

Prueba de conocimientos - MIS Analyst

Sección 1:

La compañía Innova S.A. desea organizar sus datos que por 10 años han sido almacenados en Excel, archivos csv y en algunos casos en plantillas predefinidas. Los datos están en archivos separados y no se encuentran estandarizados.

Cada área ha venido recolectando sus datos de manera independiente para su uso específico y suele ser común que se encuentren deficiencias en la calidad de los datos debido a la manualidad en el tratamiento de estos.

Los archivos crecen en un volumen de 40.000 registros por mes y la base de datos que se quiere obtener será de uso exclusivo para el consumo de futuros reportes en Power BI y llamadas desde modelos de datos en Excel por el área de MIS, por lo tanto su consumo no es masivo.

La gerencia decide contratarlo como Chief Data Officer con el objetivo de ejecutar un proyecto desde cero, donde el resultado final sea un data warehouse que facilite la disponibilidad y organización de los datos para futuros proyectos de BI y BA.

Por favor proporcione un resumen en el que explique las decisiones que tomaría, bajo los siguientes parámetros:

1. ¿Qué plan de acción propondría para ejecutar el proyecto desde su inicio hasta la culminación de este?
2. ¿De las opciones disponibles en el mercado, ¿Qué infraestructura para el almacenamiento de datos propondría y por qué?
3. ¿Cuáles proyectos de BI y BA propondría?, ¿Por qué?

Sección 2:

El área de operaciones requiere llevar el control de la ocupación y facturación del arriendo del campus.

Actualmente el área no cuenta con un sistema de información que le permita rápidamente generar las alarmas o consultar históricos de ocupación y facturación con un óptimo nivel de detalle, esto ocasiona que se pierda oportunidad y calidad en el proceso y genera un riesgo financiero y operativo.

Actualmente se cuentan con las siguientes fuentes:

1. **Centros de Costos:** Codifica y agrupa las operaciones por centro de costo y grupo o área que operan en la compañía. Su actualización varía de acuerdo con la creación o modificación de áreas en la compañía.
2. **Canon de Arriendo y capacidad instalada (Tabla Arriendos Mes...):** Especifica los cánones de arriendo del campus, capacidad de las sillas y área total (mt2) por edificio y piso. Estos valores se actualizan cada año, por lo que siempre se usa la última actualización que corresponda.

Por ejemplo, si la tabla contiene dos actualizaciones, una para enero 2018 y otra para enero 2019 y se quiere consultar el valor del arriendo para junio de 2018, el valor que se debe tomar es el más reciente previo a la fecha de consulta, es decir, enero 2018.

3. **Áreas comunes:** Especifica el área (mt2) de cada zona común en cada edificio y piso. Esta área hace parte del área total del campus, por lo que el área ocupada por las sillas junto con la sumatoria de áreas comunes del edificio y piso debe ser igual al área total.

- Sillas Usadas (Tabla Seats Input Mes...):** Son reportes de uso de las sillas por centro de costo en cada edificio y piso los cuales se actualizan mensualmente (el primer día de cada mes). Es útil para determinar el canon que cada operación debe pagar. Entre más sillas use, mayor área ocupa y por lo tanto mayor es el canon a pagar.
- Áreas Comunes usadas (Tabla Áreas Comunes Mes...):** Son reportes de uso por centro de costo en cada edificio y piso. Al igual que con las sillas, se utiliza para determinar el área de zonas comunes usado por cada operación. A mayor espacio ocupado por la operación, mayor el canon a pagar por este concepto.

Total tablas: 10

Actividad

Desarrolle un proceso ETL y presente evidencia de su ejecución. Entregue una descripción de los pasos realizados para analizar y procesar la data. En caso de usar un programa especializado envíe un tutorial con las capturas de pantalla y su descripción.

El modelo de datos debe permitir consultar el siguiente reporte:

- Tabla histórica (Dic 2018 a Feb 2019) con el valor del canon de arrendamiento de las áreas comunes y sillas usadas. Separado también por edificio, piso y centro de costo con su respectivo total. Agregar cantidad de sillas usadas por centro de costo.

Nota: Los valores mostrados en la tabla de ejemplo son a los que se debe llegar con la data suministrada.

Ejemplo:

Month-Year	Building	Floor	Cost Center Id	Cost Center Standard List	Rent Seats	Rent Common Areas	Total Rent	Seats Used
Dec-18	Choachi	Floor 1	1001-001	Company	84,470,597		84,470,597	20
Dec-18	Choachi	Floor 1	1004-031	Infraestructure		5,213,660	5,213,660	
Dec-18	Choachi	Floor 2	1001-001	Company	23,165,379		23,165,379	104
Dec-18	Choachi	Floor 2	1004-031	Infraestructure		6,300,627	6,300,627	
Dec-18	Choachi	Floor 2	7004-181	Compliance	9,577,993		9,577,993	43
Dec-18	Choachi	Floor 2	7077-011	FS Codes Payable	52,344,847		52,344,847	235
Dec-18	Choachi	Floor 2	7077-017	FS Ops&S Acct - Indirect	8,018,785		8,018,785	36
Dec-18	Choachi	Floor 3	1001-001	Company	12,745,783		12,745,783	58

Entregables:

- Documento en pdf con capturas de pantalla sobre el proceso ETL.
- Los reportes con los resultados descritos en la tabla de ejemplo para los meses de los que se tiene información de sillas y áreas usadas.
- Cualquier archivo o documento con el que desee complementar la prueba.

Nota:

- Por favor confirme tan pronto reciba este email.
- Esta prueba mide la capacidad de análisis y conocimientos que tiene para procesar información no estructurada y sus habilidades para proponer soluciones eficientes. Por lo tanto, se recomienda solo hacer la transformación de la data (en caso de requerirlo) usando herramientas que permitan el procesamiento en lote.
- En caso no terminar completamente la prueba con el plazo propuesto, por favor enviar el progreso obtenido.

Éxitos – MIS Team