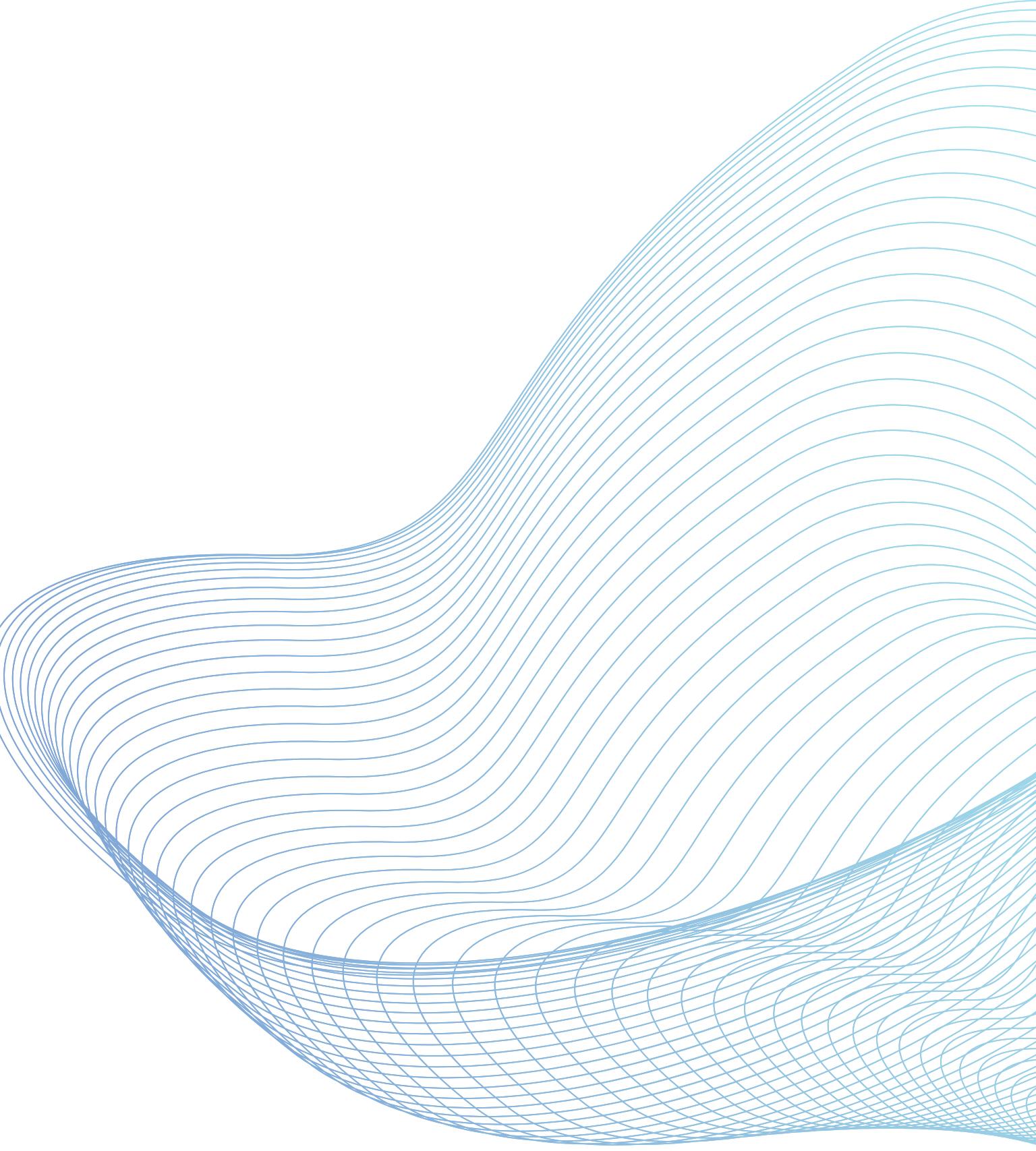


# ASSIGNMENT PRESENTATION

Mammad Aliyev - applicant  
[mammed.aliyev.e@gmail.com](mailto:mammed.aliyev.e@gmail.com)



# SECTION 1

In this section there were 8 sub-tasks to complete:

6 calculation-related questions to be answered  
2 analytical question to be answered: one is answered by using plot, the other is answer by using the analytics of the data from previous questions

Techs

Answers

Most answers we got from the Python application which were developed using such technologies(libraries) like: Numpy, Matplotlib and Pandas

7 out of 8 answers were displayed using the console of the application and 1 is answered here in one of the next slides

# QUESTION 1

1. How many unique restaurants could be found in this data set?  
(Hint: Use the [Business\_ID] column for this evaluation.)

Answer: The Number of unique restaurants: 30276

## How everything was calculated?

[unique\\_restaurants = data\['Business\\_ID'\].nunique\(\)](#)

By using the `.nunique()` function from Pandas library we can filter out the unique rows by a specific column which is 'Business\_ID' in our case

# QUESTION 2

2. Which restaurant received the highest number of reviews?

What about percentage-wise?

Answer:

Restaurant with the highest number of reviews: Mon Ami Gabi

(Business\_ID: 4bEjOyTaDG24SY5TxsaUNQ) with 856 reviews.

It accounts for 77.64% of the total reviews.

## How everything was calculated?

First, we group by Business\_ID and count reviews(review\_counts). Then, we find the ID with the most reviews(most\_reviewed\_id), get its review count(most\_reviewed\_count), and get its name from Business\_Name. Finally, we calculate its percentage of total reviews.

# QUESTION 3

3. Which cities have got at least one 5-star review in Nevada (NV) state?

Answer: The cities with at least one 5-star review: ['Las Vegas'  
'Henderson' 'Boulder City' 'Nellis']

## How everything was calculated?

First, we filter the data to include only rows where the State is 'NV' and the Avg\_Business\_Star\_Rating is 5 or higher, storing it in 'nv5startreviews'. Then, we extract the unique city names from the City column to avoid multiple city names in our result.

# QUESTION 4

4. Which city has the highest number of reviews in the Business Category of "Hotels & Travel"? What about percentage-wise?

Answer: The city with the highest number of reviews in 'Hotels & Travel' is 'Las Vegas' with 10245 reviews.

It accounts for 77.64% of the total reviews in this category.

## How everything was calculated?

First, we filter for "Hotels & Travel," count reviews per city, find the city with the most reviews (`r_city`) and its number, and then we calculate its review percentage.

# QUESTION 5

5. At what day of the week people are more likely to post their reviews?

Answer: People are more likely to post their reviews on Monday which stands for 36446 reviews.

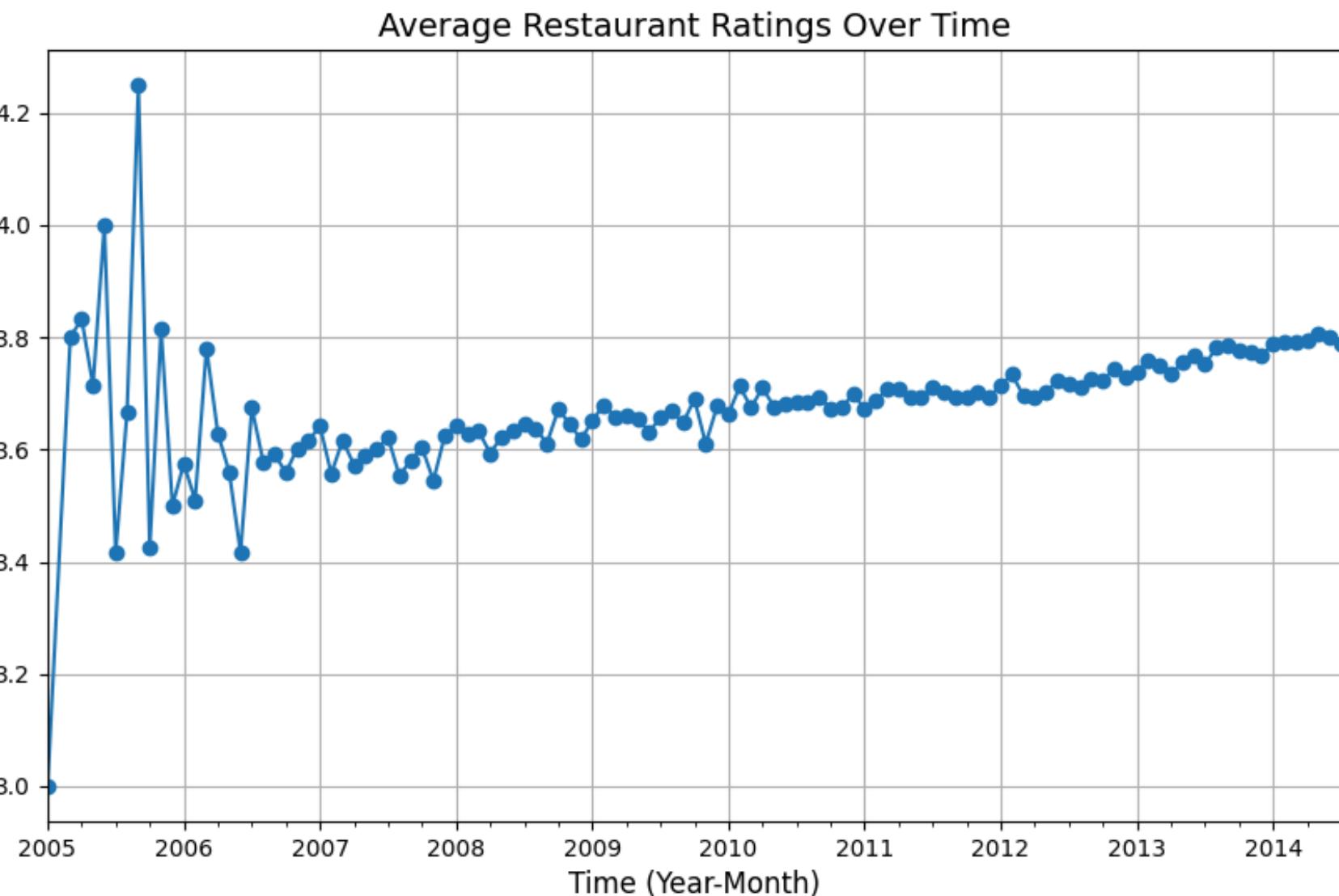
## How everything was calculated?

First, we convert Review\_Date to datetime format, extract the day of the week into a new created column, and count reviews for each day. Then, we find the day with the most reviews (highest\_day) and its number (most\_reviews).

# QUESTION 6

6. Showcase if there are any trends regarding restaurant performance as time goes by.

As shown in the plot, average restaurant ratings fluctuate initially but gradually improve over time, rising from around 3.5 in 2006 to 3.8 by 2014, showing better results over time



# QUESTION 7

Based on analyzed data showcase if there are any steps that the restaurant can take to improve their public appeal.

1

## **Focus on High-Review Days:**

Since Mondays receive the highest number of reviews, restaurants can do special offers, events and employ the best employees on this day to spread better overall opinion.

2

## **Target 5-Star Reviews:**

Studying 5-star reviews in Nevada can show what works well, like good food, atmosphere, and service, to apply in other places.

3

## **Cater to Top-Performing Locations:**

Restaurants in review-heavy cities like Las Vegas should focus on quality service and marketing to boost satisfaction.

# QUESTION 8

8. Based on this data set which user had the highest cumulative travel distance? What distance has been covered by him/her?

Answer:

User with ID: 6uYJ-ixRxPMyf-iEbhoz2g covered the highest distance of: 31766.80 km.

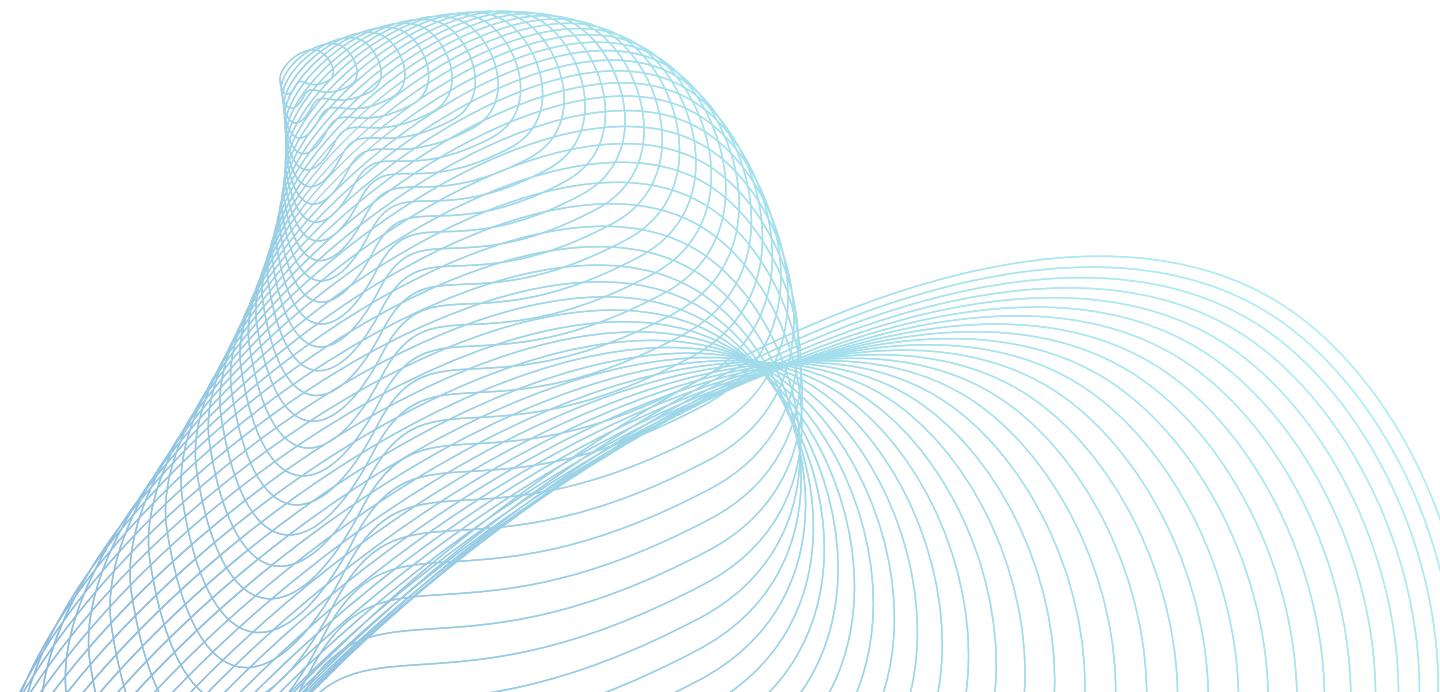
## How everything was calculated?

We decided to sort the data by User\_ID and Review\_Date, then calculated distances between consecutive locations for each user using the Haversine function. The total distances were stored in a DataFrame, and we identified the user with the highest cumulative distance.

# SECTION 2

## Part 1:

Create an algorithm that generates an NxN matrix filled with the numbers in a specific pattern starting from the center. This matrix should follow the pattern illustrated in the attached image.



We decided to construct a NxN spiral array, starting in the middle and going in the predetermined directions (up, left, down, and right), switching directions after a predetermined number of steps were completed, and increasing the number of steps every two turns.

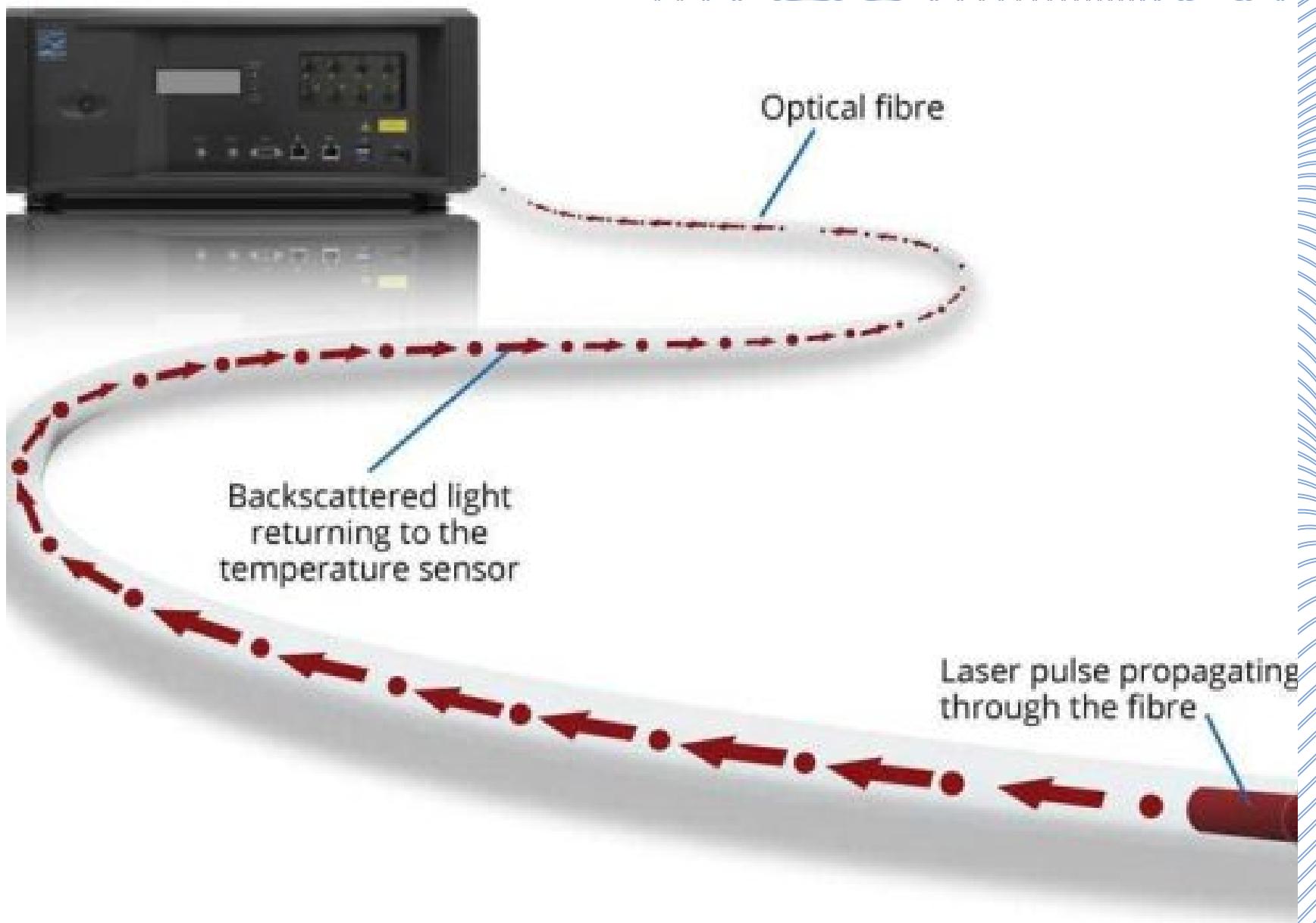
In order to add the values of the primary and secondary diagonal elements, we iterated over the array for diagonal sums.

```
Enter the N(NxN array) : 7
[43, 44, 45, 46, 47, 48, 49]
[42, 21, 22, 23, 24, 25, 26]
[41, 20, 7, 8, 9, 10, 27]
[40, 19, 6, 1, 2, 11, 28]
[39, 18, 5, 4, 3, 12, 29]
[38, 17, 16, 15, 14, 13, 30]
[37, 36, 35, 34, 33, 32, 31]
Primary diagonal elements sum: 119
Secondary diagonal elements sum: 143
```

# SECTION 3

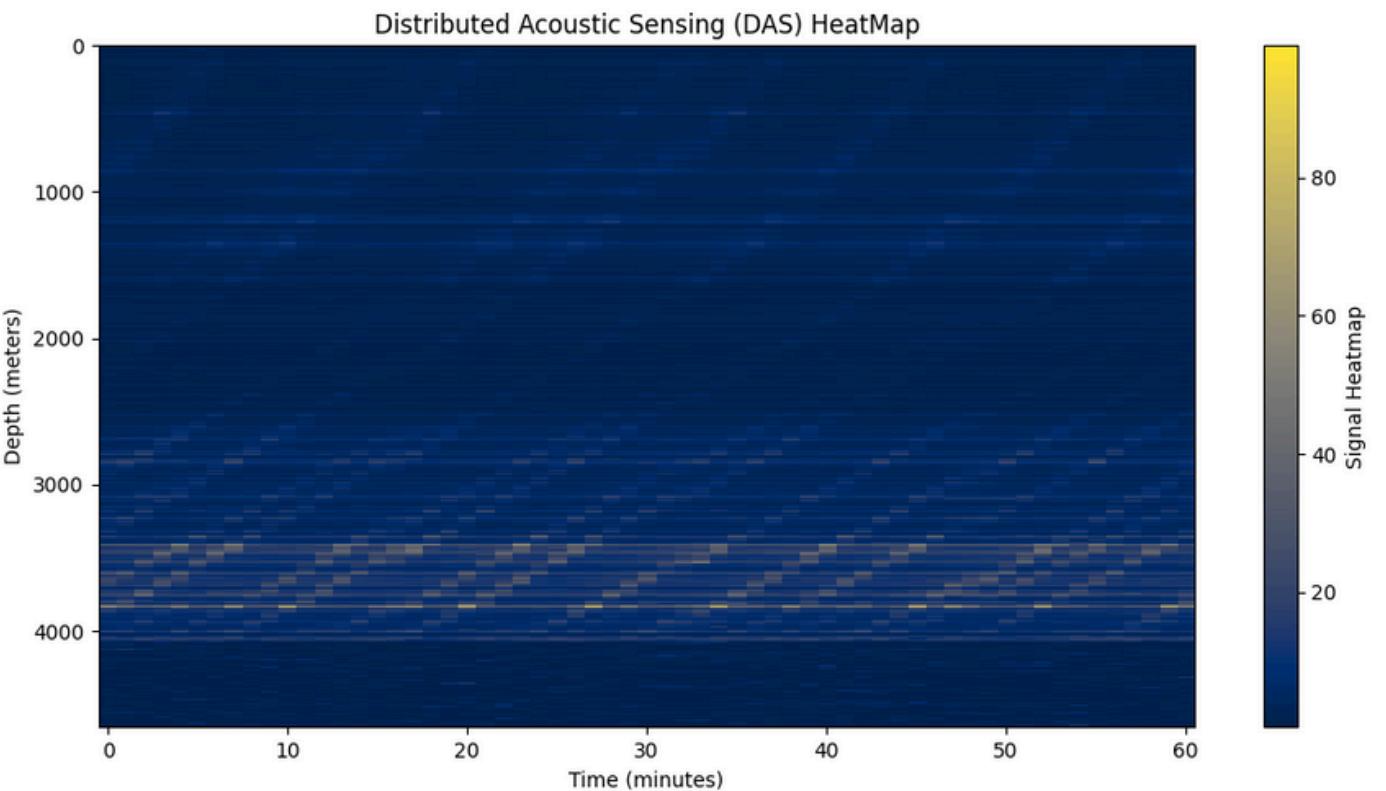
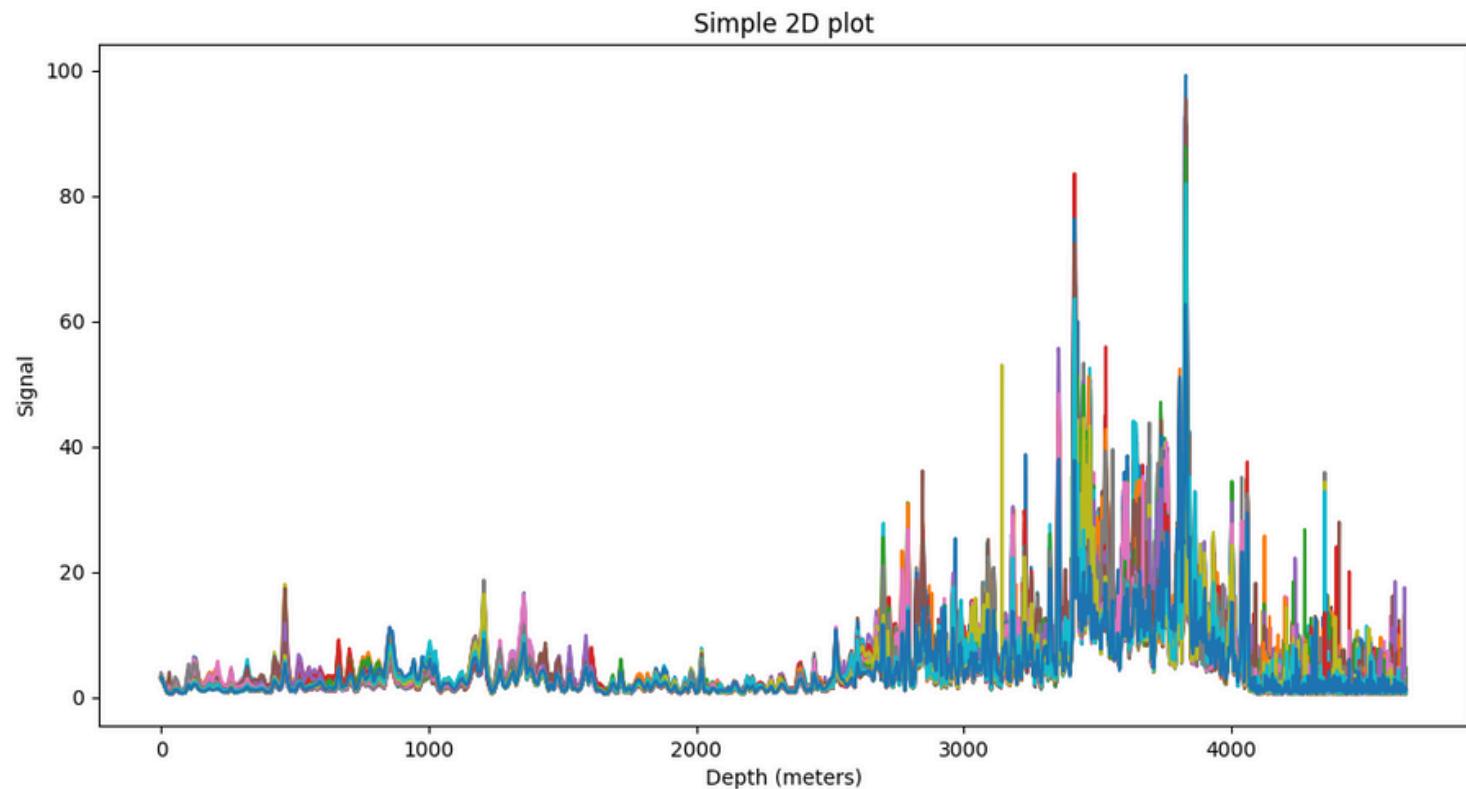
**Brief description(By my own words, very short):**

**Distributed Acoustic Sensing (DAS)** uses standard fiber-optic cables to detect and measure vibrations along the way this cable was installed. Most of the time it is being applied by Oil & Gas companies in order to monitor the damage that can be caused by something else to pipes to which these cables are attached. Everything that can possibly damage the pipe and our cable makes small or even large vibrations(information) that the device in the end of our cable can catch, process(put into our application) or keep(for later use).



# SECTION 3

The plot shows the variance of signal over depth (meters). According to the graph, the signal is rather low in the first half of the plot and increases from 3000 to 4000, indicating some anomalies or a topic that requires further investigation. We can observe almost the same thing by looking at our HeatMap

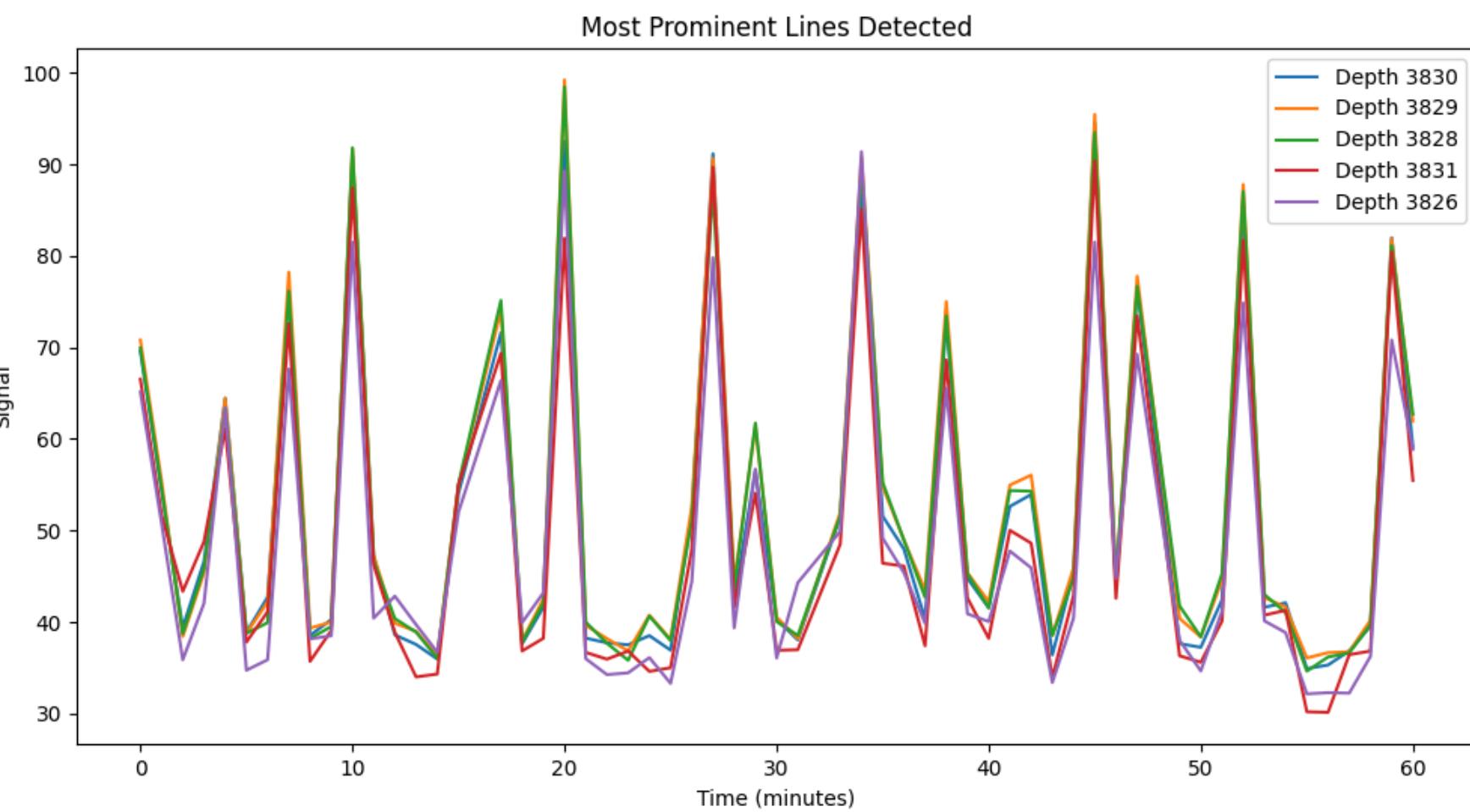


# SECTION 3

- **Data Conversion:** The DataFrame is converted to a NumPy array (`data_array`) for easier slicing and other operations.
- **Peak Detection:** A `find_peaks` function is defined to identify local maxima in each row (`signal`) where the value exceeds its nearby values and a specified threshold.
- **Prominence Calculation:** For each row, the function calculates the “prominence” by summing the heights of all peaks, storing this value in the ‘`peak_info`’ dictionary.
- **Sorting:** Depths are sorted in descending order of their peak prominence to get the most prominent rows easily(top ones).
- **Visualization:** The selected top rows are plotted to visualize the most prominent signal patterns over time.

## As a Result:

The top 10 depths with possible anomalies: [3830, 3829, 3828, 3831, 3826, 3827, 3832, 3414, 3411, 3413], while top 5 depths with possible anomalies: [3830, 3829, 3828, 3831, 3826] Excluding the ones that are far away [3414, 3411, 3413] We can indicate from here that there were some anomalies somewhere in depth of 3826-3831 meters down the surface of earth. (THE VALUES ARE APPROXIMATE)



**THANK  
YOU FOR  
YOUR  
ATTENTION**

Mammad Aliyev - applicant  
[mammed.aliyev.e@gmail.com](mailto:mammed.aliyev.e@gmail.com)  
+994 55 939 1771  
+44 7944 393 682

