In [2]:

```
import pandas as pd
import matplotlib.pyplot as plt
```

In [3]:

```
# run plots in the notebook
%matplotlib inline
url = "http://pbpython.com/extras/sample-salesv2.csv"
sales = pd.read_csv(url)
```

In [18]:

```
sales.columns = [c.strip().lower().replace(' ', '_') for c in sales.columns]
```

In [20]:

```
sales[['name', 'category','quantity','unit_price']][1:]
```

Out[20]:

	name	category	quantity	unit_price
1	Heidenreich-Bosco	Shoes	19	53.62
2	Kerluke, Reilly and Bechtelar	Shirt	12	24.16
3	Waters-Walker	Shirt	5	82.68
4	Waelchi-Fahey	Shirt	18	99.64
5	Kerluke, Reilly and Bechtelar	Shirt	17	52.82
...	...	...	...	...
995	Kihn, McClure and Denesik	Belt	11	60.50
996	Kunze Inc	Shoes	16	19.66
997	Waters-Walker	Shoes	13	90.95
998	Huel-Haag	Shoes	9	98.22
999	Heidenreich-Bosco	Shoes	14	74.83

999 rows × 4 columns

In [21]:

```
shirts = sales[sales['category'] == "Shirt"]
```

In [24]:

```
print(shirts)
```

```
     account_number                            name       sku categor
y  \
2            563905  Kerluke, Reilly and Bechtelar  AS-93055    Shir
t
3             93356                Waters-Walker  AS-93055    Shir
t
4            659366                Waelchi-Fahey  AS-93055    Shir
t
5            563905  Kerluke, Reilly and Bechtelar  AS-93055    Shir
t
9            563905  Kerluke, Reilly and Bechtelar  KV-99194    Shir
t
..              ...                         ...       ...      ..
.
982          115138                Gorczany-Hahn  VG-32047    Shir
t
984          711951             Kilback-Gerlach  VG-32047    Shir
t
986           93356                Waters-Walker  KV-99194    Shir
t
990           93356                Waters-Walker  AS-93055    Shir
t
992          850140                    Kunze Inc  KV-99194    Shir
t

     quantity  unit_price  ext_price                 date
2          12       24.16     289.92  2014-03-01 10:51:24
3           5       82.68     413.40  2013-11-17 20:41:11
4          18       99.64    1793.52  2014-01-03 08:14:27
5          17       52.82     897.94  2013-12-04 02:07:05
9          12       26.98     323.76  2014-05-20 00:21:28
..        ...         ...        ...                  ...
982        20       72.35    1447.00  2014-08-11 19:02:14
984         1       91.96      91.96  2013-10-23 15:08:31
986        20       99.74    1994.80  2014-04-18 15:39:12
990         9       68.62     617.58  2014-08-12 08:06:21
992         9       27.16     244.44  2014-02-08 06:22:34

[404 rows x 8 columns]
```

In [39]:

```
shirts = sales[sales["category"] == "Shirt"].copy()
```

In [40]:

```
shirts['shirt_sales'] = (shirts["quantity"] * shirts["unit_price"])
```

In [54]:

```
shirts_by_company = shirts.groupby('name', as_index=False).sum()
```

In [55]:

```
shirts_by_company
```

Out[55]:

	name	account_number	quantity	unit_price	ext_price	shirt_sales
0	Berge LLC	14698440	166	1226.54	9670.24	9670.24
1	Carroll PLC	6529798	257	1098.93	13717.61	13717.61
2	Cole-Eichmann	20900607	236	1226.75	14528.01	14528.01
3	Davis, Kshlerin and Reilly	4892848	161	828.51	7533.03	7533.03
4	Ernser, Cruickshank and Lind	13334256	262	1500.25	16944.19	16944.19
5	Gorczany-Hahn	2417898	237	1132.22	12576.83	12576.83
6	Hamill-Hackett	11506399	148	1091.55	8880.04	8880.04
7	Hegmann and Sons	13624546	278	1528.84	16774.47	16774.47
8	Heidenreich-Bosco	980220	92	582.24	5965.25	5965.25
9	Huel-Haag	5792340	200	1146.17	11944.01	11944.01
10	Kerluke, Reilly and Bechtelar	11842005	269	1038.53	12958.23	12958.23
11	Kihn, McClure and Denesik	20469591	288	1653.58	18956.35	18956.35
12	Kilback-Gerlach	12103167	163	1052.53	9904.85	9904.85
13	Koelpin PLC	3018885	132	786.07	7908.28	7908.28
14	Kunze Inc	21253500	260	1439.92	15638.87	15638.87
15	Kuphal, Zieme and Kub	6594962	252	1167.28	12101.14	12101.14
16	Senger, Upton and Breitenberg	15799800	144	939.38	7659.70	7659.70
17	Volkman, Goyette and Lemke	15009220	220	1136.25	12791.27	12791.27
18	Waelchi-Fahey	13187320	201	1057.67	11689.05	11689.05
19	Waters-Walker	2240544	288	1603.36	18633.71	18633.71

In [56]:

```
top_sellers = shirts_by_company.sort_values(by='name', ascending=False).head(10)
```

In [57]:

```
top_sellers
```

Out[57]:

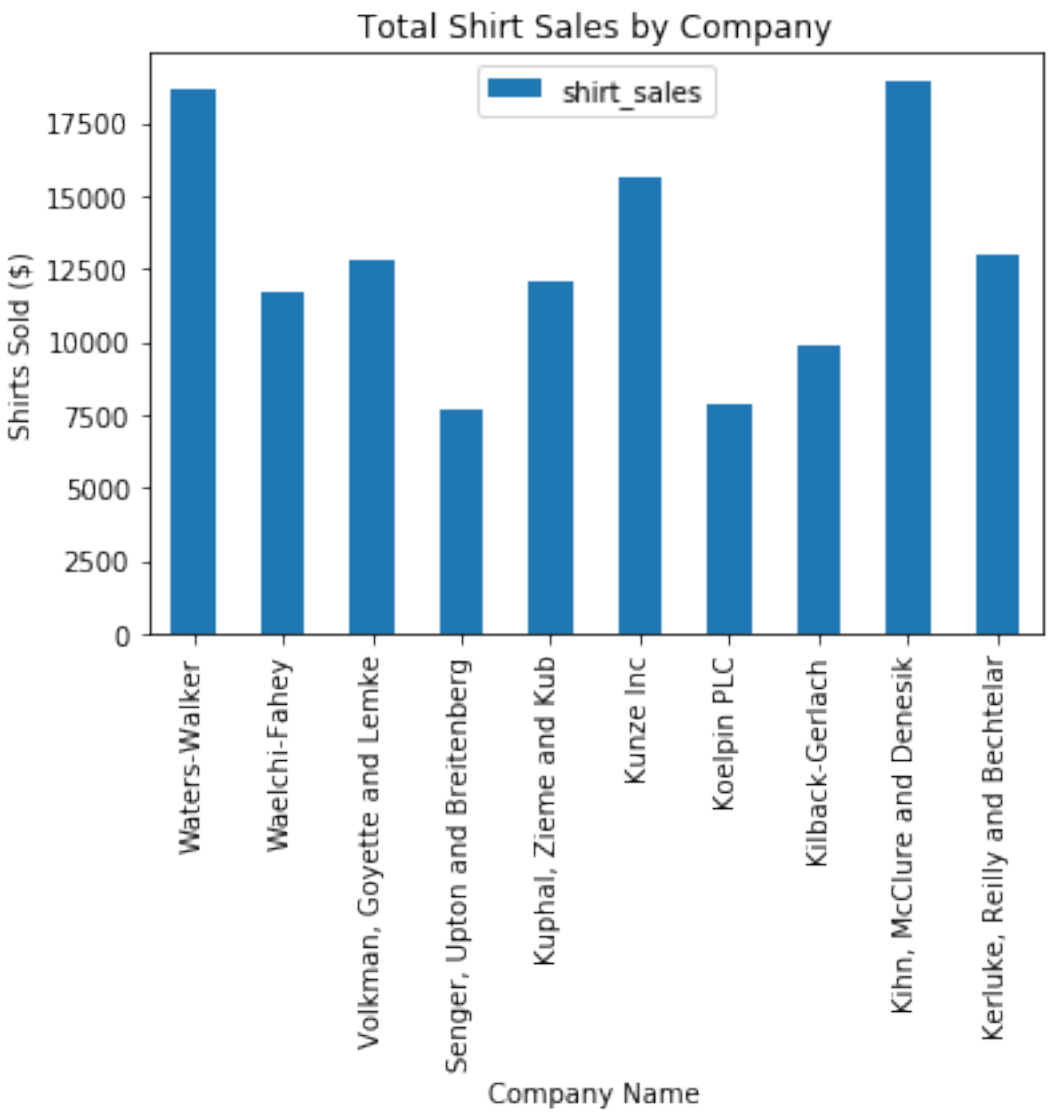
	name	account_number	quantity	unit_price	ext_price	shirt_sales
19	Waters-Walker	2240544	288	1603.36	18633.71	18633.71
18	Waelchi-Fahey	13187320	201	1057.67	11689.05	11689.05
17	Volkman, Goyette and Lemke	15009220	220	1136.25	12791.27	12791.27
16	Senger, Upton and Breitenberg	15799800	144	939.38	7659.70	7659.70
15	Kuphal, Zieme and Kub	6594962	252	1167.28	12101.14	12101.14
14	Kunze Inc	21253500	260	1439.92	15638.87	15638.87
13	Koelpin PLC	3018885	132	786.07	7908.28	7908.28
12	Kilback-Gerlach	12103167	163	1052.53	9904.85	9904.85
11	Kihn, McClure and Denesik	20469591	288	1653.58	18956.35	18956.35
10	Kerluke, Reilly and Bechtelar	11842005	269	1038.53	12958.23	12958.23

In [58]:

```
shirt_plot = top_sellers.plot(kind="bar",
                              title="Total Shirt Sales by Company",
                              x="name",
                              y="shirt_sales")
shirt_plot.set_xlabel("Company Name")
shirt_plot.set_ylabel("Shirts Sold ($)")
```

Out[58]:

```
Text(0, 0.5, 'Shirts Sold ($)')
```



In [ ]: