Customer Feedback Analysis and Improvement

Comprehensive Analysis using SQL, Data Warehousing, Python

Project Brief

This project aims to analyze customer reviews of British airline customers using data science techniques in order to gain insights used to identify areas of improvement.

The Phases

- create database using SQL database and collecting data.
- cleanse the collected data using python data processing.
- load the cleanse data into data warehouse.
- obtain insights to identify areas of improvement.

PHASE 1: CREATING DATABASE

- Database creation : using SQL server to create database used to store customer feedback.
- Table definition : defining a table within the database to hold the customer feedback: (specifying columns & data type).
- Importing the dataset : the dataset CSV file is imported to the database.

PHASE 1: CREATING DATABASE

```
CREATE TABLE [dbo].[CustomerReview1](
    [ID] [smallint] NULL,
    [OverallRating] [tinyint] NULL,
    [ReviewHeader] [nvarchar](100) NULL,
    [Name] [nvarchar](50) NULL,
    [Datetime] [date] NULL,
    [VerifiedReview] [bit] NULL,
    [ReviewBody] [varchar](max) NULL,
    [TypeOfTraveller] [nvarchar](50) NULL,
    [SeatType] [nvarchar](50) NULL,
    [Route] [nvarchar](50) NULL,
    [DateFlown] [date] NULL,
    [SeatComfort] [tinyint] NULL,
    [CabinStaffService] [tinyint] NULL,
    [GroundService] [tinyint] NULL,
    [ValueForMoney] [tinyint] NULL,
    [Recommended] [bit] NULL,
    [Aircraft] [nvarchar](50) NULL,
    [Food_Beverages] [tinyint] NULL,
    [InflightEntertainment] [tinyint] NULL,
    [Wifi_Connectivity] [tinyint] NULL
) ON [PRIMARY] TEXTIMAGE_ON [PRIMARY]
GO
```

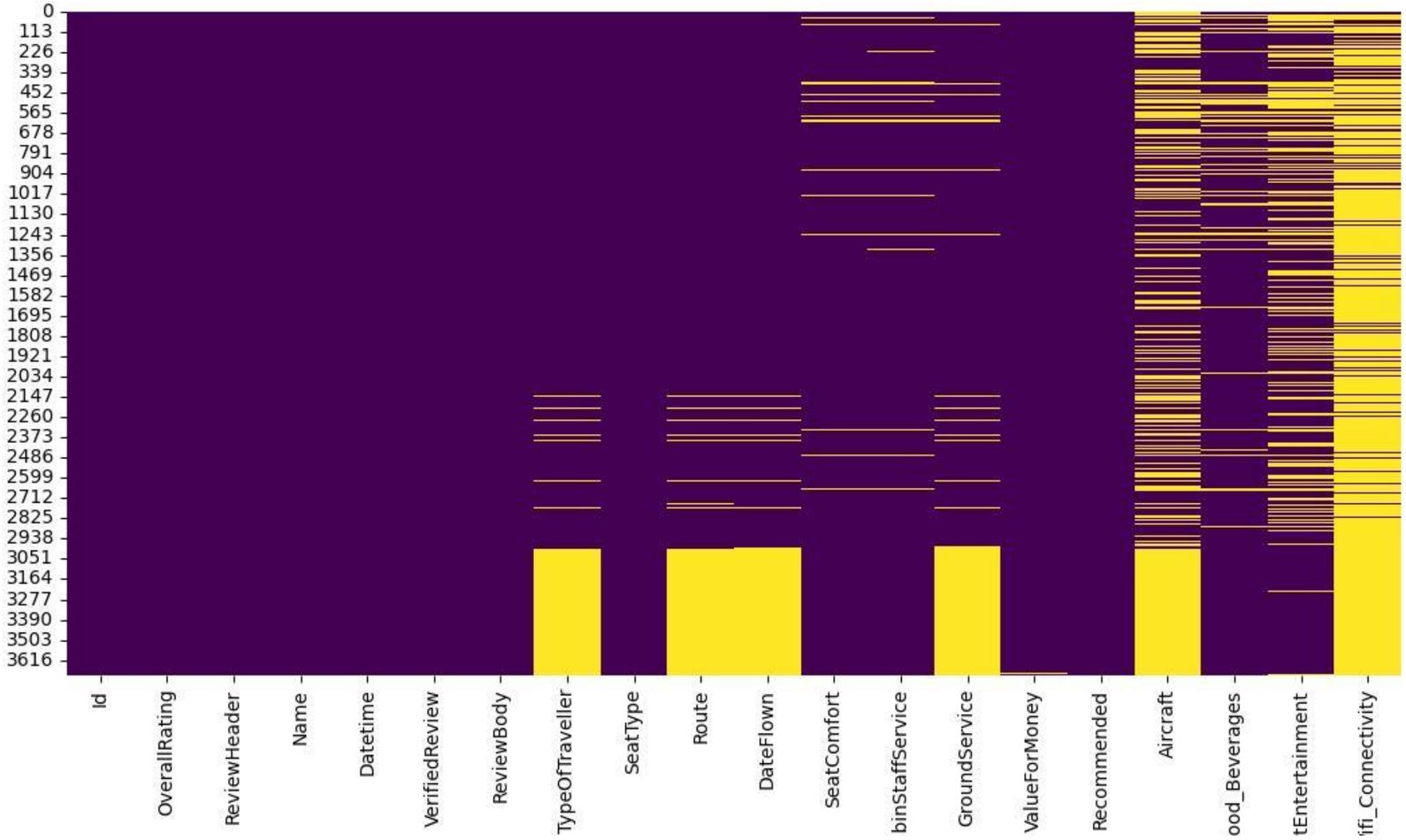
PHASE 1: CREATING DATABASE

column1	OverallRating	ReviewHeader	Name	Datetime	VerifiedReview	ReviewBody	TypeOfTraveller	SeatType	Route	DateFlown	SeatComfort	CabinStaffServ...	GroundService	ValueF
795	10	"Excellent cust...	Duane Wade	2019-08-15	False	Gatwick to Orla...	Family Leisure	Economy Class	Gatwick to Orla...	2020-07-19	4	5	5	4
796	6	"There is enoug...	Chun Sing Poon	2019-08-14	True	London to Mun...	Family Leisure	Economy Class	London to Mun...	2020-08-19	4	4	2	3
797	1	"will try to avoi...	Neil McKeganey	2019-08-14	True	London to Glas...	Business	Economy Class	London to Glas...	2020-06-19	2	2	1	1
798	1	"seats are aging...	Glenn Tong	2019-08-13	True	The seating con...	Business	Business Class	London to Hon...	2020-08-19	1	5	4	1
799	10	"the flight was ...	D Patterson	2019-08-12	True	No problems at...	Solo Leisure	Economy Class	Vienna to Heat...	2020-07-19	4	5	5	4
800	2	"A shambolic ai...	Peter Barker	2019-08-11	True	Malaga to Lond...	Couple Leisure	Business Class	Malaga to Lond...	2020-08-19	1	2	1	1
801	2	"Extremely disa...	Z Han	2019-08-10	True	Sydney to Sing...	Couple Leisure	Business Class	Sydney to Sing...	2020-08-19	1	4	1	1
802	1	"Utterly appalli...	H Lind	2019-08-09	True	London Heathr...	Solo Leisure	Economy Class	London Heathr...	2020-08-19	2	3	1	1
803	2	"Very tired and ...	Robert Davis	2019-08-09	False	Heathrow to Gl...	Solo Leisure	Economy Class	Heathrow to Gl...	2020-08-19	1	1	1	1
804	1	"awful experien...	H Lind	2019-08-09	True	A supremely a...	Solo Leisure	Economy Class	Inverness to Va...	2020-08-19	2	2	1	1
805	1	"a below avera...	Dave Matthews	2019-08-09	False	British Airways ...	Couple Leisure	Economy Class	London Heathr...	2020-08-19	2	4	1	1
806	1	"they managed...	P Garrett	2019-08-08	True	London to Nice...	Solo Leisure	Economy Class	London to Nice	2020-08-19	NULL	NULL	1	1
807	1	"time for some ...	Michael Croft	2019-08-07	True	Palma to Lond...	Couple Leisure	Business Class	Palma to Lond...	2020-08-19	1	1	1	1
808	1	"Trip was basic"	Mike Pettet	2019-08-06	True	London City to ...	Couple Leisure	Economy Class	London City to ...	2020-08-19	3	2	2	2
809	1	"split all family ...	Ilona Brown	2019-08-02	False	London to Faro...	Family Leisure	Economy Class	London to Faro	2020-08-19	NULL	NULL	NULL	1
810	3	"Staff not in hi...	M Dartira	2019-08-02	True	Calgary to Lon...	Couple Leisure	Economy Class	Calgary to Lon...	2020-07-19	3	1	5	3
811	8	"pretty good, d...	Milena Cook	2019-08-02	False	London Heathr...	Family Leisure	Economy Class	London Heathr...	2020-07-19	5	5	4	4
812	1	"due to the del...	A Almasha	2019-08-02	True	Nice to Kuwait ...	Solo Leisure	Business Class	Nice to Kuwait ...	2020-07-19	1	1	1	1
813	9	"made memora...	Peter Saunders	2019-08-01	False	Buenos Aires to...	Couple Leisure	Economy Class	Buenos Aires to...	2020-07-19	3	5	4	4
814	6	"seat is quite co...	Chun Sing Poon	2019-08-01	True	Hong Kong to ...	Family Leisure	Economy Class	Hong Kong to ...	2020-07-19	4	3	2	3
815	3	"What a useless...	Tony Palmer	2019-07-31	True	What happene...	Business	Premium Econ...	London Heathr...	2020-07-19	3	4	1	3
816	2	"are staff prepa...	K Robinson	2019-07-31	True	London Heathr...	Solo Leisure	Business Class	London Heathr...	2020-07-19	2	1	3	1
817	5	"777 aircraft is l...	R Jitene	2019-07-30	True	I was flying BA ...	Business	Economy Class	London to Delhi	2020-06-19	1	1	3	2
818	6	"Thank you so ...	Barbara Chareka	2019-07-30	True	Seattle to Durb...	Solo Leisure	Economy Class	Seattle to Durb...	2020-07-19	4	5	5	5
819	2	"Noisy cabin, ti...	A Hardein	2019-07-30	True	London Heathr...	Family Leisure	Economy Class	London Heathr...	2020-07-19	1	3	3	2
820	1	"worst business...	C Marni	2019-07-29	True	Mykonos to Lo...	Business	Business Class	Mykonos to Lo...	2020-07-19	2	1	1	1
821	8	"flight has impr...	K Lee	2019-07-29	True	Johannesburg t...	Family Leisure	Economy Class	Johannesburg ...	2020-07-19	4	5	5	5
822	5	"seat comfort h...	David Worten	2019-07-28	True	Grenada to Gat...	Family Leisure	Economy Class	Grenada to Gat...	2020-07-19	1	4	4	3
823	1	"delayed by 3 h...	M Elbadawi	2019-07-27	True	Prague to Lond...	Family Leisure	Economy Class	Prague to Lond...	2020-07-19	1	1	1	1
824	10	"professional a...	S Calito	2019-07-24	True	London to Dub...	Solo Leisure	Premium Econ...	London to Dubai	2020-07-19	5	5	5	5
825	9	"staff were help...	Karen Deitch	2019-07-21	True	Tel Aviv to Toro...	Family Leisure	Economy Class	Tel Aviv to Toro...	2020-07-19	4	5	4	5
826	3	"sent to the wr...	P Meares	2019-07-20	True	I flew yesterday...	Solo Leisure	Business Class	Heathrow to M...	2020-07-19	2	3	1	1
827	8	"team was very ...	C Marvere	2019-07-18	True	The first part fr...	Solo Leisure	Business Class	Brussels to Larn...	2020-06-19	3	4	4	4
828	2	"The meal was i...	G Wailin	2019-07-16	True	Stockholm to L...	Couple Leisure	Business Class	Stockholm to L...	2020-07-19	2	1	3	1
829	2	"no apologies f...	L Allen	2019-07-15	True	Seems like no o...	Business	Business Class	Toronto to Lon...	2020-07-19	2	1	1	1
830	7	"don't waste yo...	Paul Rogers	2019-07-14	True	Premium econ...	Family Leisure	Premium Econ...	Orlando to Lon...	2020-07-19	3	4	3	3
831	1	"Meal service h...	T Kerber	2019-07-13	True	Lisbon to Lond...	Couple Leisure	Business Class	Lisbon to Lond...	2020-07-19	1	1	2	1

PHASE 2 : DATA CLEANSING USING Python

- using ETL process to cleanse the dataset :
- Data extraction: the dataset is extracted from the database in SQL server to Python to capture the necessary data for the cleansing process
- Data transformation : the extracted data using transformation techniques in Python is now used to handle missing values and correct data types
- Data validation : after the transformation a validation techniques is applied to identify and remove duplicates, outliers, and incorrect formats
- the cleanse and transformed data will be then loaded to different database.

Dataset Before Cleansing:



- visuals of the captured data before the transformation process

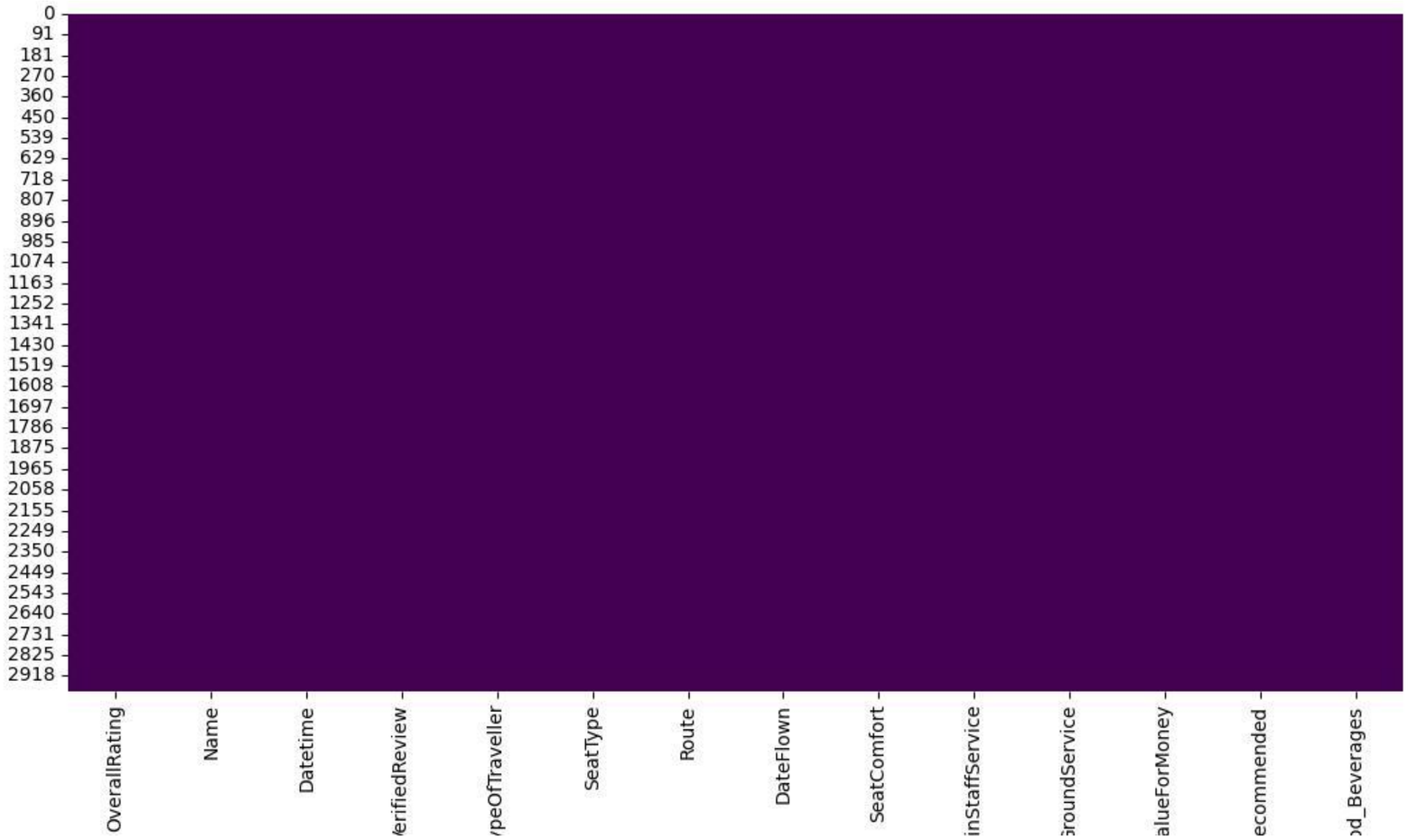
Dataset Before Cleansing:

```
Id                          0
OverallRating               5
ReviewHeader                0
Name                        0
Datetime                    0
VerifiedReview              0
ReviewBody                  0
TypeOfTraveller           771
SeatType                    2
Route                     775
DateFlown                 778
SeatComfort               116
CabinStaffService         127
GroundService             846
ValueForMoney               1
Recommended                 0
Aircraft                 1779
Food_Beverages            386
InflightEntertainment    1150
Wifi_Connectivity        3092
```

```
RangeIndex: 3701 entries, 0 to 3700
Data columns (total 20 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   Id                     3701 non-null   int64
 1   OverallRating          3696 non-null   float64
 2   ReviewHeader           3701 non-null   object
 3   Name                   3701 non-null   object
 4   Datetime               3701 non-null   object
 5   VerifiedReview         3701 non-null   bool
 6   ReviewBody             3701 non-null   object
 7   TypeOfTraveller        2930 non-null   object
 8   SeatType               3699 non-null   object
 9   Route                  2926 non-null   object
 10  DateFlown              2923 non-null   object
 11  SeatComfort            3585 non-null   float64
 12  CabinStaffService      3574 non-null   float64
 13  GroundService          2855 non-null   float64
 14  ValueForMoney          3700 non-null   float64
 15  Recommended            3701 non-null   bool
 16  Aircraft               1922 non-null   object
 17  Food_Beverages         3315 non-null   float64
 18  InflightEntertainment  2551 non-null   float64
 19  Wifi_Connectivity      609 non-null    float64
dtypes: bool(2), float64(8), int64(1), object(9)
```

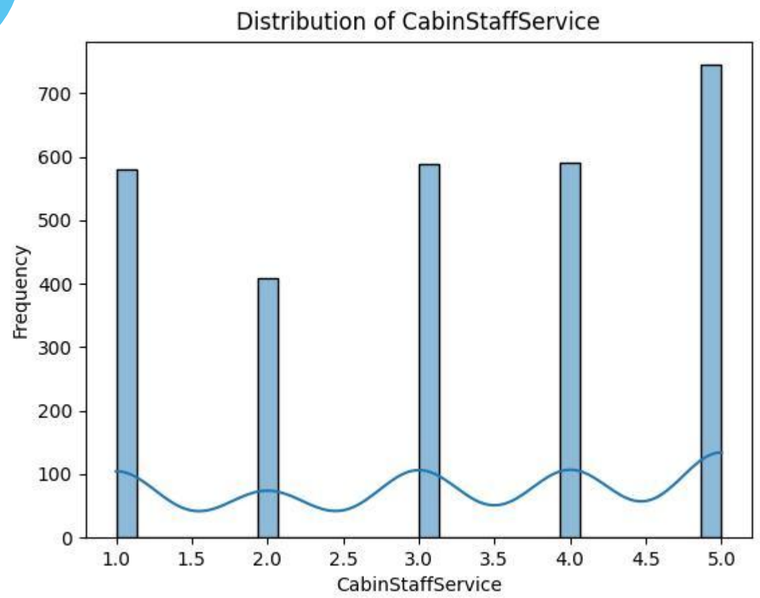
- visuals of the captured data before the transformation process

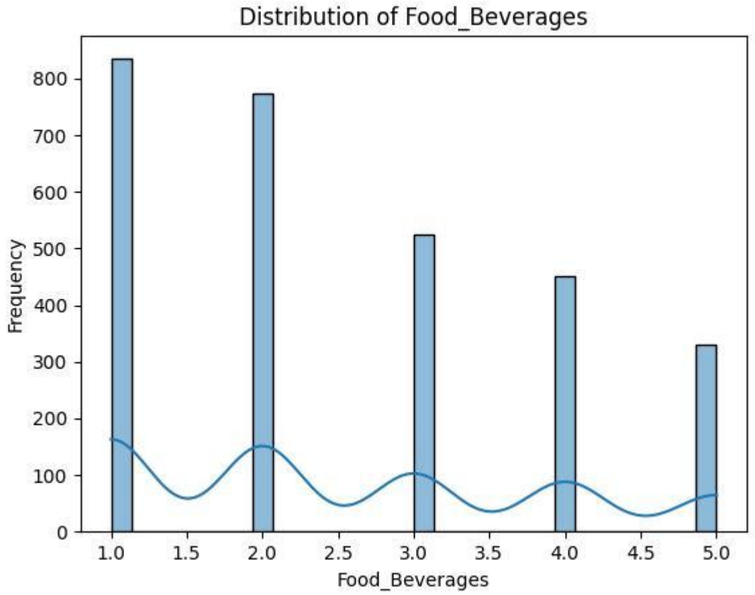
Dataset After Cleansing:

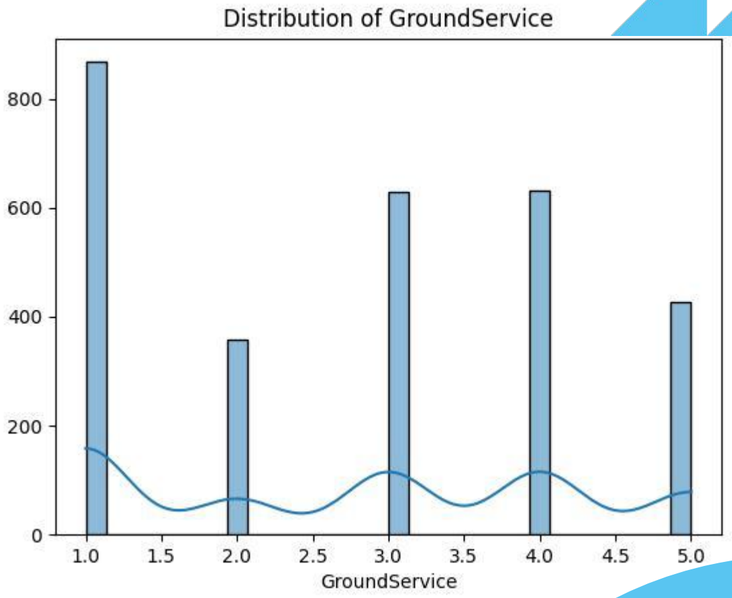


- visuals of the captured data after the transformation process

Dataset Visualisation:

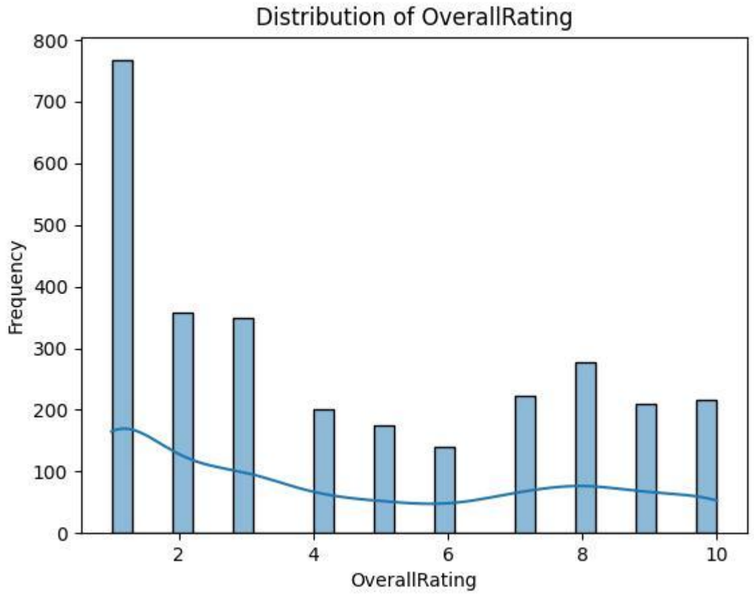


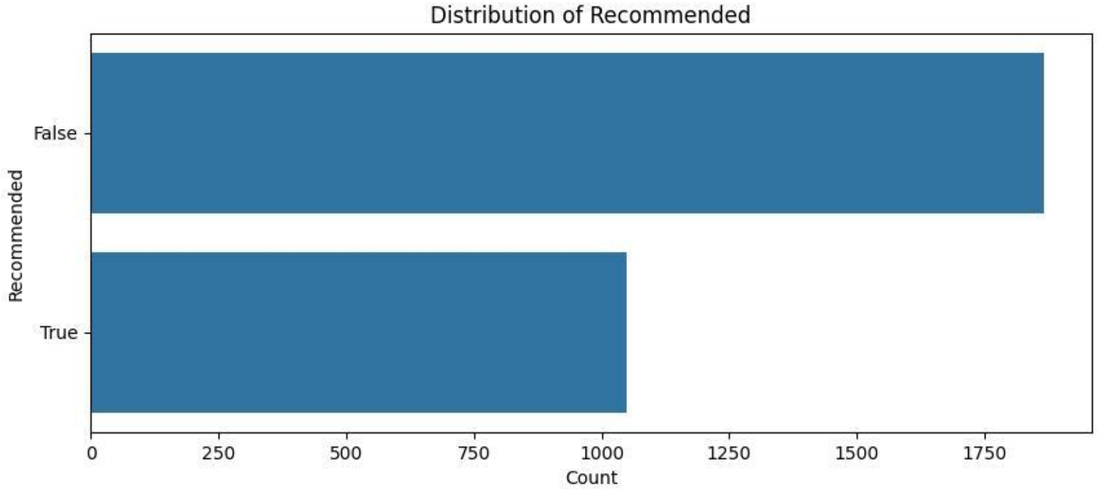




- visuals of the captured data after the transformation process

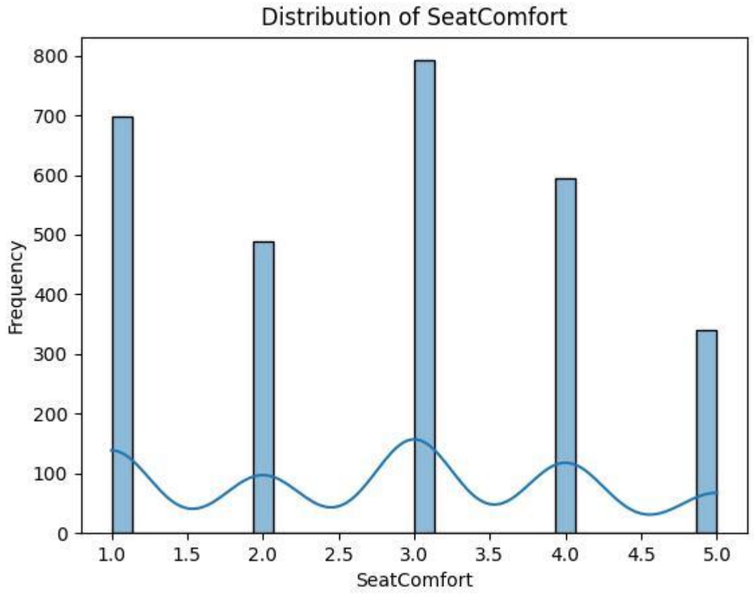
Dataset Visualisation:

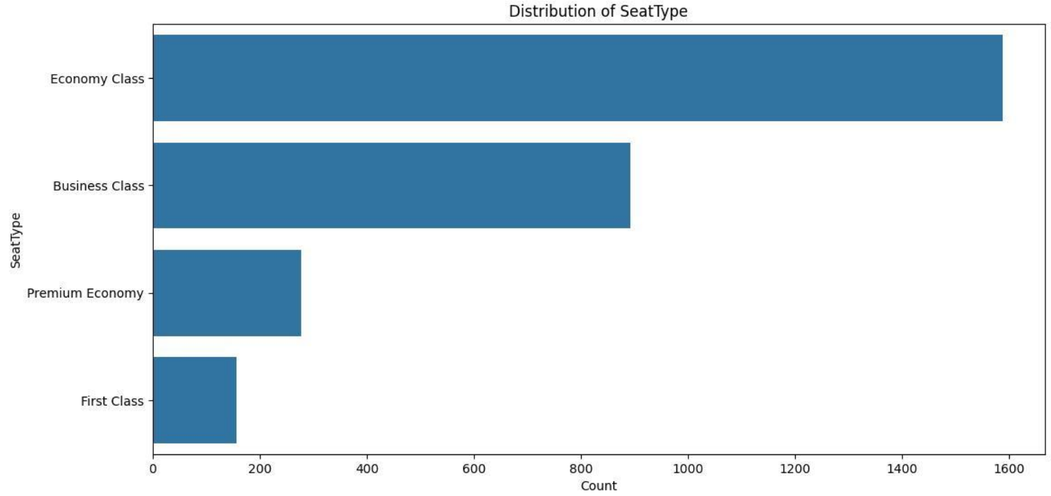




- visuals of the captured data after the transformation process

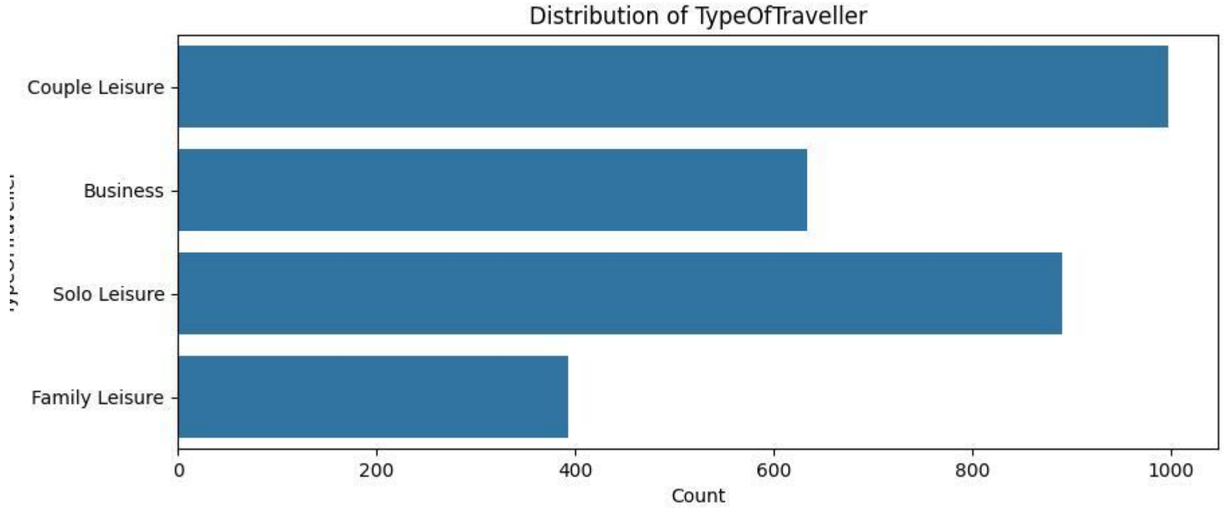
Dataset Visualisation:

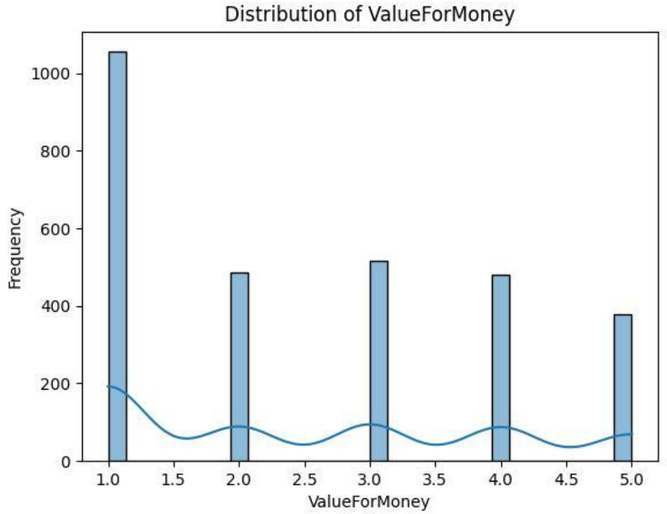




- visuals of the captured data after the transformation process

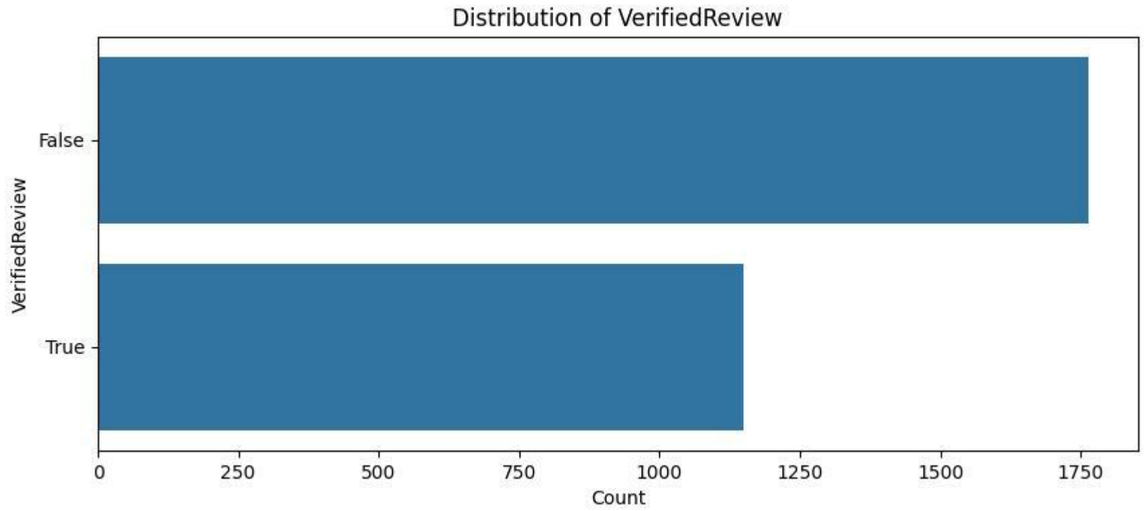
Dataset Visualisation:





- visuals of the captured data after the transformation process

Dataset Visualisation:



- visuals of the captured data after the transformation process

PHASE 3 : THE STAGING PROCESS

- Data preparation: the cleaned and transformed data is stored in staging area database to prepare the data to be transformed to data warehouse.

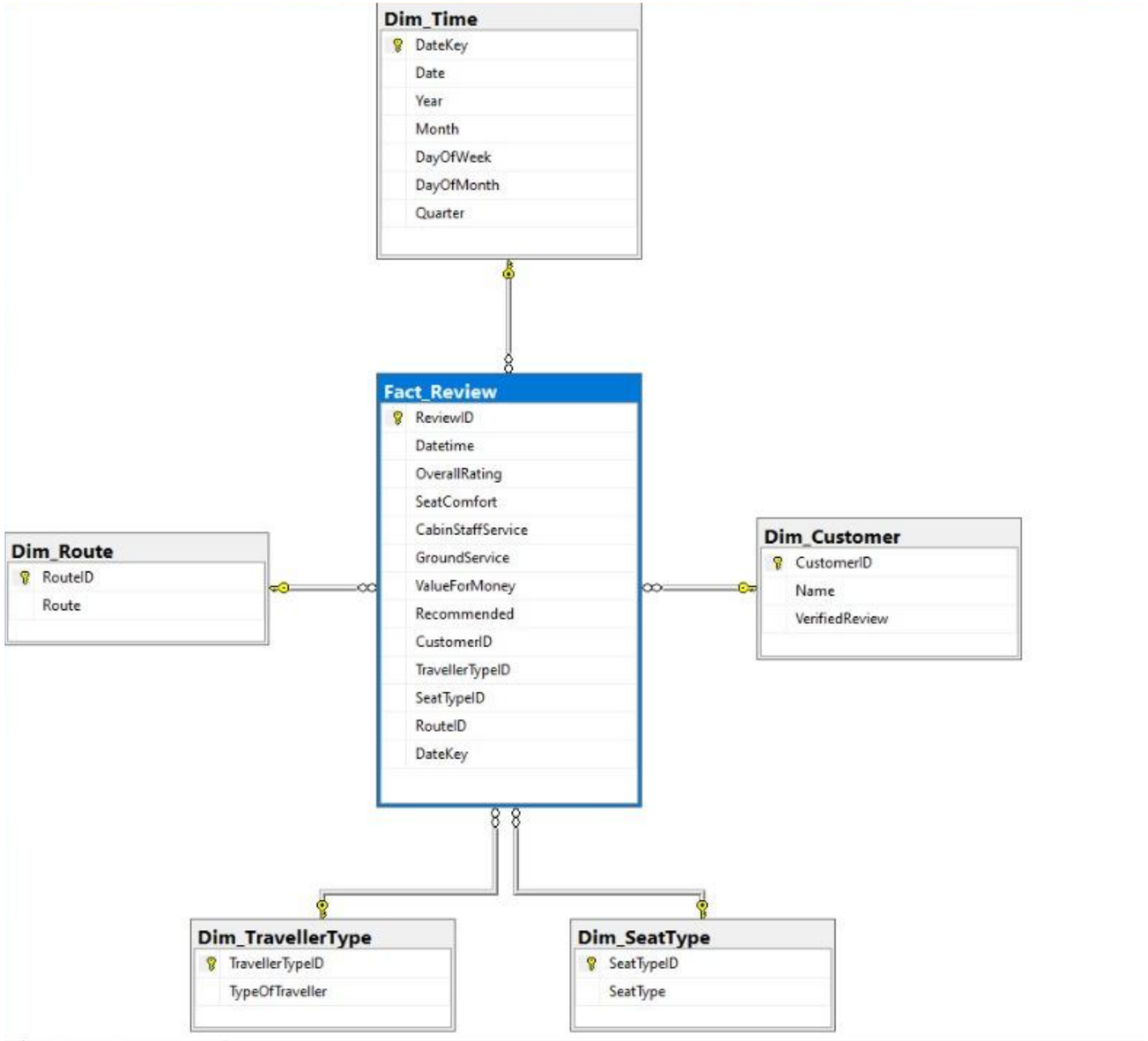
PHASE 3 : THE STAGING PROCESS

```
CREATE TABLE [dbo].[Customerreview2](
    -- Overall rating of the review
    [OverallRating] [tinyint] NULL,

    -- Name of the customer
    [Name] [nvarchar](50) NULL,

    -- Date and time of the review
    [Datetime] [date] NULL,

    -- Indicates if the review is verified
    [VerifiedReview] [bit] NULL,

    -- Type of traveller
    [TypeOfTraveller] [nvarchar](50) NULL,

    -- Type of seat
    [SeatType] [nvarchar](50) NULL,

    -- Travel route
    [Route] [nvarchar](50) NULL,

    -- Date flown
    [DateFlown] [date] NULL,

    -- Comfort rating of the seat
    [SeatComfort] [tinyint] NULL,

    -- Service rating of the cabin staff
    [CabinStaffService] [tinyint] NULL,

    -- Service rating of ground services
    [GroundService] [tinyint] NULL,

    -- Value for money rating
    [ValueForMoney] [tinyint] NULL,

    -- Indicates if the service is recommended
    [Recommended] [bit] NULL,

    -- Food and beverages rating
    [Food_Beverages] [tinyint] NULL
) ON [PRIMARY];
```

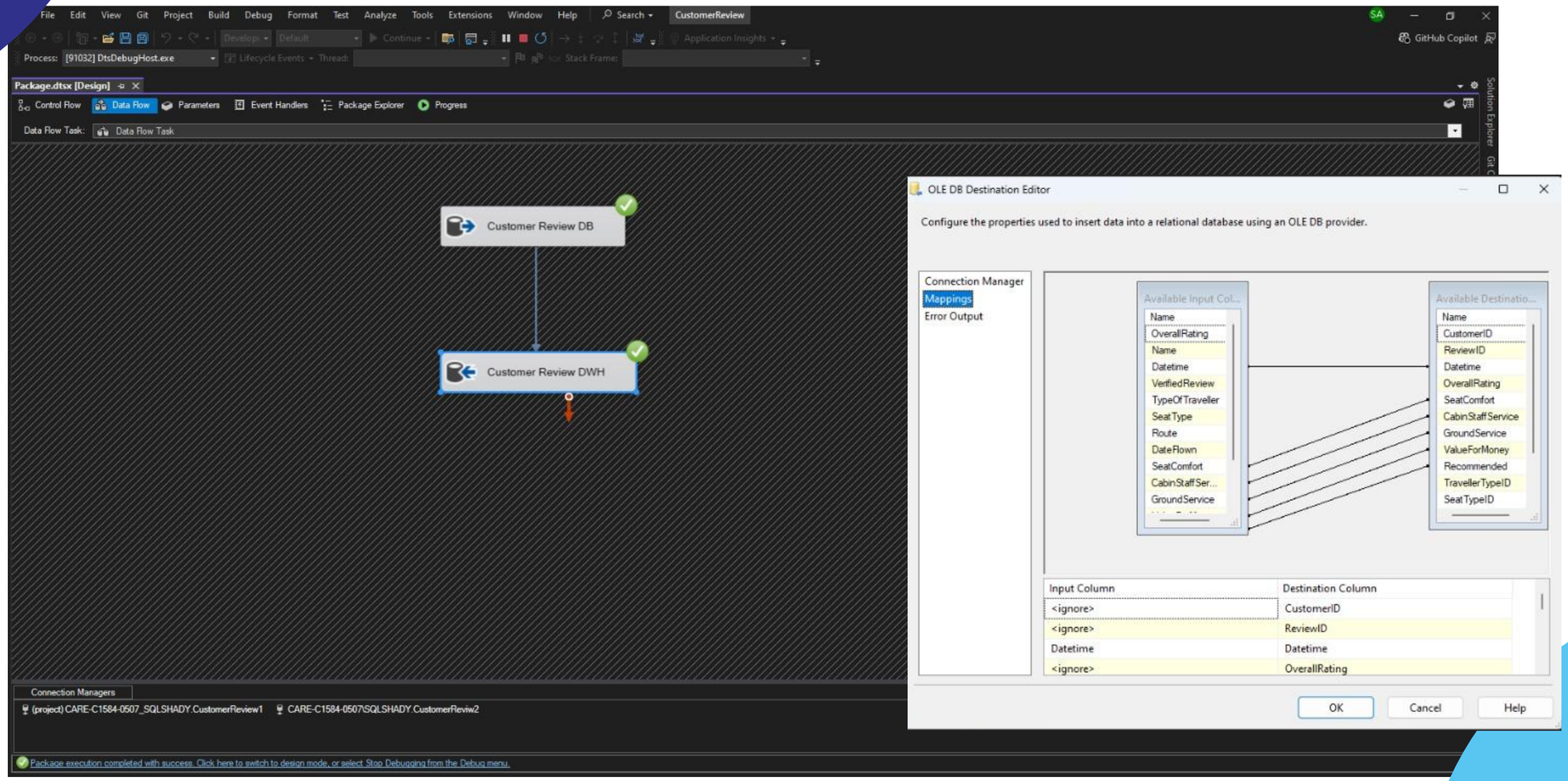
PHASE 4 : Warehousing

- Data Warehouse creation : using SQL server to create data warehouse used to customer feedback analysis.
- Table definition : defining the fact table and Dimensions tables.
- Extracting the data from database into the new data warehouse using SSIS.

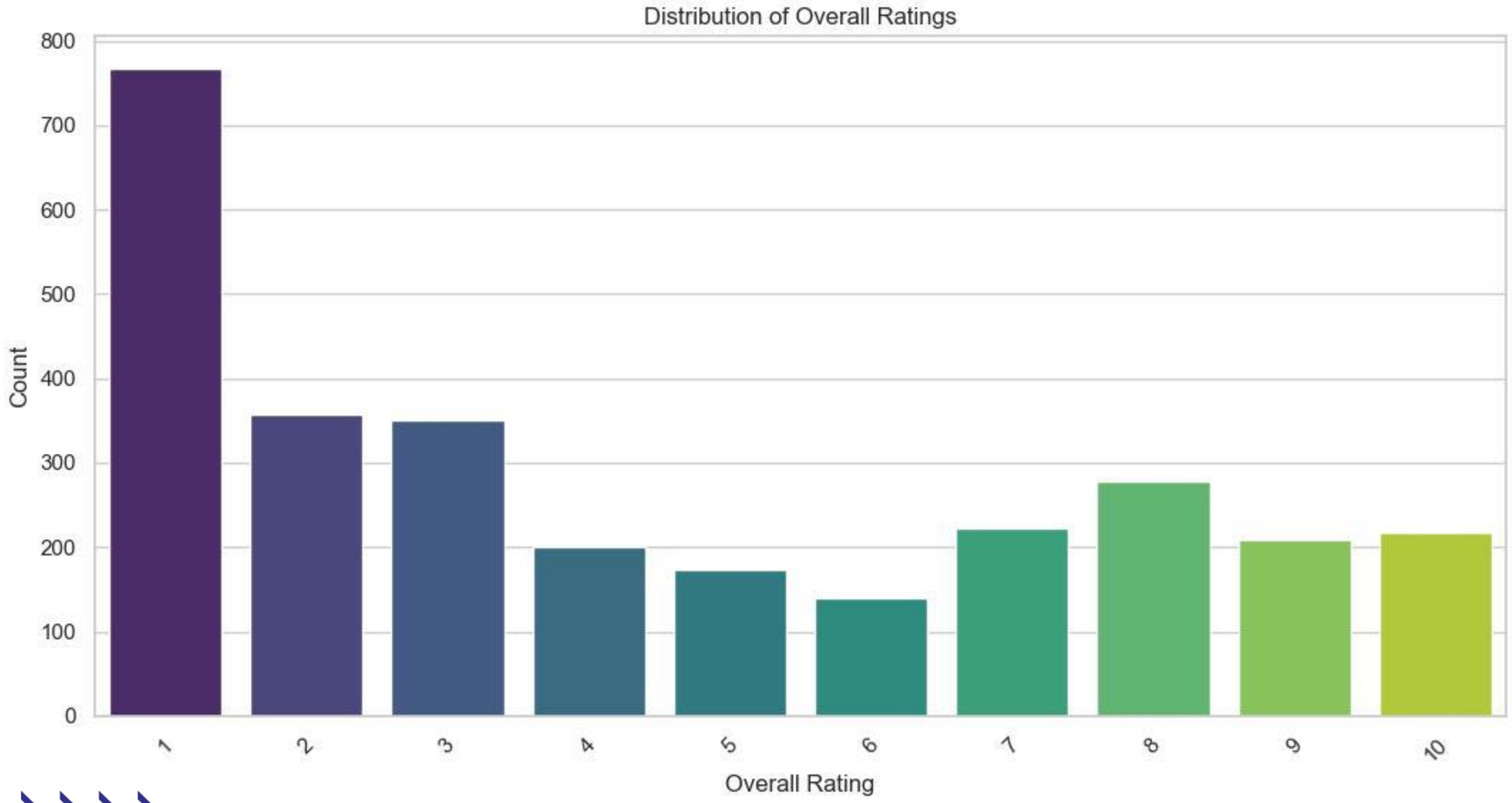
PHASE 4 : Warehousing



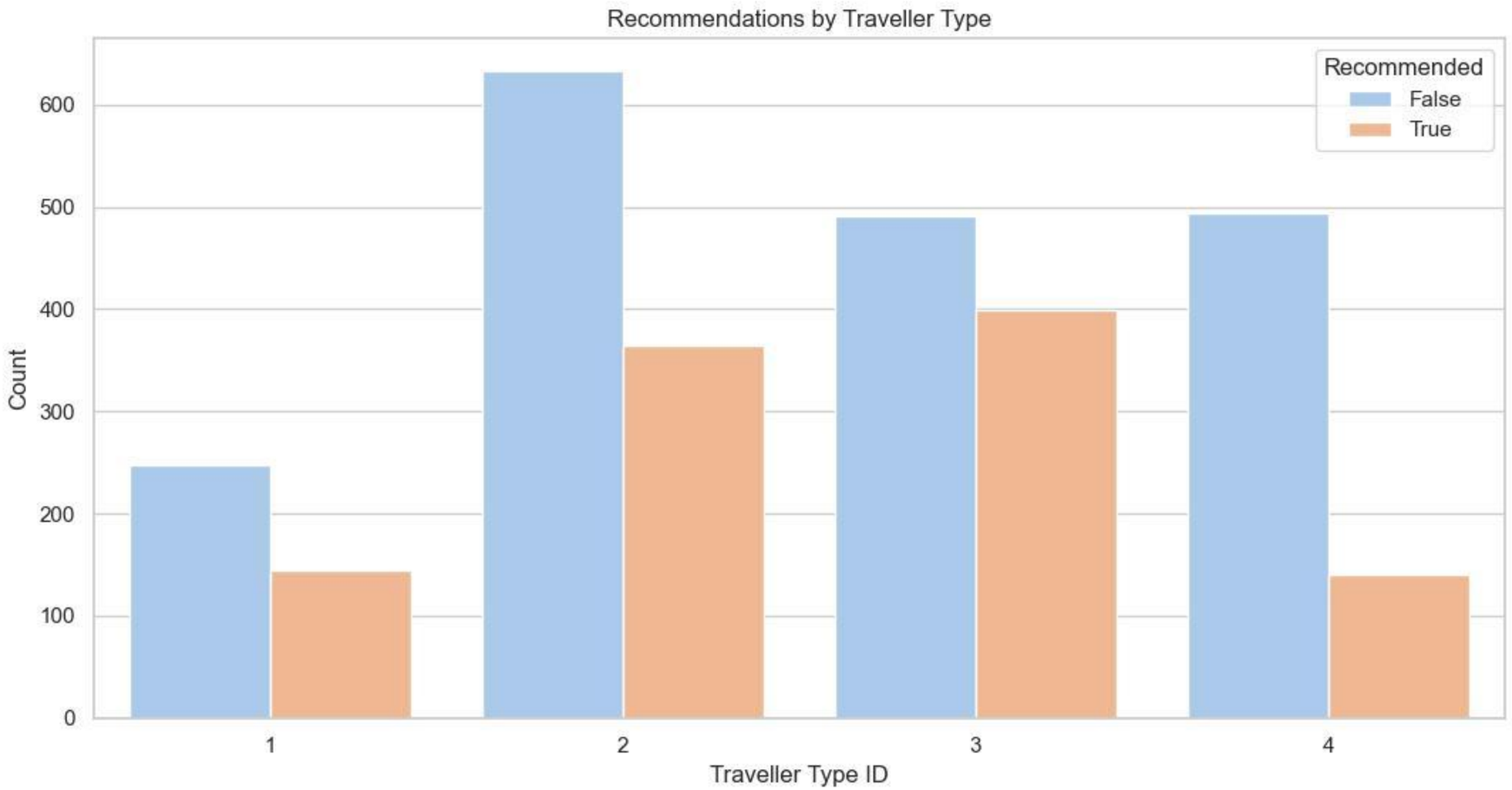
PHASE 4 : Warehousing



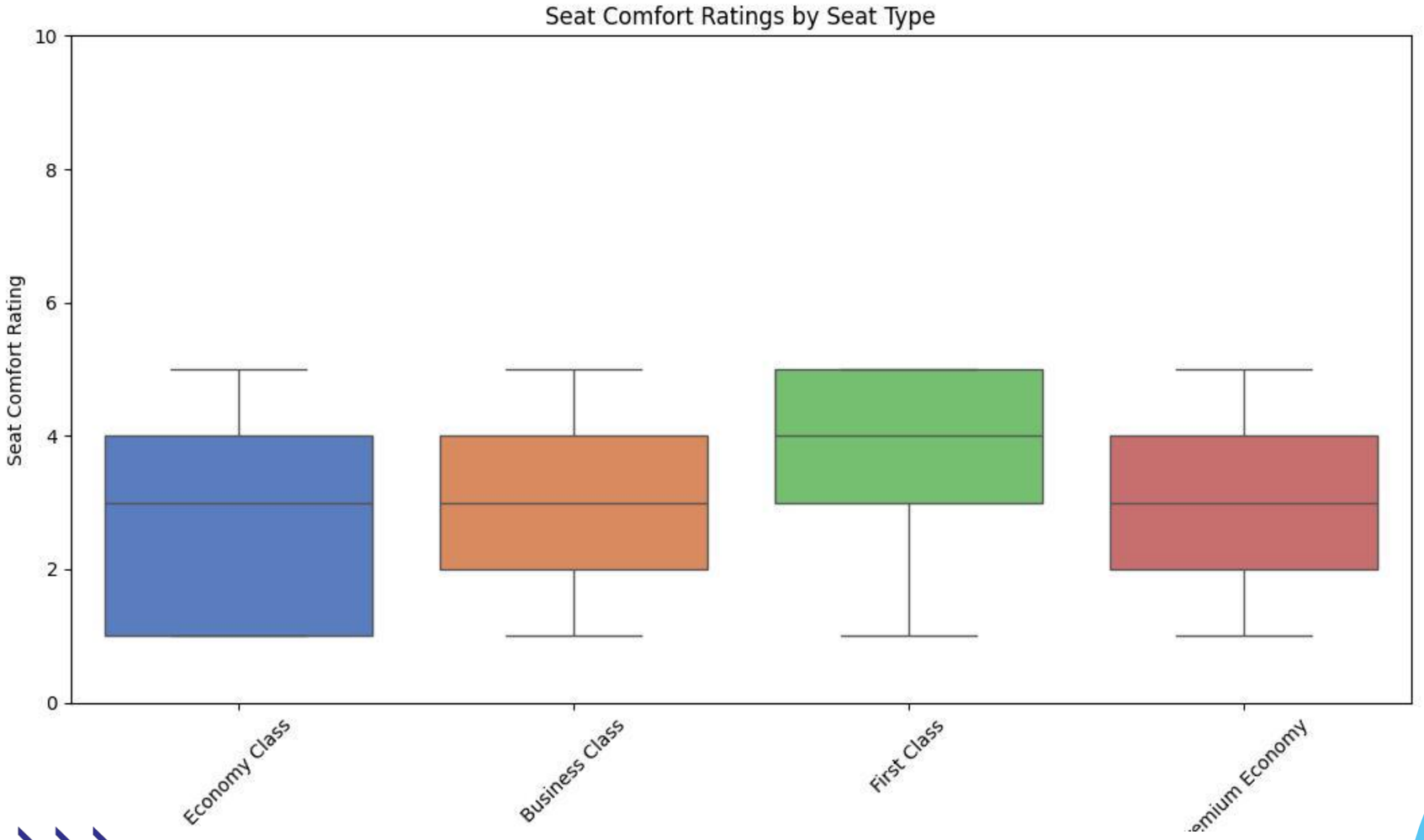
RESULT


Distribution of Overall Ratings
Count
800
700
600
500
400
300
200
100
0
1
2
3
4
5
6
7
8
9
10
Overall Rating

RESULT


Recommendations by Traveller Type
Recommended
False
True
600
500
400
300
200
100
0
Count
1
2
3
4
Traveller Type ID

RESULT


Seat Comfort Ratings by Seat Type
Seat Comfort Rating
10
8
6
4
2
0
Economy Class
Business Class
First Class
emium Economy

THANK YOU