

Tarea 1 modelos probabilísticos aplicados

Romulo Enrique Troncoso Pacheco

Septiembre 2020

1 Introducción

Los datos que se tomaron de la pagina del INEGI[1] en el apartado de agricultura fueron sobre Superficie cultivada y producción de cultivos anuales y perennes según tipo de agricultura por cultivo con representatividad en la muestra los cuales feuron recolectados en las fechas de 2016 a septiembre de 2017.

Nombre del archivo original: 'ena17agri01.xlsx'.

Con estos datos se pretende realizar una practica de graficación de datos en el lenguaje de programación R.

2 Interpretación

A este cúmulo de datos se le puede interpretar de tal manera en donde se le pueden hacer graficas de tipo circular , grafica de barras y grafica de puntos ya que por ser solo un punto relacionado con una etiqueta.

Estos datos nos muestran la producción de productos agrícolas a lo largo de un tiempo determinado (octubre de 2016 a septiembre de 2017) basándose en tres datos relevantes:

- Hectáreas sembradas.
- Hectáreas cosechadas.
- Toneladas recolectadas.

Estas agrupadas en dos especies

- Cielo abierto.
- Invernadero.

3 Riesgos

- Los valores entre los campos sean desproporcionados evitando ver la grafica de mejor manera.
- Al contar con demasiados cultivos en las graficas se podrá presentar un overlap con los nombres de los cultivos.

4 Objetivos

- Graficar los datos a cielo abierto en relación a hectáreas sembradas, hectáreas cosechadas y producción en toneladas por tipo de cultivo
- Graficar los datos en invernadero en relación a hectáreas sembradas y producción en toneladas por tipo de cultivo
- Graficar la producción total en relación al cultivo y toneladas cosechadas
- Calcular las hectáreas sembradas tanto en aire libre e invernadero
- Calcular el porcentaje de hectáreas sembradas conforme al cultivo
- Graficar en forma circular y en forma de barra los porcentajes
- Calcular las hectáreas perdidas en base a las hectáreas sembradas y las hectáreas cosechadas por cultivo
- Graficar en forma circular y en forma de barra las hectáreas perdidas

5 Prerrequisitos

5.1 Preprocesamiento

Se crearon dos archivos en formato csv ya que el archivo que se descarga de la pagina de INEGI[1] contiene datos basura que afectan a la manipulación del mismo, en los archivos csv se conservo el orden de los datos unicamente separando los datos de cielo abierto y de invernadero.

Nombre de los formatos:

- 'Cielo abierto.csv'
- 'invernadero.csv'

5.2 Paqueteria y versionado

Para la generación de las graficas se utilizo el lenguaje de programación R en su versión 4.0.2[2] en conjunto con el editor visual studio code[3].

Internamente se utilizaron las siguientes librerías:

5.2.1 readr

Descripción:

El objetivo de 'readr' es proporcionar una forma rápida y sencilla de leer datos rectangulares (como 'csv', 'tsv' y 'fwf').

Versión:

1.3.1

Repositorio:

<https://cran.r-project.org/web/packages/readr/index.html>

5.2.2 plotrix

Descripción:

Contiene una gran cantidad de gráficos, diversas funciones de etiquetado, eje y escala de color.

Versión:

3.7.8

Repositorio:

<https://cran.r-project.org/web/packages/plotrix/index.html>

6 Salida

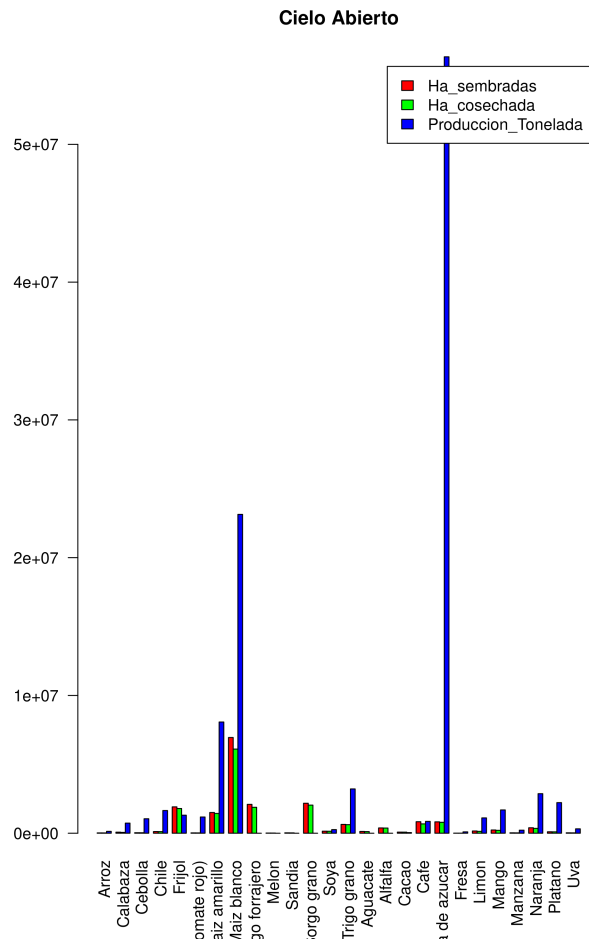


Figure 1: Relación hectáreas sembradas, hectáreas cosechadas y producción en toneladas por tipo de cultivo.

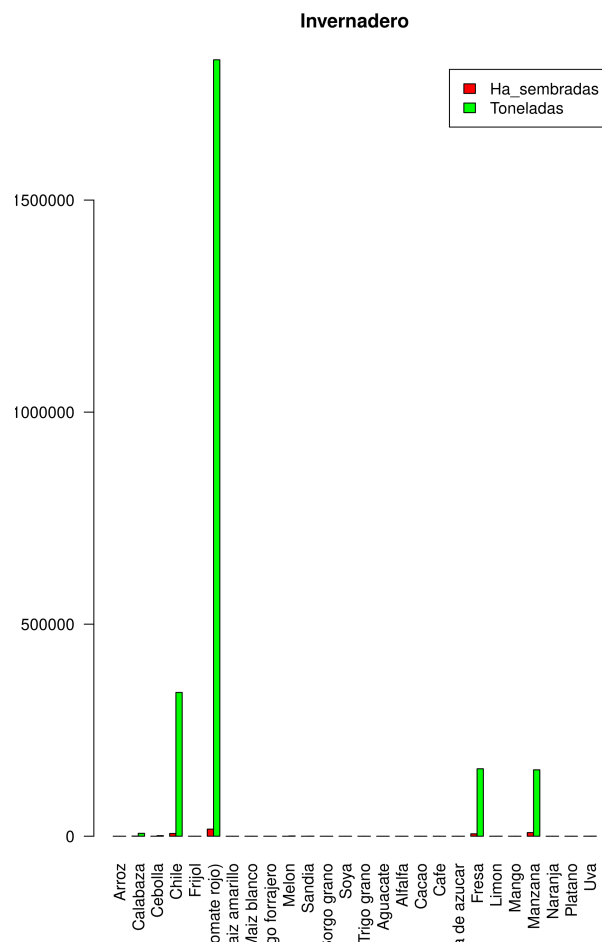


Figure 2: Relación hectáreas sembradas y producción en toneladas por tipo de cultivo.

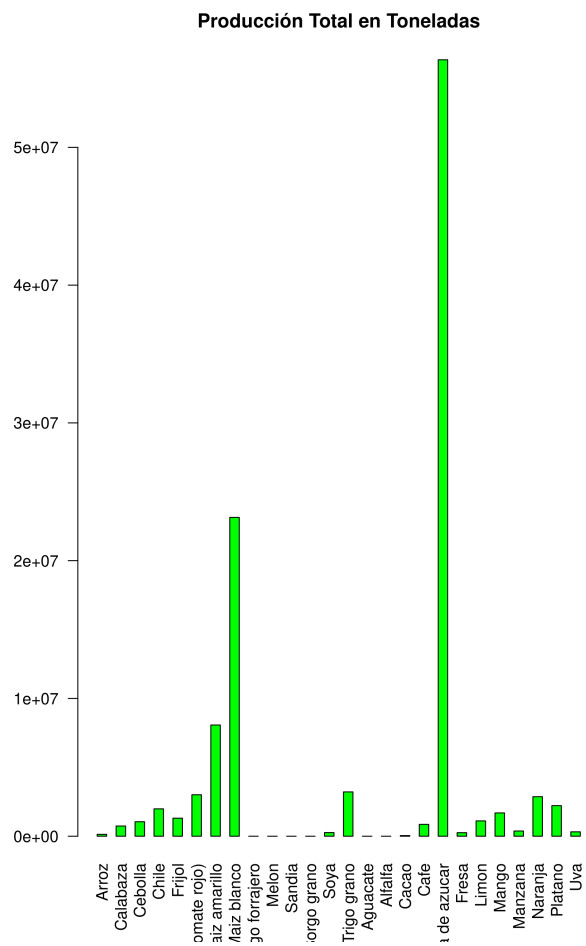


Figure 3: Muestra la cantidad de toneladas cosechadas por cultivo

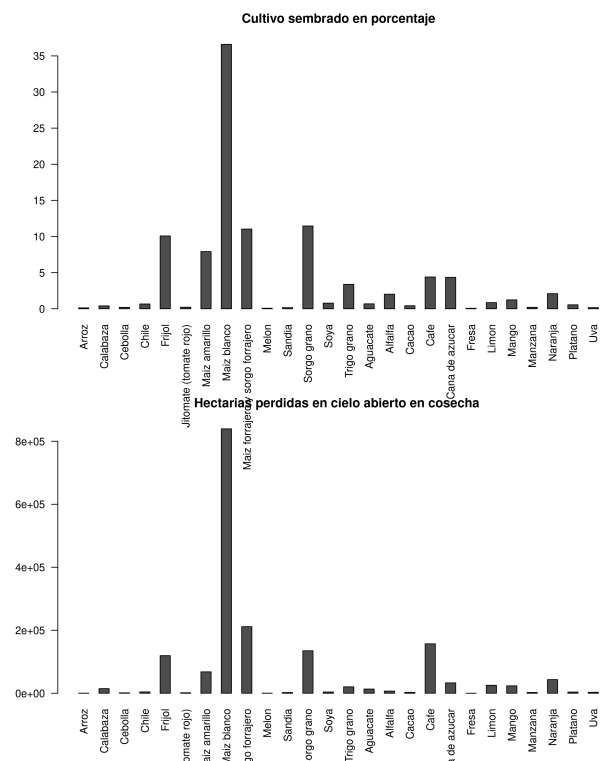
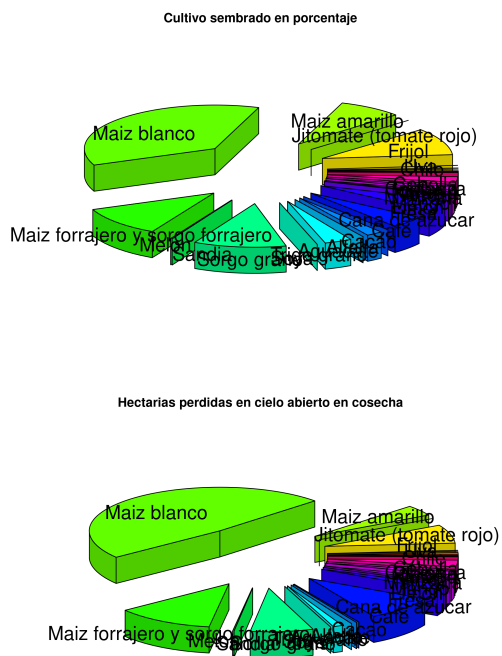


Figure 4: Muestra del porcentaje de hectáreas cosechadas por cultivo y perdida de hectáreas en forma de pastel y barras

References

- [1] INEGI. Agricultura, ganadería y pesca.
- [2] R.org. R.
- [3] Visual Studio. Visual studio code.