

Restaurant Promotional Event

Sean Youngseon Lee
Jeff Youngjae Lee
March 25, 2024

Table of Contents

01

About Us

02

Introduction

03

Data Analysis

04

Results

05

Possible
Confounding
Variable

06

Reflection



01

About Us

UC Irvine - B.S. Data Science



Sean Lee

(949) - 390 - 4811



TensorFlow Certified Developer

[LinkedIn.com/seanlee319](#)

[Github.com/seanlee319](#)

Experience

- **Dr. Yoonjin Won's Thermodynamics Lab (UCI)** - Research Assistant
 - Engaged in the **Hyundai Autonomous Vehicle Project** to analyze degradation in cameras under extreme conditions
 - Analyzed thousands of images through **MATLAB's Statistics and Machine Learning Toolbox**
 - Created pixel histograms to quantify noises in images after data extraction

Projects

- **Energy Usage Prediction in California** - Accenture
 - Led a project sponsored by Accenture to **predict the monthly energy consumption** of California
 - Utilized seasonality of average gas price by state to predict the future energy consumption of California
 - Data wrangling, data processing, SARIMAX, LSTM, Decision Tree, Linear Regression
 - Submitted a paper to "**Energy Exploration & Exploitation**" which is currently under review



Jeff Lee

(949) - 390 - 4809



TensorFlow Certified Developer

JeffTheScientist.vercel.app

LinkedIn.com/jeffyoungjaelee

Github.com/leeff0319

Vanderbilt University - B.A. Chemistry

Research Assistant @ Boston Children's Hospital

Experience

- **Korbato-Health**

- Currently working under supervision of **Dr. Ryuichiro Yagi** and **Dr. Shinichi Goto** to develop ML models to analyze ECG signals to detect rare heart diseases such as Hypertrophic Cardiomyopathy

Projects

- **Jeff The Scientist** | JeffTheScientist.vercel.app

- Developed a personal portfolio website from scratch using **HTML**, **TypeScript**, **Tailwind CSS**, **NextJS**

- **Interactive Digit Recognizer** |

- Utilized **TensorFlowJS** and **MNIST** dataset to create an interactive webpage on which one can draw digits for the ML model to read

Click Me

02

Introduction

Our Motivations

We needed a Data project for Portfolio:

- Opportunity arose at Sean's workplace
 - Owner was family friend
 - We wished to contribute while refining our **expertise in data analytics/science**
 - Pitched the idea of a **Promotional Event**:
 - Noodle orders come with a free beverage of choice (**7/16 ~ 8/31**)



Goals

Goal 1

Increase **noodle sales** by **50%** during event period



Goal 2

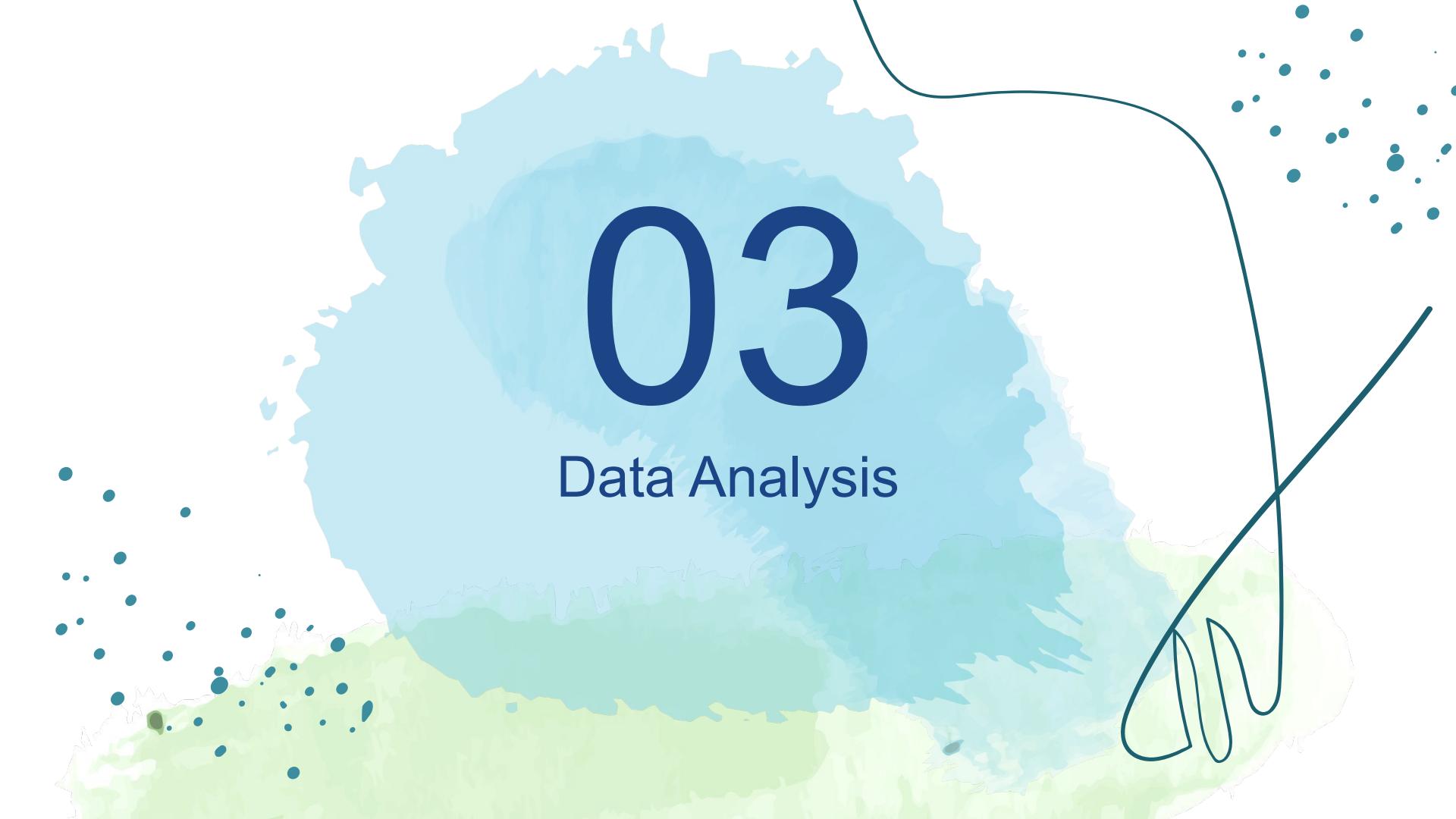
Increase **net sales** by **15%** during event period

Goal 3

Retain **noodle sales** after the event period

Timeline





03

Data Analysis

Dataset

- Granted permission to use sales data
 - But only generalized data, not the comprehensive dataset
 - Privacy/security reasons

Dates	1/18/24	1/19/24	1/20/24	1/21/24	1/22/24	1/23/24	1/24/24	1/25/24	1/26/24
Sales	5	6	7	1	2	3	4	5	6
Gross Sales	1498.23	1977.98	3050.42	3523.89	1978.52	1430.29	1948.65	1596.85	2484.78
Refunds	0	0	0	-15.99	0	0	0	0	0
Discounts & Comps	0	0	0	0	0	0	0	0	0
Net Sales	1498.23	1977.98	3050.42	3507.90	1978.52	1430.29	1948.65	1596.85	2484.78
Tax	116.14	153.31	236.41	271.86	153.34	110.85	151.02	123.76	192.59
Tips	83.39	94.44	149.04	167.57	95.28	67.08	76.98	68.56	91.68
Total	1697.76	2225.73	3435.87	3947.33	2227.14	1608.22	2176.65	1789.17	2769.05
Category Sales									
BBQ	4	7	2	13	6	3	7	0	14
Combo	10	12	21	24	10	8	15	16	12
Drink	10	12	26	40	13	15	23	19	22
Special Menu	16	26	47	52	28	25	25	26	41
Tofu Soup	40	57	91	105	60	33	50	54	55
Noodles	7	3	6	5	4	3	8	5	4
% Noodles	8.41	2.73	3.54	2.56	3.64	3.77	7.39	5.63	2.90

Fig 1. Snapshot of the data used for analysis

Analysis

Data Preprocessing

Approach

- Analyzed/organized dataset using **Excel & SQL**
- Preprocessed data using **Python**

```
# Open CSV file
with open('dataset\\Restaurant_Sales_Data-Cleaned_Data.csv', newline='') as csvfile:
    # Create CSV reader object
    csv_reader = csv.reader(csvfile, delimiter=',')

    # Skip header row
    next(csv_reader)

    # Initialize empty list
    rows = []

    # Iterate over each row and append data to list
    for row in csv_reader:
        rows.append(row)

    # Create a DataFrame from the list of rows
    df = pd.DataFrame(rows, columns=["Dates", "Net Sales", "% Noodles", "Weekdays"])

    # Print the DataFrame
    # print(df)

# Convert columns to correct format
df['Dates'] = pd.to_datetime(df['Dates'])
df['Net Sales'] = df['Net Sales'].astype(float)
df['% Noodles'] = df['% Noodles'].astype(float)

# Filter data for Event_df and nonEvent_df
Event_df = df[(df['Dates'] >= '2023-07-16') & (df['Dates'] <= '2023-08-31')]
nonEvent_df = df[(df['Dates'] < '2023-07-16') | (df['Dates'] > '2023-08-31')]

# Print the filtered DataFrame
#print(nonEvent_df)
#print(Event_df)
```

Duration

- Data Collection- 8 mo**
- Data Analysis- 2 days**

Stakeholders

- Restaurant Owner
- Manager

04

Results

Results

- Business Impact
 - **81.13% increase in Noodle Sales**
 - **7.18% increase in Net Sales**
 - Significant increase in Noodle Sales had **little to no effect on Net Sales**
 - Customers ordered noodles instead of tofu or other alternative entrees, which are similar prices

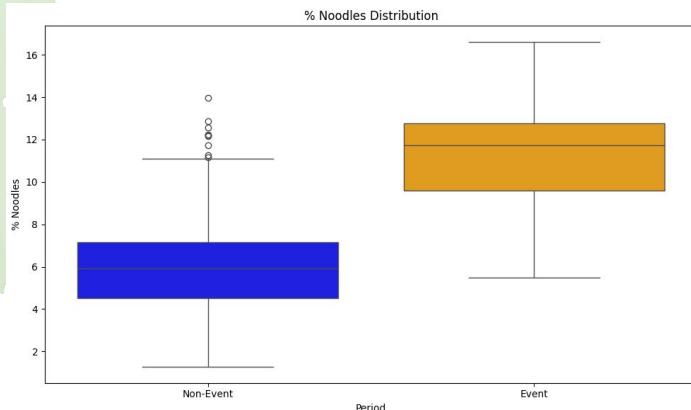


Fig 2. Bar graph of % Noodles during Non-event period vs. Event period

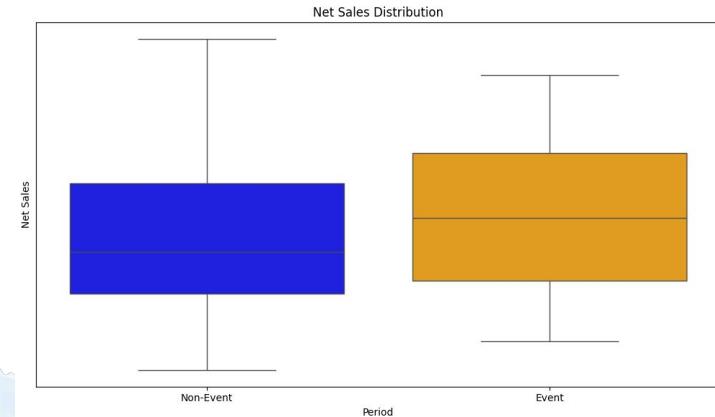


Fig 3. Bar graph of Net Sales (in \$) during Non-event period vs. Event period. **Actual numbers are omitted for privacy reasons**

Results

- Little to **no visual difference** in **Net Sales** during the event period
- Noticeable **increase** in **% Noodles** during the event period

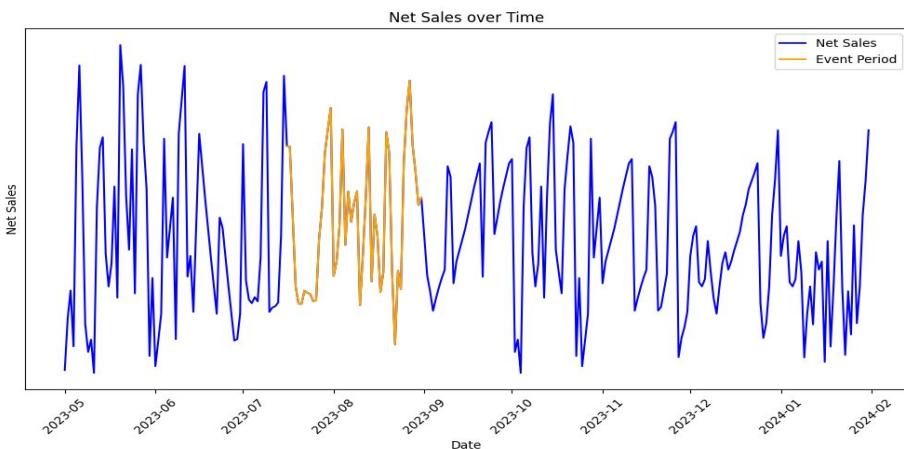


Fig 4. Graph of net sales (in \$) over recorded time period. Highlighted region represents period of promotional event. **Actual numbers are omitted for privacy reasons.**

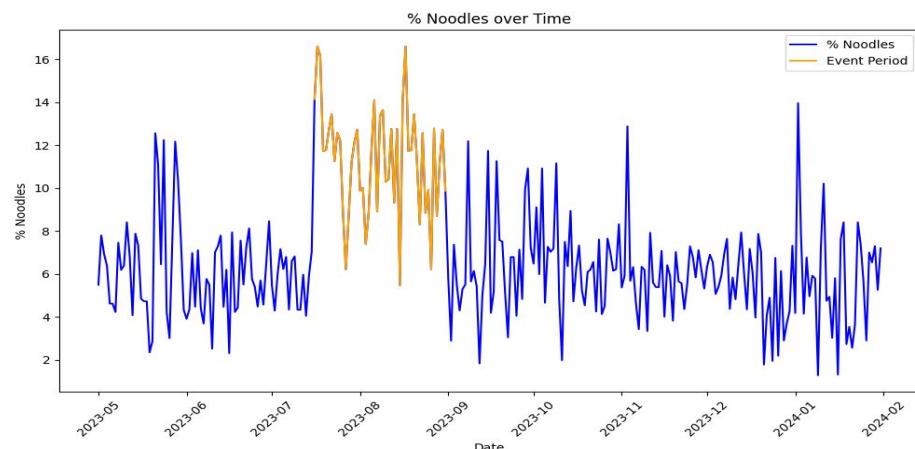


Fig 5. Graph represents the % of net sales that are composed of noodles.

Results

- More number of noodles sold during promotion **on average**
 - Shows a **similar trend** with **% Noodles over Time** graph

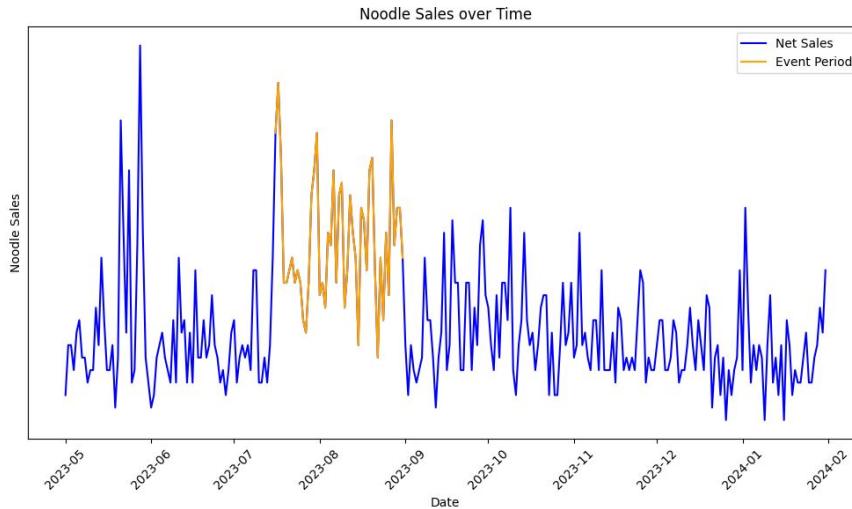


Fig 6. Graph represents the total numbers of noodles sold each day over recorded time period.

Results

- Little to no difference in % Noodle distribution and Noodle Sales after the event when compared to before the event
 - Outliers present

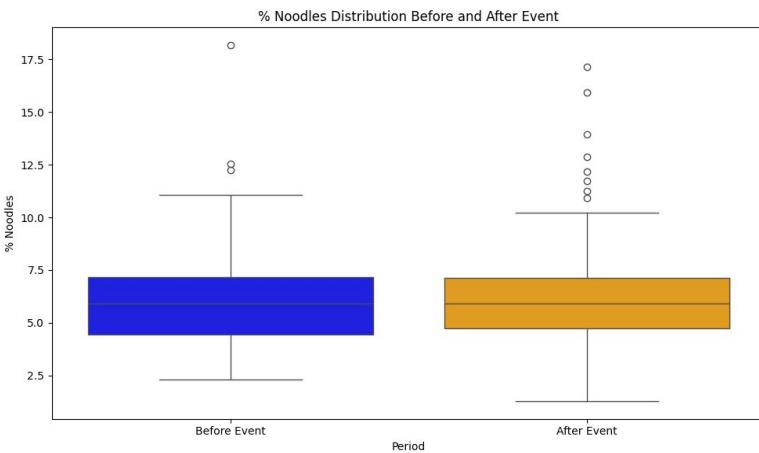


Fig 7. Bar graph of % of net sales composed by noodles before and after the promotional event. Used to evaluate noodle sale retention, or long-lasting effects of promotional event.

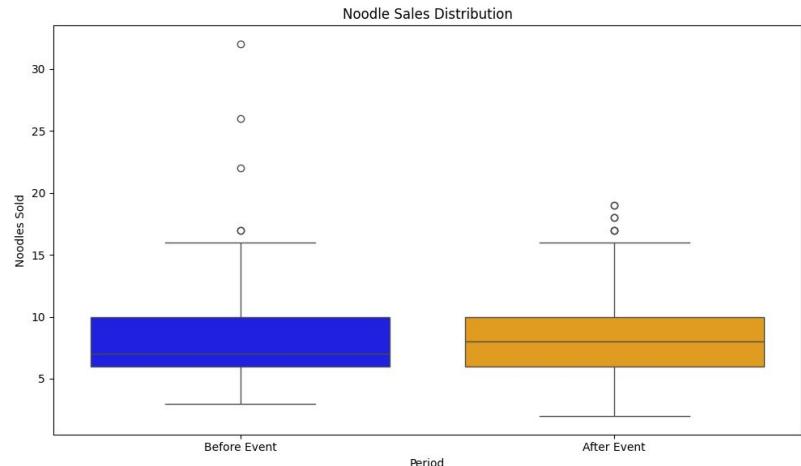


Fig 8. Bar graph shows average number of noodle sales per day before and after event. Also used to evaluate noodle sale retention, or long-lasting effects of promotional event.



05

Possible Confounding Variable

Possible Confounding Variable

- Considered **confounding variables** that could have affected the results
 - Such as **temperature**

DATE MM DD YYYY	OBS TIME HH:MM	PRECIP AMOUNT & TYPE (IN)	AIR TEMPERATURE			WIND DD SS	ETo	SOL RAD	SOIL T MAX MIN	WX	RELATIVE HUMIDITY MAX MIN	BULB TEMP WET (F)	TEMP DRY (IN)	PAN EVAP
			*	(F)	*				*					
05-01-2023	23:59		66	55					66	65				
05-02-2023	23:59		64	51					65	64				
05-03-2023	23:59		65	51					64	63				
05-04-2023	23:59		64	52					64	62				
05-05-2023	23:59		65	49					64	62				
05-06-2023	23:59		67	47					64	62				
05-07-2023	23:59		68	49					64	62				
05-08-2023	23:59		70	48					65	62				
05-09-2023	23:59		69	51					65	63				
05-10-2023	23:59		68	53					65	64				
05-11-2023	23:59		72	54 A					72	68 A				
05-12-2023	23:59		71	54					65	64				
05-13-2023	23:59		74	57					66	64				
05-14-2023	23:59		76	58					67	65				
05-15-2023	23:59		70	57					67	66				
05-16-2023	23:59		74	56 A					68	66				
05-17-2023	23:59		73	57					68	67				
05-18-2023	23:59		70	58					68	67				

Fig 9. Public dataset of temperatures in the area from <https://ipm.ucanr.edu/>.

Possible Confounding Variable

- **Higher average max air temperature resulted in higher % noodle sales**
 - Difficult to state the increase in the % noodle sales were solely due to the promotion

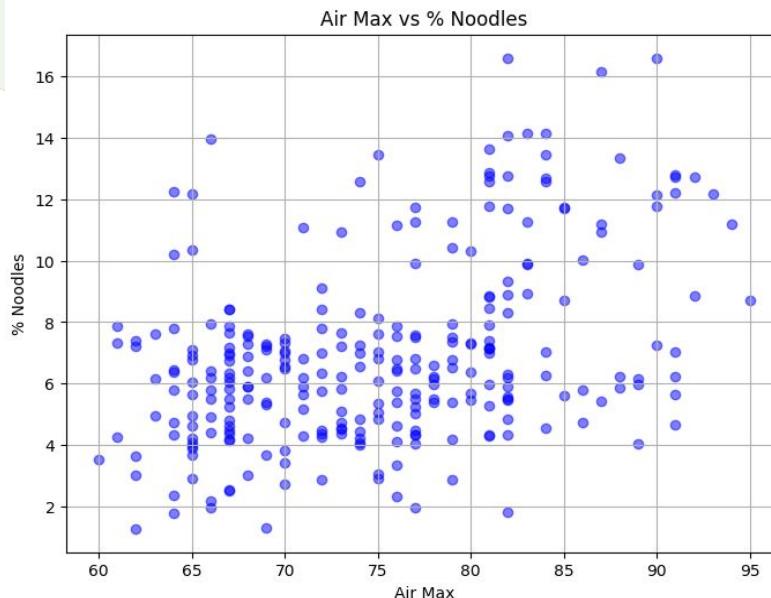


Fig 10. Scatter plot of the highest temperature measured at 1.5 meters vs. % net sales composed of noodles.

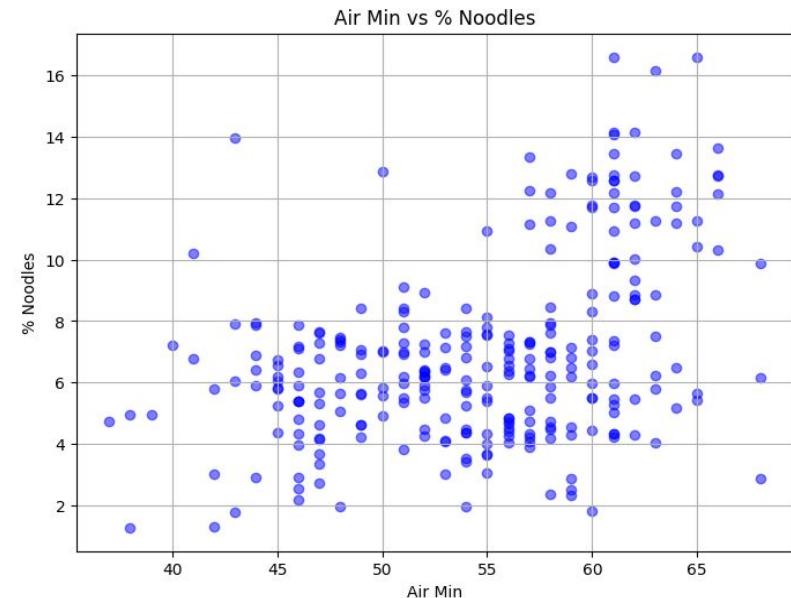


Fig 11. Scatter plot of the lowest temperature measured at 1.5 meters vs. % net sales composed of noodles.

Possible Confounding Variable

- Higher average max air temperature resulted in higher % noodle sales
 - Difficult to state the increase in the % noodle sales were solely due to the promotion

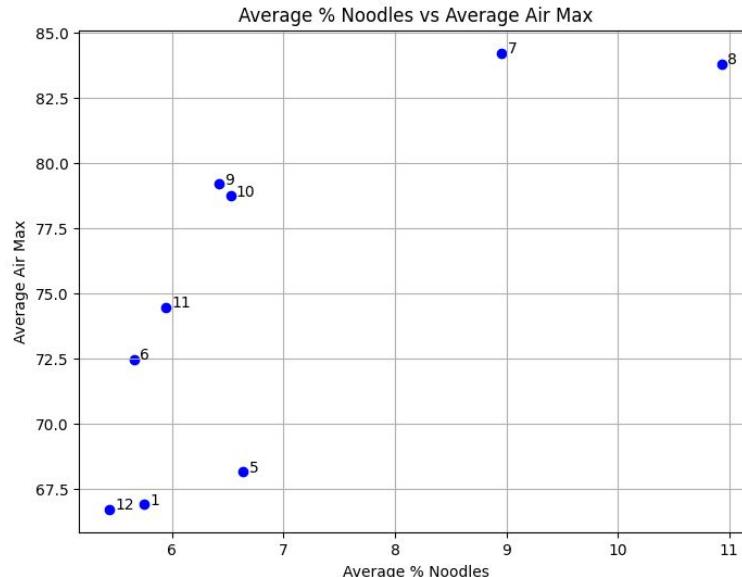


Fig 12. Scatter plot of the average highest temperature measured at 1.5 meters for each month vs. % net sales composed of noodles.



06

Reflection

Goals

Goal 1

Increase **noodle sales** by **50%** during event period



Goal 2

Increase **net sales** by **15%** during event period

Goal 3

Retain **noodle sales** after the event period

Goals Reflection

Goal 1 ✓

Increased **noodle sales** by **81.13%** during event period (Figs. 2, 5, 6)



X Goal 2

Increased **net sales** by **7.18%** during event period (Figs. 3 & 4)

X Goal 3

Unsuccessful retention of **noodle sales** after the event period (Figs. 7 & 8)

Reflection



Strengths

- **Proactive & Initiative**
 - Always looking to apply data in everyday life
- **Noticeable increase** of noodle sales during event
 - Percent Noodles over time graph (Figure)



Areas of Improvement

- **More & diverse data**
 - From different years, seasons, etc.
- **Differentiate items** within item category
- **More detailed visualization** using Tableau
 - We used Seaborn
- **Include number of orders/day**
 - Compare number of customers on days of event
 - **Customer Retention** (Returning & New)
 - Not provided for us on this dataset



Weaknesses

- **Inconclusive Results**
 - Within error bars
- **Dataset Limitation**
 - Incomplete dataset (only 8 months)
 - Confounding Variables
- **Design Flaw**



Feedback from Stakeholders

- **Insightful analysis**
 - despite limitations of given dataset
- **Positive effect** in noodle sales



THANK YOU

Do you have any questions?