**PSEUDOCÓDIGO PARA LA SELECCIÓN DE TÍTULOS DE SECCIONES.**

**INICIO**

**CREAR** DATAFRAME DE DATOS

**LIMPIAR** DATAFRAME {

QUITAR CARACTERES ESPECIALES

QUITAR ESPACIOS EN BLANCO DE MAS

QUITAR PALABRAS QUE APORTAN RUIDO AL TEXTO

DEJAR “: . \n”

DEJAR NÚMEROS ROMANOS

}

**BUSCAR TÍTULOS** CON **NÚMEROS ROMANOS {**

EXPRESIÓN REGULAR: **r'(I{1,3}|IV|V|VI{1,3}|IX|X)+[-.)][\w\s-]+[:\n]'**

**IF** COMIENZA CON NUMERO ROMANO(I,II,III,IV,V,VI,VII,VIII,IX,X) **THEN**

**IF** SIGUE UN “-” **OR** “.” **OR** “)” **THEN**

**IF** SIGUE UNA O VARIAS PALABRAS O “-“ **THEN**

**IF** FINALIZA CON “:” **OR** “\n” **OR** “.” **THEN**

**GUARDAR TITULO CON Nro. ROMANO**

}

**BUSCAR TÍTULOS** EN **MAYÚSCULAS** **{**

EXPRESIÓN REGULAR:  **r'(I{1,3}|IV|V|VI{1,3}|IX|X)[-.)][A-Z -]+[:.\n]'**

**IF** COMIENZA CON NUMERO ROMANO(I,II,III,IV,V,VI,VII,VIII,IX,X) **THEN**

**IF** SIGUE UN “-” **OR** “.” **OR** “)” **THEN**

**IF** SIGUE UNA O VARIAS PALABRAS EN MAYÚSCULAS O “-“ **THEN**

**IF** FINALIZA CON “:” **OR** “\n” **OR** “.” **THEN**

**GUARDAR TITULO EN MAYÚSCULA**

**}**

**FIN**