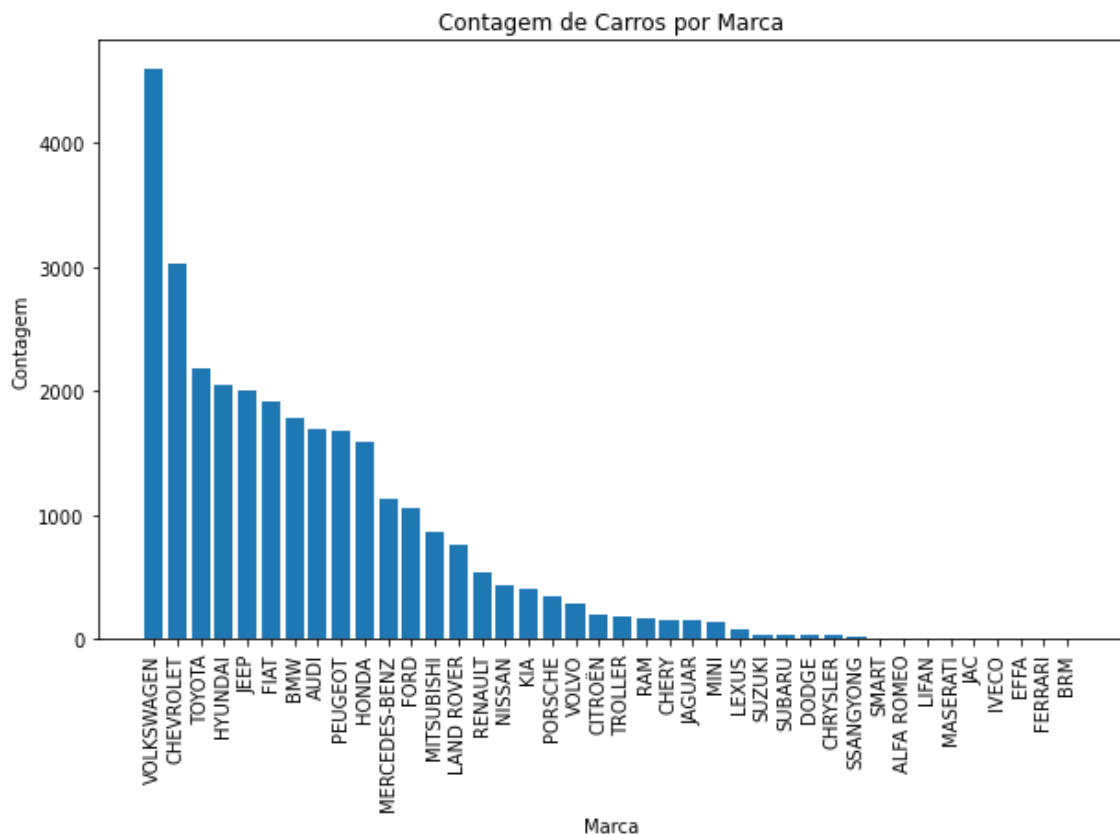


## EDA

Qual melhor estado cadastrado na base de dados para se vender um carro de marca popular e por que?

```
# Contagem de ocorrências de cada marca de carro  
train.marca.value_counts()
```

```
marca_counts = train.marca.value_counts()  
  
plt.figure(figsize=(10, 6))  
plt.bar(marca_counts.index, marca_counts.values)  
plt.xlabel('Marca')  
plt.ylabel('Contagem')  
plt.title('Contagem de Carros por Marca')  
plt.xticks(rotation=90)  
plt.show()
```



```
marcas =  
['VOLKSWAGEN', 'CHEVROLET', 'TOYOTA', 'HYUNDAI', 'FIAT', 'PEUGEOT', 'HONDA', 'FORD',  
, 'RENAULT', 'NISSAN', 'CITROEN']
```

```
# Atribui 1 às instâncias da marca popular
```

```
for i in range(0, 11):  
    marca = marcas[i]  
    train_mask = train.marca == marca  
    train['marca'][train_mask] = 1
```

```
# Atribui 0 às instâncias das marcas não populares
```

```
marca_nao_popular = train.marca.value_counts().index[1:31]  
for marca in marca_nao_popular:  
    train_mask = train.marca == marca  
    train.loc[train_mask, 'marca'] = 0
```

```
# Renomeia a coluna "marca" para "marca_popular"
```

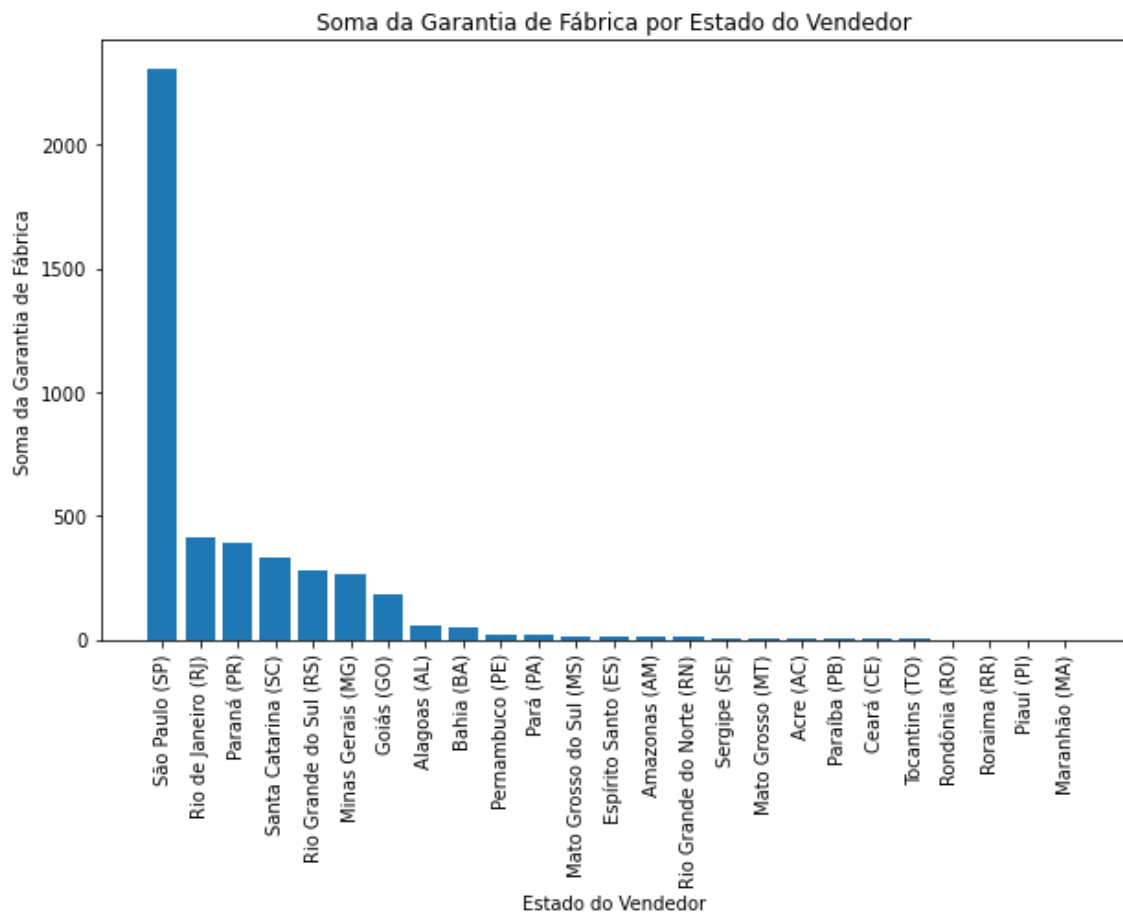
```
train.rename({"marca": "marca_popular"}, axis=1, inplace=True)
```

```
# Agrupa e soma a garantia de fábrica e o preço dos carros por estado do  
vendedor, classificando por garantia de fábrica em ordem decrescente
```

```
carros_garantia = train[['garantia_de_fábrica', 'preco',  
'estado_vendedor']].groupby('estado_vendedor').sum().sort_values('garantia_de_fáb  
rica', ascending=False)  
carros_garantia
```

```
# Criação do gráfico de barras
```

```
plt.figure(figsize=(10, 6))  
plt.bar(carros_garantia.index, carros_garantia['garantia_de_fábrica'])  
plt.xlabel('Estado do Vendedor')  
plt.ylabel('Soma da Garantia de Fábrica')  
plt.title('Soma da Garantia de Fábrica por Estado do Vendedor')  
plt.xticks(rotation=90)  
plt.show()
```

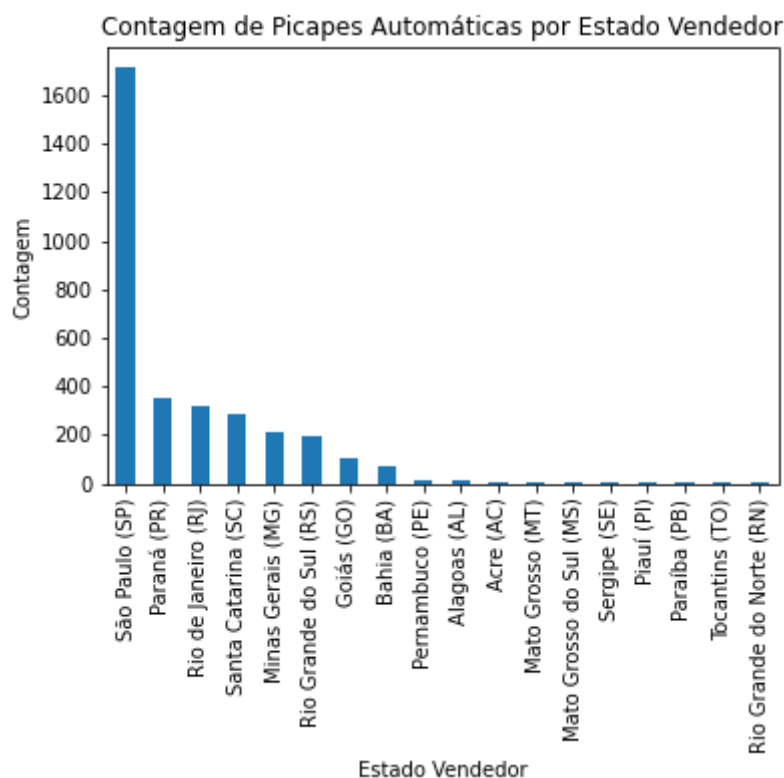


Como podemos analisar de acordo com os filtros aplicados . São Paulo lidera a venda de carros populares baseado no seu tempo de garantia e seu preço médio por venda . assim como RJ vem em segundo colocado em relação ao tempo de garantia de carros populares e seu preços médio . outros estados que não obtiveram Por outro lado, há estados com um número menor de garantias de fábrica, como Rondônia (RO), Roraima (RR), Piauí (PI) e Maranhão (MA), onde o preço é bastante baixo ou até mesmo zero.

## Qual Melhor estado para se comprar uma picape com transmissão automática e por que ?

```
train.cambio.value_counts()[0:20]
```

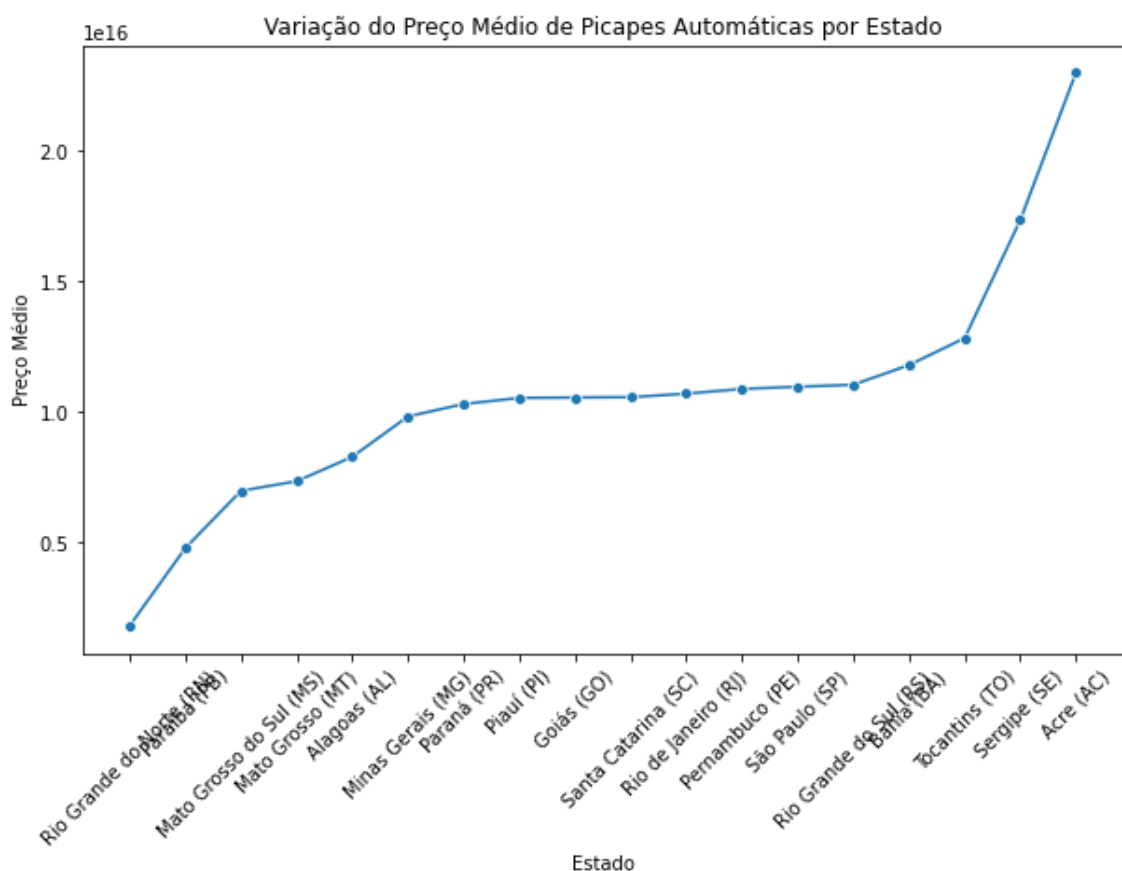
```
picape_aut = train[(train.tipo == 'Picape') & (train.cambio == 'Automática')]  
[['estado_vendedor', 'tipo']].groupby('estado_vendedor').count().sort_values('tipo', as  
cending=False)  
picape_aut
```



```
# Filtrando os dados para picapes automáticas e agrupando por estado_vendedor,  
calculando a média do preço e ordenando os valores  
picape_auto = train[(train.tipo == 'Picape') & (train.cambio ==  
'Automática')][['estado_vendedor',  
'preco']].groupby('estado_vendedor').mean().sort_values('preco').reset_index()
```

```
picape_auto.sort_values('preco')
```

```
# Gerando um gráfico de linha a partir dos dados filtrados
plt.figure(figsize=(10, 6))
sns.lineplot(x='estado_vendedor', y='preco', data=picape_auto, marker='o')
plt.title('Variação do Preço Médio de Picapes Automáticas por Estado')
plt.xlabel('Estado')
plt.ylabel('Preço Médio')
plt.xticks(rotation=45)
plt.show()
```



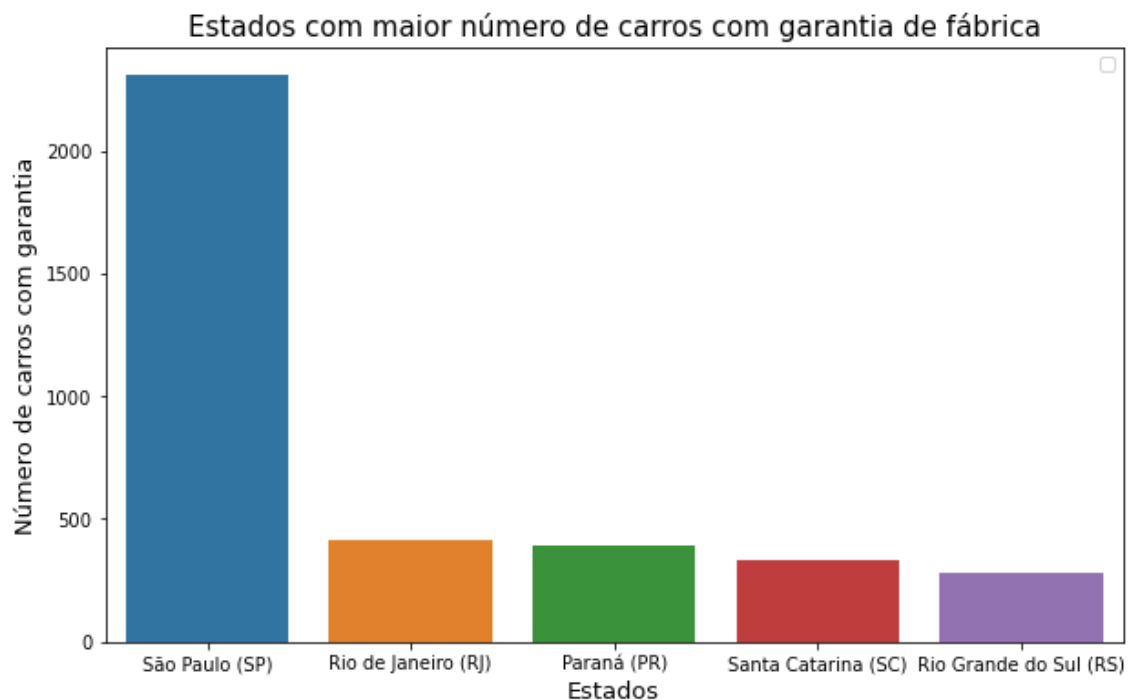
Com base nos dados fornecidos na tabela, o melhor estado para comprar uma picape automática considerando os valores de preço, pode-se observar que o estado com o menor valor de preço é o Rio Grande do Norte (RN), seguido por Paraíba (PB), Mato Grosso do Sul (MS) e Mato Grosso (MT).

**Qual melhor estado para se comprar carros que ainda estejam dentro da garantia de fábrica e por que ?**

```
garantia_de_fabrica = train
```

```
estado_garantia =  
garantia_de_fabrica[['estado_vendedor','garantia_de_fábrica','preco','revisoes_dentr  
o_agenda']].groupby('estado_vendedor').sum().sort_values('garantia_de_fábrica',asc  
ending=False)[0:5].reset_index()  
estado_garantia
```

```
plt.figure(figsize=(10, 6))  
sns.barplot(data=estado_garantia, x='estado_vendedor', y='garantia_de_fábrica')  
plt.xlabel('Estados', fontsize=13)  
plt.ylabel('Número de carros com garantia', fontsize=13)  
plt.title('Estados com maior número de carros com garantia de fábrica',  
fontsize=15)  
plt.legend(fontsize=10)  
plt.show()
```



**Usando a variável de revisões dentro da agenda**

```
train.revisoes_dentro_agenda.value_counts().index[0]
```

```
train.revisoes_dentro_agenda =  
train.revisoes_dentro_agenda.replace(train.revisoes_dentro_agenda.value_counts().  
index[0],1)  
train.revisoes_dentro_agenda = train.revisoes_dentro_agenda.fillna(0)
```

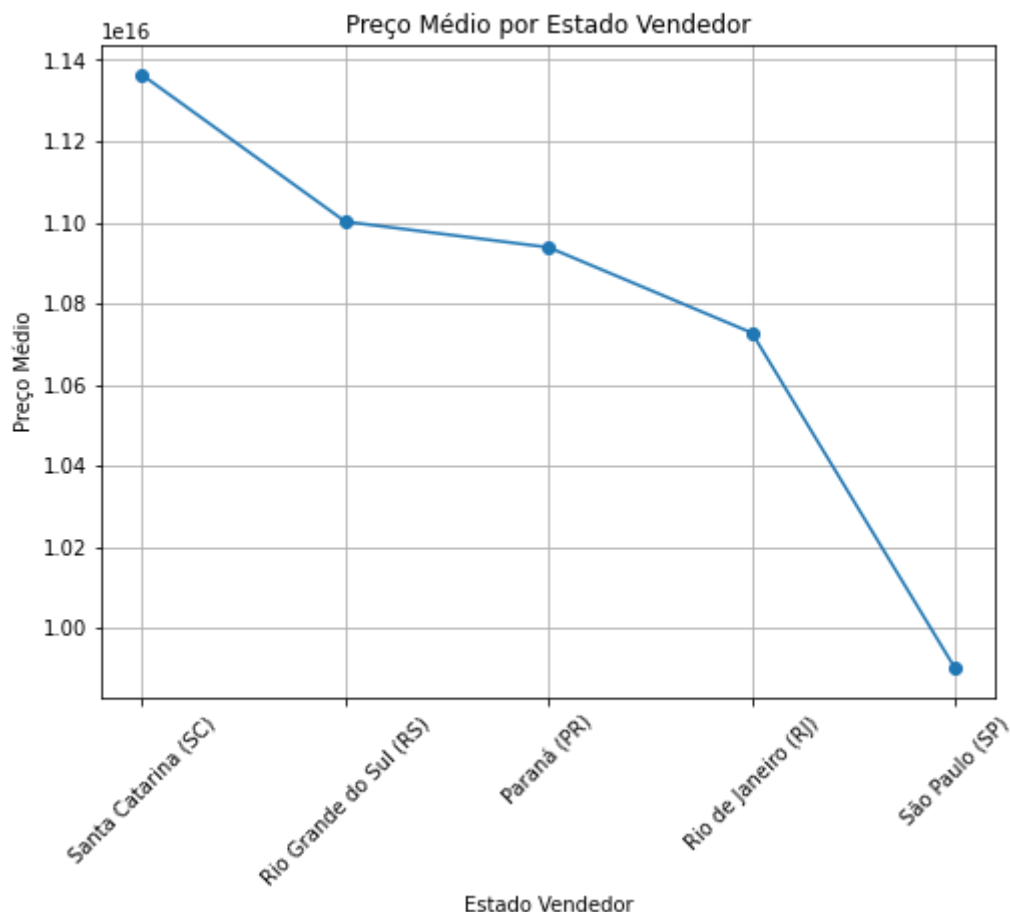
```
train[(train.garantia_de_fábrica==1)&(train.revisoes_dentro_agenda==1)][['estado_v  
endedor','garantia_de_fábrica','revisoes_dentro_agenda','preco']].groupby('estado_v  
endedor').count().sort_values('garantia_de_fábrica',ascending=False)[0:5].reset_inde  
x()
```

## Analizando a lista de estado com garantia de revisão

```
lista_estados = ['São Paulo (SP)','Paraná (PR)','Rio de Janeiro (RJ)','Santa Catarina  
(SC)','Rio Grande do Sul (RS)']  
garantia_revisao = listaFiltro(train,'estado_vendedor' ,lista_estados)
```

```
estado_preco =  
garantia_revisao[(garantia_revisao.garantia_de_fábrica==1)&(garantia_revisao.revis  
oes_dentro_agenda==1)][['estado_vendedor','preco']].groupby('estado_vendedor').m  
ean().sort_values('preco',ascending=False).reset_index()  
estado_preco
```

```
# Configuração do gráfico  
plt.figure(figsize=(8, 6))  
plt.plot(estado_preco['estado_vendedor'], estado_preco['preco'], marker='o')  
plt.xlabel('Estado Vendedor')  
plt.ylabel('Preço Médio')  
plt.title('Preço Médio por Estado Vendedor')  
plt.xticks(rotation=45)  
plt.grid(True)  
plt.show()
```



Com base no processo de análise e trabalhando com as variáveis , revisões dentro da agenda e garantia de revisão temos dados que apontam que as revisões dentro da agenda são melhores nos estados de São Paulo (SP), Rio de Janeiro (RJ), Paraná (PR), Santa Catarina (SC) e Rio Grande do Sul (RS) são os melhores estados para se comprar carros que ainda estejam dentro da garantia de fábrica e revisões dentro da agenda.

Essa conclusão é baseada no fato de que esses estados possuem os maiores números de garantias de fábrica e revisões dentro da agenda, indicando que há uma maior disponibilidade de carros com garantia de fábrica e serviços de revisão regularmente agendados.



E quando trata-se de garantia de revisão, os estados de Santa Catarina (SC), Rio Grande do Sul (RS), Paraná (PR), Rio de Janeiro (RJ) e São Paulo (SP) apresentam os maiores valores médios de preço para carros que possuem garantia de fábrica e revisões dentro da agenda.