Untitled1

May 20, 2023

[2]:
```
from matplotlib import pyplot as plot
```

[3]:
```
import numpy as np
```

[4]:
```
import pandas as pd
```

[6]:
```
document = pd.read_csv(r'C:\Users\bonin\Downloads\booking_of_hotel.csv')
```

[7]:
```
df = pd.DataFrame(document)
```

[8]:
```
df.head()
```

[8]:

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	\
0	Resort Hotel	0	342	2015	July	
1	Resort Hotel	0	737	2015	July	
2	Resort Hotel	0	7	2015	July	
3	Resort Hotel	0	13	2015	July	
4	Resort Hotel	0	14	2015	July	

	arrival_date_week_number	arrival_date_day_of_month	\
0	27	1	
1	27	1	
2	27	1	
3	27	1	
4	27	1	

	stays_in_weekend_nights	stays_in_week_nights	adults	…	deposit_type	\
0	0	0	2	…	No Deposit	
1	0	0	2	…	No Deposit	
2	0	1	1	…	No Deposit	
3	0	1	1	…	No Deposit	
4	0	2	2	…	No Deposit	

	agent	company	days_in_waiting_list	customer_type	adr	\
0	NaN	NaN	0	Transient	0.0	
1	NaN	NaN	0	Transient	0.0	
2	NaN	NaN	0	Transient	75.0	
3	304.0	NaN	0	Transient	75.0	

```
4  240.0     NaN                    0     Transient  98.0

   required_car_parking_spaces  total_of_special_requests  reservation_status  \
0                            0                          0           Check-Out
1                            0                          0           Check-Out
2                            0                          0           Check-Out
3                            0                          0           Check-Out
4                            0                          1           Check-Out

  reservation_status_date
0              2015-07-01
1              2015-07-01
2              2015-07-02
3              2015-07-02
4              2015-07-03

[5 rows x 32 columns]
```

[11]:
```
number_of_hotel_bookings = len(df.index)
print(number_of_hotel_bookings)
```

```
119390
```

[14]:
```
number_of_cancellations = len(df[df['is_canceled'] != 0])
print(number_of_cancellations)
```

```
44224
```

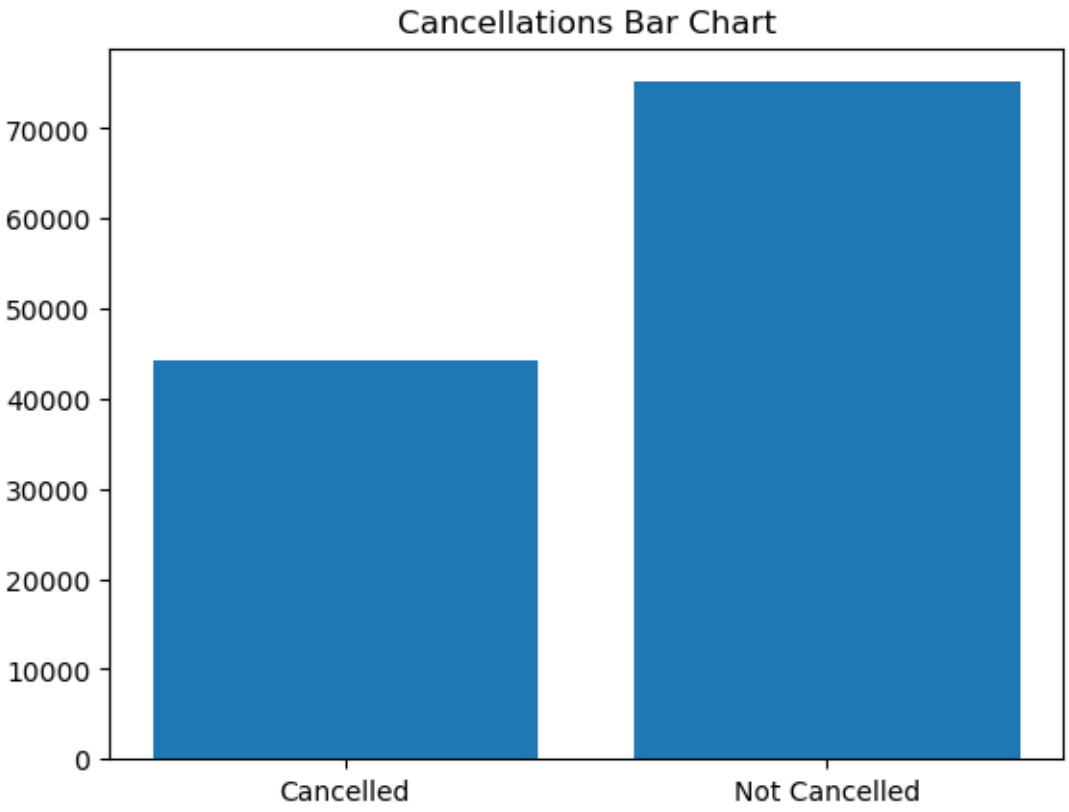
[16]:
```
number_of_noncancellations = len(df[df['is_canceled'] == 0])
print(number_of_noncancellations)
```

```
75166
```

[21]:
```
cancel_array = [number_of_cancellations, number_of_noncancellations]
```

[26]:
```
plot.bar(["Cancelled", "Not Cancelled"], cancel_array)
plot.title("Cancellations Bar Chart")
```

[26]:
```
Text(0.5, 1.0, 'Cancellations Bar Chart')
```

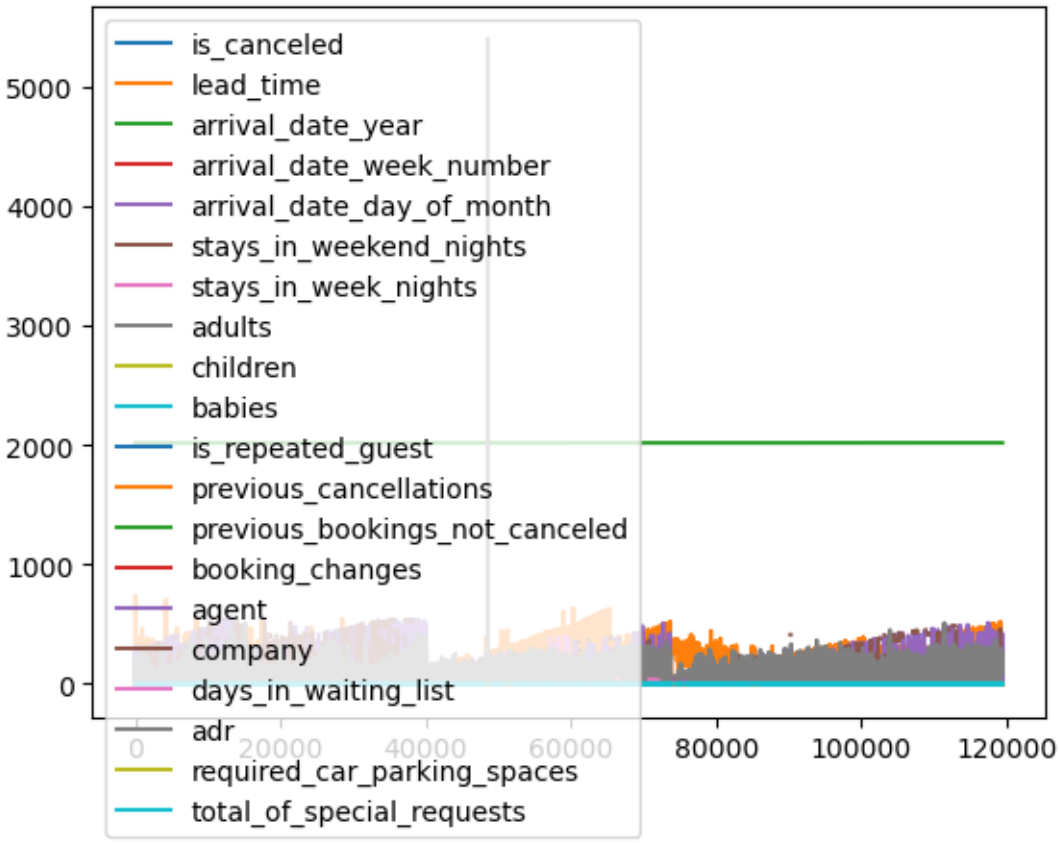


[32]:
```
most_popular_month_of_arrival = df["arrival_date_month"].mode()
print(most_popular_month_of_arrival)
```

```
0    August
Name: arrival_date_month, dtype: object
```

[51]:
```
#plot.hist(df["arrival_date_month"], rwidth=0.1)
df.plot()
```

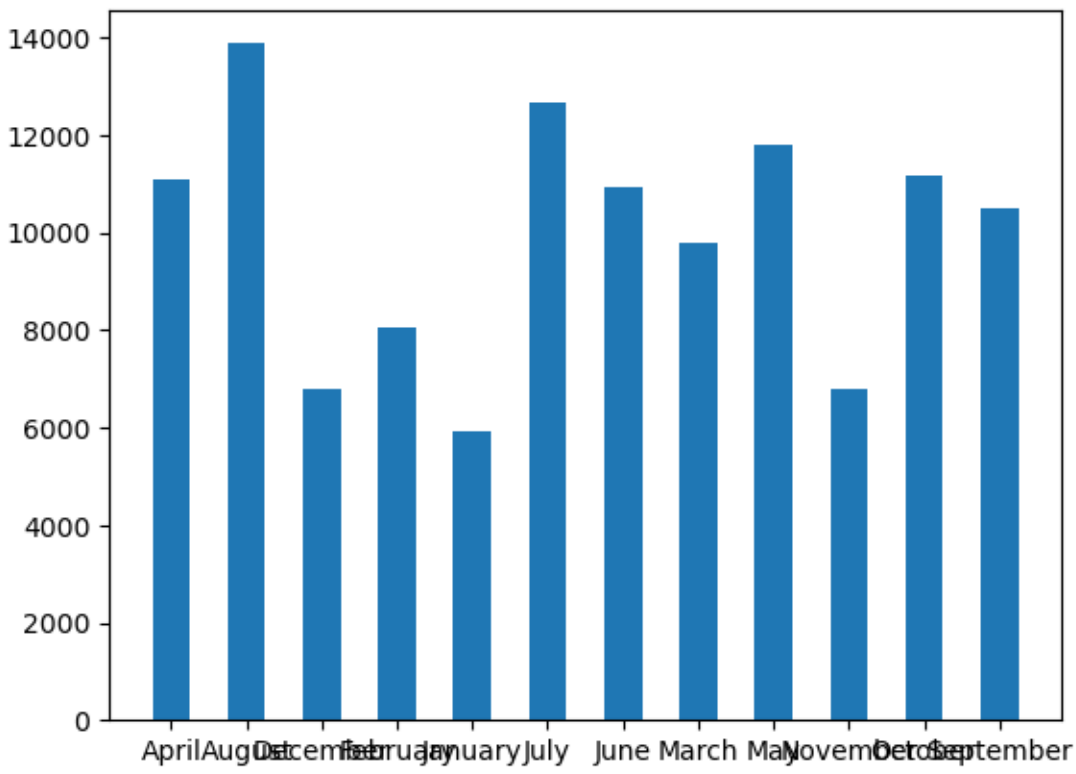
[51]: <Axes: >



[57]:
```
month_array = np.array(document.arrival_date_month)
month, count = np.unique(month_array, return_counts=True)
```

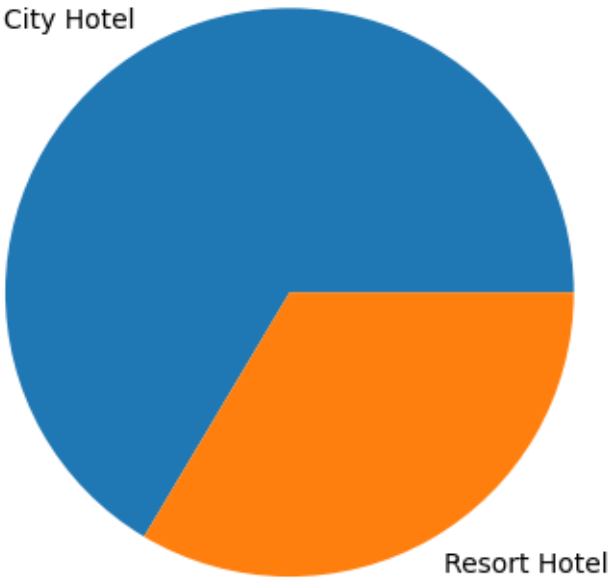
[65]:
```
plot.bar(month, count, width=0.5)
```

[65]:
```
<BarContainer object of 12 artists>
```



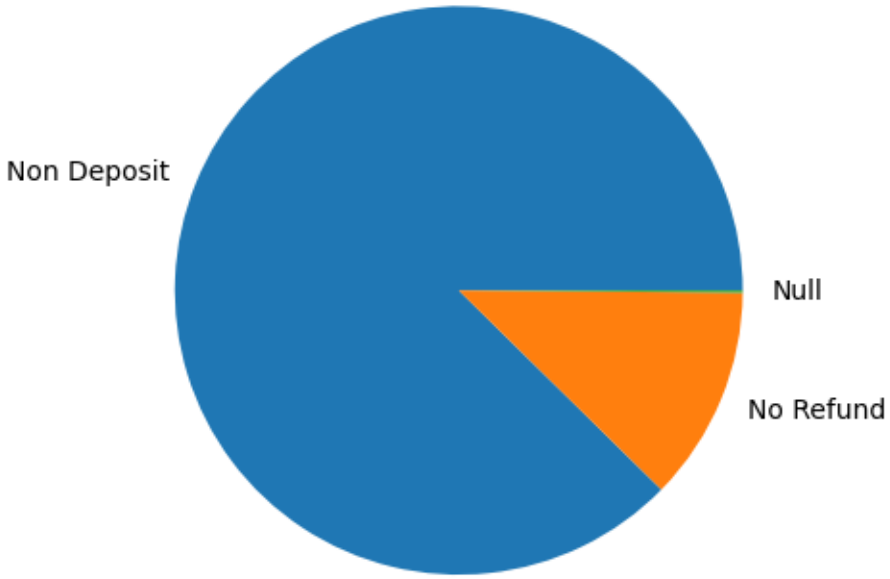
[69]:
```
plot.pie(df["hotel"].value_counts(), labels=["City Hotel", "Resort Hotel"])
```

[69]:
```
([<matplotlib.patches.Wedge at 0x2698d521d80>,
  <matplotlib.patches.Wedge at 0x2698d521c90>],
 [Text(-0.5433858480011854, 0.956416133381298, 'City Hotel'),
  Text(0.5433858480011848, -0.9564161333812983, 'Resort Hotel')])
```



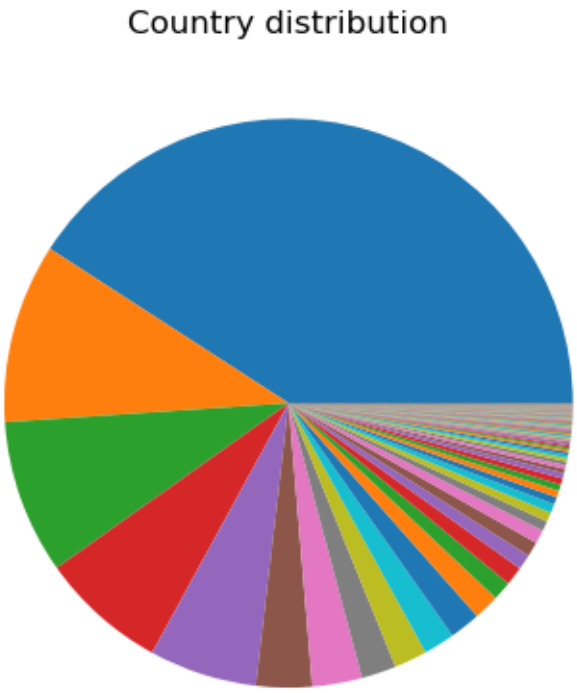
[73]:
```
plot.pie(df["deposit_type"].value_counts(), labels=["Non Deposit", "No Refund",␣
↪"Null"])
```

[73]:
```
([<matplotlib.patches.Wedge at 0x2698d690ee0>,
  <matplotlib.patches.Wedge at 0x2698d690df0>,
  <matplotlib.patches.Wedge at 0x2698d691570>],
 [Text(-1.0181924325401428, 0.4162741528343872, 'Non Deposit'),
  Text(1.0164087119406244, -0.4206106635490841, 'No Refund'),
  Text(1.0999900062128796, -0.004688947833933911, 'Null')])
```



[87]:
```
plot.pie(df["country"].value_counts())
plot.title("Country distribution")
```

[87]:
```
Text(0.5, 1.0, 'Country distribution')
```


Country distribution

[ ]: