

# Airlines\_Flights\_Data\_Analysis

```
In [1]: print("Hello Abel!")
```

Hello Abel!

```
In [2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
```

```
In [3]: airln = pd.read_csv(r"C:\Users\Abel\Desktop\Data Analytics 2025\airlines_flights_data
```

```
In [4]: airln
```

Out[4]:

	index	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	
	0	0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai
	1	1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai
	2	2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai
	3	3	Vistara	UK-995	Delhi	Morning	zero	Afternoon	Mumbai
	4	4	Vistara	UK-963	Delhi	Morning	zero	Morning	Mumbai
	...	...	...	...	...	...	...	...	...
	300148	300148	Vistara	UK-822	Chennai	Morning	one	Evening	Hyderabad
	300149	300149	Vistara	UK-826	Chennai	Afternoon	one	Night	Hyderabad
	300150	300150	Vistara	UK-832	Chennai	Early_Morning	one	Night	Hyderabad
	300151	300151	Vistara	UK-828	Chennai	Early_Morning	one	Evening	Hyderabad
	300152	300152	Vistara	UK-822	Chennai	Morning	one	Evening	Hyderabad

300153 rows × 12 columns

## Processing & EDA

```
In [5]: airln.drop(columns = "index", inplace = True)
```

```
In [6]: airln
```

Out[6]:

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy
1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy
2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy
3	Vistara	UK-995	Delhi	Morning	zero	Afternoon	Mumbai	Economy
4	Vistara	UK-963	Delhi	Morning	zero	Morning	Mumbai	Economy
...	...	...	...	...	...	...	...	...
300148	Vistara	UK-822	Chennai	Morning	one	Evening	Hyderabad	Business
300149	Vistara	UK-826	Chennai	Afternoon	one	Night	Hyderabad	Business
300150	Vistara	UK-832	Chennai	Early_Morning	one	Night	Hyderabad	Business
300151	Vistara	UK-828	Chennai	Early_Morning	one	Evening	Hyderabad	Business
300152	Vistara	UK-822	Chennai	Morning	one	Evening	Hyderabad	Business

300153 rows × 11 columns



```
In [7]: airln.describe()
```

Out[7]:

	duration	days_left	price
count	300153.000000	300153.000000	300153.000000
mean	12.221021	26.004751	20889.660523
std	7.191997	13.561004	22697.767366
min	0.830000	1.000000	1105.000000
25%	6.830000	15.000000	4783.000000
50%	11.250000	26.000000	7425.000000
75%	16.170000	38.000000	42521.000000
max	49.830000	49.000000	123071.000000

```
In [8]: airln[airln['duration'] == 49.830000]
```

```
Out[8]:
```

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class
193889	Air_India	AI-672	Chennai	Evening	two_or_more	Evening	Bangalore	Economy
194359	Air_India	AI-672	Chennai	Evening	one	Evening	Bangalore	Economy

```
In [9]: airln[airln['days_left'] == 15]
```

```
Out[9]:
```

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class
2504	Vistara	UK-975	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy
2505	Vistara	UK-953	Delhi	Night	zero	Night	Mumbai	Economy
2506	Vistara	UK-927	Delhi	Morning	zero	Morning	Mumbai	Economy
2507	Vistara	UK-993	Delhi	Afternoon	zero	Afternoon	Mumbai	Economy
2508	Vistara	UK-951	Delhi	Afternoon	zero	Evening	Mumbai	Economy
...	...	...	...	...	...	...	...	...
298581	Vistara	UK-834	Chennai	Evening	one	Afternoon	Hyderabad	Business
298582	Vistara	UK-834	Chennai	Evening	one	Evening	Hyderabad	Business
298583	Vistara	UK-834	Chennai	Evening	one	Evening	Hyderabad	Business
298584	Air_India	AI-430	Chennai	Morning	one	Night	Hyderabad	Business
298585	Air_India	AI-440	Chennai	Early_Morning	one	Night	Hyderabad	Business

6340 rows × 11 columns

```
In [10]: airln.groupby('airline')['price'].mean()
```

```
Out[10]:
```

airline	price
AirAsia	4091.072742
Air_India	23507.019112
GO_FIRST	5652.007595
Indigo	5324.216303
SpiceJet	6179.278881
Vistara	30396.536302

Name: price, dtype: float64

```
In [11]: airline.isnull().sum()
```

```
Out[11]: airline          0
flight          0
source_city     0
departure_time  0
stops          0
arrival_time    0
destination_city 0
class          0
duration        0
days_left      0
price          0
dtype: int64
```

## Q.1. what are the airlines with their frequency

```
In [12]: airline
```

```
Out[12]:
```

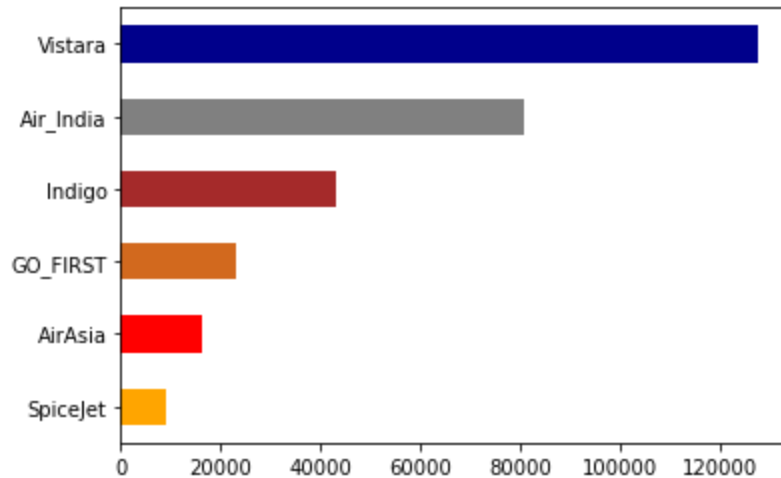
	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy
1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy
2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy
3	Vistara	UK-995	Delhi	Morning	zero	Afternoon	Mumbai	Economy
4	Vistara	UK-963	Delhi	Morning	zero	Morning	Mumbai	Economy
...	...	...	...	...	...	...	...	...
300148	Vistara	UK-822	Chennai	Morning	one	Evening	Hyderabad	Business
300149	Vistara	UK-826	Chennai	Afternoon	one	Night	Hyderabad	Business
300150	Vistara	UK-832	Chennai	Early_Morning	one	Night	Hyderabad	Business
300151	Vistara	UK-828	Chennai	Early_Morning	one	Evening	Hyderabad	Business
300152	Vistara	UK-822	Chennai	Morning	one	Evening	Hyderabad	Business

300153 rows × 11 columns

```
In [13]: airline['airline'].nunique(), airline.airline.unique()
```

```
Out[13]: (6,
array(['SpiceJet', 'AirAsia', 'Vistara', 'GO_FIRST', 'Indigo',
      'Air_India'], dtype=object))
```

```
In [14]: airln['airline'].value_counts(ascending = True).plot.barh(color = ['orange', 'red', 'choc', 'blue', 'gray'],
plt.show()
```



## 2. Show Bar Graphs representing the Departure Time & Arrival Time.

```
In [15]: airln.head(3)
```

```
Out[15]:
```

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	durati
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2
1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2
2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy	2

```
In [16]: pd.DataFrame({
    "Dep. Counts": airln['departure_time'].value_counts(ascending = True),
    "Arri. Counts": airln['arrival_time'].value_counts()
})
```

Out[16]:

	Dep. Counts	Arri. Counts
Afternoon	47794	38139
Early_Morning	66790	15417
Evening	65102	78323
Late_Night	1306	14001
Morning	71146	62735
Night	48015	91538

```

In [17]: plt.figure(figsize = (12,3))

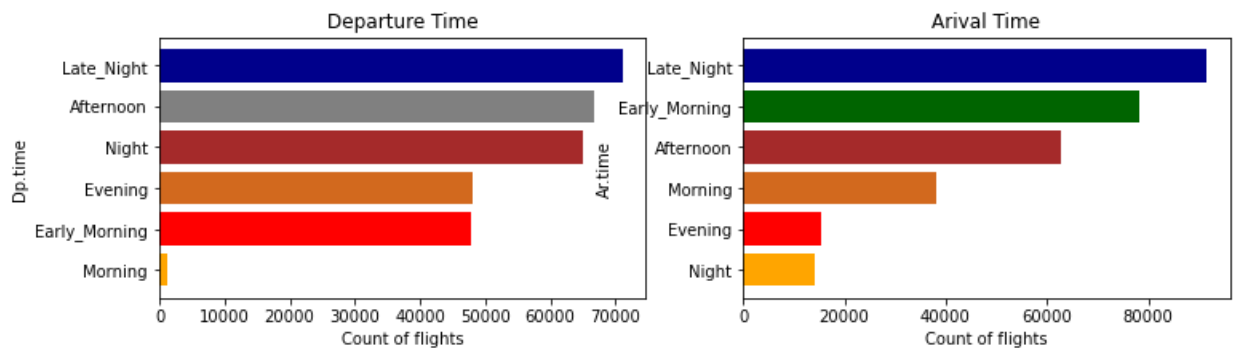
plt.subplot(1,2,1)

plt.barh(airln['departure_time'].value_counts().index ,
          airln['departure_time'].value_counts(ascending=True).values, color = ['orange',
plt.title('Departure Time')
plt.ylabel('Dp.time')
plt.xlabel('Count of flights')

plt.subplot(1,2,2)

plt.barh(airln['arrival_time'].value_counts().index ,
          airln['arrival_time'].value_counts(ascending=True).values, color = ['orange',
plt.title('Arival Time')
plt.ylabel('Ar.time')
plt.xlabel('Count of flights')
plt.show()

```



### 3. Show Bar Graphs representing the Source City & Destination City

```

In [18]: airln.head()

```

Out[18]:

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	durati
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2
1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2
2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy	2
3	Vistara	UK-995	Delhi	Morning	zero	Afternoon	Mumbai	Economy	2
4	Vistara	UK-963	Delhi	Morning	zero	Morning	Mumbai	Economy	2

In [19]: `airln['source_city'].nunique() , airln['source_city'].unique()`

Out[19]: (6,  
array(['Delhi', 'Mumbai', 'Bangalore', 'Kolkata', 'Hyderabad', 'Chennai'],  
dtype=object))

In [20]: `airln['destination_city'].nunique() , airln['destination_city'].unique()`

Out[20]: (6,  
array(['Mumbai', 'Bangalore', 'Kolkata', 'Hyderabad', 'Chennai', 'Delhi'],  
dtype=object))

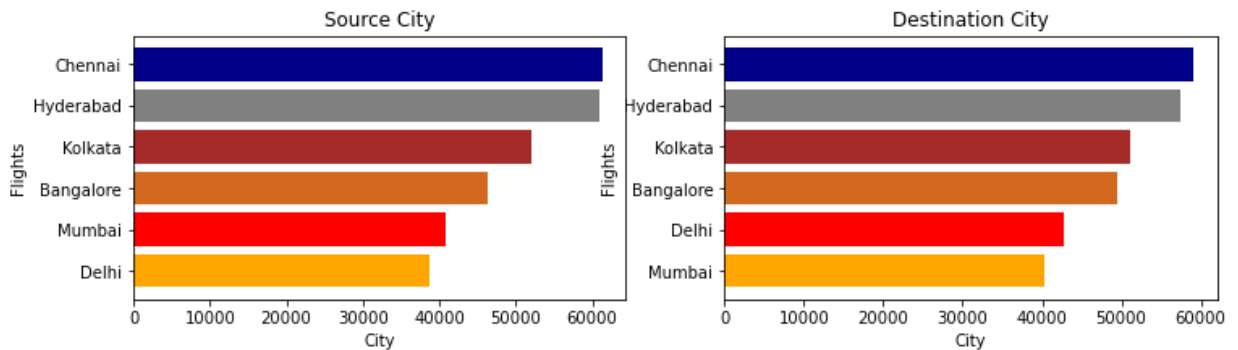
In [21]: `pd.DataFrame({  
 "Sourc. Counts": airln['source_city'].value_counts(ascending = True),  
 "Dest. Counts": airln['destination_city'].value_counts()  
})`

Out[21]:

	Sourc. Counts	Dest. Counts
<b>Bangalore</b>	52061	51068
<b>Chennai</b>	38700	40368
<b>Delhi</b>	61343	57360
<b>Hyderabad</b>	40806	42726
<b>Kolkata</b>	46347	49534
<b>Mumbai</b>	60896	59097

In [22]: `plt.figure(figsize = (12,3))  
plt.subplot(1,2,1)  
  
plt.barh(airln['source_city'].value_counts().index ,  
 airln['source_city'].value_counts(ascending = True).values, color = ['orange'  
plt.title('Source City')  
plt.ylabel('Flights')  
plt.xlabel('City')  
  
plt.subplot(1,2,2)  
plt.barh(airln['destination_city'].value_counts().index ,  
 airln['destination_city'].value_counts(ascending = True).values, color = ['or`

```
plt.title('Destination City')
plt.ylabel('Flights')
plt.xlabel('City')
plt.show()
```



## 4. Does price varie with airlines?

```
In [23]: airln.head(3)
```

```
Out[23]:
```

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	durati
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2
1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2
2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy	2

```
In [24]: airln.groupby('airline')['price'].mean().sort_values(ascending = True)
```

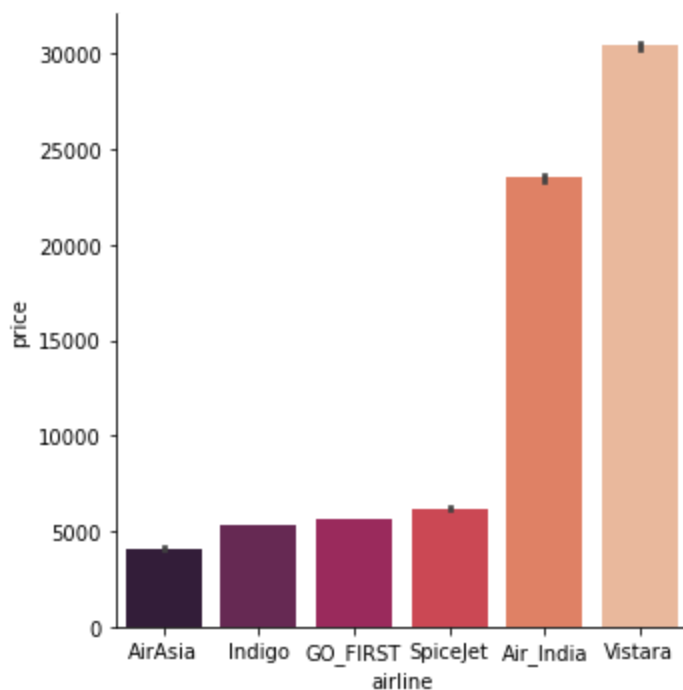
```
Out[24]:
```

airline	price
AirAsia	4091.072742
Indigo	5324.216303
GO_FIRST	5652.007595
SpiceJet	6179.278881
Air_India	23507.019112
Vistara	30396.536302

Name: price, dtype: float64

```
In [25]: order = airln.groupby('airline')['price'].mean().sort_values().index
sns.catplot( data = airln, x = 'airline', y='price', kind='bar', palette = 'rocket', or
plt.show()
```





## 5. Does ticket price change based on the departure time and arrival time?

In [26]: `airln.head(3)`

Out[26]:

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	durati
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2
1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2
2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy	2

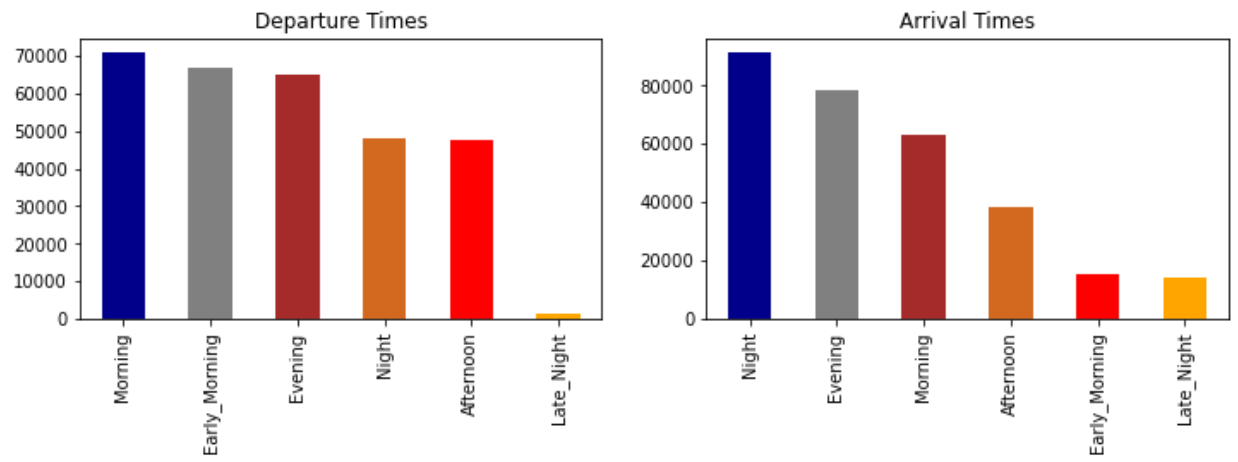
```
In [27]: dep_avg = airln.groupby('departure_time')['price'].mean()
arr_avg = airln.groupby('arrival_time')['price'].mean()

comparison = pd.DataFrame({
    "Avg Price by Departure Time": dep_avg,
    "Avg Price by Arrival Time": arr_avg
})
comparison
```

Out[27]:

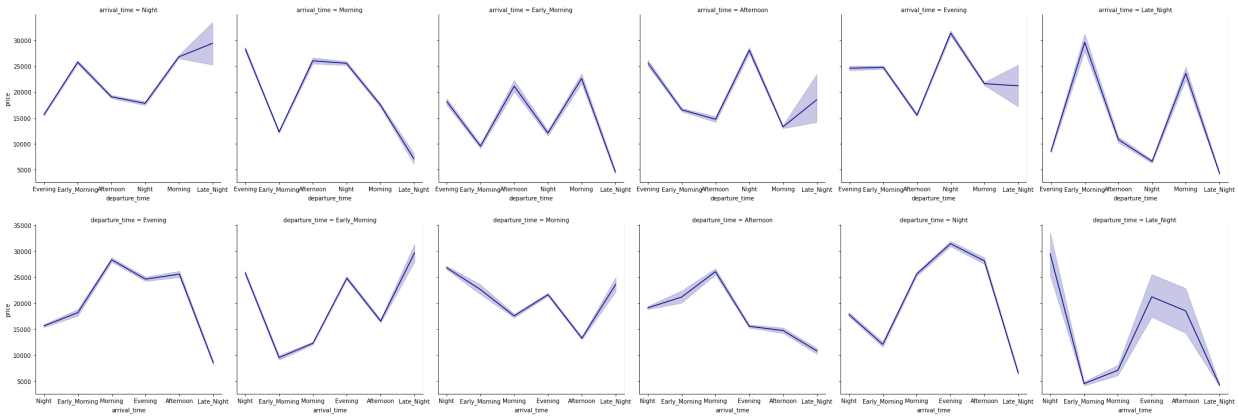
	Avg Price by Departure Time	Avg Price by Arrival Time
Afternoon	18179.203331	18494.598993
Early_Morning	20370.676718	14993.139521
Evening	21232.361894	23044.371615
Late_Night	9295.299387	11284.906078
Morning	21630.760254	22231.076098
Night	23062.146808	21586.758341

```
In [28]: fig, axes = plt.subplots(1, 2, figsize=(12,3))
airln['departure_time'].value_counts().plot(kind='bar', ax=axes[0], title="Departure Times")
airln['arrival_time'].value_counts().plot(kind='bar', ax=axes[1], title="Arrival Times")
plt.show()
```



```
In [29]: sns.relplot(data=airln, x='departure_time', y='price', col = 'arrival_time', kind = 'line')
sns.relplot(data=airln, x='arrival_time', y='price', col = 'departure_time', kind = 'line')
```

Out[29]: <seaborn.axisgrid.FacetGrid at 0x19e307300a0>



## 6. How does price change with change in Source and Destination?

```
In [30]: airln.head(1)
```

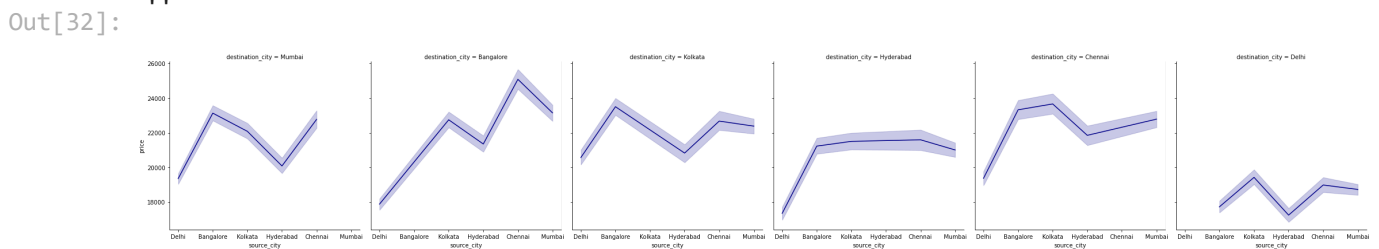
```
Out[30]:
```

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	duration
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2.1

```
In [31]: airln['source_city'].nunique() , airln['source_city'].unique()
```

```
Out[31]: (6,
array(['Delhi', 'Mumbai', 'Bangalore', 'Kolkata', 'Hyderabad', 'Chennai'],
      dtype=object))
```

```
In [32]: sns.relplot(data = airln, x = 'source_city', y = 'price', col = 'destination_city' , kind = 'line')
```



## 7. How is the price affected when tickets are bought in just 1 or 2 days before departure?

```
In [33]: airln.head(1)
```

```
Out[33]:
```

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	duration
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2.1

```
In [34]: airln['days_left'].nunique() , airln['days_left'].unique()
```

```
Out[34]: (49,
array([ 1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16, 17,
        18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34,
        35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49],
      dtype=int64))
```

```
In [35]: filtered = airln[airln['days_left'].isin([1, 2, 7, 15, 30, 45, 49])]
result = filtered.groupby('days_left')['price'].mean()
print(result)
```

```

days_left
1      21591.867151
2      30211.299801
7      25588.367351
15     21952.540852
30     19567.580834
45     19199.876307
49     18992.971888
Name: price, dtype: float64

```

```

In [36]: filtered = airln[airln['days_left'].isin([1, 2])]
result = filtered.groupby('days_left')['price'].mean()
print(result)

```

```

days_left
1      21591.867151
2      30211.299801
Name: price, dtype: float64

```

```

In [37]: airln['is_last_minute'] = airln['days_left'].isin([1, 2])
comparison = airln.groupby('is_last_minute')['price'].mean()
print(comparison)

```

```

is_last_minute
False      20757.498484
True       27421.169326
Name: price, dtype: float64

```

```

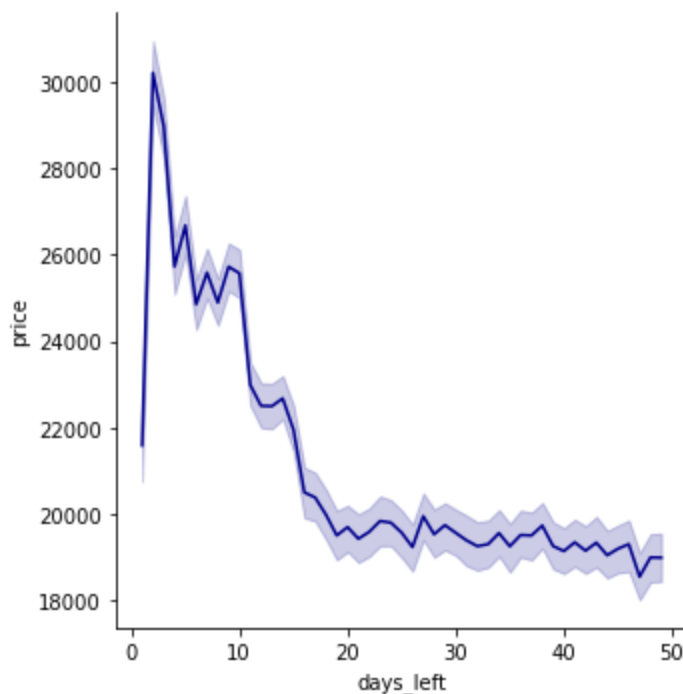
In [38]: sns.relplot(data = airln, x = 'days_left', y = 'price' , kind = 'line' , color = 'darkblue'
;

```

```

Out[38]:

```



## 8. How does the ticket price vary between Economy and Business class?

```
In [39]: airln.head(1)
```

```
Out[39]:
```

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	duration
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2.1

```
In [40]: airln['class'].nunique() , airln['class'].unique()
```

```
Out[40]: (2, array(['Economy', 'Business'], dtype=object))
```

```
In [41]: airln.groupby('class')['price'].mean()
```

```
Out[41]:
```

class	price
Business	52540.081124
Economy	6572.342383

Name: price, dtype: float64

```
In [42]: plt.figure(figsize=(8,2))
airln.groupby('class')['price'].mean().sort_values(ascending = True).plot.bar(color =

plt.title('Class - Price' )
plt.xlabel('Class')
plt.ylabel('Price')
plt.show()
```



## 9. What will be the Average Price of Vistara airline for a flight from Delhi to Hyderabad in Business Class?

```
In [43]: airln.head(3)
```

Out[43]:

	airline	flight	source_city	departure_time	stops	arrival_time	destination_city	class	durati
0	SpiceJet	SG-8709	Delhi	Evening	zero	Night	Mumbai	Economy	2
1	SpiceJet	SG-8157	Delhi	Early_Morning	zero	Morning	Mumbai	Economy	2
2	AirAsia	I5-764	Delhi	Early_Morning	zero	Early_Morning	Mumbai	Economy	2

In [44]:

```
filter = airln[(airln['airline'] == 'Vistara') & (airln['source_city'] == 'Delhi') &
               (airln['destination_city'] == 'Hyderabad') & (airln['class'] == 'Busine
filter['price'].mean()
```

Out[44]: 47939.840361445786

In [ ]: