

Sílabo

170332 - Programación para la Ciencia de Datos

I. Información general

Nombre del Curso: Programación para la Ciencia de Datos

Código del curso: 170332

Departamento Académico: Ingeniería

Créditos: 4

Horas Teoría: 3

Horas Práctica: 2

Periodo Académico: 2023-01-PRE

Sección: A

Modalidad: Presencial

Idioma: Español

Docente: UWE ROJAS VILLANUEVA

Email docente: u.rojasv@up.edu.pe

II. Introducción

El presente curso brindará a los estudiantes una visión a un nivel básico e intermedio de conceptos relativos a Algoritmos y Estructuras de Datos utilizando el lenguaje R. Se presentan temas de análisis de complejidad de algoritmos. Se explicará cómo las técnicas aprendidas serán particularmente útiles en la aplicabilidad de conceptos computacionales al interior de problemas relativos a sistemas de información en las organizaciones.

III. Logro de aprendizaje final del curso

IV. Unidades de aprendizaje

Introducción a los algoritmos y al lenguaje de programación R

Logro de Aprendizaje / propósito de la unidad:

Conocer la definición formal y matemática de un algoritmo e implementación en R

Contenidos:

- Algoritmos y pseudocódigo
- Estructuras secuenciales
- Estructuras condicionales
- Estructuras repetitivas

Unidad de aprendizaje 2: Estructura de control: Secuencia, selección y repetición

Logro de Aprendizaje / propósito de la unidad:

Conocer y aplicar las definiciones de condicionales y repetición: if- else, while, for



Contenidos:

- Funciones
- Listas
- Vectores
- Matrices

Unidad de aprendizaje 3: Escritura de funciones

Logro de Aprendizaje / propósito de la unidad:

Aplicar la definición de una función, parámetros de paso en una función, funciones con retorno y sin retorno y funciones con recursividad.

Contenidos:

- Definición de una función
- Definición de una función con retorno
- Definición de una función sin retorno
- Definición de una función recursiva

Unidad de aprendizaje 4: Estructura de Datos:

Logro de Aprendizaje / propósito de la unidad:

Conocer y aplicar la definición de Vectores, Listas, Matrices, Data Frame.

Contenidos:

Operaciones con vectores, matrices, listas

- Creación de un vector
- Creación de una matriz y operaciones
- Implementación de un vector y una matriz utilizando funciones

Unidad de aprendizaje 5: Análisis de Datos

Logro de Aprendizaje / propósito de la unidad:

Estudio y análisis de datos importados desde tablas

Contenidos:

- Data Frames
- Strings (Cadenas de caracteres)

- Expresiones Regulares
- Librería dplyr
- Filtros

Unidad de aprendizaje 6: Visualización de Datos

Logro de Aprendizaje / propósito de la unidad:

Conocer la importación de archivos con extensión CSV, EXCEL, XML, JSON, Gráficos con Histogramas, Gráficos Circulares, Gráficos de Barras, diagramas de cajas

Contenidos:

- Instalación de Base de Datos MySQL
- Tablas MySQL
- Llaves Primarias y Secundarias (Primary key and Foreign Key)
- Interconexión MySQL y R (Rstudio)
- Lectura de datos de tipo Excel, CSV
- Gráficos usando histogramas
- Gráficos Circulares
- Diagramas de Cajas
- Interfaz utilizando Shiny con R

V. Estrategias Didácticas

- Participación activa de parte de los estudiantes mediante la solución compartida de problemas.
- Explicación de la aplicabilidad de los conceptos aprendidos a situaciones reales.
- Método expositivo por parte del docente.

Implementación de los algoritmos y conceptos aprendidos haciendo uso del lenguaje R.

VI. Sistemas de evaluación

Consideraciones para las evaluaciones

SISTEMA DE EVALUACIÓN

Nota de trabajo 40%

Examen Parcial 25%

Examen Final 35%

Cálculo de la Nota de Trabajo:

Promedio trabajo escrito = 40% Nota Primera Entrega + 60% Nota Segunda Entrega

Factor_3: Calculado según la tabla y en función de la nota_de_sustentación_individual

Factor_4: Refleja la nota de contribución (en cantidad y calidad) dada por el resto del grupo con un mínimo de 0.8 y un máximo de 1.05

Nota Proyecto Empresarial = Promedio trabajo escrito * Factor_3 * Factor_4

Factor_2: Calculado según la tabla y en función de la Nota_Proyecto_Empresarial

Factor_1: Calculado según la tabla y en función de la Nota_Proyecto_Empresarial

Nota de Trabajo = [Promedio_de_Prácticas individuales_1_y_2] * Factor_1 * Factor 2

Notas:

1) PCs, Exámenes son individuales y miden competencias (teoría no puede ser más de 5 puntos y no puede ser selección múltiple debe ser minicaso o ejemplificación de la teoría).

2) No hay puntos extras en PCs ni en exámenes bajo ningún motivo cualquier punto extra será considerado en "Otras actividades"

3) La tabla es una equivalencia entre nota vigesimal y un factor de 0 a 1.

Nombre evaluación	%	Fecha	Criterios	Comentarios
1. Trabajos	40			
1.1. Primer Entregable	40			
1.2. Segundo Entregable	60			
1.3. Practicas Calificadas				
1.3.1. Primera Practica				
1.3.2. Segunda Practica				
1.3.3. Tercera Practica				
1.3.4. Cuarta Practica				



2. Examen Parcial	25			
3. Examen Final	35			

VII. Cronograma referencial de actividades

Unidades de aprendizaje	Contenidos y actividades a realizar	Recursos y materiales	Evaluaciones
Semana 1: del 20/03/2023 al 25/03/2023			
<ul style="list-style-type: none"> Introducción a los algoritmos y al lenguaje de programación R 	<p>Contenidos:</p> <p>Definición formal y matemática de un algoritmo, instalación de RStudio IDE. Definición formal y matemática de un algoritmo, instalación de RStudio IDE.</p>		
Semana 2: del 27/03/2023 al 01/04/2023			
<ul style="list-style-type: none"> Introducción a los algoritmos y al lenguaje de programación R 	<p>Fundamentos de Lenguaje de programación R. Sintaxis básica, tipos de datos, estructura de datos, variables, keywords, operadores.</p>		
Semana 3 con feriados el jueves 06, viernes 07 y sábado 08: del 03/04/2023 al 08/04/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 2: Estructura de control: Secuencia, selección y repetición Unidad de aprendizaje 3: Escritura de funciones 	<p>Definición de condicionales y repetición como son if- else, while, for</p>		
Semana 4: del 10/04/2023 al 15/04/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 2: Estructura de control: Secuencia, selección y repetición Unidad de aprendizaje 3: Escritura de funciones 	<p>Definición de una función, parámetros de paso en una función, funciones con retorno y sin retorno, funciones con recursividad</p>		

Unidades de aprendizaje	Contenidos y actividades a realizar	Recursos y materiales	Evaluaciones
Semana 5: del 17/04/2023 al 22/04/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 4: Estructura de Datos: 	Vectores, Listas, Arrays		
Semana 6: del 24/04/2023 al 29/04/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 3: Escritura de funciones Unidad de aprendizaje 4: Estructura de Datos: 	Métodos de Ordenamiento: Algoritmos de selección, inserción, burbuja, etc.		
Semana 7: del 01/05/2023 al 06/05/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 3: Escritura de funciones Unidad de aprendizaje 4: Estructura de Datos: 	Matrices, Operaciones con matrices, Data Frame, Factors,		
Semana 8 de exámenes parciales: del 08/05/2023 al 13/05/2023			
Semana 9: del 15/05/2023 al 20/05/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 5: Análisis de Datos 	<ul style="list-style-type: none"> - Data Frames - Verificación de datos repetidos, datos nulos, etc - Funciones select mutate 		
Semana 10: del 22/05/2023 al 27/05/2023			

Unidades de aprendizaje	Contenidos y actividades a realizar	Recursos y materiales	Evaluaciones
<ul style="list-style-type: none"> Unidad de aprendizaje 5: Análisis de Datos 	<ul style="list-style-type: none"> - Filtros (funcion filter) - libreria dplyr - agrupamiento (group_by) - sumariacion (resumen de los datos) 		
Semana 11: del 29/05/2023 al 03/06/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 5: Análisis de Datos 	<ul style="list-style-type: none"> - muestreo de datos (sample). - expresiones regulares (para cadenas) - funciones apply 		
Semana 12: del 05/06/2023 al 10/06/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 6: Visualización de Datos 	Conexión de una base datos libre MySQL con R studio		
Semana 13: del 12/06/2023 al 17/06/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 6: Visualización de Datos 	Visualización de Tablas Excel, CSV, txt, etc. cargados al lenguaje R		
Semana 14: del 19/06/2023 al 24/06/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 6: Visualización de Datos 	Calculo de variables estadísticas y análisis modelos de regresión lineal y multiple		
Semana 15 con feriado jueves 29: del 26/06/2023 al 01/07/2023			
<ul style="list-style-type: none"> Unidad de aprendizaje 6: Visualización de Datos 	Introducción a machine learning - predicción de resultados via modelos de regresión lineal y múltiple, graficos		
Semana 16 de exámenes finales: del 03/07/2023 al 08/07/2023			

VIII. Referencias bibliográficas