

```
In [1]: import matplotlib.pyplot as plt
import pandas as pd
import os
```

## Gather data from csv files

```
In [2]: # Get list of files in directory
csv_files = [csv for csv in os.listdir() if csv.startswith('cart')]
```

```
In [3]: # Create list of df
df_list= [pd.read_csv(x, index_col='Item') for x in csv_files]
```

```
In [4]: # Strip unwanted characters; convert to float
for df in df_list:
    df['Price'] = [x.strip('$') for x in df['Price']]
    df['Price'] = [x.strip('c') for x in df['Price']]

    for x in df['Price']:
        df['Price'] = df['Price'].astype(float)
```

```
In [5]: # Iterate over df_list and change Price label according to Date
for df in df_list:
    date = pd.to_datetime(df['Date']).max().strftime('%m-%d-%Y') # grab max date
    df.rename(columns={'Price': date}, inplace=True) # change column label to date
    df.drop(columns='Date', inplace=True) # Remove columns no longer needed
```

```
In [6]: df_list[0].head()
```

```
Out[6]:
```

	<b>04-03-2020</b>
<b>Item</b>	
<b>President's Choice</b> Cremini Mushrooms, Whole227 g	2.28
<b>President's Choice</b> Green Beans340 g	3.98
<b>Dole</b> Ultimate Caesar Kit252 g	4.98
<b>Dole</b> Sunflower Bacon Crunch Chopped Salad Kit346 g	4.98
<b>Farmer's Market</b> FM Grape Tomato 907g908 g	4.98

## Create empty dataframe

```
In [7]: df_main = pd.DataFrame(columns=['Item'])
```

```
In [8]: for df in df_list:  
        df_main = pd.concat(df_list)
```

In [9]: df\_main

Out[9]:

	04-03-2020	04-07-2020	03-16-2020	04-11-2020	03-20-2020	03-26-2020	04-08-2020	04-17-2020	03-27-2020	04-22-2020	04-16-2020	03-17-2020
Item												
<b>President's Choice</b> <b>Cremini Mushrooms, Whole</b> 227 g	2.28	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
<b>President's Choice</b> <b>Green Beans</b> 340 g	3.98	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
<b>Dole</b> <b>Ultimate Caesar Kit</b> 252 g	4.98	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
<b>Dole</b> <b>Sunflower Bacon Crunch Chopped Salad Kit</b> 346 g	4.98	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
<b>Farmer's Market</b> <b>FM Grape Tomato</b> 907g908 g	4.98	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...
<b>Tomato on the Vine Red</b>	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5.45
<b>White Mushrooms</b>	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.26
<b>Italian Parsley</b> 1 ea	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.48
<b>Cilantro</b> 1 ea	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.98
<b>President's Choice</b> <b>White Mushrooms, Whole</b> 227 g	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2.28

115740 rows × 12 columns

```
In [10]: df_main.iloc[0]
```

```
Out[10]: 04-03-2020    2.28
         04-07-2020    NaN
         03-16-2020    NaN
         04-11-2020    NaN
         03-20-2020    NaN
         03-26-2020    NaN
         04-08-2020    NaN
         04-17-2020    NaN
         03-27-2020    NaN
         04-22-2020    NaN
         04-16-2020    NaN
         03-17-2020    NaN
         Name: President's ChoiceCremini Mushrooms, Whole227 g, dtype: float64
```

## Visualization

```
In [11]: df_main.iloc[1:].plot()
```

```
Out[11]: <matplotlib.axes._subplots.AxesSubplot at 0x7f59ee8dac18>
```

