

Files

- ..
- drive
- sample\_data
- costco\_warehouses.json

Read the data

- data set is costco\_warehouses.json

```
import pandas as pd
df = pd.read_json("costco_warehouses.json")
df.head()
```

	stlocID	displayName	identifier	phone	fax	address1	city	state	country	zipCode	...	tireCenterHours	gasPrices	coreServices
0	1	1	1	(206) 622-3136		4401 4TH AVE S	SEATTLE	WA	US	98134-2389	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouseid": "1", "regular": "2.799", "diesel": "2.799", "prem..."}, {"name": "Gas Station", "localizedName": "Gas"}	
1	2	2	2	(503) 252-2243		4849 NE 138TH AVE	PORLAND	OR	US	97230-3401	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouseid": "2", "regular": "2.899", "diesel": "2.899", "prem..."}, {"name": "Gas Station", "localizedName": "Gas"}	
2	6	6	6	(206) 575-9191		400 COSTCO DR STE 150	TUKWILA	WA	US	98188-4808	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouseid": "6", "regular": "2.959", "diesel": "2.959", "prem..."}, {"name": "Gas Station", "localizedName": "Gas"}	
3	8	8	8	(425) 827-1693		8629 120TH AVE NE	KIRKLAND	WA	US	98033-5865	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouseid": "8", "regular": "2.739", "diesel": "2.739", "prem..."}, {"name": "Gas Station", "localizedName": "Gas"}	
4	9	9	9	(503) 644-7615		15901 SW JENKINS RD	ALOHA	OR	US	97006-5098	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouseid": "9", "regular": "2.759", "diesel": "2.759", "prem..."}, {"name": "Gas Station", "localizedName": "Gas"}	

5 rows × 41 columns

Some Understanding of the dataset

```
[16] df.shape
(782, 41)
```

```
[17] df.columns
```

Index(['stlocID', 'displayName', 'identifier', 'phone', 'fax', 'address1', 'city', 'state', 'country', 'zipCode', 'manager', 'openDate', 'tiresDepartmentPhone', 'distance', 'latitude', 'longitude', 'parentGeoNodeID', 'active', 'languageID', 'hasGasDepartment', 'hasTiresDepartment', 'hasFoodDepartment', 'hasHearingDepartment', 'hasPharmacyDepartment', 'hasOpticalDepartment', 'hasBusinessDepartment', 'hasPhotoCenterDepartment', 'warehouseHours', 'pharmacyHours', 'upcomingHolidays', 'gasStationHours', 'tireCenterHours', 'gasPrices', 'coreServices', 'specialtyDepartments', 'locationName', 'isShipToWarehouse', 'isWarehousePickup', 'enableShipToHome', 'isBusinessWarehouse', 'closeDate'], dtype='object')
[18] df.dtypes

```
stlocID          int64
displayName      int64
identifier      int64
phone           object
fax             object
address1        object
city            object
state            object
country          object
zipCode          object
manager          object
openDate         object
tiresDepartment object
distance        float64
latitude         float64
longitude        float64
parentGeoNodeID int64
active           int64
languageID      int64
hasGasDepartment bool
hasTiresDepartment bool
hasFoodDepartment bool
hasHearingDepartment bool
hasPharmacyDepartment bool
hasOpticalDepartment bool
hasBusinessDepartment bool
hasPhotoCenterDepartment bool
warehouseHours   object
pharmacyHours    object
upcomingHolidays object
gasStationHours  object
tireCenterHours  object
gasPrices         object
coreServices     object
specialtyDepartments object
locationName     object
```

```

isShipToWarehouse      bool
isWarehousePickup     bool
enableShipToHome       bool
isBusinessWarehouse   bool
closeDate             object
dtype: object

```

```

[19] df.isnull().sum()
os

stlocID              0
displayName          0
identifier           0
phone                0
fax                  0
address1             0
city                 0
state                0
country              0
zipCode              0
manager               0
openDate              0
tiresDepartmentPhone 0
distance              0
latitude              0
longitude             0
parentGeoNodeID      0
active               0
languageID           0
hasGasDepartment     0
hasTiresDepartment   0
hasFoodDepartment    0
hasHearingDepartment 0
hasPharmacyDepartment 0
hasOpticalDepartment 0
hasBusinessDepartment 0
hasPhotoCenterDepartment 0
warehouseHours        0
pharmacyHours         0
upcomingHolidays      0
gasStationHours       0
tireCenterHours       0
gasPrices             203
coreServices          0
specialtyDepartments 0
locationName          0
isShipToWarehouse     0
isWarehousePickup     0
enableShipToHome       0
isBusinessWarehouse   0
closeDate             781
dtype: int64

```

#### ▼ drop closeDate column

```

[20] df.drop('closeDate', axis=1, inplace = True)
df.head()

```

	stlocID	displayName	identifier	phone	fax	address1	city	state	country	zipCode	...	gasStationHours	tireCenterHours	gasPri
0	1	1	1	(206) 622-3136		4401 4TH AVE S	SEATTLE	WA	US	98134-2389	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouse": "1", "regular": "2.7", "premium": "2.7"}
1	2	2	2	(503) 252-2243		4849 NE 138TH AVE	PORLTND	OR	US	97230-3401	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouse": "2", "regular": "2.8", "premium": "2.8"}
2	6	6	6	(206) 575-9191		400 COSTCO DR STE 150	TUKWILA	WA	US	98188-4808	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouse": "6", "regular": "2.9", "premium": "2.9"}
3	8	8	8	(425) 827-1693		8629 120TH AVE NE	KIRKLAND	WA	US	98033-5865	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouse": "8", "regular": "2.7", "premium": "2.7"}
4	9	9	9	(503) 644-7615		15901 SW JENKINS RD	ALOHA	OR	US	97006-5098	...	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	[{"title": "Mon-Fri", "code": "open", "time": "..."}]	{"warehouse": "9", "regular": "2.7", "premium": "2.7"}

5 rows × 40 columns

#### ▼ How many Costco store in each country

```

[21] df['country'].value_counts()
os

```

US	543
CA	99
MX	39
UK	29
JP	26
KO	16
TW	13
AU	12
ES	2
IS	1
FR	1

```
CH      1  
Name: country, dtype: int64
```

```
✓ [22] df.groupby(['country','state']).count()
```

country	state	stlocID	displayName	identifier	phone	fax	address1	city	zipCode	manager	openDate	...	gasStationHours	tireCenterHours	gasP
AU		6	6	6	6	6	6	6	6	6	6	6	6	6	6
	ACT	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	New South Wales	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	Queensland	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Victoria	2	2	2	2	2	2	2	2	2	2	2	2	2	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
US	UT	11	11	11	11	11	11	11	11	11	11	11	11	11	11
	VA	17	17	17	17	17	17	17	17	17	17	17	17	17	17
	VT	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	WA	32	32	32	32	32	32	32	32	32	32	32	32	32	32
	WI	9	9	9	9	9	9	9	9	9	9	9	9	9	9

116 rows × 38 columns



◀ Exploring Costco stores in Canada

```
✓ [23] df_CA = df[df['country'] == 'CA']  
df_CA.head()
```

	stlocID	displayName	identifier	phone	fax	address1	city	state	country	zipCode	...	gasStationHours	tireCenterHours	gasP
16	51	51	51	(604) 420-2668		3550 BRIGHTON AVE	BURNABY	BC	CA	V5A 4W3	...	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]	
17	54	54	54	(604) 270-3647		9151 BRIDGEPORT RD	RICHMOND	BC	CA	V6X 3L9	...	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]	
18	55	55	55	(604) 596-7435		7423 KING GEORGE BLVD	SURREY	BC	CA	V3W 5A8	...	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]	
19	56	56	56	(403) 299-1610		2853 32 ST NE	CALGARY	AB	CA	T1Y 6T7	...	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]	
20	57	57	57	(204) 788-4415		1315 ST JAMES ST	WINNIPEG	MB	CA	R3H 0K9	...	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]	

5 rows × 40 columns



◀ How many Costco store in each Province

```
✓ [24] df_CA['state'].value_counts()
```

```
ON    35  
QC    21  
AB    17  
BC    14  
MB     3  
SK     3  
NB     3  
NS     2  
NL     1  
Name: state, dtype: int64
```

```
✓ [25] #!pip install pywaffle
```

```
from pywaffle import Waffle  
import matplotlib.pyplot as plt  
  
# Data  
states_df = df_CA['state']  
value = states_df.value_counts()  
  
value_dict = value.to_dict()  
#store each state and its count in a dictionary as this will help with the labels and legend for the waffle chart
```



#### ▼ How many pick up locations are there

```

[26] df_CA.groupby(["isWarehousePickup"]).count()

      stlocID displayName identifier phone fax address1 city state country zipCode ... upcomingHolidays gasStationHours
isWarehousePickup
  False          1           1       1   1   1       1       1       1       1       1   ...
  True          98          98      98   98   98      98      98      98      98   98   ...
2 rows x 39 columns

```

#### ▼ Where is this one store that doesn't have pick up

```

[27] df_CA_No_PickUp = df_CA[df_CA["isWarehousePickup"] == False]
df_CA_No_PickUp

      stlocID displayName identifier phone fax address1 city state country zipCode ... gasStationHours tireCenterHours g
  314      595        595     595 (647) 484-5840      50 THERMOS SCARBOROUGH ON CA M1L 0E6 ...          0          0
1 rows x 40 columns

```

#### ▼ How many enableShipToHome Costco Stores are there

```

[28] df_CA.groupby(["enableShipToHome"]).count()

      stlocID displayName identifier phone fax address1 city state country zipCode ... upcomingHolidays gasStationHours
enableShipToHome
  False          99           99       99   99   99       99       99       99       99   ...
1 rows x 39 columns

```

# Seems like All the costco stores in Canada do not have enableShipToHome feature

#### ▼ View the Canada costco data again for further analysis

[29] df\_CA.head()

	stlocID	displayName	identifier	phone	fax	address1	city	state	country	zipCode	...	gasStationHours	tireCenterHours	gasP
16	51	51	51	(604) 420-2668		3550 BRIGHTON AVE	BURNABY	BC	CA	V5A 4W3	...	[]	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]
17	54	54	54	(604) 270-3647		9151 BRIDGEPORT RD	RICHMOND	BC	CA	V6X 3L9	...	[]	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]
18	55	55	55	(604) 596-7435		7423 KING GEORGE BLVD	SURREY	BC	CA	V3W 5A8	...	[]	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]
19	56	56	56	(403) 299-1610		2853 32 ST NE	CALGARY	AB	CA	T1Y 6T7	...	[]	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]
20	57	57	57	(204) 788-4415		1315 ST JAMES ST	WINNIPEG	MB	CA	R3H 0K9	...	[]	[]	[{"title": "Mon-Fri.", "code": "open", "time": "..."}]

5 rows × 40 columns

[30] df\_CA.columns

```
Index(['stlocID', 'displayName', 'identifier', 'phone', 'fax', 'address1', 'city', 'state', 'country', 'zipCode', 'manager', 'openDate', 'tiresDepartmentPhone', 'distance', 'latitude', 'longitude', 'parentGeoNodeID', 'active', 'languageID', 'hasGasDepartment', 'hasTiresDepartment', 'hasFoodDepartment', 'hasHearingDepartment', 'hasPharmacyDepartment', 'hasOpticalDepartment', 'hasBusinessDepartment', 'hasPhotoCenterDepartment', 'warehouseHours', 'pharmacyHours', 'upcomingHolidays', 'gasStationHours', 'tireCenterHours', 'gasPrices', 'coreServices', 'specialtyDepartments', 'locationName', 'isShipToWarehouse', 'isWarehousePickup', 'enableShipToHome', 'isBusinessWarehouse'], dtype='object')
```

#### ▼ How many store has a Gas Department

[31] df\_CA['hasGasDepartment'].value\_counts()

```
True    65
False   34
Name: hasGasDepartment, dtype: int64
```

View dataframe of the stores that has gas department along with the gasStationHours  
and the gasPrices

[32] df\_CA\_has\_GasDepartment = df\_CA[df\_CA["hasGasDepartment"] == True]

df\_CA\_has\_GasDepartment[['hasGasDepartment', 'gasStationHours', 'gasPrices']]

hasGasDepartment	gasStationHours	gasPrices
72	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "152", "regular": ".909", "pre...}
73	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "153", "regular": ".909", "pre...}
74	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "154", "regular": ".909", "pre...}
76	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "156", "regular": ".909", "pre...}
77	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "158", "diesel": ".019", "regu...}
...	...	...
672	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "1273", "regular": ".869", "pr...}
683	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "1296", "regular": ".909", "pr...}
693	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "1324", "regular": ".019", "pr...}
703	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "1345", "regular": ".909", "pr...}
706	["{"title": "Mon-Fri.", "code": "open", "time": "..."}"]	{"warehouseid": "11359", "diesel": ".019", "regu...}

65 rows × 3 columns

#### ▼ View location of Canada Costco stores, only the ones that has gas stations

[38] def folium\_deepnote\_show(m):
 data = m.get\_root().render()
 data\_fixed\_height = data.replace('width: 100%;height: 100%', 'width: 100%').replace('height: 100.0%', 'height: 609px;', 1)
 display(HTML(data\_fixed\_height))

▼ To view the map click this link and scroll down to the end of the file

<https://colab.research.google.com/drive/1mJH2LLVqwBJqniE7ElstYC1155bEsoj?usp=sharing>

```
✓ [40] import folium
      from IPython.core.display import display, HTML

      df_CA_has_GasDepartment_forMap = df_CA_has_GasDepartment[["latitude", "longitude", "locationName"]]

      map = folium.Map(
          location=[df_CA_has_GasDepartment_forMap.latitude.mean(),
                    df_CA_has_GasDepartment_forMap.longitude.mean()], zoom_start=5, control_scale=True)

      for index, location_info in df_CA_has_GasDepartment_forMap.iterrows():
          folium.Marker([location_info["latitude"], location_info["longitude"]], popup=location_info["locationName"]).add_to(map)

      folium_static.show(map)
```

```
✓ 0s
  ⏪ #!pip install streamlit-folium
  #Here I am trying to make the map visible in the github as I am getting an error
  # Error is : Make this Notebook Trusted to load map:File -> Trust Notebook
  ...
  import streamlit as st
  from streamlit_folium import st_folium
  import folium

  df_CA_has_GasDepartment_forMap = df_CA_has_GasDepartment[["latitude", "longitude", "locationName"]]
  map = folium.Map(
      location=[df_CA_has_GasDepartment_forMap.latitude.mean(),
                df_CA_has_GasDepartment_forMap.longitude.mean()], zoom_start=5, control_scale=True)

  for index, location_info in df_CA_has_GasDepartment_forMap.iterrows():
      folium.Marker([location_info["latitude"], location_info["longitude"]], popup=location_info["locationName"], tooltip = "Tool Tip").add_to(map)

  st_data = st_folium(map , width = 725)'''
```

```
'\nimport streamlit as st\nfrom streamlit_folium import st_folium\nimport folium\n\ndf_CA_has_GasDepartment = df_CA_has_GasDepartment[["latitude", "longitude", "locationName"]]\nmap = folium.Map(\n    location=[\n        df_CA_has_GasDepartment_forMap.latitude.mean(),\n        df_CA_has_GasDepartment_forMap.longitude.mean()],\n    zoom_start=5,\n    control_scale=True)\n\nfor index, location_info in df_CA_has_GasDepartment_forMap.iterrows():\n    folium.Marker([\n        location_info["latitude"],\n        location_info["longitude"]],\n        popup=location_info["locationName"],\n        tooltip = "Tool Map").add_to(map)\n\nst_data = st_folium(map , width = 725)'
```