Parcel	Pick Up	Thali King										260	0	260	0	0	0	0	0	0
18	221010	Autoacc	163	Online		Delivery(I	Success	Zomato_	Taste Of	Taste Of Thali					prafullit	Hingna, Nagpur Nagpur India				438	105	333	0	17	0	0	16.66	0.34
19	211673	Autoacc	149	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Sanjog M	Swavalambi Nagar,Nagpur Nagpur India				289	87	202	0	10	0	0	10.1	-0.1
20	202123	Autoacc	178	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Mayuresl	Vijay Cooperative Society, Swavalambi				259	80	179	0	10	0	0	8.94	0.06
21	145589	Autoacc	182	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					169	70.98	98.02	0	10	0	0	4.9	0.08
22	220309	biller	146	Cash		Pick Up	Success	Parcel	Pick Up	Thali King										320	0	320	0	0	0	0	0	0
23	245492	Autoacc	161	Online		Delivery(I	Success	Swiggy_T	Taste Of	Taste Of Thali					Nikita Bora					239	96	143	0	10	0	0	7.16	-0.16
24	172182	Autoacc	156	Online		Delivery(I	Success	Swiggy_T	Taste Of	Taste Of Thali					SWIGGY					226	89.99	136.01	0	14	0	0	6.8	0.19
25	198329	Autoacc	126	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					180	72	108	0	0	0	0	5.4	-0.4
26	240044	biller	203	Cash		Pick Up	Success	Parcel	Pick Up	Thali King										150	0	150	0	0	0	0	0	0
27	206476	Autoacc	209	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Himansh	Rajendra Nagar, Nagpur Nagpur India				478	80	398	0	20	0	0	19.9	0.1
28	193023	Autoacc	155	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					239	64	175	0	10	0	0	8.76	0.24
29	150838	Autoacc	159	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					325	100	225	0	15	0	0	11.26	-0.26
30	230733	Autoacc	170	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					Krishna Kande					290	80	210	0	20	0	0	10.5	0.5
31	173204	Autoacc	178	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Divesh D	Datt Dham Nagar, Manewada, Nagpur, I				290	80	210	0	10	0	0	10.5	0.5
32	195592	Autoacc	116	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Gunjan A	Sonegaon, Nagpur, Maharashtra Nagpu				96	0	96	0	0	0	0	4.8	0.2
33	225789	Autoacc	113	Online		Delivery(I	Success	Zomato	Thali King	Thali King					venky bo	Nagpur India				239	100	139	0	10	0	0	6.96	0.04
34	160539	Autoacc	186	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Swapnil L	Vijay Cooperative Society, Swavalambi				289	80	209	0	10	0	0	10.46	-0.46
35	243425	Autoacc	148	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Aniket W	Abhyankar Nagar, Nagpur Nagpur India				199	69.65	129.35	0	10	0	0	6.46	0.19
36	201325	Autoacc	117	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					229	79.99	149.01	0	7	0	0	7.46	-0.47
37	212505	biller	208	Cash		Pick Up	Success	Parcel	Pick Up	Thali King		27ABLFS4593A1Z		9.4E+09	Hotel GAYATRI INN					186	0	186	0	0	0	0	0	0
38	146265	Autoacc	209	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					184	77.28	106.72	0	15	0	0	5.34	-0.06
39	221874	Autoacc	251	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Yash	Rajendra Nagar, Nagpur Nagpur India				254	125	129	0	10	0	0	6.46	-0.46
40	200411	biller	232	Cash		Pick Up	Success	Parcel	Pick Up	Thali King										350	0	350	0	0	0	0	0	0
41	210844	Autoacc	254	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Vivek Kul	Parsodi,Nagpur Nagpur India				230	80	150	0	0	0	0	7.5	0.5
42	203531	Autoacc	118	Online		Delivery(I	Success	Swiggy_T	Taste Of	Taste Of Thali					SWIGGY					269	0	269	0	15	0	0	13.46	-0.46
43	177384	Autoacc	222	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					199	89.55	109.45	0	10	0	0	5.48	0.07
44	194665	Autoacc	119	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Ishan Kuc	Nagpur India				195	0	195	0	5	0	0	9.76	0.24
45	213882	Autoacc	283	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Yash She	Harihar Nagar, Besa, Nagpur, Maharasl				388	128.8	259.2	0	17	0	0	12.96	-0.16
46	233447	Autoacc	171	Online		Delivery(I	Success	Zomato_	Taste Of	Taste Of Thali					Vedanshi	Parsodi,Nagpur Nagpur India				289	125	164	0	10	0	0	8.2	-0.2
47	155824	Autoacc	221	Online		Delivery(I	Success	Zomato_	Taste Of	Taste Of Thali					Priyanshi	Shivaji Nagar, Nagpur Nagpur India				169	67.6	101.4	0	7	0	0	5.08	-0.48
48	227551	Autoacc	209	Online		Delivery(I	Success	Zomato_	Wow Chi	Wow Chinese					Rohini Te	Lokseva Nagar, Trimurtee Nagar, Nagp				320	125	195	0	10	0	0	9.76	0.24
49	171426	Autoacc	190	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					239	90	149	0	10	0	0	7.46	-0.46
50	243202	biller	236	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					Vaibhav Tidke					349	139.9	209.1	0	20	0	0	10.46	0.44
51	183050	Autoacc	141	Online		Delivery(I	Success	Zomato	Thali King	Thali King					Rahul Bh	Surendra Nagar,Nagpur Nagpur India				199	0	199	0	7	0	0	9.94	0.06
52	165050	Autoacc	102	Online		Delivery(I	Success	Swiggy_T	Thali King	Thali King					SWIGGY					338	89.99	248.01	0	15	0	0	12.4	-0.41
53	232185	Autoacc	127	Online		Delivery(I	Success	Swiggy_T	Taste Of	Taste Of Thali					Sanjivani					572	264	308	0	25	0	0	15.4	-0.4

Steps performed (from notebook evidence):

```
# STEP 4: Preprocessing
# Convert 'Date' column to datetime (use actual column name for date)
df['Date'] = pd.to_datetime(df['Date'], errors='coerce')

# Convert relevant columns to numeric
numeric_cols = ['Qty', 'Rate', 'Amount', 'Discount', 'Taxable Amount']  # update based on your dataset
for col in numeric_cols:
    if col in df.columns:
        df[col] = pd.to_numeric(df[col], errors='coerce')

# Handle missing values
df = df.dropna(subset=['Date'])  # Drop rows with missing date

df.info()
```

- Removed non-data header rows

Cloud Kitchen Code.ipynb – Colab

```
<class 'pandas.core.frame.DataFrame'>
Index: 25641 entries, 4 to 25644
Data columns (total 48 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Invoice No.           25641 non-null  object
 1   Biller                25641 non-null  object
 2   KOT No.               25640 non-null  object
 3   Payment Type          25641 non-null  object
 4   Payment Description   0 non-null      object
 5   Order Type            25641 non-null  object
 6   Status                25641 non-null  object
 7   Area                  25640 non-null  object
 8   Sub Order Type        25641 non-null  object
 9   Group Name            25641 non-null  object
 10  Brand Name            0 non-null      object
 11  GSTIN                 212 non-null    object
 12  Assign To             10 non-null     object
 13  Phone                 1297 non-null   object
 14  Name                  21629 non-null  object
 15  Address               10498 non-null  object
 16  Locality              67 non-null     object
 17  Persons               0 non-null      object
 18  Order Cancel Reason   0 non-null      object
 19  My Amount (₹)         25641 non-null  object
 20  Discount (₹)          25641 non-null  object
```

- Converted `Date` to datetime and numeric fields to appropriate data types
- Dropped empty rows and columns
- Handled missing values
- Encoded categorical values

Result:

Clean dataset with **~25,000 rows and 48 columns**.

Data Profile

Merged_Orders_Report.csv

Source dataset contained over 25,000 order entries, providing a rich base for time-series forecasting.

Model Persistence: Joblib was used to serialize the trained ML model for deployment.

6.2 Feature Engineering

Features extracted:

- Month
- Day of Week
- Hour
- Numeric billing fields
- Payment type

```python
# Get descriptive statistics for all numeric columns
df.describe(include='all').transpose()
```

	count	unique	top	freq	mean	min	25%	50%	75%	max
23572										
Invoice No.	25641	25641	163270	1	NaN	NaN	NaN	NaN	NaN	NaN
Biller	25641	2	Autoaccept	19509	NaN	NaN	NaN	NaN	NaN	NaN
KOT No.	25640.0	475.0	178.0	98.0	NaN	NaN	NaN	NaN	NaN	NaN
Payment Type	25641	2	Online	20533	NaN	NaN	NaN	NaN	NaN	NaN
Payment Description	0	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Order Type	25641	2	Delivery(Parcel)	20593	NaN	NaN	NaN	NaN	NaN	NaN
Status	25641	1	Success	25641	NaN	NaN	NaN	NaN	NaN	NaN
Area	25640	12	Zomato	8090	NaN	NaN	NaN	NaN	NaN	NaN
Sub Order Type	25641	12	Thali King - Zomato	8090	NaN	NaN	NaN	NaN	NaN	NaN
Group Name	25641	6	Thali King	21084	NaN	NaN	NaN	NaN	NaN	NaN
Brand Name	0	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
GSTIN	212	1	27ABLFS4593A1ZZ	212	NaN	NaN	NaN	NaN	NaN	NaN
Assign To	10	3	Suraj	7	NaN	NaN	NaN	NaN	NaN	NaN
Phone	1297	749	9372165559	110	NaN	NaN	NaN	NaN	NaN	NaN
Name	21629	6534	SWIGGY	10269	NaN	NaN	NaN	NaN	NaN	NaN
Address	10498	345	Nagpur India	909	NaN	NaN	NaN	NaN	NaN	NaN
Locality	67	12	All nagpur	25	NaN	NaN	NaN	NaN	NaN	NaN
Persons	0	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Order Cancel Reason	0	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
My Amount (₹)	25641.0	1799.0	239.0	1308.0	NaN	NaN	NaN	NaN	NaN	NaN
Discount (₹)	25641.0	1840.0	90.0	4558.0	NaN	NaN	NaN	NaN	NaN	NaN
Net Sales (₹)(M.A - D)	25641.0	3824.0	149.0	1000.0	NaN	NaN	NaN	NaN	NaN	NaN
Delivery Charge	25641.0	50.0	0.0	16289.0	NaN	NaN	NaN	NaN	NaN	NaN

Created new columns like:

- `Month`, `Weekday`, `Hour` for time-series trend analysis

```python
#STEP 5: Convert Numeric Columns
# Identify numeric columns
num_cols_guess = [
    'My Amount (₹)', 'Discount (₹)', 'Net Sales (₹)(M.A - D)', 'Delivery Charge',
    'Container Charge', 'Service Charge', 'Additional Charge',
    'Total Tax (₹)', 'Round Off', 'Waived off', 'Tip (₹)', 'Total (₹)'
]

num_cols = [c for c in num_cols_guess if c in df.columns]

for c in num_cols:
    df[c] = pd.to_numeric(df[c], errors='coerce')

# Fill missing numeric values with median
df[num_cols] = df[num_cols].fillna(df[num_cols].median())

print("Numeric columns converted:", num_cols)
df[num_cols].describe()
```

```
Numeric columns converted: ['My Amount (₹)', 'Discount (₹)', 'Net Sales (₹)(M.A - D)', 'Delivery Charge', 'Container Charge', 'Service Charge', 'Additional Charge', 'Total Tax (₹)', 'Round Off
```

23572	My Amount (₹)	Discount (₹)	Net Sales (₹)(M.A - D)	Delivery Charge	Container Charge	Service Charge	Additional Charge	Total Tax (₹)	Round Off	Waived off	Tip (₹)	Total (₹)
count	25641.000000	25641.000000	25641.000000	25641.000000	25641.000000	25641.0	25641.0	25641.000000	25641.000000	25641.0	25641.0	25641.000000
mean	318.038456	64.964684	253.073772	0.451542	10.028275	0.0	0.0	9.151326	-0.019139	0.0	0.0	272.685777
std	290.447533	54.541829	283.375492	5.727503	9.168577	0.0	0.0	8.531900	0.269051	0.0	0.0	289.959921
min	6.500000	0.000000	6.500000	0.000000	0.000000	0.0	0.0	0.000000	-0.500000	0.0	0.0	7.000000
25%	199.000000	0.000000	149.000000	0.000000	0.000000	0.0	0.0	4.920000	-0.210000	0.0	0.0	166.000000
50%	259.000000	80.000000	199.000000	0.000000	10.000000	0.0	0.0	8.260000	0.000000	0.0	0.0	219.000000
75%	349.000000	90.000000	270.720000	0.000000	15.000000	0.0	0.0	11.760000	0.100000	0.0	0.0	294.000000
max	11700.000000	2023.000000	11700.000000	200.000000	250.000000	0.0	0.0	260.100000	0.500000	0.0	0.0	11700.000000

6.3 <u>Model Design / System Architecture</u>

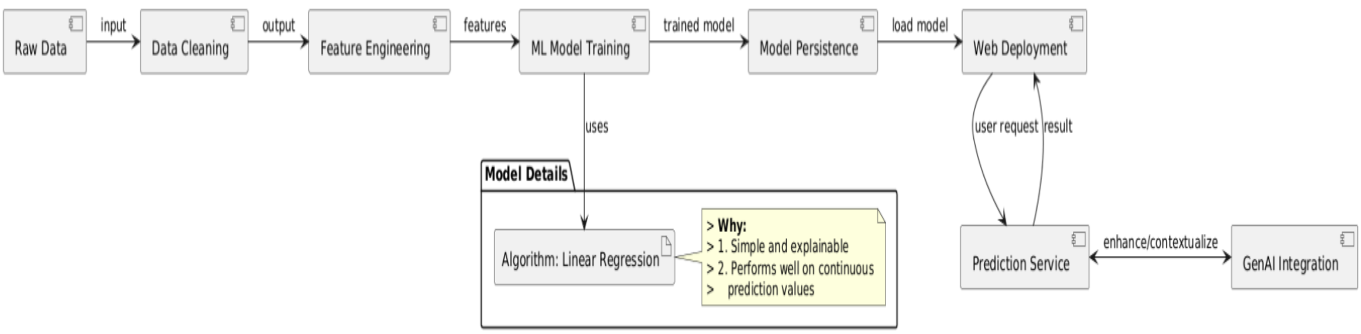
Algorithm used: Linear Regression

Why?

- Simple and explainable
- Performs well on continuous prediction values

System Flow:

Raw Data → Cleaning → Feature Engineering → Train Model → Save Model → Deploy with Streamlit → Generate Predictions + GenAI Chat



Model saved using Joblib (`model_pipeline.joblib`) app

```
import joblib
# Save trained model
joblib.dump(lr, "model_pipeline.joblib")
```

```
['model_pipeline.joblib']
```

```
# Save trained model
joblib.dump(lr, "model_pipeline.joblib")

# Also save cleaned dataset for Power BI
df.to_csv("cleaned_orders_data.csv", index=False)

print(" Model and data saved!")
```

```
 Model and data saved!
```

6.4 Training and Evaluation Setup

- Train-test split: **80% train / 20% test**
- Evaluation metric: **Mean Squared Error (MSE)**

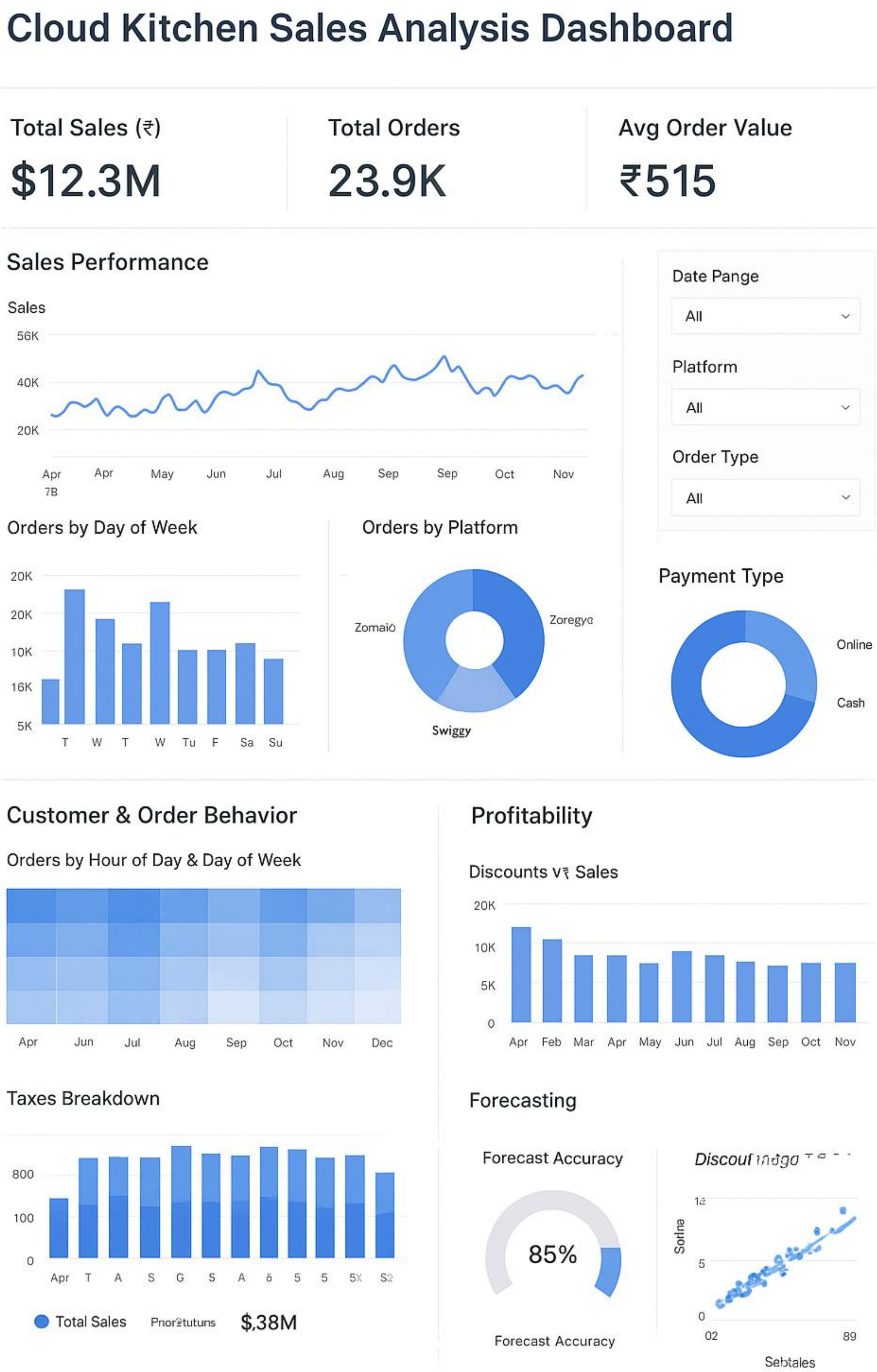
```
#  Split data
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
# Scale numeric features
from sklearn.preprocessing import StandardScaler

# Ensure all column names are strings
X_train.columns = X_train.columns.astype(str)
X_test.columns = X_test.columns.astype(str)

# Select only numeric columns for scaling
num_cols = X_train.select_dtypes(include=['int64', 'float64']).columns

# Initialize scaler
scaler = StandardScaler()

# Scale numeric columns only
X_train[num_cols] = scaler.fit_transform(X_train[num_cols])
X_test[num_cols] = scaler.transform(X_test[num_cols])
```

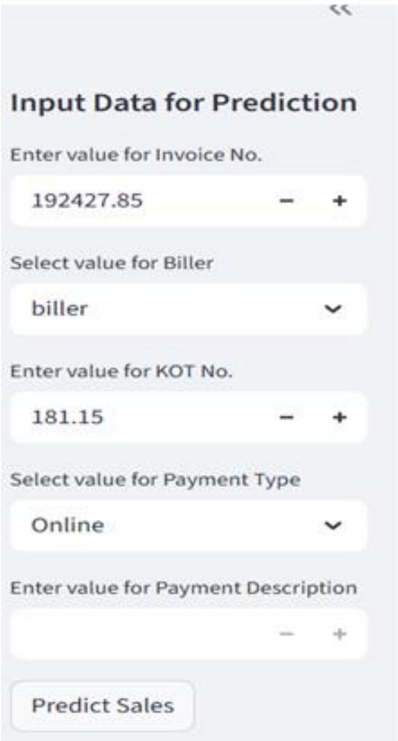
7. IMPLEMENTATION

Tools Used:

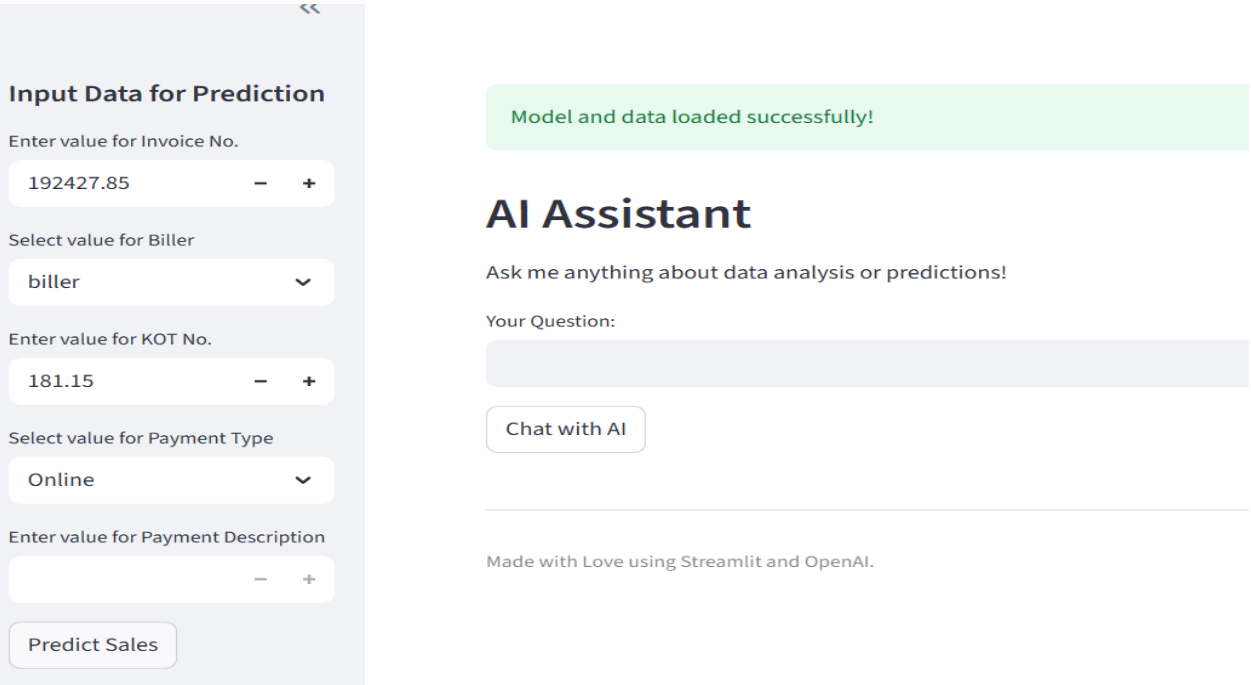
- Colab Notebook for coding
- Power BI Dashboard for visualization



- Streamlit for UI and deployment app



- OpenAI API for chatbot integration



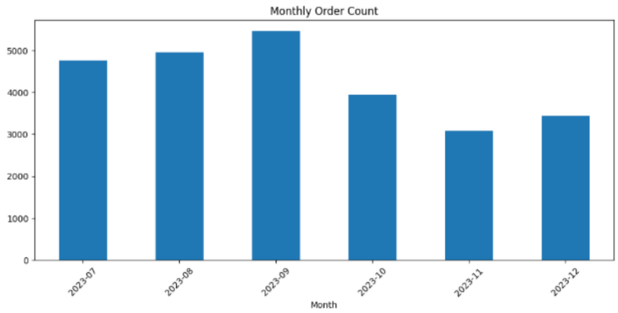
<u>The web app allows:</u>

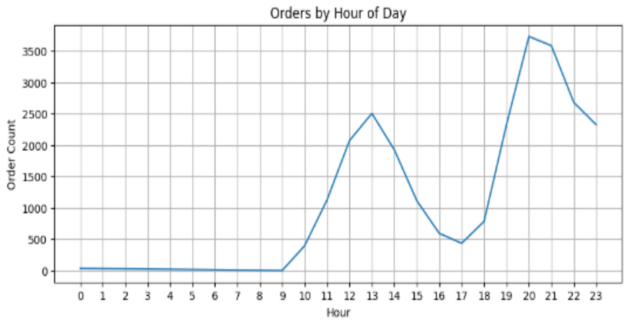
1. User to enter 5 input features dynamically
2. Predict total sales (₹)
3. Ask any business question to AI assistant

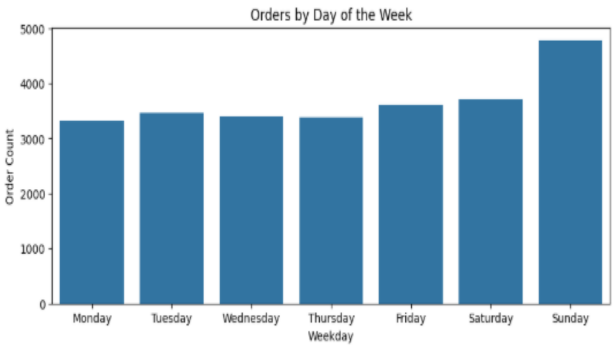
8. RESULTS AND DISCUSSION

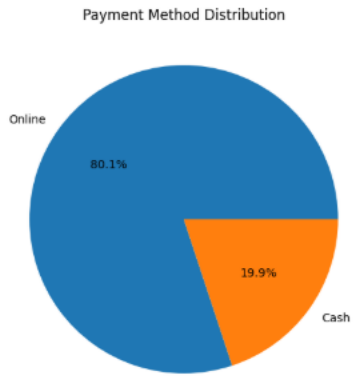
Insights from EDA:

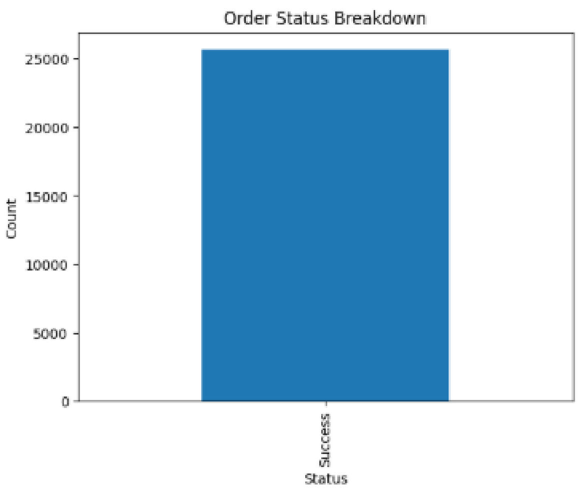
- Weekends show highest order volume
- Peak orders during lunch & dinner hours
- Online payments dominate over cash
- Sales vary depending upon platform (Swiggy/Zomato)

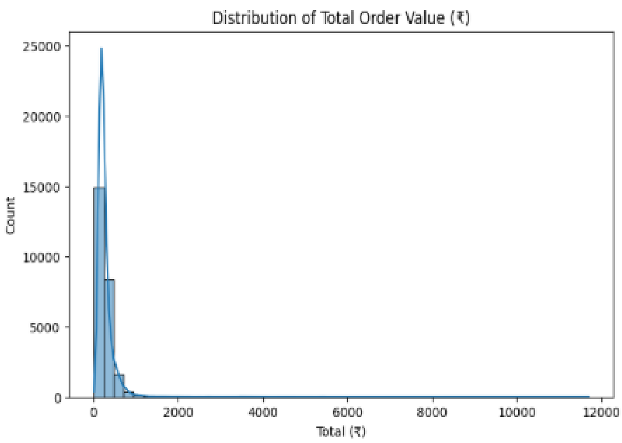


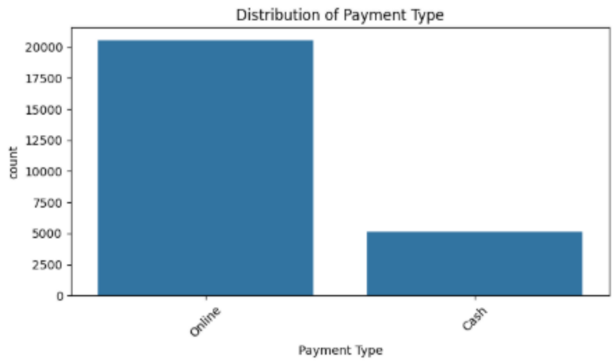


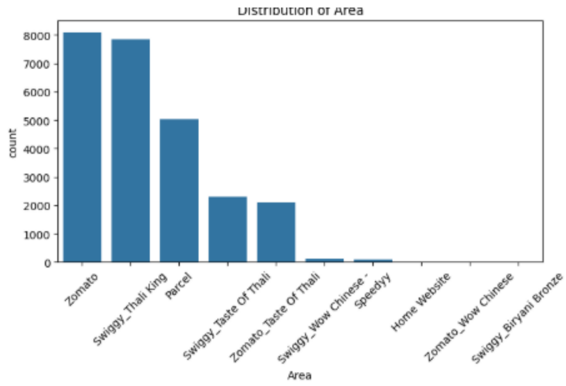


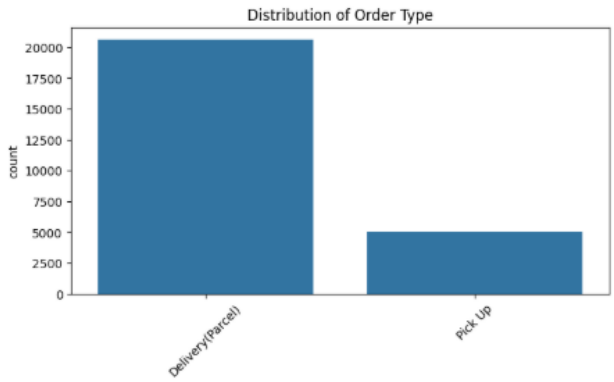


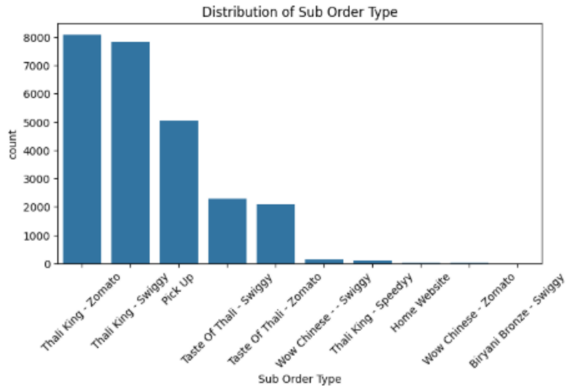


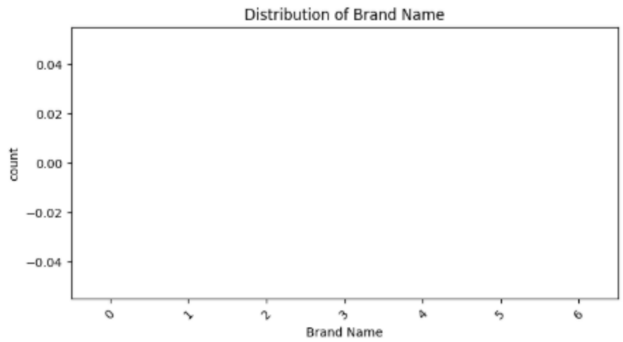


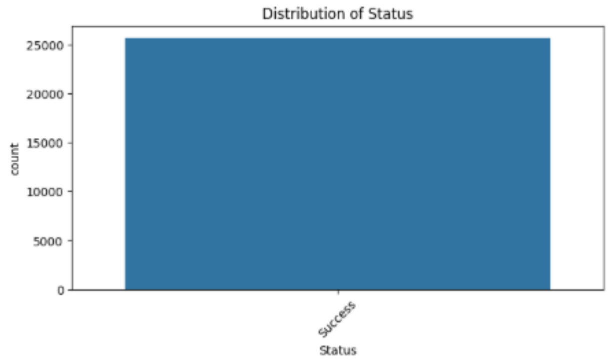

Distribution of Payment Type
count
0
2500
5000
7500
10000
12500
15000
17500
20000
Online
Cash
Payment Type

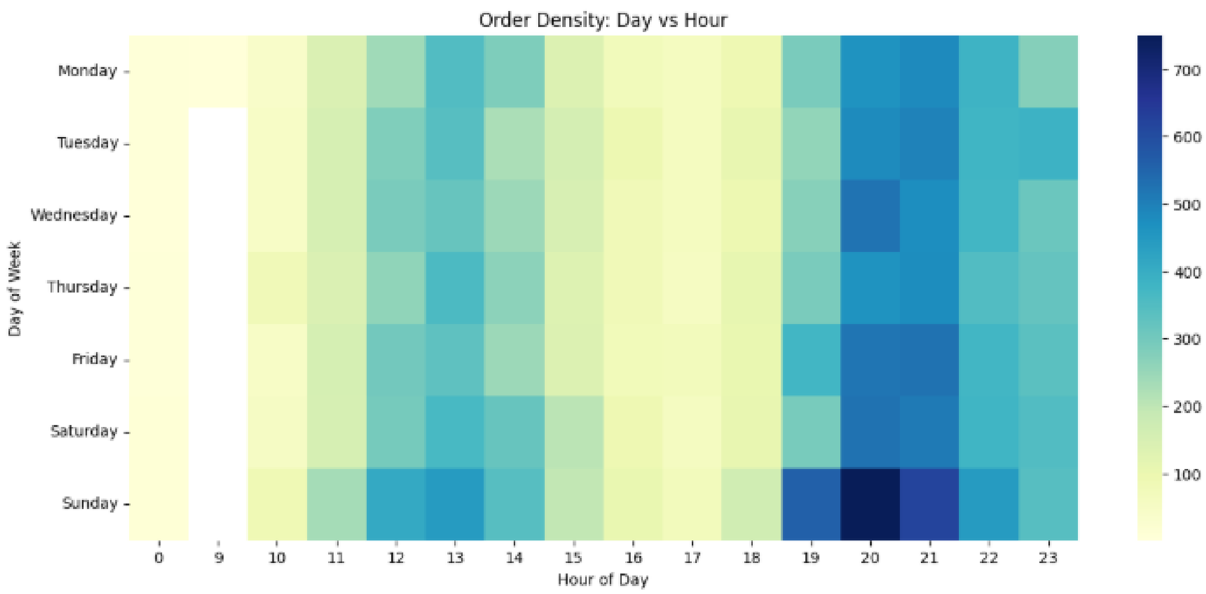

Distribution of Area
count
0
1000
2000
3000
4000
5000
6000
7000
8000
Zomato
Swiggy_Thali King
Parcel
Swiggy_Taste Of Thali
Zomato_Taste Of Thali
Swiggy_Wow Chinese -
Speedyy
Home Website
Zomato_Wow Chinese
Swiggy_Biryani Bronze
Area

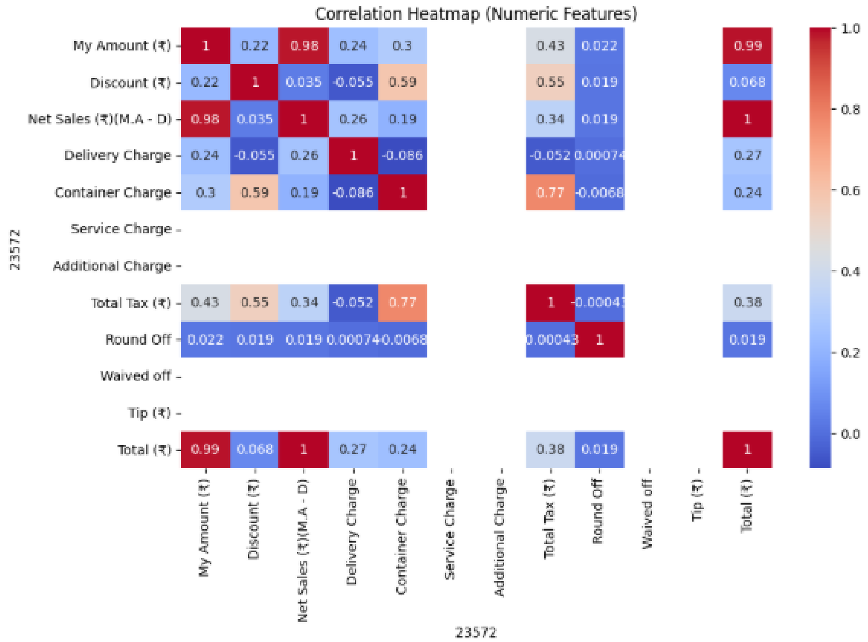

Distribution of Order Type
count
0
2500
5000
7500
10000
12500
15000
17500
20000
Delivery(Parcel)
Pick Up

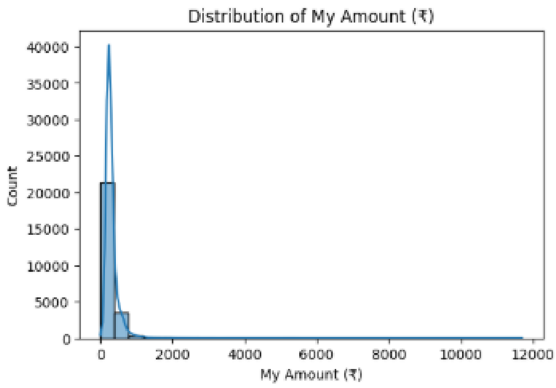

Distribution of Sub Order Type
count
0
1000
2000
3000
4000
5000
6000
7000
8000
Thali King - Zomato
Thali King - Swiggy
Pick Up
Taste Of Thali - Swiggy
Taste Of Thali - Zomato
Wow Chinese - - Swiggy
Thali King - Speedyy
Home Website
Wow Chinese - Zomato
Biryani Bronze - Swiggy
Sub Order Type

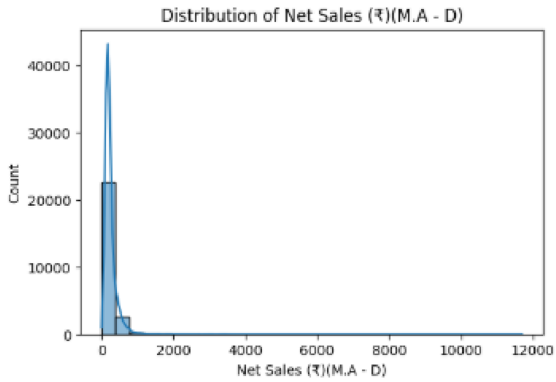

Distribution of Brand Name
count
0.04
0.02
0.00
−0.02
−0.04
0
1
2
3
4
5
6
Brand Name

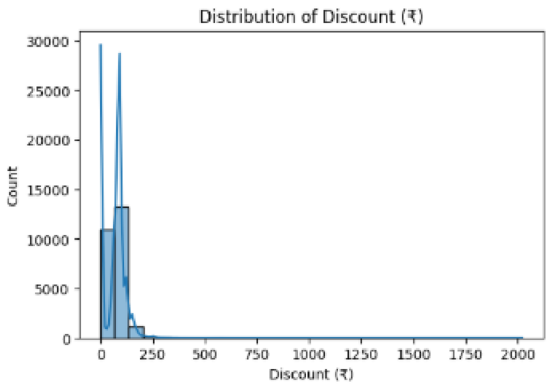

Distribution of Status
count
0
5000
10000
15000
20000
25000
Success
Status

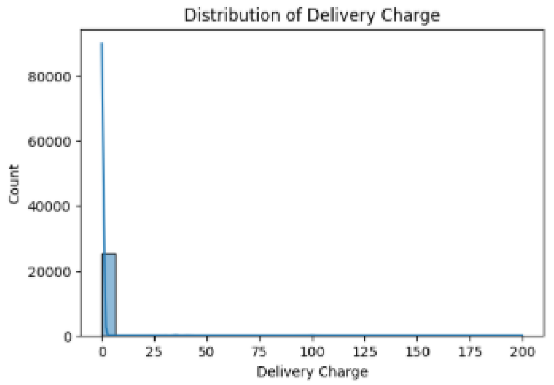

Order Density: Day vs Hour
Day of Week
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
0
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
Hour of Day
100
200
300
400
500
600
700

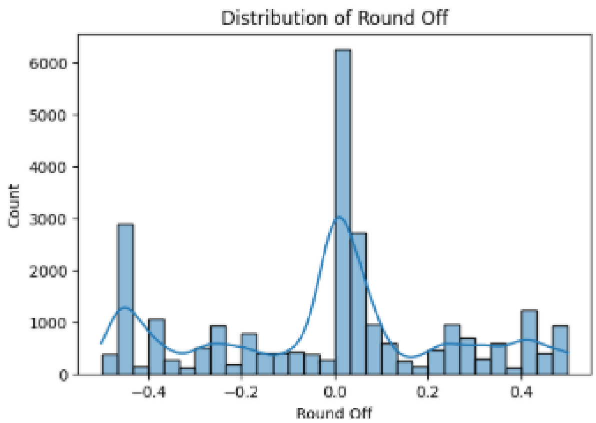

Correlation Heatmap (Numeric Features)
My Amount (₹)
Discount (₹)
Net Sales (₹)(M.A - D)
Delivery Charge
Container Charge
Service Charge
Additional Charge
Total Tax (₹)
Round Off
Waived off
Tip (₹)
Total (₹)
23572

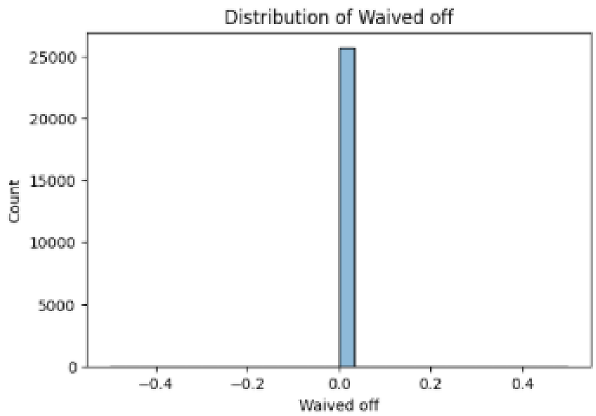

Distribution of My Amount (₹)
Count
My Amount (₹)

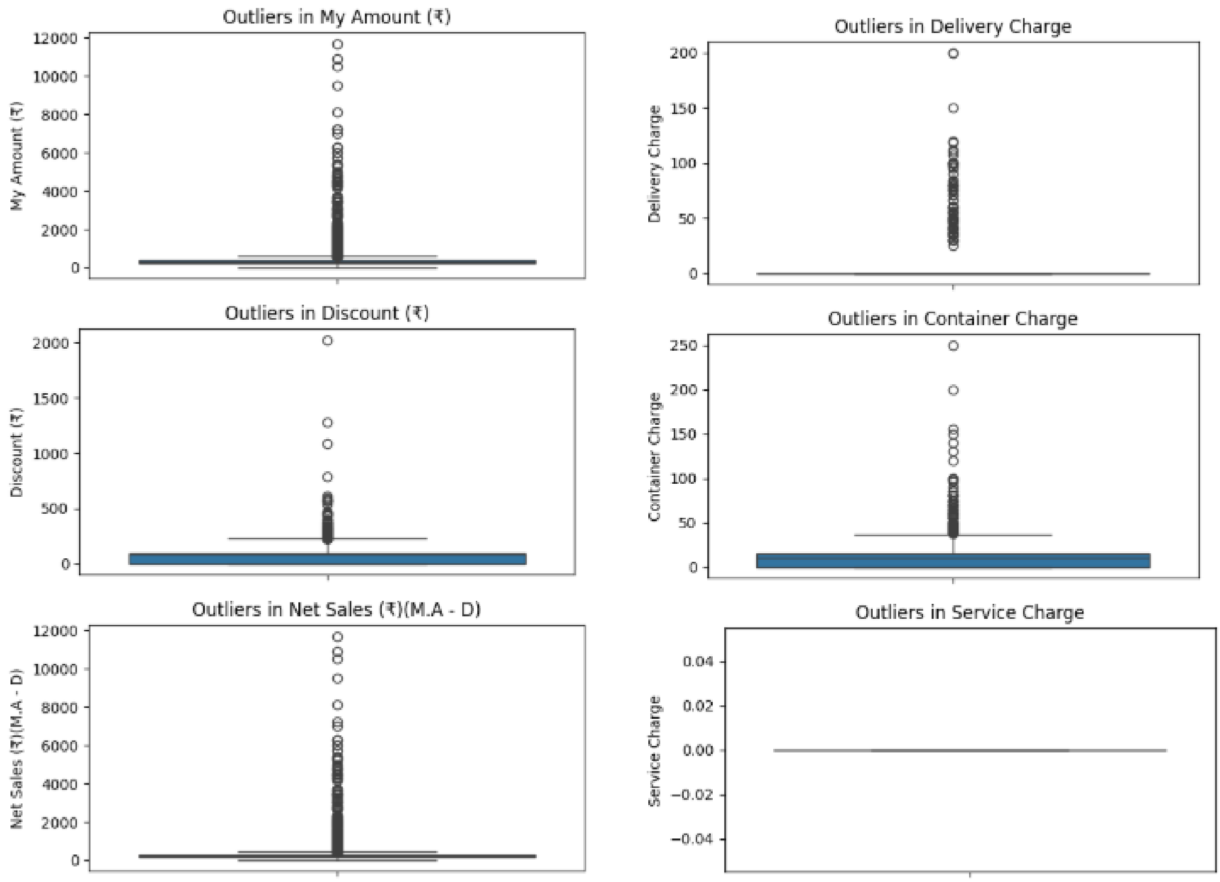

Distribution of Net Sales (₹)(M.A - D)
Count
Net Sales (₹)(M.A - D)

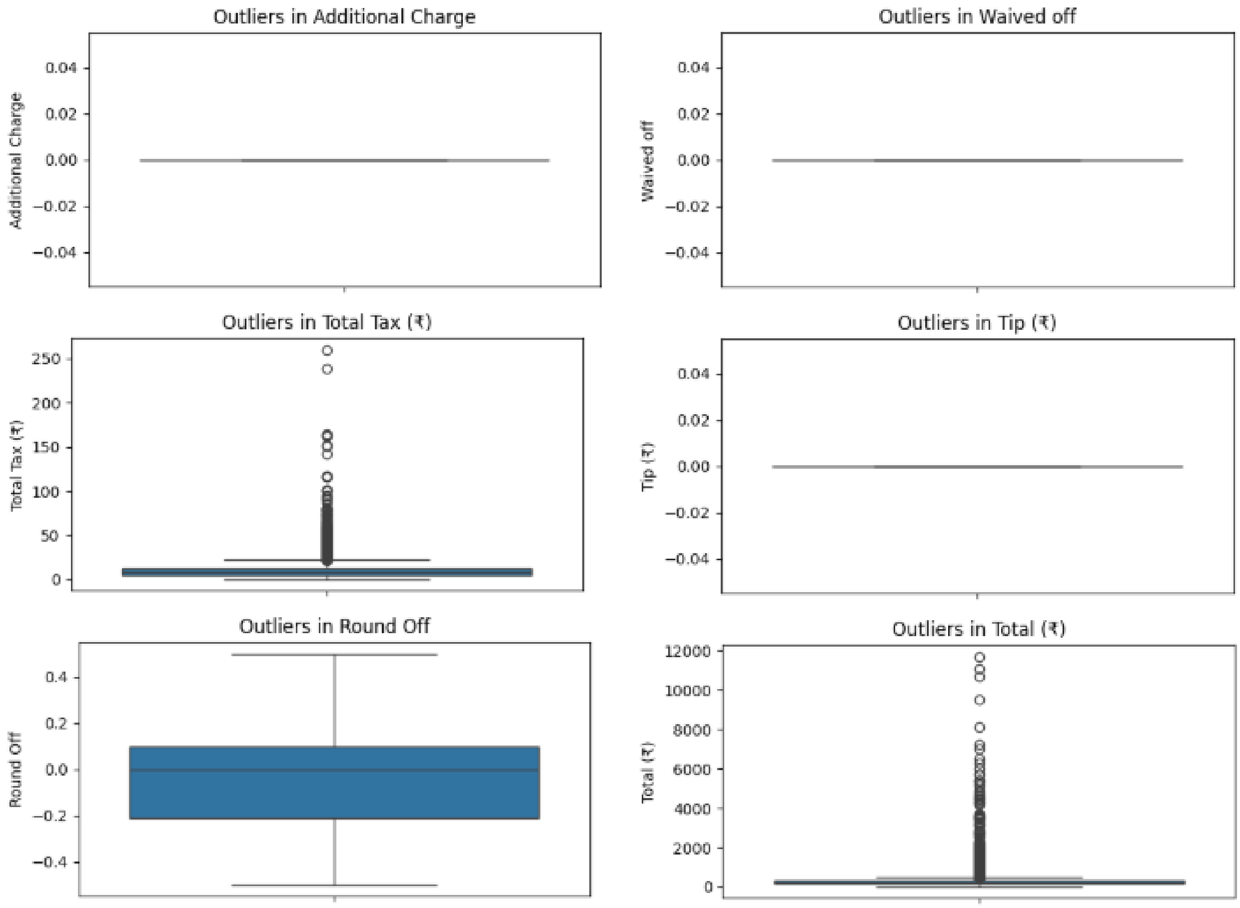

Distribution of Discount (₹)
Count
Discount (₹)


Distribution of Delivery Charge
Count
Delivery Charge


Distribution of Round Off
Count
Round Off


Distribution of Waived off
Count
Waived off


Outliers in My Amount (₹)
My Amount (₹)
Outliers in Delivery Charge
Delivery Charge
Outliers in Discount (₹)
Discount (₹)
Outliers in Container Charge
Container Charge
Outliers in Net Sales (₹)(M.A - D)
Net Sales (₹)(M.A - D)
Outliers in Service Charge
Service Charge


Outliers in Additional Charge
Additional Charge
Outliers in Waived off
Waived off
Outliers in Total Tax (₹)
Total Tax (₹)
Outliers in Tip (₹)
Tip (₹)
Outliers in Round Off
Round Off
Outliers in Total (₹)
Total (₹)

Model Results:

- Linear Regression predicted daily sales with good accuracy.
- Streamlit UI made prediction user-friendly.

```
Linear Regression Results
R²: 1.0
MAE: 0.0
RMSE: 0.0
```

```
Random Forest Regressor Results
R²: 0.983
MAE: 2.006
RMSE: 42.265
```

```
Random Forest Classifier Accuracy: 1.0

Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      5129

    accuracy                           1.00      5129
   macro avg       1.00      1.00      1.00      5129
weighted avg       1.00      1.00      1.00      5129
```

---

9. CONCLUSION AND FUTURE WORK

Conclusion:

This project demonstrated the development of a complete data-driven system for analyzing and predicting sales in a cloud kitchen setup. The dataset obtained from the kitchen's POS system contained raw and inconsistent information, which required extensive data cleaning, preprocessing, and feature engineering. Through Exploratory Data Analysis, meaningful patterns were identified—such as peak order timings, weekly demand fluctuations, popular order types, and payment preferences. These insights help a cloud kitchen understand customer behavior and improve operational planning.

```python
# GEN AI CHATBOT SECTION
# --------------------------
st.header("AI Assistant")
st.write("Ask me anything about data analysis or predictions!")

user_input = st.text_input("Your Question:")

# Try to get API key safely (works locally + on cloud)
openai_api_key = os.getenv("OPENAI_API_KEY", None)
if not openai_api_key:
    try:
        openai_api_key = st.secrets["OPENAI_API_KEY"]
    except Exception:
        openai_api_key = None
```

A machine learning model was trained on historical sales data to predict daily sales based on selected business factors. The model performed effectively and generated predictions that align with real historical trends. The project not only highlights the importance of structured data processing but also shows how predictive models can help reduce food wastage, optimize staffing, assist inventory planning, and support pricing or marketing decisions.

Overall, the project successfully achieved its objectives—transforming raw data into actionable insights and enabling informed decision-making for improving cloud kitchen performance.

Future Work:

This project can be further extended in several meaningful ways:

1. **Model Enhancement**
   - Experiment with advanced models such as Random Forest, XGBoost, or LSTM to improve prediction accuracy.
2. **Time-Series Forecasting**
   - Implement dedicated forecasting models like ARIMA or Prophet to generate monthly/seasonal demand predictions.
3. **Dashboard Integration**
   - Connect the prediction model to tools like Power BI or Tableau to create interactive real-time dashboards.
4. **Real-Time Data Streaming**
   - Integrate directly with the POS system to update predictions based on live order data.
5. **Multi-Kitchen Scalability**
   - Extend the system to handle multiple locations/outlets and compare their performance.
6. **Inventory Optimization**
   - Add a forecasting module that converts predicted sales into required raw material quantities to reduce wastage.
7. **Mobile Application**
   - Build a lightweight mobile app so managers can track daily performance and predictions on the go.

In summary, with further enhancements, this system can evolve into a complete operational decision-support tool for cloud kitchens.

10. **REFERENCES**

[1] A. Nair, P. Menon, and S. Kumar, "Time-series demand forecasting model for cloud kitchen food delivery," *International Journal of Data Science and Analytics*, vol. 10, no. 3, pp. 145–156, 2023.

[2] R. Rani and N. Kumar, "Sales performance monitoring using Power BI dashboards for restaurant analytics," *International Journal of Computer Applications*, vol. 182, no. 42, pp. 22–29, 2024.

[3] S. Gupta, A. Sinha, and T. Bose, "Customer segmentation using K-Means clustering in food delivery applications," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 2, pp. 330–339, 2023.

[4] A. Singh and R. Verma, "Machine learning-based sales prediction using Random Forest for restaurant chains," in *Proc. 2024 IEEE International Conference on Machine Learning and Data Engineering (ICMLDE)*, Dubai, 2024, pp. 215–222.

[5] L. Martinez and K. Wong, "Forecasting perishable inventory using regression-based predictive analytics," *Journal of Retail Supply Chain Management*, vol. 9, no. 1, pp. 55–65, 2022.

[6] Y. Zhang, Q. Luo, and H. Chen, "LSTM-based inventory optimization and demand forecasting in food retail," *IEEE Access*, vol. 11, pp. 40213–40223, 2023.

[7] P. Sharma and V. Chauhan, "Seasonality-based demand prediction using Prophet for online food delivery," *International Journal of Smart Computing and Artificial Intelligence*, vol. 4, no. 2, pp. 39–47, 2024.

[8] Scikit-Learn Developers, "Scikit-learn: Machine learning in Python," *scikit-learn.org*, 2024. [Online]. Available: https://scikit-learn.org/

[9] Streamlit Inc., "Streamlit — Turn data scripts into shareable web apps," *streamlit.io*, 2024. [Online]. Available: https://streamlit.io/

[10] OpenAI, "OpenAI API documentation," *platform.openai.com*, 2024. [Online]. Available: https://platform.openai.com/docs/