**Guía para nombrar los archivos de la automatización**

Este documento tiene como objetivo proporcionar pautas claras y prácticas recomendadas para nombrar archivos y carpetas de manera coherente. Al seguir estas directrices, no solo mejoraremos la legibilidad y la comprensión de nuestros sistemas automatizados, sino que también facilitaremos la colaboración y el mantenimiento a lo largo del tiempo.

La correcta organización de archivos y carpetas es importante por la gestión eficiente y la comprensión clara de cualquier proyecto. En la automatización es importante tener un orden en la estructura de archivos y carpetas, ya que esto no solo facilita la identificación y el acceso rápido a los recursos relevantes, sino que también promueve la coherencia en todo el sistema.

A continuación, establecemos los criterios de orden y designación de los nombres de las carpetas y archivos.

**1. Organización de carpetas**

La estructura de carpetas debe ser lógica y fácil de entender para cualquier persona que acceda a los archivos. Es preciso tener una jerarquía clara y consistente. Se recomienda tener una carpeta principal que aloje dentro de ella las carpetas que funcionarás como insumo y salida de los proyectos.

Para mantener el código organizado, es necesario:

Estructura Modular: Divide el código en módulos o paquetes lógicos. Cada módulo debe tener su propia carpeta y archivo principal.

Subcarpetas Descriptivas: Utiliza subcarpetas con nombres significativos para categorizar el código.

* **Carpeta Principal:** Se deberá crear una carpeta con el nombre del proceso que se está automatizando, y con la especificación del piloto. Esta carpeta alojará en orden todas las subcarpetas que contendrán los diferentes archivos según corresponda.

Ejemplo: **EMMET\_Automatización**

* **Carpetas secundarias:** El proyecto de automatización deberá contar con carpetas secundarias que ayuden a mantener en orden los diferentes archivos que se usan.

Algunas consideraciones clave para manejar:

• Versionado: Si los datos cambian con el tiempo, considera la posibilidad de incluir versiones de datos en subcarpetas con nombres como v1, v2, etc.

• Evitar Duplicación: No dupliques datos en la carpeta de datos si no es necesario. Utiliza enlaces simbólicos o referencias si es posible.

• Incluye diccionarios o información de la metadata

**data:** En esta carpeta se almacenan los archivos de entrada para el proyecto. Contendrá los archivos globales, conjuntos de datos u otro tipo de archivo que requiera los códigos de automatización en cada fase. Adicionalmente, contiene subcarpetas por años, y estas a su vez alojan las carpetas especificadas por meses, solo con las tres iniciales de cada mes.

Ejemplo: Si es de su interés**,** observar los archivos del mes de noviembre del año 2023, está será la dirección de ruta, que deberá tomar:

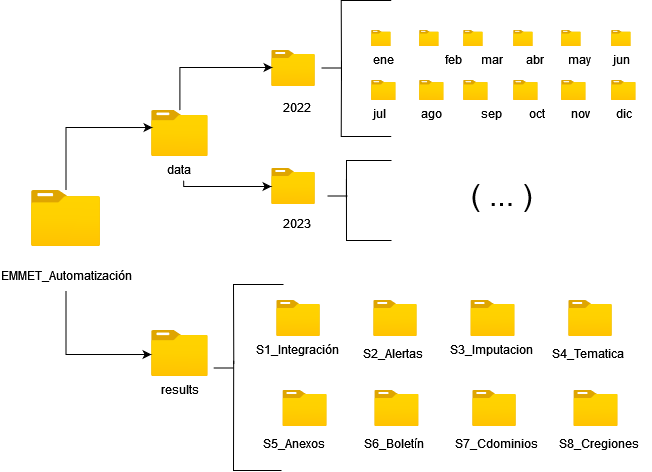
**EMMET\_Automatización > data > 2023 > nov**

**results:** Esta carpeta alojará en ella las carpetas de los diferentes procesos que se llevaron a cabo en la automatización. Estas deberán nombrarse de acuerdo a los procesos que recrean, con el prefijo de “S” y el respectivo número según el orden, permitiendo que se alojen de forma organizada (S1\_Integración, S2\_ Alertas, S3\_ Imputación, … ) :

Ejemplo: Si desea ingresar a los archivos alojados en la carpeta S2\_Alertas, está será la dirección de ruta, que deberá tomar:

**EMMET\_Automatización > results > S2\_Alertas**

A continuación, se anexa una imagen para describir gráficamente el contenido de las carpetas y como deben nombrar cada una de estas:



**2. Nombre de archivos:**

Los nombres de archivos deben ser informativos y descriptivos para que los usuarios puedan identificar rápidamente su contenido. Aquí hay algunas pautas a considerar:

• Brevedad y Claridad: Utilizar nombres concisos pero descriptivos que reflejen el contenido del archivo.

• Evitar Caracteres Especiales: Limitar el uso de caracteres especiales y espacios en blanco en los nombres de archivo para evitar posibles problemas de compatibilidad.

• Consistencia: Mantener la consistencia en la estructura y el formato de los nombres de archivo en toda la organización. No duplicar los nombres creando archivos tales como “Integración(1)”, “Integración -Copia”, etc

Adicionalmente, debe ser específico en cuanto al uso que se le puede dar al archivo, es decir, a qué proceso pertenece, y de cuál puede ser insumo:

Aclaramos las especificaciones que deben tener en cuenta:

* Nombre del proceso
* Base a la que se hace referencia
* Proceso del cual es resultado
* Iniciales del mes al cual pertenece
* Año al que hace referencia

Ejemplo: EMMET\_PANEL\_trabajo\_original\_nov2023.

Archivo resultante del proceso de integración de bases en donde se trabaja con la base panel del proceso ejecutado para el mes de noviembre del año 2023.

**3. Uso de la librería:**

Este paquete tiene como objetivo optimizar los tiempos del proceso que se realiza para la entrega de productos de la encuesta mensual manufacturera con enfoque territorial, ( EMMET )

3.1. Instalación

Para instalar la librería “EMMET\_Automatización” se debe utilizar los siguientes comandos en R: remotes::install\_github("NataliArteaga/DANE.EMMET")

library(DANE.EMMET)

3.2. Uso

El primer paso para usar la librería es correr la función inicial, con el siguiente comando

f0\_inicial(directorio,anio,mes)

Donde el directorio será la ruta en donde tenemos alojadas las carpetas de doc y mes y anio son el periodo de interés a correr. Esta función lo que hará será instalar los paquetes necesarios para el correcto funcionamiento de la librería, rectificar si las carpetas necesarias ya están creadas, en caso contrario crearlas en el directorio suministrado.

La ejecución de la siguiente función se realiza ejecutando el siguiente comando:

f1\_integración(directorio,anio,mes)

Esta función integra todas las bases que son necesarias para la ejecución de las demás funciones.

Las de mas funciones requieres de los mismos argumentos y se ejecutan de igual forma:

f2\_alertas(directorio,anio,mes)

Esta función tiene como objetivo identificar las variables dentro de cada establecimiento que pueden llegar a ser casos de imputación, ya sea imputación por deuda o por casos especiales.

f3\_imputación(directorio,anio,mes)

Esta función realiza la imputación por deuda y por casos especiales en variables de capítulo 2 y capítulo 3.

f4\_tematica(directorio,anio,mes)

Esta función construye la base temática. La base temática expone los datos procesados de acuerdo con la metodología de la operación, presentando los datos reales a partir de los nominales. También incluye la información ponderada y agrega variables de la identificación de los dominios por los cuales se publica. Con esto, se puede presentar la base temática como la base final de cuadros. Esta función toma como insumo la base Panel con los casos de imputación aplicados y exporta un archivo con extensión .xlsx.

f5\_anacional (directorio,anio,mes)

Esta función crea el archivo de Anexo Nacional. Utiliza la base de datos Tematica como insumo y genera cada una de las hojas del reporte de Anexo Nacional. Los cuadros de salida o anexos estadísticos de EMMET muestran información complementaria a la registrada en el boletín de prensa con el fin de brindar información a un nivel más desagregado tanto a nivel nacional como desagregado a nivel de departamentos, áreas metropolitanas y principales ciudades del país. Los resultados se presentan con variaciones anuales, año corrido y doce meses, junto con sus respectivas contribuciones, según dominios, por las principales variables que se recolectan en el proceso: Producción (real y nominal), ventas (real y nominal) y empleo, desagregando a su vez esta última variable por área funcional y tipo de contrato. De igual manera se presentan los resultados de sueldos causados y horas totales trabajadas para los dominios nacionales.

f6\_aterritorial(directorio,anio,mes)

Esta función crea el archivo de Anexo Territorial utilizando la base de datos Tematica como insumo. El reporte generado contiene información complementaria a la registrada en el boletín de prensa con el objetivo de proporcionar información a un nivel más desagregado tanto a nivel nacional como a nivel departamental, áreas metropolitanas y principales ciudades del país. Los resultados se presentan con variaciones anuales, año corrido y doce meses, junto con sus respectivas contribuciones, según dominios, por las principales variables que se recolectan en el proceso: Producción (real y nominal), ventas (real y nominal) y empleo. También se incluyen resultados de sueldos causados y horas totales trabajadas para los dominios nacionales

f7\_cdominios(directorio,anio,mes)

Esta función crea un archivo de excel que contiene las variaciones y contribuciones anuales desagregadas por dominio 39. Por cada dominio se crea una hoja en donde se consigna las variaciones y contribuciones de las variables de producción, ventas, empleados, sueldos y horas. Adicionalmente se crea una hoja en donde se presentan los valores de deflactores y se calcula la variación anual y mensuals por cada dominio.

f8\_ cregiones (directorio,anio,mes)

Esta función crea un archivo de excel que contiene las variaciones y contribuciones anuales desagregadas por cada región. Se crea una hoja por cada desagregación en donde se consigna las variaciones y contribuciones de las variables de producción, ventas, empleados, sueldos y horas.

f8\_ boletin(directorio,anio,mes)