**Análisis de los cultivos de coca en Colombia**

Héctor Alonso Bohórquez Suárez

Servicio Nacional de Aprendizaje SENA.

Programación para la Analítica de Datos

**Tabla de Contenidos**

[Capítulo 1 Planteamiento del problema 1](#_Toc99814586)

[Capítulo 2 Pregunta del problema 2](#_Toc99814587)

[Capítulo 3 Tipo de investigación 3](#_Toc99814588)

[Capítulo 4 Objetivos generales 4](#_Toc99814589)

[Capítulo 5 Objetivos específicos 5](#_Toc99814590)

[Capítulo 6 Metodología. 6](#_Toc99814591)

[Capítulo 7 Resultado. 8](#_Toc99814592)

[Fase 1 Obtención de los datos. 8](#_Toc99814593)

[Fase 1.1 Vista previa de los datos. 9](#_Toc99814594)

[Fase 1.2 Ingresar a Jupiter Notebook. 10](#_Toc99814595)

[Fase 1.3 Crear carpeta en Jupiter Notebook. 11](#_Toc99814596)

[Fase 1.4 Crear Notebook en Jupiter Notebook. 12](#_Toc99814597)

[Fase 2 Refinamiento de los datos. 13](#_Toc99814598)

[Fase 2.1 Leer los datos. 14](#_Toc99814599)

[Fase 2.2 Descripción de los datos. 15](#_Toc99814600)

[Fase 2.3 Explorar los datos. 16](#_Toc99814601)

[Fase 2.4 Tipos de datos. 17](#_Toc99814602)

[Fase 2.5 Convertir datos a tipo numérico. 18](#_Toc99814603)

[Fase 2.6 Comprobar datos numéricos. 19](#_Toc99814604)

[Fase 2.7 Verificar datos nulos. 20](#_Toc99814605)

[Fase 2.8 Rellenar datos nulos. 21](#_Toc99814606)

[Fase 2.9 Comprobar datos nulos. 22](#_Toc99814607)

[Fase 2.10 Eliminar columnas. 23](#_Toc99814608)

[Fase 2.11 Agrupar filas. 24](#_Toc99814609)

[Fase 3 Análisis de los datos. 25](#_Toc99814610)

[Fase 3.1 Cantidad de hectáreas por año. 26](#_Toc99814611)

[Fase 3.2 Selección de columnas. 27](#_Toc99814612)

[Fase 3.3 Renombrar filas. 28](#_Toc99814613)

[Fase 3.4 Promedio de los datos. 29](#_Toc99814614)

[Fase 3.5 Valores máximos. 30](#_Toc99814615)

[Fase 3.6 Valores mínimos. 31](#_Toc99814616)

[Fase 3.7 Suma de los datos. 32](#_Toc99814617)

[Fase 4 Visualización de los datos. 33](#_Toc99814618)

[Fase 4.1 Cultivos de coca por departamento. 34](#_Toc99814619)

[Fase 4.2 Datos estadísticos por departamento. 35](#_Toc99814620)

[Fase 4.3 Cultivos de coca por regiones. 36](#_Toc99814621)

[Fase 4.4 Datos estadísticos por regiones. 37](#_Toc99814622)

[Fase 4.5 Cultivos de coca del año 2020. 38](#_Toc99814623)

[Fase 4.6 Cultivos de coca por años. 39](#_Toc99814624)

[Capítulo 8 Conclusión. 40](#_Toc99814625)

[Bibliografía 41](#_Toc99814626)

**Tabla de Ilustraciones**

[Ilustración 1 8](#_Toc99814627)

[Ilustración 2 9](#_Toc99814628)

[Ilustración 3 10](#_Toc99814629)

[Ilustración 4 11](#_Toc99814630)

[Ilustración 5 12](#_Toc99814631)

[Ilustración 6 13](#_Toc99814632)

[Ilustración 7 14](#_Toc99814633)

[Ilustración 8 15](#_Toc99814634)

[Ilustración 9 16](#_Toc99814635)

[Ilustración 10 17](#_Toc99814636)

[Ilustración 11 18](#_Toc99814637)

[Ilustración 12 19](#_Toc99814638)

[Ilustración 13 20](#_Toc99814639)

[Ilustración 14 21](#_Toc99814640)

[Ilustración 15 22](#_Toc99814641)

[Ilustración 16 23](#_Toc99814642)

[Ilustración 17 24](#_Toc99814643)

[Ilustración 18 25](#_Toc99814644)

[Ilustración 19 26](#_Toc99814645)

[Ilustración 20 27](#_Toc99814646)

[Ilustración 21 28](#_Toc99814647)

[Ilustración 22 29](#_Toc99814648)

[Ilustración 23 30](#_Toc99814649)

[Ilustración 24 31](#_Toc99814650)

[Ilustración 25 32](#_Toc99814651)

[Ilustración 26 34](#_Toc99814652)

[Ilustración 27 35](#_Toc99814653)

[Ilustración 28 36](#_Toc99814654)

[Ilustración 29 37](#_Toc99814655)

[Ilustración 30 38](#_Toc99814656)

[Ilustración 31 39](#_Toc99814657)

# Capítulo 1 Planteamiento del problema

La problemática de la siembra de los cultivos de coca en Colombia se ha convertido en la principal fuente de financiamiento de los grupos delincuenciales al margen de la ley los cuales secuestran, matan, reclutan niños, desplazan muchas familias cada año, ejecutan acciones terroristas en contra de la población civil, perjudican la vida de muchas personas debido a las graves consecuencias del consumo de las drogas y en el proceso de elaboración de la cocaína contaminan los ríos y el medio ambiente.

# Capítulo 2 Pregunta del problema

¿Como ha evolucionado la cantidad de hectáreas de los cultivos de coca en Colombia desde el año 2015 hasta el 2020?

# Capítulo 3 Tipo de investigación

# 

La investigación de los datos es de tipo cuantitativa ya que tenemos que encontrar valores numéricos y estadísticos de la cantidad de hectáreas de los cultivos de coca que se puedan contar, verificar y que nos den información acerca de las cantidades e información que pueda ser medida y escrita con números que nos permitan realizar operaciones matemáticas, estadísticas y de visualización con los datos.

“Método de investigación descriptiva: Se realiza previo al análisis de datos con el objetivo de encontrar factores y características importantes del fenómeno de estudio.

La investigación cuantitativa proporciona una descripción y exploración de fenómenos en situaciones y características de la vida real. Se describe la correlación de individuos, situaciones o grupos particulares.

La investigación cuantitativa consiste en recolectar y analizar datos numéricos. Este método es ideal para identificar tendencias y promedios, realizar predicciones, comprobar relaciones y obtener resultados generales de poblaciones grandes”, Según un nuevo estudio... (Ortega, 2022b).

# Capítulo 4 Objetivos generales

* Analizar la cantidad de hectáreas por departamentos de los cultivos de coca desde el año 2015 hasta el 2020.

# Capítulo 5 Objetivos específicos

* Extraer los datos que contengan la información de la cantidad de hectáreas de los cultivos de coca de la página web: (Detección de Cultivos de Coca (hectáreas) | Datos Abiertos Colombia, 2021).
* Refinar los datos de la cantidad de cultivos de coca y prepararlos para poder analizarlos utilizando técnicas de refinación y limpieza de datos.
* Analizar los datos por medio de la estadística descriptiva para resumir, y presentar los datos que describen el comportamiento de los cultivos de coca.
* Visualizar por medio de gráficos el comportamiento de los cultivos de coca por departamentos y regiones.

# Capítulo 6 Metodología.

La metodología que se va a implementar para analizar el comportamiento de la cantidad de hectáreas de los cultivos de coca en Colombia es la siguiente:

**Fase1 Obtención de datos**: en este proceso para obtener los datos debemos buscar una fuente confiable que contenga la información que queremos obtener e importar las librerías que nos permitan extraer, reunir y procesar los datos en un formato por ejemplo ya sea tipo Json, Excel, csv. La extracción de la información se hará de la página web: (Detección *de Cultivos de Coca (hectáreas) | Datos Abiertos Colombia*, 2021).

**Fase 2 Refinamiento de datos**: en este proceso debemos explorar los datos para verificar los tipos de datos utilizando técnicas de refinación y limpieza de datos empleadas por José de Jesús Fernández García en el video: (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

A continuación, se detalla los pasos a tener presente en esta fase del proyecto:

* + Verificar los tipos de datos si estos son de tipo Object se procede a convertir a datos numéricos para poder realizar operaciones con ellos.
  + Luego verificar la cantidad de datos nulos que tengan las columnas y los reemplazamos por el promedio de los datos de la columna o por ceros según sea el caso.
  + A continuación, limpiar los datos eliminando filas o columnas que no nos aportan información valiosa.

**Fase 3 Análisis de los datos**: en este proceso debemos analizar qué tipo de relación hay entre los datos utilizando la librería Pandas (*Pandas Documentation — Pandas 1.4.1 Documentation*, 2022) Esta librería Pandas es un paquete de Python que proporciona estructuras de datos similares a los dataframes de R. Pandas depende de Numpy, la librería que añade un potente tipo matricial a Python.

También utilizaremos la biblioteca NumPy que es un proyecto de código abierto que tiene como objetivo habilitar la computación numérica con Python.  (*NumPy*, 2022).

* Debemos analizar la cantidad de hectáreas de coca por departamentos.
* Debemos analizar la cantidad de hectáreas de coca por regiones.
* Analizar en qué años ha habido un aumento o disminución considerable de los cultivos de coca.

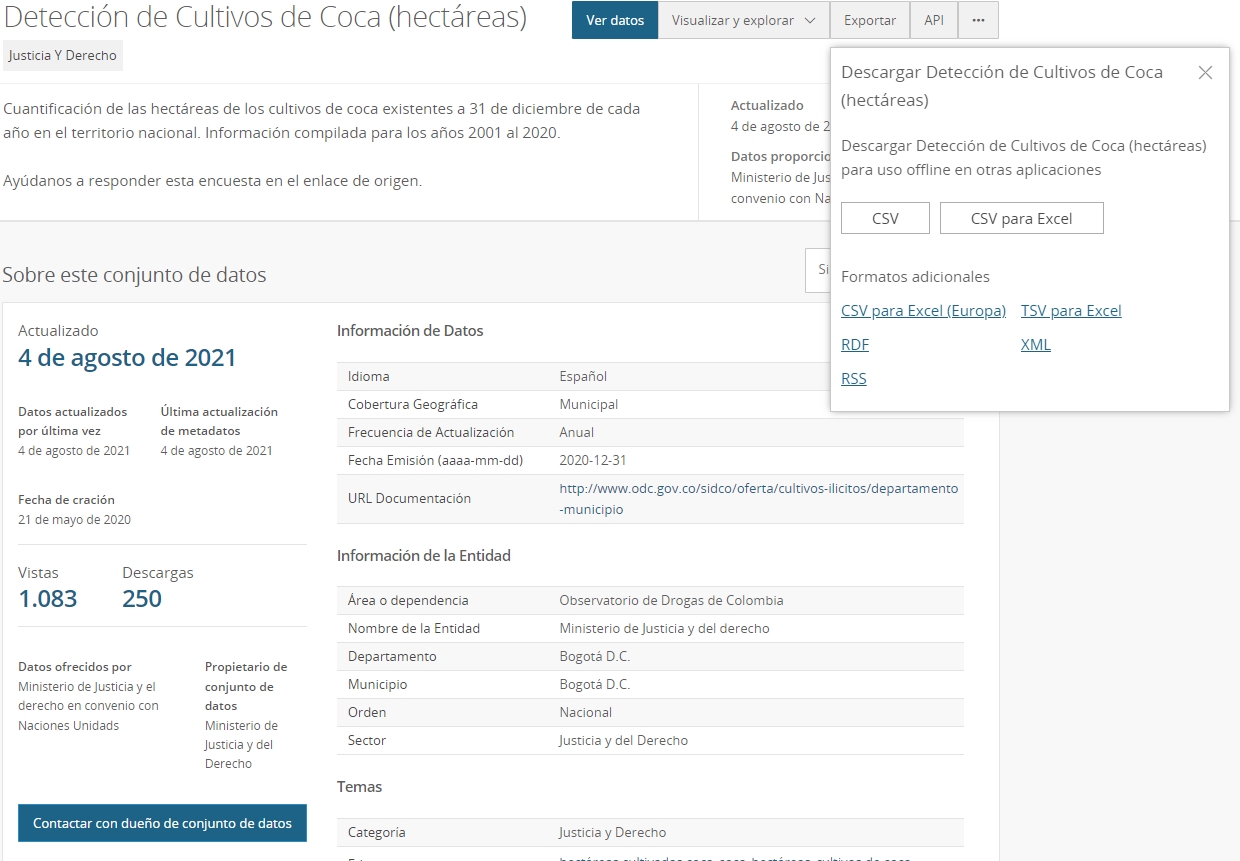
**Fase 4 Visualización de los datos**: en este proceso debemos visualizar los datos atreves de imágenes o gráficos que nos permitan tener una visión clara, dinámica, visible y sencilla de los datos utilizando las técnicas de visualización de datos empleadas por  
Frank Andrade (*Gráficos Interactivos en Python con Pandas y Plotly | Curso de Visualización de Datos en Python*, 2021).

# Capítulo 7 Resultado.

## Fase 1 Obtención de los datos.

Ilustración 1

Página web: Cultivos de coca.



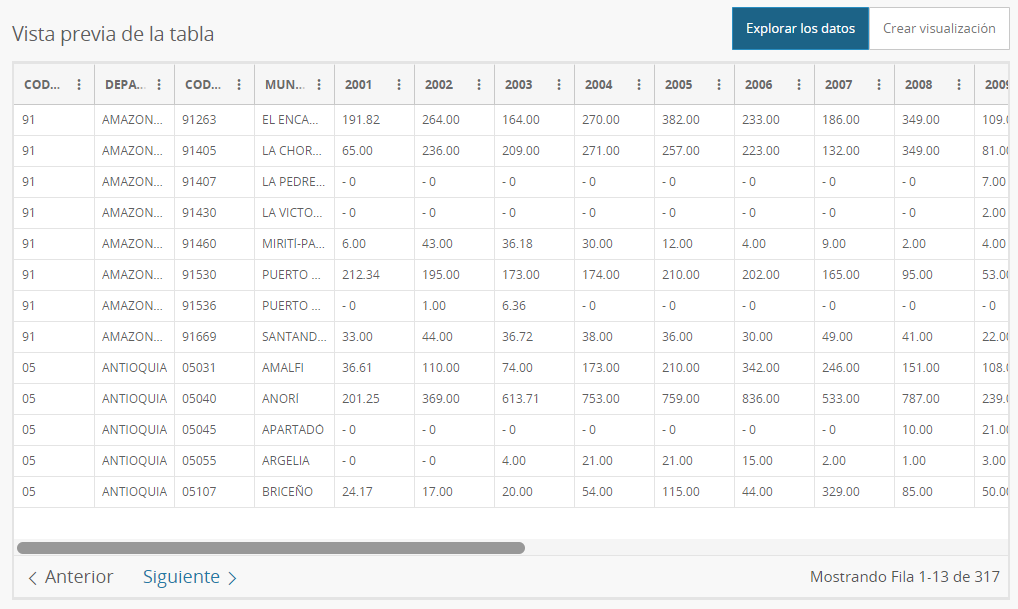
Fuente: Elaboración Propia.

En esta fase se obtuvieron los datos de la página web: (Detección *de Cultivos de Coca (hectáreas) | Datos Abiertos Colombia*, 2021). Los datos fueron descargados en formato CSV.

## Fase 1.1 Vista previa de los datos.

Ilustración 2

Vista previa de la tabla: Cultivos de coca.



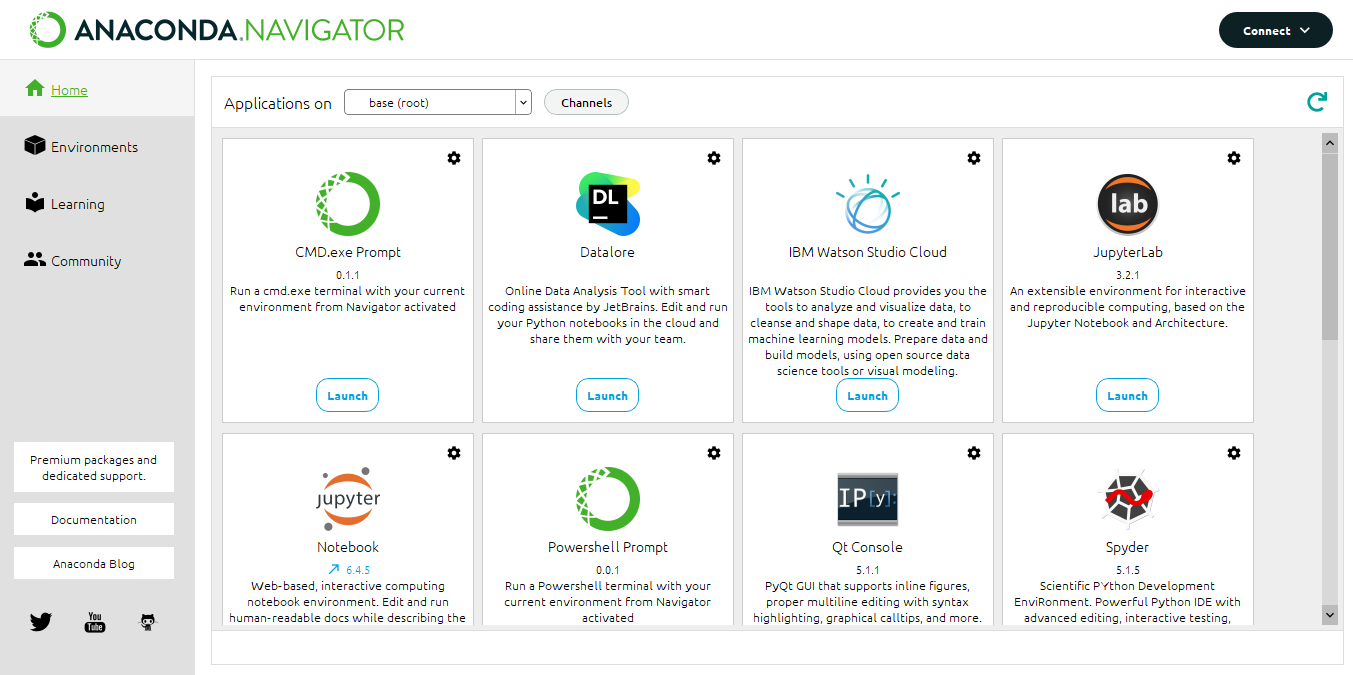
Fuente: Elaboración Propia.

En esta fase, nos desplazamos hacia abajo y podemos observar una vista previa de la tabla que contiene la información de las hectáreas de coca, los departamentos y los municipios.

## Fase 1.2 Ingresar a Jupiter Notebook.

Ilustración 3

Ingresar a Jupiter Notebook desde Anaconda Navigator.



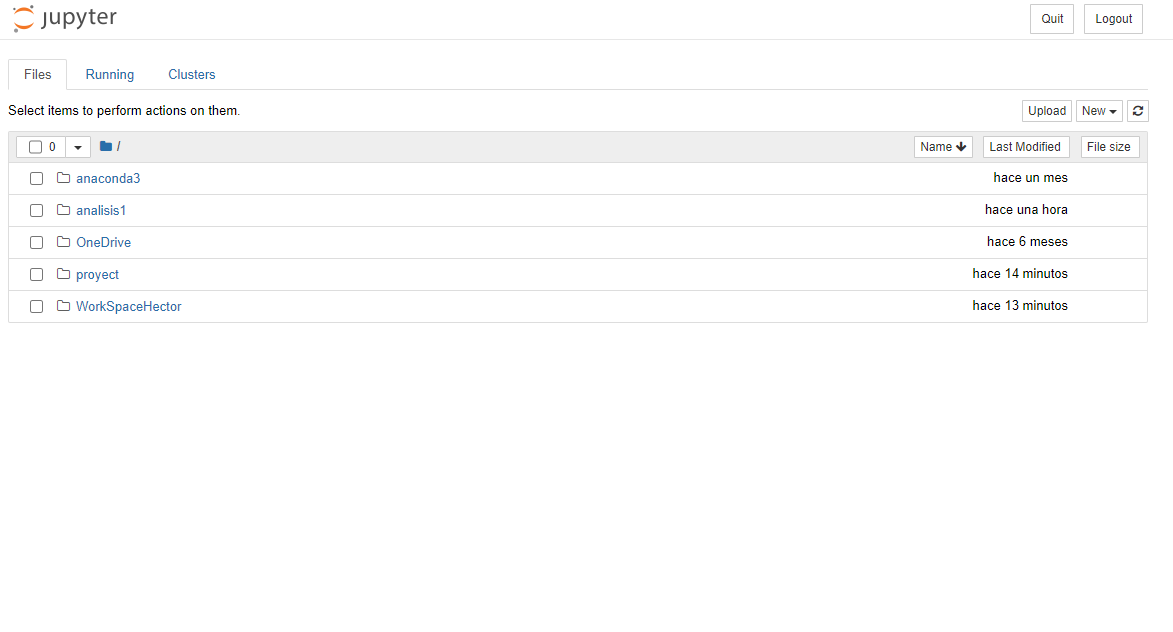
Fuente: Elaboración Propia.

En esta fase ingresamos a nuestro entorno de trabajo Jupiter Notebook que es una herramienta de trabajo para el análisis de datos de Anaconda Navigator (*Anaconda | Individual Edition*, 2022).

## Fase 1.3 Crear carpeta en Jupiter Notebook.

Ilustración 4

Crear carpeta en Jupiter Notebook



Fuente: Elaboración Propia.

En esta fase cuando entremos a Jupyter Notebook creamos una carpeta de trabajo dando click en el botón New y le ponemos un nombre, en mi caso la llamare WorkSpaceHector (*Project Jupyter*, 2022).

## Fase 1.4 Crear Notebook en Jupiter Notebook.

Ilustración 5

Crear notebook.



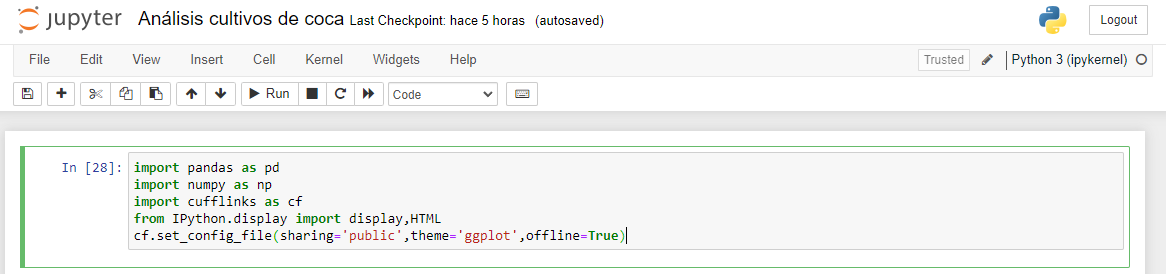
Fuente: Elaboración Propia.

En esta fase cuando estemos dentro de nuestra carpeta de trabajo cargamos nuestros datos previamente descargado y renombrado cultivosDeCoca.csv y creamos un notebook que en mi caso llamare Análisis cultivos de coca (*Project Jupyter*, 2022).

## Fase 2 Refinamiento de los datos.

Ilustración 6

Importar librerías.



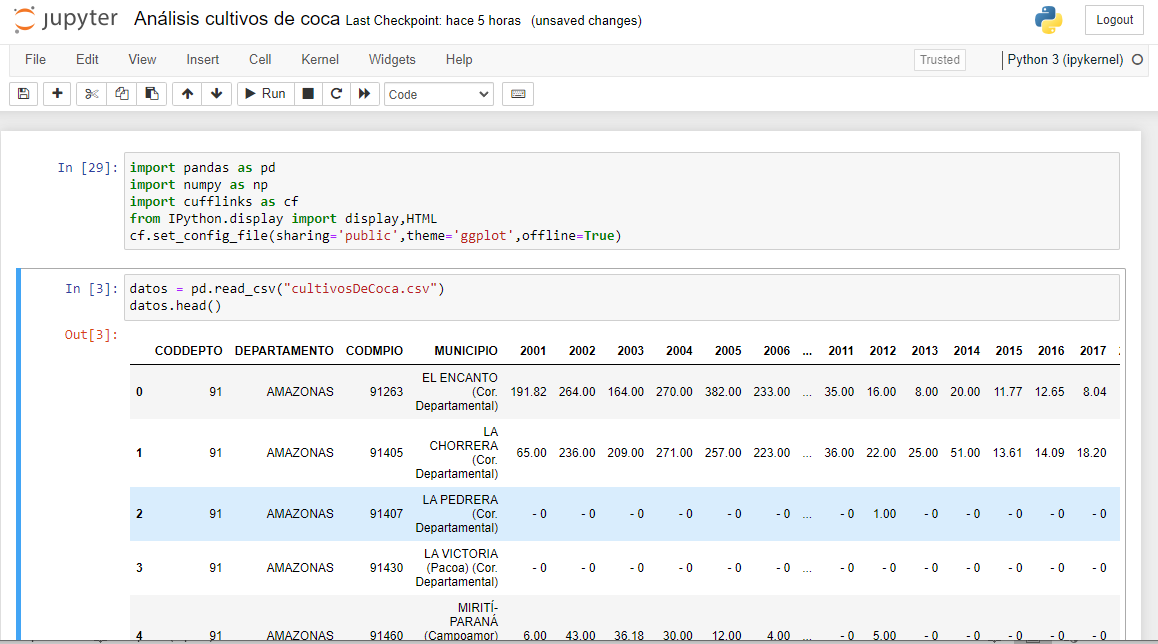
Fuente: Elaboración Propia.

En esta fase cuando estemos en nuestro notebook empezamos importando las librerías. (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

## Fase 2.1 Leer los datos.

Ilustración 7

Leer los datos.



Fuente: Elaboración propia.

En esta fase leemos nuestros datos almacenándolos en una variable, en mi caso la llamare datos. (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

df.read\_csv: lee un archivo de valores separados por comas (csv) en DataFrame.

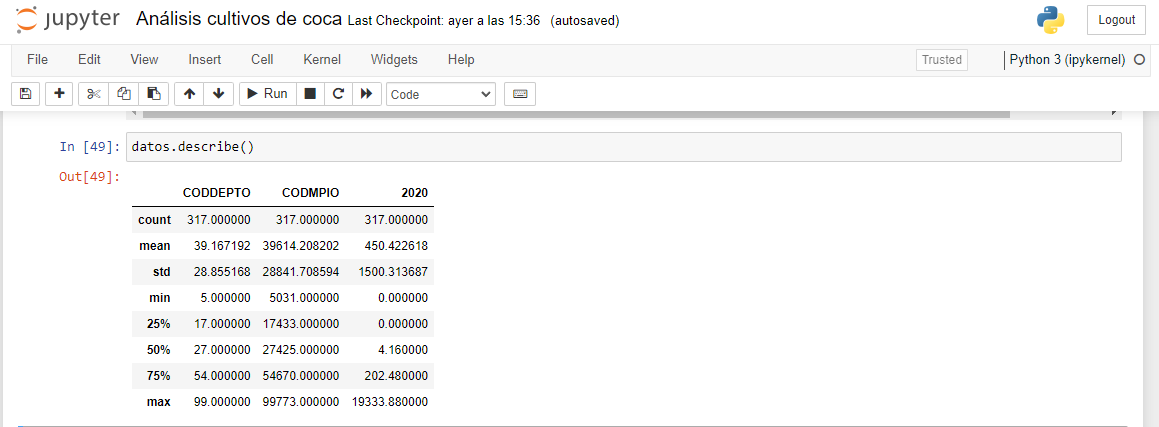
df.head(n) : Devuelve las n primeras filas del DataFrame df.

df.tail(n) : Devuelve las n últimas filas del DataFrame df. (Alberca, 2021).

## Fase 2.2 Descripción de los datos.

Ilustración 8

Descripción de los datos.



Fuente: Elaboración Propia.

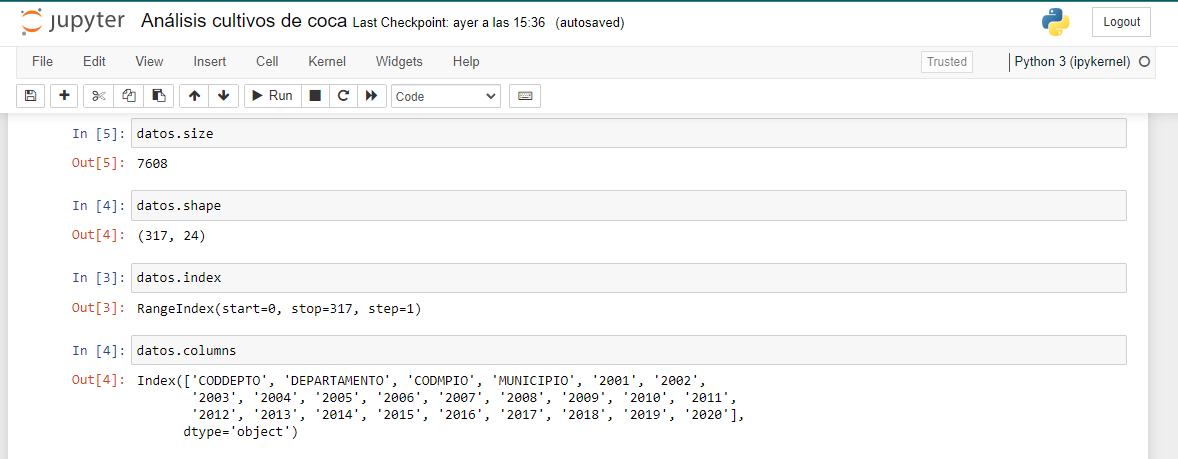
En esta fase observamos el promedio de los datos que tenemos pandas.DataFrame.describe: Genera estadísticas descriptivas.

Las estadísticas descriptivas incluyen aquellas que resumen la tendencia central, la dispersión y la forma de la distribución de un conjunto de datos, excluyendo NaNlos valores. (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

## Fase 2.3 Explorar los datos.

Ilustración 9

Explorar los datos.



Fuente: Elaboración Propia.

En esta fase exploramos nuestro dataframe para conocer los datos (*Preprocesamiento y*

*análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

df.size : Devuelve el número de elementos del DataFrame.

df.shape : Devuelve una tupla con el número de filas y columnas del DataFrame df.

df.index : Devuelve una lista con los nombres de las filas del DataFrame df.

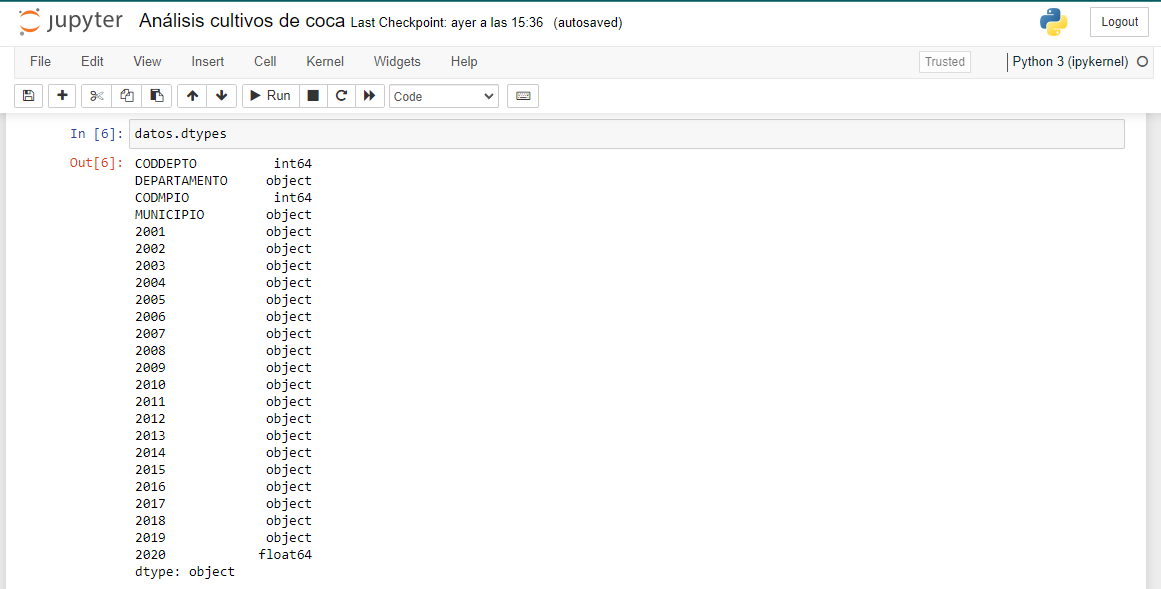
df.columns : Devuelve una lista con los nombres de las columnas del DataFrame df

(Alberca, 2021).

## Fase 2.4 Tipos de datos.

Ilustración 10

Tipo de datos.



Fuente: Elaboración Propia.

En esta fase observamos que tipo de datos tenemos (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

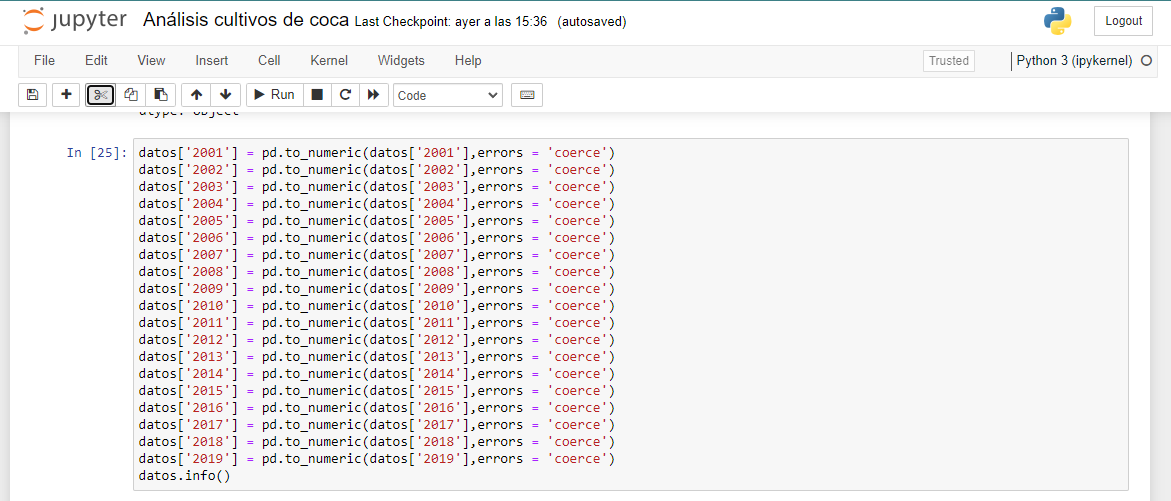
df.dtypes : Devuelve una serie con los tipos de datos de las columnas del DataFrame df.

(Alberca, 2021).

## Fase 2.5 Convertir datos a tipo numérico.

Ilustración 11

Convertir datos a numérico.



Fuente: Elaboración Propia.

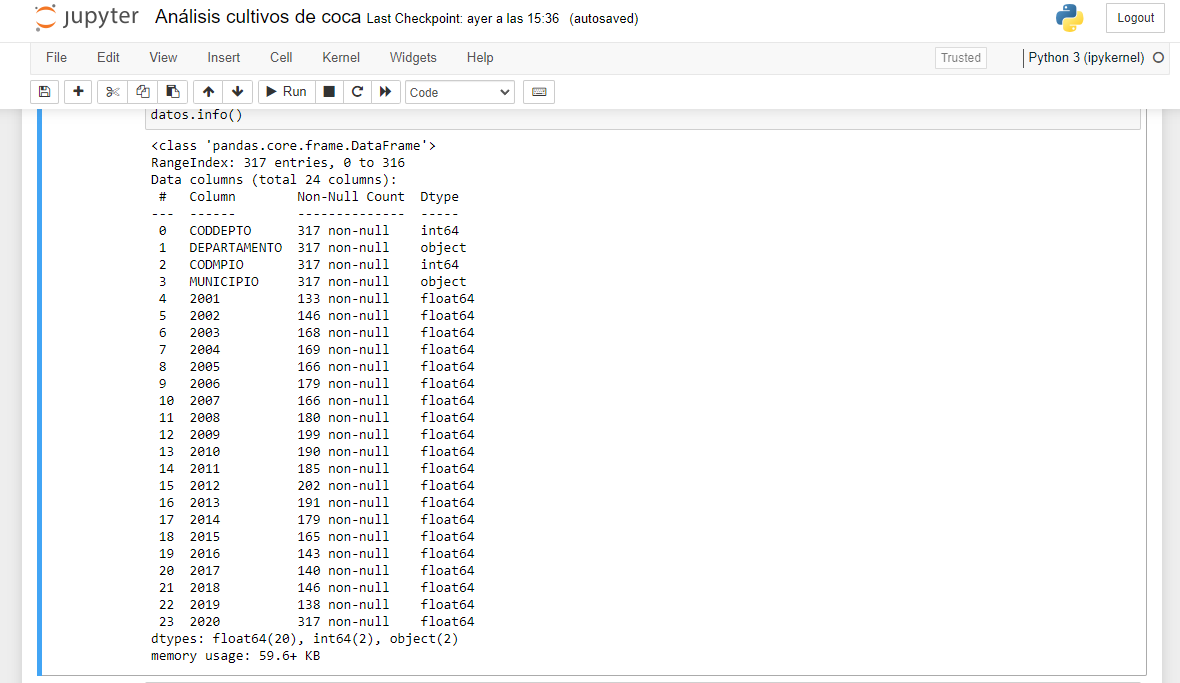
En esta fase observamos que la mayoría de los datos que tenemos son de tipo Object los cuales tenemos que cambiarlos a tipo numérico para poder trabajar con ellos (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

pandas to\_numeric: Convierte el argumento a un tipo numérico (Alberca, 2021).

## Fase 2.6 Comprobar datos numéricos.

Ilustración 12

Verificar datos.



Fuente: Elaboración Propia.

En esta fase verificamos nuestro cambio de datos:

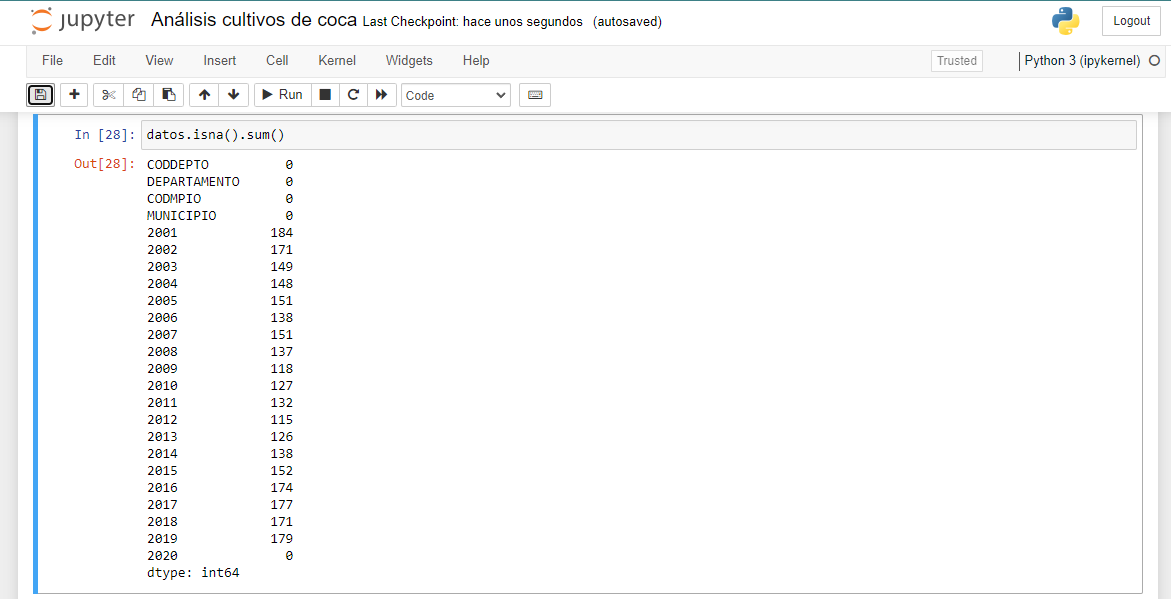
df.info() : Devuelve información (número de filas, número de columnas, índices, tipo de

las columnas y memoria usado) sobre el DataFrame df (Alberca, 2021).

## Fase 2.7 Verificar datos nulos.

Ilustración 13

Verificación de datos nulos.



Fuente: Elaboración Propia.

En esta fase verificamos la cantidad de datos nulos que tenemos (*Preprocesamiento y*

*Análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

df.isna().sum(): Detectar valores faltantes.

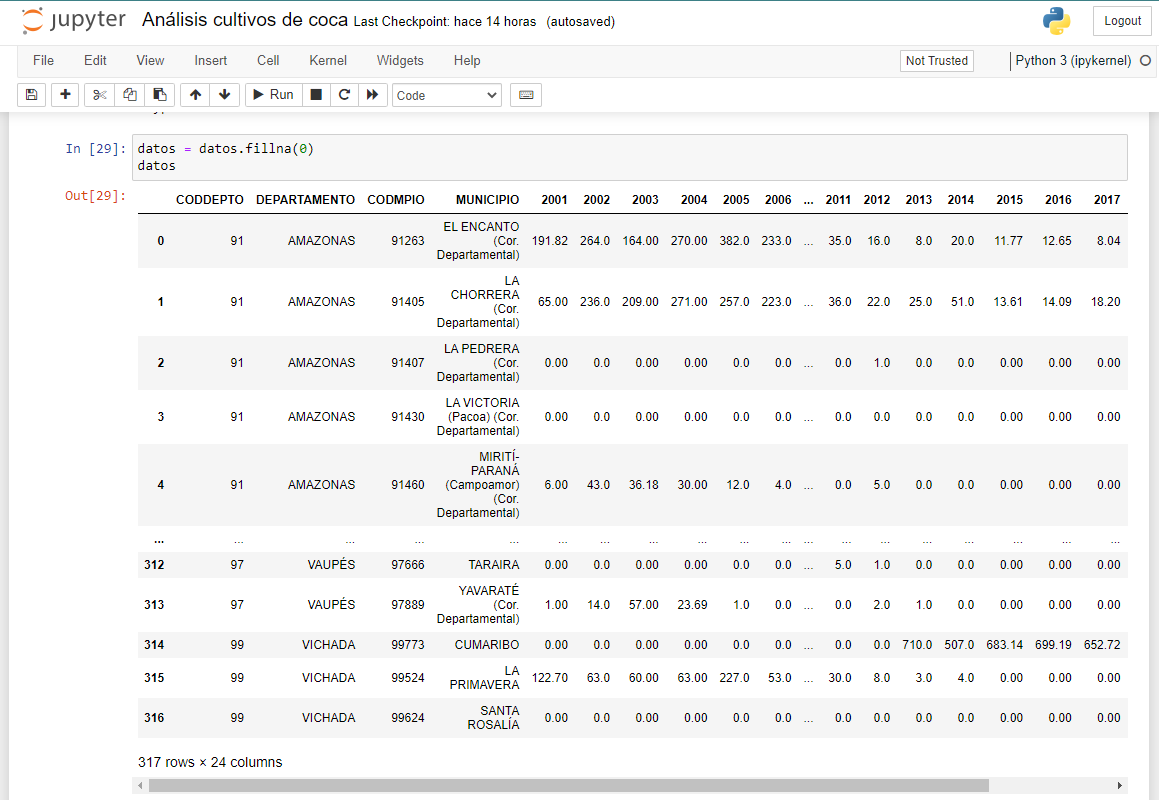
Devuelve un objeto booleano del mismo tamaño que indica si los valores son NA

(Alberca, 2021).

## Fase 2.8 Rellenar datos nulos.

Ilustración 14

Rellenar los valores nulos.



Fuente: Elaboración Propia.

En esta fase procedemos a reemplazar los valores nulos por 0 (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

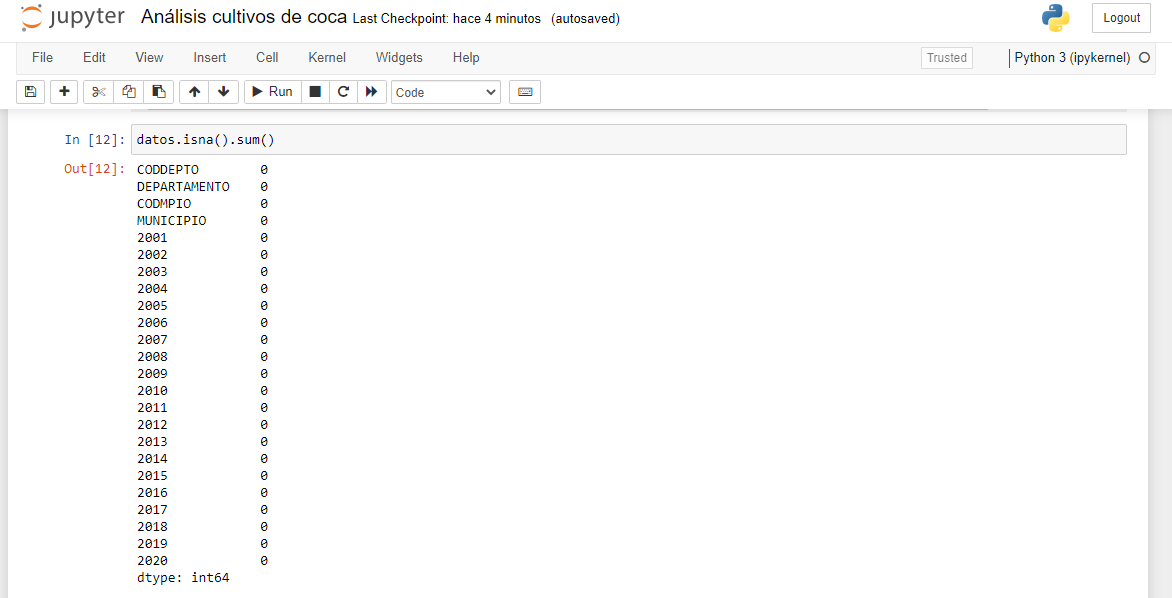
df.fillna(): Rellena los valores NA/NaN utilizando el método especificado (Alberca,

2021).

## Fase 2.9 Comprobar datos nulos.

Ilustración 15

Comprobar valores nulos.



Fuente: Elaboración Propia.

En esta fase comprobamos que ya no tengamos datos nulos (*Preprocesamiento y análisis*

*dedatos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

df.isna().sum(): Detectar valores faltantes.

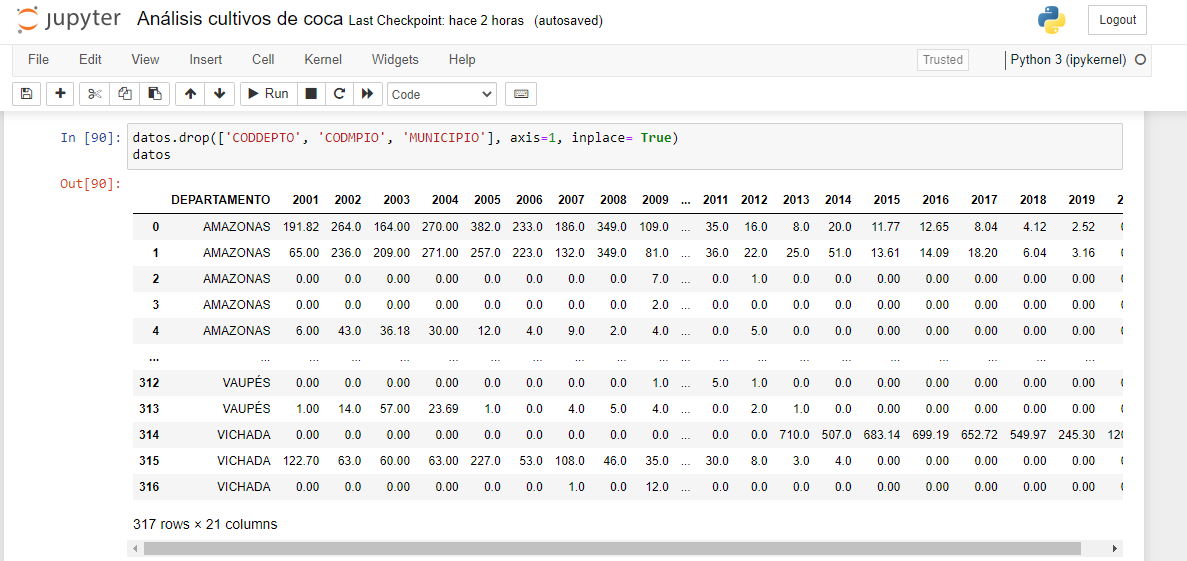
Devuelve un objeto booleano del mismo tamaño que indica si los valores son NA

(Alberca, 2021).

## Fase 2.10 Eliminar columnas.

Ilustración 16

Eliminar columnas.



Fuente: Elaboración Propia.

En esta fase eliminamos las filas o columnas que no nos aportan información valiosa

(*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

df.drop: Suelta etiquetas específicas de filas o columnas.

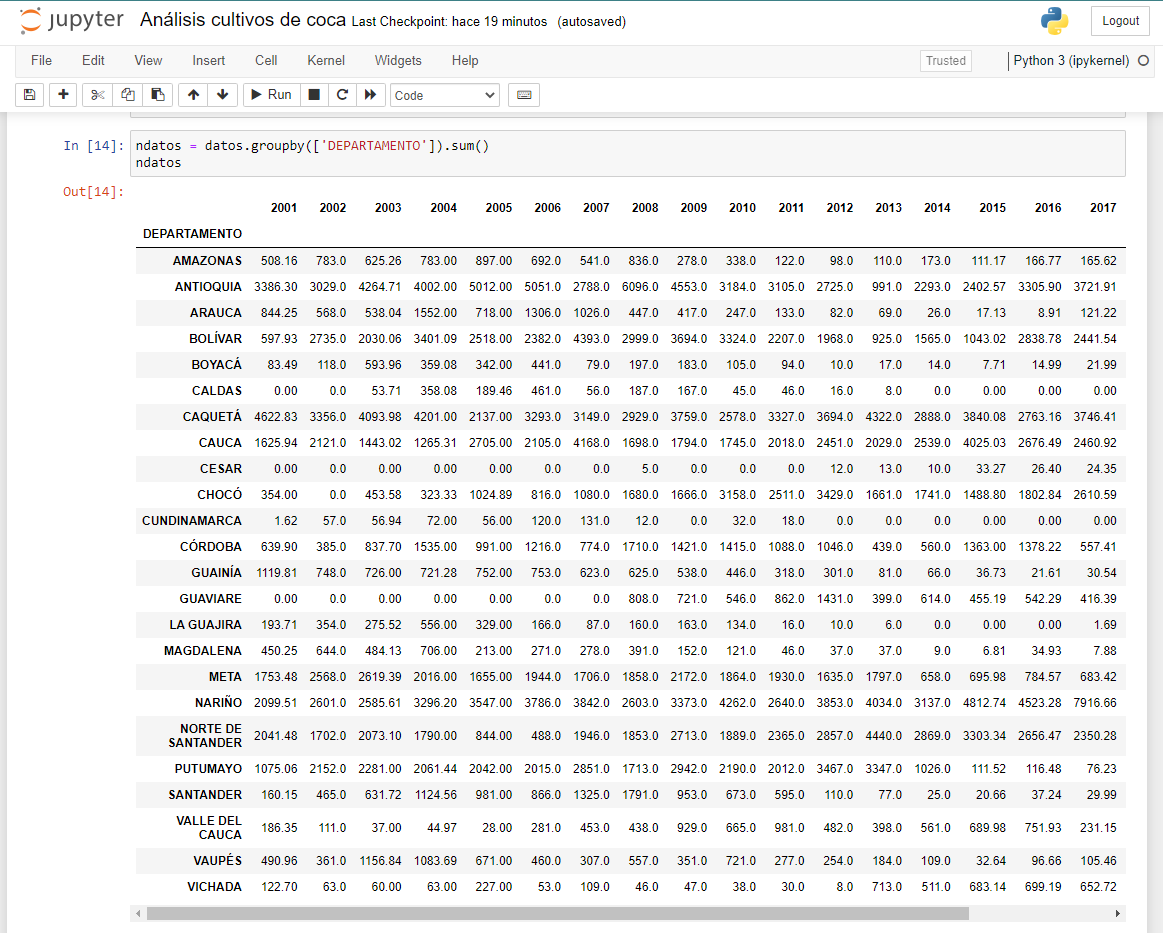
Elimine filas o columnas especificando nombres de etiquetas y ejes correspondientes, o especificando directamente nombres de índices o columnas. Cuando se utiliza un índice múltiple, las etiquetas en diferentes niveles se pueden eliminar especificando el nivel

(Alberca, 2021).

## Fase 2.11 Agrupar filas.

Ilustración 17

Agrupar filas.



Fuente: Elaboración Propia.

En esta fase agrupamos las filas o columnas que deseamos.

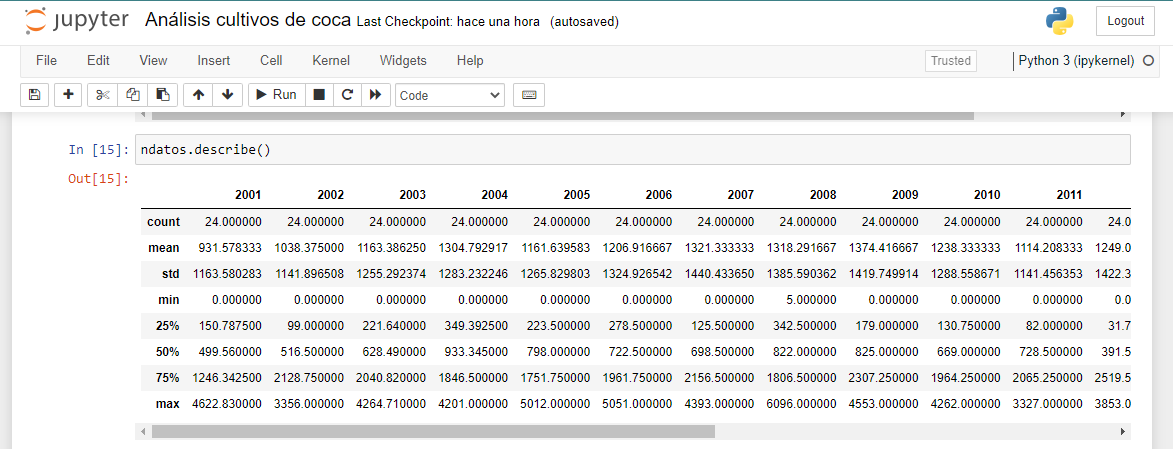
pandas.DataFrame.groupby: agrupa DataFrame usando un mapeador o por una serie de

columnas (Alberca, 2021).

## Fase 3 Análisis de los datos.

Ilustración 18

Descripción de los datos.



Fuente: Elaboración Propia.

En esta fase usamos el método describe() que nos genera una serie de datos estadísticos descriptivos y podemos observar que ya lo hace con las columnas que eran del tipo object (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

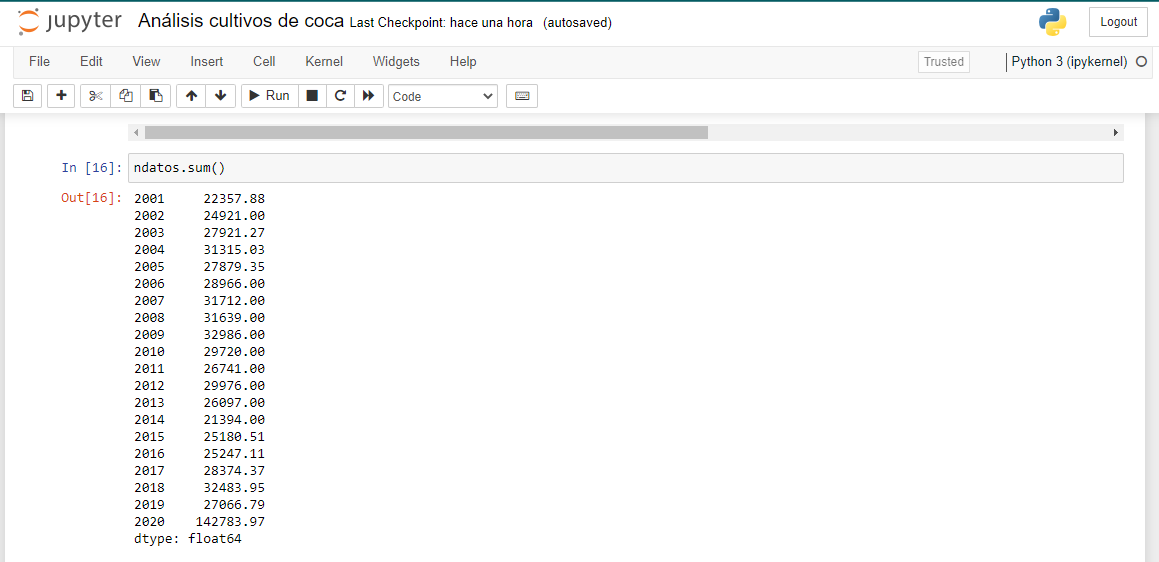
pandas.DataFrame.describe: Genera estadísticas descriptivas.

Las estadísticas descriptivas incluyen aquellas que resumen la tendencia central, la dispersión y la forma de la distribución de un conjunto de datos, excluyendo NaNlos valores. (*Pandas.DataFrame.Describe — Pandas 1.4.1 Documentation*, 2022).

## Fase 3.1 Cantidad de hectáreas por año.

Ilustración 19

Suma de hectáreas.



Fuente: Elaboración Propia.

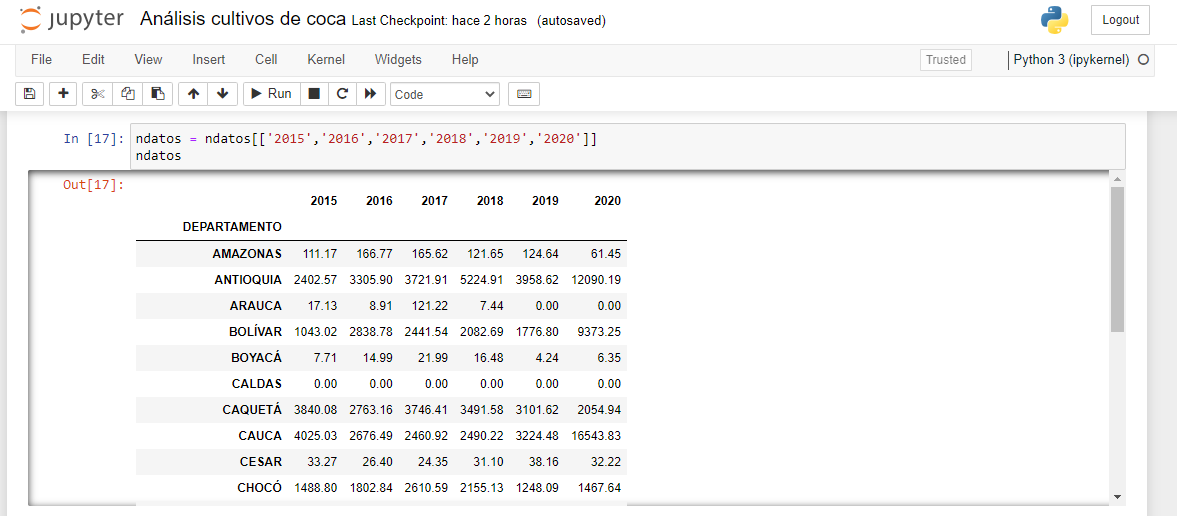
En esta afse sumamos la cantidad de hectáreas que hay en cada año de nuestras columnas (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

df.sum(): devuelve la suma de los valores sobre el eje solicitado. (*Pandas.DataFrame.Sum — Pandas 1.4.1 Documentation*, 2022).

## Fase 3.2 Selección de columnas.

Ilustración 20

Selección de columnas.



Fuente: Elaboración Propia.

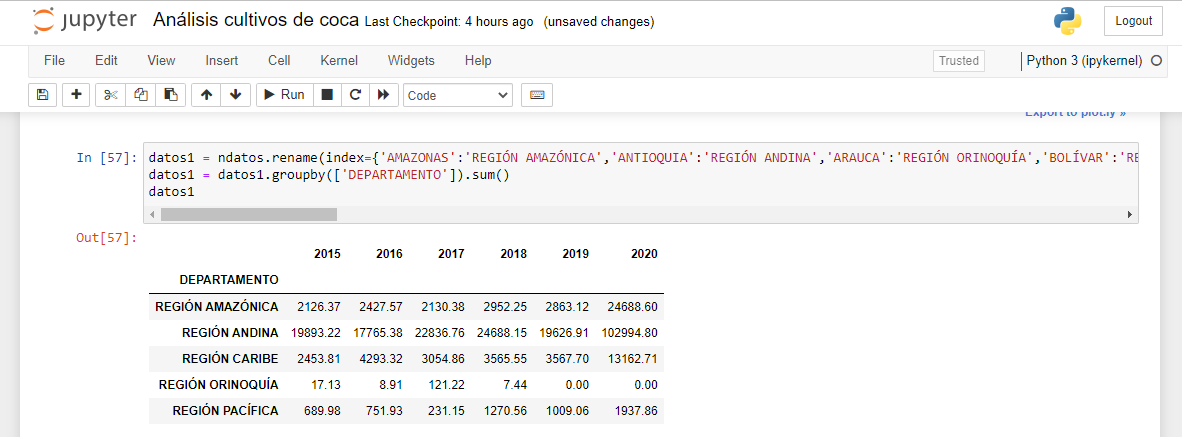
En esta fase seleccionamos las columnas que queremos analizar (*Preprocesamiento y*

*análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

## Fase 3.3 Renombrar filas.

Ilustración 21

Renombrar departamentos.



Fuente: Elaboración Propia.

En esta fase renombramos los departamentos por el nombre de cada región y luego los agrupamos con el método rename (*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*, 2020).

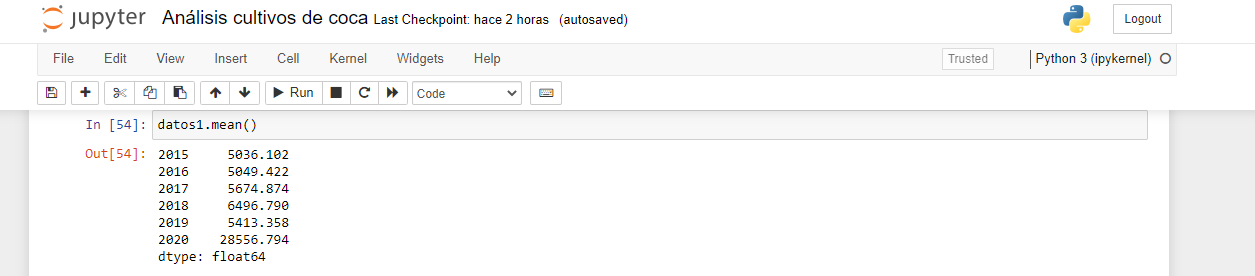
Alterar las etiquetas de los ejes.

Los valores de función/dictado deben ser únicos (1 a 1). Las etiquetas que no estén contenidas en un dictado/serie se dejarán como están. Las etiquetas adicionales enumeradas no arrojan un error. (*Pandas Documentation — Pandas 1.4.1 Documentation*, 2022).

## Fase 3.4 Promedio de los datos.

Ilustración 22

Promedio de los datos:



Fuente: Elaboración Propia.

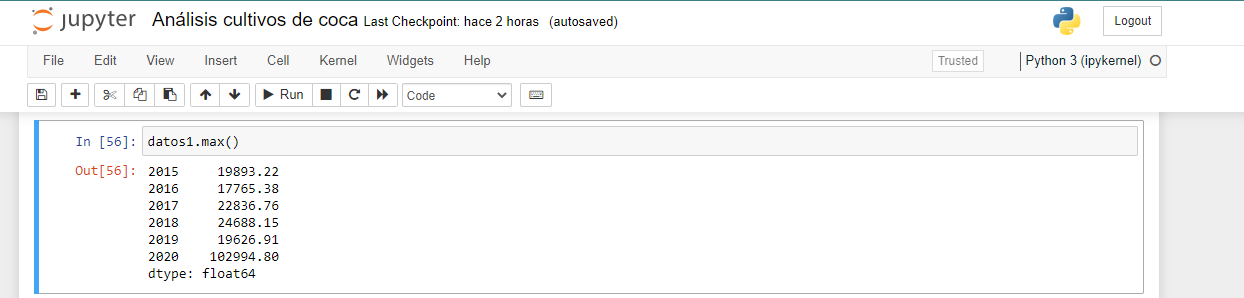
En esta fase obtenemos el promedio de la cantidad de cultivos de coca desde el año 2015 hasta el año 2020.

pandas.DataFrame.Mean: Devuelve la media de los valores sobre el eje solicitado. (*Pandas.DataFrame.Mean — Pandas 1.4.1 Documentation*, 2022).

## Fase 3.5 Valores máximos.

Ilustración 23

Valores máximos.



Fuente: Elaboración Propia.

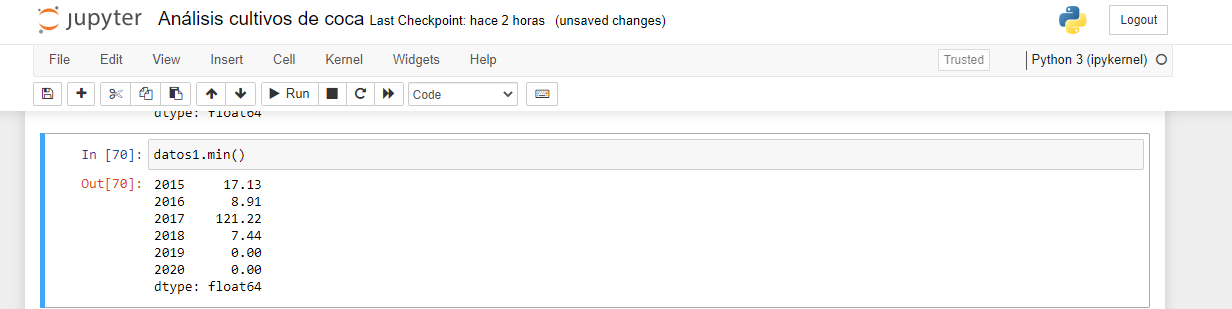
En esta fase obtenemos los valores máximos de la cantidad de cultivos de coca desde el año 2015 hasta el año 2020.

pandas.DataFrame.Max: Devuelve el valor máximo de los valores sobre el eje solicitado. (*Pandas.DataFrame.Max — Pandas 1.4.1 Documentation*, 2022).

## Fase 3.6 Valores mínimos.

Ilustración 24

Valores mínimos.



Fuente: Elaboración Propia.

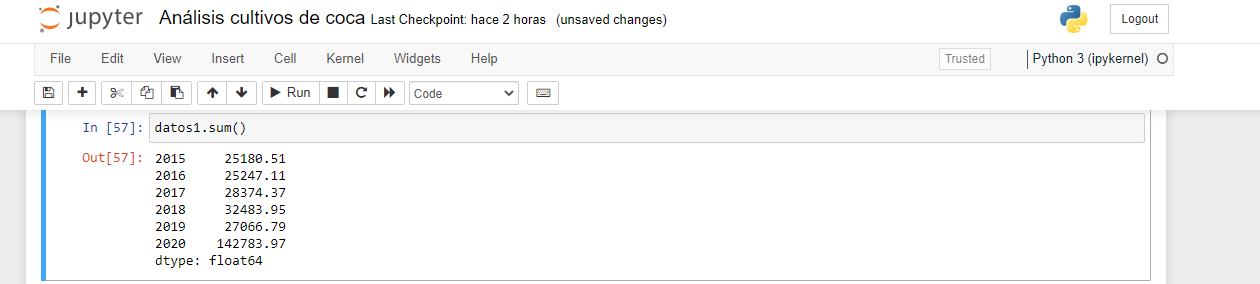
En esta fase obtenemos los valores mínimos de la cantidad de cultivos de coca desde el año 2015 hasta el año 2020.

pandas.DataFrame.Min: Devuelve el valor máximo de los valores sobre el eje solicitado. (*Pandas.DataFrame.Min — Pandas 1.4.1 Documentation*, 2022).

## Fase 3.7 Suma de los datos.

Ilustración 25

Suma de los datos.

****

Fuente: Elaboración Propia.

En esta fase obtenemos la suma de la cantidad de cultivos de coca desde el año 2015 hasta el año 2020.

pandas.DataFrame.Sum: Devuelve la suma de los valores sobre el eje solicitado. (*Pandas.DataFrame.Sum — Pandas 1.4.1 Documentation*, 2022).

## Fase 4 Visualización de los datos.

En esta fase procedemos a visualizar nuestros datos.

iplot: es trama interactiva. Plotly toma el código de Python y crea hermosas tramas de JavaScript. Le permiten tener mucho control sobre cómo se ven estos gráficos y le permiten hacer zoom, mostrar información al pasar el mouse y alternar los datos para verlos en el gráfico.

kind='line': Esta función es útil para trazar líneas utilizando los valores de DataFrame como coordenadas.

xTitle: Añade un título con el contenido de la cadena título al eje x.

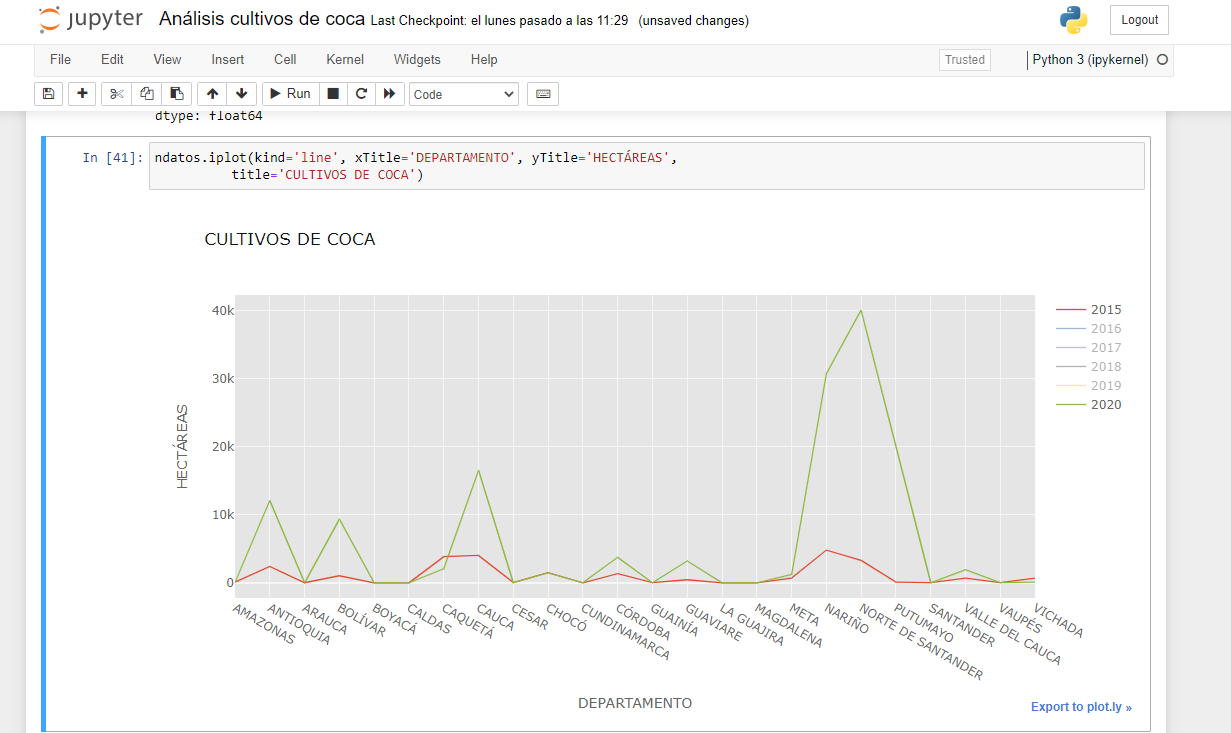
yTitle: Añade un título con el contenido de la cadena título al eje y.

title: Añade un título a la gráfica. (*Gráficos Interactivos en Python con Pandas y Plotly | Curso de Visualización de Datos en Python*, 2021b).

## Fase 4.1 Cultivos de coca por departamento.

Ilustración 26

Cultivos de coca por departamento.



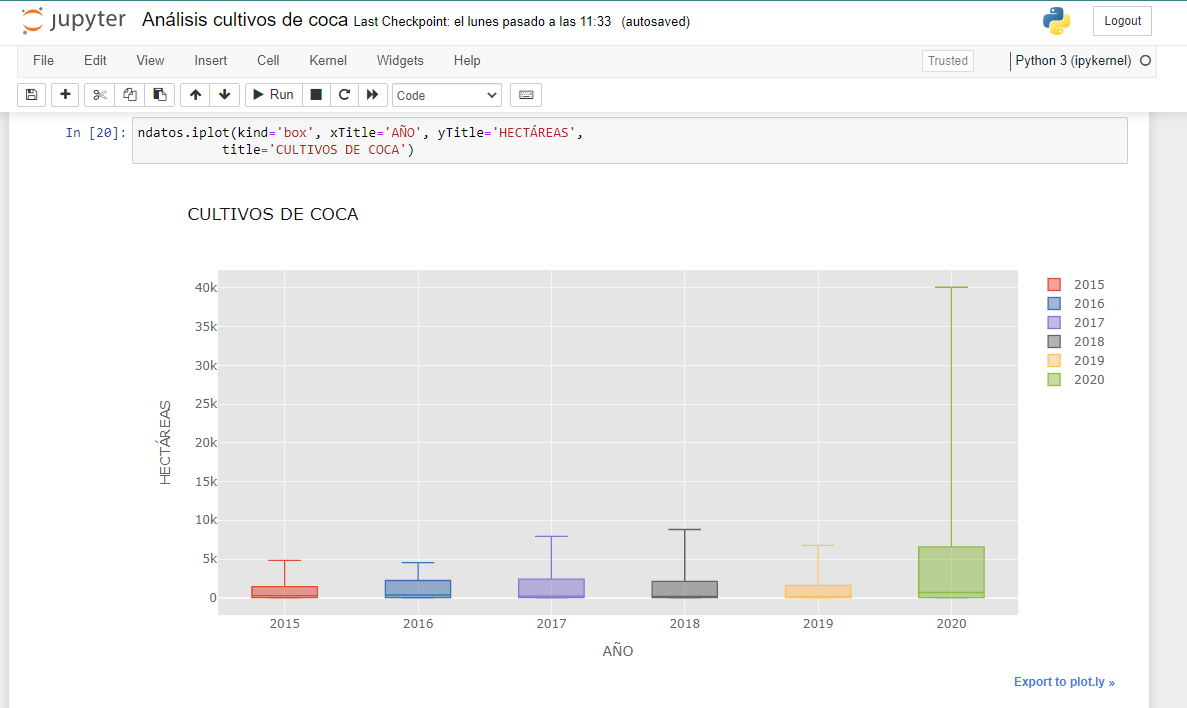
Fuente: Elaboración Propia.

Esta grafica muestra la cantidad de hectáreas de los cultivos de coca por departamento desde el año 2015 hasta el año 2020.

## Fase 4.2 Datos estadísticos por departamento.

Ilustración 27

Datos estadísticos por departamento.



Fuente: Elaboración Propia.

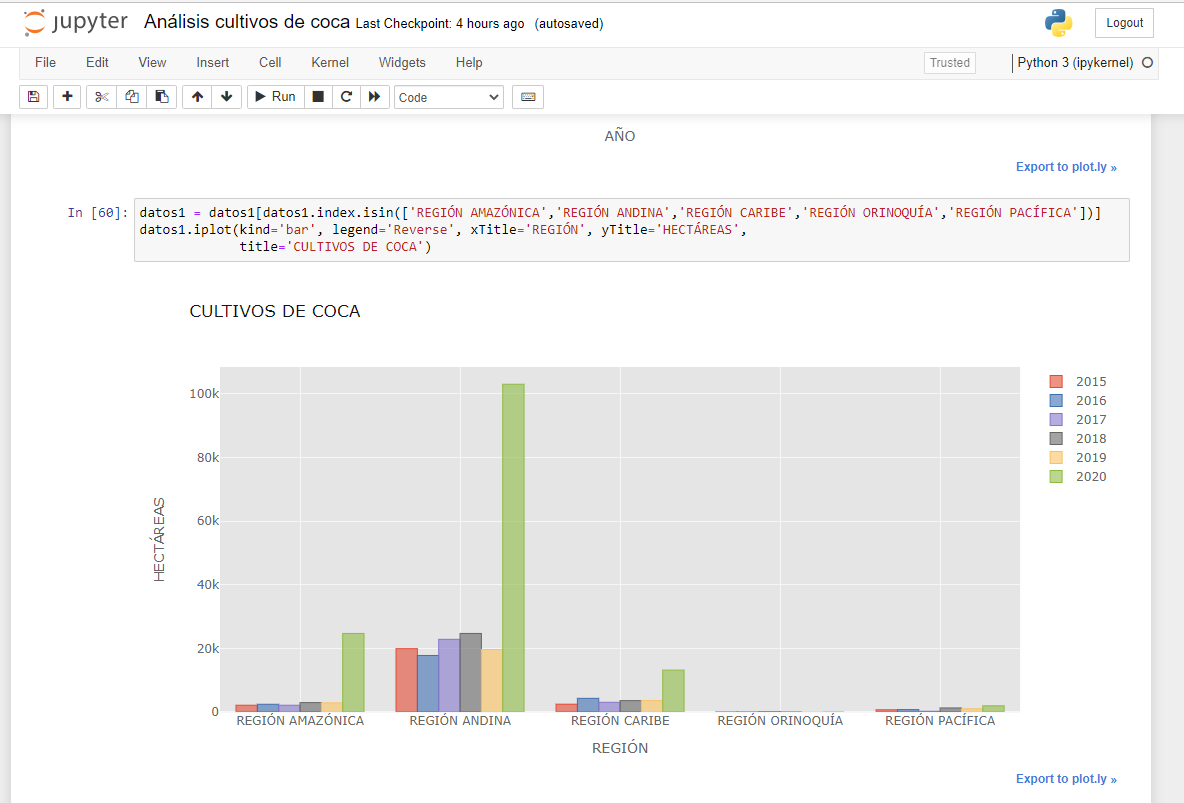
Esta grafica nos muestra los datos estadísticos de la cantidad de hectáreas de coca por departamentos desde el año 2015 hasta el año 2020.

Un diagrama de caja es una representación estadística de la distribución de una variable a través de sus cuartiles. Los extremos de la caja representan los cuartiles inferior y superior, mientras que la mediana (segundo cuartil) está marcada por una línea dentro de la caja. (*Gráficos Interactivos en Python con Pandas y Plotly | Curso de Visualización de Datos en Python*, 2021b).

## Fase 4.3 Cultivos de coca por regiones.

Ilustración 28

Cultivos de coca por regiones.



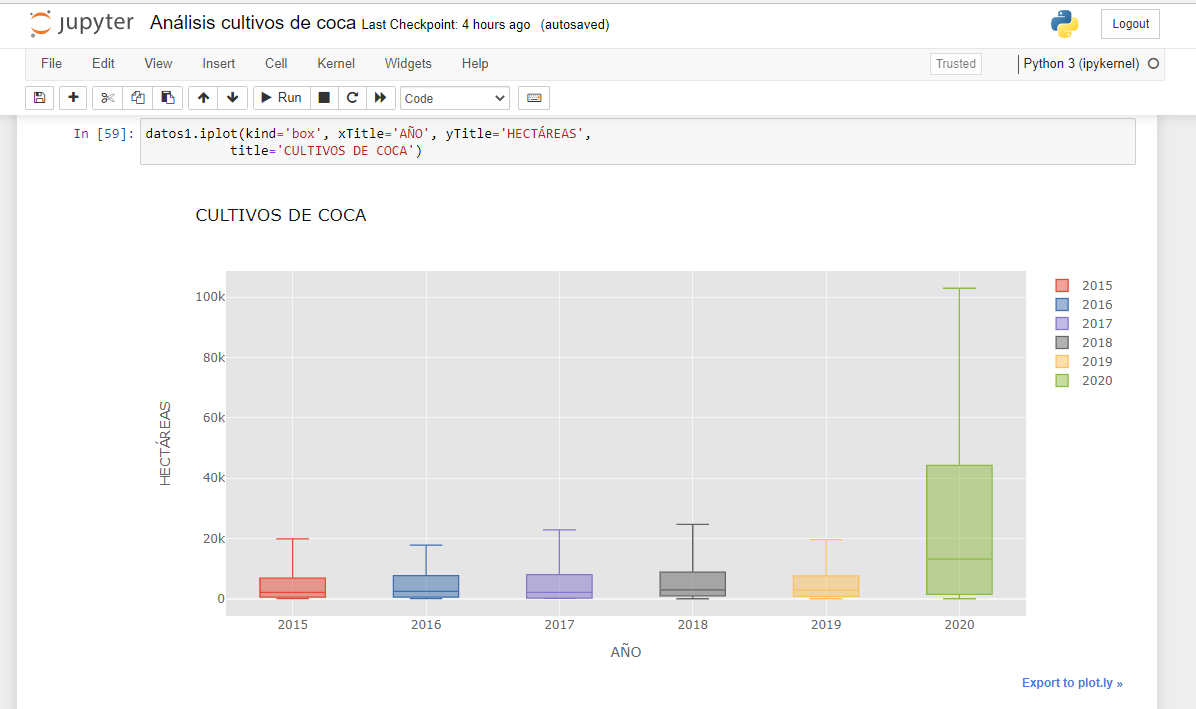
Fuente: Elaboración Propia.

En esta grafica podemos ver la cantidad de hectáreas de coca por regiones desde el año 2015 hasta el año 2020. (*Gráficos Interactivos en Python con Pandas y Plotly | Curso de Visualización de Datos en Python*, 2021b).

## Fase 4.4 Datos estadísticos por regiones.

Ilustración 29

Datos estadísticos por regiones.



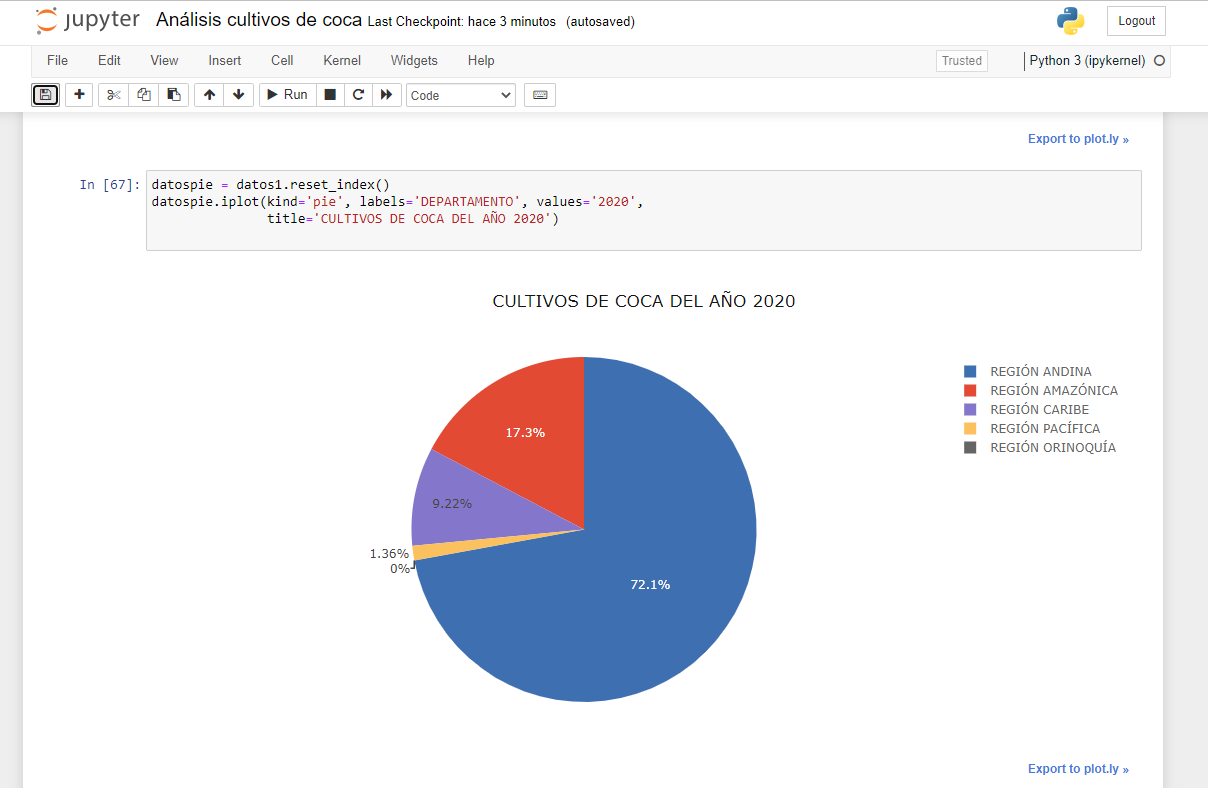
Fuente: Elaboración Propia.

Esta grafica nos muestra los datos estadísticos de la cantidad de hectáreas de coca por regiones desde el año 2015 hasta el año 2020. (*Gráficos Interactivos en Python con Pandas y Plotly | Curso de Visualización de Datos en Python*, 2021b).

## Fase 4.5 Cultivos de coca del año 2020.

Ilustración 30

Cultivos de coca del año 2020.



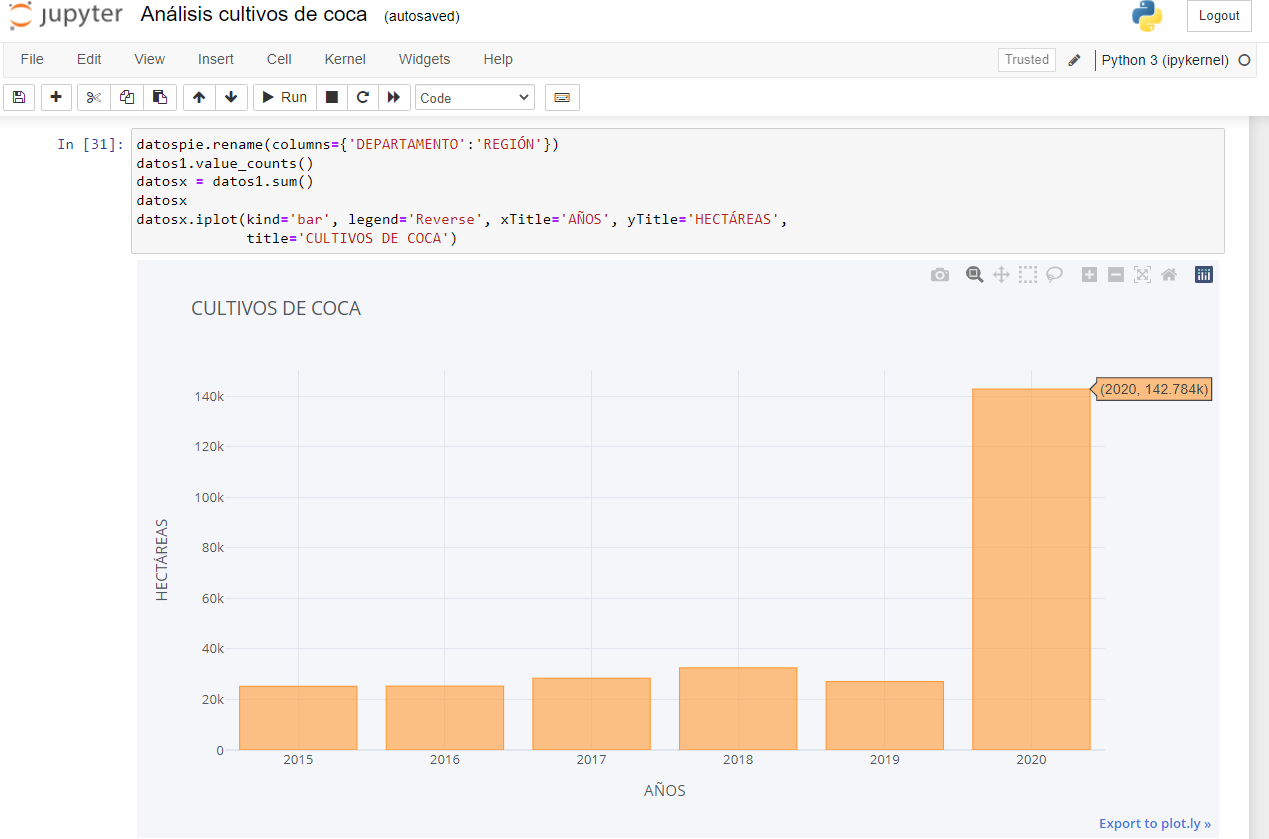
Fuente: Elaboración propia

En esta grafica podemos observar la cantidad de hectáreas de coca por regiones del año 2020. (*Gráficos Interactivos en Python con Pandas y Plotly | Curso de Visualización de Datos en Python*, 2021b).

## Fase 4.6 Cultivos de coca por años.

Ilustración 31

Cultivos de coca por años.



Fuente: Elaboración propia

En esta grafica podemos observar la cantidad de hectáreas de coca por años. (*Gráficos Interactivos en Python con Pandas y Plotly | Curso de Visualización de Datos en Python*, 2021b).

# Capítulo 8 Conclusión.

Como conclusión del proyecto podemos decir que el proyecto fue viable porque logramos analizar la evolución de la cantidad de los cultivos de coca en Colombia desde el año 2015 hasta el año 2020 por departamentos y por regiones.

Con base a este análisis podemos saber que la cantidad de cultivos de coca en Colombia han aumentado y en el año 2020 ha habido un aumento considerable en esta, siendo la región andina la que más hectáreas de coca presento con 102.994 hectárea.

# Bibliografía

Ortega, C. (2022b, febrero 25). *Investigación cuantitativa. Qué es y cómo realizarla*. QuestionPro. https://www.questionpro.com/blog/es/que-es-la-investigacion-cuantitativa/

*Detección de Cultivos de Coca (hectáreas) | Datos Abiertos Colombia*. (2021, 4 agosto). Página Web. https://www.datos.gov.co/Justicia-y-Derecho/Detecci-n-de-Cultivos-de-Coca-hect-reas-/acs4-3wgp

*Anaconda | Individual Edition*. (2022, 7 marzo). Anaconda. Recuperado 7 de marzo de 2022, de https://www.anaconda.com/products/individual

*Project Jupyter*. (2022, 7 marzo). Home. Recuperado 7 de marzo de 2022, de https://jupyter.org/

*Preprocesamiento y análisis de datos en python con Jupyter Notebook utilizando el dataset titanic*. (2020, 27 septiembre). [Vídeo]. YouTube. https://www.youtube.com/watch?v=fVoVLzqQgtE&list=WL&index=2&t=1526s

*pandas documentation — pandas 1.4.1 documentation*. (2022, 25 febrero). Página Web. Recuperado 25 de febrero de 2022, de https://pandas.pydata.org/docs/

*NumPy*. (2022, 25 febrero). Página Wed. Recuperado 25 de febrero de 2022, de https://numpy.org/

Alberca, A. S. (2021, 14 mayo). *La librería Pandas*. Aprende con Alf. Recuperado 7 de marzo de 2022, de https://aprendeconalf.es/docencia/python/manual/pandas/

*Gráficos Interactivos en Python con Pandas y Plotly | Curso de Visualización de Datos en Python*. (2021, 22 septiembre). YouTube. Recuperado 25 de febrero de 2022, de https://www.youtube.com/watch?v=Zt3tzGjLmPA

*plotly*. (2022, 9 febrero). PyPI. Recuperado 7 de marzo de 2022, de https://pypi.org/project/plotly/

*iplot*. (2022, 7 marzo). Página web. Recuperado 7 de marzo de 2022, de https://stackoverflow.com/questions/49880314/what-is-difference-between-plot-and-iplot-in-pandas