

MSDS-593: Homework 2 Notebook

Tiance Tan

Jacek Plonowski

1) Redesign a bubble chart

```
In [ ]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.pylab as pylab

%matplotlib inline
%config InlineBackend.figure_format = 'retina'
```

```

In [ ]: data = {'Name': ['Walmart', 'Sinopec Group', 'State Grid', 'China National Petroleum',
                        'Royal Dutch Shell', 'Saudi Aramco', 'Volkswagen', 'BP', 'Amazon.com', 'Toyota Motor'],
                'Revenue': [523964, 407009, 383906, 379130, 352106, 329784, 282760, 282616, 280522, 275288],
                'Profits': [14881.00, 6793.20, 7970.00, 4443.20, 15842.00, 88210.90, 15542.00, 4026.00, 11588.00, 19096.20]}

fig, ax = plt.subplots(nrows=1, ncols=1, figsize=(4,4))

x1, y1 = [data["Profits"][0], data['Revenue'][0]], [10,10]
x2, y2 = [data["Profits"][1], data['Revenue'][1]], [9,9]
x3, y3 = [data["Profits"][2], data['Revenue'][2]], [8,8]
x4, y4 = [data["Profits"][3], data['Revenue'][3]], [7,7]
x5, y5 = [data["Profits"][4], data['Revenue'][4]], [6,6]
x6, y6 = [data["Profits"][5], data['Revenue'][5]], [5,5]
x7, y7 = [data["Profits"][6], data['Revenue'][6]], [4,4]
x8, y8 = [data["Profits"][7], data['Revenue'][7]], [3,3]
x9, y9 = [data["Profits"][8], data['Revenue'][8]], [2,2]
x10, y10 = [data["Profits"][9], data['Revenue'][9]], [1,1]

plt.plot(x1,y1,x2,y2,x3,y3,x4,y4,x5,y5,
         x6,y6,x7,y7,x8,y8,x9,y9,x10,y10,marker = 'o',c = 'black')

ax.tick_params(axis='x', colors='dimgrey')

ax.spines['top'].set_visible(False)
ax.spines['right'].set_visible(False)
ax.spines['bottom'].set_visible(False)
ax.spines['left'].set_visible(False)
ax.set_yticks([])
ax.set_xticks([])

for i in range(1,11):
    plt.axhline(y=i, xmin=0, xmax=1.2, linestyle=':', c = 'grey')

ax.text(-0.8, 0.93, 'Walmart', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold',transform=ax.transAxes)
ax.text(-0.8, 0.83, 'Sinopec Group', fontsize = 'large', c = 'Black',
        fontweight='bold',fontname="Times New Roman",transform=ax.transAxes)
ax.text(-0.8, 0.73, 'State Grid', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold', transform=ax.transAxes)
ax.text(-0.8, 0.63, 'China National Petroleum', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold', transform=ax.transAxes)
ax.text(-0.8, 0.53, 'Royal Dutch Shell', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold', transform=ax.transAxes)

```

```

ax.text(-0.8, 0.43, 'Saudi Aramco', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold', transform=ax.transAxes)
ax.text(-0.8, 0.33, 'Volkswagen', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold', transform=ax.transAxes)
ax.text(-0.8, 0.23, 'BP', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold', transform=ax.transAxes)
ax.text(-0.8, 0.13, 'Amazon.com', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold', transform=ax.transAxes)
ax.text(-0.8, 0.03, 'Toyota Motor', fontsize = 'large', c = 'Black',
        fontname="Times New Roman",fontweight='bold', transform=ax.transAxes)

ax.text(-0.8, 1.2, 'Revenue and profits data for the top 10 global companies',
        fontsize = 'large', c = 'white',
        fontname="Times New Roman",fontweight='bold',
        transform=ax.transAxes, bbox=dict(facecolor='black', alpha=0.7))

ax.text(-0.8, 1.1, 'Billions of Dollars', fontsize = 'medium', c = 'Black', transform=ax.transAxes)
ax.text(0, 1.1, 'Profits ← Revenue', fontsize = 'medium', c = 'Black', transform=ax.transAxes)
ax.text(0.6, 1.1, 'Source: Fortune', fontsize = 'medium', c = 'Black', transform=ax.transAxes)

ax.text(0.02, 0.98, '14.9', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.01, 0.88, '6.8', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.01, 0.78, '8.0', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.01, 0.68, '4.4', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.01, 0.58, '15.8', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.1, 0.48, '88.2', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.01, 0.38, '15.5', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.01, 0.28, '4.0', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.01, 0.18, '11.6', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.01, 0.08, '19.1', fontsize = 'small', c = 'Black', transform=ax.transAxes)

ax.text(0.9, 0.98, '524.0', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.7, 0.88, '407.0', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.65, 0.78, '383.9', fontsize = 'small', c = 'Black', transform=ax.transAxes)

```

```

ax.text(0.62, 0.68, '379.1', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.57, 0.58, '352.1', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.54, 0.48, '329.8', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.45, 0.38, '282.8', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.45, 0.28, '282.6', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.45, 0.18, '280.5', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.44, 0.08, '275.3', fontsize = 'small', c = 'Black', transform=ax.transAxes)

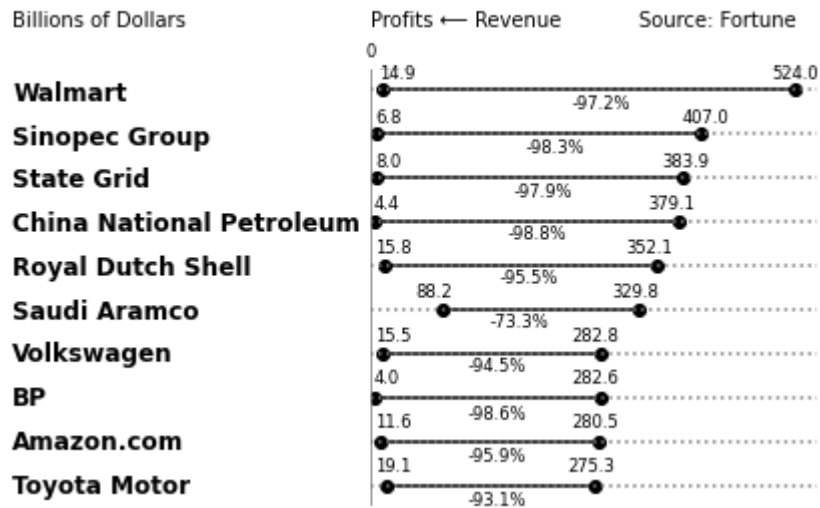
ax.text(0.45, 0.91, '-97.2%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.35, 0.81, '-98.3%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.32, 0.71, '-97.9%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.31, 0.61, '-98.8%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.29, 0.51, '-95.5%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.27, 0.41, '-73.3%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.22, 0.31, '-94.5%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.22, 0.2, '-98.6%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.22, 0.1, '-95.9%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(0.22, 0, '-93.1%', fontsize = 'small', c = 'Black', transform=ax.transAxes)
ax.text(-0.01, 1.03, '0', fontsize = 'small', c = 'Black', transform=ax.transAxes)

plt.axvline(x = 0, c = 'grey')
plt.xlim(-1000, 550000)

plt.show()

```

Revenue and profits data for the top 10 global companies



2) Decluttering

```
In [ ]: months = pd.Series(['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun',
                           'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec'])

sales_18 = pd.Series([38, 47, 49, 42, 42, 38, 43, 52, 49, 67, 74, 86
])
sales_19 = pd.Series([101, 78, 91, 77, 78, 71, 78, 62, 68, 50, 43, 41
])
```

```
In [ ]: fig, ax = plt.subplots()

ax.plot(months, sales_18, label='2018')
ax.plot(months, sales_19, label='2019')

ax.set_ylabel('Quantity Sold (Thousands)')
ax.grid(alpha=0.25)

ax.spines['top'].set_visible(False)
ax.spines['right'].set_visible(False)

plt.title('Vehicles Sold Per Month in 2018 and 2019')
ax.legend(loc=1)
plt.show()
```

