

ANEXOS

LICENCIA DE USO - AUTORIZACIÓN DE LOS AUTORES

Actuando en nombre propio identificado (s) de la siguiente forma:

Nombre Completo Wilson Vargas Martinez

Tipo de documento de identidad: C.C. T.I. C.E. Número: 1031999935

Nombre Completo Maslon Felipe Rodriguez Santos

Tipo de documento de identidad: C.C. T.I. C.E. Número: 1016085824

Nombre Completo Oscar Samuel Salgado Nope

Tipo de documento de identidad: C.C. T.I. C.E. Número: 103237160

Nombre Completo _____

Tipo de documento de identidad: C.C. T.I. C.E. Número: _____

El (Los) suscritor(s) en calidad de autor (es) del trabajo de tesis, monografía o trabajo de grado, documento de investigación, denominado:

Proyecto de transformación del modelo de gestión integral de procesos y datos hacia tecnologías emergentes

Dejo (dejamos) constancia que la obra contiene información confidencial, secreta o similar: SI NO
(Si marqué (marcamos) SI, en un documento adjunto explicaremos tal condición, para que la Universidad EAN mantenga restricción de acceso sobre la obra)

Por medio del presente escrito autorizo (autorizamos) a la Universidad EAN, a los usuarios de la Biblioteca de la Universidad EAN y a los usuarios de bases de datos y sitios webs con los cuales la Institución tenga convenio, a ejercer las siguientes atribuciones sobre la obra anteriormente mencionada:

- A. Conservación de los ejemplares en la Biblioteca de la Universidad EAN.
- B. Comunicación pública de la obra por cualquier medio, incluyendo Internet.
- C. Reproducción bajo cualquier formato que se conozca actualmente o que se conozca en el futuro.
- D. Que los ejemplares sean consultados en medio electrónico.
- E. Inclusión en bases de datos o redes o sitios web con los cuales la Universidad EAN tenga convenio con las mismas facultades y limitaciones que se expresan en este documento.
- F. Distribución y consulta de la obra a las entidades con las cuales la Universidad EAN tenga convenio.

Con el debido respeto de los derechos patrimoniales y morales de la obra, la presente licencia se otorga a título gratuito, de conformidad con la normalidad vigente en la materia y teniendo en cuenta que la Universidad EAN busca difundir y promover la formación académica, la enseñanza y el espíritu investigativo y emprendedor.

Manifiesto (manifestamos) que la obra objeto de la presente autorización es original, el (los) suscritos es (son) el (los) autor (es) exclusivo (s), fue producto de mi (nuestro) ingenio y esfuerzo personal y la realizo (zamos) sin violar o usurpar derechos de autor de terceros, por lo tanto la obra es de exclusiva autoría y tengo (tenemos) la titularidad sobre la misma. En vista de lo expuesto, asumo (asumimos) la total responsabilidad sobre la elaboración, presentación y contenidos de la obra, eximiendo de cualquier responsabilidad a la Universidad EAN por estos aspectos.

En constancia suscribimos el presente documento en la ciudad de Bogotá D.C.,

NOMBRE COMPLETO: Wilson Vargas Martínez
FIRMA: Wilson Vargas
DOCUMENTO DE IDENTIDAD: 1022974935
FACULTAD: Ingeniería
PROGRAMA ACADÉMICO: Especialización en Gerencia de Proyectos

NOMBRE COMPLETO: Maximiliano Rodríguez Santos
FIRMA: Maximiliano Rodríguez Santos
DOCUMENTO DE IDENTIDAD: 1016028824
FACULTAD: Ingeniería
PROGRAMA ACADÉMICO: Especialización en Gerencia de Proyectos

NOMBRE COMPLETO: Oscar Samuel Salguero Noya
FIRMA: Oscar S
DOCUMENTO DE IDENTIDAD: 1032397160
FACULTAD: Ingeniería
PROGRAMA ACADÉMICO: Especialización en Gerencia de Proyectos

NOMBRE COMPLETO: _____
FIRMA: _____
DOCUMENTO DE IDENTIDAD: _____
FACULTAD: _____
PROGRAMA ACADÉMICO: _____

Fecha de firma: 06/12/2019

Para efectos de trabajos documentales, investigaciones futuras y demás archivos construidos por parte de la Universidad EAN, a continuación, se exponen la codificación que en los programas tanto R como Python son referencia para la solución el problema de recolección de datos en documentos Excel, bases sueltas, o datos en bruto.

CÓDIGO R

```
library(readxl)
Base_Principal <- read_excel("Análisis aportes seminario/Base Principal.xlsx")
View(Base_Principal)
```

```
library(readxl)
Basecartera_Secundaria <- read_excel("Análisis aportes seminario/Basecartera Secundaria.xlsx")
View(Basecartera_Secundaria)
```

```

# Conversion matrices
DatosCOP<-as.data.frame(Base_Principal)
DatosCOPSEC<-as.data.frame(Basecartera_Secundaria)
#Procesado de colocaciones

DatosCOP$`Total Colocaciones`<-as.character(DatosCOP$`Total Colocaciones`)
DatosCOP$`Total Colocaciones`<-gsub(",", ".", DatosCOP$`Total Colocaciones`)
DatosCOP$`Total Colocaciones`<-gsub("-1",NA, DatosCOP$`Total Colocaciones`)
DatosCOP$`Total Colocaciones`<-as.numeric(DatosCOP$`Total Colocaciones`)
Colmedio<-mean(DatosCOP$`Total Colocaciones`, na.rm=TRUE)
Colmedio
DatosCOP$`Total Colocaciones`[is.na(DatosCOP$`Total Colocaciones`)]<-Colmedio
hist(DatosCOP$`Total Colocaciones`)
boxplot(DatosCOP$`Total Colocaciones`)
summary(DatosCOP$`Total Colocaciones`)

#Procesado de captaciones
DatosCOP$`Total Captaciones`<-as.character(DatosCOP$`Total Captaciones`)
DatosCOP$`Total Captaciones`<-gsub(",", ".", DatosCOP$`Total Captaciones`)
DatosCOP$`Total Captaciones`<-gsub("-1",NA, DatosCOP$`Total Captaciones`)
DatosCOP$`Total Captaciones`<-as.numeric(DatosCOP$`Total Captaciones`)
Capmedio<-mean(DatosCOP$`Total Captaciones`, na.rm=TRUE)
Capmedio
DatosCOP$`Total Captaciones`[is.na(DatosCOP$`Total Captaciones`)]<-Capmedio
hist(DatosCOP$`Total Captaciones`)
boxplot(DatosCOP$`Total Captaciones`)
summary(DatosCOP$`Total Captaciones`)

#Procesado de mora aportes
DatosCOP$`Mora Aportes`<-as.character(DatosCOP$`Mora Aportes`)
DatosCOP$`Mora Aportes`<-gsub(",", ".", DatosCOP$`Mora Aportes`)
DatosCOP$`Mora Aportes`<-gsub("-1",NA, DatosCOP$`Mora Aportes`)
DatosCOP$`Mora Aportes`<-as.numeric(DatosCOP$`Mora Aportes`)
MorAPmedio<-mean(DatosCOP$`Mora Aportes`, na.rm=TRUE)
MorAPmedio
DatosCOP$`Mora Aportes`[is.na(DatosCOP$`Mora Aportes`)]<-MorAPmedio
hist(DatosCOP$`Mora Aportes`)
boxplot(DatosCOP$`Mora Aportes`)
summary(DatosCOP$`Mora Aportes`)

#Procesado de mora cartera
DatosCOP$`Mora Cartera`<-as.character(DatosCOP$`Mora Cartera`)

```

```

DatosCOP$`Mora Cartera`<-gsub(",", ".", DatosCOP$`Mora Cartera`)
DatosCOP$`Mora Cartera`<-gsub("-1",NA, DatosCOP$`Mora Cartera`)
DatosCOP$`Mora Cartera`<-as.numeric(DatosCOP$`Mora Cartera`)
MorCarmedio<-mean(DatosCOP$`Mora Cartera`, na.rm=TRUE)
MorCarmedio
DatosCOP$`Mora Cartera`[is.na(DatosCOP$`Mora Cartera`)]<-MorCarmedio
hist(DatosCOP$`Mora Cartera`)
boxplot(DatosCOP$`Mora Cartera`)
summary(DatosCOP$`Mora Cartera`)

```

#Procesado de edad

```

DatosCOP$`Edad`<-as.character(DatosCOP$`Edad`)
DatosCOP$`Edad`<-gsub(",", ".", DatosCOP$`Edad`)
DatosCOP$`Edad`<-gsub("-1",NA, DatosCOP$`Edad`)
DatosCOP$`Edad`<-as.numeric(DatosCOP$`Edad`)
Edmedio<-mean(DatosCOP$`Edad`, na.rm=TRUE)
Edmedio
DatosCOP$`Edad`[is.na(DatosCOP$`Edad`)]<-Edmedio
hist(DatosCOP$`Edad`)
boxplot(DatosCOP$`Edad`)
summary(DatosCOP$`Edad`)

```

#Procesado de saldo

```

DatosCOPSEC$`Saldo`<-as.character(DatosCOPSEC$`Saldo`)
DatosCOPSEC$`Saldo`<-gsub(",", ".", DatosCOPSEC$`Saldo`)
DatosCOPSEC$`Saldo`<-gsub("-1",NA, DatosCOPSEC$`Saldo`)
DatosCOPSEC$`Saldo`<-as.numeric(DatosCOPSEC$`Saldo`)
Salmedio<-mean(DatosCOPSEC$`Saldo`, na.rm=TRUE)
Salmedio
DatosCOPSEC$`Saldo`[is.na(DatosCOPSEC$`Saldo`)]<-Salmedio
hist(DatosCOPSEC$`Saldo`)
boxplot(DatosCOPSEC$`Saldo`)
summary(DatosCOPSEC$`Saldo`)

```

#Procesado de monto desembolsado

```

DatosCOPSEC$`MontDesembol`<-as.character(DatosCOPSEC$`MontDesembol`)
DatosCOPSEC$`MontDesembol`<-gsub(",", ".", DatosCOPSEC$`MontDesembol`)
DatosCOPSEC$`MontDesembol`<-gsub("-1",NA, DatosCOPSEC$`MontDesembol`)
DatosCOPSEC$`MontDesembol`<-as.numeric(DatosCOPSEC$`MontDesembol`)
Salmedio<-mean(DatosCOPSEC$`MontDesembol`, na.rm=TRUE)
Salmedio
DatosCOPSEC$`MontDesembol`[is.na(DatosCOPSEC$`MontDesembol`)]<-Salmedio
hist(DatosCOPSEC$`MontDesembol`)
boxplot(DatosCOPSEC$`MontDesembol`)

```

```
summary(DatosCOPSEC$`MontDesembol`)
```

```
DatosCOPjoin<-merge(DatosCOP,DatosCOPSEC, by="Cedula",all.x = TRUE)
```

```
DatosCOPjoin
```

```
f <- filter(DatosCOPjoin,Ciudad == "BARRANQUILLA")
```

```
Comparativo <- select.list(f,Edad) %>% unlist
```

```
mean(Comparativo)
```

```
sd(Comparativo)
```

```
View(Comparativo)
```

```
hist(Comparativo)
```

CÓDIGO PYTHON

```
#!/usr/bin/env python
```

```
# coding: utf-8
```

```
# ## Analisis Datos con PYTHON
```

```
#Se prepara el entorno de trabajo
```

```
get_ipython().run_line_magic('matplotlib', 'inline')
```

```
import matplotlib
```

```
import numpy as np
```

```
import matplotlib.pyplot as plt
```

```
import seaborn as sns
```

```
import pandas as pd
```

```
excel_file1 = (r"C:\Users\wilso\Documents\Analisis aportes seminario/Base Principal.xlsx")
```

```
Principal = pd.read_excel(excel_file1,parse_dates=[0])
```

```
excel_file2 = (r"C:\Users\wilso\Documents\Analisis aportes seminario/Basecartera  
Secundaria.xlsx")
```

```
Secundaria = pd.read_excel(excel_file2,parse_dates=[0])
```

```
Principal.head()
```

```
#Cantidad de filas y columnas.
```

```
Principal.shape
```

```
#Descripción en general de las variables
```

```
Principal.info()
```

```
Secundaria.head()
```

```
#Cantidad de filas y columnas.
```

```
Secundaria.shape
```

```
#Descripción en general de las variables
```

```
Secundaria.info()
```

```
#CRUZAR TABLAS
```

```
merged_left = pd.merge(left=Principal,right=Secundaria, how='left', left_on='Cedula',  
right_on='Cedula')
```

```
merged_left.info()
```

```
merged_left.head()
```

```
Grupo_1=merged_left.groupby('Fecha Afiliacion')
```

```
Grupo_1.fillna(value=0)
```

```
Grupo_1['Saldo'].agg([np.size,sum,max,min]).dropna(how='any')
```

```
Grupo_1['MontDesembol'].agg([np.size,sum,max,min]).dropna(how='any')
```

```
Grupo_1['Saldo'].agg([max,min]).sort_values(by='Fecha Afiliacion',ascending=True)[-  
10:].plot(kind='bar')
```

```
Grupo_1['MontDesembol'].agg([max,min]).sort_values(by='Fecha  
Afiliacion',ascending=True)[-10:].plot(kind='bar')
```

```
#Seleccionar campos para trabajar con muestras mas pequeñas
```

```
COP1=Principal[['Fecha Afiliacion','Ciudad','Total Captaciones','Total Colocaciones']]
```

```
COP2=Secundaria[['FAper','Ofi','MontDesembol','Saldo']]
```

```
#La cantidad de registros para cada fecha
```

```
Grupo_3.size().sort_values(ascending=False)
```

```
Grupo_3.sum().sort_values(by='Fecha Afiliacion').dropna(how='any')
```