

week9solutions

December 6, 2018

```
In [42]: # Exc 1.a - Read in data
import numpy as np
deaths = np.genfromtxt("data/Deaths.csv", delimiter=",", names=True, dtype=None)
spending = np.genfromtxt("data/Household_Expenditure.csv", delimiter=",", names=True)

/usr/lib/python3/dist-packages/ipykernel_launcher.py:3: VisibleDeprecationWarning: Reading uni
This is separate from the ipykernel package so we can avoid doing imports until

In [3]: # Can access data in arrays like we would with lists
for row in spending:
    print("Year: %i, Percentage spent on alcohol: %0.2f" %(row[0], row[3]))

In [105]: # Exc 1.b - Simple Graph
# Let's plot the total amount spent, and the percentage spent, on alcohol
import matplotlib.pyplot as plt

# Easier to see what's happening if we make new lists
year = []
total = []
percentage = []

for row in spending:
    year.append(row[0])
    total.append(row[2])
    percentage.append(row[3])

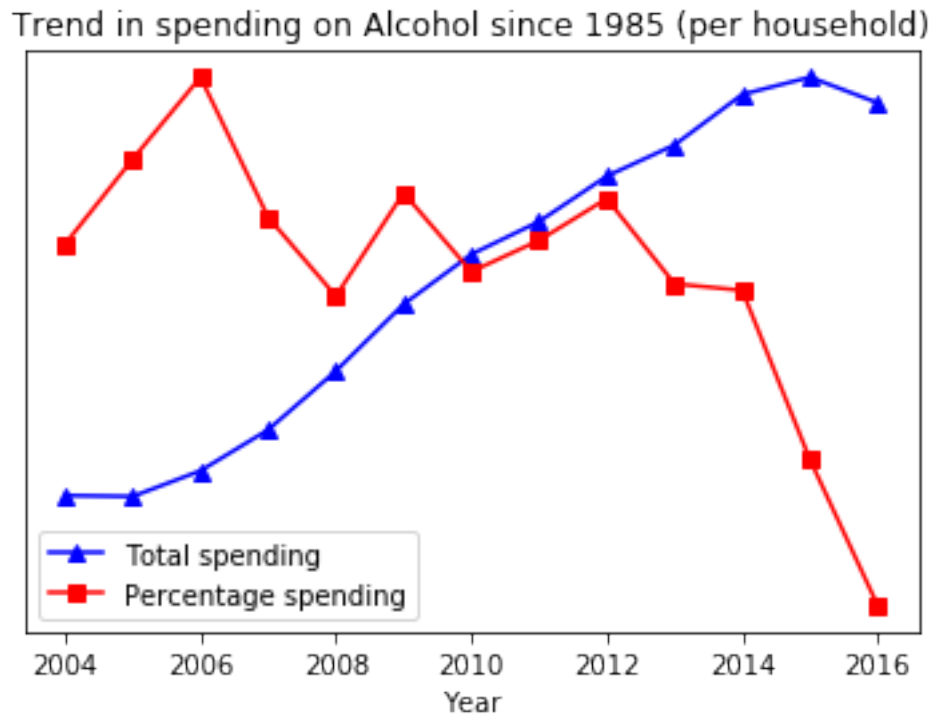
# CHALLENGE: Normalize total spending data, as we only care about the trend
total /= np.max(total)
percentage /= np.max(percentage)

# Plot the data
plt.plot(year, total, label="Total spending", color='b', marker="^")
plt.plot(year, percentage, label="Percentage spending", color='r', marker="s")

# Make it pretty
plt.xlabel("Year")
plt.yticks([]) # No ticks
```

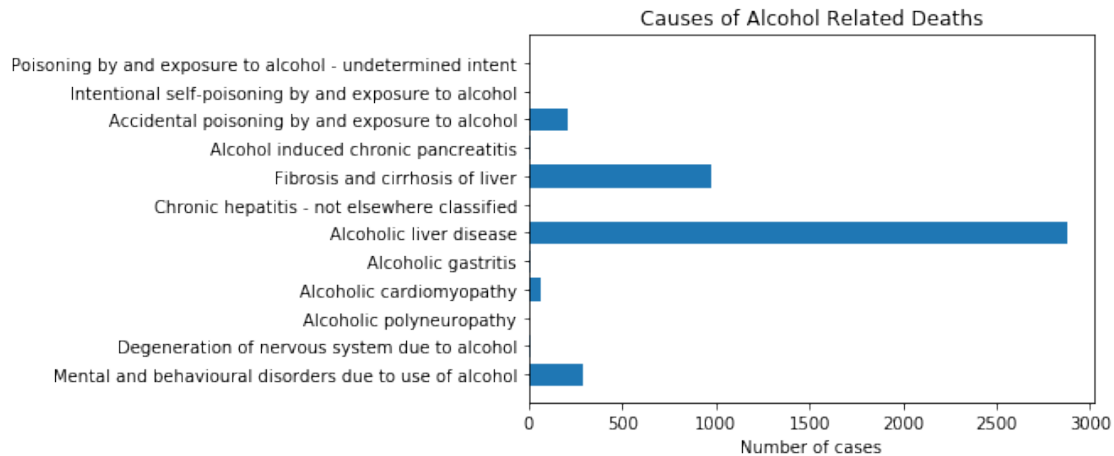
```
plt.title("Trend in spending on Alcohol since 1985 (per household)")
plt.legend() # Draw a legend

plt.show()
# Could also save: plt.savefig("myplot.pdf") etc.
```



```
In [118]: # Exc. 1c Bar Chart
# Plot causes of deaths as a bar chart (2015 data)
labels = []
counts = []
for row in deaths:
    label = row[0]
    labels.append(label)
    counts.append(row[2])

plt.barh(labels, counts)
plt.xlabel("Number of cases")
plt.title("Causes of Alcohol Related Deaths")
plt.show()
```



```

In [87]: # Exc. 2
         # Alcohol Price
         # Drug Prescriptions
         alcohol_prices = np.genfromtxt("data/Alcohol_Prices.csv", delimiter=",", names=True)
         drug_prescription = np.genfromtxt("data/Drugs_Prescribed.csv", delimiter=",", names=True)

In [116]: # Exc. 2a Plot the data
          # Be careful to ensure same number of data points from both
          # This is not needed, but requested
          years = []
          disposable_income = []
          national_drug_prescription = []

          for row in alcohol_prices:
              if row[0] >= 2004: # Only want data since 2004
                  years.append(row[0])
                  disposable_income.append(row[5])

          for row in drug_prescription:
              national_drug_prescription.append(row[1])

          # Normalize data
          disposable_income /= np.max(disposable_income)
          national_drug_prescription /= np.max(national_drug_prescription)

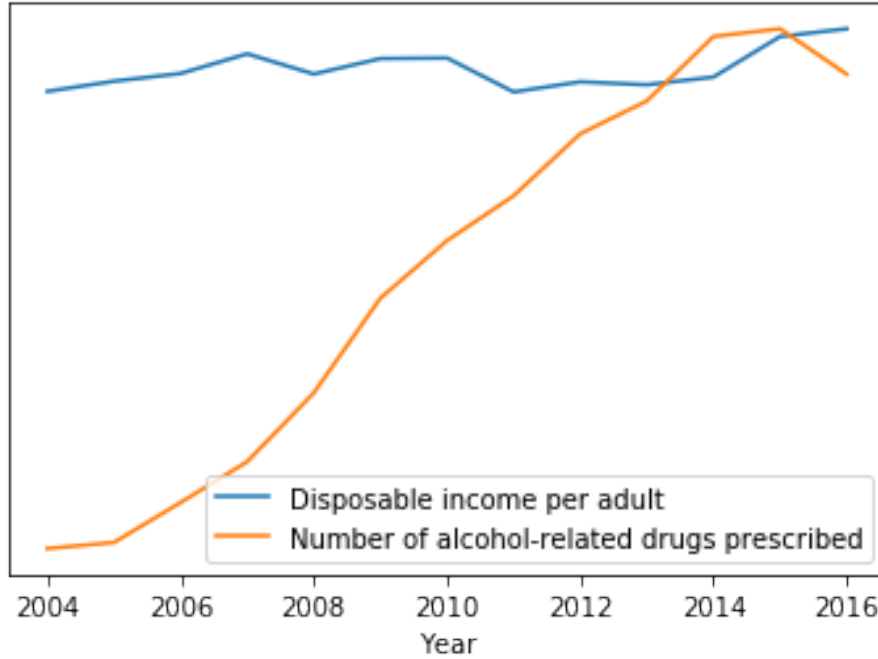
          plt.plot(years, disposable_income, label="Disposable income per adult")
          plt.plot(years, national_drug_prescription, label="Number of alcohol-related drugs p

          plt.xlabel("Year")
          plt.yticks([])

```

```
plt.title("Are more people ending up in A&E because they have more money?")
plt.legend()
plt.show()
```

Are more people ending up in A&E because they have more money?



In []: