



Programa de asignatura por competencias de educación superior

Sección I. Identificación del Curso

Tabla 1. Identificación de la Planificación del Curso.

Actualización:	Marzo 12, 2024				
Carrera:	Ingeniería en Desarrollo de Software	Asignatura:	Análisis de datos		
Academia:	Ciencias Computacionales y Programación /	Clave:	19SDS28		
Módulo formativo:	Programación aplicada	Seriación:	- -		
Tipo de curso:	Presencial	Prerrequisito:	19SDS15 - Estructuras de datos y algoritmia		
Semestre:	Séptimo	Créditos:	6.75	Horas semestre:	108 horas
Teoría:	2 horas	Práctica:	2 horas	Trabajo indpt.:	2 horas
				Total x semana:	6 horas

Sección II. Objetivos educacionales

Tabla 2. Objetivos educacionales

Objetivos educacionales		Criterios de desempeño	Indicadores
1	Los egresados gestionarán recursos relacionados con el desarrollo de software en alguna organización.	Los egresados podrán aplicar metodologías en el desarrollo de proyectos en el contexto laboral.	20% de los egresados aplicarán metodologías en el desarrollo de software en su contexto laboral.
2	Los egresados diseñarán e implementarán soluciones innovadoras mediante el uso de tecnologías de la información.	Los egresados participarán activamente en el ciclo de desarrollo e integración continuos	25% de los egresados desempeñarán labores de desarrollo e integración continuos.
3	Los egresados desarrollarán conocimiento especializado que les permite enfocarse en un área del conocimiento específico del desarrollo de software.	Los egresados desempeñarán actividades orientadas al aseguramiento de los activos de información de manera resiliente, la gestión de la infraestructura de redes y comunicaciones, o integrando hardware y software para crear soluciones IoT; así como el uso de inteligencia artificial para gestionar datos y reconocer patrones que determinen oportunidades de negocio en las organizaciones.	5% de los egresados desempeñarán labores en desarrollo de soluciones IoT.
4	Los egresados serán capaces de emprender un negocio basado en el desarrollo de un producto o servicio de tecnologías de la información, aportando valor a la generación de empleos e incrementar el bienestar económico y social, de forma ecológica y sustentable.	Los egresados serán capaces de emprender un negocio basado en el desarrollo propio de un producto o servicio de tecnologías de la información.	2% de los egresados tendrán participación en el acta constitutiva de una empresa creada a partir del desarrollo de software para ofrecer un producto o servicio.



Atributos de egreso de plan de estudios		Criterios de desempeño	Componentes
1	Desarrollar una experimentación adecuada para recopilar, almacenar y analizar grandes cantidades de información basándose en el juicio ingenieril para crear productos o servicios innovadores mediados por software.	- Comprenderá los elementos necesarios para realizar el análisis de datos estructurados y no estructurados para contribuir en la solución de problemas en el campo de la ingeniería.	<p>Análisis de Datos estructurados y no estructurados.</p> <ol style="list-style-type: none"> 1. Formato de archivos. <ol style="list-style-type: none"> 1.1 CSV. 1.2 Excel. 1.3 JSON. 1.4 SQL. 1.5 XML. 1.6 BSON. 2. Procesamiento de Información. <ol style="list-style-type: none"> 2.1 Series. 2.2 DataFrames. 2.3 Inner y outer join. 2.4 Concatenar. 2.5 Eliminar. 2.6 Agregación. 2.7 Agrupación. 2.8 Iterar. 3.- Manipulación de datos. <ol style="list-style-type: none"> 3.1 Convertir tipos de datos. 3.2 Variables categóricas. 3.3 Eliminar duplicados. 3.4 Datos faltantes. 3.5 Imputación. 3.6 Normalización. 3.7 Datos atípicos. 4. Visualización de datos.



Continuación: Tabla 2. Objetivos educacionales (continuación)

No.	Atributos de egreso de plan de estudios	Criterios de desempeño	Componentes
			4.1 Estadísticas de los datos. 4.2 Histogramas. 4.3 Distribution plot. 4.4 Boxplot. 4.5 Violinplot. 4.6 Barplot. 4.7 Heatmap. 4.8 Visualización en altas dimensiones. 4.9 Visualización geográfica. 5. Datos no estructurados. 5.1 Audio. 5.2 Texto. 5.3 Imágenes. 5.4 Data Lake. 6.- Análisis de componentes principales (PCA). 6.1 Factorización no negativa. 6.2 Análisis de componentes independientes (ICA). 6.3 Reducción de dimensión supervisada: Análisis de discriminante lineal (LDA). 7.- Modelo autorregresivo. 7.1 Definición. 7.2 Ejemplos. 7.3 Autorelación. 7.4 Estacionalidad. 7.5 Modelo Arima.
2		- Reconocerá los tipos de datos estructurados y no estructurados para darles el tratamiento adecuado para la presentación de la información.	Análisis de Datos estructurados y no estructurados. 1. Formato de archivos. 1.1 CSV.



Continuación: Tabla 2. Objetivos educacionales (continuación)

No.	Atributos de egreso de plan de estudios	Criterios de desempeño	Componentes
	<p>Reconocer la mejora continua como parte de su desarrollo profesional para mantener un perfil actualizado en desarrollo de software para el diseño e implementación de productos y servicios basados en tecnologías con las tendencias emergentes.</p>		<p>1.2 Excel. 1.3 JSON. 1.4 SQL. 1.5 XML. 1.6 BSON. 2. Procesamiento de Información. 2.1 Series. 2.2 DataFrames. 2.3 Inner y outer join. 2.4 Concatenar. 2.5 Eliminar. 2.6 Agregación. 2.7 Agrupación. 2.8 Iterar. 3.- Manipulación de datos. 3.1 Convertir tipos de datos. 3.2 Variables categóricas. 3.3 Eliminar duplicados. 3.4 Datos faltantes. 3.5 Imputación. 3.6 Normalización. 3.7 Datos atípicos. 4. Visualización de datos. 4.1 Estadísticas de los datos. 4.2 Histogramas. 4.3 Distribution plot. 4.4 Boxplot. 4.5 Violinplot. 4.6 Barplot. 4.7 Heatmap.</p>



Continuación: Tabla 2. Objetivos educacionales (continuación)

No.	Atributos de egreso de plan de estudios	Criterios de desempeño	Componentes
			4.8 Visualización en altas dimensiones. 4.9 Visualización geográfica. 5. Datos no estructurados. 5.1 Audio. 5.2 Texto. 5.3 Imágenes. 5.4 Data Lake. 6.- Análisis de componentes principales (PCA). 6.1 Factorización no negativa. 6.2 Análisis de componentes independientes (ICA). 6.3 Reducción de dimensión supervisada: Análisis de discriminante lineal (LDA). 7.- Modelo autorregresivo. 7.1 Definición. 7.2 Ejemplos. 7.3 Autorelación. 7.4 Estacionalidad. 7.5 Modelo Arima.

Sección III. Atributos de la asignatura

Tabla 3. Atributos de la asignatura

Problema a resolver		
Permitir dar una mejor manipulación e interpretación de la información, a través de la conceptualización de temas de exploración y análisis de datos como base de la ciencia de datos.		
Atributos (competencia específica) de la asignatura		
Conocer los conceptos exploración y análisis de datos para aplicarlos en la comprensión de la literatura concerniente a ciencia de datos.		
Aportación a la competencia específica		Aportación a las competencias transversales
Saber	Saber hacer	Saber Ser
- Identificar, distinguir y asimilar los diferentes conceptos de exploración y análisis de datos como antesala para la ciencia de datos.	- Realizar ejercicios de exploración y análisis de datos como base de la ciencia de datos.	- Demuestra un compromiso ético y formal en la realización y entrega de ejercicios y proyectos durante el curso.
Producto integrador de la asignatura, considerando los avances por unidad		
Conceptualización y realización de ejercicios de exploración y análisis de datos enfocados en el tema de ciencia de datos, para un correcto entendimiento de los temas.		

Sección IV. Desglose específico por cada unidad formativa

Tabla 4.1. Desglose específico de la unidad "Importando datos."

Número y nombre de la unidad: 1. Importando datos.				
Tiempo y porcentaje para esta unidad:		Teoría: 4 horas	Práctica: 4 horas	Porcentaje del programa: 11.11%
Aprendizajes esperados: Reconocer algunos de los diferentes formatos de datos estructurados existentes para manipular la información a futuro.				
Temas y subtemas (secuencia)	Criterios de desempeño	Estrategias didácticas	Estrategias de evaluación	Producto Integrador de la unidad (Evidencia de aprendizaje de la unidad)
1. Formato de archivos. 1.1 CSV. 1.2 Excel. 1.3 JSON. 1.4 SQL. 1.5 XML. 1.6 BSON.	Saber: - Asimilar de forma adecuada la importación de datos en diferentes formatos. Saber hacer: - Implementar de forma correcta la importación de datos en diferentes formatos. Ser: Entrega en tiempo y forma las actividades asignadas durante el periodo del curso.	- Investigación. - Aprendizaje colaborativo. - Aprendizaje basado en problemas. - Aprendizaje basado en proyectos. - Explicación por parte del docente. - Alumno enseña a otro compañero.	Evaluación diagnóstica: - Rescatar conocimientos previos. Evaluación formativa: -Retroalimentación de trabajos. -Autoevaluación. -Coevaluación. -Evaluaciones en base a TIC's.: formativas con retroalimentación y automáticas. Evaluación sumativa: -Prueba problemas. -Prueba escrita. -Pruebas orales. -Evaluación mediante procesos de investigación. -Evaluación basada en proyectos.	Portafolio de evidencias de prácticas realizadas, así como actividades de aprendizaje realizados durante la unidad.
Bibliografía				



Continuación: Tabla 4.1. Desglose específico de la unidad "Importando datos."

Bibliografía

- Dale, K. (2022). Data Visualization with Python and JavaScript. 2nd ed. O'Reilly Media, Inc.USA CA.
- Doble, M. (2020). The Data Visualization Workshop. Packt Publishing Ltd. UK
- Pajankar, A. (2022). Hands-on Matplotlib: Learn Plotting and Visualizations with Python 3. Apress. India
- Phuong, V. (2017). Python: Data Analytics and Visualization. Packt Publishing Ltd. UK

Sección IV. Desglose específico por cada unidad formativa

Tabla 4.2. Desglose específico de la unidad "Manipulación de Datos"

Número y nombre de la unidad: 2. Manipulación de Datos							
Tiempo y porcentaje para esta unidad:		Teoría:	5 horas	Práctica:	5 horas	Porcentaje del programa:	13.89%
Aprendizajes esperados: Dar tratamiento adecuado a los datos estructurados para su posterior utilización.							
Temas y subtemas (secuencia)	Criterios de desempeño	Estrategias didácticas	Estrategias de evaluación	Producto Integrador de la unidad (Evidencia de aprendizaje de la unidad)			
2. Procesamiento de Información. 2.1 Series. 2.2 DataFrames. 2.3 Inner y outer join. 2.4 Concatenar. 2.5 Eliminar. 2.6 Agregación. 2.7 Agrupación. 2.8 Iterar.	Saber: - Identificar conceptos concernientes al procesamiento de información. - Asimilar de forma adecuada la manera de procesar la información para darles una interpretación. Saber hacer: - Implementa de forma correcta la manera de procesar la información para darles una interpretación. Ser: Entrega en tiempo y forma las actividades	- Investigación. - Aprendizaje colaborativo. - Aprendizaje basado en problemas. - Aprendizaje basado en proyectos. - Explicación por parte del docente. - Alumno enseña a otro compañero.	Evaluación formativa: -Retroalimentación de trabajos. -Autoevaluación. -Coevaluación. -Evaluaciones en base a TIC's.: formativas con retroalimentación y automáticas. Evaluación sumativa: -Prueba problemas. -Prueba escrita. -Pruebas orales. -Evaluación mediante procesos de investigación. -Evaluación basada en proyectos.	Portafolio de evidencias de prácticas realizadas, así como actividades de aprendizaje realizados durante la unidad.			



Continuación: Tabla 4.2. Desglose específico de la unidad "Manipulación de Datos"

Temas y subtemas (secuencia)	Criterios de desempeño	Estrategias didácticas	Estrategias de evaluación	Producto Integrador de la unidad
	asignadas durante el periodo del curso.			

Bibliografía

- Dale, K. (2022). Data Visualization with Python and JavaScript. 2nd ed. O'Reilly Media, Inc. USA CA.
- Doble, M. (2020). The Data Visualization Workshop. Packt Publishing Ltd. UK
- Pajankar, A. (2022). Hands-on Matplotlib: Learn Plotting and Visualizations with Python 3. Apress. India
- Phuong, V. (2017). Python: Data Analytics and Visualization. Packt Publishing Ltd. UK

Sección IV. Desglose específico por cada unidad formativa

Tabla 4.3. Desglose específico de la unidad "Tratamiento de Información."

Número y nombre de la unidad: 3. Tratamiento de Información.							
Tiempo y porcentaje para esta unidad:		Teoría:	6 horas	Práctica:	6 horas	Porcentaje del programa:	16.67%
Aprendizajes esperados: Depurar los datos para generalizarlos con el preprocesamiento de información.							
Temas y subtemas (secuencia)	Criterios de desempeño	Estrategias didácticas	Estrategias de evaluación	Producto Integrador de la unidad (Evidencia de aprendizaje de la unidad)			
3.- Manipulación de datos. 3.1 Convertir tipos de datos. 3.2 Variables categóricas. 3.3 Eliminar duplicados. 3.4 Datos faltantes. 3.5 Imputación. 3.6 Normalización. 3.7 Datos atípicos.	Saber: - Asimilar de forma adecuada la manera de manipular datos para una mejor interpretación Saber hacer: - Implementar de forma correcta la manipulación de datos para una mejor interpretación. Ser: Entrega en tiempo y forma las actividades asignadas durante el periodo del curso.	- Investigación. - Aprendizaje colaborativo. - Aprendizaje basado en problemas. - Aprendizaje basado en proyectos. - Explicación por parte del docente. - Alumno enseña a otro compañero.	Evaluación formativa: -Retroalimentación de trabajos. -Autoevaluación. -Coevaluación. -Evaluaciones en base a TIC's.: formativas con retroalimentación y automáticas. Evaluación sumativa: -Prueba problemas. -Prueba escrita. -Pruebas orales. -Evaluación mediante procesos de investigación. -Evaluación basada en proyectos.	Portafolio de evidencias de prácticas realizadas, así como actividades de aprendizaje realizados durante la unidad.			
Bibliografía							
<ul style="list-style-type: none"> - Dale, K. (2022). Data Visualization with Python and JavaScript. 2nd ed. O'Reilly Media, Inc.USA CA. - Doble, M. (2020). The Data Visualization Workshop. Packt Publishing Ltd. UK - Pajankar, A. (2022). Hands-on Matplotlib: Learn Plotting and Visualizations with Python 3. Apress. India - Phuong, V. (2017). Python: Data Analytics and Visualization. Packt Publishing Ltd. UK 							

Sección IV. Desglose específico por cada unidad formativa

Tabla 4.4. Desglose específico de la unidad "Visualización de datos."

Número y nombre de la unidad: 4. Visualización de datos.							
Tiempo y porcentaje para esta unidad:		Teoría:	6 horas	Práctica:	6 horas	Porcentaje del programa:	16.67%
Aprendizajes esperados: Mostrar en los diferentes escenarios la información preprocesada para ser analizada en sus ámbitos.							
Temas y subtemas (secuencia)	Criterios de desempeño	Estrategias didácticas	Estrategias de evaluación	Producto Integrador de la unidad (Evidencia de aprendizaje de la unidad)			
4. Visualización de Datos. 4.1 Estadísticas de los datos. 4.2 Histogramas. 4.3 Distribution plot. 4.4 Boxplot. 4.5 Violinplot. 4.6 Barplot. 4.7 Heatmap. 4.8 Visualización en altas dimensiones. 4.9 Visualización geográfica.	Saber: - Asimilar de forma adecuada las diferentes formas de visualizar información. Saber hacer: - Implementar de forma correcta las diferentes formas de visualizar información. Ser: Entrega en tiempo y forma las actividades asignadas durante el periodo del curso.	- Investigación. - Aprendizaje colaborativo. - Aprendizaje basado en problemas. - Aprendizaje basado en proyectos. - Explicación por parte del docente. - Alumno enseña a otro compañero.	Evaluación formativa: -Retroalimentación de trabajos. -Autoevaluación. -Coevaluación. -Evaluaciones en base a TIC's.: formativas con retroalimentación y automáticas. Evaluación sumativa: -Prueba problemas. -Prueba escrita. -Pruebas orales. -Evaluación mediante procesos de investigación. -Evaluación basada en proyectos.	Portafolio de evidencias de prácticas realizadas, así como actividades de aprendizaje realizados durante la unidad.			
Bibliografía							
<ul style="list-style-type: none"> - Dale, K. (2022). Data Visualization with Python and JavaScript. 2nd ed. O'Reilly Media, Inc.USA CA. - Doble, M. (2020). The Data Visualization Workshop. Packt Publishing Ltd. UK - Pajankar, A. (2022). Hands-on Matplotlib: Learn Plotting and Visualizations with Python 3. Apress. India - Phuong, V. (2017). Python: Data Analytics and Visualization. Packt Publishing Ltd. UK 							

Sección IV. Desglose específico por cada unidad formativa

Tabla 4.5. Desglose específico de la unidad "Datos no Estructurados."

Número y nombre de la unidad: 5. Datos no Estructurados.							
Tiempo y porcentaje para esta unidad:		Teoría:	5 horas	Práctica:	5 horas	Porcentaje del programa:	13.89%
Aprendizajes esperados: Reconocer los datos no estructurados en sus diferentes presentaciones para darles el tratamiento adecuado.							
Temas y subtemas (secuencia)	Criterios de desempeño	Estrategias didácticas	Estrategias de evaluación	Producto Integrador de la unidad (Evidencia de aprendizaje de la unidad)			
5. Datos no estructurados. 5.1 Audio. 5.2 Texto. 5.3 Imágenes. 5.4 Data Lake.	<p>Saber:</p> <ul style="list-style-type: none"> - Asimila de forma adecuada los datos no estructurados para su correcto procesamiento. <p>Saber hacer:</p> <ul style="list-style-type: none"> - Implementa de forma correcta los datos no estructurados para su correcto procesamiento. <p>Ser:</p> <ul style="list-style-type: none"> - Entrega en tiempo y forma las actividades asignadas durante el periodo del curso. 	<ul style="list-style-type: none"> - Investigación. - Aprendizaje colaborativo. - Aprendizaje basado en problemas. - Aprendizaje basado en proyectos. - Explicación por parte del docente. - Alumno enseña a otro compañero. 	<p>Evaluación formativa:</p> <ul style="list-style-type: none"> -Retroalimentación de trabajos. -Autoevaluación. -Coevaluación. -Evaluaciones en base a TIC's.: formativas con retroalimentación y automáticas. <p>Evaluación sumativa:</p> <ul style="list-style-type: none"> -Prueba problemas. -Prueba escrita. -Pruebas orales. -Evaluación mediante procesos de investigación. -Evaluación basada en proyectos. 	<p>Portafolio de evidencias de prácticas realizadas, así como actividades de aprendizaje realizados durante la unidad.</p>			
Bibliografía							
<ul style="list-style-type: none"> - Elder, J. (2012). Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications. Academic Press. USA - Isson, J. (2018). Unstructured Data Analytics. Wiley USA - Inmon, W. (2007). Tapping into Unstructured Data. Prentice-Hall. USA 							

Sección IV. Desglose específico por cada unidad formativa

Tabla 4.6. Desglose específico de la unidad "Reducción de dimensión."

Número y nombre de la unidad: 6. Reducción de dimensión.							
Tiempo y porcentaje para esta unidad:		Teoría:	5 horas	Práctica:	5 horas	Porcentaje del programa:	13.89%
Aprendizajes esperados: Conocer y utilizar los métodos para la reducción de las dimensiones de la información.							
Temas y subtemas (secuencia)	Criterios de desempeño	Estrategias didácticas	Estrategias de evaluación	Producto Integrador de la unidad (Evidencia de aprendizaje de la unidad)			
<p>6.- Análisis de componentes principales (PCA).</p> <p>6.1 Factorización no negativa.</p> <p>6.2 Análisis de componentes independientes (ICA).</p> <p>6.3 Reducción de dimensión supervisada: Análisis de discriminante lineal (LDA).</p>	<p>Saber:</p> <ul style="list-style-type: none"> - Asimilar de forma las diferentes maneras de reducir la dimensionalidad de un conjunto de datos. <p>Saber hacer:</p> <ul style="list-style-type: none"> - Implementar de forma correcta las diferentes maneras de reducir la dimensionalidad de un conjunto de datos. <p>Ser:</p> <ul style="list-style-type: none"> - Entrega en tiempo y forma las actividades asignadas durante el periodo del curso. 	<ul style="list-style-type: none"> - Investigación. - Aprendizaje colaborativo. - Aprendizaje basado en problemas. - Aprendizaje basado en proyectos. - Explicación por parte del docente. - Alumno enseña a otro compañero. 	<p>Evaluación formativa:</p> <ul style="list-style-type: none"> -Retroalimentación de trabajos. -Autoevaluación. -Coevaluación. -Evaluaciones en base a TIC's.: formativas con retroalimentación y automáticas. <p>Evaluación sumativa:</p> <ul style="list-style-type: none"> -Prueba problemas. -Prueba escrita. -Pruebas orales. -Evaluación mediante procesos de investigación. -Evaluación basada en proyectos. 	<p>Portafolio de evidencias de prácticas realizadas, así como actividades de aprendizaje realizados durante la unidad.</p>			
Bibliografía							
<ul style="list-style-type: none"> - Deisenroth, M. et al. (2020). Mathematics for Machine Learning. Cambridge University Press. UK - Unpingco, J. (2019). Python for Probability, Satatistics, and Machine Learning 2nd Ed. Springer. USA - Bishop, C. (2006).Pattern Recognition and Machine Learning, MIT Press. USA 							

Sección IV. Desglose específico por cada unidad formativa

Tabla 4.7. Desglose específico de la unidad "Introducción a las series de tiempo"

Número y nombre de la unidad: 7. Introducción a las series de tiempo							
Tiempo y porcentaje para esta unidad:		Teoría:	5 horas	Práctica:	5 horas	Porcentaje del programa:	13.89%
Aprendizajes esperados: Conocer las series de tiempo en base a su comportamiento para ser representados de manera visual.							
Temas y subtemas (secuencia)	Criterios de desempeño	Estrategias didácticas	Estrategias de evaluación	Producto Integrador de la unidad (Evidencia de aprendizaje de la unidad)			
6. Modelo autorregresivo. 7.1 Definición. 7.2 Ejemplos. 7.3 Autorelación. 7.4 Estacionalidad. 7.5 Modelo Arima.	Saber: - Asimilar de forma adecuada el comportamiento de una serie para identificar patrones y realizar previsiones. Saber hacer: - Implementar de forma correcta el comportamiento de una serie para identificar patrones y realizar previsiones. Ser: Entrega en tiempo y forma las actividades asignadas durante el periodo del curso.	- Investigación. - Aprendizaje colaborativo. - Aprendizaje basado en problemas. - Aprendizaje basado en proyectos. - Explicación por parte del docente. - Alumno enseña a otro compañero.	Evaluación formativa: -Retroalimentación de trabajos. -Autoevaluación. -Coevaluación. -Evaluaciones en base a TIC's.: formativas con retroalimentación y automáticas. Evaluación sumativa: -Prueba problemas. -Prueba escrita. -Pruebas orales. -Evaluación mediante procesos de investigación. -Evaluación basada en proyectos.	Portafolio de evidencias de prácticas realizadas, así como actividades de aprendizaje realizados durante la unidad.			
Bibliografía							
<ul style="list-style-type: none"> - Shumway, R. (S/F). Time Series Analysis and Its Applications with R Examples 2nd Ed Springer. USA. - Joseph, M. (2022). Modern Time Series Forecasting with Python. Packt Publishing. UK. - Pal, A. (2017). Practical Time Series Analysis. Packt Publishing Ltd. UK 							



V. Perfil docente

Tabla 5. Descripción del perfil docente

Perfil deseable docente para impartir la asignatura
<p>Carrera(s): Ingeniería en Desarrollo de Software, Ingeniería en Computación o carrera afín. o carrera afín</p> <ul style="list-style-type: none">- Relacionada con la materia de análisis de datos enfocada a la Ciencia de Datos.- Experiencia mínima de dos años- Mínimo Maestría relacionada con el área de conocimiento.