

Manejo de datos en R con dplyr

Obtener la información utilizando la librería dplyr en R deberá adjuntar tanto el código utilizado como los resultados obtenidos.

La base de datos "Frijol.xlsx", utilizada para este estudio, contiene información sobre el uso de tecnología y servicios en el sector agrícola. Estos datos han sido obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP) de la SAGARPA, y corresponden a los años agrícolas 2017, 2018 y 2019. Las variables de estudio incluidas son las siguientes:

- Cultivo: Tipo de cultivo (Frijol)
- Variedad: Variedad del cultivo.
- Año: Año agrícola (2015,2018 y2019)
- Estado
- Municipio
- Modalidad: Riego (Requieren el suministro de importantes cantidades de agua a través de diversos métodos artificiales de riego) o Temporal (Depende del comportamiento de las lluvias durante el ciclo de producción y de la capacidad del suelo para captar el agua y conservar la humedad).
- SupSembrada: Superficie en la cual ya se depositó la semilla de cualquier cultivo, previa preparación del suelo (ha).
- SupCosechada: Es el área de la cual se obtuvo producción (ha).
- SupFertilizada: Superficie en la que se aplica, al menos una vez, algún químico para el mejoramiento del desarrollo del cultivo (ha).
- SupMecanizada : Superficie en la que se utiliza, al menos una vez, algún tipo de maquinaria y equipo en la siembra y cosecha de la superficie cultivada (ha).
- SupSanidadVegetal: Superficie sembrada en la cual se realizan actividades orientadas a la prevención, control y erradicación de plagas o enfermedades que afectan los cultivos(ha).
- SupSemillaMejorada: Superficie sembrada con semilla obtenida de una modificación genética natural y que no lleva ninguna alteración transgénica (ha).
- SupSemillaCriolla: Área sembrada en la cual se utiliza semilla recolectada o comprada de cosechas anteriores (ha).

- ProdObt: Es la cantidad de producto que se levantó en determinada superficie cosechada (ton).

Utilizando la librería dplyr obtenga la siguiente información:

1. Obtener una tabla con las variables: estado, municipio, modalidad y variedad donde la modalidad de riego es por temporal para el municipio Acatzingo.
2. Obtener la tabla con las variables: Nombre del estado, Superficie sembrada, variedad, año agrícola y modalidad de riego donde la superficie sembrada es mayor a 20000 hectáreas en el estado de Sinaloa.
3. Obtener el valor máximo y mínimo de la superficie cosechada para la variedad “Frijol bayo”.
4. Obtener una tabla con las variables: superficie sembrada, superficie cosechada y variedad donde la superficie sembrada sea diferente a la superficie cosechada.
5. Obtener el promedio de la superficie sembrada por estado.
6. Obtener maximo y promedio para cada una de las variedades de frijo.
7. Utilizando la función “mutate” crear una nueva variable con Rendimiento Obtenido (Productividad obtenida dividida por la superficie cosechada Ton/ha) para cada una de las observaciones del conjunto de datos.
8. Calcula el número de observaciones en la variable “Superficie Con Sanidad Vegetal” que tienen valores NA (Not Available). (Utiliza la función table con la opción useNA = c(“always”))
9. Por el método de imputación de sustitución por media reemplaza los valores “NA” de la variable “Superficie Con Sanidad Vegetal”.