Exercise set 0

Problem 1

Task a

Load the data set in x.csv:

```
import pandas as pd

x = pd.read_csv("x.csv")
x
```

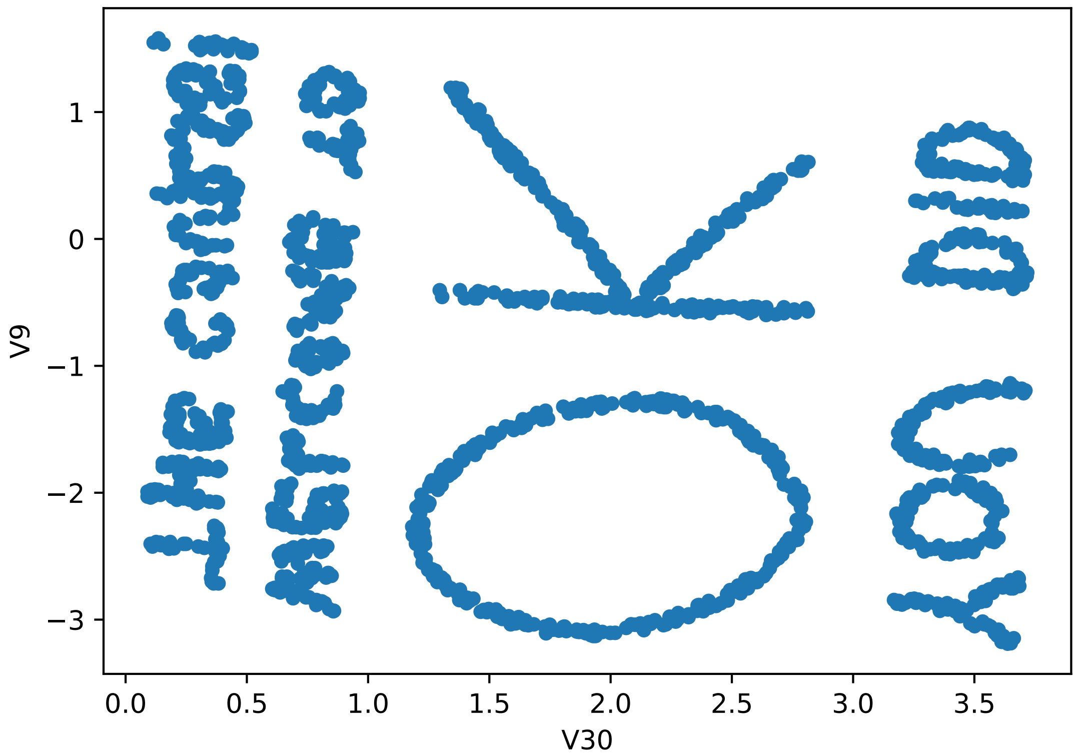
```
##             V1         V2        V3  ...       V30       V31        V32
## 0     0.641520   8.875380 -0.228512  ...  3.365227 -1.371321 -10.625266
## 1    -0.838917   9.017552 -0.881894  ...  0.279617  0.473437  -9.965022
## 2    -0.036031  10.924529  1.498176  ...  0.400544  1.438860  -9.887093
## 3     0.169189  10.273841  0.725405  ...  3.483404 -0.361023  -9.553151
## 4     0.200184  10.307359  2.231854  ...  2.404839  0.052692 -11.492905
## ...        ...        ...       ...  ...       ...       ...        ...
## 2043  1.191834  11.186569  2.703404  ...  1.673791  1.136313 -10.657274
## 2044 -0.344549  10.056357  0.970595  ...  0.408082 -0.220003  -8.859562
## 2045  0.738210  10.772708  3.186998  ...  3.584401 -0.422226 -11.176845
## 2046  0.000416   9.066987  0.717677  ...  1.532421 -0.125529 -10.787022
## 2047  1.736197  10.597908  1.774258  ...  1.396373  1.117738  -7.948153
##
## [2048 rows x 32 columns]
```

Finding the two variables having the largest variances:

```
import matplotlib.pyplot as plt

largest = x.var().idxmax()
largest_df = x.pop(largest).to_frame()

large = x.var().idxmax()
large_df = x.pop(large).to_frame()

new_x = large_df.join(largest_df)
```

Plotting variables:

```
new_x.plot(x=large, y=largest, kind='scatter')
plt.show()
```



Problem 2

Task a

Let's break down B:

$$B = \sum_{j=1}^{n} \lambda_j x_j x_j^T = \begin{pmatrix} x_1 & x_2 & ... & x_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & & & \\ 0 & 0 & ... & \lambda_n \end{pmatrix} \begin{pmatrix} x_1^T \\ x_2^T \\ ... \\ x_n^T \end{pmatrix}$$

For λ_1 and x_1 to be the eigenvalues/-vectors of matrix B we must prove that

$$B_i x_i = \lambda_i x_i \; \forall i = 1..n$$

Now we have the following:

$$B_i x_i = \lambda_i \begin{pmatrix} x_{1i}x_{1i} & ... & x_{1i}x_{ni} \\ ... & ... & ... \\ x_{ni}x_{1i} & ... & ...x_{ni}x_{ni} \end{pmatrix} \begin{pmatrix} x_{1i} \\ ... \\ x_{ni} \end{pmatrix} = \lambda_i \begin{pmatrix} x_{1i} \sum_{j=1}^{n} x_{ji}^2 \\ ... \\ x_{ni} \sum_{j=1}^{n} x_{ji}^2 \end{pmatrix}$$

The last vector is orthonormal giving it the sum of 1 and proving that the vector is x_i. This proves that

$$B_i x_i = \lambda_i x_i \; \forall i = 1, ..., n$$

And so λ_i and x_i are eigenvalues/-vectors of B.

Task b

$$\begin{vmatrix} 1 & 2 \\ 2 & 3.14159 \end{vmatrix} \to \begin{vmatrix} -\lambda+1 & 2 \\ 2 & -\lambda+3.14159 \end{vmatrix}$$

With calculator: Eigenvalues: $\lambda_1 = \frac{\sqrt{205864077281}+414159}{200000}$ $\lambda_2 = -\frac{\sqrt{205864077281}+414159}{200000}$ Eigenvectors: $x_1 = \frac{\sqrt{205864077281}+214159}{400000} * x_2$ $x_2 = x_1$

Problem 3

Task a

In order for E to be a linear operator, we need to show that: $E[X+Y] = E[X] + E[Y]$ and $E[tX] = tE[X]$.

Now

$$E[X+Y] = \sum_{\omega\in\Omega}(X+Y)(\omega)P(\omega) = \sum_{\omega\in\Omega} X(\omega)P(\omega) + \sum_{\omega\in\Omega} Y(\omega)P(\omega)$$

This gives us:

$$E[X+Y] = E[X] + E[Y]$$

$$E[tX] = \sum_{\omega\in\Omega} tX(\omega)P(\omega) = t\sum_{\omega\in\Omega} X(\omega)P(\omega)$$

Which gives us:

$$E[tX] = tE[X]$$

According to this proof we can declare that E is a linear operator.

Task b

$\mathrm{Var}(X) = E[(X-\mu)^2] = E[x^2-2\mu x+\mu^2] = E[x^2]-E[\mu(2x-\mu)] = E[x^2]-\mu E[2x-\mu] = E[x^2]-u*(2u-u) = E[X^2] - E[X]^2$

Problem 4

Task a

$$P(X\cap Y) = P(Y|X)P(X)$$

Formula for conditional probability:

$$P(X|Y) = \frac{P(X\cap Y)}{P(Y)} = \frac{P(Y|X)P(X)}{P(Y)}$$

Task b

Allergic:

Y	N
0.2	0.8

Test result:

Allergic	Pos	Neg
1	0.85	0.15
0	0.23	0.77

Calculating ratio:

$$R = \frac{P(''Allergic''|''Pos'')}{P(''Notallergic''|''Pos'')} = \frac{P(''Pos''|''Allergic'')P(''Allergic'')}{P(''Pos''|''Notallergic'')P(''Notallergic'')} = \frac{0.85 * 0.2}{0.23 * 0.8} = 0.17/0.184 = 0.92391...$$

$$P(''Allergic''|''Pos'') = \frac{R}{1+R} = 0.4802259...$$

Probability is about 48%

Problem 5

Task a

$$f'(x) = 4ax^3 + b = 0$$

$$f''(x_o) = 0, x_0 = \frac{-b}{4a}$$

(Minimum) Following:

$$f_{min} = f(x_o) = \frac{3b}{4}(\frac{-b}{4a})^{1/3} + c$$

Task b

Only condition is $a > 0$.

Problem 6

Task a

```
function fibonacci(n):
  counter = 0
  fib1 = 0
  fib2 = 1
  while counter < n:
    counter = counter+1
    print(fib1)
    sum = fib1 + fib2
    fib1 = n2
    fib2 = sum
```

Task b

O(n)