Importing necessary modules

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
import seaborn as sns
```

Task1:To find which cars give the best mileage

In [2]:
```
car=pd.read_excel(r"K:\Desktop\NIIT\tables\DS1_C4_S5_Car_Data_Challenge.xlsx")
for item in car.columns:
    print(item)
```

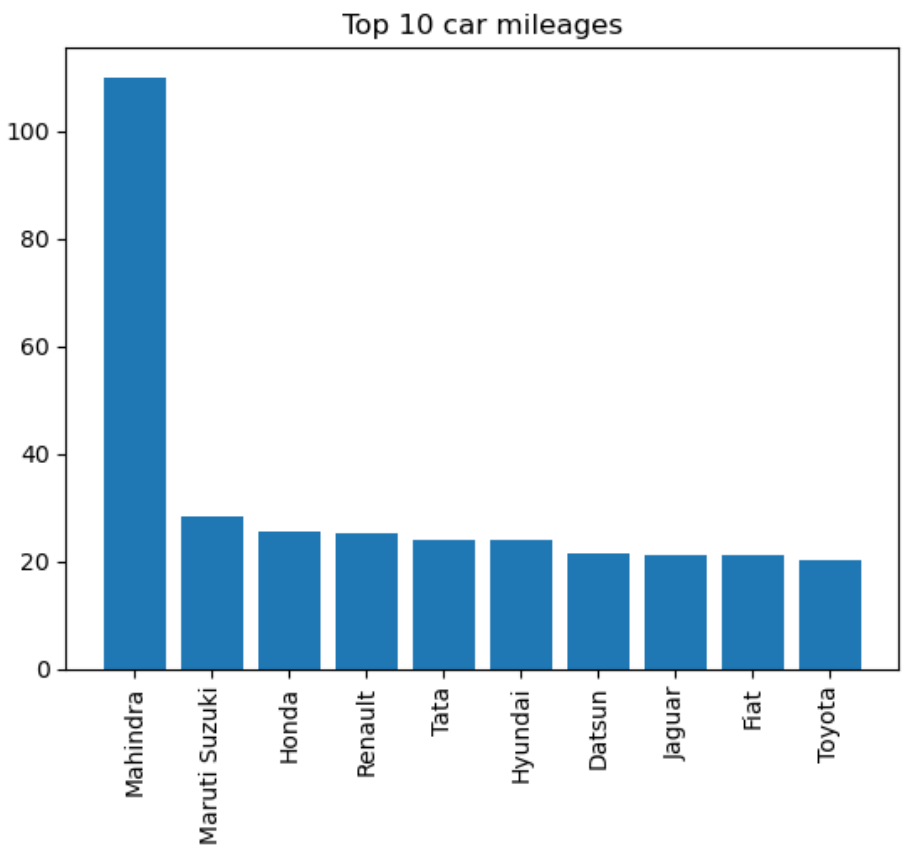
```
Wheelbase
Wheels_Size
Start_/_Stop_Button
12v_Power_Outlet
Audiosystem
Aux-in_Compatibility
Average_Fuel_Consumption
Basic_Warranty
Bluetooth
Boot-lid_Opener
Boot_Space_litre
CD_/_MP3_/_DVD_Player
Central_Locking
Child_Safety_Locks
Clock
Cup_Holders
Distance_to_Empty
Door_Pockets
Engine_Malfunction_Light
Extended_Warranty
```

Data preprocessing

Task1: To find the car manufacturer that gives the best mileage

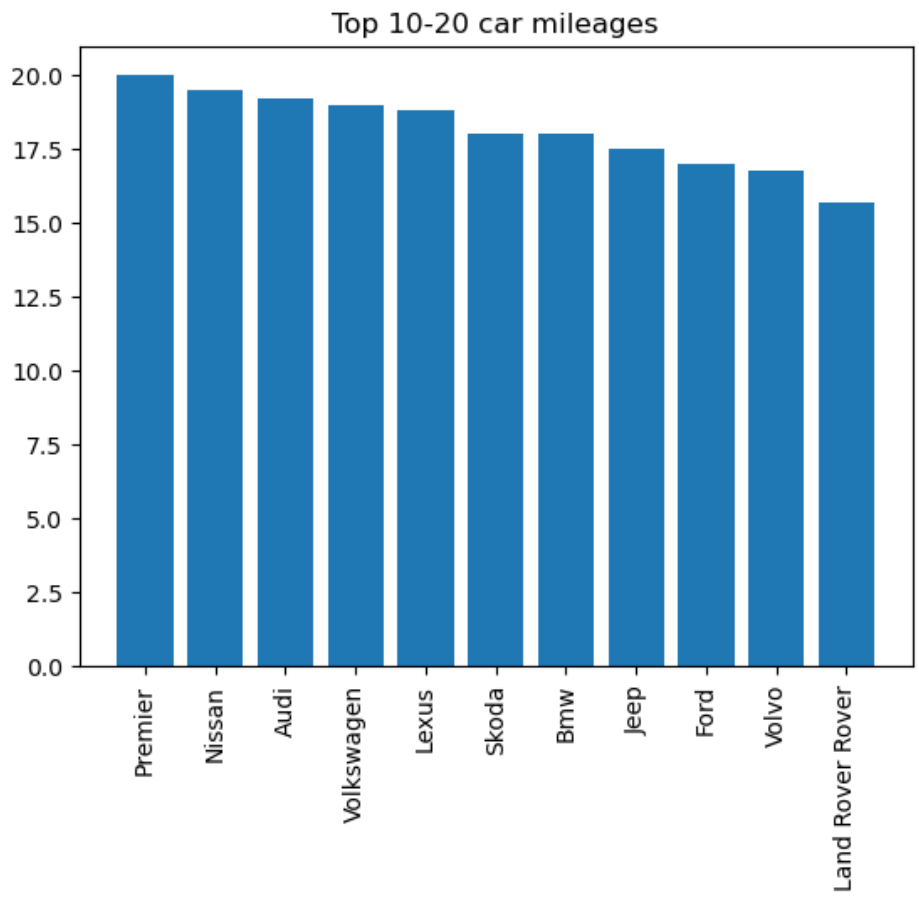
In [3]:
```
median=car["City_Mileage_km_litre"].median()
median
car['City_Mileage_km_litre'].fillna(median,inplace=True)
mileage=car.groupby("Make").max().City_Mileage_km_litre.sort_values(ascending=False)[0:10]

plt.bar(mileage.index,mileage)
plt.xticks(rotation=90)
plt.title("Top 10 car mileages")
plt.show()
```



Task2: To find next top 10 car makers

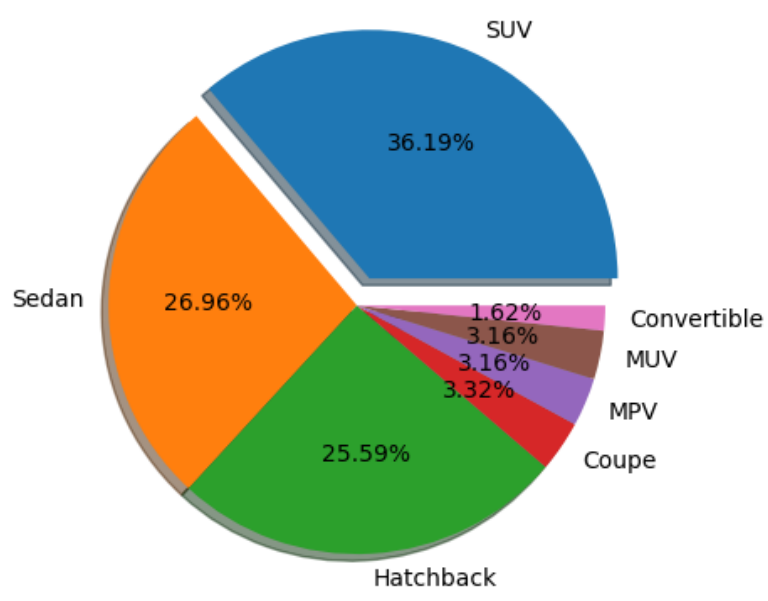
In [4]:
```
mileage2=car.groupby("Make").max().City_Mileage_km_litre.sort_values(ascending=False)[10:21]
plt.bar(mileage2.index,mileage2)
plt.xticks(rotation=90)
plt.title("Top 10-20 car mileages")
plt.show()
```



Task3: To find which body type contributes most to market

In [5]:
```
mode=car.Make.mode()
print(mode)
car.fillna("Maruti Suzuki")
body=car.value_counts("Body_Type")[0:7]
plt.pie(body,labels=body.index,autopct="%.2f%%",explode=[0.12,0,0,0,0,0,0],shadow=True);
plt.show()
print("Suv body type contributes the most ")
```

```
0    Maruti Suzuki
Name: Make, dtype: object
```



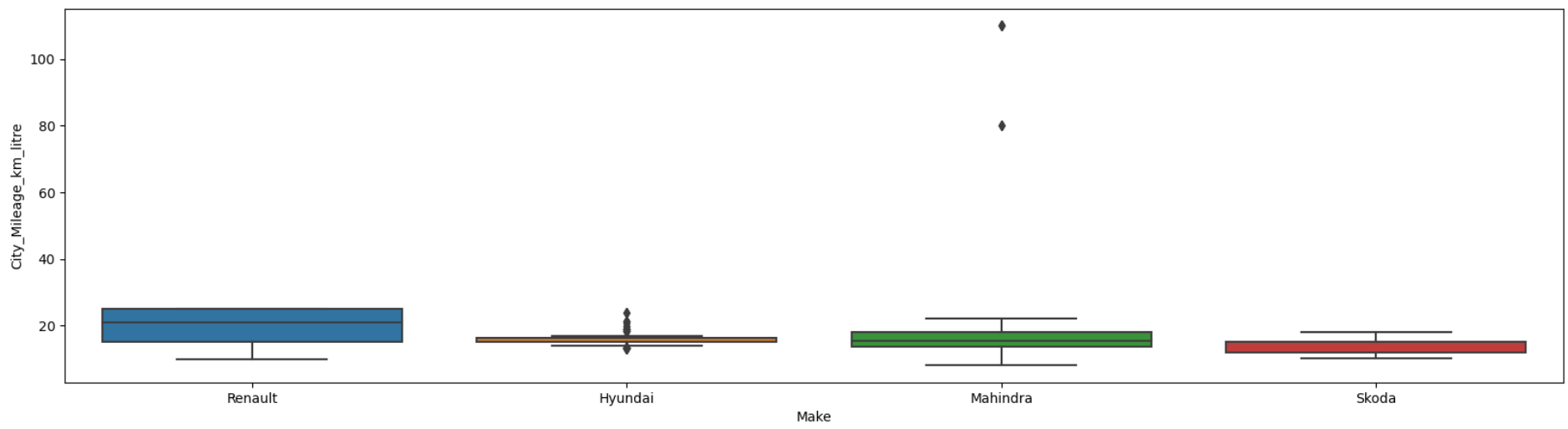

```
Suv body type contributes the most
```

Task4: To compare the central measures of tendency in hyundai mahindra renault and skoda

In [6]:
```
german=car[(car.Make=="Skoda")|(car.Make=="Mahindra")|(car.Make=="Hyundai")|(car.Make=="Renault")].loc[:,["Make","City_Mileage_km_litre"]]

pd.pivot_table(german,index="Make",values="City_Mileage_km_litre",aggfunc=["mean","median","min","max"])
plt.figure(figsize=(20,5))
sns.boxplot(x=german.Make,y=german.City_Mileage_km_litre,data=german)
```

Out[6]: <AxesSubplot: xlabel='Make', ylabel='City_Mileage_km_litre'>



Task5: To plot mileages of Toyota Mahindra and Volkswagen Hyundai

In [12]:
```
japan=car[(car.Make=="Volkswagen")|(car.Make=="Mahindra")|(car.Make=="Hyundai")|(car.Make=="Toyota")].loc[:,["Make","City_Mileage_km_litre","H
japan_df=pd.pivot_table(japan,index="Make",values=["Highway_Mileage_km_litre","City_Mileage_km_litre"],aggfunc=["mean","median"])
japan_df["Make"]=japan_df.index
japan_df

japan_df.plot(x="Make",kind="bar",stacked=True)
plt.legend(loc=(1.1, 0.2))
plt.show()
```