데이터융합SW과
김규석 교수

Data Analysis with Java

데이터 분석 프로그래밍03

Objective of Today’s Class

AES256

- Practicing Encryption and Decryption

Pearson Correlation Coefficient

- Measuring how strong a relationship is between two variables

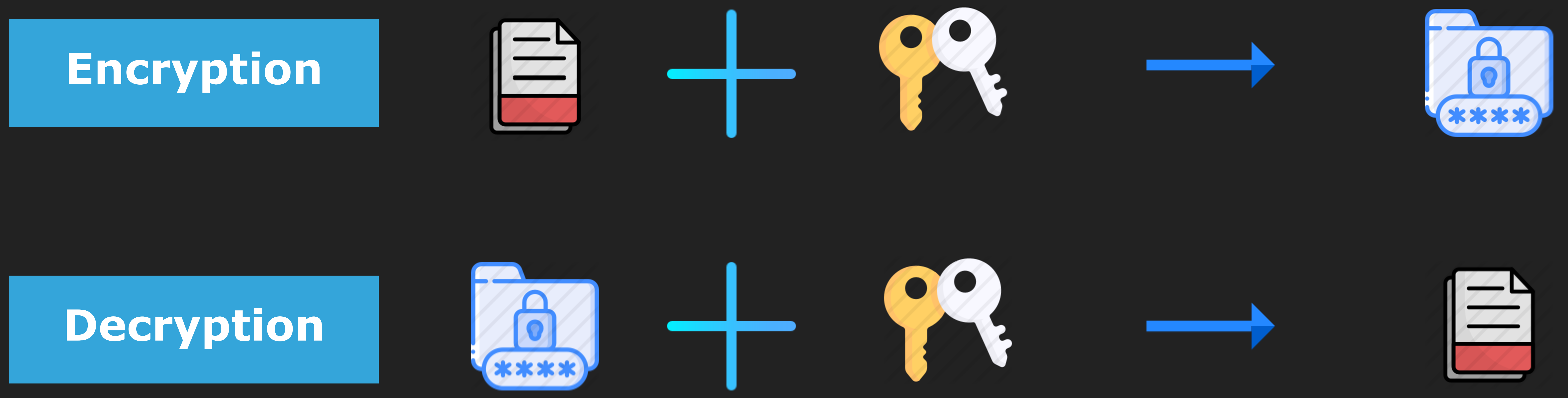
AES256(Cont’d)

Encryption

- A process which transforms the original information into an unrecognizable form

Decryption

- A process of converting encoded/encrypted data in a form that is readable and understood by a human or a computer



Encryption

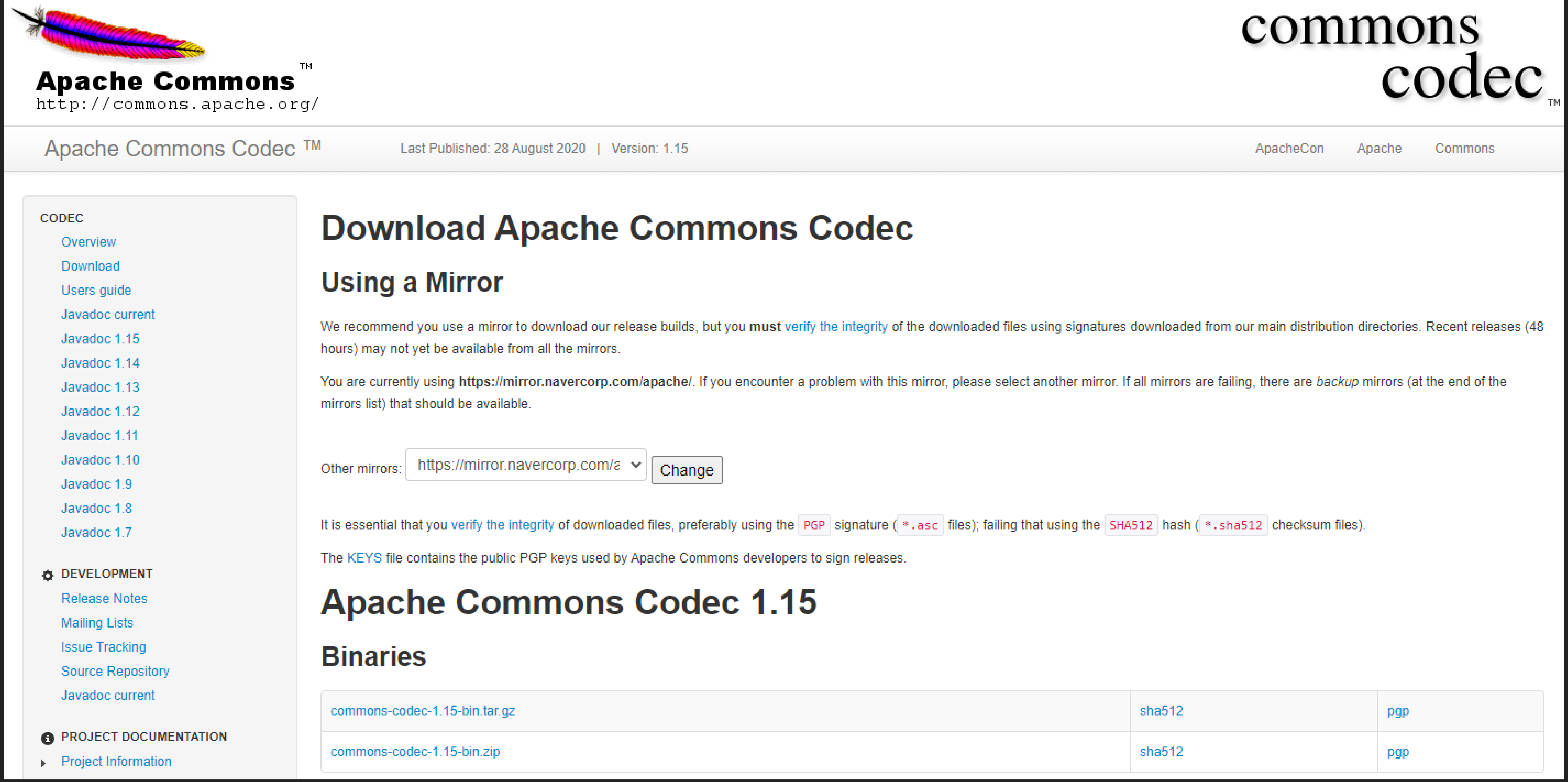
Decryption

Source : https://www.guru99.com/difference-encryption-decryption.html / image: Flaticon.com

AES256(Cont'd)

Download the Related Library and Add it to the Project

- https://commons.apache.org/proper/commons-codec/download_codec.cgi



Last Published: 28 August 2020 | Version: 1.15

Download Apache Commons Codec

Using a Mirror

We recommend you use a mirror to download our release builds, but you **must** verify the integrity of the downloaded files using signatures downloaded from our main distribution directories. Recent releases (48 hours) may not yet be available from all the mirrors.

You are currently using **https://mirror.navercorp.com/apache/**. If you encounter a problem with this mirror, please select another mirror. If all mirrors are failing, there are *backup* mirrors (at the end of the mirrors list) that should be available.

Other mirrors: https://mirror.navercorp.com/a

It is essential that you verify the integrity of downloaded files, preferably using the `PGP` signature (`*.asc` files); failing that using the `SHA512` hash (`*.sha512` checksum files).

The KEYS file contains the public PGP keys used by Apache Commons developers to sign releases.

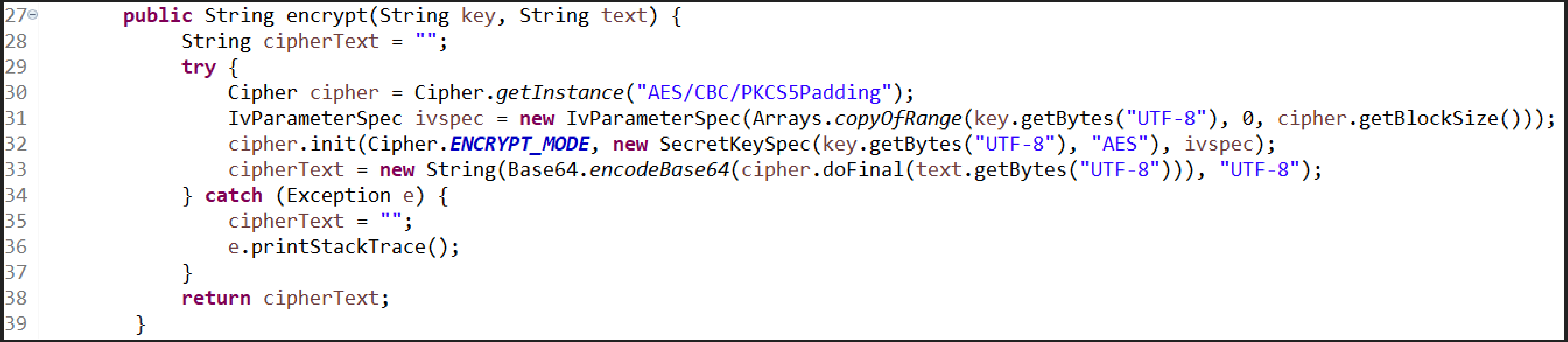
Apache Commons Codec 1.15

Binaries

commons-codec-1.15-bin.tar.gz	sha512	pgp
commons-codec-1.15-bin.zip	sha512	pgp

AES256(Cont’d)

Encryption

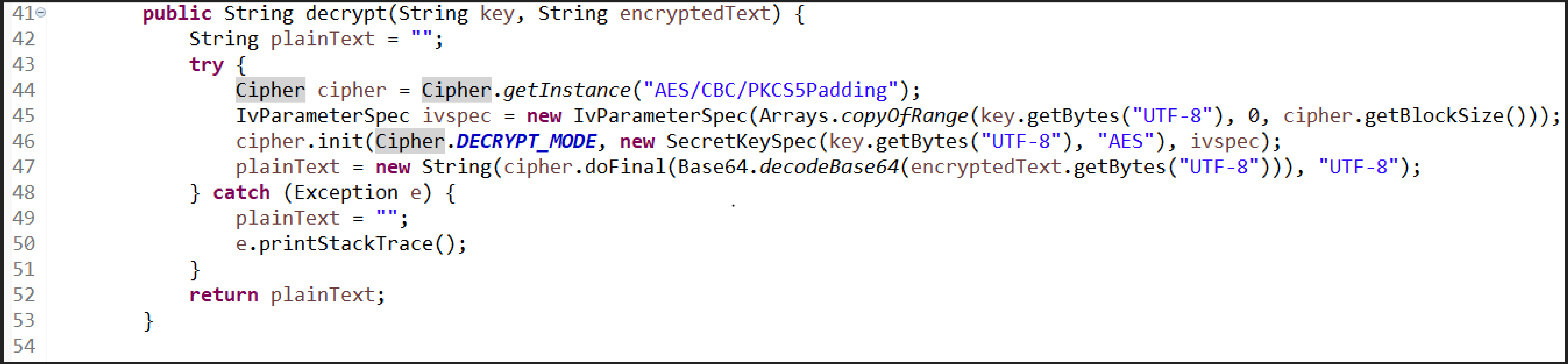


```
    public String encrypt(String key, String text) {
        String cipherText = "";
        try {
            Cipher cipher = Cipher.getInstance("AES/CBC/PKCS5Padding");
            IvParameterSpec ivspec = new IvParameterSpec(Arrays.copyOfRange(key.getBytes("UTF-8"), 0, cipher.getBlockSize()));
            cipher.init(Cipher.ENCRYPT_MODE, new SecretKeySpec(key.getBytes("UTF-8"), "AES"), ivspec);
            cipherText = new String(Base64.encodeBase64(cipher.doFinal(text.getBytes("UTF-8"))), "UTF-8");
        } catch (Exception e) {
            cipherText = "";
            e.printStackTrace();
        }
        return cipherText;
    }
```

P1 : Encrypt “Hello World”

AES256

Decryption



```
    public String decrypt(String key, String encryptedText) {
        String plainText = "";
        try {
            Cipher cipher = Cipher.getInstance("AES/CBC/PKCS5Padding");
            IvParameterSpec ivspec = new IvParameterSpec(Arrays.copyOfRange(key.getBytes("UTF-8"), 0, cipher.getBlockSize()));
            cipher.init(Cipher.DECRYPT_MODE, new SecretKeySpec(key.getBytes("UTF-8"), "AES"), ivspec);
            plainText = new String(cipher.doFinal(Base64.decodeBase64(encryptedText.getBytes("UTF-8"))), "UTF-8");
        } catch (Exception e) {
            plainText = "";
            e.printStackTrace();
        }
        return plainText;
    }

```

P2 : Decrypt the encrypted text of “Hello World”

P3

Decrypt Text

- The encoded text is **ruDZ3CTS5Md3+ipVKt20hQ==**

- Decrypt the text above

- Hint1, the key starts with “aaaaaaaaaaaaaaaaaaaaaaaaaaaaa” <- 29
  so, find out the last three characters

- Hint2, the last three characters contain numeric characters only

P4

Login Module

- Connect Java and MySQL with using JDBC

- The table scheme should contain the following columns (no – int, name – varchar, password – varchar)

- Insert some of login information in advance into the table

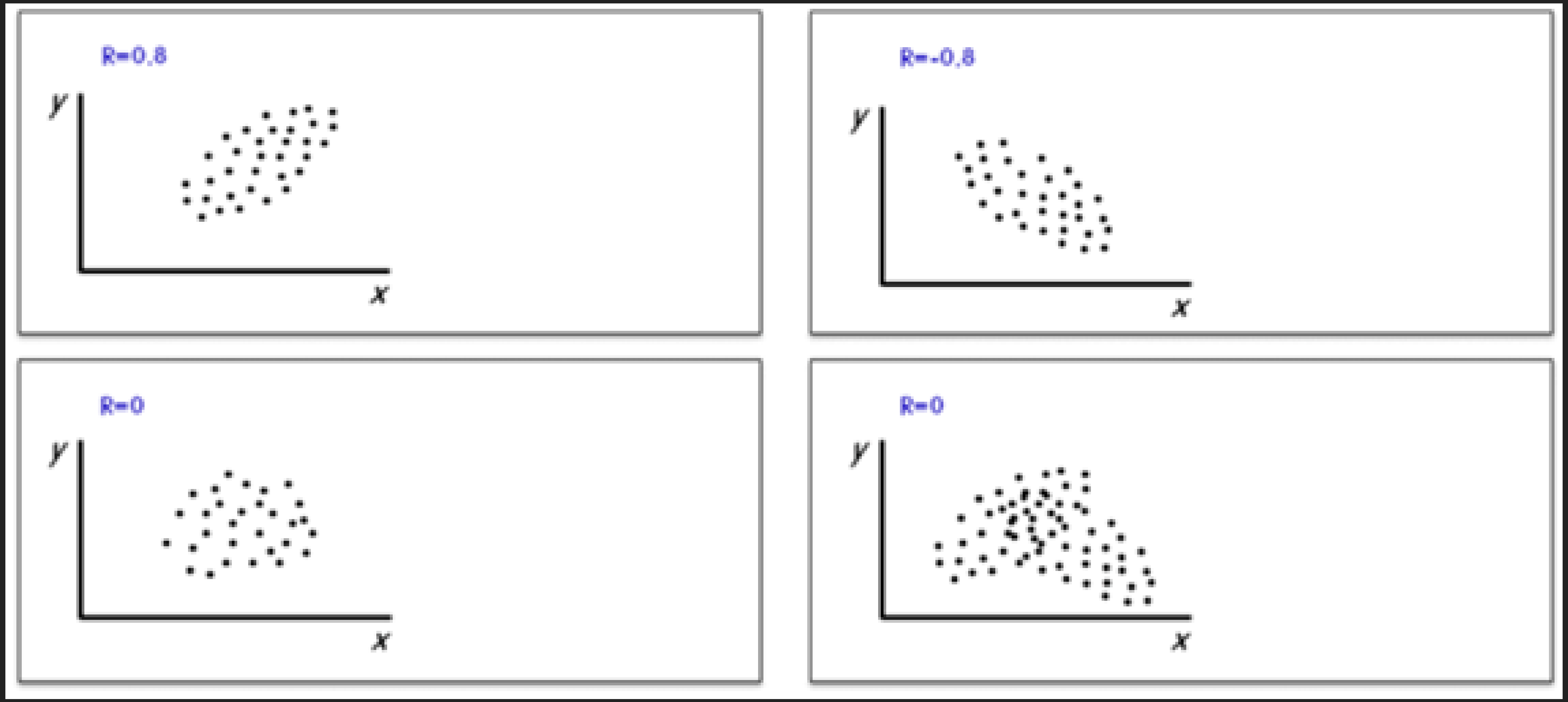
1. Input a name and a password from the console
2. Check if the names and the passwords are the same
3. The password from the console and the one from the table should be encrypted by AES256

Pearson Correlation Coefficient(Cont’d)

PCC(Pearson Correlation Coefficient)

- A measure of linear correlation between two sets of data

- The coefficient has a value between -1 and 1

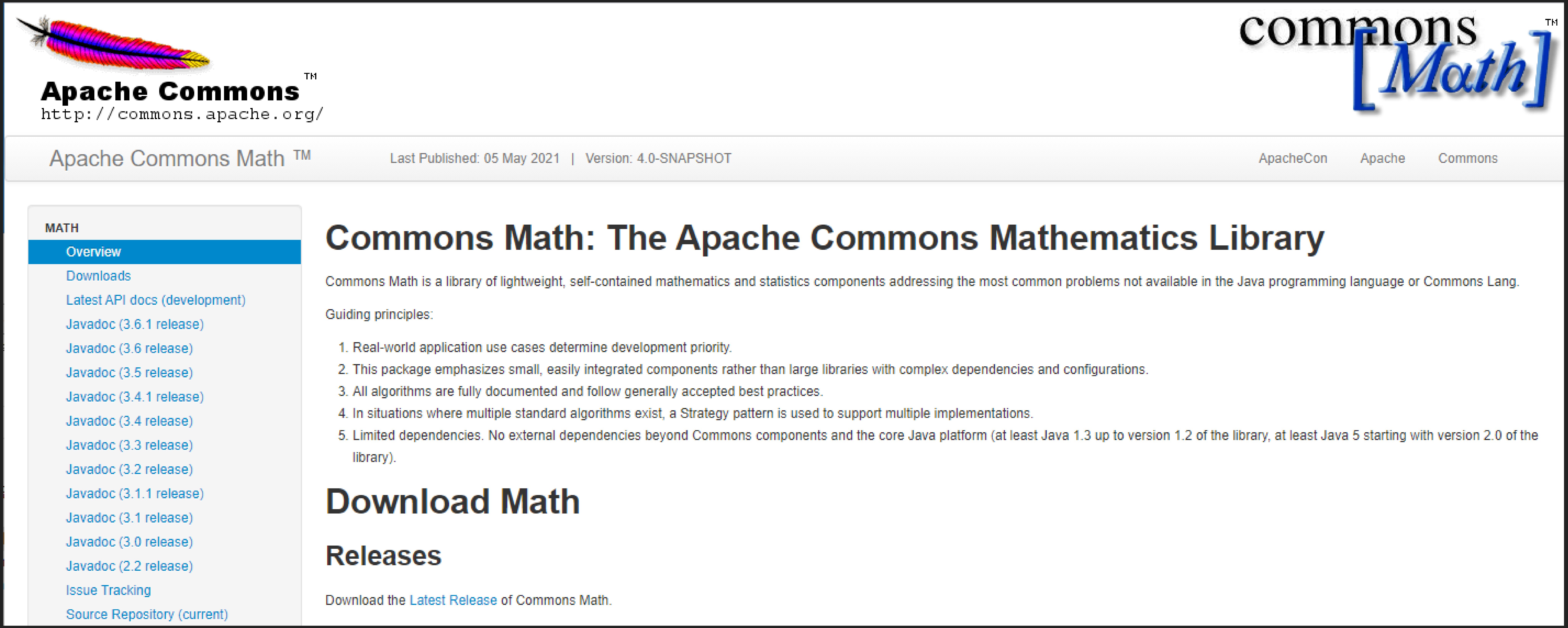


Source : https://en.wikipedia.org/wiki/Pearson_correlation_coefficient / https://www.researchgate.net/figure/Meaning-of-Pearson-correlation-coefficient-value-r_tbl1_299402589

Pearson Correlation Coefficient(Cont’d)

Download Apache Math Library and Add it to the Project

- https://commons.apache.org/proper/commons-math



Commons Math: The Apache Commons Mathematics Library

Commons Math is a library of lightweight, self-contained mathematics and statistics components addressing the most common problems not available in the Java programming language or Commons Lang.

Guiding principles:

1. Real-world application use cases determine development priority.
2. This package emphasizes small, easily integrated components rather than large libraries with complex dependencies and configurations.
3. All algorithms are fully documented and follow generally accepted best practices.
4. In situations where multiple standard algorithms exist, a Strategy pattern is used to support multiple implementations.
5. Limited dependencies. No external dependencies beyond Commons components and the core Java platform (at least Java 1.3 up to version 1.2 of the library, at least Java 5 starting with version 2.0 of the library).

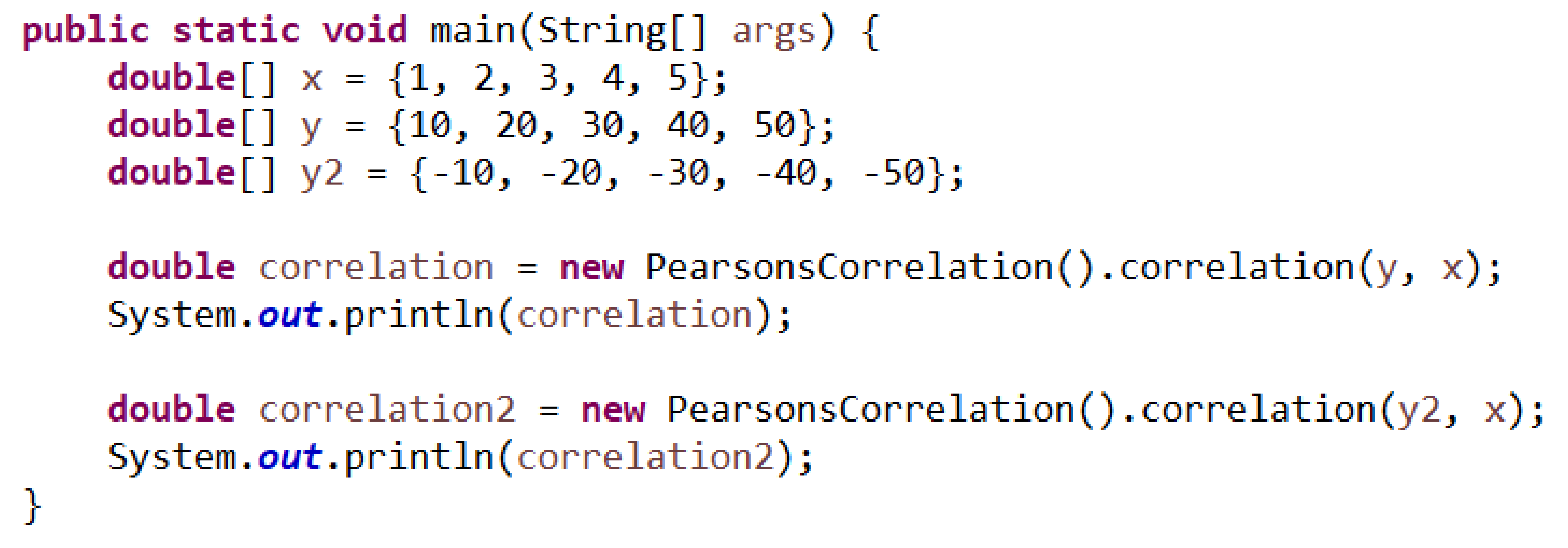
Download Math

Releases

Download the Latest Release of Commons Math.

Pearson Correlation Coefficient

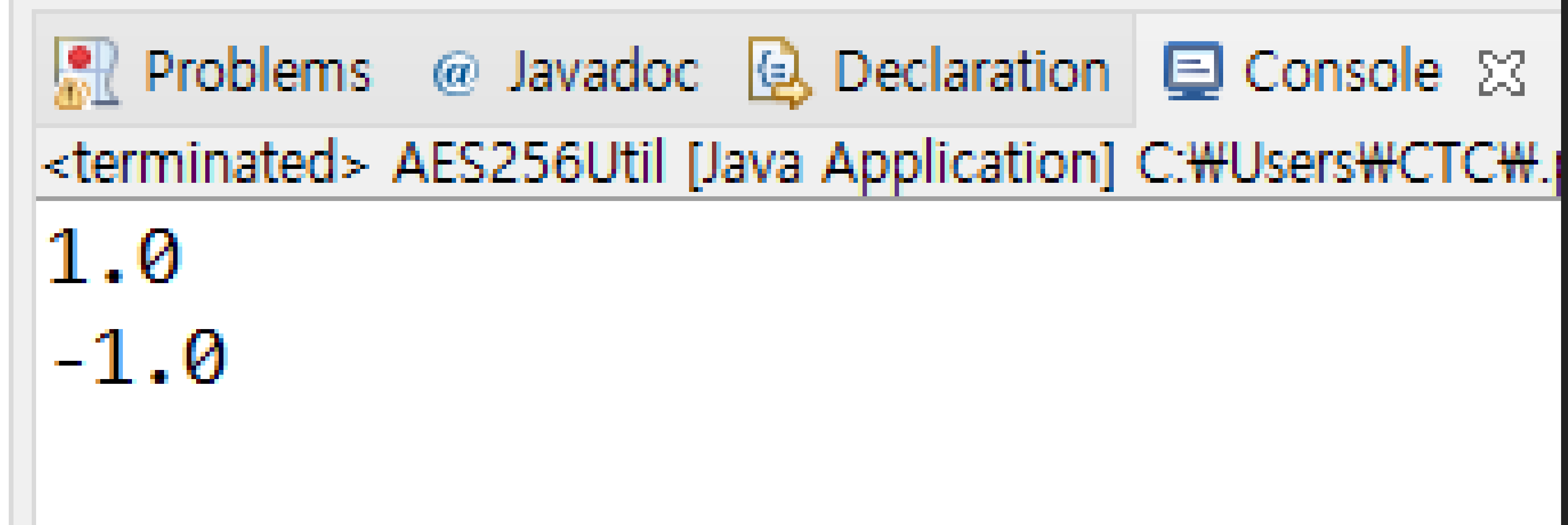
Get the value r



```
public static void main(String[] args) {
    double[] x = {1, 2, 3, 4, 5};
    double[] y = {10, 20, 30, 40, 50};
    double[] y2 = {-10, -20, -30, -40, -50};

    double correlation = new PearsonsCorrelation().correlation(y, x);
    System.out.println(correlation);

    double correlation2 = new PearsonsCorrelation().correlation(y2, x);
    System.out.println(correlation2);
}
```



```
<terminated> AES256Util [Java Application] C:₩Users₩CTC₩.
1.0
-1.0
```

P5

Practice for PCC

- Collect two sets of data and restore them in a CSV file
  e.g. population – housing price

- Sample size should be more than 100

- Get a PCC value between them

P6

Regression Analysis

- Execute the regression analysis in Excel
- The number of independent variables should be more than 5
- Explain the result