

**Máster universitario de Ciencia de Datos**

**Práctica 3**

**Diseño y uso de bases de datos analíticas – Explotación de datos.**

Autor:

Mario Ubierna San Mamés

|  |
| --- |
| Índice de Contenido |

[Índice de Contenido 3](#_Toc72748456)

[Índice de tablas 4](#_Toc72748457)

[Índice de ilustraciones 5](#_Toc72748458)

[1. Introducción 6](#_Toc72748459)

[1.1. Presentación 6](#_Toc72748460)

|  |
| --- |
| Índice de tablas |

**No se encuentran elementos de tabla de ilustraciones.**

|  |
| --- |
| Índice de ilustraciones |

**No se encuentran elementos de tabla de ilustraciones.**

|  |
| --- |
| Introducción |

## Presentación

A partir de la solución oficial de la segunda práctica (PRA2) el estudiante debe implementar los cubos multidimensionales necesarios para la explotación de la información y el posterior análisis de datos. De este modo se facilitará la toma de decisiones de los usuarios potenciales.

Así pues, esta actividad tiene el objetivo de implementar un modelo multidimensional online analytical processing (OLAP) para el análisis multidimensional de la información con el fin de responder a las preguntas definidas en el análisis de requerimientos.

Adicionalmente, se facilitará junto a este enunciado el fichero «*export\_DW\_COVID.sql*», que contiene los scripts de generación y carga de todas las tablas planteadas en la solución, para que el estudiantado pueda partir de la misma base.

## Descripción

Más concretamente esta tercera parte del caso práctico consiste en diseñar un modelo OLAP para el análisis multidimensional de la información disponible en el almacén de datos que permita dar respuesta a las siguientes cuestiones:

* Análisis de las provincias con mayor porcentaje de movilidad según datos móviles.
* Análisis del porcentaje de la población que evitaba las aglomeraciones según la comunidad autónoma.
* Análisis del promedio de sanciones por habitante.
* Evolución de las llamadas de urgencia al 112 en Cataluña por tipología de llamada.
* Análisis de las llamadas de urgencia frente al porcentaje de la población que evitaba las aglomeraciones entre los meses comprendidos entre marzo y junio de 2020 en Cataluña, desglosado por provincia.
* Determinación del día de la semana con menor número de denuncias.
* Análisis de las diez fechas (*top ten*) con mayor número de llamadas de urgencia al 112 con tipología de tránsito registrada, tanto en época de *COVID* como antes.

|  |
| --- |
| Creación del modelo OLAP |

En este apartado vamos a crear tanto el proyecto en *Visual Studio*, como observar el origen de datos, la creación de los cubos y la decisión sobre jerarquías/dimensiones/atributos relacionados.

## Creación del proyecto