

**CURSO:** CC50 – ADMINISTRACION DE LA INFORMACION

**CLASE:** TA1 #3 (TRABAJO ACADEMICO GRUPAL)

**TEMA:** ANÁLISIS EXPLORATORIO DE UN CONJUNTO DE DATOS EN R/RSTUDIO

**PROFESOR/A:** Ing. PATRICIA REYES SILVA

---

### **Objetivo**

Realizar un análisis exploratorio de un conjunto de datos, creando visualizaciones y obteniendo inferencias básicas utilizando R/RStudio como herramienta de software.

### **Competencias**

Serán evaluadas las competencias, según rubrica adjunta EA-RUBCC50-2101-CC51.

### **Conjunto de Datos**

El conjunto de datos motivo de análisis se denomina: **Hotel booking demand** y se obtiene desde este enlace: <https://www.kaggle.com/jessemostipak/hotel-booking-demand>

En este conjunto de datos se recopilan datos de un hotel urbano y otro de tipo resort. Incluye información de cuándo se realizó la reserva, la duración de la estadía, la cantidad de espacios de estacionamiento disponibles, cantidad de huéspedes adultos, niños y / o bebés, entre otros datos.

El conjunto de datos original proviene del documento: [Hotel booking demand datasets](#)

### **Documento Entregable**

El grupo de estudiantes entregará un único documento de tipo PDF, desarrollando los siguientes temas en este orden propuesto:

1. CASO DE ANALISIS

Explicación sobre el origen de los datos (procedencia de los datos, autor/autores, fecha, país, etc.)

Casos de uso aplicables (describir, por ejemplo: ¿Para quién sería importante el análisis de estos datos?, ¿Quien o quienes se benefician?)

2. CONJUNTO DE DATOS (DATA SET)

❖ Descripción de la estructura de los datos

3. ANÁLISIS DE DATOS EXPLORATORIO

Descripción de instrucciones ejecutadas en R/RStudio y resultados obtenidos para:

- ❖ CARGAR DATOS
- ❖ INSPECCIONAR DATOS
- ❖ VISUALIZACION GRAFICA (mínimo tres visualizaciones)

#### 4. CONCLUSIONES PRELIMINARES

##### **Consideraciones adicionales**

- En cuanto al PREPROCESAMIENTO DE DATOS (p.e. limpieza/depuración de datos), se identifican las tareas o técnicas a aplicar mas no se ejecutan.
- Se recomienda se familiaricen en el uso de Github.com para mantener un control del código, datos y resultados.
- Se evaluará el orden dentro de la organización del documento, así como la correcta redacción y gramática.

**Fecha límite de entrega:** 05/05/2021

Nombre TA1-3-RUBCC50-2101-CC51

Descripción Rubrica correspondiente a la tercera TA1 (Tarea academica) del curso Administracion de la informacion CC50-2101-CC51

Detalle de las rúbricas

Niveles de rendimiento

Criterios	No logrado	En proceso	Logrado
<b>Origen y Adquisición de los datos</b>  <b>Ponderación 25,00%</b>	<b>10,00 a 30,00 %</b> Reconoce la existencia de un conjunto de datos por analizar.	<b>30,00 a 70,00 %</b> Entiende y describe la estructura de un conjunto de datos.	<b>70,00 a 100,00 %</b> Investiga y describe la fuente del conjunto de datos y lo pone en contexto. Identifica la Fase del Ciclo de datos que se desarrolla
<b>Análisis exploratorio inicial</b>  <b>Ponderación 40,00%</b>	<b>10,00 a 30,00 %</b> Utiliza una herramienta de software para analizar un conjunto de datos.	<b>30,00 a 70,00 %</b> Es capaz de realizar operaciones básicas con un conjunto de datos.	<b>70,00 a 100,00 %</b> Hace uso de visualizaciones básicas a fin de encontrar inferencias a partir de los datos observados. Identifica que tareas de pre-procesamiento de datos son necesarias aplicar.
<b>Publicacion preliminar de resultados</b>  <b>Ponderación 35,00%</b>	<b>10,00 a 30,00 %</b> Describe en un documento la estructura de un conjunto de datos.	<b>30,00 a 70,00 %</b> Describe cada una de la operaciones aplicadas sobre un conjunto de datos.	<b>70,00 a 100,00 %</b> Describe cada una de las visualizaciones, y en base a ellas concluye señalando las inferencias observadas.

[Ver elementos asociados](#)