I. Abstract

1. The purpose of the analysis

The analysis aims to address challenges faced by a small production company that sells four primary products at varying prices across different timeframes. Issues such as frequent price adjustments, incorrect user input, and system upgrades have made tracking price records for each product difficult to follow. The specific objectives of the analysis are:

- Determine the revenue generated by each product at different price points.
- Calculate the total revenue generated by each individual product.

2. The expectation of the analysis:

By conducting this analysis, the company aims to gain insights into the financial performance of each product, consider the dynamic pricing landscape, and address challenges related to pricing adjustments and system upgrades.

Also, the sample dataset was recorded within a month only, and it is expected that the solution applied to the dataset could also be applied to similar datasets with large numbers of records.

II. Dataset exploration

1. Dataset dictionaries

There are two datasets which were recorded in September 2018 containing essential information for analysis:

Prices Dataset:

- Product ID: Unique identifier for each product.
- Old Price: Initial selling price before adjustments.
- New Price: Selling price after adjustments.
- Updated at: Timestamp indicating when the new price became effective.

	product_id	old_price	new_price	updated_at
0	64	270000	239000	9/10/18 16:37
1	3954203	60000	64000	9/11/18 11:54
2	3954203	60500	57500	9/17/18 22:59
3	3954203	64000	60500	9/15/18 3:49
4	3954203	68800	60000	9/10/18 16:32
5	3998909	15500	16500	9/16/18 5:09
6	3998909	17000	15500	9/13/18 6:43
7	3998909	19000	17000	9/10/18 16:35
8	4085861	53500	52000	9/17/18 22:59
9	4085861	53500	67000	9/12/18 3:51
10	4085861	58000	53500	9/17/18 3:35
11	4085861	60000	53500	9/11/18 8:51
12	4085861	62500	58000	9/15/18 3:51
13	4085861	67000	62500	9/13/18 6:43

Sales Dataset:

- Product ID: Unique identifier for each product.
- Order at: Timestamp indicating when the order was processed.
- Quantity order: Number of units of the product ordered.

```
     product_id     ordered_at  quantity_ordered
0       3998909  9/18/18 17:51                 1
1       3998909  9/18/18 12:52                 1
2       3998909  9/18/18 11:33                 1
3       3998909  9/18/18 18:47                 1
4       3998909  9/18/18 17:36                 1
..          ...            ...               ...
170     4085861  9/13/18 23:29                 1
171          64  9/11/18 19:32                 1
172          64  9/14/18 20:10                 1
173          64  9/12/18 16:00                 1
174          64  9/12/18 13:23                 1

[175 rows x 3 columns]
```

1. Dataset understanding

It is necessary to determine the sales price of each order in sales table following the price dataset. Two scenario could be happened:

Post-Update Orders:

- The product was ordered following the most recent price update.
- This indicates that the order was placed after the latest price adjustment for the product. Pre-Update Orders:

Pre-Update Orders:

- The product was ordered before the latest price update.
- Depending on the timeframe of the order, the corresponding price can be determined.

2. Workflow

1. Handling DateTime Data:

To maintain consistency in data types, the 'order at' attribute in the sales table and 'updated at' in the price table need to be converted to the same datetime type.

```
#Convert datetime
df_sales['ordered_at'] = pd.to_datetime(df_sales['ordered_at'])
df_sales['ordered_at'] = df_sales['ordered_at'].dt.strftime('%Y-%m-%d %H:%M:%S')
df_prices['updated_at'] = pd.to_datetime(df_prices['updated_at'])
```

2. Utilizing Pandas Functions:

In line with the analysis objectives and dataset characteristics, various Pandas functions have been employed, including:

- Rank
- Join Tables
- Calculation (Sum, Group By)
- Union

These Pandas functions are crucial in achieving the analysis goals, facilitating tasks such as ranking, table merging, summarization through calculations, and combining datasets for comprehensive insights.

III. Results and Discussions:

1. Results:

- The revenue generated by each product at different price points in September 2018:

Product Id	Final Price	Total Order	Revenue
64	239000	4	956000
3954203	57500	1	57500
3954203	60000	3	180000
3954203	64000	10	640000
3998909	15500	1	15500
3998909	16500	14	231000
3998909	17000	2	34000
4085861	52000	20	1040000
4085861	53500	40	2140000
4085861	58000	38	2204000
4085861	62500	29	1812500
4085861	67000	13	871000
4085861	60000	3	180000

- Total revenue for each product in September 2018:

Product Id	Revenue
64	956000
3954203	877500
3998909	280500
4085861	8247500

2. Discussion:

Based on the analysis, the highest purchase patterns for each product have been identified and highlighted in the first table. This insight can serve as a valuable reference for the company to determine optimal pricing strategies. Additionally, this solution can be extended to analyse sales and prices on a broader scale, including monthly, quarterly, or yearly intervals, providing the company with a versatile tool for decision-making across different timeframes.