# Caso - Automatización de métricas

**Problemática**

Actualmente el equipo de Awareness Security de Seguridad Informática, realiza el seguimiento del estado de capacitaciones de ciberseguridad de los colaboradores internos y externos (siendo los colaboradores contratados de proveedores) de todo Mercado Libre.

Para poder tomar decisiones y armar la estrategia de awareness de este año, se necesita obtener la métrica de colaboradores capacitados y una distinción del estado de cada unidad de negocio, también conocida como BU, y su evolución en el último año.

**Objetivo**

Armar un script en Python, Go o Java que pueda consumir los datos de estado de capacitaciones de los colaboradores (internos y externos) desde una base de datos relacional,elabora un cálculo mensual del porcentaje de finalización de la capacitación en ciberseguridad correspondiente al último año y súbelo a una plataforma o aplicación donde los usuarios finales puedan acceder y visualizar fácilmente estos datos. elabora un cálculo mensual del porcentaje de finalización de la capacitación en ciberseguridad correspondiente al último año y súbelo a una plataforma o aplicación donde los usuarios finales puedan acceder y visualizar fácilmente estos datos.. La plataforma destino puede ser DataFlow, Pentaho, Apache Nifi, o alguna plataforma de características similares a la que puedan otorgarse permisos para visualizar los resultados de la ejecución del script.

Se debe entonces incluir:

* Crear la instancia de base de datos relacional con un conjunto de datos (no usar únicamente los datos del ejemplo dado) a partir del cual se pueda realizar un cálculo del porcentaje por mes en el último año.
* Recorrer la fuente de datos de usuarios y capacitaciones para obtener los datos necesarios para el cálculo de la métrica.
* Procesar los datos y realizar el cálculo del dato porcentual de usuarios capacitados por cada BU (unidad de negocio).
* Cargar las métricas obtenidas en un histórico (puede ser una tabla de datos relacionales también, con destino en otra base de datos). En caso de considerarlo apropiado, se pueden agregar datos al histórico.

Se debe considerar que sólo interesa la métrica de los usuarios activos que completaron la capacitación en cada mes (descartar usuarios inactivos a pesar de haber realizado el curso previamente). Además, para realizar la segmentación por unidad de negocio, tener en cuenta que las mismas son: *Mercado Libre*, *Mercado Pago* y *Mercado Envíos*.

Se muestra a continuación un *ejemplo* de la estructura de las tablas de la base de datos relacional que tiene los datos de usuarios y capacitaciones a ser consumidos:

TABLA DE USUARIOS

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ID | USERNAME | START\_DATE | END\_DATE | BUSINESS\_UNIT | MANAGER | LAST\_UPDATE | IS\_EXTERNAL |
| 123 | ext\_jperez | 12-01-2024 | 18-04-2024 | Mercado Libre | lsalme | 18-04-2024 | true |
| 124 | lsalme | 15-03-2024 | null | Mercado Pago | ifrancis | 01-01-2025 | false |
| 125 | ifernandez | 12-09-2024 | null | Mercado Envío | asuarez | 12-09-2024 | false |

TABLA DE CAPACITACIONES POR USUARIO

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | FK\_USERNAME | FK\_TRAINING | END\_DATE | ASSIGNMENT\_DATE |
| 5 | lsalme | 2 | null | 15-03-2024 |
| 87 | ifernandez | 2 | 15-09-2024 | 12-09-2024 |
| 10 | ext\_jperez | 2 | 18-01-2024 | 12-01-2024 |
| 11 | ext\_jperez | 5 | 19-01-2024 | 12-01-2024 |

TABLA DE CAPACITACIONES

|  |  |  |  |
| --- | --- | --- | --- |
| ID | NAME | LINK | CREATION\_DATE |
| 2 | Ciberseguridad | https://meli.ciberseguridad.training/.com | 01-12-2023 |
| 3 | Código de Ética | https://meli.etica.training/.com | 03-12-2023 |
| 5 | Onboarding | https://meli.onboarding.training/.com | 10-10-2022 |

## **Entregables**

Se espera:

* Código fuente (link al repositorio en Github)
* Armar un README.md en el repositorio del código fuente que contenga las instrucciones sobre cómo realizar la ejecución del programa (incluida cualquier aplicación o librería a instalar para el correcto funcionamiento del programa)
* Descripción de la aplicación realizada, descripción de elección de herramientas, problemas y soluciones con los que se encontró al realizar la misma.
* Credenciales para acceso o permisos asignados a la plataforma de ETL utilizada para aplicar el script y cargar el histórico, para su visualización.

## **Consideraciones**

A tener en cuenta:

* Se podrán crear todas las funciones complementarias que se consideren necesarias para un correcto funcionamiento de la aplicación.
* Se recomienda modularizar y aplicar buenas prácticas de programación para un mejor entendimiento del código.
* Se considerará durante la evaluación el manejo de errores, de autenticación y la documentación del código.
* Se debe generar un dataset con la estructura de datos presentada de 200 usuarios, con variaciones en los datos para poder obtener al menos un valor por métrica indicada (un valor para cada unidad de negocio entre 10% y 100% de completitud de cursos).