

Exploratory Data analysis on Daily Public Transport Passenger Boardings By Ticket Type

Information about the dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 396 entries, 0 to 395
Data columns (total 3 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   Date        396 non-null   object
1   MyWay       396 non-null   int64
2   Paper Ticket 396 non-null   int64
dtypes: int64(2), object(1)
memory usage: 9.4+ KB
None
```

Filtered dataset for 1 year (2023-04-01 to 2024-04-30)

Data for the period from 1st April 2023 to 30th April 2024:

	Date	MyWay	Paper Ticket
0	2023-04-01	23147	3027
1	2023-04-02	17043	2187
2	2023-04-03	63842	6507
3	2023-04-04	68322	6687
4	2023-04-05	67852	6742
..
391	2024-04-26	49214	5496
392	2024-04-27	24302	3118
393	2024-04-28	18294	2521
394	2024-04-29	57028	5792
395	2024-04-30	74119	9885

[396 rows x 3 columns]

Determining outliers using IQR

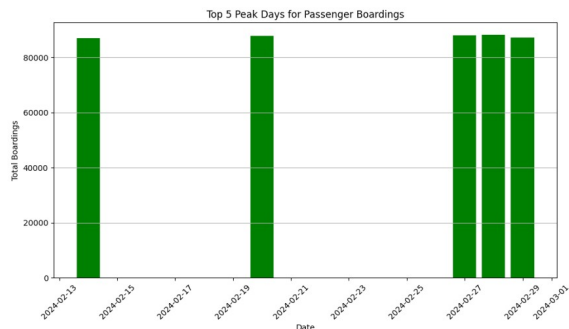
```
Interquartile Range (IQR): 36225.75
Lower Bound for Outliers: -36300.125
Upper Bound for Outliers: 108602.875
Outliers:
Empty DataFrame
Columns: [Date, MyWay, Paper Ticket]
Index: []
```

There is no outliers in the Filtered_dataset.

TOP 5 peak days in the year:

Top 5 Peak Days for Passenger Boardings:

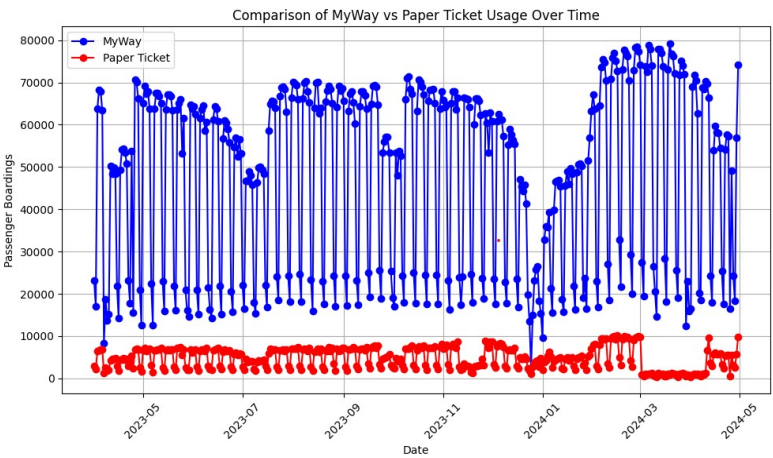
Date	Total Boardings
2024-02-28	88293
2024-02-27	87902
2024-02-26	87801
2024-02-25	87255
2024-02-14	86998



The peak days are not in week ends, they are in week days.

MyWay consistently outperforms paper tickets in terms of popularity. Its average usage surpasses that of paper tickets consistently across the entire timeframe depicted in the graph.

- Both MyWay and paper ticket usage exhibit fluctuations over time, with noticeable peaks and valleys occurring throughout the weeks.

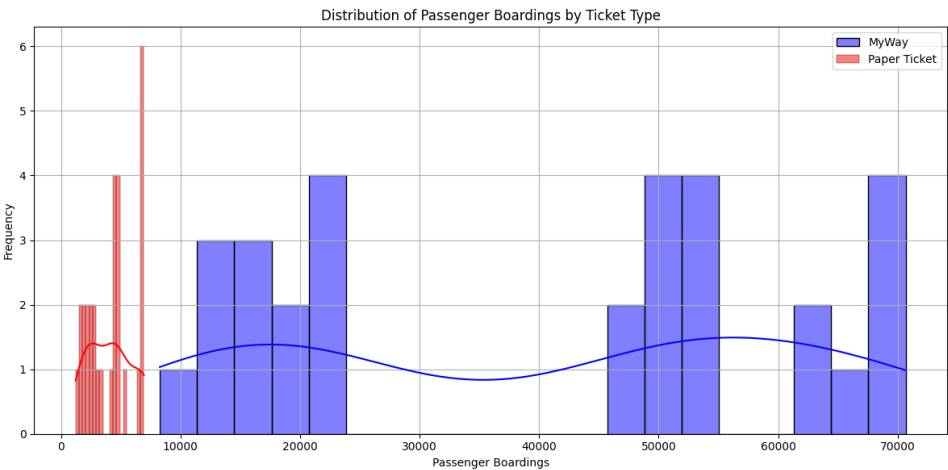


MyWay emerges as the predominant ticket type, evident from the tallest bar on the graph corresponding to its frequency.

- Although MyWay dominates, paper tickets retain significance, as indicated by the second-tallest bar.
- The data underscores MyWay's preference, yet acknowledges the continued usage of traditional paper tickets.

Summary Statistics:

	MyWay	Paper Ticket
count	30.000000	30.000000
mean	40178.700000	4139.566667
std	21662.719293	1864.342071
min	8263.000000	1215.000000
25%	18038.500000	2508.250000
50%	48848.000000	4436.000000
75%	54264.250000	5259.750000
max	70668.000000	6939.000000



Correlation Matrix:

	MyWay	Paper Ticket
MyWay	1.000000	0.970836
Paper Ticket	0.970836	1.000000