



MTA Subway Data Analysis For MA Company.

The Presentation outlines

Backstory

- About the company
- Our client idea
- Our analysis objectives

Data Preparing

- What is cleaning steps
- What is the added feature

Questions to Answers

- The data analysis

What About Sales

- The additional dataset
description
- The data analysis

Backstory

MA Company
for marketing



What About
Sales
Questions to
Answers
Data Preparing
Backstory

Used Tools



Excel



SQLite



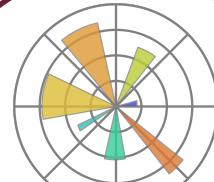
Seaborn



Pandas



SQLAlchemy



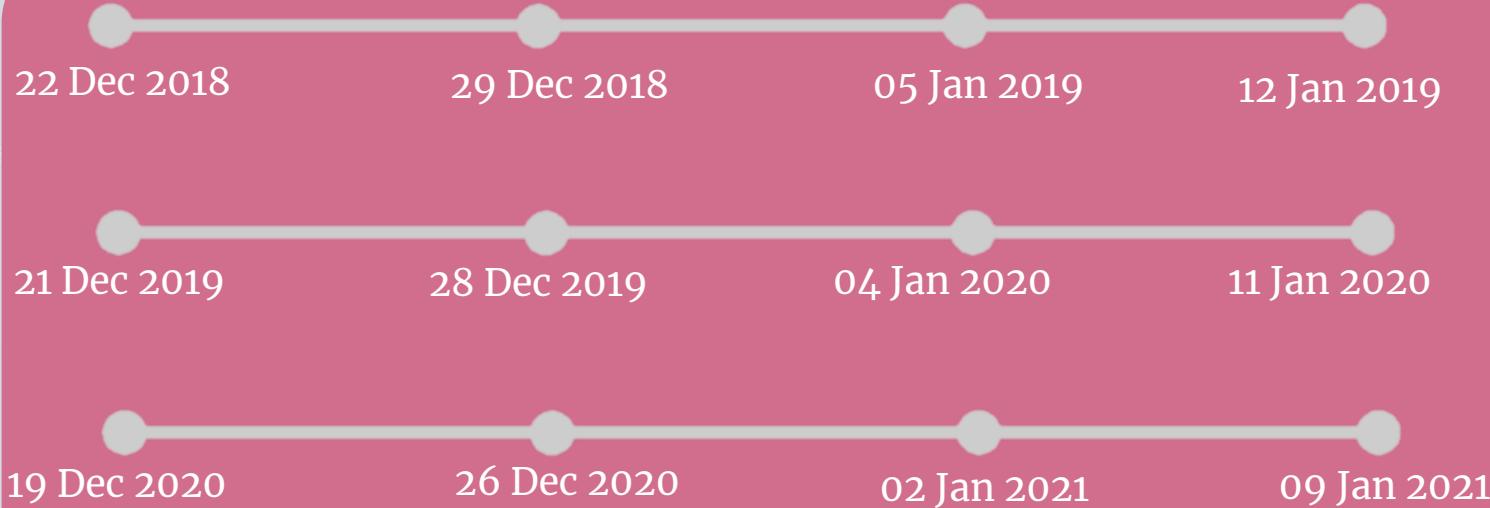
Matplotlib



NumPy

What About
Sales
Questions to
Answers
Data Preparing
Backstory

Chosen Sample

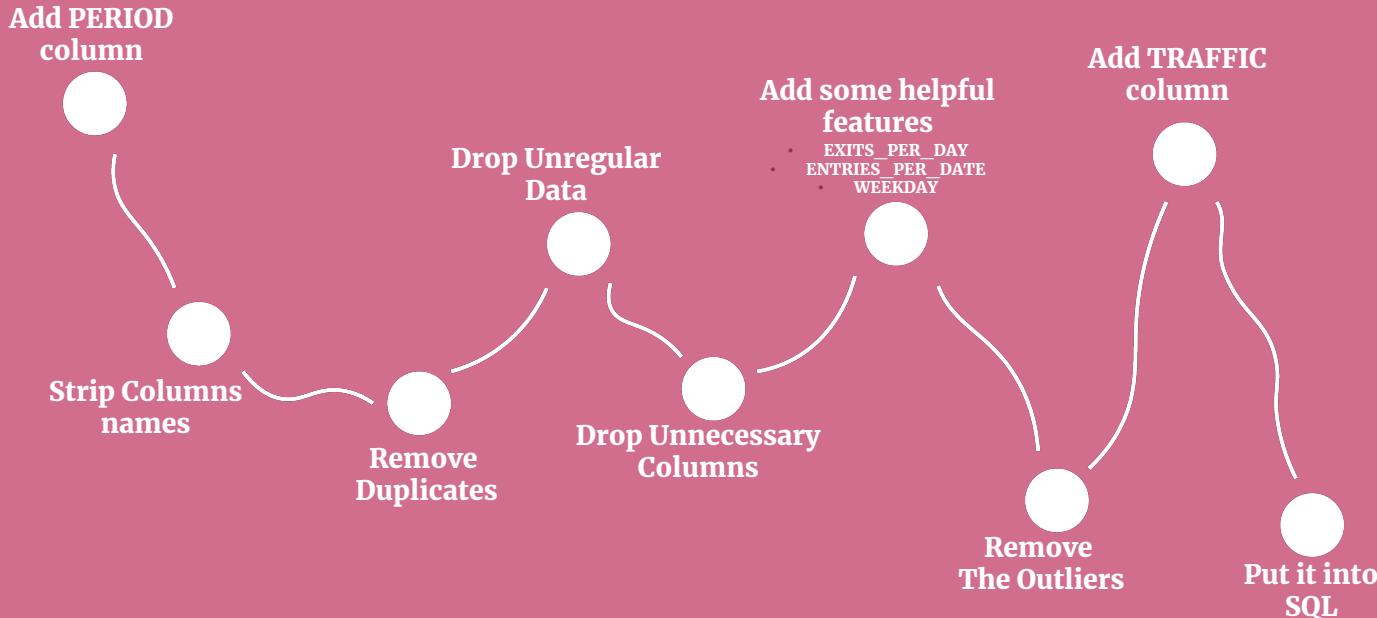


about 2,476,441 rows

11 Columns

What About Sales
Questions to Answers
Data Preparing Backstory

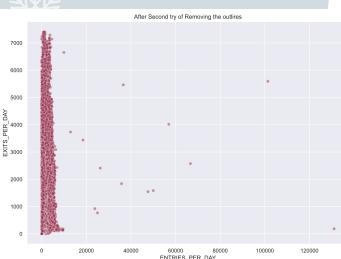
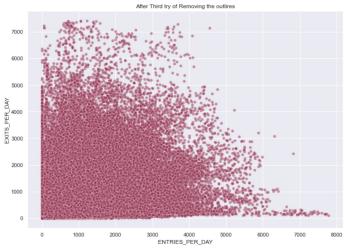
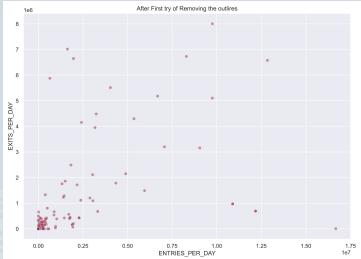
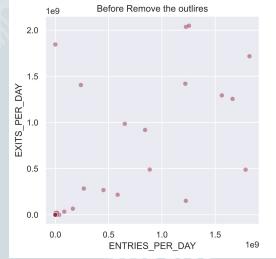
Cleaning Steps



What About Sales
Questions to Answers
Data Preparing

Backstory

Removing Outliers Steps



What About
Sales
Questions to
Answers
Data Preparing
Backstory

Questions To Answer

- **Is the MTA subway more crowded in the Christmas or the New Year period?**
- **then, what is the most crowded station?**
- **What is the most crowded day?**

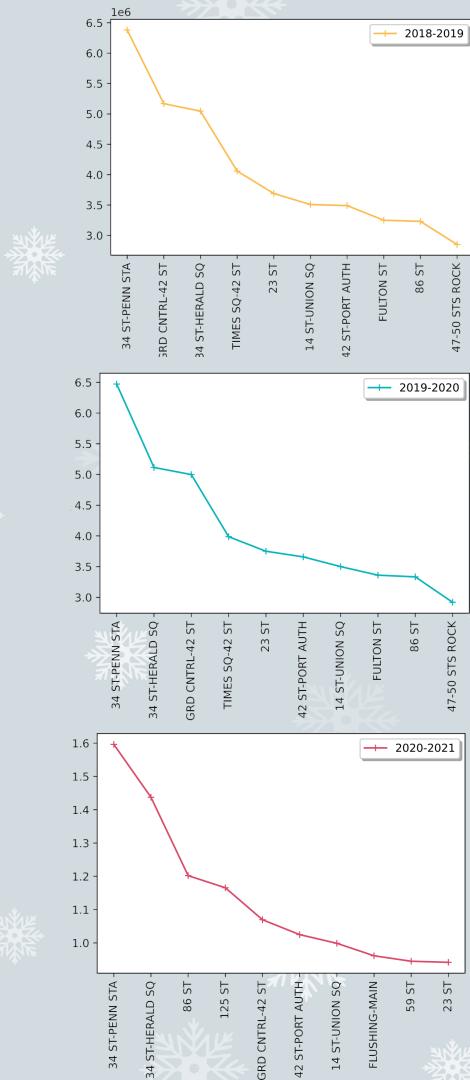
Questions to
Answers

What About
Sales

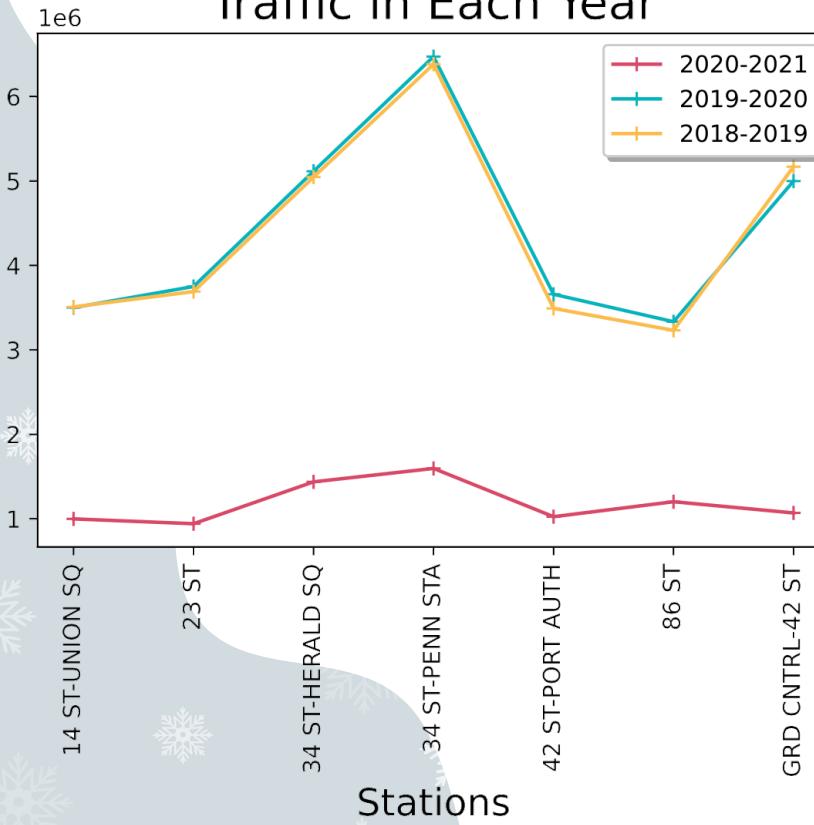
Backstory

Data Preparing

COVID-19 Pandemic Effect



Traffic in Each Year

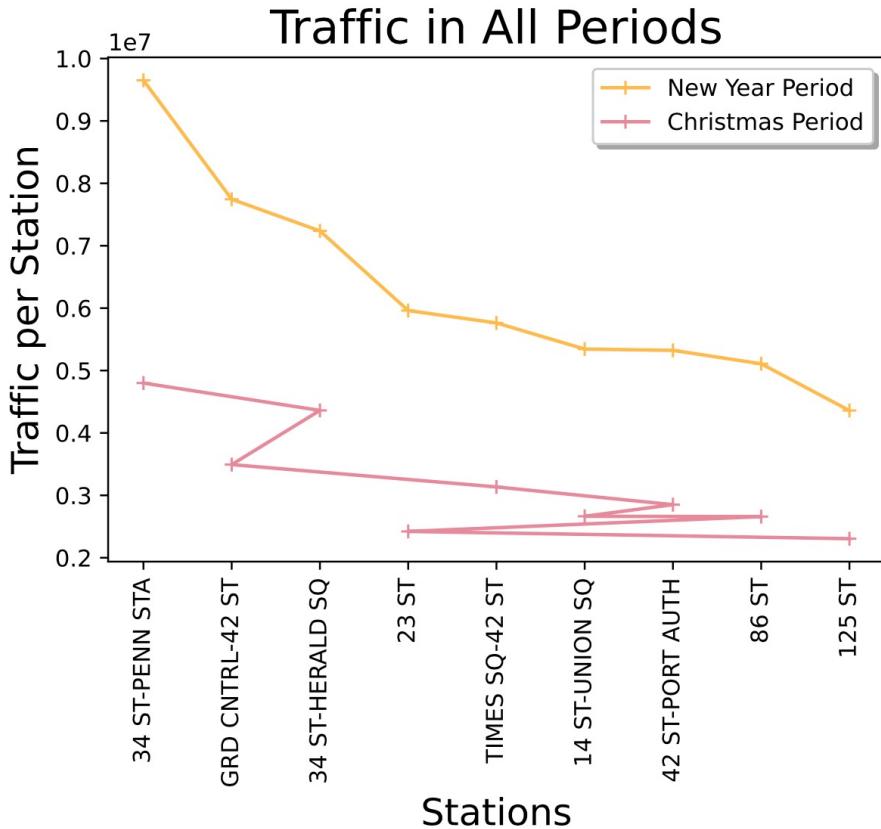


What About
Sales
Questions to
Answers

Backstory
Data Preparing

Questions To Answer

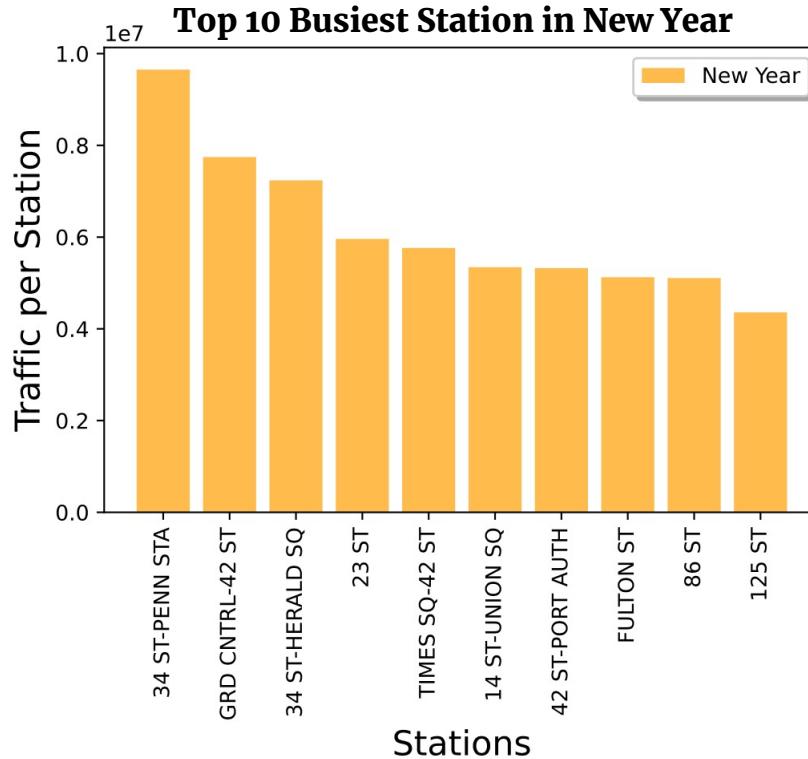
Q1: Is the MTA subway more crowded in the Christmas or the New Year period?



What About Sales
Questions to Answers
Data Preparing Backstory

Questions To Answer

Q2 : what is the most crowded station?



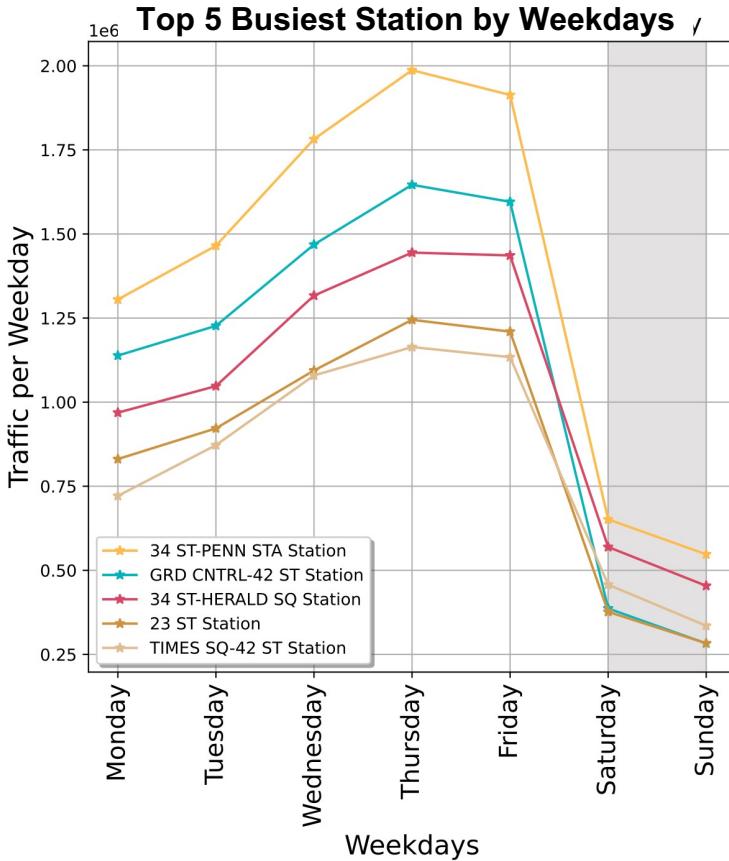
What About
Sales

Questions to
Answers

Data Preparing
Backstory

Questions To Answer

Q3: What is the most crowded day?



What About
Sales
Questions to
Answers

Data Preparing Backstory

Sales Dataset Contents and Analysis Process

Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Total Price
----------	---------	------------------	------------	------------	------------------	-------------

Analysis Process

- Data Cleaning
- Take only rows that address belongs to YN
- Add Total Price Column = quantity ordered * Price Each
- Sum the total price for each month

Order ID	Product	Quantity Ord	Price Each	Order Date	Purchase Address		
----------	---------	--------------	------------	------------	------------------	--	--

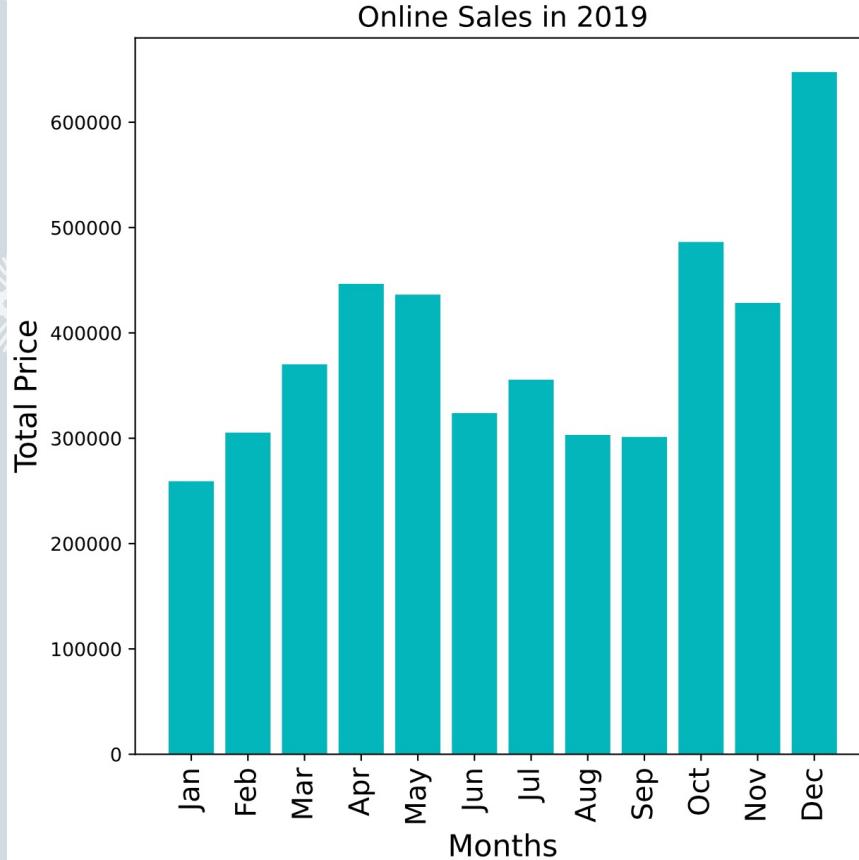
What About
Sales

Questions to
Answers

Data Preparing

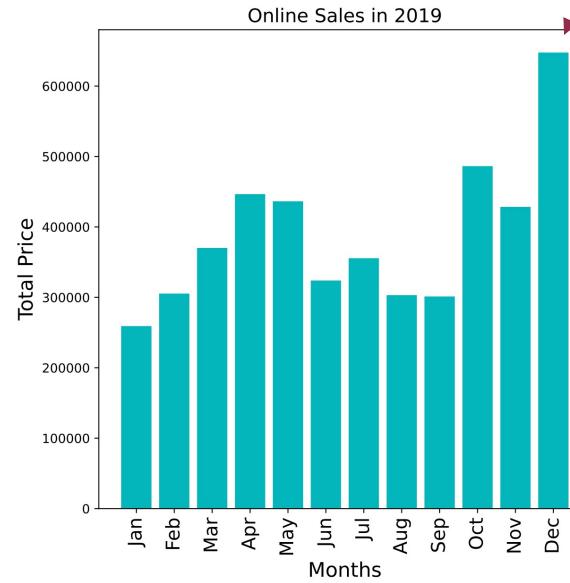
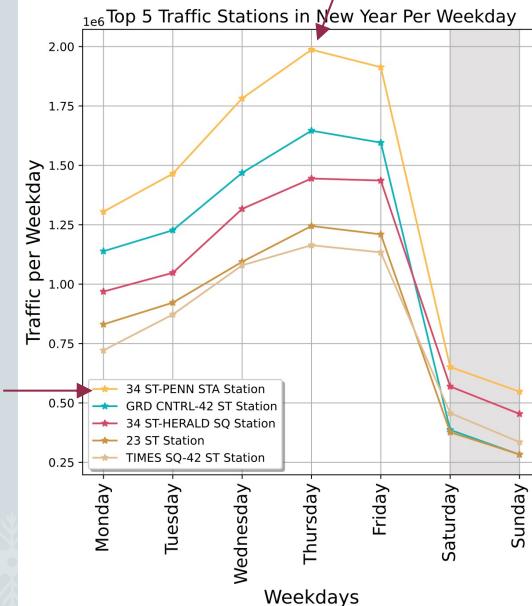
Backstory

Total Price for each month



What About Sales
Questions to Answers
Data Preparing Backstory

The Final Decision



**The sample distribution will be
on 24 Dec to 1 Jan
On Thursday
At 34 ST-PENN Station**

Thank you !

Done By: Shahad Alkaltham

