In [13]:
```
# Import necessary libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [14]:
```
data = pd.read_excel('Diva Cabs.xlsx')
```

In [15]:
```
data.head(5)
```

Out[15]:

	PARTY DETAILS	Unnamed: 1	Unnamed: 2	Unnamed: 3	SEEKING	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	OFFERING	...	Unnamed: 11	Unnamed: 12	Unnamed: 13	Unnamed: 14	Unnamed: 15	Unnamed: 16	Unnamed: 17	Unnamed: 18	Unnamed: 19	Unnamed: 20
0	Posting Dt	Posting Time	Username	Number	Car	Date	Time	Location	Destination	Car	...	Time	Location	Destination	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	2025-08-26 00:00:00	12:03:00	Srinivas Rao	9905407147	Innova Crysta	2025-08-27 00:00:00	08:00:00	Tata	Ranchi	Innova Crysta	...	15:00:00	Ranchi	Tata	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	2025-08-26 00:00:00	12:30:00	Md Arsalan	9661334122	Ertiga	2025-08-27 00:00:00	08:00:00	Tata	Any	Dzire	...	15:00:00	Tata	Ranchi	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	2025-08-26 00:00:00	13:00:00	My life	8789340832	Dzire	2025-08-26 00:00:00	17:15:00	Ranchi	Tata	Dzire	...	16:00:00	Ranchi	Tata	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	2025-08-26 00:00:00	13:10:00	Sukhwinder Singh	7903808015	Dzire	2025-08-26 00:00:00	14:15:00	Ranchi	Tata	Dzire	...	08:00:00	Ranchi	Tata	NaN	NaN	NaN	NaN	NaN	NaN	NaN

5 rows × 21 columns

In [16]:
```
data.columns = [
    'Post Date', 'Post Time', 'Username', 'Number', 'Seek Car', 'Seek Date', 'Seek Time', 'Seek Location','Seek Destination',
    'Offer Car', 'Offer Date', 'Offer Time', 'Offer Location', 'Offer Destination', 'Unnamed:14','Unnamed:15','Unnamed:16','Unnamed:17','Unnamed:18','Unnamed: 19','Unnamed:20'
]
```

In [17]:
```
data.head(5)
```

Out[17]:

	Post Date	Post Time	Username	Number	Seek Car	Seek Date	Seek Time	Seek Location	Seek Destination	Offer Car	...	Offer Time	Offer Location	Offer Destination	Unnamed:14	Unnamed:15	Unnamed:16	Unnamed:17	Unnamed:18	Unnamed: 19	Unnamed:20
0	Posting Dt	Posting Time	Username	Number	Car	Date	Time	Location	Destination	Car	...	Time	Location	Destination	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	2025-08-26 00:00:00	12:03:00	Srinivas Rao	9905407147	Innova Crysta	2025-08-27 00:00:00	08:00:00	Tata	Ranchi	Innova Crysta	...	15:00:00	Ranchi	Tata	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	2025-08-26 00:00:00	12:30:00	Md Arsalan	9661334122	Ertiga	2025-08-27 00:00:00	08:00:00	Tata	Any	Dzire	...	15:00:00	Tata	Ranchi	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	2025-08-26 00:00:00	13:00:00	My life	8789340832	Dzire	2025-08-26 00:00:00	17:15:00	Ranchi	Tata	Dzire	...	16:00:00	Ranchi	Tata	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	2025-08-26 00:00:00	13:10:00	Sukhwinder Singh	7903808015	Dzire	2025-08-26 00:00:00	14:15:00	Ranchi	Tata	Dzire	...	08:00:00	Ranchi	Tata	NaN	NaN	NaN	NaN	NaN	NaN	NaN

5 rows × 21 columns

In [18]:
```
data = data.drop(index=0).reset_index(drop=True)
```

In [19]:
```
data = data.drop(columns=[
    'Unnamed:14', 'Unnamed:15', 'Unnamed:16',
    'Unnamed:17', 'Unnamed:18', 'Unnamed: 19', 'Unnamed:20'
])
```

In [20]:
```
print(data.head())
```

```
            Post Date Post Time          Username      Number       Seek Car  \
0  2025-08-26 00:00:00  12:03:00      Srinivas Rao  9905407147  Innova Crysta
1  2025-08-26 00:00:00  12:30:00        Md Arsalan  9661334122         Ertiga
2  2025-08-26 00:00:00  13:00:00          My life   8789340832          Dzire
3  2025-08-26 00:00:00  13:10:00  Sukhwinder Singh  7903808015          Dzire
4  2025-08-26 00:00:00  13:26:00           Travels  7209231174          Dzire

            Seek Date Seek Time Seek Location Seek Destination  \
0  2025-08-27 00:00:00  08:00:00          Tata           Ranchi
1  2025-08-27 00:00:00  08:00:00          Tata              Any
2  2025-08-26 00:00:00  17:15:00        Ranchi             Tata
3  2025-08-26 00:00:00  14:15:00        Ranchi             Tata
4  2025-08-26 00:00:00  15:00:00          Tata           Ranchi

       Offer Car           Offer Date Offer Time Offer Location  \
0  Innova Crysta  2025-08-26 00:00:00   15:00:00         Ranchi
1          Dzire  2025-08-26 00:00:00   15:00:00           Tata
2          Dzire  2025-08-26 00:00:00   16:00:00         Ranchi
3          Dzire  2025-08-27 00:00:00   08:00:00         Ranchi
4          Dzire  2025-08-26 00:00:00   16:00:00         Ranchi

  Offer Destination
0              Tata
1            Ranchi
2              Tata
3              Tata
4              Tata
```

In [21]:
```
data.info()
data
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 189 entries, 0 to 188
Data columns (total 14 columns):
 #   Column             Non-Null Count  Dtype 
---  ------             --------------  ----- 
 0   Post Date          189 non-null    object
 1   Post Time          188 non-null    object
 2   Username           182 non-null    object
 3   Number             166 non-null    object
 4   Seek Car           157 non-null    object
 5   Seek Date          178 non-null    object
 6   Seek Time          176 non-null    object
 7   Seek Location      176 non-null    object
 8   Seek Destination   176 non-null    object
 9   Offer Car          154 non-null    object
 10  Offer Date         181 non-null    object
 11  Offer Time         179 non-null    object
 12  Offer Location     179 non-null    object
 13  Offer Destination  178 non-null    object
dtypes: object(14)
memory usage: 20.8+ KB
```

Out[21]:

	Post Date	Post Time	Username	Number	Seek Car	Seek Date	Seek Time	Seek Location	Seek Destination	Offer Car	Offer Date	Offer Time	Offer Location	Offer Destination
0	2025-08-26 00:00:00	12:03:00	Srinivas Rao	9905407147	Innova Crysta	2025-08-27 00:00:00	08:00:00	Tata	Ranchi	Innova Crysta	2025-08-26 00:00:00	15:00:00	Ranchi	Tata
1	2025-08-26 00:00:00	12:30:00	Md Arsalan	9661334122	Ertiga	2025-08-27 00:00:00	08:00:00	Tata	Any	Dzire	2025-08-26 00:00:00	15:00:00	Tata	Ranchi
2	2025-08-26 00:00:00	13:00:00	My life	8789340832	Dzire	2025-08-26 00:00:00	17:15:00	Ranchi	Tata	Dzire	2025-08-26 00:00:00	16:00:00	Ranchi	Tata
3	2025-08-26 00:00:00	13:10:00	Sukhwinder Singh	7903808015	Dzire	2025-08-26 00:00:00	14:15:00	Ranchi	Tata	Dzire	2025-08-27 00:00:00	08:00:00	Ranchi	Tata
4	2025-08-26 00:00:00	13:26:00	Travels	7209231174	Dzire	2025-08-26 00:00:00	15:00:00	Tata	Ranchi	Dzire	2025-08-26 00:00:00	16:00:00	Ranchi	Tata
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
184	2025-09-02 00:00:00	08:09:00	Sujit Frnds	9122007630	Dzire	2025-09-02 00:00:00	12:00:00	XLRI	Airport	Dzire/Ertiga	2025-09-02 00:00:00	09:00:00	Tata	Ranchi Airport
185	2025-09-02 00:00:00	08:49:00	Murli Manohar Ray	9155540003	Dzire	2025-09-02 00:00:00	04:00:00	Ranchi	Tata	Dzire	2025-09-02 00:00:00	10:00:00	Ranchi	Tata
186	2025-09-02 00:00:00	09:32:00	Rishi Tour and Travel	8651568297	Dzire/Ertiga	2025-09-02 00:00:00	12:00:00	Ranchi	Tata	Sedan	2025-09-02 00:00:00	11:45:00	Jamshedpur	Ranchi
187	2025-09-02 00:00:00	09:48:00	Rajiv Kumar Singh	7903817646	Dzire/Ertiga	2025-09-02 00:00:00	07:00:00	Ranchi	Tata	Dzire/Ertiga	2025-09-02 00:00:00	10:30:00	Ranchi AIrport	Tata
188	2025-09-02 00:00:00	10:05:00	Kuldeep Singh	7992356915	Dzire	2025-09-02 00:00:00	5:00:00	Ranchi	Tata	Dzire	2025-09-02 00:00:00	03:30:00	Ranchi	Tata

189 rows × 14 columns

In [22]:
```
data.shape
```

Out[22]:
```
(189, 14)
```

Data Prepocessing

In [26]:
```
# Force both to string before concatenation
data['seek_dt'] = pd.to_datetime(
    data['Seek Date'].astype(str) + " " + data['Seek Time'].astype(str),
    errors='coerce'
)

data['offer_dt'] = pd.to_datetime(
    data['Offer Date'].astype(str) + " " + data['Offer Time'].astype(str),
    errors='coerce'
)

data['post_dt'] = pd.to_datetime(
    data['Post Date'].astype(str) + " " + data['Post Time'].astype(str),
    errors='coerce'
)
```

```
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\2441106594.py:2: UserWarning: Could not infer format, so each element will be parsed individually, falling back to `dateutil`. To ensure parsing is consistent and as-expected, please specify a format.
  data['seek_dt'] = pd.to_datetime(
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\2441106594.py:7: UserWarning: Could not infer format, so each element will be parsed individually, falling back to `dateutil`. To ensure parsing is consistent and as-expected, please specify a format.
  data['offer_dt'] = pd.to_datetime(
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\2441106594.py:12: UserWarning: Could not infer format, so each element will be parsed individually, falling back to `dateutil`. To ensure parsing is consistent and as-expected, please specify a format.
  data['post_dt'] = pd.to_datetime(
```

In [27]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 189 entries, 0 to 188
Data columns (total 17 columns):
 #   Column             Non-Null Count  Dtype         
---  ------             --------------  -----         
 0   Post Date          189 non-null    object        
 1   Post Time          188 non-null    object        
 2   Username           182 non-null    object        
 3   Number             166 non-null    object        
 4   Seek Car           157 non-null    object        
 5   Seek Date          178 non-null    object        
 6   Seek Time          176 non-null    object        
 7   Seek Location      176 non-null    object        
 8   Seek Destination   176 non-null    object        
 9   Offer Car          154 non-null    object        
 10  Offer Date         181 non-null    object        
 11  Offer Time         179 non-null    object        
 12  Offer Location     179 non-null    object        
 13  Offer Destination  178 non-null    object        
 14  seek_dt            175 non-null    datetime64[ns]
 15  offer_dt           178 non-null    datetime64[ns]
 16  post_dt            187 non-null    datetime64[ns]
dtypes: datetime64[ns](3), object(14)
memory usage: 25.2+ KB
```

In [28]:

```
data['Post Date'].fillna("NOT",inplace = True)
data['Username'].fillna("NOT",inplace = True)
data['Number'].fillna("NOT",inplace = True)
data['Seek Car'].fillna("NOT",inplace = True)
data['Seek Location'].fillna("NOT",inplace = True)
data['Seek Destination'].fillna("NOT",inplace = True)
data['Offer Car'].fillna("NOT",inplace = True)
data['Offer Location'].fillna("NOT",inplace = True)
data['Offer Destination'].fillna("NOT",inplace = True)
```

```
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:1: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Post Date'].fillna("NOT",inplace = True)
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:2: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Username'].fillna("NOT",inplace = True)
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:3: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Number'].fillna("NOT",inplace = True)
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:4: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Seek Car'].fillna("NOT",inplace = True)
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:5: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Seek Location'].fillna("NOT",inplace = True)
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:6: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Seek Destination'].fillna("NOT",inplace = True)
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:7: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Offer Car'].fillna("NOT",inplace = True)
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:8: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Offer Location'].fillna("NOT",inplace = True)
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3514184000.py:9: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  data['Offer Destination'].fillna("NOT",inplace = True)
```

In [29]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 189 entries, 0 to 188
Data columns (total 17 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   Post Date          189 non-null    object
 1   Post Time          188 non-null    object
 2   Username           189 non-null    object
 3   Number             189 non-null    object
 4   Seek Car           189 non-null    object
 5   Seek Date          178 non-null    object
 6   Seek Time          176 non-null    object
 7   Seek Location      189 non-null    object
 8   Seek Destination   189 non-null    object
 9   Offer Car          189 non-null    object
 10  Offer Date         181 non-null    object
 11  Offer Time         179 non-null    object
 12  Offer Location     189 non-null    object
 13  Offer Destination  189 non-null    object
 14  seek_dt            175 non-null    datetime64[ns]
 15  offer_dt           178 non-null    datetime64[ns]
 16  post_dt            187 non-null    datetime64[ns]
dtypes: datetime64[ns](3), object(14)
memory usage: 25.2+ KB
```

In [30]:
```
print("Missing values (non-date/time columns):")
print(data.drop(columns=['Seek Time','Offer Time']).isnull().sum())
```

```
Missing values (non-date/time columns):
Post Date             0
Post Time             1
Username              0
Number                0
Seek Car              0
Seek Date            11
Seek Location         0
Seek Destination      0
Offer Car             0
Offer Date            8
Offer Location        0
Offer Destination     0
seek_dt              14
offer_dt             11
post_dt               2
dtype: int64
```

In [32]:
```
data = data.dropna(subset=['Seek Time','Offer Time'])
print("Shape after dropping:", data.shape)
```

```
Shape after dropping: (167, 17)
```

In [33]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 167 entries, 0 to 188
Data columns (total 17 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   Post Date          167 non-null    object
 1   Post Time          167 non-null    object
 2   Username           167 non-null    object
 3   Number             167 non-null    object
 4   Seek Car           167 non-null    object
 5   Seek Date          167 non-null    object
 6   Seek Time          167 non-null    object
 7   Seek Location      167 non-null    object
 8   Seek Destination   167 non-null    object
 9   Offer Car          167 non-null    object
 10  Offer Date         167 non-null    object
 11  Offer Time         167 non-null    object
 12  Offer Location     167 non-null    object
 13  Offer Destination  167 non-null    object
 14  seek_dt            167 non-null    datetime64[ns]
 15  offer_dt           167 non-null    datetime64[ns]
 16  post_dt            167 non-null    datetime64[ns]
dtypes: datetime64[ns](3), object(14)
memory usage: 23.5+ KB
```

In [34]:
```
data['seek_hour'] = data['seek_dt'].dt.hour
data['offer_hour'] = data['offer_dt'].dt.hour
```

In [35]:
```
data.head()
```

Out[35]:

	Post Date	Post Time	Username	Number	Seek Car	Seek Date	Seek Time	Seek Location	Seek Destination	Offer Car	Offer Date	Offer Time	Offer Location	Offer Destination	seek_dt	offer_dt	post_dt	seek_hour	offer_hour
0	2025-08-26 00:00:00	12:03:00	Srinivas Rao	9905407147	Innova Crysta	2025-08-27 00:00:00	08:00:00	Tata	Ranchi	Innova Crysta	2025-08-26 00:00:00	15:00:00	Ranchi	Tata	2025-08-27 08:00:00	2025-08-26 15:00:00	2025-08-26 12:03:00	8	15
1	2025-08-26 00:00:00	12:30:00	Md Arsalan	9661334122	Ertiga	2025-08-27 00:00:00	08:00:00	Tata	Any	Dzire	2025-08-26 00:00:00	15:00:00	Tata	Ranchi	2025-08-27 08:00:00	2025-08-26 15:00:00	2025-08-26 12:30:00	8	15
2	2025-08-26 00:00:00	13:00:00	My life	8789340832	Dzire	2025-08-26 00:00:00	17:15:00	Ranchi	Tata	Dzire	2025-08-26 00:00:00	16:00:00	Ranchi	Tata	2025-08-26 17:15:00	2025-08-26 16:00:00	2025-08-26 13:00:00	17	16
3	2025-08-26 00:00:00	13:10:00	Sukhwinder Singh	7903808015	Dzire	2025-08-26 00:00:00	14:15:00	Ranchi	Tata	Dzire	2025-08-27 00:00:00	08:00:00	Ranchi	Tata	2025-08-26 14:15:00	2025-08-27 08:00:00	2025-08-26 13:10:00	14	8
4	2025-08-26 00:00:00	13:26:00	Travels	7209231174	Dzire	2025-08-26 00:00:00	15:00:00	Tata	Ranchi	Dzire	2025-08-26 00:00:00	16:00:00	Ranchi	Tata	2025-08-26 15:00:00	2025-08-26 16:00:00	2025-08-26 13:26:00	15	16

In [36]:
```
def time_period(h):
    if 5 <= h <= 11: return "Morning"
    elif 12 <= h <= 16: return "Afternoon"
    elif 17 <= h <= 20: return "Evening"
    else: return "Night"
data['seek_period'] = data['seek_hour'].apply(time_period)
data['offer_period'] = data['offer_hour'].apply(time_period)
```

In [37]:
```
data.head()
```

Out[37]:

	Post Date	Post Time	Username	Number	Seek Car	Seek Date	Seek Time	Seek Location	Seek Destination	Offer Car	...	Offer Time	Offer Location	Offer Destination	seek_dt	offer_dt	post_dt	seek_hour	offer_hour	seek_period	offer_period
0	2025-08-26 00:00:00	12:03:00	Srinivas Rao	9905407147	Innova Crysta	2025-08-27 00:00:00	08:00:00	Tata	Ranchi	Innova Crysta	...	15:00:00	Ranchi	Tata	2025-08-27 08:00:00	2025-08-26 15:00:00	2025-08-26 12:03:00	8	15	Morning	Afternoon
1	2025-08-26 00:00:00	12:30:00	Md Arsalan	9661334122	Ertiga	2025-08-27 00:00:00	08:00:00	Tata	Any	Dzire	...	15:00:00	Tata	Ranchi	2025-08-27 08:00:00	2025-08-26 15:00:00	2025-08-26 12:30:00	8	15	Morning	Afternoon
2	2025-08-26 00:00:00	13:00:00	My life	8789340832	Dzire	2025-08-26 00:00:00	17:15:00	Ranchi	Tata	Dzire	...	16:00:00	Ranchi	Tata	2025-08-26 17:15:00	2025-08-26 16:00:00	2025-08-26 13:00:00	17	16	Evening	Afternoon
3	2025-08-26 00:00:00	13:10:00	Sukhwinder Singh	7903808015	Dzire	2025-08-26 00:00:00	14:15:00	Ranchi	Tata	Dzire	...	08:00:00	Ranchi	Tata	2025-08-26 14:15:00	2025-08-27 08:00:00	2025-08-26 13:10:00	14	8	Afternoon	Morning
4	2025-08-26 00:00:00	13:26:00	Travels	7209231174	Dzire	2025-08-26 00:00:00	15:00:00	Tata	Ranchi	Dzire	...	16:00:00	Ranchi	Tata	2025-08-26 15:00:00	2025-08-26 16:00:00	2025-08-26 13:26:00	15	16	Afternoon	Afternoon

5 rows × 21 columns

In [39]:
```
# Calculate duration (offer - seek). If negative, assume next day
data['duration'] = data['offer_dt'] - data['seek_dt']
data.loc[data['duration'] < pd.Timedelta(0), 'duration'] += pd.Timedelta(days=1)
data
```

Out[39]:

	Post Date	Post Time	Username	Number	Seek Car	Seek Date	Seek Time	Seek Location	Seek Destination	Offer Car	...	Offer Location	Offer Destination	seek_dt	offer_dt	post_dt	seek_hour	offer_hour	seek_period	offer_period	duration
0	2025-08-26 00:00:00	12:03:00	Srinivas Rao	9905407147	Innova Crysta	2025-08-27 00:00:00	08:00:00	Tata	Ranchi	Innova Crysta	...	Ranchi	Tata	2025-08-27 08:00:00	2025-08-26 15:00:00	2025-08-26 12:03:00	8	15	Morning	Afternoon	0 days 07:00:00
1	2025-08-26 00:00:00	12:30:00	Md Arsalan	9661334122	Ertiga	2025-08-27 00:00:00	08:00:00	Tata	Any	Dzire	...	Tata	Ranchi	2025-08-27 08:00:00	2025-08-26 15:00:00	2025-08-26 12:30:00	8	15	Morning	Afternoon	0 days 07:00:00
2	2025-08-26 00:00:00	13:00:00	My life	8789340832	Dzire	2025-08-26 00:00:00	17:15:00	Ranchi	Tata	Dzire	...	Ranchi	Tata	2025-08-26 17:15:00	2025-08-26 16:00:00	2025-08-26 13:00:00	17	16	Evening	Afternoon	0 days 22:45:00
3	2025-08-26 00:00:00	13:10:00	Sukhwinder Singh	7903808015	Dzire	2025-08-26 00:00:00	14:15:00	Ranchi	Tata	Dzire	...	Ranchi	Tata	2025-08-26 14:15:00	2025-08-27 08:00:00	2025-08-26 13:10:00	14	8	Afternoon	Morning	0 days 17:45:00
4	2025-08-26 00:00:00	13:26:00	Travels	7209231174	Dzire	2025-08-26 00:00:00	15:00:00	Tata	Ranchi	Dzire	...	Ranchi	Tata	2025-08-26 15:00:00	2025-08-26 16:00:00	2025-08-26 13:26:00	15	16	Afternoon	Afternoon	0 days 01:00:00
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
184	2025-09-02 00:00:00	08:09:00	Sujit Frnds	9122007630	Dzire	2025-09-02 00:00:00	12:00:00	XLRI	Airport	Dzire/Ertiga	...	Tata	Ranchi Airport	2025-09-02 12:00:00	2025-09-02 09:00:00	2025-09-02 08:09:00	12	9	Afternoon	Morning	0 days 21:00:00
185	2025-09-02 00:00:00	08:49:00	Murli Manohar Ray	9155540003	Dzire	2025-09-02 00:00:00	04:00:00	Ranchi	Tata	Dzire	...	Ranchi	Tata	2025-09-02 04:00:00	2025-09-02 10:00:00	2025-09-02 08:49:00	4	10	Night	Morning	0 days 06:00:00
186	2025-09-02 00:00:00	09:32:00	Rishi Tour and Travel	8651568297	Dzire/Ertiga	2025-09-02 00:00:00	12:00:00	Ranchi	Tata	Sedan	...	Jamshedpur	Ranchi	2025-09-02 12:00:00	2025-09-02 11:45:00	2025-09-02 09:32:00	12	11	Afternoon	Morning	0 days 23:45:00
187	2025-09-02 00:00:00	09:48:00	Rajiv Kumar Singh	7903817646	Dzire/Ertiga	2025-09-02 00:00:00	07:00:00	Ranchi	Tata	Dzire/Ertiga	...	Ranchi Alrport	Tata	2025-09-02 07:00:00	2025-09-02 10:30:00	2025-09-02 09:48:00	7	10	Morning	Morning	0 days 03:30:00
188	2025-09-02 00:00:00	10:05:00	Kuldeep Singh	7992356915	Dzire	2025-09-02 00:00:00	5:00:00	Ranchi	Tata	Dzire	...	Ranchi	Tata	2025-09-02 05:00:00	2025-09-02 03:30:00	2025-09-02 10:05:00	5	3	Morning	Night	0 days 22:30:00

167 rows × 22 columns

In [42]:
```
# Create up_trip and down_trip keys
data['up_trip'] = data['Seek Location'].str.upper().str.strip() + " - " + data['Seek Destination'].str.upper().str.strip()
```

```
data['down_trip'] = data['Offer Location'].str.upper().str.strip() + " - " + data['Offer Destination'].str.upper().str.strip()
data
```

Out[42]:

	Post Date	Post Time	Username	Number	Seek Car	Seek Date	Seek Time	Seek Location	Seek Destination	Offer Car	...	seek_dt	offer_dt	post_dt	seek_hour	offer_hour	seek_period	offer_period	duration	up_trip	down_trip
0	2025-08-26 00:00:00	12:03:00	Srinivas Rao	9905407147	Innova Crysta	2025-08-27 00:00:00	08:00:00	Tata	Ranchi	Innova Crysta	...	2025-08-27 08:00:00	2025-08-26 15:00:00	2025-08-26 12:03:00	8	15	Morning	Afternoon	0 days 07:00:00	TATA - RANCHI	RANCHI - TATA
1	2025-08-26 00:00:00	12:30:00	Md Arsalan	9661334122	Ertiga	2025-08-27 00:00:00	08:00:00	Tata	Any	Dzire	...	2025-08-27 08:00:00	2025-08-26 15:00:00	2025-08-26 12:30:00	8	15	Morning	Afternoon	0 days 07:00:00	TATA - ANY	TATA - RANCHI
2	2025-08-26 00:00:00	13:00:00	My life	8789340832	Dzire	2025-08-26 00:00:00	17:15:00	Ranchi	Tata	Dzire	...	2025-08-26 17:15:00	2025-08-26 16:00:00	2025-08-26 13:00:00	17	16	Evening	Afternoon	0 days 22:45:00	RANCHI - TATA	RANCHI - TATA
3	2025-08-26 00:00:00	13:10:00	Sukhwinder Singh	7903808015	Dzire	2025-08-26 00:00:00	14:15:00	Ranchi	Tata	Dzire	...	2025-08-26 14:15:00	2025-08-27 08:00:00	2025-08-26 13:10:00	14	8	Afternoon	Morning	0 days 17:45:00	RANCHI - TATA	RANCHI - TATA
4	2025-08-26 00:00:00	13:26:00	Travels	7209231174	Dzire	2025-08-26 00:00:00	15:00:00	Tata	Ranchi	Dzire	...	2025-08-26 15:00:00	2025-08-26 16:00:00	2025-08-26 13:26:00	15	16	Afternoon	Afternoon	0 days 01:00:00	TATA - RANCHI	RANCHI - TATA
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
184	2025-09-02 00:00:00	08:09:00	Sujit Frnds	9122007630	Dzire	2025-09-02 00:00:00	12:00:00	XLRI	Airport	Dzire/Ertiga	...	2025-09-02 12:00:00	2025-09-02 09:00:00	2025-09-02 08:09:00	12	9	Afternoon	Morning	0 days 21:00:00	XLRI - AIRPORT	TATA - RANCHI AIRPORT
185	2025-09-02 00:00:00	08:49:00	Murli Manohar Ray	9155540003	Dzire	2025-09-02 00:00:00	04:00:00	Ranchi	Tata	Dzire	...	2025-09-02 04:00:00	2025-09-02 10:00:00	2025-09-02 08:49:00	4	10	Night	Morning	0 days 06:00:00	RANCHI - TATA	RANCHI - TATA
186	2025-09-02 00:00:00	09:32:00	Rishi Tour and Travel	8651568297	Dzire/Ertiga	2025-09-02 00:00:00	12:00:00	Ranchi	Tata	Sedan	...	2025-09-02 12:00:00	2025-09-02 11:45:00	2025-09-02 09:32:00	12	11	Afternoon	Morning	0 days 23:45:00	RANCHI - TATA	JAMSHEDPUR - RANCHI
187	2025-09-02 00:00:00	09:48:00	Rajiv Kumar Singh	7903817646	Dzire/Ertiga	2025-09-02 00:00:00	07:00:00	Ranchi	Tata	Dzire/Ertiga	...	2025-09-02 07:00:00	2025-09-02 10:30:00	2025-09-02 09:48:00	7	10	Morning	Morning	0 days 03:30:00	RANCHI - TATA	RANCHI AIRPORT - TATA
188	2025-09-02 00:00:00	10:05:00	Kuldeep Singh	7992356915	Dzire	2025-09-02 00:00:00	5:00:00	Ranchi	Tata	Dzire	...	2025-09-02 05:00:00	2025-09-02 03:30:00	2025-09-02 10:05:00	5	3	Morning	Night	0 days 22:30:00	RANCHI - TATA	RANCHI - TATA

167 rows × 24 columns

In [44]:
```
# Shortest duration by up_trip / down_trip
shortest_up = data.loc[data.groupby('up_trip')['duration'].idxmin(), ['up_trip','Username','duration']]
shortest_down = data.loc[data.groupby('down_trip')['duration'].idxmin(), ['down_trip','Username','duration']]
print("\nShortest Up Trips:")
print(shortest_up.head())
print("\nShortest Down Trips:")
print(shortest_down.head())
```

```
Shortest Up Trips:
                  up_trip          Username        duration
16   DHANBAD/BOKARO - TATA       Dinesh Cabs 0 days 04:00:00
145      HAZARIBAGH - TATA  Cab Service Only 0 days 02:00:00
54     JAMSHEDPUR - RANCHI   Jay Ram Tarvels 0 days 08:00:00
21           KOLKATA - TATA   Gayatri Travels 0 days 06:00:00
125    RANCHI - HAZARIBAGH   Mohammad Shahid 0 days 18:00:00

Shortest Down Trips:
               down_trip               Username        duration
116      DHANBAD - RANCHI      Ravi Kumar Thakur 0 days 04:00:00
16         DHANBAD - TATA            Dinesh Cabs 0 days 04:00:00
54   JAMSHEDPUR - DHANBAD        Jay Ram Tarvels 0 days 08:00:00
142      JAMSHEDPUR - NOT                 Rupam 0 days 01:00:00
169   JAMSHEDPUR - RANCHI  Rishi Tour and Travel 0 days 22:00:00
```

In [48]:
```
# Duplicate detection (same username + same trips + same duration)
Operator_routs = data[data.duplicated(['Username','up_trip','down_trip','duration'], keep=False)]
Operator_status = data[data.duplicated(['Username','seek_dt','offer_dt'], keep=False)]
print("\nOperator Offers same routs")
print(Operator_routs)
print("\nOperator Offers same timings")
print(Operator_status)
```

Operator Offers same routs

	Post Date	Post Time	Username	Number	\
11	2025-08-26 00:00:00	22:17:00	Srinivas Rao	NOT	
17	2025-08-26 00:00:00	23:35:00	Srinivas Rao	NOT	
23	2025-08-27 00:00:00	07:11:00	Ravi Kumar Thakur	NOT	
27	2025-08-27 00:00:00	07:22:00	Ravi Kumar Thakur	NOT	
33	2025-07-27 00:00:00	08:09:00	Niaz Ali	NOT	
46	2025-08-27 00:00:00	10:13:00	Niaz Ali	7050480613	
107	2025-08-29 00:00:00	07:15:00	Car Taxi	6205382947	
112	2025-08-29 00:00:00	07:44:00	Car Taxi	6205382947	
167	2025-09-01 00:00:00	09:45:00	Sunil Kumar Singh	7992418286	
175	2025-09-01 00:00:00	11:28:00	Ravi Kumar Thakur	9263355987	
179	2025-09-01 00:00:00	11:44:00	Sunil Kumar Singh	7992418286	
181	2025-09-01 00:00:00	12:02:00	Ravi Kumar Thakur	9263355987	

	Seek Car	Seek Date	Seek Time	Seek Location	\
11	Swift Dzire	2025-08-26 00:00:00	11:00:00	Tata	
17	Swift Dzire	2025-08-26 00:00:00	11:00:00	Tata	
23	Dzire	2025-08-27 00:00:00	02:00:00	Tata	
27	Ertiga	2025-08-27 00:00:00	02:00:00	Tata	
33	Dzire	2025-08-27 00:00:00	06:00:00	Ranchi	
46	Dzire	2025-08-27 00:00:00	06:00:00	Ranchi	
107	Dzire	2025-08-29 00:00:00	10:30:00	Tata	
112	Dzire	2025-08-29 00:00:00	10:30:00	Tata	
167	Dzire	2025-09-01 00:00:00	01:00:00	Tata	
175	Dzire	2025-09-01 00:00:00	06:00:00	Ranchi	
179	Dzire	2025-09-01 00:00:00	01:00:00	Tata	
181	Dzire	2025-09-01 00:00:00	06:00:00	Ranchi	

	Seek Destination	Offer Car	...	seek_dt	\
11	Ranchi	Swift Dzire	...	2025-08-26 11:00:00	
17	Ranchi	Dzire	...	2025-08-26 11:00:00	
23	Ranchi	Ertika	...	2025-08-27 02:00:00	
27	Ranchi	Ertiga	...	2025-08-27 02:00:00	
33	Tata	Dzire	...	2025-08-27 06:00:00	
46	Tata	Dzire	...	2025-08-27 06:00:00	
107	Ranchi	Dzire	...	2025-08-29 10:30:00	
112	Ranchi	Ertica	...	2025-08-29 10:30:00	
167	Ranchi	Dzire	...	2025-09-01 01:00:00	
175	Tata	Dzire	...	2025-09-01 06:00:00	
179	Ranchi	Dzire	...	2025-09-01 01:00:00	
181	Tata	Ertiga	...	2025-09-01 06:00:00	

	offer_dt	post_dt	seek_hour	offer_hour	seek_period	\
11	2025-08-26 09:00:00	2025-08-26 22:17:00	11	9	Morning	
17	2025-08-26 09:00:00	2025-08-26 23:35:00	11	9	Morning	
23	2025-08-27 12:00:00	2025-08-27 07:11:00	2	12	Night	
27	2025-08-27 12:00:00	2025-08-27 07:22:00	2	12	Night	
33	2025-08-27 04:00:00	2025-07-27 08:09:00	6	4	Morning	
46	2025-08-27 04:00:00	2025-08-27 10:13:00	6	4	Morning	
107	2025-08-29 10:30:00	2025-08-29 07:15:00	10	10	Morning	
112	2025-08-29 10:30:00	2025-08-29 07:44:00	10	10	Morning	
167	2025-09-01 03:00:00	2025-09-01 09:45:00	1	3	Night	
175	2025-09-01 04:00:00	2025-09-01 11:28:00	6	4	Morning	
179	2025-09-01 03:00:00	2025-09-01 11:44:00	1	3	Night	
181	2025-09-01 04:00:00	2025-09-01 12:02:00	6	4	Morning	

	offer_period	duration	up_trip	down_trip
11	Morning	0 days 22:00:00	TATA - RANCHI	RANCHI HATIA - TATA
17	Morning	0 days 22:00:00	TATA - RANCHI	RANCHI HATIA - TATA
23	Afternoon	0 days 10:00:00	TATA - RANCHI	RANCHI - TATA
27	Afternoon	0 days 10:00:00	TATA - RANCHI	RANCHI - TATA
33	Night	0 days 22:00:00	RANCHI - TATA	TATA - RANCHI
46	Night	0 days 22:00:00	RANCHI - TATA	TATA - RANCHI
107	Morning	0 days 00:00:00	TATA - RANCHI	RANCHI AIRPORT - TATA
112	Morning	0 days 00:00:00	TATA - RANCHI	RANCHI AIRPORT - TATA
167	Night	0 days 02:00:00	TATA - RANCHI	TATA - RANCHI
175	Night	0 days 22:00:00	RANCHI - TATA	TATA - RANCHI
179	Night	0 days 02:00:00	TATA - RANCHI	TATA - RANCHI
181	Night	0 days 22:00:00	RANCHI - TATA	TATA - RANCHI

[12 rows x 24 columns]

Operator Offers same timings

	Post Date	Post Time	Username	Number	\
11	2025-08-26 00:00:00	22:17:00	Srinivas Rao	NOT	
17	2025-08-26 00:00:00	23:35:00	Srinivas Rao	NOT	
21	2025-08-27 00:00:00	06:47:00	Gayatri Travels	NOT	
23	2025-08-27 00:00:00	07:11:00	Ravi Kumar Thakur	NOT	
27	2025-08-27 00:00:00	07:22:00	Ravi Kumar Thakur	NOT	
33	2025-07-27 00:00:00	08:09:00	Niaz Ali	NOT	
38	2025-08-27 00:00:00	08:46:00	Gayatri Travels	901313153	
46	2025-08-27 00:00:00	10:13:00	Niaz Ali	7050480613	
73	2025-08-28 00:00:00	10:42:00	Ali Bhai Jaan	7903634718	
81	2025-08-28 00:00:00	11:49:00	Abhijit sahu	7903828230	
83	2025-08-28 00:00:00	12:05:00	Abhijit sahu	7903828230	
86	2025-08-28 00:00:00	12:30:00	Ali Bhai Jaan	7903634718	

```
103  2025-08-29 00:00:00  06:33:00            Wasimh           9709108878
107  2025-08-29 00:00:00  07:15:00          Car Taxi           6205382947
112  2025-08-29 00:00:00  07:44:00          Car Taxi           6205382947
113  2025-08-29 00:00:00  07:49:00            Wasimh           9709108878
167  2025-09-01 00:00:00  09:45:00  Sunil Kumar Singh          7992418286
172  2025-09-01 00:00:00  10:19:00  Sunil Kumar Singh          7992418286
175  2025-09-01 00:00:00  11:28:00  Ravi Kumar Thakur          9263355987
179  2025-09-01 00:00:00  11:44:00  Sunil Kumar Singh          7992418286
180  2025-09-01 00:00:00  11:53:00  Sunil Kumar Singh          7992418286
181  2025-09-01 00:00:00  12:02:00  Ravi Kumar Thakur          9263355987

        Seek Car            Seek Date Seek Time   Seek Location  \
11   Swift Dzire  2025-08-26 00:00:00  11:00:00            Tata
17   Swift Dzire  2025-08-26 00:00:00  11:00:00            Tata
21        Ertiga  2025-08-27 00:00:00  04:00:00         Kolkata
23         Dzire  2025-08-27 00:00:00  02:00:00            Tata
27        Ertiga  2025-08-27 00:00:00  02:00:00            Tata
33         Dzire  2025-08-27 00:00:00  06:00:00          Ranchi
38        Ertiga  2025-08-27 00:00:00  04:00:00         Kolkata
46         Dzire  2025-08-27 00:00:00  06:00:00          Ranchi
73         Dzire  2025-08-28 00:00:00  07:00:00  Ranchi Airport
81         Dzire  2025-08-28 00:00:00  01:00:00          Ranchi
83         Dzire  2025-08-28 00:00:00  01:00:00          Ranchi
86         Dzire  2025-08-28 00:00:00  07:00:00  Ranchi Airport
103        Dzire  2025-08-29 00:00:00  10:00:00            Tata
107        Dzire  2025-08-29 00:00:00  10:30:00            Tata
112        Dzire  2025-08-29 00:00:00  10:30:00            Tata
113        Dzire  2025-08-29 00:00:00  10:00:00          Ranchi
167        Dzire  2025-09-01 00:00:00  01:00:00            Tata
172        Dzire  2025-09-01 00:00:00  06:00:00          Ranchi
175        Dzire  2025-09-01 00:00:00  06:00:00          Ranchi
179        Dzire  2025-09-01 00:00:00  01:00:00            Tata
180        Dzire  2025-09-01 00:00:00  06:00:00          Ranchi
181        Dzire  2025-09-01 00:00:00  06:00:00          Ranchi

    Seek Destination    Offer Car  ...             seek_dt  \
11           Ranchi  Swift Dzire  ... 2025-08-26 11:00:00
17           Ranchi        Dzire  ... 2025-08-26 11:00:00
21             Tata       Innova  ... 2025-08-27 04:00:00
23           Ranchi       Ertika  ... 2025-08-27 02:00:00
27           Ranchi      Ertiga   ... 2025-08-27 02:00:00
33             Tata        Dzire  ... 2025-08-27 06:00:00
38             Tata       Innova  ... 2025-08-27 04:00:00
46             Tata        Dzire  ... 2025-08-27 06:00:00
73             Tata       Ertiga  ... 2025-08-28 07:00:00
81             Tata       Ertiga  ... 2025-08-28 01:00:00
83             Tata       Ertiga  ... 2025-08-28 01:00:00
86             Tata       Ertiga  ... 2025-08-28 07:00:00
103          Ranchi        Dzire  ... 2025-08-29 10:00:00
107          Ranchi        Dzire  ... 2025-08-29 10:30:00
112          Ranchi       Ertica  ... 2025-08-29 10:30:00
113            Tata        Dzire  ... 2025-08-29 10:00:00
167          Ranchi        Dzire  ... 2025-09-01 01:00:00
172            Tata        Dzire  ... 2025-09-01 06:00:00
175            Tata        Dzire  ... 2025-09-01 06:00:00
179          Ranchi        Dzire  ... 2025-09-01 01:00:00
180            Tata        Dzire  ... 2025-09-01 06:00:00
181            Tata       Ertiga  ... 2025-09-01 06:00:00

              offer_dt             post_dt seek_hour offer_hour seek_period  \
11  2025-08-26 09:00:00 2025-08-26 22:17:00        11          9     Morning
17  2025-08-26 09:00:00 2025-08-26 23:35:00        11          9     Morning
21  2025-08-27 10:00:00 2025-08-27 06:47:00         4         10       Night
23  2025-08-27 12:00:00 2025-08-27 07:11:00         2         12       Night
27  2025-08-27 12:00:00 2025-08-27 07:22:00         2         12       Night
33  2025-08-27 04:00:00 2025-07-27 08:09:00         6          4     Morning
38  2025-08-27 10:00:00 2025-08-27 08:46:00         4         10       Night
46  2025-08-27 04:00:00 2025-08-27 10:13:00         6          4     Morning
73  2025-08-28 04:00:00 2025-08-28 10:42:00         7          4     Morning
81  2025-08-28 05:00:00 2025-08-28 11:49:00         1          5       Night
83  2025-08-28 05:00:00 2025-08-28 12:05:00         1          5       Night
86  2025-08-28 04:00:00 2025-08-28 12:30:00         7          4     Morning
103 2025-08-29 08:00:00 2025-08-29 06:33:00        10          8     Morning
107 2025-08-29 10:30:00 2025-08-29 07:15:00        10         10     Morning
112 2025-08-29 10:30:00 2025-08-29 07:44:00        10         10     Morning
113 2025-08-29 08:00:00 2025-08-29 07:49:00        10          8     Morning
167 2025-09-01 03:00:00 2025-09-01 09:45:00         1          3       Night
172 2025-09-01 04:00:00 2025-09-01 10:19:00         6          4     Morning
175 2025-09-01 04:00:00 2025-09-01 11:28:00         6          4     Morning
179 2025-09-01 03:00:00 2025-09-01 11:44:00         1          3       Night
180 2025-09-01 04:00:00 2025-09-01 11:53:00         6          4     Morning
181 2025-09-01 04:00:00 2025-09-01 12:02:00         6          4     Morning

    offer_period        duration               up_trip             down_trip
11       Morning 0 days 22:00:00          TATA - RANCHI    RANCHI HATIA - TATA
17       Morning 0 days 22:00:00          TATA - RANCHI    RANCHI HATIA - TATA
21       Morning 0 days 06:00:00         KOLKATA - TATA         KOLKATA - TATA
```

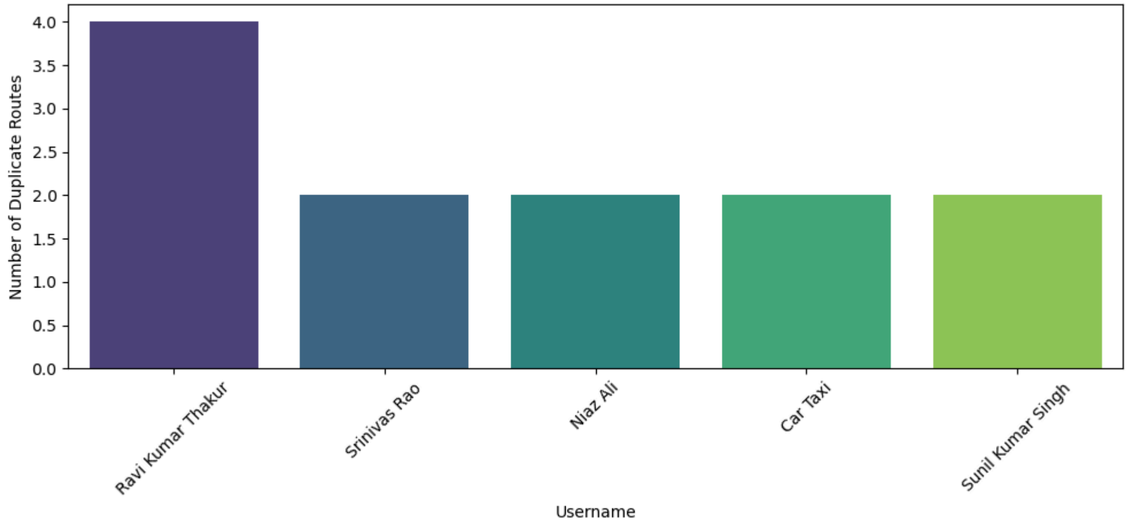
23	Afternoon	0 days 10:00:00	TATA - RANCHI	RANCHI - TATA	
27	Afternoon	0 days 10:00:00	TATA - RANCHI	RANCHI - TATA	
33	Night	0 days 22:00:00	RANCHI - TATA	TATA - RANCHI	
38	Morning	0 days 06:00:00	KOLKATA - TATA	KOLKATA - RANCHI	
46	Night	0 days 22:00:00	RANCHI - TATA	TATA - RANCHI	
73	Night	0 days 21:00:00	RANCHI AIRPORT - TATA	RANCHI - TATA	
81	Morning	0 days 04:00:00	RANCHI - TATA	RANCHI - TATA	
83	Morning	0 days 04:00:00	RANCHI - TATA	RANCHI - RANCHI	
86	Night	0 days 21:00:00	RANCHI AIRPORT - TATA	RANCHI AIRPORT - TATA	
103	Morning	0 days 22:00:00	TATA - RANCHI	RANCHI AIRPORT - TATA	
107	Morning	0 days 00:00:00	TATA - RANCHI	RANCHI AIRPORT - TATA	
112	Morning	0 days 00:00:00	TATA - RANCHI	RANCHI AIRPORT - TATA	
113	Morning	0 days 22:00:00	RANCHI - TATA	RANCHI - TATA	
167	Night	0 days 02:00:00	TATA - RANCHI	TATA - RANCHI	
172	Night	0 days 22:00:00	RANCHI - TATA	RANCHI - TATA	
175	Night	0 days 22:00:00	RANCHI - TATA	TATA - RANCHI	
179	Night	0 days 02:00:00	TATA - RANCHI	TATA - RANCHI	
180	Night	0 days 22:00:00	RANCHI - TATA	RANCHI AIRPORT - TATA	
181	Night	0 days 22:00:00	RANCHI - TATA	TATA - RANCHI	

[22 rows x 24 columns]

Data Visualization

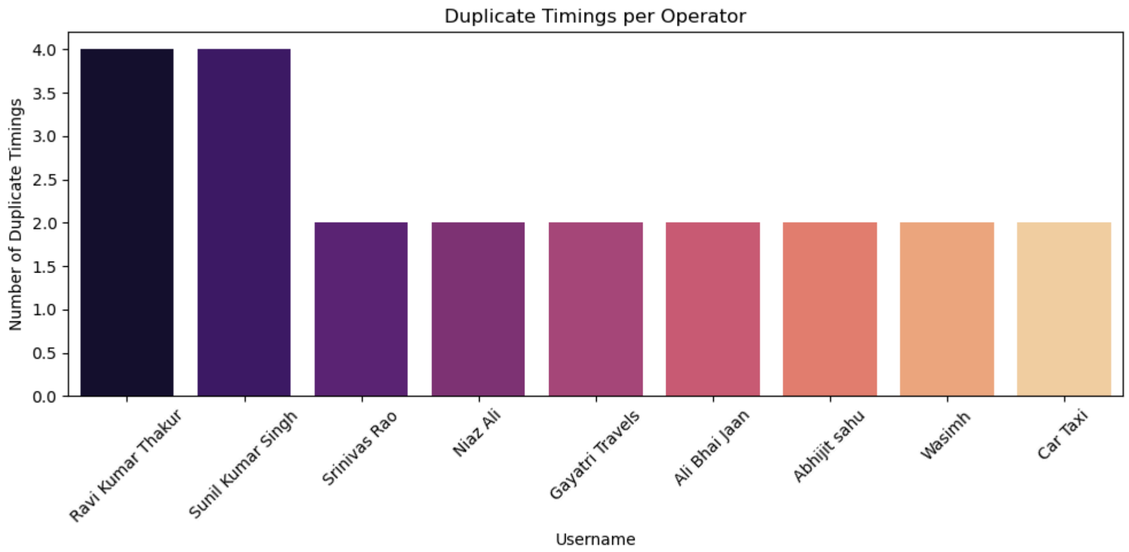
In [49]:
```
# Count duplicates by Username for routes
dup_route_counts = Operator_routs['Username'].value_counts()

plt.figure(figsize=(10,5))
sns.barplot(x=dup_route_counts.index, y=dup_route_counts.values, palette="viridis")
plt.xticks(rotation=45)
plt.title("Duplicate Routes per Operator")
plt.xlabel("Username")
plt.ylabel("Number of Duplicate Routes")
plt.tight_layout()
plt.show()
```

```
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\728053615.py:5: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

  sns.barplot(x=dup_route_counts.index, y=dup_route_counts.values, palette="viridis")
```

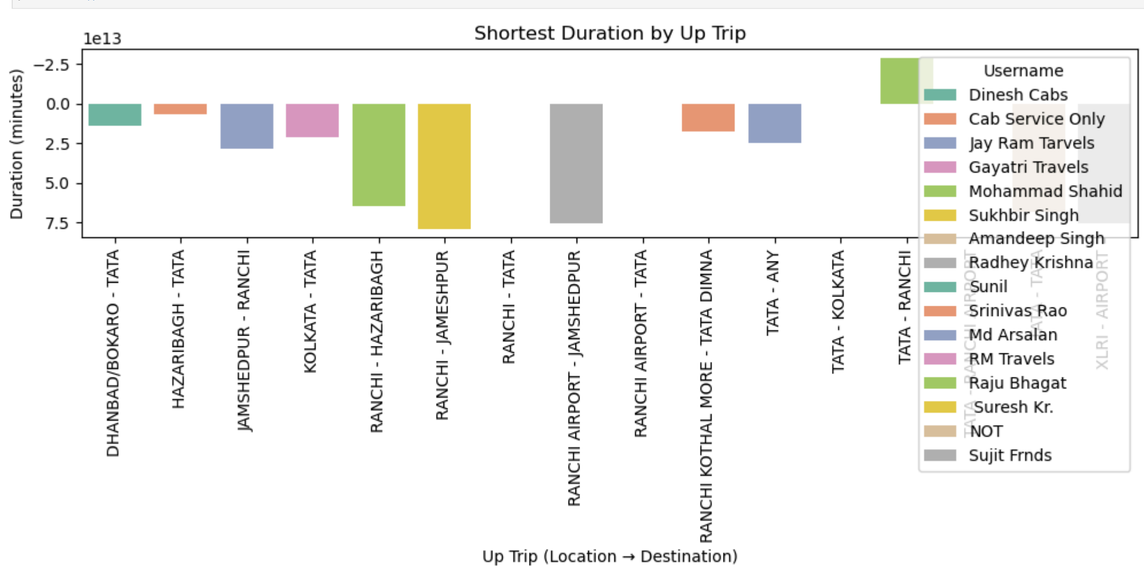


In [50]:
```
# Count duplicates by Username for timings
dup_time_counts = Operator_status['Username'].value_counts()

plt.figure(figsize=(10,5))
sns.barplot(x=dup_time_counts.index, y=dup_time_counts.values, palette="magma")
plt.xticks(rotation=45)
plt.title("Duplicate Timings per Operator")
plt.xlabel("Username")
plt.ylabel("Number of Duplicate Timings")
plt.tight_layout()
plt.show()
```

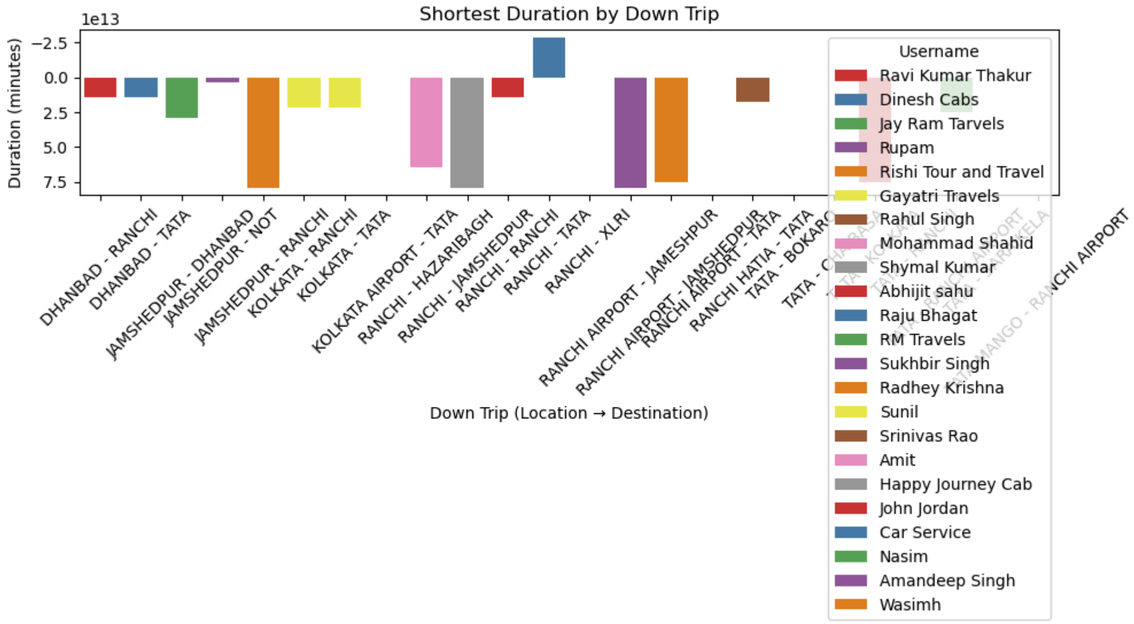
```
C:\Users\MALINI\AppData\Local\Temp\ipykernel_3476\3858970559.py:5: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

  sns.barplot(x=dup_time_counts.index, y=dup_time_counts.values, palette="magma")
```



In [60]:
```python
# Shortest Up Trips
plt.figure(figsize=(10,5))
sns.barplot(data=shortest_up, x="up_trip", y="duration", hue="Username", dodge=False, palette="Set2")
plt.xticks(rotation=90)
plt.title("Shortest Duration by Up Trip")
plt.xlabel("Up Trip (Location → Destination)")
plt.ylabel("Duration (minutes)")
plt.tight_layout()
plt.show()
```



In [54]:
```python
# Shortest Down Trips
plt.figure(figsize=(10,5))
sns.barplot(data=shortest_down, x="down_trip", y="duration", hue="Username", dodge=False, palette="Set1")
plt.xticks(rotation=45)
plt.title("Shortest Duration by Down Trip")
plt.xlabel("Down Trip (Location → Destination)")
plt.ylabel("Duration (minutes)")
plt.tight_layout()
plt.show()
```

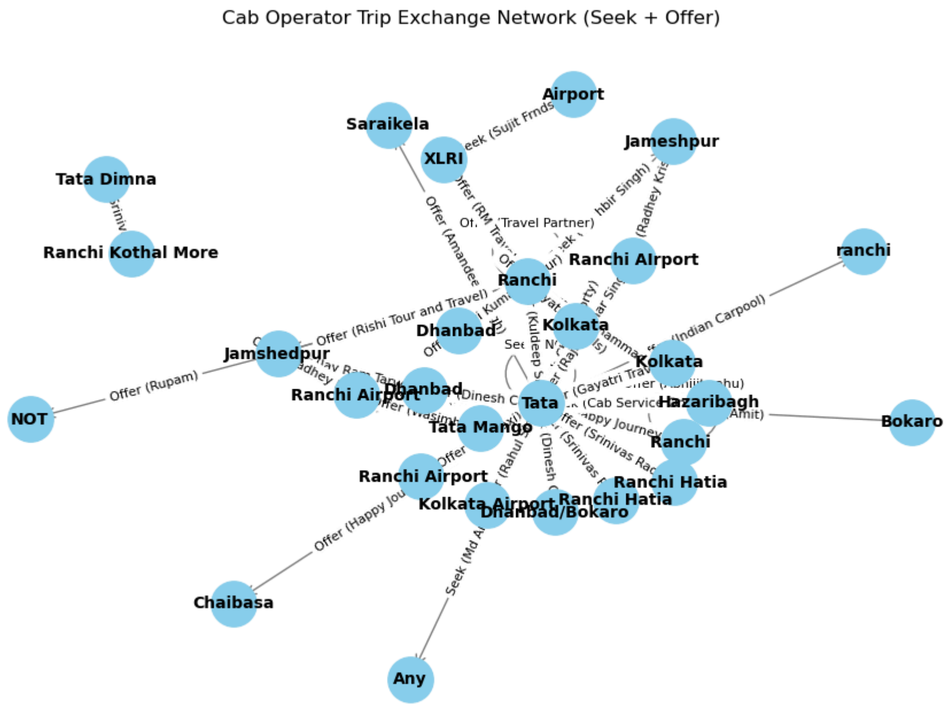


In [59]:

```python
import networkx as nx

# 1. Build graph: each edge is a trip (Seek or Offer)
G = nx.DiGraph()

# Add SEEK trips
for _, row in data.iterrows():
    if pd.notnull(row['Seek Location']) and pd.notnull(row['Seek Destination']):
        G.add_edge(
            row['Seek Location'],
            row['Seek Destination'],
            label=f"Seek ({row['Username']})"
        )

# Add OFFER trips
for _, row in data.iterrows():
    if pd.notnull(row['Offer Location']) and pd.notnull(row['Offer Destination']):
        G.add_edge(
            row['Offer Location'],
            row['Offer Destination'],
            label=f"Offer ({row['Username']})"
        )

# 2. Find chains of length >= 3 (4+ nodes)
chains = []
for source in G.nodes():
    for target in G.nodes():
        if source != target:  # avoid same start and end
            for path in nx.all_simple_paths(G, source=source, target=target, cutoff=5):
                if len(path) >= 4:  # means 3+ edges
                    chains.append(path)

print("\nDetected Chains (3-4+ trips):")
for chain in chains[:5]:  # show only first 5
    print(" → ".join(chain))

# 3. Visualize Graph
plt.figure(figsize=(12, 8))
pos = nx.spring_layout(G, k=0.5, iterations=50)

# Draw nodes
nx.draw_networkx_nodes(G, pos, node_size=800, node_color="skyblue")

# Draw edges
nx.draw_networkx_edges(G, pos, arrowstyle="->", arrowsize=15, edge_color="gray")

# Draw labels
nx.draw_networkx_labels(G, pos, font_size=10, font_weight="bold")

# Edge labels (optional: show Seek/Offer + Username)
edge_labels = nx.get_edge_attributes(G, 'label')
nx.draw_networkx_edge_labels(G, pos, edge_labels=edge_labels, font_size=8)

plt.title("Cab Operator Trip Exchange Network (Seek + Offer)")
```

```
plt.axis("off")
plt.show()
```

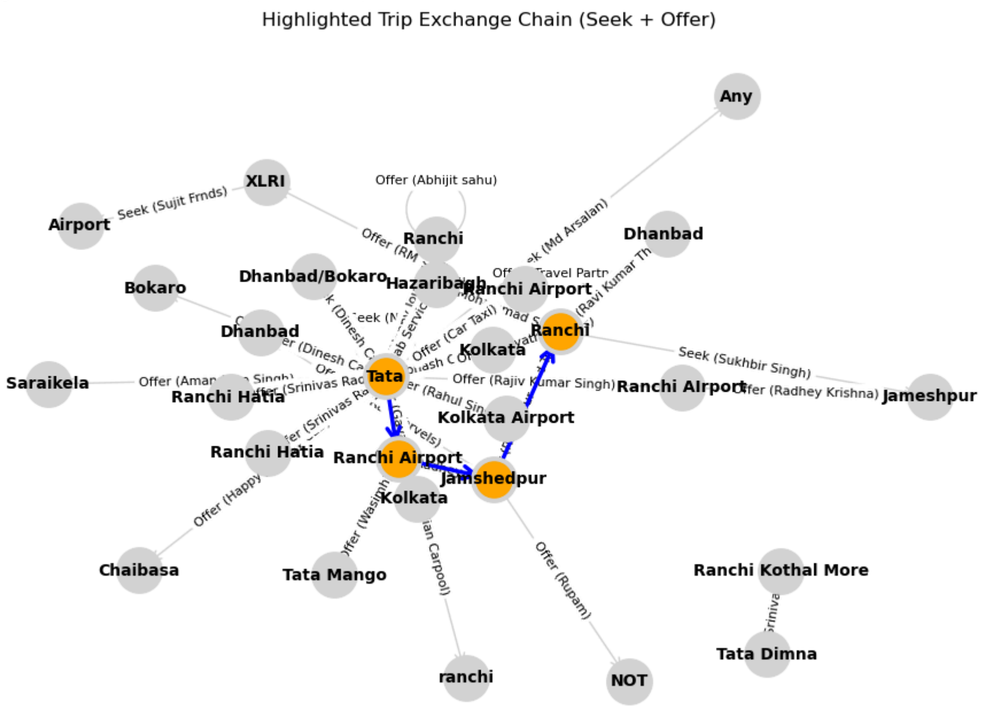
```
Detected Chains (3–4+ trips):
Tata → Ranchi Airport → Jamshedpur → Ranchi
Tata → Kolkata → Ranchi → Jamshedpur
Tata → Ranchi Airport → Jamshedpur → Ranchi → Hazaribagh
Tata → Kolkata → Ranchi → Hazaribagh
Tata → Ranchi Airport → Jamshedpur → Ranchi → Jameshpur
```



In [75]:
```
import networkx as nx
import matplotlib.pyplot as plt
import pandas as pd

# 1. Build graph: each edge is a trip (Seek or Offer)
G = nx.DiGraph()

# Add SEEK trips
for _, row in data.iterrows():
    if pd.notnull(row['Seek Location']) and pd.notnull(row['Seek Destination']):
        G.add_edge(
            row['Seek Location'],
            row['Seek Destination'],
            label=f"Seek ({row['Username']})"
        )

# Add OFFER trips
for _, row in data.iterrows():
    if pd.notnull(row['Offer Location']) and pd.notnull(row['Offer Destination']):
        G.add_edge(
            row['Offer Location'],
            row['Offer Destination'],
            label=f"Offer ({row['Username']})"
        )

# 2. Find chains of length >= 3 (4+ nodes)
chains = []
for source in G.nodes():
    for target in G.nodes():
        if source != target:
            for path in nx.all_simple_paths(G, source=source, target=target, cutoff=5):
                if len(path) >= 4:
                    chains.append(path)

print("\nDetected Chains (3–4+ trips):")
for chain in chains[:5]:
    print(" → ".join(chain))

# Pick one chain to highlight (first one if available)
```

```
highlight_chain = chains[0] if chains else []

# 3. Visualize Graph
plt.figure(figsize=(12, 8))
pos = nx.spring_layout(G, k=0.5, iterations=50)

# Draw all nodes & edges in gray
nx.draw_networkx_nodes(G, pos, node_size=800, node_color="lightgray")
nx.draw_networkx_edges(G, pos, arrowstyle="->", arrowsize=15, edge_color="lightgray")
nx.draw_networkx_labels(G, pos, font_size=10, font_weight="bold")

# Draw edge labels (Seek/Offer + Username)
edge_labels = nx.get_edge_attributes(G, 'label')
nx.draw_networkx_edge_labels(G, pos, edge_labels=edge_labels, font_size=8)

# Highlight the chain in color (red nodes + blue edges)
if highlight_chain:
    chain_edges = list(zip(highlight_chain[:-1], highlight_chain[1:]))  # consecutive pairs
    nx.draw_networkx_nodes(G, pos, nodelist=highlight_chain, node_size=500, node_color="orange")
    nx.draw_networkx_edges(G, pos, edgelist=chain_edges, arrowstyle="->", arrowsize=20, edge_color="blue", width=2.5)

    print("\nHighlighted Chain:", " → ".join(highlight_chain))

plt.title("Highlighted Trip Exchange Chain (Seek + Offer)")
plt.axis("off")
plt.show()
```

```
Detected Chains (3–4+ trips):
Tata → Ranchi Airport → Jamshedpur → Ranchi
Tata → Kolkata → Ranchi → Jamshedpur
Tata → Ranchi Airport → Jamshedpur → Ranchi → Hazaribagh
Tata → Kolkata → Ranchi → Hazaribagh
Tata → Ranchi Airport → Jamshedpur → Ranchi → Jameshpur

Highlighted Chain: Tata → Ranchi Airport → Jamshedpur → Ranchi
```



In [ ]: