
Software “R” como herramienta en la enseñanza de la estadística descriptiva

The software "R" as tool in the teaching of the Descriptive Statistic

João Kumbo

Sección de Matemáticas del Instituto de Ciencias de la Educación de Uíge, Angola.

ORCID: <http://orcid.org/0000-0003-3804-4159>

Correo electrónico:

ndonga28@hotmail.com

Recibido: 05/02/2021

Aceptado: 18/07/2021

Resumen: Se abordan las implicaciones que un ambiente computacional puede proporcionar al proceso de enseñanza-aprendizaje de la Estadística en los cursos de nivel superior. Se realizó una investigación descriptiva de naturaleza cuantitativa, en los cursos de licenciatura en Enseñanza de Ciencias, en la unidad curricular Estadística Aplicada a la Educación. Se introdujo la computación a través del Software “R”. Los resultados mostraron que el proceso de enseñanza-aprendizaje de la Estadística por medio computacional favorece el tratamiento didáctico-pedagógico, de los contenidos.

Palabras clave: Enseñanza Superior; Estadística descriptiva; Tecnologías; Ciencias de la Educación.

Abstract: They discuss the implications that a computational environment can provide to the process of teaching learning of the Statistics in the courses of superior level. Curricular accomplished a descriptive research of quantitative nature, in the courses of bachelor's degree in Enseñanza of Ciencias, in the unit itself Applied Statistics Education. the r introduced the computation through the Software itself. The results evidenced than the process of teaching the didactic pedagogic, contentses treatment favors learning of the Statistics for computational means.

Keywords: Higher education; Descriptive statistics; Technologies; Educational Sciences.

Introducción

En la actualidad casi todas las esferas de la actividad humana están influenciadas por la Estadística. Todos los días tanto diarios, revistas, televisión, e internet presentan datos estimados, gráficos, tablas, proyecciones y resultados de expedientes, aproximando así la Estadística a la vida cotidiana. (Tomás, 2006)

Según João, “la Estadística Descriptiva consiste en la recolección, presentación, análisis e interpretación de datos numéricos a través de la creación de instrumentos adecuados: cuadros, gráficos e indicadores numéricos”. (2017, p. 23)

Para este autor la Estadística se encuentra actualmente integrada con varios niveles de

profundidad en los planos curriculares de las diferentes instituciones de Enseñanza Superior (João, 2017). Sin embargo, en el Instituto Superior de Ciencias de la Educación (ISCED), Uíge, es semestral y se encuadra en el 2º año de licenciatura en todos los cursos. En esta integración, y teniendo en cuenta el mundo actual, las herramientas computacionales, como parte de las TIC, tiene un rol decisivo en el proceso de enseñanza-aprendizaje de todas las disciplinas del currículo, particularmente de la Estadística Descriptiva. Sin embargo, el proceso de enseñanza-aprendizaje de la Estadística en diversos cursos de nivel superior se encuentra desvinculado de esta realidad.

Como la educación actual las futuras generaciones no pueden estar al margen de la sociedad, cada vez más, el proceso de enseñanza-aprendizaje de la Estadística en la escuela es foco de docentes universitarios e investigadores de esta área en lo que se refiere las actuales propuestas didáctico-metodológicas a ser utilizadas en la sala de clase a modo de desarrollar en los alumnos la capacidad y comprensión de análisis de datos, la resolución de problemas y casos de estudios. Todas estas propuestas están apoyadas en las nuevas tecnologías de la información y las comunicaciones (TIC), donde las herramientas computacionales tienen un rol de considerable significado en la educación, y para lo cual el docente necesita de preparación especializada para su aplicación acertada.

No obstante, la enseñanza de la Estadística Descriptiva en el Instituto Superior de Ciencias de la Educación ISCED/Uíge es formal, limitado solo a reproducir fórmulas matemáticas y construir gráficos de forma manual, sin utilizar herramientas computacionales. Esta realidad constituye una preocupación para los alumnos, principalmente para los cursos de Enseñanza de Ciencias Sociales y Psicopedagógicas, pues muchos de ellos, provienen de cursos de Ciencias Sociales y Humanas en la Enseñanza Secundaria; para éstos la Matemática no es profundizada, pues dejan de entrar en contacto con ella muy temprano. Y cuando se encuentran con la Estadística en los niveles subsiguientes, la preocupación es mayor. Es para este componente de la formación que el presente estudio va dirigido, proponiendo el Software “R” como herramienta en la enseñanza-aprendizaje de la Estadística.

Uno de los softwares con más resultados en el proceso de enseñanza-aprendizaje de la Estadística es la “R”, por su lenguaje de programación simple y estar disponible de forma

gratuita en internet -consultar sitio: (<https://www.r-project.org/>) -. La aplicación de este software, además de la motivación que genera su uso, disminuye el tiempo dedicado a resolver un determinado ejercicio, y agota menos a la persona al no tener que realizar cálculos aburridos y utilizar instrumentos de dibujo para construir tablas, cuadros y/o gráficos.

A pesar de su carácter gratuito, la “R” es una herramienta poderosa, con buenas capacidades a nivel de programación y contiene un conjunto amplio - en constante crecimiento - de paquetes que añaden bastante potencialidad a la ya poderosa versión base de la “R”. (Pereira y Piario, 2012)

Estos autores definen el software “R” como: “ un conjunto integrado de herramientas computacionales que permiten la manipulación, análisis de datos, el cálculo numérico y la producción de gráficos” (Pereira y Piario, 2012, p.14).

Y por otra parte plantea (João, 2017), que el referido software

Trae nuevas formas de explorar la Estadística Descriptiva, proporciona mayor rapidez en la resolución de problemas y permite la comparación activa de soluciones. Además, abre camino a un conjunto de usuarios en los medios académicos, empresarial y administrativo en la utilización de la estadística como herramienta eficaz en la respuesta a sus problemas. (p. 12)

Según Carvalho (2015), la Informática desde que surgió, está en alianza con la Estadística, ofreciendo herramientas que realizan cálculos lentos y complejos y facilita la construcción de escenarios alternativos para la mejor comprensión de los datos y de las estructuras que se asocian.

La creciente necesidad de convertir datos en información comprensible para quien la quiera, ha traído consigo una mayor utilización de la Estadística en los más variados dominios, exigiendo un buen conocimiento de sus métodos y técnicas (Figueredo, Figueredo, Ramos, y Paulo, 2017).

En el presente estudio, decidimos utilizar el Software “R” para el proceso de enseñanza-aprendizaje de la Estadística Descriptiva, como manifestación del impacto de las TIC en el contexto de la Enseñanza Superior, sin alterar la forma de trabajar en las salas de clases, y

recapitular el tratamiento de los componentes didácticos. El problema científico fue formulado como: ¿Qué implicaciones puede proporcionar un ambiente computacional, para el proceso de enseñanza-aprendizaje de la Estadística Descriptiva, en los cursos de nivel superior? Con esta investigación se pretende: Proponer un estudio estadístico por medio de un ambiente computacional, investigando y discutiendo las implicaciones que éste puede proporcionar para el proceso de enseñanza-aprendizaje de la Estadística Descriptiva en los cursos de nivel superior.

Para alcanzar este objetivo general, fue necesario alcanzar los siguientes objetivos parciales: diagnosticar el estado actual de desarrollo de habilidades informáticas generales y específicas de los docentes; elevar el nivel de preparación de los docentes en el uso del Software “R”, en las clases de Estadística; evaluar el desempeño de los docentes en las salas de clases después de la realización del seminario metodológico y finalmente, evaluar el estado final del desarrollo de las habilidades de los docentes después de la intervención en la práctica educativa.

La realización y concreción de los objetivos fue posible gracias a la utilización de los métodos de investigación científica de los niveles teórico (histórico-lógico, análisis-síntesis, inductivo-deductivo, modelación), empírico (observación, expediente, tests, revisión bibliográfica), y estadístico-matemático (técnicas propias de la Estadística Descriptiva, construcciones de tablas y/o gráficos). Es una investigación descriptiva de naturaleza cuantitativa, y la recogida de datos fue hecha a través de la observación participante - registros de clases, fotografías- y actividades realizadas por los alumnos.

Desarrollo

Fundamentación y contexto

El estudio fue realizado en el Instituto Superior de Ciencias de la Educación ISCED/Uíge; en la unidad curricular Estadística aplicada a la Educación. El estudio se concentra en tres momentos fundamentales, que son: determinación del estado inicial de la temática objeto de estudio, intervención en la práctica educativa del proceso de enseñanza-aprendizaje de la Estadística Descriptiva y determinación del estado final de la temática después de la referida intervención.

Determinación del estado inicial de la temática objeto de estudio

En esta fase, se realizaron tres actividades fundamentales, la primera fue la organización de bibliografías sobre la temática objeto de estudio; esta actividad permitió fundamentar teóricamente los conceptos utilizados en el estudio, teniendo como base la revisión bibliográfica. La segunda actividad es la observación de clases a los docentes que participaron en la investigación; sin embargo, se observó las clases y se constató que ninguno de ellos utilizaba herramientas computacionales en la enseñanza de la Estadística Descriptiva, éste era formal, la pizarra y los cuadernos estaban llenos de fórmulas, tablas y gráficos hechos manualmente.

Los aportes teóricos presentados, motivaron la organización de la tercera actividad desde el primer momento que es la formulación y aplicación de un cuestionario sobre las implicaciones que un ambiente computacional puede proporcionar para la enseñanza de la Estadística, con objetivos principales de: diagnosticar el estado actual de desarrollo de habilidades de los docentes, en cuanto la utilización de medios computacionales de forma general (manejar accesorios, navegar por el sistema operativo) o específicamente - interactuar con los softwares educativos en las construcciones de tablas y/o gráficos-; analizar las concepción de los docentes en cuanto a la incorporación de estas herramientas en el proceso de enseñanza-aprendizaje de la Estadística.

Intervención en la práctica educativa de la estadística descriptiva

En esta fase se realizaron dos actividades investigativas. La primera de es el desarrollo del seminario metodológico sobre el uso del Software “R”, en las clases de Estadística. Tuvo como objetivo, elevar el nivel de preparación de los docentes en el uso del Software “R”, en

las clases. Fueron realizadas tres sesiones de formación, durante tres días. Las sesiones abordaron los siguientes temas:

La primera sesión, trató la introducción al Software “R”. En esta sesión se introdujo las nociones básicas del programa desde la instalación hasta al modo de uso. La segunda sesión trató las construcciones de tablas de distribución de frecuencias y gráficos, haciendo un estudio comparativo entre las construcciones de tablas y gráficos hechos manualmente y por medio del Software “R”. La tercera sesión abordó, medidas de localización, de dispersión y de asimetría. Se planteó que para construir un código en “R”, con varias instrucciones, es conveniente abrir una ventana (New script) donde serán editados y guardados todas las instrucciones útiles a la resolución del problema que se trate. En ciertas computadoras, las instrucciones de la ventana pueden ser enviadas para la consola R usando F5 o Ctrl + R. En los ejemplos se presentan algunas de las actividades ofrecidas en las sesiones:

Ejemplo 1: Con el objetivo de divulgar uno de sus productos, una industria entrevistó 600 personas, para saber que vehículo de información, es más utilizado por ellas. Dentro de los entrevistados, 72 prefirieron Diario, 276 Radio, 42 Revista y 210 Televisión, construya: una tabla de distribución de frecuencias, gráficos de barras y de sectores correspondientes a esta información. (ver Tabla 1, gráfico 1 y gráfico 2).

Tabla 1- Resultados sobre la utilización de los vehículos de información

Veiculos	fi	fr	fac	frac	frap
Jornal	72	0.12	72	0.12	12
Rádio	276	0.46	348	0.58	58
Revista	42	0.07	390	0.65	65
Televisão	210	0.35	600	1	100

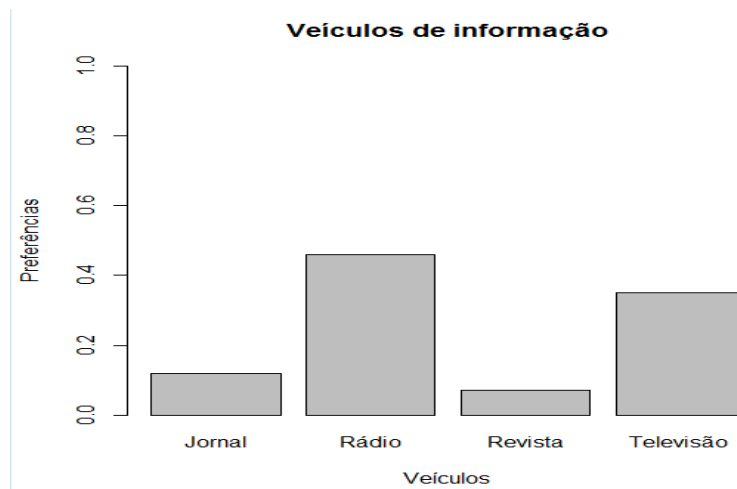


Gráfico 1- Representación por barras

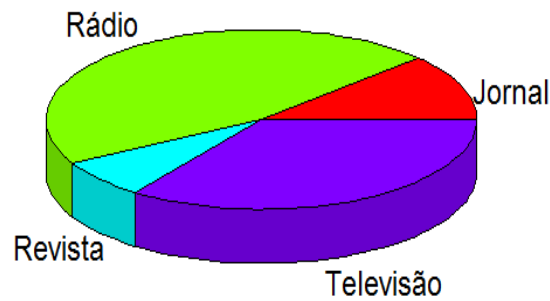


Gráfico 2- Representación por sectores

Tabla 2- Comando de "R" utilizados en las construcciones

```
#Construção da tabela

veiculos=c("Jornal", "Rádio", "Revista", "Televisão")

fi=c(72, 276, 42, 210)

cbind(veiculos, fi, fi=fi/sum(fi), fac=cumsum(fi), fap=cumsum(fi/sum(fi)*100))

#Representação gráfica

veiculos<- c(rep(1, 72),rep(2, 276),rep(3, 42),rep(4, 210))

veiculos.nom<- c("Jornal","Rádio","Revista","Televisão")

barplot(table(veiculos)/sum(table(veiculos)), ylim=c(0,1), names=veiculos.nom, xlab="veiculos",
ylab="Preferências", main="veiculos de informação")

pie3D(table(veiculos)/sum(table(veiculos)), labels=veiculos.nom, explode=0.0). Carregando o pacote
"plotrix".
```

Ejemplo 2: Considere los datos obtenidos por las medidas de las alturas de 208 individuos:

```
151 152 154 155 158 159 159 160 161 161 170 170 171 166 166 166
167 167 167 167 167 168 168 171 171 171 161 162 163 163 163 164
165 165 165 166 172 172 172 169 169 169 169 169 170 170 170 170
170 173 173 174 168 168 168 168 168 168 168 168 169 169 174 174
175 182 182 183 184 185 186 187 188 190 190 175 175 175 170 170
171 171 171 171 172 172 172 173 176 176 176 173 173 174 174 174
175 175 175 175 176 177 177 177 176 176 176 177 177 177 177 178
178 178 178 178 177 179 179 180 180 180 180 181 181 181 182 180
179 179 151 152 154 155 158 159 159 160 161 161 180 180 180 166
166 166 167 167 167 167 167 168 181 181 181 161 162 163 163 163
```

164 165 165 165 166 181 181 181 169 169 169 169 169 170 170 170
 170 170 181 181 181 168 168 168 168 168 168 168 168 169 169 176
 178 182 182 182 183 184 185 186 187 188 190 190 176 178 182 168

a) Presentar los datos en una tabla de distribución de frecuencias;

b) Calcular: media, mediana, moda, o desviación media, varianza y patrón de desvío.

Tabla 3- Resultados de las medidas de las alturas

Classlimits	f	rf	rf(%)	cf	cf(%)
[151, 156)	8	0.04	3.85	8	3.85
[156, 161)	8	0.04	3.85	16	7.69
[161, 166)	22	0.11	10.58	38	18.27
[166, 171)	66	0.32	31.73	104	50.00
[171, 176)	33	0.16	15.87	137	65.87
[176, 181)	36	0.17	17.31	173	83.17
[181, 186)	25	0.12	12.02	198	95.19
[186, 191)	10	0.05	4.81	208	100.00

Fuente: Elaboración propia

Tabla 4- Comandos de "R" utilizados en la construcción y cálculos de medidas de localización y de dispersión

```
#Carregar o pacote "fith" (frequency, distribution table histogram and polygon)  
alturas = c(151, 152, 154, 155, ..., 182)  
fht(alturas, start=151, end=191, h=5)  
  
# Cálculo das medidas de tendência central e de dispersão:  
  
Altura média é dada pelo comando: mean(alturas).  
  
Altura mediana é dada pelo comando: median(alturas).  
  
A moda é dada pelo comando: mfv(alturas), carregando o pacote "modeest".  
  
O desviomédio é dado pelo comando: aad(alturas), carregando o pacote "lsr".  
  
A variância é dada pelo comando: var(alturas)  
  
O desviopadrão é dado pelo comando: sd(alturas).
```

Fuente: Elaboración propia

La segunda actividad de este momento, es la organización de grupos de trabajo con el propósito de planificar clases utilizando el Software “R”. Se formaron cuatro grupos de trabajo, designados por A, B, C y D, con dos elementos cada uno; cuyos objetivos principales fueron: verificar hasta qué punto fue provechoso el seminario metodológico ofrecido y evaluar el desempeño de los docentes en las clases utilizando Software “R”.

Fueron seleccionados los cursos de Enseñanza de Historia y Filosofía para la implementación de las clases; cuya selección fue hecha en base a las observaciones efectuadas durante la visita a clases de los docentes participantes en el estudio. Se seleccionaron los cursos que mayores dificultades presentaban en el aprendizaje de la Estadística.

Las clases se dan en la sala de Informática del ISCED/Uíge, con una duración de 90 minutos. Fueron dadas ocho clases en general, cuatro de ellas en el curso de licenciatura en Enseñanza de la Historia en el período diurno, con la participación de 20 alumnos y las otras en el curso de Licenciatura en Enseñanza de la Filosofía período post-laboral con una participación de 19 alumnos. Todas las clases contaron con la participación activa del investigador de este estudio, y durante la observación el objetivo del investigador fue en

torno al desempeño de los docentes investigados, por esta razón, en este estudio no serán presentados los resultados sobre el desempeño de los alumnos.

Determinación del estado final después de la intervención en la práctica educativa

Esta fase fue una actividad investigativa, que consistió en la formulación y aplicación de un cuestionario semejante al aplicado en el estado inicial, para comparar las condiciones de los docentes antes y después de la intervención en la práctica educativa. El objetivo principal de esta actividad es diagnosticar el estado final de desarrollo de habilidades de los docentes en la utilización del Software “R”, en las clases de la Estadística.

Evaluación de la implementación de la práctica y principales resultados

Desde el inicio de este estudio, se constató una enseñanza de la Estadística desvinculado del panorama actual. Se concibió la investigación desde la determinación del estado inicial (real) que presenta la temática objeto de estudio, trabajando para su transformación a un estado final (deseado). Para esta transformación se utilizó el seminario metodológico y el trabajo de cada grupo que participó en el estudio. La discusión y aprobación de los resultados de este estudio, son fundamentadas en el análisis de los tres momentos, que versaron sobre la descripción de la práctica educativa y su implementación. El cuestionario dirigido a los docentes participantes en la pesquisa, antes de la intervención en la práctica educativa y después de esta, permitió lograr los resultados presentados en las tablas y gráficos a continuación:

Tabla 5- Evaluación de las habilidades informáticas de los docentes

Tipos de habilidades	Habilidades	Resultados iniciales (RI)					Resultados finales (RF)				
		1	2	3	4	5	1	2	3	4	5
		B	R	M	A	M. A	B	R	M	A	M. A
Habilidades generales con el uso de herramientas de cómputo	1. Manipular accesorios.	8							2	4	2
	2. Navegar por el sistema operativo.	8							3	2	3
	3. Interactuar con softwares	8							3	3	2
Habilidades específicas en el uso del Software “R”	1. Construir tablas	8						1	2	4	1
	2. Construir gráficos de barras	8						2	3	1	2
	3. Construir gráficos de sectores.	8						1	4	2	1

Observación: **B** significa evaluación baja, **R**_ regular, **M**_ media, **A** _ alta e **M.A** _ muy alta.

Fuente: Elaboración propia

La tabla tres muestra la comparación de los resultados iniciales y finales de las habilidades informáticas de los docentes participantes en el estudio. Sin embargo, al observar éstos resultados se observa que el nivel de desarrollo de sus habilidades en la ejecución de las tareas propuestas mejoró considerablemente. En las habilidades generales, donde se analizó la forma de manejar accesorios, navegar por el sistema operativo e interactuar con el Software “R”, la evaluación final de los ocho docentes fue: media, alta y muy alta. En las habilidades específicas, donde fue evaluada las construcciones de tablas de distribución de frecuencias, gráficos de barras y sectores, la evaluación final fue: regular, media, alta y muy alta.

Tabla 6- Criterios iniciales y finales de los docentes en relación al uso del computador en la enseñanza de la Estadística

Preguntas	Criterios iniciales (CI)			Criterios finales (CF)		
	Si	No	Depende	Si	No	Depende
¿Es posible motivar a los alumnos en el aprendizaje de la Estadística Descriptiva, conciliando las construcciones gráficas hechas manualmente con las herramientas de cómputo?	2	4	2	7	0	1
¿La utilización de herramientas de cómputo desarrollaría una enseñanza de Estadística Descriptiva relacionada con la realidad?	3	2	3	5	1	2
¿La implementación de herramientas de cómputo, ayudaría a mejorar las condiciones de las instituciones de enseñanza superior en el país?	2	1	5	4	1	3

Gráficos correspondiente

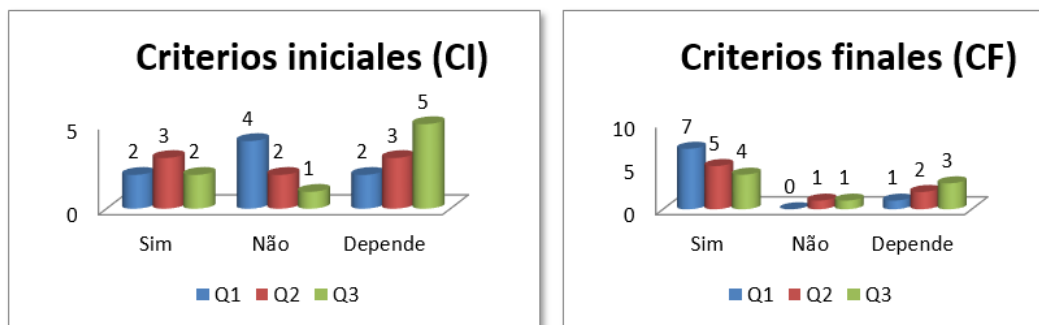


Gráfico 3- Criterios iniciales y finales de los docentes en relación al uso de la computadora

Los resultados presentados en la tabla 5, muestran el cambio significativo en los criterios iniciales y finales, fruto del trabajo realizado con ellos en el seminario metodológico y durante las clases de experimentación impartidas por docentes experimentados en la materia objeto de estudio.

Tabla 7- Criterios iniciales y finales de los docentes en relación a las principales razones de no utilización de herramientas computacionales en la enseñanza de la Estadística

Pregunta4.	Cantidad de docentes	
	Criterios iniciales (CI)	Criterios finales (CF)
¿Cuáles son las principales razones de no utilización de herramientas computacionales en la enseñanza de Estadística Descriptiva?		
a) No hace falta en el aprendizaje de los alumnos.	5	2
b) Falta de condiciones en las instituciones.	3	1
c) Falta de dominio de los Softwares adecuados para el efecto.	8	3
d) El exceso de alumnos en las aulas.	2	1

Fuente: Elaboración propia

Los resultados iniciales y finales muestran un cambio positivo de criterios, de pensamientos y el significativo nivel de desarrollo de las habilidades logradas por los docentes con relación al uso de las herramientas computacionales en la enseñanza de Estadística Descriptiva.

Desempeño de los docentes en las clases de experimentación

De forma general, las clases fueron exploratorias, mostrando las diversas potencialidades que hacen de la “R” herramienta adecuada para apoyar el trabajo del docente en la enseñanza de la Estadística. Los docentes con un nivel de conocimiento sobre el uso de la computadora -integrantes de los grupos A, B y D-, mostraban desarrollo en la perspectiva tecnológica y valoraban la experimentación, la exploración y el papel activo de los alumnos en el aprendizaje, contrariamente de aquéllos que poco dominaban el uso de la computadora -integrantes del grupo C-.

Sin embargo, los resultados de esta experiencia, mostraron que cuando existen condiciones mínimas en las instituciones de la Enseñanza Superior, la utilización de herramientas computacionales en el proceso de enseñanza-aprendizaje de la Estadística Descriptiva es un camino viable y útil.

Conclusiones

El presente estudio es fruto de un trabajo investigativo de una Maestría que apunta en esa dirección, es la primera investigación de esta naturaleza en la Sección de Enseñanza de Matemáticas del ISCED/Uíge. Es una fuente de información para futuros trabajos en esta

dirección. Un estudio apoyado en la experiencia internacional, mostraron que es aconsejable la introducción de las nuevas tecnologías por vía científica, para perfeccionar la enseñanza de las disciplinas con el uso de herramientas computacionales.

La utilización del Software “R”, alteró la enseñanza y dinámica de las clases de Estadística Descriptiva en el ISCED/Uíge, tornando el proceso más fácil y natural. Además, aportó en la mejoría de condiciones de laboratorio de Informática del ISCED/Uíge e incentivó a los docentes en la implementación de otros métodos de enseñanza en las disciplinas que necesitan la computadora como medio auxiliar.

La integración de estos resultados muestra que es posible y necesario elevar el nivel de preparación de los docentes, lo cual promueve una mejoría en la eficacia y dinamismo en la conciliación de las resoluciones y construcciones hechas manualmente en la Estadística.

Este estudio, constituye un ejemplo de utilización exitosa de la herramienta computacional “R” en una institución de la Enseñanza Superior en Angola, al transformar la enseñanza de la Estadística Descriptiva en contenidos atractivos y motivando a los alumnos en el aprendizaje de esta disciplina, particularmente los del curso de licenciatura en enseñanza de las ciencias sociales.

La utilización de herramientas computacionales en el análisis estadístico es de extrema importancia, pues la rapidez y eficiencia en la resolución de casos fue notoria. Los docentes que inicialmente mostraban desinterés en el uso del Software “R”, por el hecho de tener poco dominio en el uso de computadora en las clases, quedaron motivados en frecuentar curso de informática en el objetivo de utilizarla en sus clases.

Referencias Bibliográficas

Carvalho, A. (2015). *Exercícios de Excel para Estatística*. FCA.

Figueredo, F.; Figueredo, A.; Ramos, A.; y Paulo, T. (2017). *Inferência Estatística - Problemas resolvidos e propostos com aplicações com R*. Escolar Editora.

João, K. (2017). *Excel e R como ferramenta no ensino da Estatística*. Porto.

Pereira, M. y Piario, H. (2012). *Introdução à Estatística com R e SPSS*. Chiado Editora.

Tomás, M. T. (2006). *Matemática 11ª Classe*. Texto Editores.