

RAILWAY

UK

ANALYSIS



Agenda analyst

- 1 Berapa banyak pembelian online dan offline
- 2 jam sibuk dan hari sibuk (Tertinggi)
- 3 staisun terendah dan tertinggi
- 4 kategori membership pengguna kereta
- 5 Performna Negatif
- 6 alasan delay
- 7 Berapa banyak alasan refund
- 8 performa stasiun positif



Cleaning data



Ubah data yang berantakan menjadi data yang ter-struktur & sistematis, bersihkan dari outliers, huruf berantakan, dan perhatikan data null

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
	Transaction ID,Date of Purchase,Time of Purchase,Purchase Type,Payment Method,Railcard,Ticket Class,Ticket Type,Price,Departure Station,Arrival Destination,Date of Journey,Departure Time,Arrival Time,Actual A																					
1	Transaction ID,Date of Purchase,Time of Purchase,Purchase Type,Payment Method,Railcard,Ticket Class,Ticket Type,Price,Departure Station,Arrival Destination,Date of Journey,Departure Time,Arrival Time,Actual A	bda8a6ba-b3dc-4677-b176,2023-12-08,12:41:11,Online,Contactless,Adult,Standard,Advance,43,London Paddington,Liverpool Lime Street,2024-01-01,11:00:00,13:30:00,On Time,,No	b0cdd1b0-f214-4197-be53,2023-12-16,11:23:01,Station,Credit Card,Adult,Standard,Advance,23,London Kings Cross,York,2024-01-01,09:45:00,11:35:00,11:40:00,Delayed,Signal Failure,No	f3ba7a96-f713-40d9-9629,2023-12-19,19:51:27,Online,Credit Card,None,Standard,Advance,3,Liverpool Lime Street,Manchester Piccadilly,2024-01-02,18:15:00,18:45:00,18:45:00,On Time,,No	b2471f11-4fe7-4c87-8ab4,2023-12-20,23:00:36,Station,Credit Card,None,Standard,Advance,13,London Paddington,Reading,2024-01-01,21:30:00,22:30:00,22:30:00,On Time,,No	2be00b45-0762-485e-a7a3,2023-12-27,18:22:56,Online,Contactless,None,Standard,Advance,76,Liverpool Lime Street,London Euston,2024-01-01,16:45:00,19:00:00,19:00:00,On Time,,No	4e1cd88e-3d95-44ef-99fa,2023-12-30,07:56:06,Online,Credit Card,None,Standard,Advance,35,London Kings Cross,York,2024-01-01,06:15:00,08:05:00,08:05:00,On Time,,No	1c74479d-85a4-4ba1-a607,2023-12-31,00:02:01,Station,Credit Card,Adult,Standard,Advance,2,London Euston,Oxford,2024-01-01,22:30:00,23:40:00,23:40:00,On Time,,No	febfb8ab-f808-46fa-bf2b,2023-12-31,01:35:18,Station,Contactless,Disabled,Standard,Advance,2,Liverpool Lime Street,Manchester Piccadilly,2024-01-01,00:00:00,00:30:00,00:30:00,On Time,,No	01df916f-4291-41ec-a37d,2023-12-31,01:43:09,Station,Credit Card,None,Standard,Advance,37,London Euston,York,2024-01-01,00:00:00,01:50:00,02:07:00,Delayed,Signal Failure,No	a8cedeba-1923-459d-b046,2023-12-31,03:05:52,Online,Credit Card,None,Standard,Advance,13,London Paddington,Reading,2024-01-01,01:30:00,02:30:00,02:30:00,On Time,,No	b3e5ca7d-e76c-49f2-b49f,2023-12-31,03:26:37,Online,Contactless,None,Standard,Advance,8,York,Durham,2024-01-01,01:45:00,02:35:00,02:35:00,On Time,,No	6c63f7ac-d590-4356-9eaa,2023-12-31,03:52:11,Online,Contactless,Adult,Standard,Advance,8,London Paddington,Reading,2024-01-01,02:15:00,03:15:00,03:15:00,On Time,,No	2e7add75-5664-41aa-9468,2023-12-31,05:55:22,Online,Contactless,None,Standard,Advance,3,Manchester Piccadilly,Liverpool Lime Street,2024-01-01,04:15:00,04:45:00,04:45:00,On Time,,No	7ed9b545-eb6f-49b2-9b5a,2023-12-31,06:44:35,Online,Contactless,None,Standard,Advance,3,Manchester Piccadilly,Liverpool Lime Street,2024-01-01,05:00:00,05:30:00,05:30:00,On Time,,No	2e05e2a6-88a4-40fb-bacc,2023-12-31,08:05:50,Online,Credit Card,Disabled,Standard,Advance,15,Birmingham New Street,London St Pancras,2024-01-01,06:30:00,07:50:00,07:50:00,On Time,,No	8a18d3b4-995e-49bf-93a3,2023-12-31,08:16:53,Online,Credit Card,None,Standard,Advance,13,London Paddington,Reading,2024-01-01,07:45:00,08:45:00,08:45:00,On Time,,No	7493a611-342a-4b17-90dc,2023-12-31,08:23:15,Online,Credit Card,None,Standard,Advance,13,London Paddington,Reading,2024-01-01,07:45:00,08:45:00,08:45:00,On Time,,No	054676ac-a976-4909-a26e,2023-12-31,09:09:20,Online,Credit Card,None,Standard,Advance,8,London St Pancras,Birmingham New Street,2024-01-01,07:30:00,08:50:00,08:50:00,On Time,,No	6b62b452-c491-468d-b39c,2023-12-31,09:12:21,Online,Credit Card,None,Standard,Advance,35,London Kings Cross,York,2024-01-01,07:30:00,09:20:00,09:20:00,On Time,,No	85e38992-6c6c-4569-914e,2023-12-31,10:42:22,Online,Credit Card,None,Standard,Advance,35,London Kings Cross,York,2024-01-01,09:00:00,10:50:00,10:50:00,On Time,,No	8dfbf0fc-aea0-424f-b30e,2023-12-31,11:57:15,Station,Debit Card,Adult,Standard,Advance,7,Birmingham New Street,Manchester Piccadilly,2024-01-01,11:15:00,12:35:00,13:06:00,Delayed,Technical Issue,Yes



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Trans -	Date of Purchase -	Time -	Purchase Type -	Payment Method -	Radical -	Ticket Class -	Ticket Type -	Price -	Departure Station -	Arrival Destination -	Date of Journey -	Departure Time -	Arrival -	Actual -	Journey -	Reason for Delay -	Refund Request -	
2	d4d9b61	08/12/2023	12:41:11	Online	Contactless	Adult	Standard	Advance	43.00	London Paddington	Liverpool Lime Street	01/01/2024	11:00:00	13:30:00	13:30:00	On Time	No		
3	b0ddfb1b	16/12/2023	11:23:01	Station	Credit Card	Adult	Standard	Advance	23.00	London Kings Cross	York	01/01/2024	09:45:00	11:35:00	11:40:00	Delayed	Signal Failure		
4	f3ba7a9	13/12/2023	19:51:27	Online	Credit Card	None	Standard	Advance	3.00	Liverpool Lime Street	Manchester Piccadilly	02/01/2024	18:15:00	18:45:00	18:45:00	On Time	No		
5	b247f911	20/12/2023	*****	Station	Credit Card	None	Standard	Advance	13.00	London Paddington	Reading	01/01/2024	21:30:00	*****	*****	On Time	No		
6	2b0d4e9	27/12/2023	18:22:56	Online	Contactless	None	Standard	Advance	7.60	Liverpool Lime Street	London Euston	01/01/2024	19:45:00	19:00:00	19:00:00	On Time	No		
7	4e1ddc8	30/12/2023	18:00:00	Online	Credit Card	None	Standard	Advance	35.00	London Kings Cross	York	01/01/2024	08:15:00	*****	*****	On Time	No		
8	1744749	31/12/2023	00:02:01	Station	Credit Card	Adult	Standard	Advance	2.00	London Euston	Oxford	01/01/2024	22:30:00	*****	*****	On Time	No		
9	1eb78dab	31/12/2023	01:35:18	Station	Contactless	Disabled	Standard	Advance	2.00	Liverpool Lime Street	Manchester Piccadilly	01/01/2024	00:00:00	*****	*****	On Time	No		
10	01d3f61	31/12/2023	01:43:09	Station	Credit Card	None	Standard	Advance	37.00	London Euston	York	01/01/2024	00:00:00	01:50:00	01:50:00	Delayed	Signal Failure		
11	a8cebd8	31/12/2023	*****	Online	Credit Card	None	Standard	Advance	13.00	London Paddington	Reading	01/01/2024	01:30:00	*****	*****	On Time	No		
12	b3e5ca1	31/12/2023	*****	Online	Contactless	None	Standard	Advance	8.00	York	Durham	01/01/2024	01:45:00	*****	*****	On Time	No		
13	b6337a	31/12/2023	03:52:11	Online	Contactless	Adult	Standard	Advance	8.00	London Paddington	Reading	01/01/2024	02:15:00	03:15:00	03:15:00	On Time	No		
14	2e7add1	31/12/2023	*****	Online	Contactless	None	Standard	Advance	3.00	Manchester Piccadilly	Liverpool Lime Street	01/01/2024	04:15:00	*****	*****	On Time	No		
15	7ed3b54	31/12/2023	*****	Online	Contactless	None	Standard	Advance	3.00	Manchester Piccadilly	Liverpool Lime Street	01/01/2024	05:00:00	*****	*****	On Time	No		
16	e05e2d1	31/12/2023	*****	Online	Credit Card	Disabled	Standard	Advance	15.00	Birmingham New Street	London St Pancras	01/01/2024	06:30:00	*****	*****	On Time	No		
17	8a183b3	31/12/2023	08:16:53	Online	Credit Card	None	Standard	Advance	13.00	London Paddington	Reading	01/01/2024	07:45:00	*****	*****	On Time	No		
18	7493aa1	31/12/2023	08:23:15	Online	Credit Card	None	Standard	Advance	13.00	London Paddington	Reading	01/01/2024	07:45:00	*****	*****	On Time	No		
19	0546781	31/12/2023	*****	Online	Credit Card	None	Standard	Advance	8.00	London Paddington	Birmingham New Street	01/01/2024	07:30:00	*****	*****	On Time	No		
20	b6b2645	31/12/2023	09:12:21	Online	Credit Card	None	Standard	Advance	35.00	London Kings Cross	York	01/01/2024	07:30:00	*****	*****	On Time	No		
21	8543695	31/12/2023	10:42:22	Online	Credit Card	None	Standard	Advance	35.00	London Kings Cross	York	01/01/2024	09:00:00	10:50:00	10:50:00	On Time	No		
22	8dbbf0f6	31/12/2023	11:57:15	Station	Debit Card	Adult	Standard	Advance	7.00	Birmingham New Street	Manchester Piccadilly	01/01/2024	11:15:00	12:35:00	13:06:00	Delayed	Technical Issue		
23	a478935	31/12/2023	12:11:47	Station	Credit Card	Disabled	Standard	Advance	2.00	Liverpool Lime Street	Manchester Piccadilly	01/01/2024	10:30:00	11:00:00	11:00:00	On Time	No		
24	d4e4b34	31/12/2023	13:33:39	Online	Credit Card	Senior	Standard	Advance	23.00	London Kings Cross	York	01/01/2024	09:30:00	11:20:00	11:20:00	On Time	No		
25	8912601	31/12/2023	14:23:09	Station	Contactless	None	Standard	Advance	7.00	London Euston	Birmingham New Street	01/01/2024	12:45:00	14:05:00	14:05:00	On Time	No		
26	f000134	31/12/2023	14:53:44	Online	Credit Card	None	Standard	Advance	7.00	London Euston	Birmingham New Street	01/01/2024	13:15:00	14:35:00	14:35:00	On Time	No		
27	9424a49	31/12/2023	15:19:53	Online	Contactless	None	Standard	Advance	86.00	Manchester Piccadilly	London Paddington	01/01/2024	13:45:00	16:00:00	16:00:00	On Time	No		
28	74462221	31/12/2023	15:53:46	Online	Credit Card	Senior	First Class	Advance	34.00	Oxford	Bristol Temple Meads	01/01/2024	14:15:00	15:30:00	15:54:00	Delayed	Signal Failure		
29	7267e26	31/12/2023	16:44:05	Station	Credit Card	None	Standard	Advance	13.00	London Paddington	Reading	01/01/2024	15:00:00	16:00:00	16:00:00	On Time	No		
30	3c375a4	31/12/2023	18:01:58	Station	Contactless	None	Standard	Advance	7.00	London Euston	Birmingham New Street	01/01/2024	16:30:00	17:50:00	17:50:00	On Time	No		
31	5b06388	31/12/2023	18:39:00	Online	Credit Card	None	First Class	Advance	57.00	London Kings Cross	York	01/01/2024	17:00:00	18:50:00	18:50:00	On Time	No		
32	e406161	31/12/2023	19:33:15	Station	Debit Card	None	Standard	Advance	8.00	London St Pancras	Birmingham New Street	01/01/2024	17:45:00	19:05:00	19:05:00	On Time	No		
33	401046	*****	19:45:52	Online	Credit Card	None	Standard	Advance	13.00	Liverpool Lime Street	Reading	01/01/2024	18:15:00	19:35:00	19:15:00	On Time	No		

from this

to this

kategorikan juga beserta tipe datanya, dari setiap kolom kategori

explore data

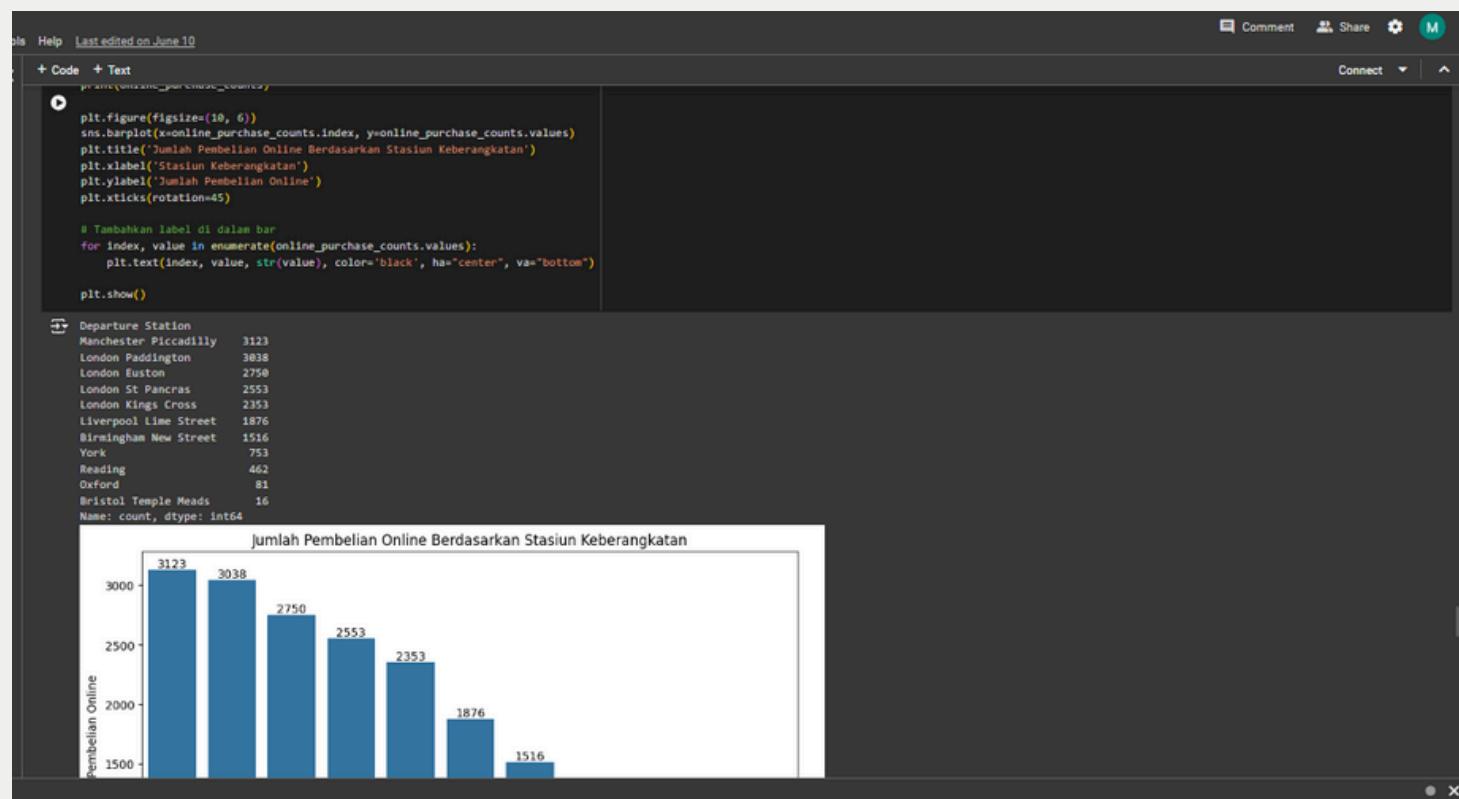
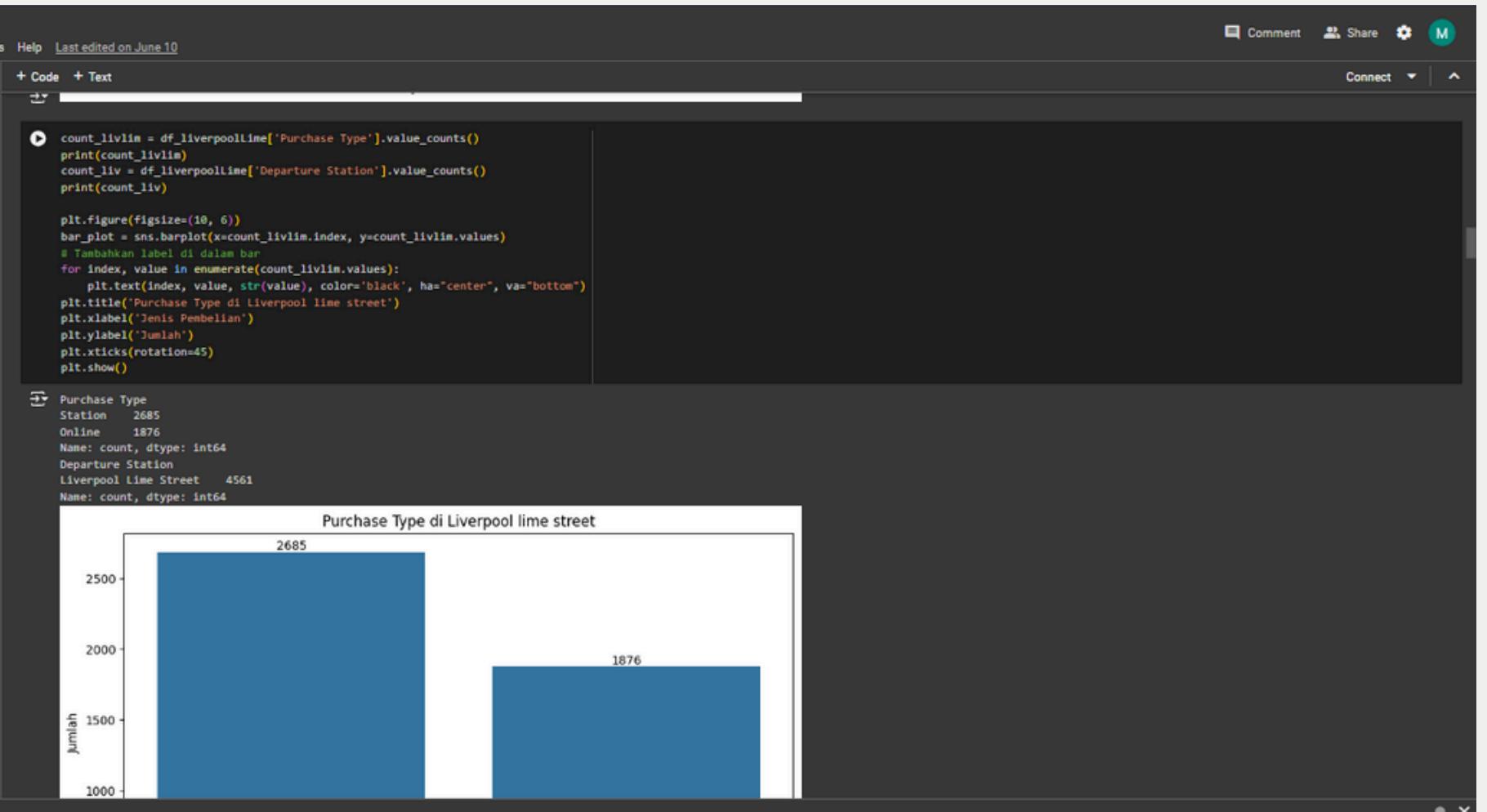


```
Tools Help Last edited on June 10
+ Code + Text
Comment Share M
Import Library
[ ] import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
import statsmodels.api as sm
import scipy.stats as stats

Upload data excel
[ ] from google.colab import files
uploaded = files.upload()
Choose File: No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving railwayuk.xlsx to railwayuk.xlsx

Membaca file
[ ] df = pd.read_excel('railwayuk.xlsx')

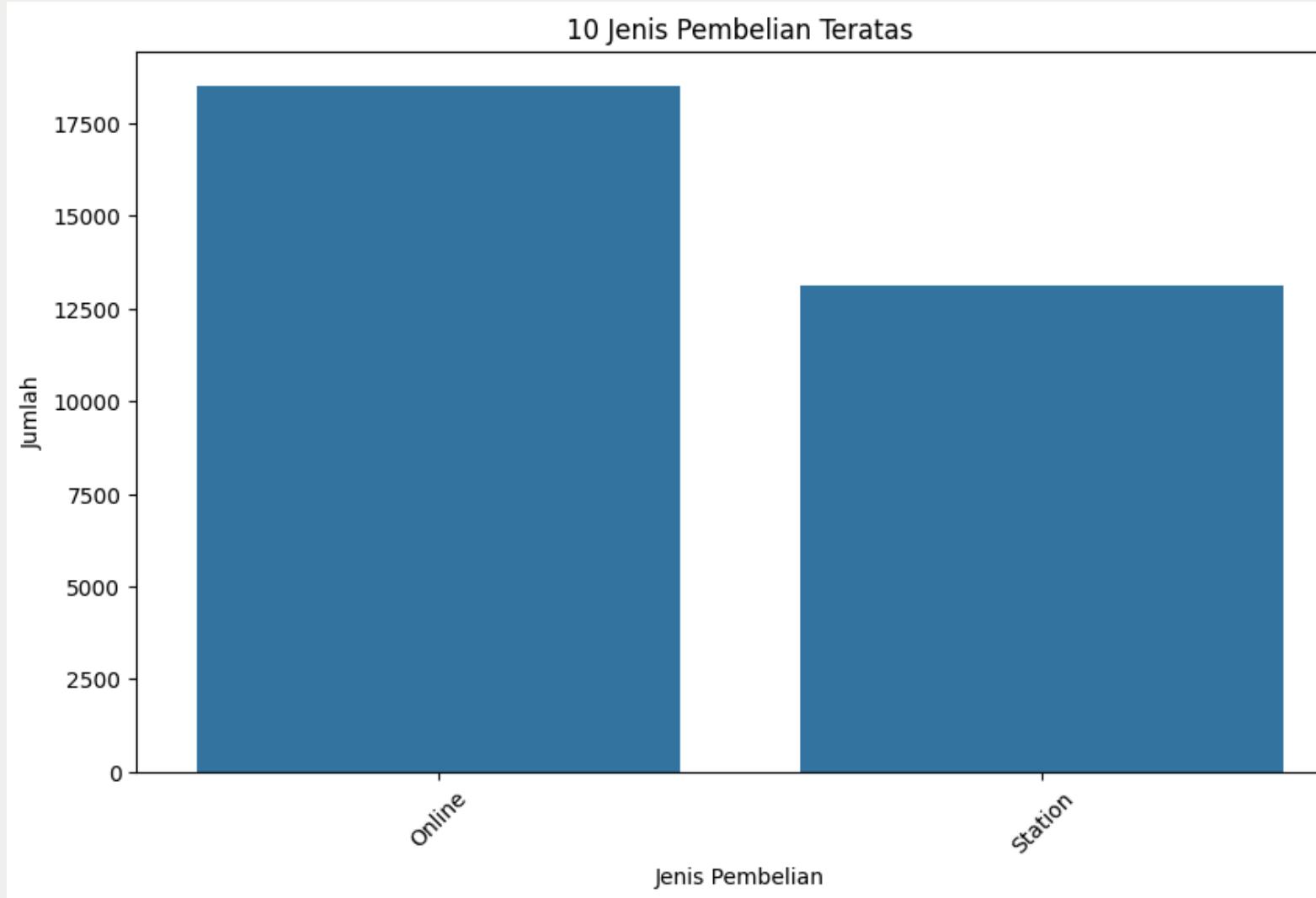
Filter station
[ ] df_londonPad = df[df['Departure Station'] == 'London Paddington']
df_manchesterPic = df[df['Departure Station'] == 'Manchester Picadilly']
df_londonEus = df[df['Departure Station'] == 'London Euston']
df_liverpoolLime = df[df['Departure Station'] == 'Liverpool Lime Street']
df_londonKings = df[df['Departure Station'] == 'London Kings Cross']
df_londonPan = df[df['Departure Station'] == 'London St Pancras']
df_birminghamNew = df[df['Departure Station'] == 'Birmingham New Street']
df_york = df[df['Departure Station'] == 'York']
df_reading = df[df['Departure Station'] == 'Reading']
```



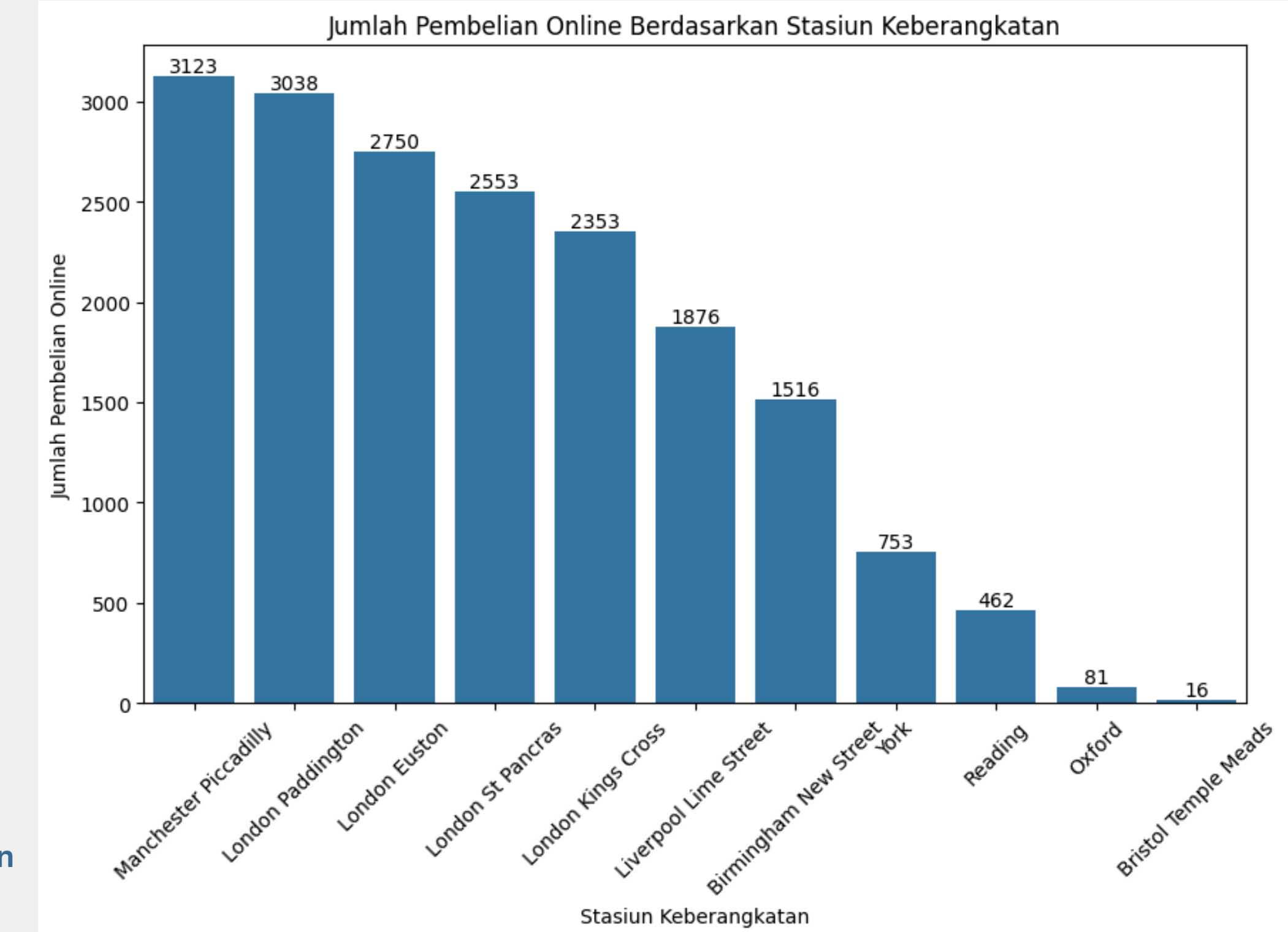
Eksplor data dengan sintax bahasa program python, melihat sub insight, measures, dan count dari beberapa data yang di cari

1

Berapa banyak pembelian online dan offline



- pembelian di dominasi oleh pembelian ticket kereta via online di atas 1700+ pembelian ticket, sedangkan untuk pembelian ticket di stasiun mencapai 1250+-, selisih 40%--
- untuk Pembelian tiket online kereta terbanyak di keberangkatan stasiun manchester picadilly, dan terendah di stasiun Bristol temple meads

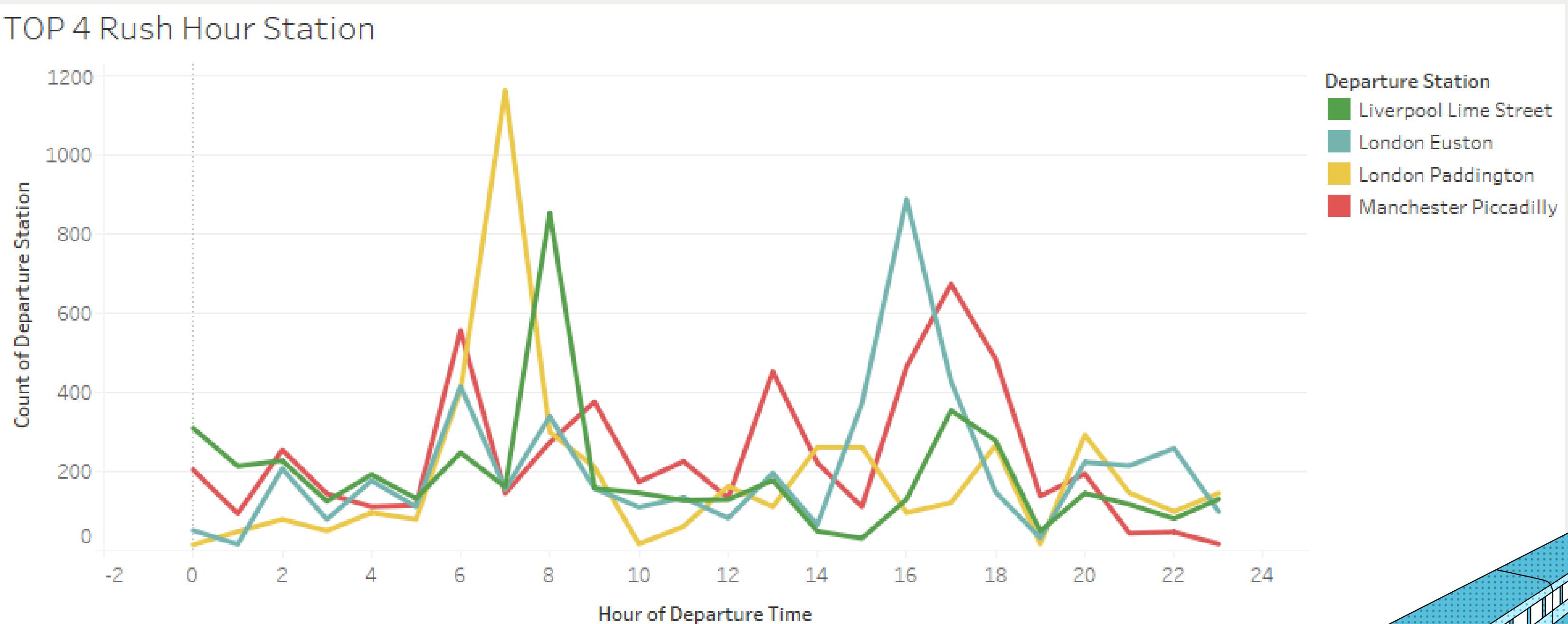


2

jam sibuk dan hari sibuk
(Tertinggi)

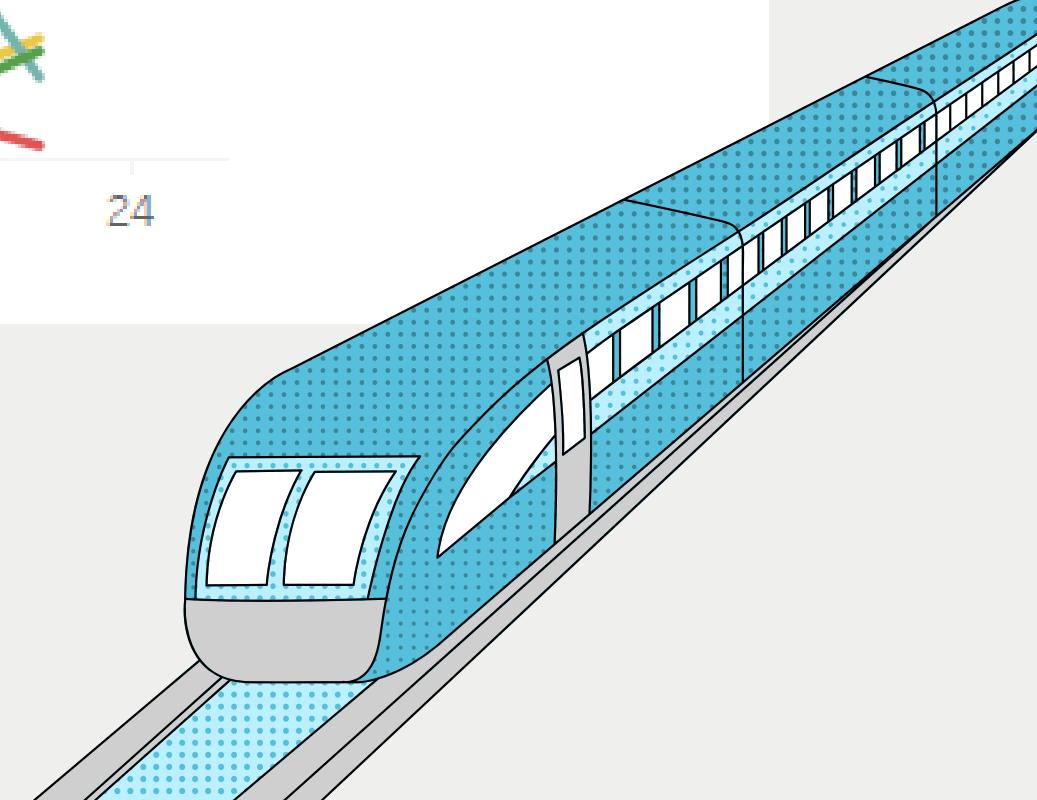
Kategori pertama 4 stasiun dengan jumlah pemberangkatan terbanyak di UK (top 4 Rush Hour)

TOP 4 Rush Hour Station



Tren Jam sibuk Top 4 Stasiun Teramai Di Inggris januari-April 2024:

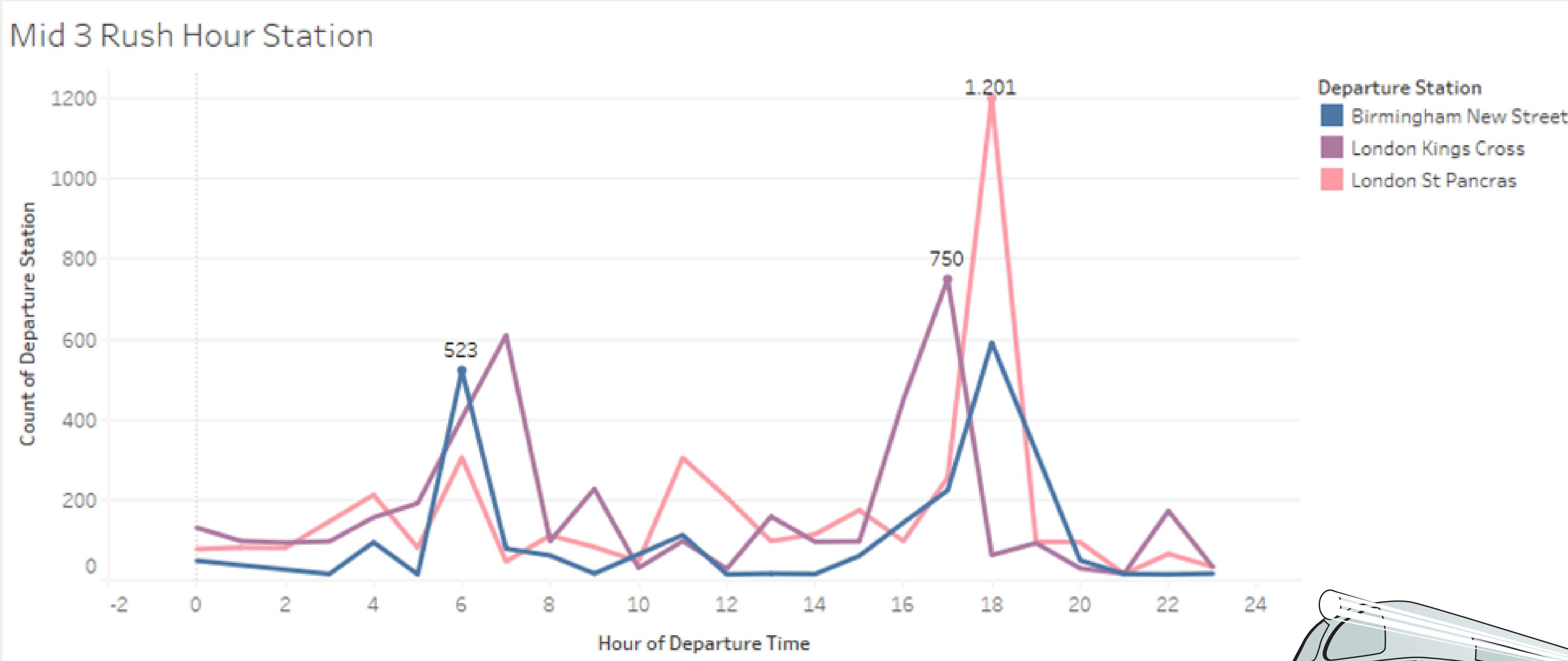
1. Manchester picadilly :5650 Keberangkatan : Jam sibuk Keberangkatan Jam 17
2. London Euston :4954 Keberangkatan : Jam sibuk Keberangkatan Jam 16
3. Liverpool Lime Street : 4561 Keberangkatan : Jam sibuk Keberangkatan Jam 8
4. London Paddington :4500 Keberangkatan : Jam sibuk Keberangkatan Jam 7



2

jam sibuk dan hari sibuk
(Tertinggi)

Kategori pertama 4 stasiun dengan jumlah pemberangkatan terbanyak di UK (Mid 3 Rush Hour)

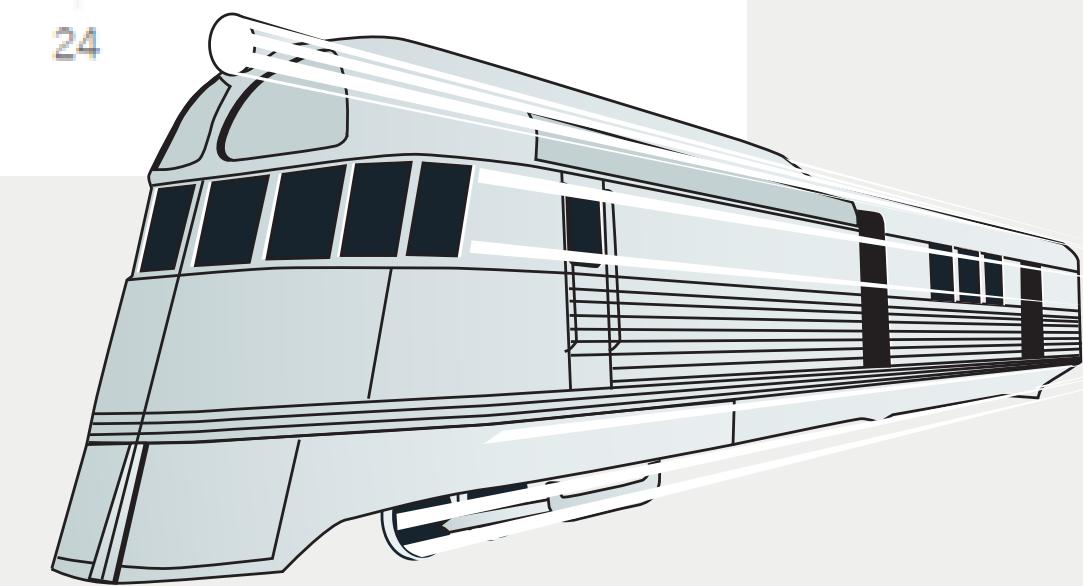


Tren Jam sibuk Mid 3 Stasiun Teramai Di Inggris januari-April 2024:

5. Londong Kings Cross :4229 Keberangkatan : Jam sibuk Keberangkatan Jam 17

6. London St Pancras :3891 Keberangkatan : Jam sibuk Keberangkatan Jam 18

7. Birmingham New Street : 2136 Keberangkatan : Jam sibuk Keberangkatan Jam 6

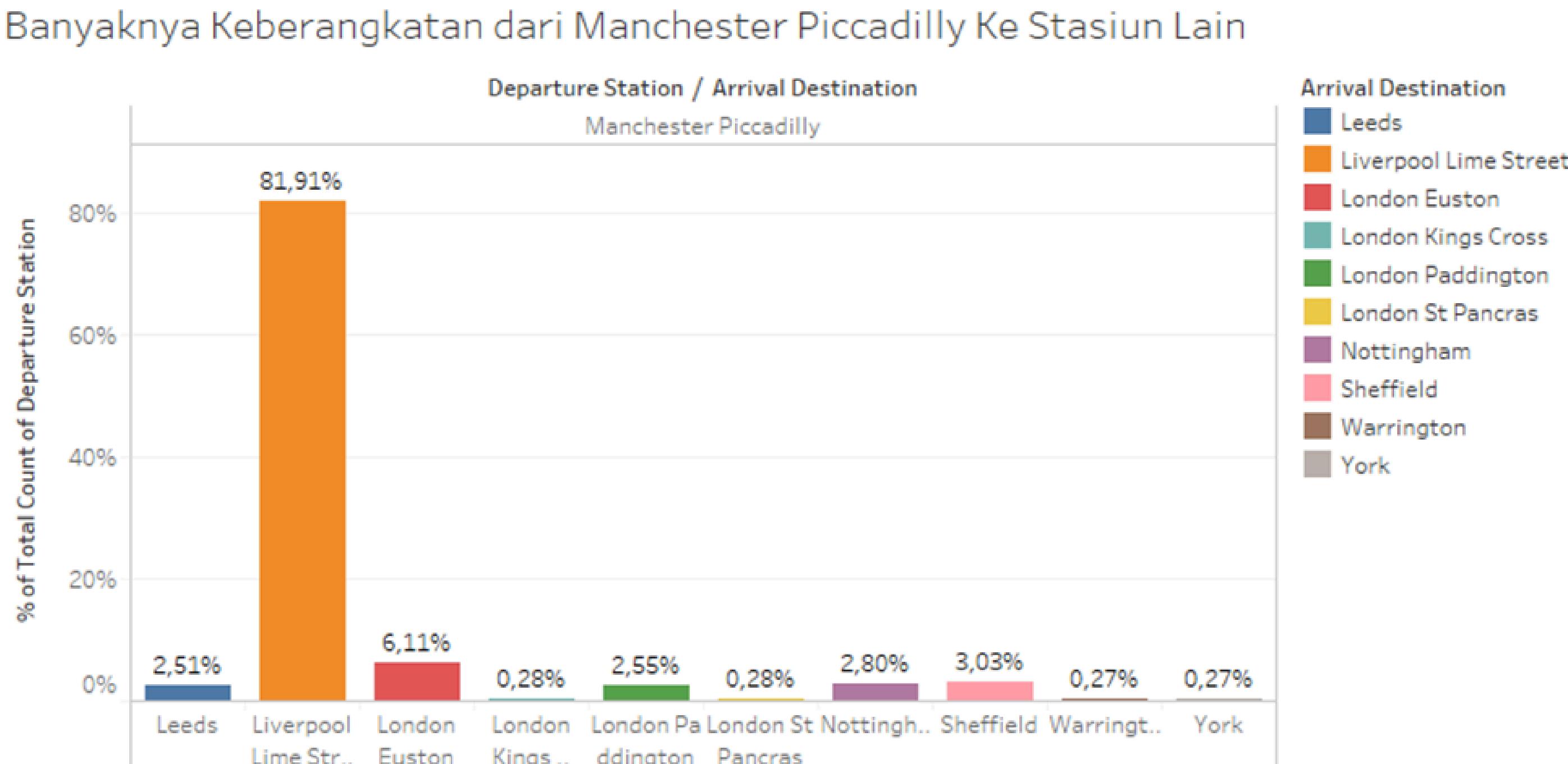


3

stasiun pemberangkatan
terendah dan tertinggi

Stasiun Manchester Picadilly

Banyaknya Keberangkatan dari Manchester Piccadilly Ke Stasiun Lain



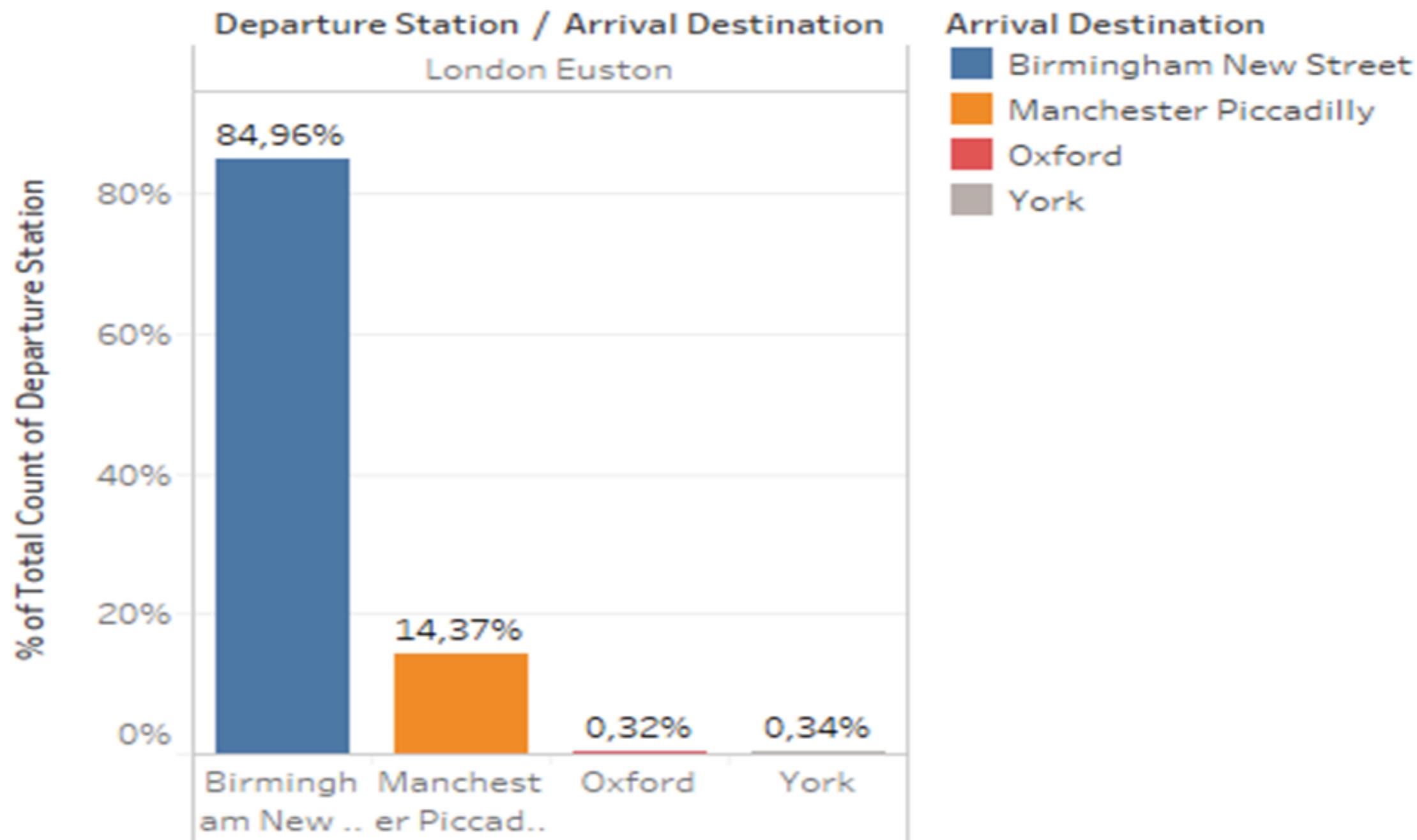
Pemberangkatan dari stasiun manchester picadilly, terbanyak menuju kedatangan ke stasiun liverpool lime street dengan persentase 81,91%

3

stasiun pemberangkatan terendah dan tertinggi

Stasiun London Euston

Banyaknya Keberangkatan dari London Euston Ke Stasiun Lain



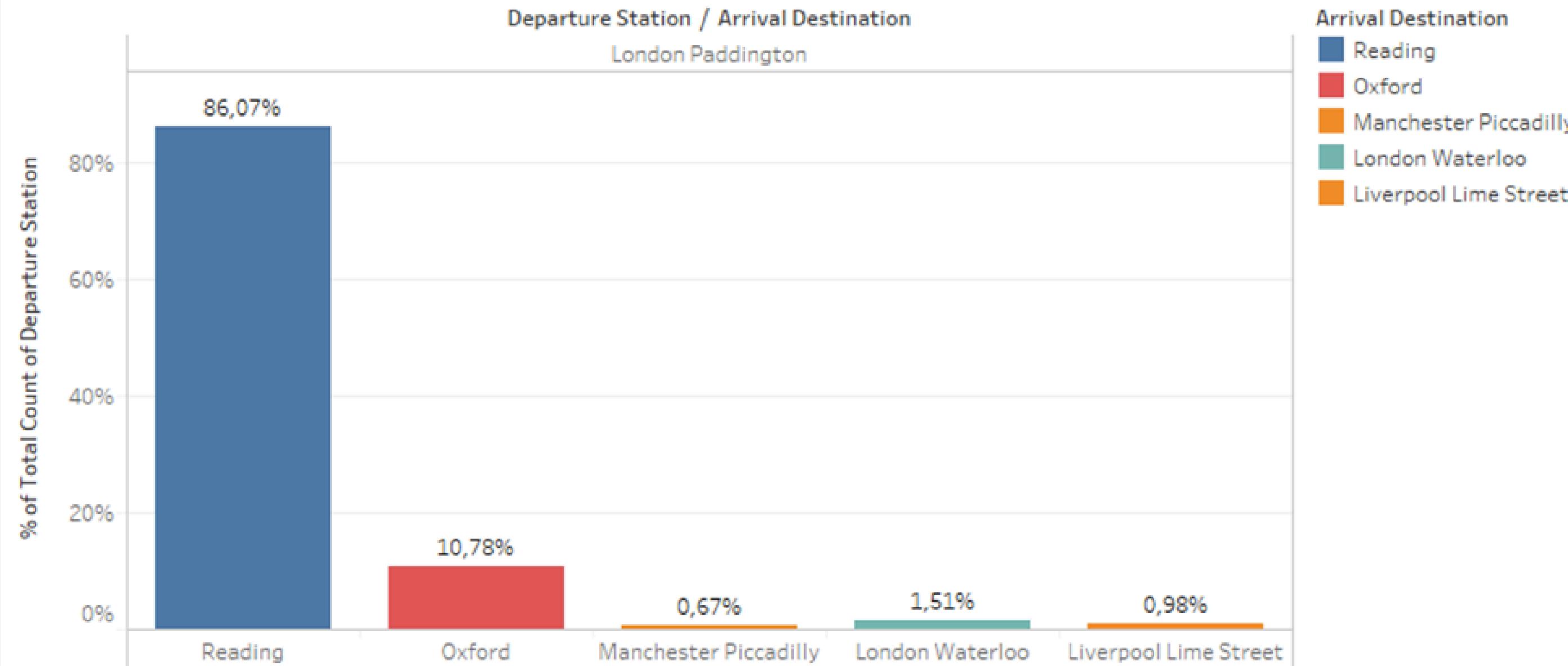
Pemberangkatan dari stasiun London Euston, terbanyak menuju kedatangan ke stasiun Birmingham New Street dengan persentase 84,96%

3

stasiun pemberangkatan
terendah dan tertinggi

Stasiun London Paddington

Banyaknya Keberangkatan dari London Paddington Ke Stasiun Lain



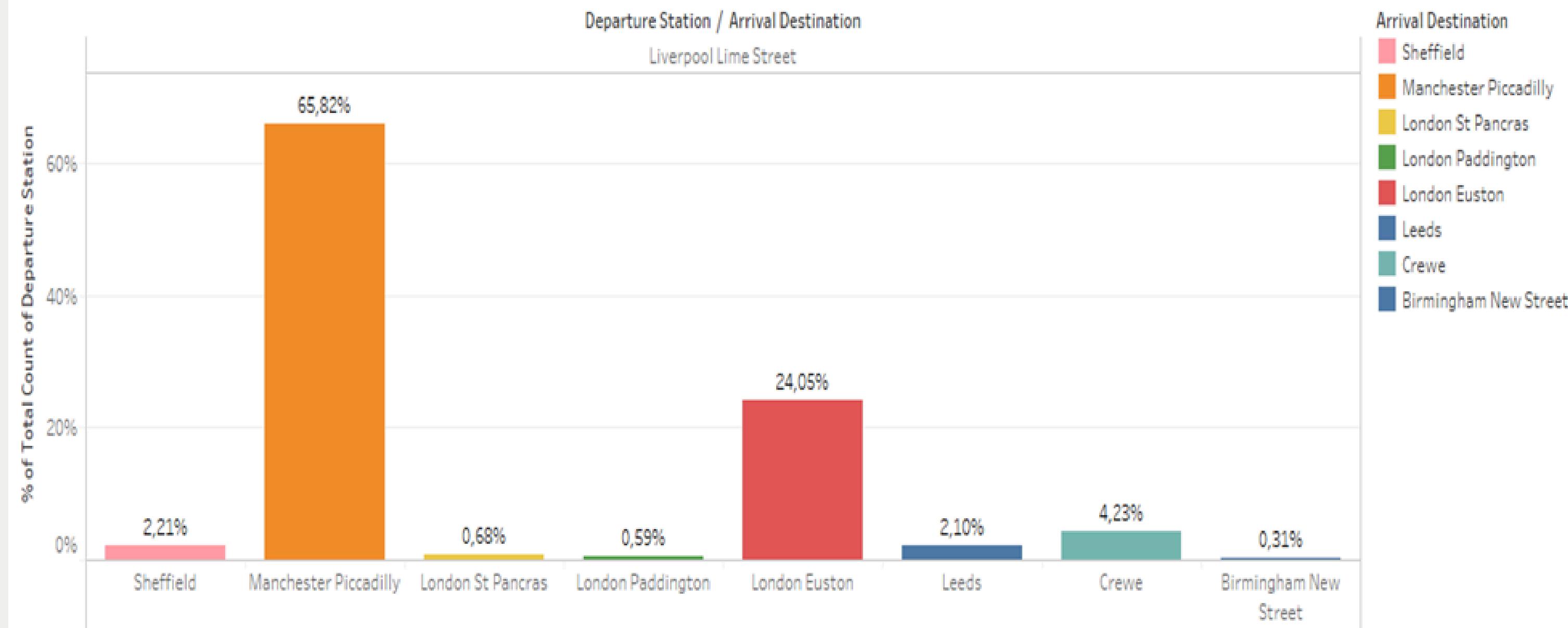
Pemberangkatan dari stasiun London Paddington, terbanyak menuju kedatangan ke stasiun Reading dengan persentase 86,07%

3

stasiun pemberangkatan
terendah dan tertinggi

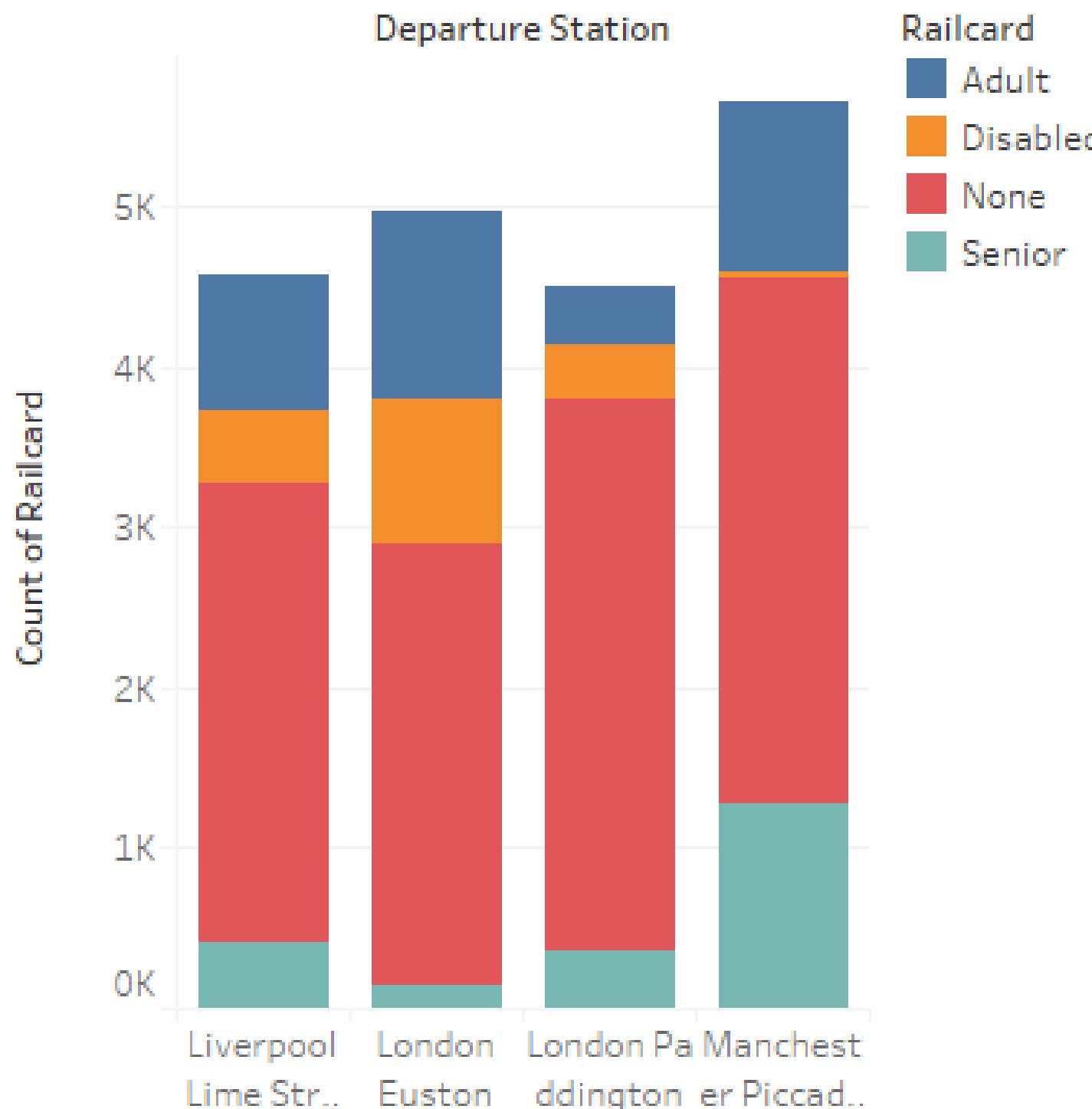
Stasiun Liverpool Lime Street

Banyaknya Keberangkatan dari Liverpool Lime Street Ke Stasiun Lain



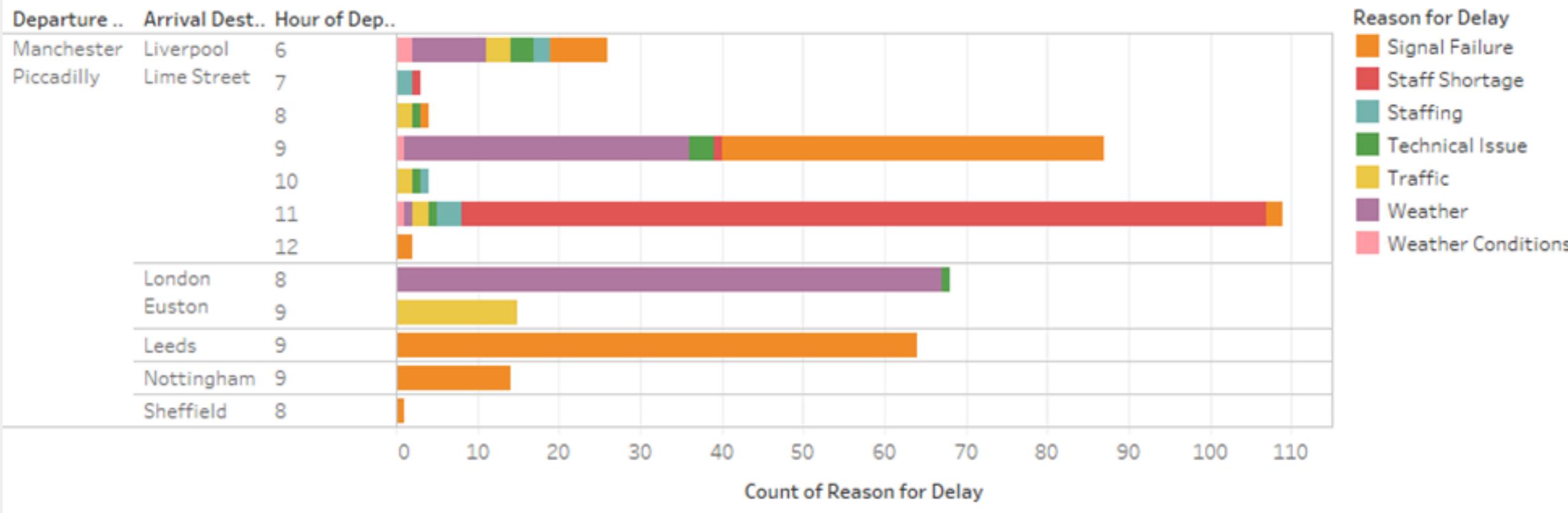
Pemberangkatan dari stasiun Liverpool Lime Street, terbanyak menuju kedatangan ke stasiun Manchester Piccadilly dengan persentase 66,82%

Penumpang Berdasarkan keterangan kartu member



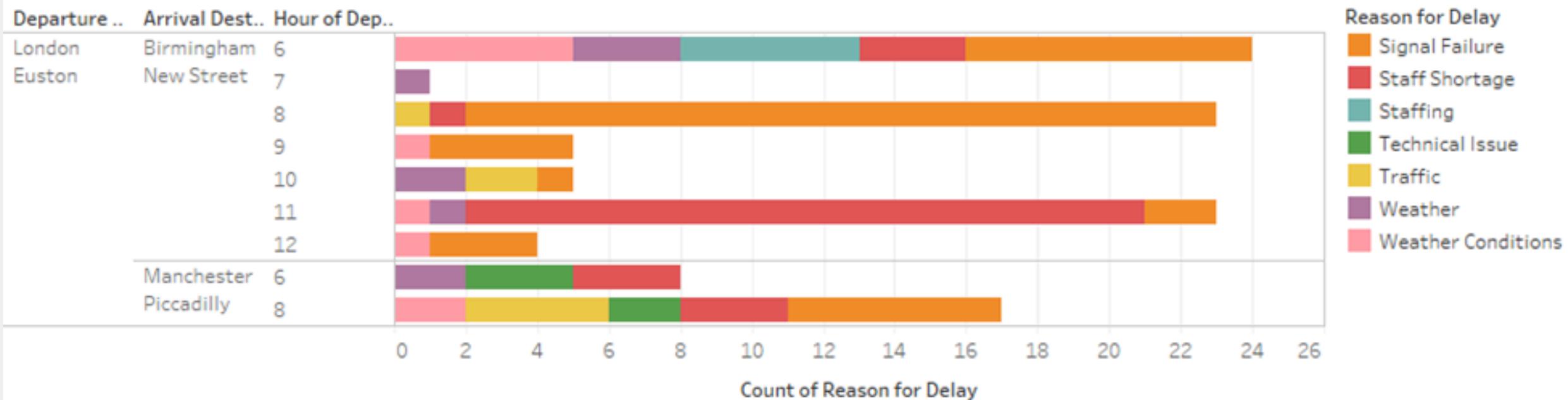
- dapat disimpulkan bahwa kategori pengguna kereta terbanyak oleh kategori non (kanak-kanak 5-11 Tahun & Remaja 12-25 Tahun)
- Untuk pengguna disabilitas kereta, cukup banyak di minati di stasiun london euston
- dan kategori dewasa & lanjut usia cukup merata di keempat stasiun

Performa Delay Stasiun Manchester Picadilly Base on Morning Departure 6-12 A.M



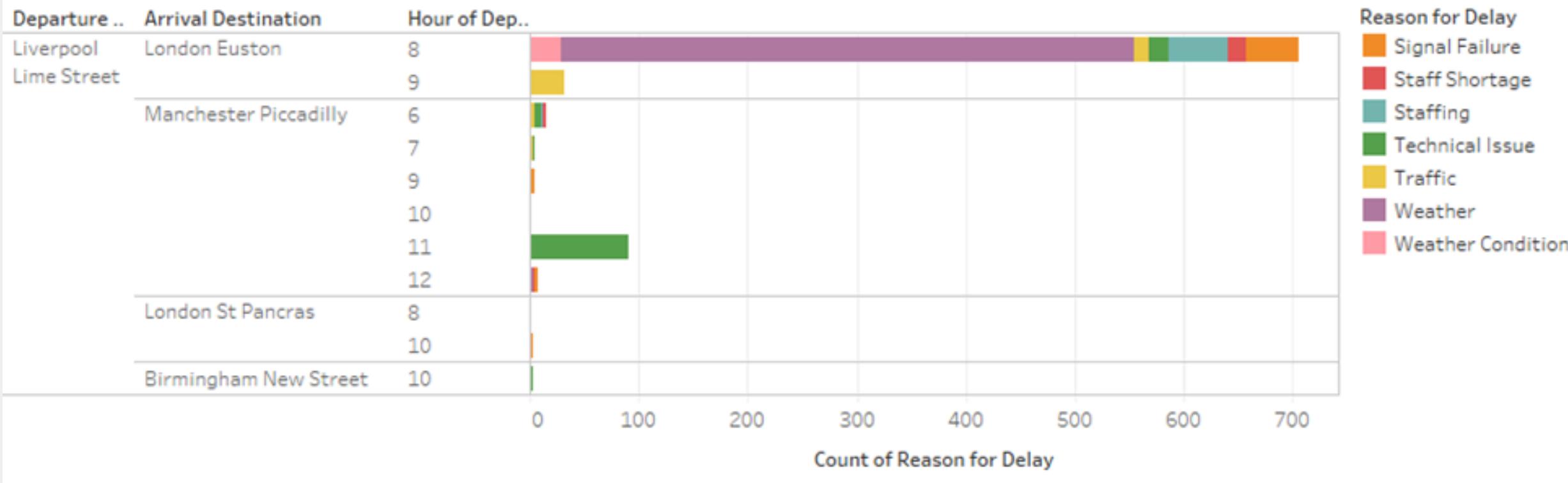
- Stasiun Manchester Picadilly banyak terhambat (Delay) karena kurangnya sumber daya manusia (man power), kemungkinan tidak sebanding antara staff dan tingkat overload penumpang, kebanyakan saat hendak menuju stasiun liverpool lime street. dan terjadi pada pukul 09.00-- pagi

Performa Delay Stasiun London Euston Base on Morning Departure 6-12 A.M



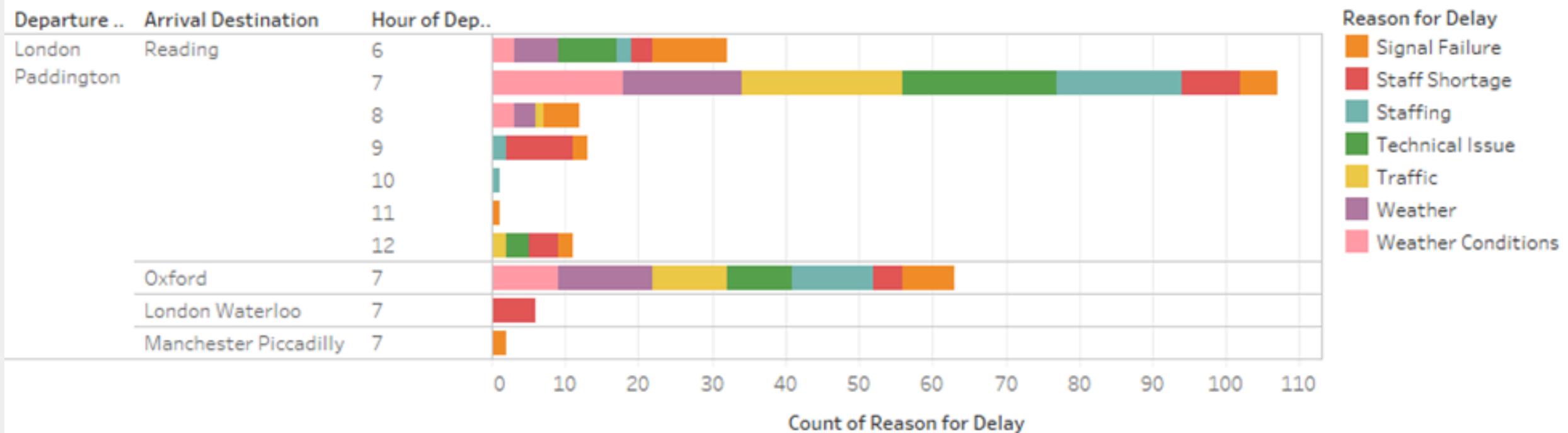
- Stasiun London euston banyak keterlambatan kebanyakan dikarenakan kegagalan sinyal kereta baik menuju birminghman new street maupun manchester picadilly, dan kebanyakan terjadi di jam 08.00-- pagi

Performa Delay Stasiun Liverpool Lime Street Base on Morning Departure 6-12 A.M



- Stasiun Liverpool lime street kebanyakan mengalami keterlambatan paling banyak pada pukul 08.00+ saat hendak menuju london euston karena cuaca buruk

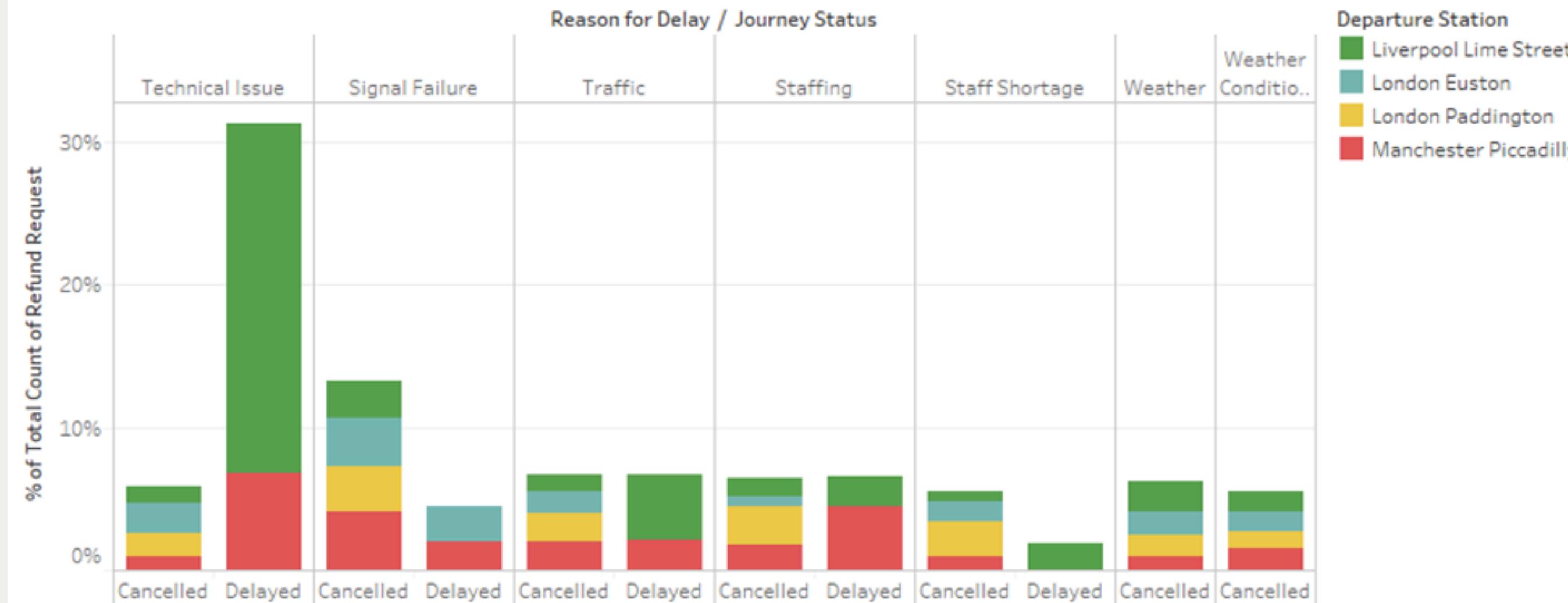
Performa Delay Stasiun London Paddington Base on Morning Departure 6-12 A.M



- Stasiun London Paddington mengalami kendala yang beragam kebanyakan pukul 07.00 Pagi secara merata baik dari masalah kondisi cuaca, cuacu buruk, masalah lalu lintas, masalah teknis, teknis pegawai,& kurangnya SDM

Dikarenakan performa negatif sering terjadi maka akan di kaitkan dengan variabel permintaan refund dari pengguna kereta

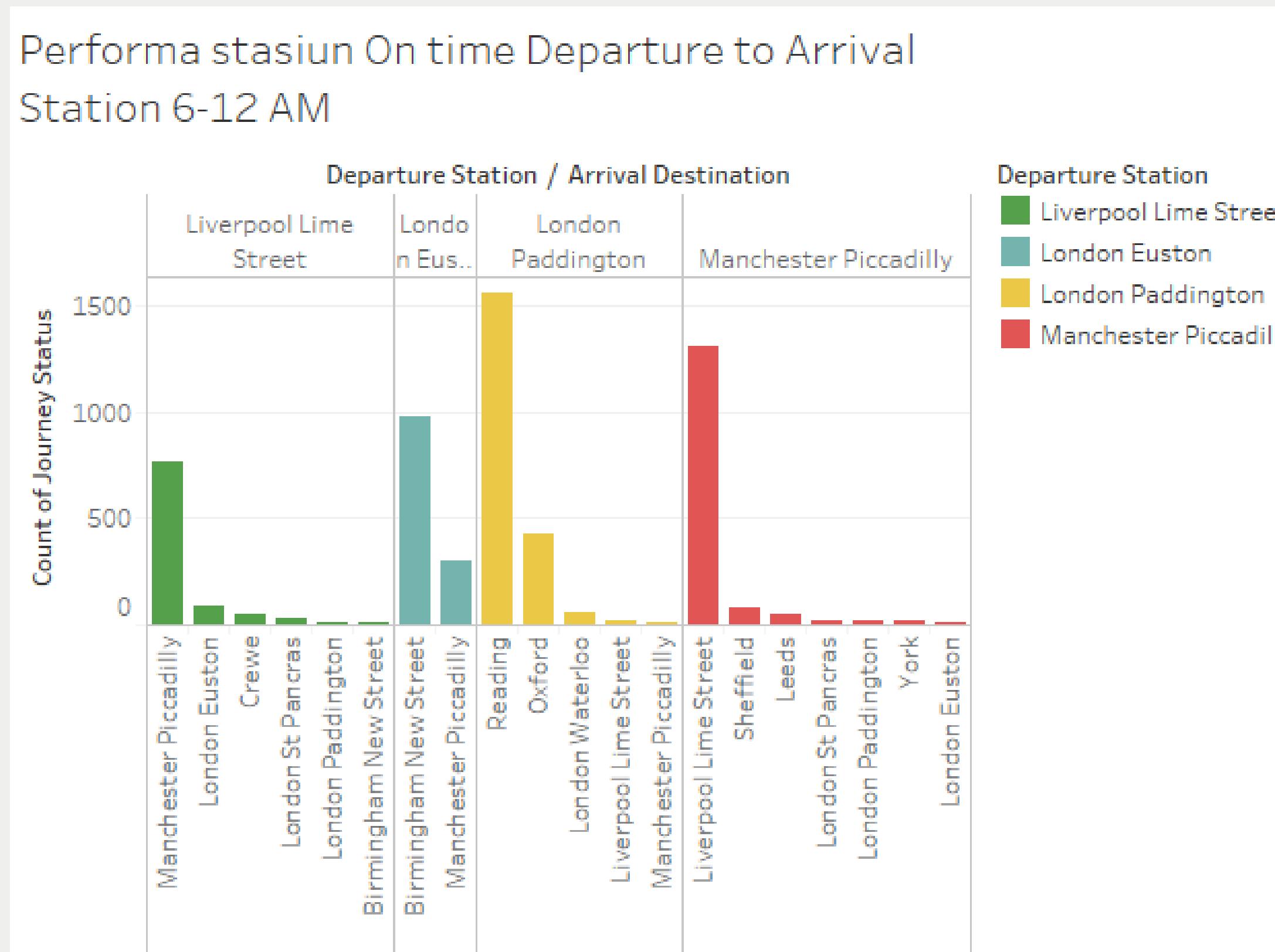
Permintaan Refund Karena Delayed dan Canceled perjalanan, berdasarkan masalah Alasan keterlambatan dari 4 Stasiun Teramai di Inggris



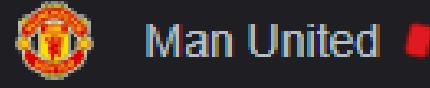
- Ternyata keterkaitan dengan permasalahan keterlambatan cukup signifikan dengan pembatalan dan juga permintaan refund, terutama terjadi di stasiun liverpool lime street karena masalah teknis

- lalu pembatal secara merata banyak di lakukan serta permintaan refund baik karena masalah, kesalahan sinyal kereta, lalu lintas, teknis pegawai, kekurangan SDM, cuaca buruk, dan kondisi cuaca

Disamping itu ada performa positif dari stasiun kereta



- Dengan ini terlihat bahwa dari ke 4 Stasiun memiliki performa yang baik kepada stasiun tujuan tertentu, untuk kasus manchester picadilly menuju liverpool lime street terpantau performa tidak bisa di prediksi baik kemungkinan delay atau ontime, sedangkan sebaliknya jika dari liverpool lime street menuju manchester picadilly terpantau cukup on time dalam kedatangannya



Man United



Liverpool

FT
17 Mar

4

3

Man United

Liverpool

FT
7 Apr

2

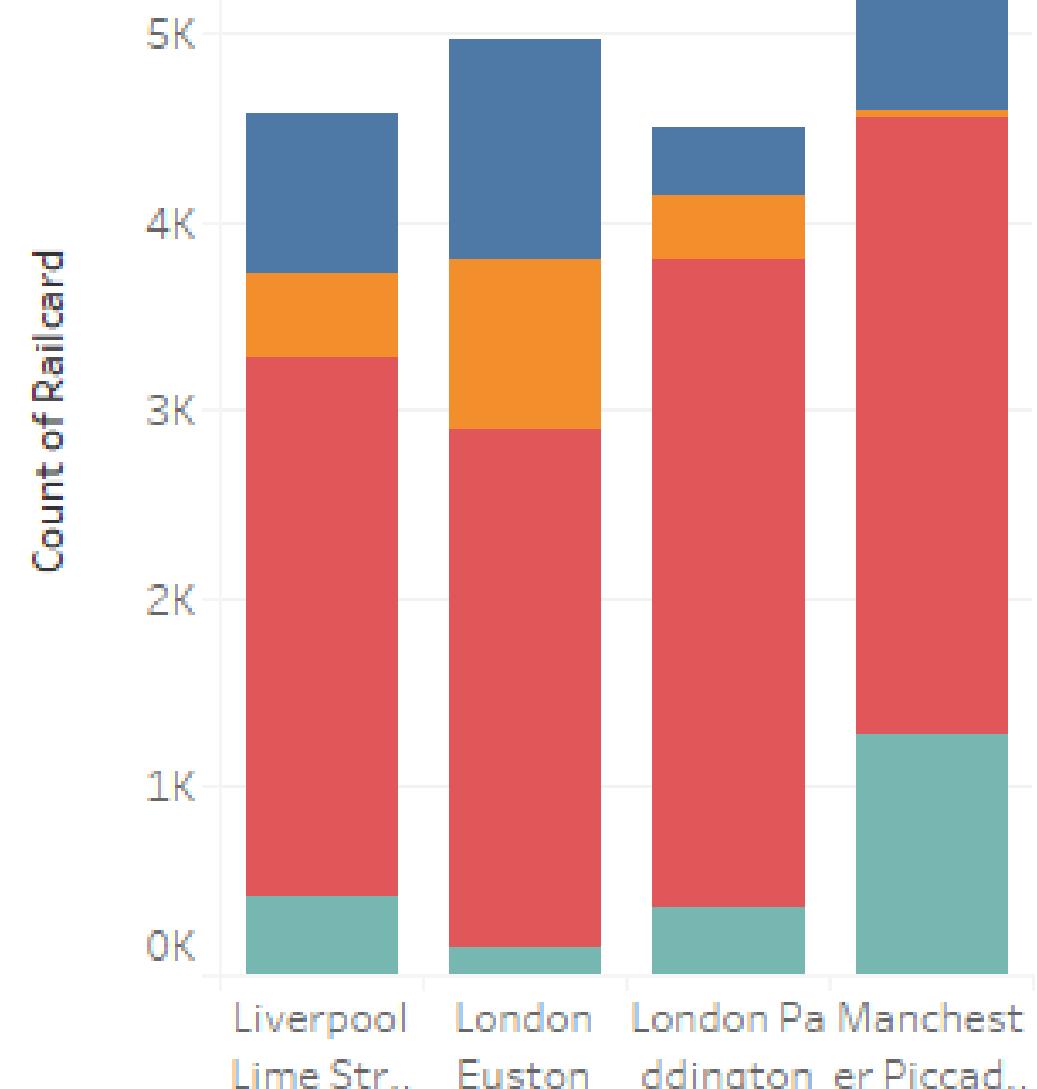
2

Penumpang Berdasarkan keterangan kartu member

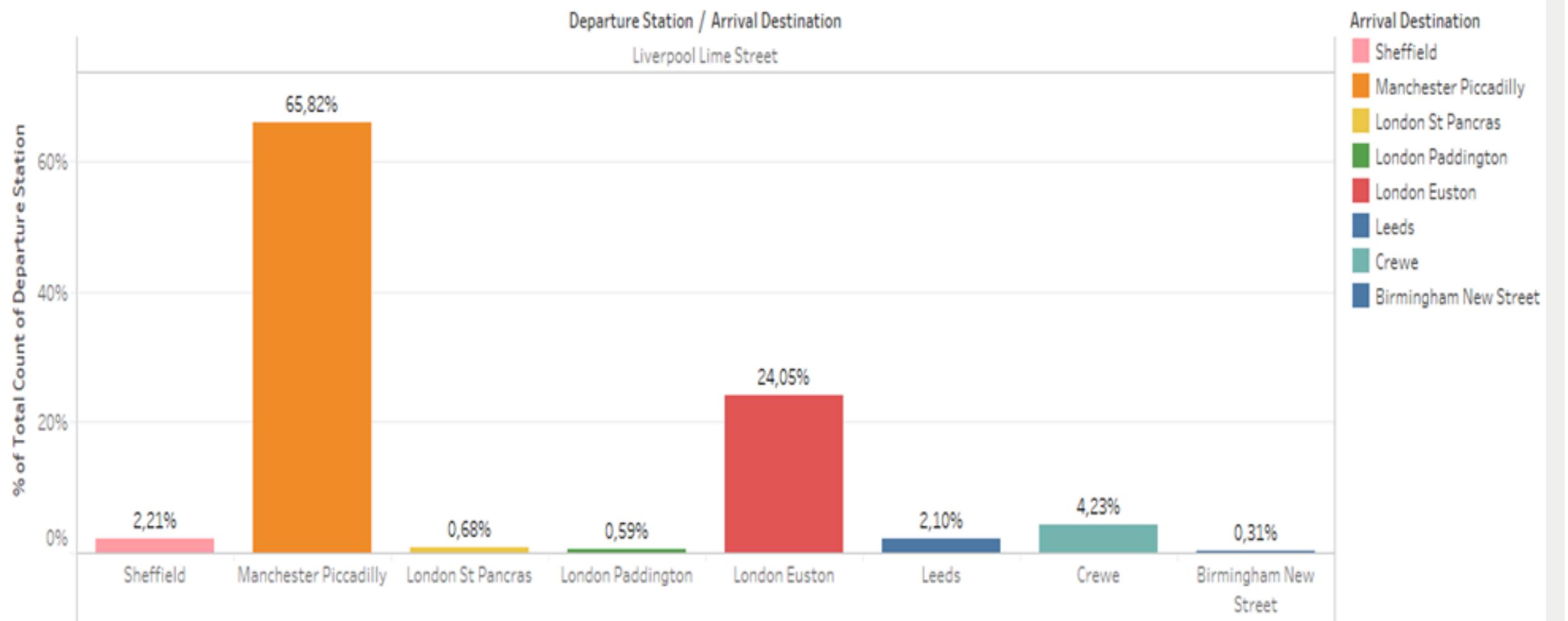
Departure Station

Railcard

- Adult
- Disabled
- None
- Senior



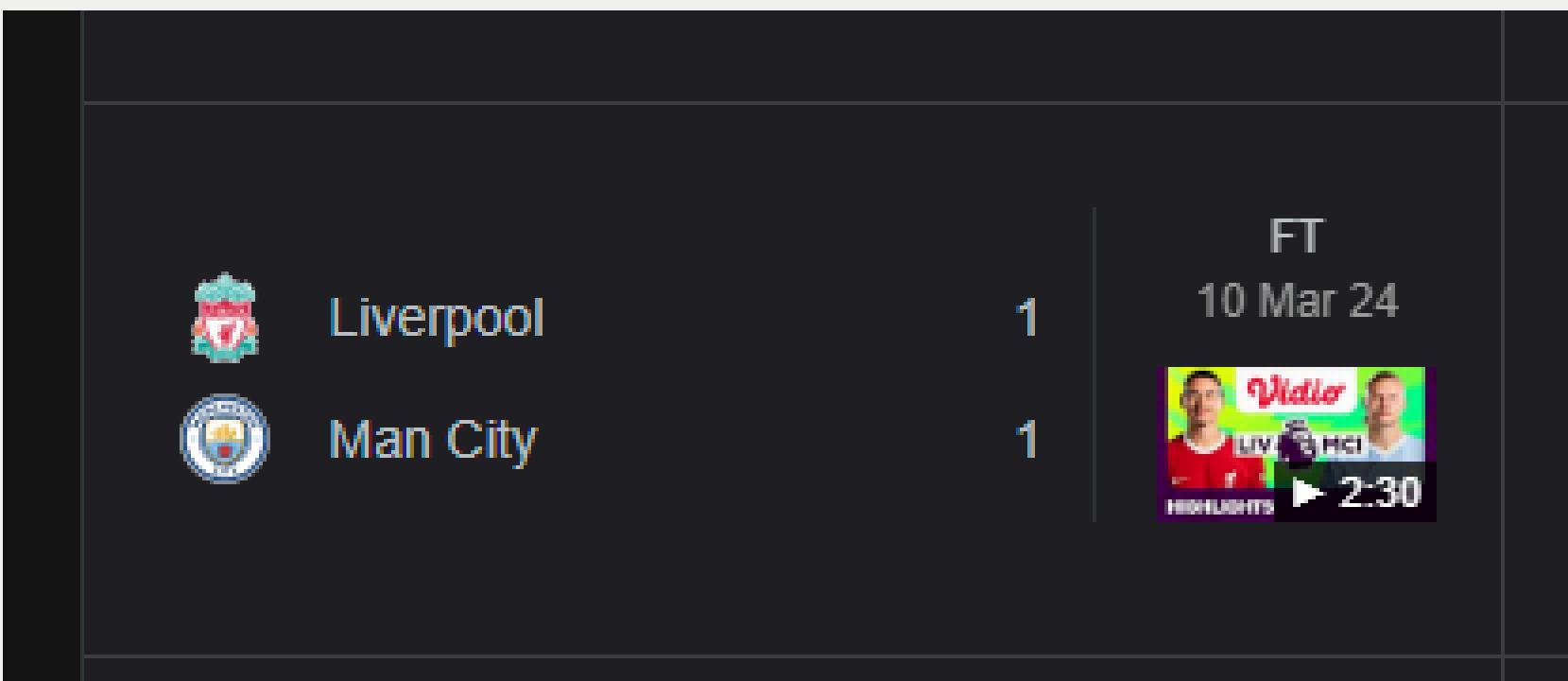
Banyaknya Keberangkatan dari Liverpool Lime Street Ke Stasiun Lain



- Dikarenakan masyarakat Inggris sangat menyukai sepak bola, banyak supporter datang ke markas tim tamu menggunakan transportasi yang mudah yaitu kereta dan dapat pula pola yang disimpulkan kemungkinan lonjakan penumpang akan sama ketika terjadi adanya pertandingan sepakbola

- Jika di lihat dari grafik januari-april dan kenyataan di real live bahwa kenaikan pengguna dari kota Liverpool ke kota Manchester dan keberangkatan terbanyak dan pengguna non yaitu kalangan remaja, probabilitas yang kemungkinan terjadi dikarenakan pertandingan antara MU vs Liverpool di Kota manchester sebanyak 2 kali di Bulan Maret & bulan April





- Sebaliknya di lihat dari grafik bahwa kemungkinan dari 2 pertandingan MU vs Liverpool 2x lipat supporter liverpool pulang pergi dari kota liverpool ke kota manchester, ditambah dengan terjadinya pertandingan lain di bulan maret antara manchester city vs Liverpool, yang mana pendukung city menuju ke kota liverpool sehingga lonjakan pengguna kereta 3x lipat dari biasanya

Banyaknya Keberangkatan dari Manchester Piccadilly Ke Stasiun Lain

