Avalanche Test

This document summarizes the analysis done on the Avalanche test. We study various statistical metrics as a function of Hash-Function parameters.

# **Coefficient of Variation (CoV):**

We use the coefficient of variation to study the relationship between the sample size and the standard deviation around the sample mean. Since the mean value for each experiment (simulation run) is different, we use the Coefficient of Variation to study this relationship.

You can learn more about the Coefficient of Variation [here](https://en.wikipedia.org/wiki/Coefficient_of_variation).

The Coefficient of Variation is calculated by dividing the sample standard deviation by the mean value. We study the CoV as function the following variables for both 512 and 128 Hash-Functions.

* **Sample size**: Here the sample size refers to the number of random strings used per run.
* **String length**: This refers to the length of the random string used in each run. String length is measured in the number characters.
* **Pre-Processing Parameters**: For both 512 and 128 Hash-Functions, we measure the effect of bitwise operation parameters in the Pre-Processing Function on the CoV value.
* **Compression Parameters**: For both 512 and 128 Hash-Functions, we measure the effect of bitwise operation parameters in the Compression Function on the CoV value.

# **Study Case 1:**

In this section, we look at the relationship between the sample size and string length of the CoV values for both 512 and 128 Hash-Functions. The following figures show the Coefficient of Variation as a function of sample size. As expected, the CoV decreases with increase in sample size. This indicates the tightening of the Standard Deviation around the mean.