Sergio Hervás Aragón

Sergiohervas9@gmail.com

Brief description

This project is a case study of a Business Intelligence (BI) solution for a telecommunications multinational. It involves designing the data analysis system architecture, modeling the database, loading extracted data, and performing data exploitation to support decision-making.

Final practice

The job of the data scientist

Index

Content

[1. Analysis of initial data and proposed data model. 2](#_Toc188477352)

[1.1. Understanding initial data 2](#_Toc188477353)

[1.2. Data processing 3](#_Toc188477354)

# Analysis of initial data and proposed data model.

## Understanding initial data

Dataset data:

* **Fecha de creación**: Specifies the time when the ticket is created. It will be used to perform the temporal analysis.
* **Numero de incidente**: Unique ticket identifier.
* **Descripción**: Brief description of the request or incident.
* **Servicio**: The different services that the company provides to its users and clients.
* **Tipo de servicio**: Specify whether it is a request or an incident:
  + **Incidencias**: Restoration of infrastructure or restoration of service to user.
  + **Peticiones**: Service request by user
* **Prioridad**: Specifies the priority with which the ticket should be treated.
  + There are four priorities, ordered from most important to least: Critical, High, Medium, and Low.
  + Only incidents can have critical priority and should only be considered in production environments.
* **Estado**: Indicates the status of the ticket.
  + A ticket is considered finished if it is in Closed or Resolved status.
  + A ticket is considered open if it is pending or assigned.
  + Cancelled tickets do not count.
* **Torre**: It is a functional grouping of services. Each service belongs to a management tower.
* **Entorno**: Indicates the environment that the ticket affects. Production environments are all those that contain a PRO.
* **Estado cumplimiento**: Indicates whether the ticket has been resolved within the time required by the client or not.
  + Within the service objective: ticket resolved appropriately
  + Service objectives not met: ticket resolved late

**Duración días**: Number of days it took to resolve the ticket.

## Data processing

An analysis of the data is performed to verify nulls, and/or data that may be erroneous, in a first instance of .isnull():

Imagen que contiene texto, grande

Descripción generada automáticamente

With the empty data, proceed by columns to:

* Descripción: Nulls are filled with the value “No description”.
* Servicio: Nulls are filled with the value “Unspecified”.
* Torre: Nulls are filled with the value “Unspecified”.
* Entorno: Nulls are filled with the value “Unspecified”.
* Estado cumplimiento SLA: Nulls are filled with the value “Unspecified”.

Imagen que contiene texto, tabla, calle, grande

Descripción generada automáticamente