

Preguntas

1. Cuáles son los patrones típicos encontrados en los datos de las tarjetas de proximidad? Cómo se ve un día típico en los empleados de GASTech.
2. Describa hasta diez de los patrones más interesantes que aparecen en los datos de edificio, describa qué es lo más notable de dicho patrón y explique qué significa.
3. Describa hasta diez anomalías notables o eventos inusuales que usted ve en los datos. Priorice esos inconvenientes qué tan peligrosos o serios son para las operaciones en el edificio.
4. Describa hasta cinco relaciones que se pueden ver entre las tarjetas de proximidad y los datos de los elementos del edificio, si encuentra alguna relación causal, por ejemplo, un evento del edificio o una condición que lleva a que el personal cambie su comportamiento en las operaciones, describa cuál fue la causa y el efecto, cuál fue la evidencia que soporta dicho evento y su nivel de confianza sobre la afirmación de la relación.

WHAT

Para este reto se tienen varios archivos como fuente, se realizará la descripción de cada uno de estos.

Empleados

Este dataset es una tabla que tiene como ítems cada empleado de la empresa GASTech; contiene los siguientes atributos:

- **Last name:** Atributo categórico con el apellido del empleado.
- **First name:** Atributo categórico con el nombre del empleado.
- **Department:** Atributo categórico con el departamento donde trabaja el empleado.
- **Office:** Atributo ordenado secuencial que indica la oficina donde el empleado trabaja.

Sensores de Hazium

Este dataset es una tabla que tiene como ítems el registro de cada sensor para hazium, teniendo a disposición 4 sensores que están separados en 4 archivos. Para tratar un solo dataset, se agrega el nombre del archivo como un atributo categórico y se genera un único archivo, el cual contiene los siguientes atributos:

- **Timestamp:** Atributo ordinal secuencial que contiene la fecha y hora en la que fue registrada la lectura.
- **Hazium concentration:** atributo cuantitativo secuencial que contiene la medición de la concentración de hazium en el ambiente.
- **Sensor:** Atributo categórico que indica el sensor con el cual se realizó la medición. Este atributo codifica en su valor tanto el piso como la zona del edificio que se encuentra

mediendo, no se separa en 2 atributos categóricos por contar con medidas de sólo 4 sensores y no vale la pena comparar niveles del piso 1, por ejemplo.

Sensores de proximidad

Este dataset es una tabla que contiene como ítems el registro de cada vez que un empleado pasa a 4 piés de de un sensor de proximidad; contiene los siguientes atributos:

- **Timestamp:** Atributo ordinal que contiene la fecha y hora del evento.
- **Type:** Atributo categórico con el tipo de sensor que realizó la medición.
- **Prox-id:** Atributo categórico con el identificador del sensor que realizó la medición.
- **Floor:** Atributo categórico con el piso donde se encuentra el sensor.
- **Zone:** Atributo categórico con la zona donde se realizó la medición.
- **X:** Atributo cuantitativo secuencial con la posición horizontal del sensor en el edificio, cuando la medición haya sido realizada por un sensor móvil.
- **Y:** Atributo cuantitativo secuencial con la posición vertical del sensor en el edificio, cuando la medición haya sido realizada por sensor móvil.

Sensores del edificio

Este dataset es una tabla que contiene como ítems el registro de la medición de los sensores en cada una de las áreas de ventilación, calefacción y aire acondicionado; contiene los siguientes atributos:

- **Timestamp:** Atributo ordenado secuencial con fecha y hora del momento en que se realizó la medición.
- **Drybulb temperature:** Atributo cuantitativo secuencial que mide la temperatura del aire.
- **Water heater tank temperature:** Atributo cuantitativo divergente que mide la temperatura del agua del tanque de la calefacción.
- **Water heater gas rate:** Atributo cuantitativo divergente que mide la proporción de gas en el calentador.
- **Supply Side Inlet Mass Flow Rate:** Atributo cuantitativo divergente que mide la proporción de aire que ingresa al edificio.
- **Supply Side Inlet Temperature:** Atributo cuantitativo divergente que mide la temperatura del aire que ingresa al edificio.
- **Supply Side Outlet Temperature:** Atributo cuantitativo divergente que mide la temperatura del aire que sale del edificio.
- **HVAC Electric Demand Power:** Atributo cuantitativo secuencial que mide la cantidad de energía gastada por el sistema del edificio.
- **Total Electric Demand Power:** Atributo cuantitativo secuencial que mide la cantidad de energía gastada por el edificio en total.
- **Loop Temp Schedule:** Atributo cuantitativo secuencial que indica los ciclos en que se controla la temperatura del edificio.
- **Water Heater Setpoint:** Atributo cuantitativo divergente que indica la temperatura a la cual fue puesta el calentador de agua.
- **DELI-FAN Power:** Atributo cuantitativo secuencial que indica la cantidad de energía consumida por el ventilador de la cafetería.

- **Pump Power:** Atributo cuantitativo secuencial que indica la cantidad de energía consumida por la bomba de agua.

WHY

Sensores de Hadium

- T1. Buscar valores atípicos particularmente altos de concentraciones de Hadium que puedan ser potencialmente peligrosos para las personas en el edificio (**Browse, Outliers**).
- T2. Identificar tendencias en las mediciones de Hadium entre las diferentes zonas y pisos del edificio (**Identify, Trends**).
- T3. Descubrir características de los datos según el comportamiento de estos (**Discover, Features**).
- T4. Derivar los datos para unificar el dataset y codificar los sensores por zona y nivel del edificio (**Derive, Features**).

Temperatura general del edificio

- T1. Descubrir características de los datos según los registros en el tiempo (**Discover, Features**).
- T2. Identificar tendencias en las temperaturas que se encuentran registradas en el edificio (**Identify, Trends**).

Consumo de energía en el edificio

- T1. Descubrir características de los datos según los registros en el tiempo (**Discover, Features**).
- T2. Identificar tendencias en el consumo de energía registrado en el edificio. (**Identify, Trends**).

Tarjetas de proximidad estáticas

- T1. Descubrir características de los datos según los registros por zona (**Discover, Features**).
- T2. Identificar tendencias en las visitas de las diferentes zonas en el edificio (**Identify, Trends**).
- T3. Explorar las diferentes zonas para encontrar cuáles son las más visitadas por piso (**Explore, Features**).
- T4. Derivar los datos para obtener el conteo por zona durante las 2 semanas de los datos (**Derive, Features**).

HOW

Hazium

Para la visualización de los niveles de Hazium se emplea el modismo de **line chart**, el cual tiene como marca líneas. En el eje **Y** se ubica el atributo cuantitativo del **nivel de concentración del químico**, empleando la posición como canal, obteniendo una manera efectiva de expresar este atributo. En el eje **X** se ubica el atributo ordenado secuencial del **tiempo de medición del sensor**, empleando la posición y alineación, obteniendo así el canal más efectivo que, junto con la alineación, permite comparar e identificar tendencias. Por último, se tiene el atributo categórico **código**, el cual hace referencia a nivel y zona del edificio; fue codificado mediante una escala de colores empleando tonos para diferenciar los diferentes sensores.

Temperatura general del edificio

Para la visualización de las temperaturas en el edificio se emplea el modismo **line chart**, teniendo en el eje **Y** el atributo cuantitativo divergente **temperatura**, en el eje **X** se tiene el **tiempo de la medición de temperatura en el sensor**; se usa la posición tanto horizontal como vertical para expresar los valores de temperatura en el tiempo, teniendo así una forma de identificar tendencias como el canal más efectivo. También se emplea el color en una escala tonal para identificar los diferentes sensores, siendo un atributo categórico.

Consumo de energía del edificio

Se emplea el modismo **line chart**, teniendo en el eje **Y** el atributo cuantitativo secuencial **consumo de energía** y en el eje **X** el atributo ordinal secuencial del **tiempo**; para la posición, tanto horizontal como vertical expresando los valores del consumo de energía en el tiempo, obteniendo de esta manera un canal efectivo; se emplea **align** para visualizar tendencias, así como **yuxtaposición** para separar en dos visualizaciones los sensores que tienen valores considerablemente más pequeños a los demás y así emplear una escala diferente en el eje **Y**, teniendo la capacidad de identificar mejor los patrones de los datos.

Tarjetas de proximidad estáticas

Se emplea el modismo **bar chart** agregando small multiples para realizar comparaciones entre los diferentes pisos del edificio, ubicado en el eje **Y** el atributo cuantitativo de **cantidad de registros del sensor** y en el eje **X** se ubica el atributo categórico de la **zona donde el sensor realizó la medición**. Se emplea **yuxtaposición** para visualizar cada piso y sus diferentes zonas; así mismo, la visualización cuenta con separación entre las zonas para poder identificarlas y compararlas, como marcas se tienen líneas y como canal se emplea la posición vertical expresando la cantidad de visitas recibidas en una zona.