# Homework 1 Compression and encryption

Files are compressed using a variety of techniques. An image file is compressed differently   
than a folder containing text files.

1. Compression reduces the size of a file by removing unnecessary data.
2. Explain the difference between lossy and lossless compression: [3]

Lossy causes some of the data from the file to be permanently lost, which may result in a decrease in quality if the file is a video/image for example, or some words being lost if it is a text file. Lossless however doesn’t lose the data, the file before it was compressed is the exact same as after it is decompressed. This is because in lossless, an algorithm is used to store frequently used data as one so you aren’t manually storing every single piece of repeated data over and over, when you can just group them into one if you know they are repeated in that order all the time. This also makes lossless harder to do because it requires more complexity whereas lossy is far simpler.

1. Explain why Run-Length Encoding works well to compress a sound file without   
   losing any of the original data: [3]

It is fast and efficient because it compresses by grouping the data. For example, similar sounds will be stored as the same type of data so that when a sound sample is heard that has been repeated many times contiguously, all these samples will be grouped and stored as one instead of individual samples one after the other. This is more efficient since for example, the sound samples are all of the same frequency. Instead of storing it like sound1 sound1 sound1 sound1 … it groups it together like 4 sound1s, saving a lot of space.

1. Sound files can also be encoded using lossy compression techniques. Describe why these techniques reduce files more significantly than RLE but can have a detrimental effect on the quality of the file: [3]

Lossy causes permanent loss of data, if you are losing sound samples you reduce the data required to store those sound samples, and you will lose a lot because there will usually be plenty of irrelevant/negligible sounds or sounds of frequencies not able to be heard by human ears in a sound file, which all get removed freeing up a lot of space, but this decreases quality of the sound because you have deleted so many sound samples. Although some of the quieter or less prominent sounds are deleted, they could have had a major impact on the sound to feel smoother and better.

1. Encryption is used to obscure data from being understood without authorisation.   
   Files can be encrypted using various methods.

Two methods of encryption are the Caesar or Vernam cipher.

1. Explain why the Caesar cipher can be easily cracked. [1]

Algorithms can be used to easily crack it since Caesar cipher will always have a pattern and frequency analysis can be used on it to find repeated characters and identify the shift that has been used, once this is done the whole cipher falls apart

1. The Vernam cipher is proven to be unbreakable. Explain why. [2]

A one time key is used. The key is also truly random so no discernable pattern exists and techniques to break the cipher are therefore unusable since the right conditions aren’t present.

3. a) Explain the difference between symmetric and asymmetric encryption. [2]

b) Explain briefly how asymmetric encryption is used. [2]

4. a) Explain what is meant by **hashing**. [1]

b) Explain why passwords and PINs may be hashed, and how a user’s password   
can then be verified. [3]

[Total 20 Marks]