Sales Transactions Dataset for 'A Pizza restaurant

In [ ]:
```
!pip install faker
import pandas as pd
import numpy as np
from faker import Faker
from random import choice, randint, uniform
```

In [8]:
```
fake = Faker()
```

In [12]:
```
num_days = 365  # Number of days for data rane
daily_customers_range = (50, 200)  # Minimum and maximum number of customers per

# List of menu items with ID and prices
menu_items = [
    {"Item_ID": 1, "Item_Name": "Margherita", "Price": 32},
    {"Item_ID": 2, "Item_Name": "Diavola", "Price": 35},
    {"Item_ID": 3, "Item_Name": "Vegana", "Price": 36},
    {"Item_ID": 4, "Item_Name": "Amatriciana", "Price": 37},
    {"Item_ID": 5, "Item_Name": "Carbonara", "Price": 38},
    {"Item_ID": 6, "Item_Name": "Bolognese", "Price": 39},
    {"Item_ID": 7, "Item_Name": "Focaccia", "Price": 15},
    {"Item_ID": 8, "Item_Name": "Coca-Cola", "Price": 10},
    {"Item_ID": 9, "Item_Name": "Tiramisu", "Price": 18},
    {"Item_ID": 10, "Item_Name": "Espresso", "Price": 12},
    {"Item_ID": 11, "Item_Name": "Water", "Price": 5}

]
# Operating hours of the pizzeria (15:00 - 21:00)
operating_hours = (15, 21)
```

In [14]:
```
# Random time within the operating hours
def random_time(start_hour, end_hour):
    hour = randint(start_hour, end_hour - 1)
    minute = randint(0, 59)
    second = randint(0, 59)
    return f"{hour:02d}:{minute:02d}:{second:02d}"
```

Dataset Description

The dataset contains the following columns:

- **Transaction_ID**: A unique identifier for each transaction.
- **Date**: The date of the transaction.
- **Time**: The time of the transaction.
- **Item_ID**: A unique identifier for the menu item.
- **Item_Name**: The name of the menu item sold.
- **Quantity**: The quantity of the item sold in the transaction.
- **Price**: The total price for the items sold.
- **Payment_Method**: The method of payment used.

In [16]:
```
data = []
transaction_id = 1

# Loop through each day of the year to generate daily transactions
for day in pd.date_range("2023-01-01", "2023-12-31"):
    num_customers = randint(*daily_customers_range)
    # Generate transactions for each customer
    for _ in range(num_customers):
        transaction_time = random_time(*operating_hours)
        item = choice(menu_items)
        quantity = randint(1, 3)
        total_price = item["Price"] * quantity

        # Append the transaction data to the list
        data.append({
            "Transaction_ID": transaction_id,
            "Date": day.strftime("%Y-%m-%d"),
            "Time": transaction_time,
            "Item_ID": item["Item_ID"],
            "Item_Name": item["Item_Name"],
            "Quantity": quantity,
            "Price": total_price,
            "Payment_Method": choice(["Cash", "Credit Card", "Mobile Payment"])
        })

        transaction_id += 1 # Increment transaction ID for the next transaction


sales_data = pd.DataFrame(data)
```

In [18]:
```
sales_data.to_csv("sales_transactions.csv", index=False)
```

Out[18]:

	Transaction_ID	Date	Time	Item_ID	Item_Name	Quantity	Price	Payment_Meth
0	1	2023-01-01	18:28:53	1	Margherita	1	32	Mobile Paym
1	2	2023-01-01	17:34:20	9	Tiramisu	1	18	Credit C
2	3	2023-01-01	19:48:57	8	Coca-Cola	2	20	C
3	4	2023-01-01	16:43:30	2	Diavola	3	105	Mobile Paym
4	5	2023-01-01	17:23:35	7	Focaccia	1	15	Credit C