

## DATA MINING

### I. INFORMACIÓN GENERAL

<b>Curso:</b>	Data Mining y Análisis	<b>Semestre:</b>	2017-I
<b>Código:</b>	GES258	<b>Ciclo:</b>	Electivo
<b>N° de Créditos:</b>	3	<b>Profesor/es:</b>	Franco Riva
<b>N° horas de teoría:</b>	03 horas a la semana	<b>Horarios:</b>	1000
<b>Curso(s) Pre-requisito(s):</b>		<b>Carácter:</b>	Obligatorio o electivo

### II. FUNDAMENTACIÓN

El curso de Data Mining y Análisis de Datos otorga al alumno una poderosa herramienta que sirve en diversos ámbitos como la investigación de mercado, análisis de inventarios, gestión de personas, entre otros, además de facilitar el aprendizaje del software estadístico SPSS.

### III. SUMILLA

Data mining es una herramienta poderosa para describir patrones y relaciones en los datos. En el curso, los estudiantes aprenden a aplicar los principios de data mining para manejar y analizar conjuntos grandes y complejos de datos, incluyendo aquellos que se encuentran en la web. El alumno desarrollará la habilidad de convertir los datos, a partir de su análisis y evaluación, en información valiosa para la toma de decisiones en las organizaciones públicas, privadas y sociales.

Los temas del curso abordan el uso de las principales bases de datos del país y el uso de árboles de decisión, redes neuronales, reglas de asociación, clustering, métodos de casos, visualización de datos, entre otros.

### IV. RESULTADOS DE APRENDIZAJE

Al finalizar el curso, el estudiante:

- Formula un proyecto en función de un diagnóstico integral, delimitando productos, resultados e impacto, de manera cuantitativa y cualitativa.
- Planifica las actividades de un proyecto de forma secuencial y tomando en cuenta los recursos y el contexto interno y externo de la organización
- Gestiona el ciclo del proyecto aplicando metodologías y herramientas adecuadas para el diseño, monitoreo y evaluación en coherencia con la naturaleza del proyecto.

- Evalúa la viabilidad en sus distintas dimensiones, a partir del análisis sustantivo del contexto y del proyecto.
- Sustenta sus decisiones en base al análisis económico-financiero de la organización y su entorno, utilizando modelos y herramientas apropiados.
- Analiza los procesos de la organización identificando oportunidades de mejora que permitan hacerlos más eficientes y sostenibles.
- Organiza y recopila información para comprender las necesidades y expectativas del público objetivo, aplicando herramientas de investigación apropiadas.
- Diseña propuestas de valor alineadas a las necesidades del público objetivo tomando en cuenta las alternativas existentes.

## V. CONTENIDO

### Unidad 1: Estadística Base

- Entorno SPSS
- Estadísticas Descriptivas
- Pruebas T
- Regresión lineal Simple
- Gráficas

### Unidad 2: Análisis comparativo

- Pruebas de Kolmogorov-Smirnov y Levene
- Tablas cruzadas
- Regresión Lineal Múltiple
- Colinealidad
- Análisis de residuales
- ANOVA
- MANOVA
- Diseño Factorial
- Análisis de mediciones repetidas
- Análisis de efectos mixtos
- Regresión Logística

### Unidad 3: Análisis avanzado

- Análisis de componentes principales
- Análisis Factorial
- Análisis de Clústers
- Análisis de Funciones Discriminantes

## VI. METODOLOGÍA

- Se utilizarán clases teórico-prácticas para el análisis de datos mediante el uso de laboratorios multimedia y el software estadístico SPSS, los cuales estarán basados en el método del caso y la resolución de ejercicios.
- Las clases serán dictadas íntegramente en laboratorios para poder trabajar y analizar los casos que se irán presentando durante el curso. Estos casos servirán como base contextual para introducir los contenidos del curso y se desarrollarán en la pizarra como base de conocimiento.
- Debido a la complejidad teórica de los temas a desarrollar, el curso se enfocará en su uso y aplicación, sin dejar de lado una breve explicación.

- El curso está basado en un entorno de análisis de datos reales y es debido a eso que la principal metodología de evaluación será un trabajo integrador que permita a los alumnos consolidar el conocimiento y desarrollarlo en una aplicación real y tangible de resultados.
- La participación en clase será clave para el desarrollo del curso debido al uso del método del caso. Es por esto que se espera una alta participación del alumno y por ello también se evaluarán constantemente los aportes.

## VII. RECURSOS

- Software estadístico SPSS
- Laboratorios Multimedia
- Casos

## VIII. EVALUACIÓN

Tipo de Evaluación	Ponderación	Unidades	Producto
Evaluación Continua	20%	1, 2 y 3	
Promedio de Prácticas Calificadas	20%	1 y 2	Dos prácticas en aula de hora y media de duración. Se utilizarán para verificar que se manejen las herramientas adecuadas y se aprenda correctamente su uso.
Trabajo Aplicativo Final	60%	2 y 3	Trabajo aplicativo de acuerdo a lo indicado. Primera entrega: 20%(equivalente a examen parcial), Segunda Entrega 30%(Equivalente a examen final), Exposición: 10%

Respecto de la calificación, de acuerdo al Artículo 45° del Reglamento de la Facultad de Gestión y Alta Dirección:

Las calificaciones se expresa en escala de cero a veinte e incluyen los dos primeros decimales truncados. En el caso del promedio final del curso la fracción del medio punto o más será elevada al entero superior. La nota mínima es 11.

Respecto de la asistencia, de acuerdo al Artículo 33° del Reglamento de la Facultad de Gestión y Alta Dirección:

la asistencia puntual a clases es obligatoria. La inasistencia al treinta por ciento o más de las clases programadas en el horario regular del curso durante el semestre inhabilita al alumno para rendir la evaluación que el profesor indique en el sílabo o en su defecto, la evaluación final del curso. El profesor podrá prohibir el ingreso al aula a todo alumno cuya impuntualidad sea superior a 10 minutos contados desde la hora programada para el inicio de clase.

## IX. BIBLIOGRAFÍA

Landau, S., & Everitt, B. (2004). *A Handbook of Statistical Analyses using SPSS*. Boca Raton, FL: Chapman & Hall.

## **X. POLÍTICAS SOBRE EL PLAGIO**

Para la corrección y evaluación de todos los trabajos del curso se va a tomar en cuenta el debido respeto a los derechos de autor, castigando severamente cualquier indicio de plagio con la nota CERO (00). Estas medidas serán independientes del proceso administrativo de sanción que la facultad estime conveniente de acuerdo a cada caso en particular. Para obtener más información, referirse a los siguientes sitios en internet:

- [www.pucp.edu.pe/documento/pucp/plagio.pdf](http://www.pucp.edu.pe/documento/pucp/plagio.pdf)
- [http://www.pucp.edu.pe/puntoedu/images/documentos/institucionales/guia\\_pucp\\_para\\_el\\_registro\\_y\\_citado\\_de\\_fuentes\\_documentales\\_2009.pdf](http://www.pucp.edu.pe/puntoedu/images/documentos/institucionales/guia_pucp_para_el_registro_y_citado_de_fuentes_documentales_2009.pdf)

## **XI. ACTUALIZACIÓN**

Marzo 2017