**Similitud basada en correlación:** A veces es más natural o conveniente trabajar con una medida de similitud entre los registros y no de distancia, la cual mide antisimilitudes.

A popular similarity measure is the square of the Pearson correlation coefficient, r2ij , where the correlation coefficient is defined by

A black line with black text

Description automatically generated

**Distancia estadística:** Esta metrica tiene en cuenta la correlación entre medidas, con esto, las medidas que están altamente correlacionadas con otras no contribuyen tanto como las que no están correlacionadas o lo están a medias. La distancia estadística entre registros i y j está definida como:

A black and white text

Description automatically generated

A close up of text

Description automatically generated

**Distancia de Manhattan:**

A math equation with black text

Description automatically generated

**Máxima distancia coordinada:**

**A white background with black text

Description automatically generated**

**Medidas de distancia para datos categóricos**

Para medidas con registros binarios tiene más sentido utilizar medidas de similitud que medidas de distancia.

A white paper with black text

Description automatically generated

**Medidas de distancia para datos mezclados:**

A math formula with text

Description automatically generated with medium confidence

A math equations on a white background

Description automatically generated