Data Visualization With Matplotlib

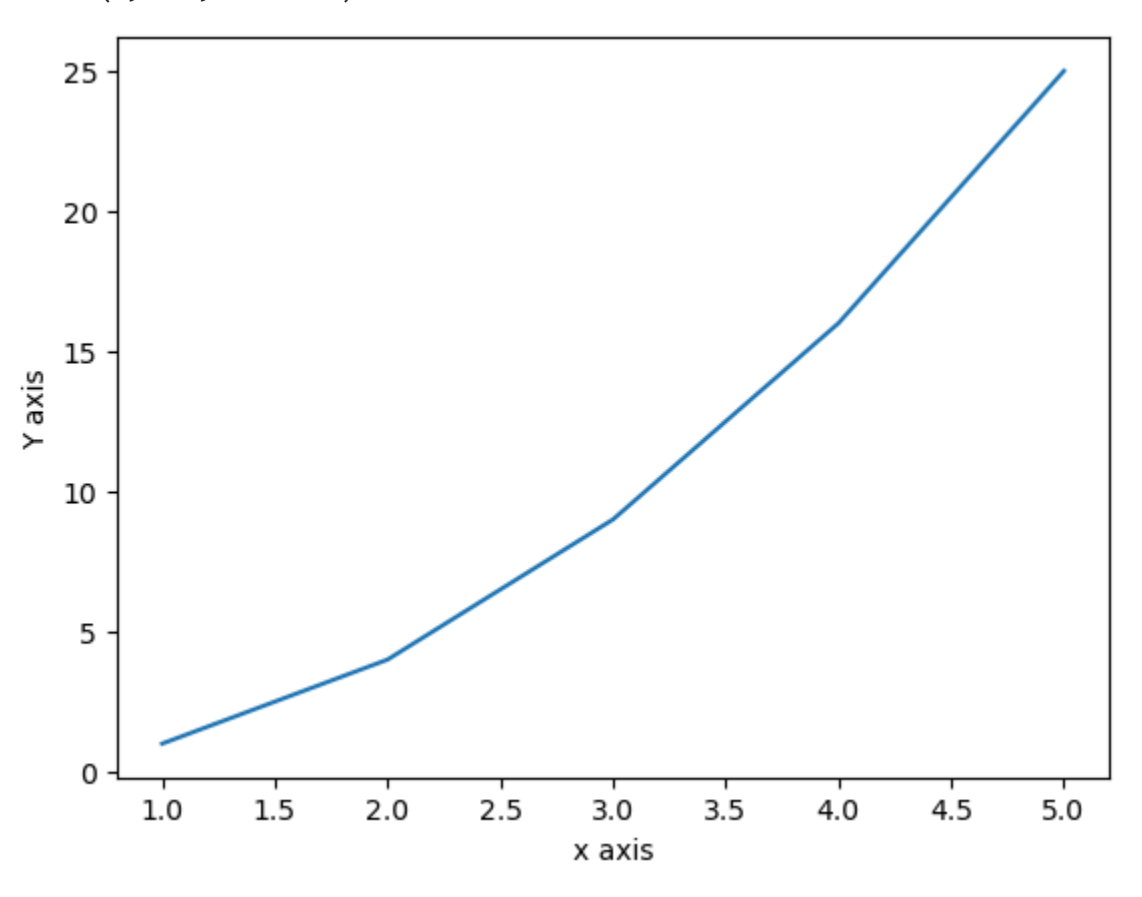
Matplotlib is a powerful plotting library for Python that enables the creation of static, animated, and interactive visualizations. It is widely used for data visualization in data science and analytics. In this lesson, we will cover the basics of Matplotlib, including creating various types of plots and customizing them.

In [2]:
```
!pip install matplotlib
```
```
Requirement already satisfied: matplotlib in e:\software\ide\anaconda\lib\site-packages (3.8.4)
Requirement already satisfied: contourpy>=1.0.1 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (1.2.0)
Requirement already satisfied: cycler>=0.10 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (4.51.0)
Requirement already satisfied: kiwisolver>=1.3.1 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: numpy>=1.21 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (1.26.4)
Requirement already satisfied: packaging>=20.0 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (23.2)
Requirement already satisfied: pillow>=8 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (10.3.0)
Requirement already satisfied: pyparsing>=2.3.1 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in e:\software\ide\anaconda\lib\site-packages (from matplotlib) (2.9.0.post0)
Requirement already satisfied: six>=1.5 in e:\software\ide\anaconda\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
```

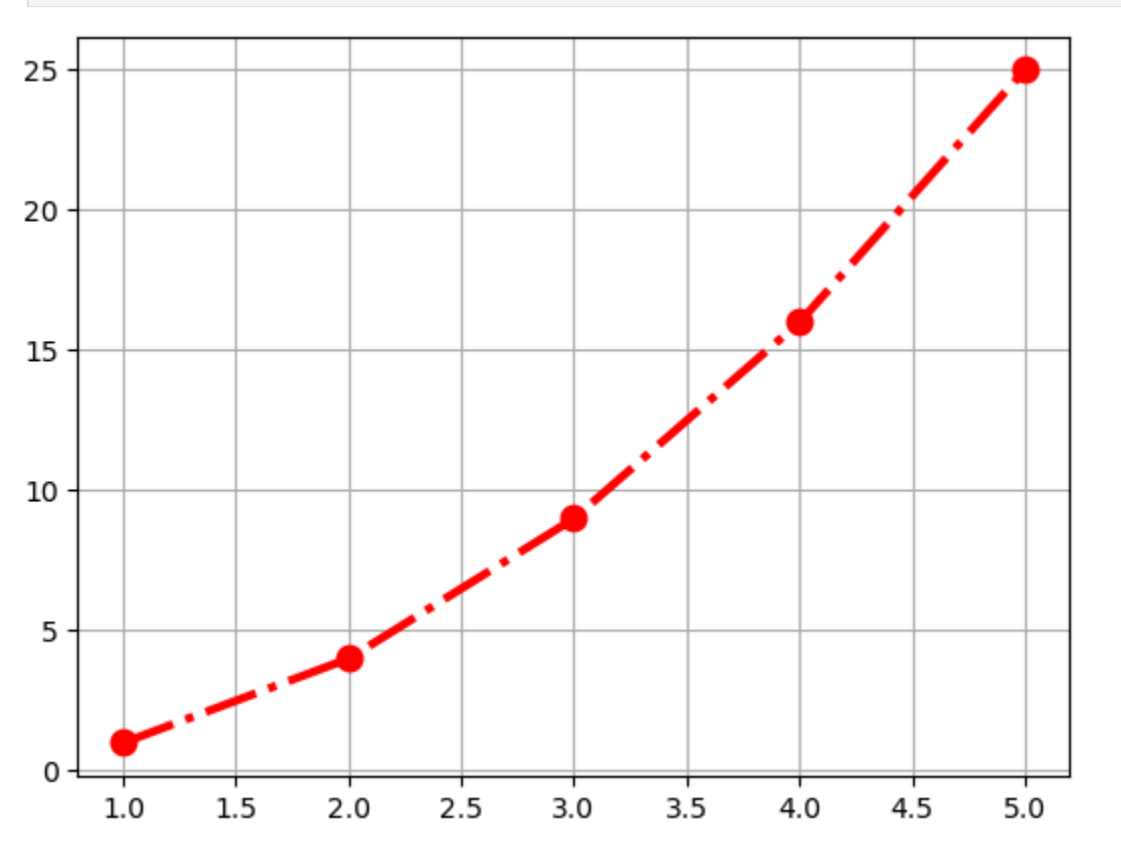
In [4]:
```
import matplotlib.pyplot as plt
```

In [8]:
```
x=[1,2,3,4,5]
y=[1,4,9,16,25]
#create a linee plot

plt.plot(x,y)
plt.xlabel("x axis")
plt.ylabel("Y axis")
```

Out[8]: `Text(0, 0.5, 'Y axis')`

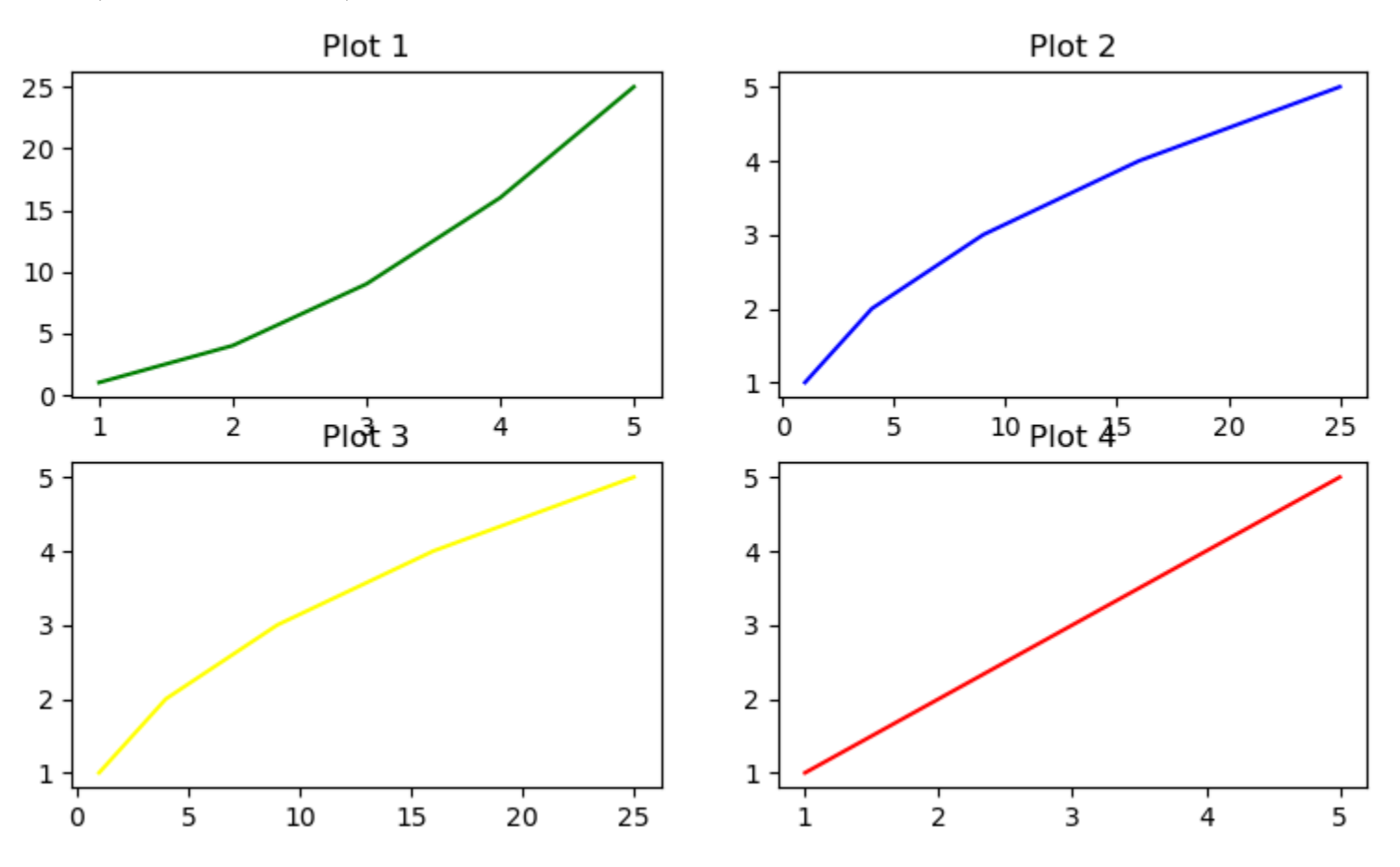


In [22]:
```
plt.plot(x,y,color='red',linestyle='-.',marker='o',linewidth=3,markersize=9)
plt.grid(True)
## customized plot
```



In [49]:
```
#multiple plots
#sample data
x = [1, 2, 3, 4, 5]
y1 = [1, 4, 9, 16, 25]
y2 = [1, 2, 3, 4, 5]
plt.figure(figsize=(9,5))
plt.subplot(2,2,1)  #(row,column,block)
plt.plot(x,y1, color='green')
plt.title("Plot 1")

plt.subplot(2,2,2)  #(row,column,block)
plt.plot(y1,x, color='blue')
plt.title("Plot 2")

plt.subplot(2,2,3)  #(row,column,block)
plt.plot(y1,y2, color='yellow')
plt.title("Plot 3")

plt.subplot(2,2,4)  #(row,column,block)
plt.plot(y2,x, color='red')
plt.title("Plot 4")
```

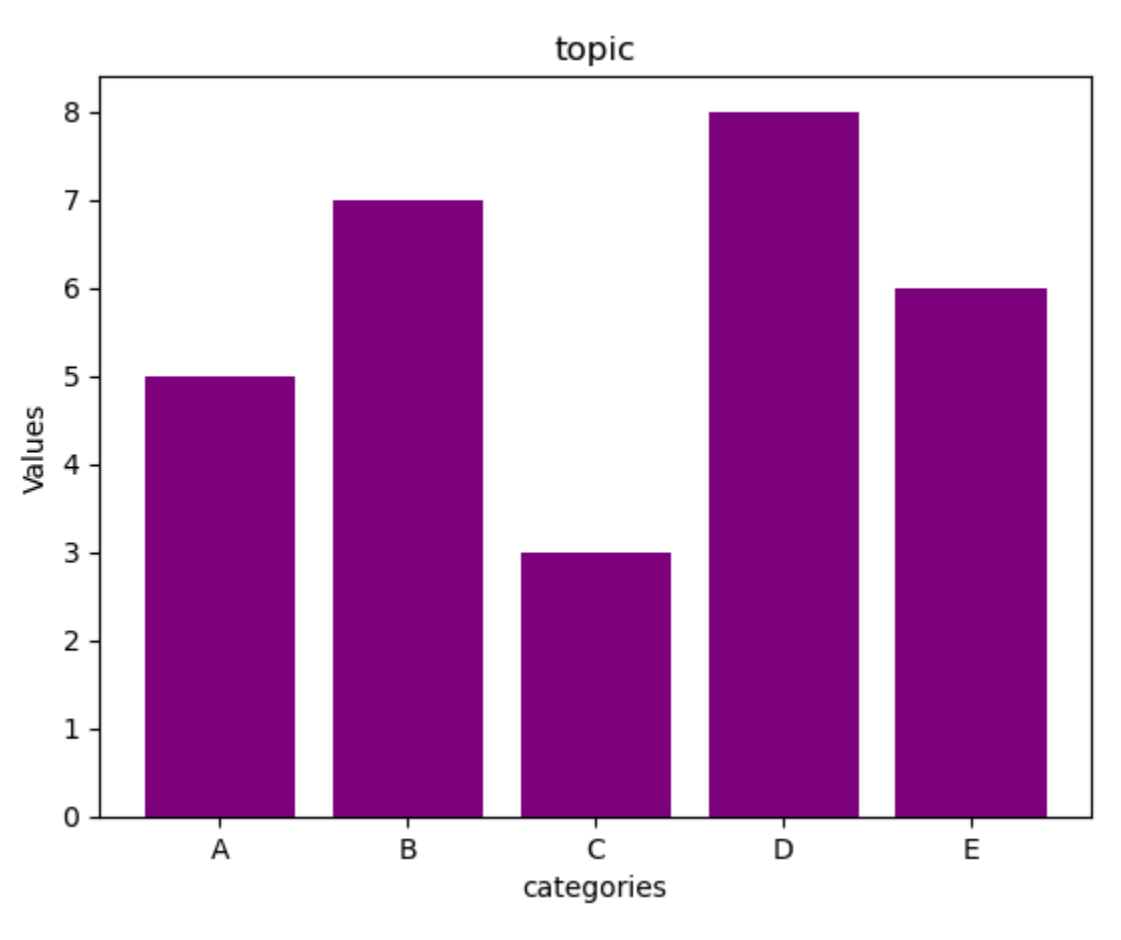
Out[49]: `Text(0.5, 1.0, 'Plot 4')`



BARPLOT

In [54]:
```
###Bar Plor
categories=['A', 'B', 'C', 'D', 'E']
values=[5,7,3,8,6]
##create a bar plot
plt.bar(categories, values, color='purple')
plt.xlabel("categories")
plt.ylabel("Values")
plt.title("topic")
```

Out[54]: `Text(0.5, 1.0, 'topic')`

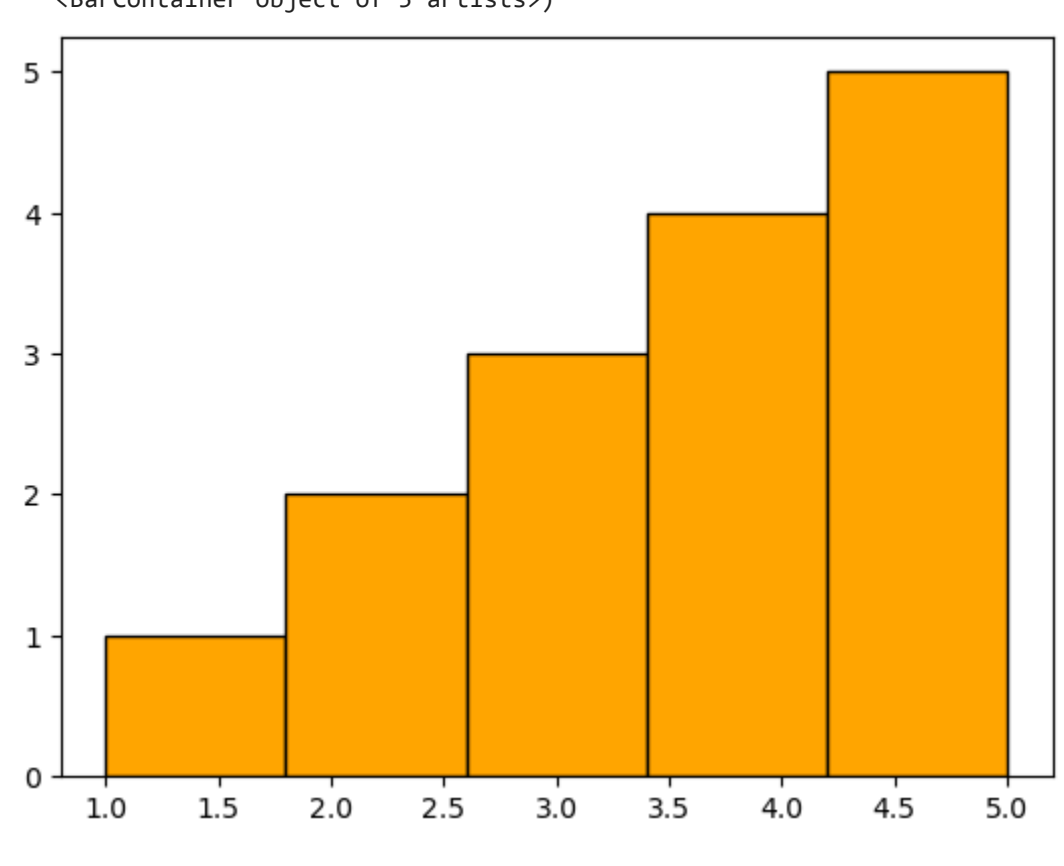


Histograms

Histograms are used to represent the distribution of a dataset. They divide the data into bins and count the number of data points in each bin.

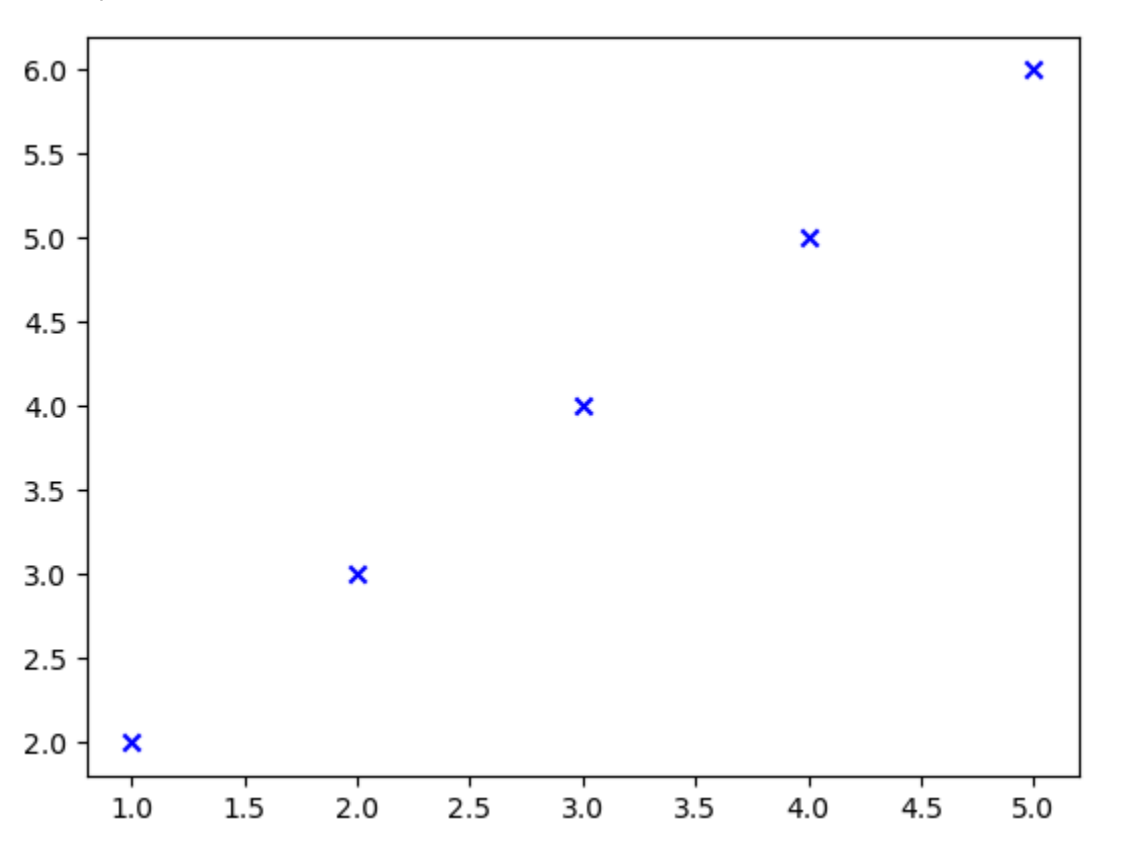
In [57]:
```
# Sample data
data = [1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5]
##create a histogram
plt.hist(data, bins=5, color='orange', edgecolor='black')
```

Out[57]:
```
(array([1., 2., 3., 4., 5.]),
 array([1. , 1.8, 2.6, 3.4, 4.2, 5. ]),
 <BarContainer object of 5 artists>)
```



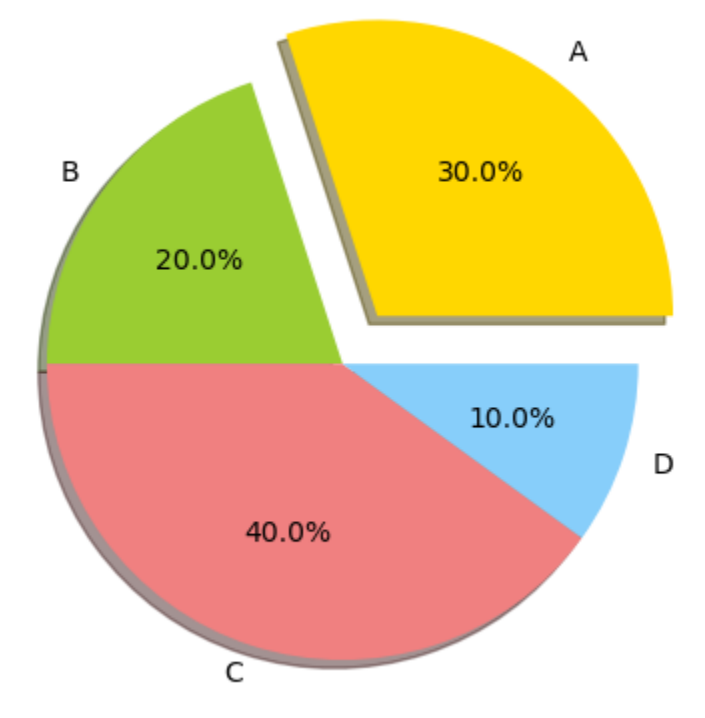
In [59]:
```
##create a scatter plot
# Sample data
x = [1, 2, 3, 4, 5]
y = [2, 3, 4, 5, 6]
plt.scatter(x, y, color="blue", marker='x')
```

Out[59]: `<matplotlib.collections.PathCollection at 0x1ad33f172c0>`



In [71]:
```
### pie chart
labels=['A', 'B', 'C', 'D']
sizes=[30, 20,40,10]
colors=['gold', 'yellowgreen', 'lightcoral', 'lightskyblue']
explode=(0.2,0,0,0)
##create a pie chart
plt.pie(sizes, explode=explode, labels=labels, colors=colors, autopct="%1.1f%%",shadow=True)
```

Out[71]:
```
([<matplotlib.patches.Wedge at 0x1ad385fa450>,
  <matplotlib.patches.Wedge at 0x1ad385fa390>,
  <matplotlib.patches.Wedge at 0x1ad385fb080>,
  <matplotlib.patches.Wedge at 0x1ad385fb830>],
 [Text(0.764120788592483, 1.051722121304293, 'A'),
  Text(-0.8899187482945419, 0.6465637025335369, 'B'),
  Text(-0.3399185762739153, -1.046162206115244, 'C'),
  Text(1.0461622140716127, -0.3399185517867209, 'D')],
 [Text(0.47022817759537416, 0.6472136131103341, '30.0%'),
  Text(-0.4854102263424773, 0.3526711104728383, '20.0%'),
  Text(-0.1854101325130447, -0.5706339306083149, '40.0%'),
  Text(0.5706339349481523, -0.18541011915639322, '10.0%')])
```



In [73]:
```
## Sales Data Visualization
import pandas as pd
sales_data_df=pd.read_csv('data.csv')
print(sales_data_df.head(5))
```
```
   VIN (1-10)    County     City State  Postal Code  Model Year     Make  \
0  5YJ3E1EBXK      King  Seattle    WA      98178.0        2019    TESLA
1  5YJYGDEE3L    Kitsap  Poulsbo    WA      98370.0        2020    TESLA
2  KM8KRDAF5P    Kitsap   Olalla    WA      98359.0        2023  HYUNDAI
3  5UXTA6C0XM    Kitsap  Seabeck    WA      98380.0        2021      BMW
4  JTMAB3FV7P  Thurston  Rainier    WA      98576.0        2023   TOYOTA

       Model                   Electric Vehicle Type  \
0    MODEL 3          Battery Electric Vehicle (BEV)
1    MODEL Y          Battery Electric Vehicle (BEV)
2    IONIQ 5          Battery Electric Vehicle (BEV)
3         X5  Plug-in Hybrid Electric Vehicle (PHEV)
4  RAV4 PRIME  Plug-in Hybrid Electric Vehicle (PHEV)

   Clean Alternative Fuel Vehicle (CAFV) Eligibility  Electric Range  \
0            Clean Alternative Fuel Vehicle Eligible           220.0
1            Clean Alternative Fuel Vehicle Eligible           291.0
2  Eligibility unknown as battery range has not b...             0.0
3            Clean Alternative Fuel Vehicle Eligible            30.0
4            Clean Alternative Fuel Vehicle Eligible            42.0

   Base MSRP  Legislative District  DOL Vehicle ID  \
0        0.0                  37.0       477309682
1        0.0                  23.0       109705683
2        0.0                  26.0       230390492
3        0.0                  35.0       267929112
4        0.0                   2.0       236505139

             Vehicle Location                             Electric Utility  \
0  POINT (-122.23825 47.49461)  CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)
1  POINT (-122.64681 47.73689)                        PUGET SOUND ENERGY INC
2  POINT (-122.54729 47.42602)                        PUGET SOUND ENERGY INC
3  POINT (-122.81585 47.64509)                        PUGET SOUND ENERGY INC
4  POINT (-122.68993 46.88897)                        PUGET SOUND ENERGY INC

   2020 Census Tract
0       5.303301e+10
1       5.303509e+10
2       5.303509e+10
3       5.303509e+10
4       5.306701e+10
```

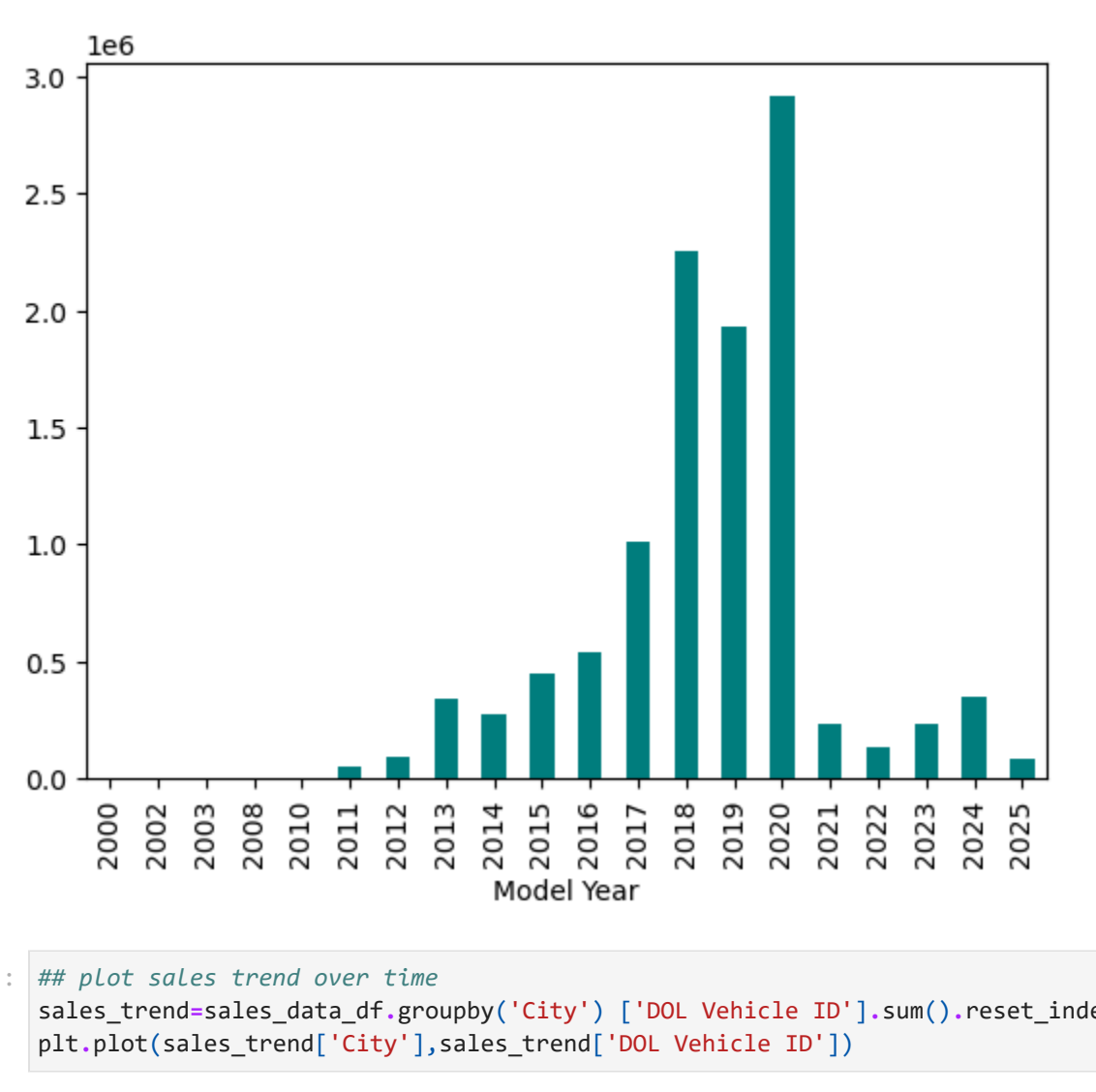
In [77]:
```
sales_data_df.info
```

Out[77]:
```
<bound method DataFrame.info of         VIN (1-10)     County          City State  Postal Code  Model Year  \
0       5YJ3E1EBXK       King       Seattle    WA      98178.0        2019
1       5YJYGDEE3L     Kitsap       Poulsbo    WA      98370.0        2020
2       KM8KRDAF5P     Kitsap        Olalla    WA      98359.0        2023
3       5UXTA6C0XM     Kitsap       Seabeck    WA      98380.0        2021
4       JTMAB3FV7P   Thurston       Rainier    WA      98576.0        2023
...            ...        ...           ...   ...          ...         ...
235687  1C4RJXN62R     Pierce        Tacoma    WA      98407.0        2024
235688  5YJSA1E28J  Snohomish      Stanwood    WA      98292.0        2018
235689  3FA6P0SU2F       King       Redmond    WA      98052.0        2015
235690  WA1BCBFZ6P  Snohomish  Lake Stevens    WA      98258.0        2023
235691  WBY33AW03P       King      Issaquah    WA      98027.0        2023

           Make       Model                   Electric Vehicle Type  \
0         TESLA     MODEL 3          Battery Electric Vehicle (BEV)
1         TESLA     MODEL Y          Battery Electric Vehicle (BEV)
2       HYUNDAI     IONIQ 5          Battery Electric Vehicle (BEV)
3           BMW          X5  Plug-in Hybrid Electric Vehicle (PHEV)
4        TOYOTA  RAV4 PRIME  Plug-in Hybrid Electric Vehicle (PHEV)
...         ...         ...                                     ...
235687     JEEP    WRANGLER  Plug-in Hybrid Electric Vehicle (PHEV)
235688    TESLA     MODEL S          Battery Electric Vehicle (BEV)
235689     FORD      FUSION  Plug-in Hybrid Electric Vehicle (PHEV)
235690     AUDI          Q4          Battery Electric Vehicle (BEV)
235691      BMW          I4          Battery Electric Vehicle (BEV)

        Clean Alternative Fuel Vehicle (CAFV) Eligibility  Electric Range  \
0                 Clean Alternative Fuel Vehicle Eligible           220.0
1                 Clean Alternative Fuel Vehicle Eligible           291.0
2       Eligibility unknown as battery range has not b...             0.0
3                 Clean Alternative Fuel Vehicle Eligible            30.0
4                 Clean Alternative Fuel Vehicle Eligible            42.0
...                                                   ...             ...
235687              Not eligible due to low battery range            21.0
235688            Clean Alternative Fuel Vehicle Eligible           249.0
235689              Not eligible due to low battery range            19.0
235690  Eligibility unknown as battery range has not b...             0.0
235691  Eligibility unknown as battery range has not b...             0.0

        Base MSRP  Legislative District  DOL Vehicle ID  \
0             0.0                  37.0       477309682
1             0.0                  23.0       109705683
2             0.0                  26.0       230390492
3             0.0                  35.0       267929112
4             0.0                   2.0       236505139
...           ...                   ...             ...
235687        0.0                  27.0       267004272
235688        0.0                  10.0       198794410
235689        0.0                  48.0       261742288
235690        0.0                  44.0       253474854
235691        0.0                   5.0       229583490

                   Vehicle Location  \
0       POINT (-122.23825 47.49461)
1       POINT (-122.64681 47.73689)
2       POINT (-122.54729 47.42602)
3       POINT (-122.81585 47.64509)
4       POINT (-122.68993 46.88897)
...                             ...
235687  POINT (-122.51134 47.29238)
235688  POINT (-122.37265 48.24159)
235689  POINT (-122.13158 47.67858)
235690  POINT (-122.06402 48.01497)
235691   POINT (-122.03439 47.5301)

                                    Electric Utility  2020 Census Tract
0           CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)       5.303301e+10
1                             PUGET SOUND ENERGY INC       5.303509e+10
2                             PUGET SOUND ENERGY INC       5.303509e+10
3                             PUGET SOUND ENERGY INC       5.303509e+10
4                             PUGET SOUND ENERGY INC       5.306701e+10
...                                              ...                ...
235687  BONNEVILLE POWER ADMINISTRATION||CITY OF TACOM...       5.305306e+10
235688  BONNEVILLE POWER ADMINISTRATION||PUD 1 OF SNOH...       5.306105e+10
235689      PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)       5.303303e+10
235690                            PUGET SOUND ENERGY INC       5.306105e+10
235691      PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)       5.303303e+10

[235692 rows x 17 columns]>
```

In [81]:
```
## plot total sales by products
total_sales_by_product=sales_data_df.groupby('Model Year') ['Electric Range'].sum()
print(total_sales_by_product)
```
```
Model Year
2000        406.0
2002        190.0
2003         95.0
2008       4620.0
2010       5345.0
2011      48289.0
2012      91530.0
2013     337358.0
2014     272114.0
2015     451192.0
2016     538470.0
2017    1008843.0
2018    2257714.0
2019    1934820.0
2020    2914562.0
2021     232276.0
2022     135158.0
2023     235345.0
2024     348349.0
2025      85376.0
Name: Electric Range, dtype: float64
```

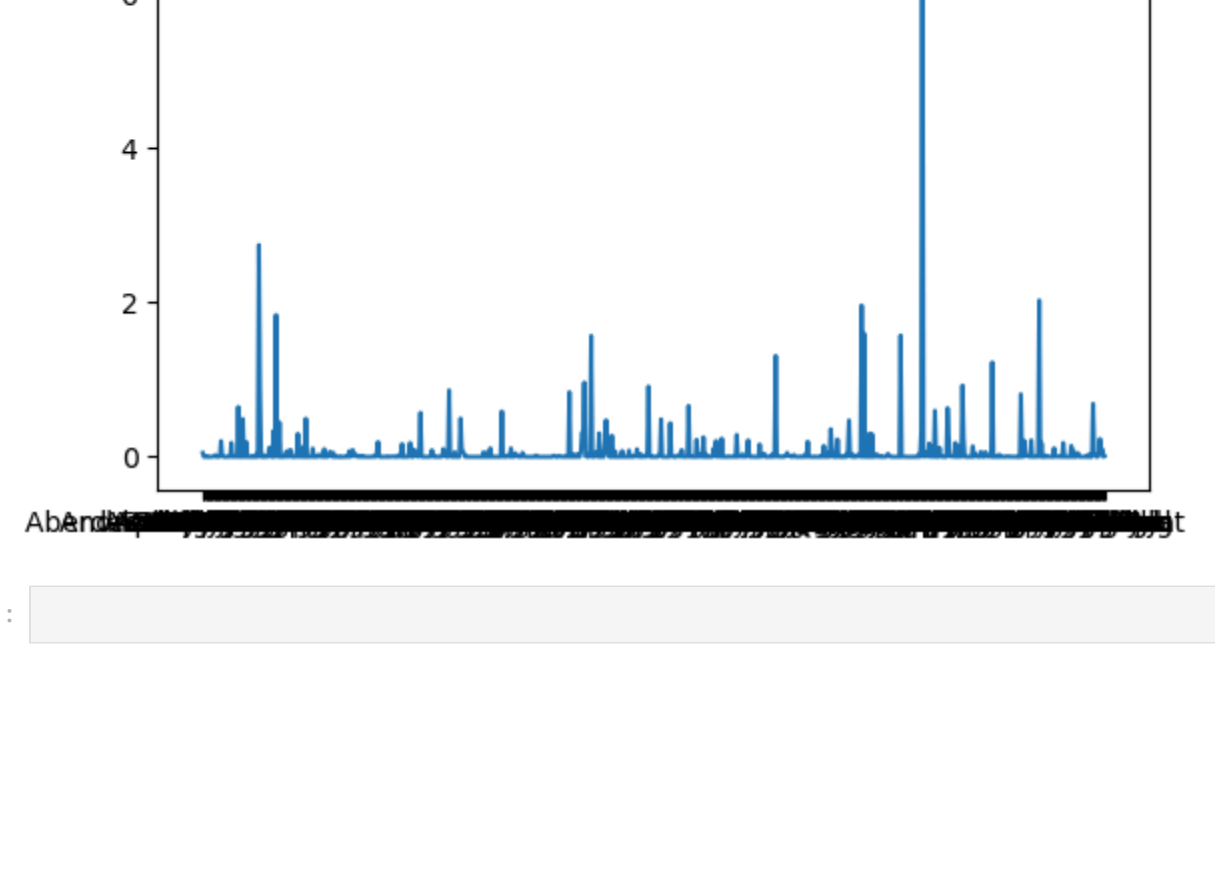
In [83]:
```
total_sales_by_product.plot(kind='bar',color='teal')
```

Out[83]: `<Axes: xlabel='Model Year'>`



In [85]:
```
## plot sales trend over time
sales_trend=sales_data_df.groupby('City') ['DOL Vehicle ID'].sum().reset_index()
plt.plot(sales_trend['City'],sales_trend['DOL Vehicle ID'])
```

Out[85]: `[<matplotlib.lines.Line2D at 0x1ad385fa420>]`



In [ ]: