

# Softwares para la visualización estadística

Camila Acosta Ramirez

2020-12-22



# Contents

<b>Portada</b>	<b>5</b>
<b>1 Introducción</b>	<b>7</b>
<b>2 Tableau</b>	<b>9</b>
2.1 Generalidades . . . . .	9
2.2 Instalación de Tableau Desktop Public . . . . .	20
2.3 Forma de navegación . . . . .	21
2.4 Flujo de trabajo . . . . .	25
<b>3 Power BI</b>	<b>125</b>
3.1 Generalidades . . . . .	125
3.2 Instalación de Power BI Desktop . . . . .	130
3.3 Forma de navegación . . . . .	130
3.4 Flujo de trabajo . . . . .	132
<b>4 Flourish</b>	<b>147</b>
4.1 Example one . . . . .	147
4.2 Example two . . . . .	147
<b>5 Conclusiones</b>	<b>149</b>



# **Portada**

Espacio para la portada ....



# Chapter 1

## Introducción

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 1. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter 2.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

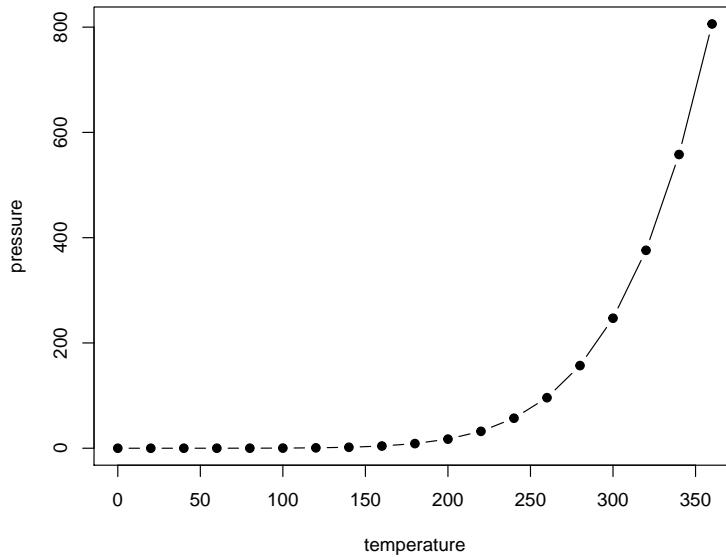


Figure 1.1: Here is a nice figure!

Table 1.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 1.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 1.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the `bookdown` package (Xie, 2020) in this sample book, which was built on top of R Markdown and `knitr` (Xie, 2015).

# Chapter 2

## Tableau

### 2.1 Generalidades

#### 2.1.1 ¿Qué es Tableau?

Tableau es una plataforma de análisis visual que transforma la forma en que usamos los datos para resolver problemas, lo que permite a las personas y organizaciones aprovechar al máximo sus datos. Esta plataforma hace que sea más fácil para las personas explorar y administrar datos, y más rápido para descubrir y compartir información que puede cambiar las empresas y el mundo, todo lo creado por esta plataforma está impulsado por ayudar a las personas a ver y comprender los datos, porque sus productos están diseñados para poner al usuario en primer lugar, ya sea un analista, un científico de datos, un estudiante, un profesor, un ejecutivo o un usuario empresarial. Desde la conexión hasta la colaboración, Tableau es la plataforma de análisis de un extremo a otro más potente, seguro y flexible.

Tableau se fundó en 2003 como resultado de un proyecto de informática en Stanford que tenía como objetivo mejorar el flujo de análisis y hacer que los datos fueran más accesibles para las personas a través de la visualización. Los cofundadores Chris Stolte, Pat Hanrahan y Christian Chabot desarrollaron y patentaron la tecnología fundamental de Tableau, VizQL, que expresa visualmente los datos al traducir las acciones de arrastrar y soltar en consultas de datos a través de una interfaz intuitiva. Desde su fundación han invertido continuamente en investigación y desarrollo, creando así soluciones para ayudar a cualquier persona que trabaje con datos a obtener respuestas más rápido y descubrir información no anticipada, este desarrollo e investigación incluye hacer que el aprendizaje automático, las estadísticas, el lenguaje natural y la preparación inteligente de datos sean más útiles para aumentar la creatividad humana en el análisis.

### 2.1.2 Principales ventajas de Tableau

- Puedes ver y entender tus datos

Es la misión de la compañía, “ayudar a las personas a ver y comprender sus datos”. Con Tableau está cambiando la forma en la que las personas resuelven sus preguntas, analizando sus datos de forma rápida, sencilla y visual. Tableau es una herramienta revolucionaria que está permitiendo acceder y analizar los datos -que son el petróleo del SXXI- a todas las personas, democratizando el análisis de datos de forma visual.

- Adaptable a diferentes situaciones y entornos

Existen diversas formas de utilizar Tableau, de forma individual puedes utilizar Tableau Desktop en tu ordenador para diseñar las visualizaciones de datos. Si necesitas un entorno para organizaciones o empresas, Tableau Server ofrece un entorno colaborativo y seguro al que puedes acceder simplemente con un navegador web. También existe Tableau Online que es equivalente, pero toda la plataforma funciona en la nube de Tableau, sin necesidad de tener infraestructura propia.

- Rápido y fácil, es posible utilizar Tableau para:
  - Crear dashboards e informes visuales.
  - Navegar y visualizar datos de múltiples formas.
  - Tener un autoservicio de BI.
  - Realizar algunos análisis estadísticos, ver tendencias y pronósticos.

No es complejo de utilizar ya que Tableau está diseñado para que sea fácil de usar, enfocado al autoservicio y no requiere de usuarios técnicos.

- Es compatible con múltiples fuentes de datos, Tableau soporta diferentes fuentes de datos y puede conectar a más de 40 diferentes fuentes, algunos ejemplos son:
  - Ficheros Excel, CSV, PDF, etc.
  - Bases de datos relacionales como SQL Server, MySQL, etc.
  - Fuentes OLAP: Microsoft Analysis Services, SAP Hana, etc.
  - Fuentes online como Google Analytics, etc.
  - Conectores a servicios web.
- Juega con tus bases de datos:

Además de conectarse a fuentes de datos diferentes, en Tableau puedes conectarte a diferentes vistas de datos a la vez, crear extracciones, hacer transformaciones, unir y dividir datos, combinar diferentes fuentes de datos, crear grupos y conjuntos, etc.

- No necesitas programar:

Todas las funcionalidades de Tableau funcionan con arrastrar y soltar, incluso la creación de cálculos (que tienen su propio asistente de ayuda) puedes hacerla así, simplemente con el ratón. Eso no quita que, si lo deseas, puedas integrar y

utilizar Tableau con herramientas y lenguajes más complejos como Python o R para la analítica de datos.

- Tiene una comunidad enorme:

En Internet es posible encontrar multitud de usuarios que comparten su trabajo y aprender de ellos, ver y compartir visualizaciones en las galerías de Tableau Public, formarte en la plataforma gratuita de Tableau Training, resolver dudas en Tableau Community, conocer a otros usuarios en los Tableau User Group.

### 2.1.3 Principales desventajas de Tableau

- Se requiere preparación de datos inicial.
- Las características pueden parecer demasiado especializadas y restrictivas, ya que Tableau está diseñado para un uso más amplio.
- Aunque es excelente para fines analíticos, no puede reemplazar las aplicaciones de informes financieros.
- Brinda la capacidad de establecer seguridad de “nivel bajo” en el nivel de datos, pero lo implementa de una manera un poco precaria.
- Tableau se ha especializado en el factor de facilidad de uso. Sin embargo, a medida que los usuarios van obteniendo habilidades y experiencia desean hacer más, y Tableau posee una capacidad limitada de ampliación que no siempre les permite ir a donde desean.
- Los usuarios especializados en herramientas de Inteligencia Empresarial suelen considerar mejor contar con una arquitectura abierta.

### 2.1.4 Productos de Tableau

- Tableau Public: es una plataforma gratuita en línea para explorar visualizaciones de datos y compartir con el público general.
  - Las visualizaciones publicadas en Tableau Public están disponibles para consultarlas en línea, es una plataforma de datos públicos.
  - No es posible guardar visualizaciones localmente.
  - Solo es posible la Conexión a archivos CSV, Excel, y archivos de texto.
  - No es posible cargar más de 15 millones de filas.
  - No permite conexión en tiempo real con los datos, solo es posible la conexión por medio de extracción.
  - Recursos de aprendizaje guiados.
  - Desafíos virtuales.
  - Datos de muestra.
- Tableau Desktop: es la versión profesional, de escritorio y paga de Tableau.
  - Las visualizaciones se pueden guardar localmente en nuestro computador.

- Cantidad ilimitada de datos.
- Conexión a todas las fuentes de datos, tanto locales como en la nube.
- Tableau Prep: proporciona una forma visual y directa de combinar, dar forma y limpiar datos, así como automatizar los flujos de preparación de datos, lo que le ayuda a obtener análisis y conocimientos más rápidamente, se compone de dos productos:
  - Tableau Prep Builder para crear sus flujos de datos. Si desea editar un valor, seleccione y edite directamente. Cambie su tipo de unión y vea el resultado de inmediato. Con cada acción, ve instantáneamente cambiar sus datos, incluso en millones de filas de datos. Tableau Prep Builder le brinda la libertad de reordenar los pasos y experimentar sin consecuencias. Utilice funciones inteligentes para solucionar problemas comunes de preparación de datos. Tableau Prep Builder emplea agrupaciones difusas para convertir tareas repetitivas, como agrupar por pronunciación, en operaciones de un solo clic.
  - Tableau Prep Conductor permite publicar y ejecutar flujos fácilmente en su entorno de servidor. Comparta sus fuentes de datos de forma segura con Tableau Server o Tableau Online. Cree un entorno en el que todos los miembros de su organización puedan trabajar con datos preparados y actualizados.
- Tableau Server: servidor que permite colaborar de forma segura y compartir la información a partir de los datos que ya hayamos subido a través de Tableau Desktop.
- Tableau Online: se trata de una versión de Tableau Server alojada en la nube que permite acceder a los datos sin necesidad de hacer instalaciones.
- Tableau Mobile: se trata de una aplicación complementaria gratuita para Tableau Server o Tableau Online que permite un acceso a los datos y la información guardada en nuestra cuenta.

### 2.1.5 Precios de Tableau

- Para individuos existe “Creador de Tableau”, tiene un costo de \$70 USD por usuario mensual. Incluye Tableau Desktop, Tableau Prep Builder y una licencia de Creator en Tableau Server o Tableau Online.
- Para equipos y organizaciones: Se tiene la opción de implementar con Tableau Server o implementar con Tableau Online, ambas implementaciones requieren al menos un usuario Creador de Tableau y proporcionan la opción de elegir entre dos roles de usuario, pero varían en los precios.
  - Implementación con Tableau Server:
    - \* Creador de Tableau, tiene un costo de \$70 USD por usuario mensual. Incluye Tableau Desktop, Tableau Prep Builder y una licencia de creador en Tableau Server.

- \* Explorador de Tableau, permite explorar datos confiables y responder sus propias preguntas más rápido con análisis completos de autoservicio, tiene un costo de \$35 USD por usuario mensual y se requieren mínimo 5 exploradores. Incluye una licencia de Explorer de Tableau Server.
- \* Visor de Tableau, permite ver e interactuar con paneles y visualizaciones en una plataforma segura y fácil de usar, posee un costo de \$12 USD por usuario mensual y se requieren mínimo 100 espectadores. Incluye una licencia de Viewer de Tableau Sever.
- Implementación con Tableau Online:
  - \* Creador de Tableau, permite al usuario descubrir información valiosa con un potente conjunto de productos que respaldan su flujo de trabajo de análisis de un extremo a otro, tiene un costo de \$70 USD por usuario mensual. Incluye Tableau Desktop, Tableau Prep Builder y una licencia de Creator en Tableau Online.
  - \* Explorador de Tableau, permite explorar datos confiables y responder sus propias preguntas más rápido con análisis completos de autoservicio, tiene un costo de \$42 USD por usuario mensual y se requieren mínimo 5 exploradores. Incluye una licencia de Explorer de Tableau Online.
  - \* Visor de Tableau, permite ver e interactuar con paneles y visualizaciones en una plataforma segura y fácil de usar, posee un costo de \$15 USD por usuario mensual y se requieren mínimo 100 espectadores. Incluye una licencia de Viewer de Tableau Online.

Para mayor información acerca de los precios puede visitar [Tableau Pricing](#).

### 2.1.6 Compartir el trabajo realizado en Tableau

Cuando se usa Tableau Desktop, hay varias formas de guardar y compartir el trabajo realizado:

1. Guardar automáticamente un libro de trabajo: Tableau Desktop guarda automáticamente el trabajo realizado cada pocos minutos; por lo que no se perderán horas de trabajo si Tableau Desktop se cierra inesperadamente. Esta función está habilitada de forma predeterminada. Si Tableau falla, se crea automáticamente una versión recuperada del libro de trabajo con una extensión *.twbr* y se guarda en la misma ubicación que el archivo original o en su carpeta Mi repositorio / libros de Tableau . Los libros de trabajo nuevos se guardan con el nombre “Libro1” más un ID numérico. Cuando vuelve a abrir Tableau, un cuadro de diálogo de recuperación muestra una lista de los archivos recuperados que puede seleccionar y abrir para continuar en su flujo.
2. Guardar un libro de trabajo: cuando abre Tableau Desktop, crea automáticamente un nuevo libro de trabajo. Los libros de trabajo contienen el

trabajo que crea y constan de una o más hojas de trabajo. Cada hoja de trabajo contiene una vista particular de sus datos. Para guardar un libro de trabajo de Tableau:

- Seleccione Archivo > Guardar.
  - Especifique el nombre del archivo del libro de trabajo en el cuadro de diálogo Guardar como. De forma predeterminada, Tableau guarda el archivo con la extensión .twb. De forma predeterminada, Tableau guarda su libro de trabajo en la carpeta Libros de trabajo de su repositorio Mi Tableau. Puede encontrar este repositorio en su carpeta Documentos. Sin embargo, puede guardar los libros de trabajo de Tableau en cualquier directorio que elija.
  - Para guardar una copia de un libro de trabajo que tiene abierto:
    - Seleccione Archivo > Guardar como y guarde el archivo con un nombre nuevo.
3. Guardar un libro de trabajo empaquetado: estos libros de trabajo contienen el libro de trabajo junto con una copia de cualquier fuente de datos de archivo local e imágenes de fondo. El libro de trabajo ya no está vinculado a las imágenes y las fuentes de datos originales. Estos libros de trabajo se guardan con una extensión de archivo .twbx. Otros usuarios pueden abrir el libro de trabajo empaquetado con Tableau Desktop o Tableau Reader y no necesitan acceder a las fuentes de datos que incluye el libro de trabajo.
4. Guardar un marcador: puede guardar una sola hoja de trabajo como marcador de Tableau. Cuando guarda el marcador, Tableau crea una instantánea de la hoja de trabajo. Se puede acceder a los marcadores desde cualquier libro utilizando el menú Marcadores. Cuando abre una hoja de trabajo marcada como favorita, agrega la hoja de trabajo a su libro de trabajo en el estado en que estaba cuando se marcó. Nunca se actualizará ni cambiará automáticamente. Los marcadores son convenientes cuando tiene hojas de trabajo que usa con frecuencia. Para guardar un marcador de Tableau:
- Seleccione Archivo > Marcador > Crear marcador.
  - Especifique el nombre y la ubicación del archivo de marcador en el cuadro de diálogo Crear marcador.
- Tableau guarda el archivo con la extensión .tmb. La ubicación predeterminada es la carpeta Marcadores en el repositorio de Tableau. Sin embargo, puede guardar marcadores en cualquier ubicación que elija. Los marcadores que no están almacenados en el repositorio de Tableau no aparecen en el menú Marcador.
- Es posible compartir libros de trabajo y marcadores con sus compañeros de trabajo, siempre que puedan acceder a las fuentes de datos relevantes que utiliza el libro de trabajo. Si sus compañeros de trabajo no tienen acceso a las fuentes de datos, puede guardar un libro de trabajo empaquetado.

Los campos personalizados como medidas agrupadas, campos calculados, grupos y conjuntos se guardan con libros de trabajo y marcadores.

5. Libros de trabajo empaquetados: estos libros contienen el libro de trabajo junto con una copia de cualquier fuente de datos de archivo local e imágenes de fondo. El libro de trabajo ya no está vinculado a las imágenes y las fuentes de datos originales. Estos libros de trabajo se guardan con una extensión de archivo .twbx. Otros usuarios pueden abrir el libro de trabajo empaquetado con Tableau Desktop o Tableau Reader.

- Cree un .twbx con fuentes de datos basadas en archivos
  1. Seleccione Archivo > Guardar como.
  2. Especifique un nombre de archivo para el libro empaquetado en el cuadro de diálogo Guardar como.
  3. Seleccione Libros de trabajo empaquetados de Tableau en la lista desplegable Guardar como tipo .
  4. Haga clic en Guardar . La ubicación predeterminada es la carpeta Workbooks del repositorio de Tableau. Sin embargo, puede guardar libros de trabajo empaquetados en cualquier directorio que elija.

Los siguientes archivos se incluyen en los libros de trabajo empaquetados:

- Imágenes de fondo.
- Geocodificación personalizada.
- Formas personalizadas.
- Archivos de cubo locales.
- Archivos de Microsoft Access.
- Archivos de Microsoft Excel.
- Archivos de extracción de Tableau (.hyper o .tde).
- Archivos de texto (.csv, .txt, etc.)

- Cree un .twbx con fuentes de datos no basadas en archivos Si el libro de trabajo contiene conexiones a fuentes de datos empresariales u otras fuentes de datos no basadas en archivos, como Microsoft SQL, Oracle o MySQL, los datos deben extraerse de las fuentes de datos para que se incluyan en un libro de trabajo empaquetado (.twbx ).

1. En el libro de trabajo, haga clic con el botón derecho en la fuente de datos en el panel Datos y elija Extraer datos.
2. En el cuadro de diálogo Extraer datos, haga clic en el botón Extraer para extraer todos los datos de la fuente de datos. Una vez que se completa la extracción, el icono de la fuente de datos cambia para indicar que hay una extracción activa para esa fuente de datos. En lugar de un solo cilindro, hay dos cilindros conectados por una flecha.
3. Opcional: repita los pasos anteriores para cada fuente de datos en el libro de trabajo.
4. Seleccione Archivo > Guardar como.
5. En el menú desplegable Guardar como tipo , seleccione Libro

de trabajo empaquetado de Tableau (\*.twbx). Una vez que se hayan creado los extractos para todas las fuentes de datos no basadas en archivos y se haya guardado el libro de trabajo empaquetado, puede enviar su libro de trabajo.

- Cree un .twbx con fuentes de datos de Tableau Server, si el libro de trabajo contiene conexiones a una fuente de datos de Tableau Server publicada, debe descargar una copia local de la fuente de datos de Tableau Server, tomar un extracto y luego reemplazar la conexión a la copia local para que se incluya en un libro de trabajo empaquetado. (.twbx).
  1. En el libro de trabajo, haga clic con el botón derecho en la fuente de datos publicada en el panel Datos y luego seleccione Crear copia local. Se agrega una copia de la fuente de datos publicada al panel Datos.
  2. Haga clic con el botón derecho en la copia local y seleccione Extraer datos.
  3. En el cuadro de diálogo Extraer datos, haga clic en el botón Extraer para extraer todos los datos de la fuente de datos. La creación de un extracto de la fuente de datos le permite a la persona con la que está compartiendo el libro tener acceso a una copia de la fuente de datos.
  4. En el panel Datos, haga clic con el botón derecho en la fuente de datos publicada y luego seleccione Reemplazar fuente de datos.
  5. Verifique que la fuente de datos publicada sea reemplazada por la fuente de datos local y luego haga clic en Aceptar.
  6. Haga clic con el botón derecho en la fuente de datos publicada y luego haga clic en Cerrar.
  7. Seleccione Archivo > Guardar como.
  8. En el menú desplegable Guardar como tipo , seleccione Libro de trabajo empaquetado de Tableau (\*.twbx).
- Desempaquetar un .twbx, los libros empaquetados se pueden descomprimir. En una computadora con Windows o macOS, cambie el nombre del archivo con una extensión .zip (por ejemplo, de myfile.twbx a myfile.zip) y luego haga doble clic en él. Cuando desempaquetas un libro de trabajo, obtiene un archivo de libro de trabajo normal (.twb), junto con una carpeta que contiene las fuentes de datos y las imágenes que se empaquetaron con el libro de trabajo.

Hay varias formas de obtener vistas y libros de trabajo de Tableau Desktop y convertirlos en una presentación, informe o página web.

- Copiar una vista como imagen, puede copiar rápidamente una vista individual como una imagen y pegarla en otra aplicación, como Microsoft Word o Excel. Si usa Tableau Desktop en macOS, se copia una imagen TIFF (formato de archivo de imagen con etiquetas) al portapapeles. En Windows, se copia una imagen BMP (mapa de bits).
  1. Seleccione Hoja de trabajo > Copiar > Imagen.

2. En el cuadro de diálogo Copiar imagen, seleccione los elementos que desea incluir en la imagen. Si la vista contiene una leyenda, en Opciones de imagen, seleccione el diseño de la leyenda.
  3. Haga clic en Copiar.
  4. Abra la aplicación de destino y pegue la imagen del portapapeles.
- Exportar una vista como un archivo de imagen, para crear un archivo de imagen que pueda reutilizar, exporte la vista en lugar de copiarla. Puede elegir el formato BMP, JPEG o PNG en macOS.
    1. Seleccione Hoja de trabajo > Exportar > Imagen.
    2. En el cuadro de diálogo Exportar imagen, seleccione los elementos que desea incluir en la imagen. Si la vista contiene una leyenda, en Opciones de imagen, seleccione el diseño de la leyenda.
    3. Haga clic en Guardar.
    4. En el cuadro de diálogo Guardar imagen, especifique la ubicación, el nombre y el formato del archivo. Luego haga clic en Guardar.
  - Exportar como una presentación de PowerPoint, cuando exporta un libro a formato de Microsoft PowerPoint, las hojas seleccionadas se convierten en imágenes PNG estáticas en diapositivas independientes. Si exporta una hoja de historia, todos los puntos de la historia se exportan como diapositivas independientes. Todos los filtros aplicados actualmente en Tableau se reflejan en la presentación exportada. Para exportar un libro a PowerPoint:
    1. Seleccione Archivo > Exportar como PowerPoint.
    2. Seleccione las hojas que desea incluir en la presentación. (También se pueden incluir hojas ocultas). El archivo de PowerPoint exportado refleja el nombre de archivo de su libro y la diapositiva de título indica el nombre del libro y la fecha en que se generó.
  - Exportar a PDF, para crear un archivo basado en vectores que incorpore las fuentes de Tableau, imprima en PDF. Después de personalizar el diseño de los elementos de la página mediante el cuadro de diálogo Archivo > Configurar página , elija Archivo > Imprimir en PDF.

Al crear, editar e interactuar con vistas en Tableau Server o Tableau Online, existen formas diferentes de guardar su trabajo:

1. Guardar un libro de trabajo: cuando crea un libro de trabajo nuevo o edita un libro de trabajo existente en Tableau Server o Tableau Online, puede guardar su trabajo en cualquier momento. Para guardar un libro de trabajo:
  - En el modo de edición web, seleccione Archivo > Guardar.
2. Guardar una copia de un libro de trabajo: A veces, no desea sobrescribir una vista existente con sus cambios. En casos como estos, puede guardar una copia de un libro de trabajo existente. Cuando hace esto, el libro de trabajo existente permanece sin cambios y se crea una copia para que pueda editarlo como deseé. Para guardar una copia de un libro de trabajo:
  - En el modo de edición web, seleccione Archivo > Guardar como .
  - En el cuadro de diálogo Guardar libro de trabajo que se abre, haga

lo siguiente:

- Para el nombre: introduzca un nombre para el libro.
  - Para proyecto: seleccione el proyecto en el que le gustaría guardar el libro de trabajo.
  - Haga clic en Guardar.
3. Guardar cambios como una vista personalizada: una vista personalizada no cambia la original, pero está relacionada con ella. Si la vista original se actualiza o se vuelve a publicar, la vista personalizada también se actualiza. También puede elegir si sus vistas personalizadas son visibles para otros usuarios (públicas) o solo para usted (privadas).

Cuando se usa Tableau Public solo se tiene la opción de guardar el libro de trabajo en el repositorio público, para el cual debe seguir estos pasos:

1. Con su libro de trabajo abierto en Tableau Desktop Public Edition, seleccione Archivo > Guardar en Tableau Public como...

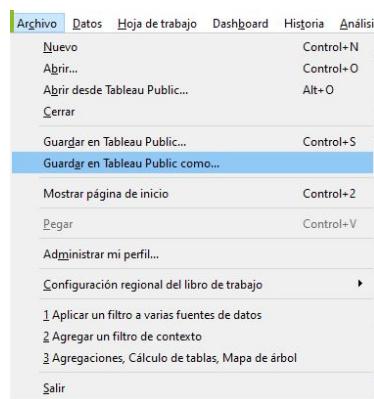


Figure 2.1: Guardar trabajo

2. Inicie sesión con su cuenta de Tableau Public. Si no tiene una cuenta, seleccione el enlace para crear una nueva.
3. Escriba un nombre para el libro y haga clic en Guardar. Cuando guarda un libro de trabajo en Tableau Public, el proceso de publicación crea un extracto de la conexión de datos.
4. Una vez publicado el libro de trabajo, se le redirige a su cuenta en el sitio web de Tableau Public.(El enlace se abre en una nueva ventana).
5. Una vez en el sitio web de la visualización haga clic en el botón con el icono de compartir y copie el enlace para agregarlo en el sitio web o donde desee publicar el trabajo.

Para obtener más información acerca de los pasos de publicación del trabajo puede visitar Ayuda de Tableau.



Figure 2.2: Iniciar sesión



Figure 2.3: Asignar nombre al libro de trabajo

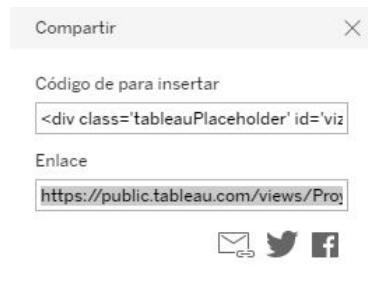


Figure 2.4: Compartir trabajo

## 2.2 Instalación de Tableau Desktop Public

El proceso de instalación de la versión de escritorio de Tableau Public se realiza mediante los siguientes pasos:

1. Dirigirse a la página principal de Tableau Public, ingrese un correo electrónico con el cual quiere vincular su descarga y cuenta de Tableau Public, luego clic en “DOWNLOAD THE APP”, inmediantamente se inicia la descarga.



Figure 2.5: Descarga de Tableau Public

2. Una vez se complete la descarga abra el archivo, lea los términos de licencia, seleccione “He leído y acepto los términos de acuerdo de licencia” y finalmente clic en “Instalar”.



Figure 2.6: Descarga de Tableau Public

3. Cuando termine el proceso de instalación se creará un acceso directo en el escritorio a la aplicación, una vez creadas las visualizaciones puede continuar con el proceso de publicación del trabajo mostrado en la sección 2.1.6.

## 2.3 Forma de navegación

Al momento de abrir Tableau esta es la pantalla con la que se encuentra, en el panel del lateral izquierdo encontrará el tipo de fuentes a las que se puede conectar, en la parte central se ubican los proyectos que ya se han realizado usando este software, en el panel lateral derecho encuentra videos paso a paso sobre conexión a datos y realización de gráficos, en la parte inferior de este panel encontrara visualizaciones alojadas en la galería de Tableau, conjunto de datos de muestra y capacitaciones.

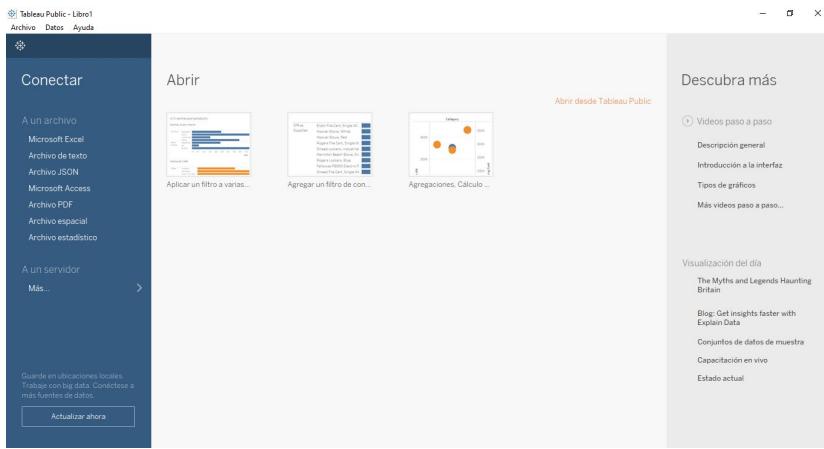


Figure 2.7: Página principal de Tableau

Sin conectarse a alguna fuente de datos puede hacer clic en el ícono de Tableau que se encuentra debajo de la pestaña archivo, se abrirá la siguiente pantalla:

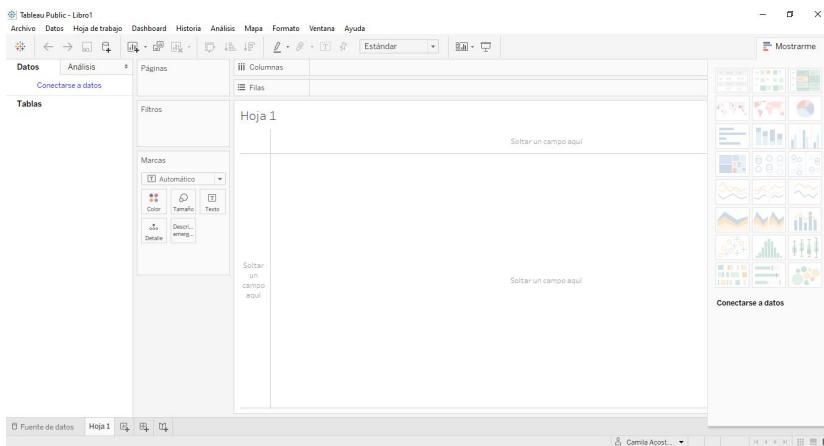


Figure 2.8: Área de trabajo

Esta es la ventana donde se pueden crear todas las visualizaciones, en la esquina superior izquierda se encuentra el nombre del libro de trabajo, recuerde que un libro de trabajo puede incluir hojas, dashboard o historias, después de esto se encuentran varias pestañas que permiten abrir libros de trabajo anteriores o guardar el trabajo que se está creando, permite conectarnos a nuevas fuentes de datos, crear hojas de trabajo dashboard e historias, editar formatos de mapas y ventanas, y finalmente una pestaña de ayuda, en la cual encontrara soporte, configuraciones de idioma y capacitaciones. Luego esta ubicada la barra de herramientas esta contiene diferentes botones como el icono de Tableau que permite navegar hacia la página de inicio que se muestra en la figura 2.7, posee botones de deshacer y rehacer, guardar y conectar a una nueva fuente de datos, también posee botones para agregar, duplicar o eliminar hojas de trabajo, intercambiar medidas, organizar de forma descendente o ascendente, opciones de texto, de tamaño y para ocultar o visualizar tarjetas.

En el panel lateral izquierdo en la pestaña datos encontrara el nombre de la fuente de datos con la que tiene conexión, en la parte tabla se ubican el nombre de todas las variables que contenga la base de datos, estas variables se dividen en dimensiones y medidas, con dimensiones se refiere a todas las variables categóricas que contenga la base y medidas se refiere a las columnas con datos numéricos, estos dos tipos de variables son las que se arrastran al lienzo en blanco que se encuentra en la mitad de la pantalla para crear las visualizaciones. Si en los datos subyacentes no se incluyen todos los campos que necesita para responder a las preguntas, puede crear nuevos campos en Tableau usando cálculos y luego guardarlos como parte de la fuente de datos. Estos campos se llaman campos calculados. También existe la posibilidad de crear conjuntos, que son campos personalizados que se crean a partir de dimensiones y especificaciones realizadas por el usuario, a demás de estos tipos de datos ya mencionados existen los parámetros que sin valores que pueden usarse como marcadores de posición en fórmulas, o sustituir valores constantes en campos calculados y filtros. En la pestaña Análisis este panel permite agregar líneas constantes, de promedio, diagramas de cajas y bigotes, pronósticos de líneas de tendencia y otros elementos a la vista, las opciones de esta pestaña se muestran en la figura 2.9, algunas de estas opciones serán exploradas en la sección 2.4.2.

Al lado derecho del panel lateral que se describió anteriormente se encuentran los estantes Páginas y Filtro y la tarjeta Marcas. El estante Páginas permite dividir una vista en una serie de páginas para que pueda analizar mejor cómo un campo específico afecta al resto de los datos en una vista. Cuando coloca una dimensión en el estante Páginas, está añadiendo una nueva fila por cada miembro de la dimensión. Cuando coloca una medida en el estante Páginas, Tableau convierte la medida automáticamente en una medida discreta. El estante Filtros le permite especificar qué datos incluir y excluir, Puede filtrar los datos usando medidas, dimensiones o ambas al mismo tiempo. Finalmente se encuentra La tarjeta Marcas que es un elemento fundamental del análisis visual en Tableau. Al arrastrar campos a distintas propiedades en la tarjeta Marcas, puede añadir contexto y detalles a las marcas de la vista, esta tarjeta sirve para definir el tipo

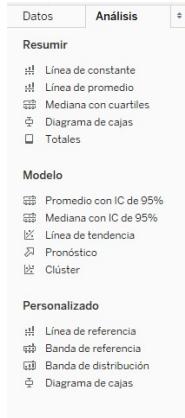


Figure 2.9: Opciones del panel análisis

de marca, esto se refiere a la forma de los datos en la visualización, las marcas disponibles se muestran en la figura 2.10, también es posible el color, el tamaño, la forma, el texto y los detalles de los datos.



Figure 2.10: Opciones de marca

En la parte central se encuentran ubicados los estantes Columnas y Filas, estos permiten dominar el eje x y eje y respectivamente. El estante Columnas crea las columnas de una tabla, mientras que el estante Filas crea las filas. Puede colocar todos los campos que quiera en estos estantes. Al colocar una dimensión en los estantes Filas o Columnas, se crean los encabezados de los miembros de dicha dimensión. Al colocar una medida en el estante Filas o Columnas, se crean ejes cuantitativos para esa medida. A medida que agrega más campos a la vista, se incluyen encabezados y ejes adicionales en la tabla y obtiene una imagen cada vez más detallada de sus datos.

La creación de vistas en Tableau es muy sencilla, existen dos opciones principales; la primera es usando los estantes de filas o columnas para añadir las medidas y las dimensiones, la segunda opción consiste en seleccionar los campos que quiere incluir en la vista y luego dar clic en el botón “Mostrarme” que se ubica en la esquina superior derecha de la barra de herramientas, le permite elegir un tipo de vista resaltando los tipos de vista que mejor se adapten a los tipos de campo que ha seleccionado de sus datos. Alrededor del tipo de gráfico más adecuado para sus datos aparece un contorno de color naranja, los tipos de vista disponibles con esta funcionalidad se presentan en la figura 2.11.



Figure 2.11: Opciones de la funcionalidad Mostrarme

En la parte inferior del área de trabajo se ubican 5 compartimentos, el primero llamado Fuente de datos, permite conectarse a una nueva fuente de datos o en el caso de no estar conectado a una lo dirige a la página principal donde se puede hacer la conexión, después se sitúa la hoja de trabajo que se está usando en el momento, a continuación, se encuentran las opciones de agregar una nueva hoja, dashboard o historia. Como se había mencionado anteriormente los libros de trabajo pueden estar compuestos de hojas, dashboards o historias. Una hoja de trabajo es donde se crean vistas de sus datos al arrastrar y soltar campos en los estantes, contiene una sola vista con estantes, tarjetas, leyendas y los paneles Datos y Análisis en la barra lateral. Un dashboard es una combinación de varias vistas que puede organizar para presentación o para supervisar. Una historia es una secuencia de vistas o dashboards que se utilizan de forma conjunta para mostrar información.

## 2.4 Flujo de trabajo

### 2.4.1 Conexión a fuentes de datos

Antes de poder crear y analizar los datos debe conectar Tableau a estos, en este caso la conexión se hará a través de un archivo de Excel, inicialmente se hará la conexión a las bases de datos de estudiantes graduados a nivel de micro datos para mostrar las funcionalidades de unión que tiene Tableau, para hacer estas conexiones debe seguir estos pasos:

1. Abrir Tableau desde el acceso directo creado en su escritorio al momento de la instalación, la pantalla que se debe ver es la mostrada en la figura 2.7.
2. Hacer clic en el botón “Microsoft Excel” si su archivo posee este formato, al hacer clic en este botón se abre una ventana que permite navegar a través de las carpetas de su equipo para ubicar la localización de las bases de datos. Debe seleccionar una de las bases y dar clic en el botón “Abrir”.

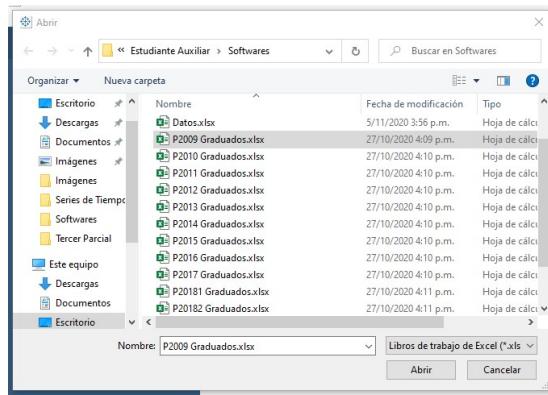


Figure 2.12: Navegación entre carpetas

Con esta conexión a la fuente de datos se obtiene la siguiente pantalla

En el panel lateral izquierdo encontrara el nombre del archivo al que se conectó en este caso “P2009 Graduados”, debajo de esto se ubican las hojas que componen el archivo para esta base solo se tiene una hoja llamada “P2009G”, luego se encuentra un botón llamado “Nueva unión”. La parte central de la conexión a datos es el lienzo en blanco dispuesto en la parte superior allí se deben arrastrar las hojas a las que se quiere conectar, en la parte inferior se encuentra una vista previa de la base de datos, los campos marcados con “#” indica que son medidas, “Abc” indica que el campo es una dimensión y finalmente las variables relacionadas con ubicaciones geográficas como latitud y longitud tiene como icono un globo terráqueo, haciendo clic sobre estos iconos se puede editar el tipo de dato, por ejemplo la variable Snies Sede Mat Tableau la tomo como

The screenshot shows the Tableau Data Source preview interface. On the left, there's a sidebar with 'Conexiones' (Connections) and 'P2009 Graduados' selected. Below it, 'Hojas' (Sheets) and 'Nueva unión' (New Union) are listed. A note says 'Usar el intérprete de datos' (Use the data interpreter) and 'Puedes que el intérprete de datos esté disponible para importar su libro de trabajo Microsoft Excel.' (The data interpreter may be available to import your Microsoft Excel workbook.) At the top right are 'Filtros' (Filters) and 'Añadir' (Add). The main area is titled 'P2009G' and shows a preview of the data with columns like 'ID', 'TÍD', 'Year', 'Pcodig', 'Semestre', 'Año', 'Tip Nivel', 'Nivel', 'Pcodig', 'Dep Nac', 'Pcodig', 'Cod Dep Nac', 'Pcodig', 'Clu Nac', 'Pcodig', 'Cod Clu Nac', and 'Lor'. A message at the bottom says 'Necesita más datos' (Needs more data) and 'Arrastre tablas para establecer relaciones entre ellas. Más información' (Drag tables to establish relationships between them. More information).

Figure 2.13: Vista previa de la conexión

una medida cuando en realidad esta variable hace referencia a la categorización establecida para las sedes de la universidad, la podemos editar dando clic en el icono “#” y en el menú desplegable seleccionar Cadena.

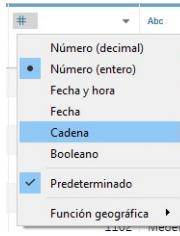


Figure 2.14: Cambiar el tipo de un campo

En la esquina superior derecha del lienzo, se observa una etiqueta llamada Filtros y un botón añadir, al hacer clic en este botón se abre un cuadro de diálogo que permite añadir, editar o eliminar filtros, a modo de ejemplo se creara un filtro que seleccione únicamente las filas en las que el campo Sede Nombre Adm sea Medellín,

- Clic en el botón añadir ubicado en la parte superior derecha del lienzo.
- En el cuadro de dialogo hacer clic en “Añadir”.
- Terminado el paso anterior se abre una nueva ventana que contiene el nombre de todas las columnas de la base de datos, aquí se debe seleccionar la columna por la que se quiere filtrar, en este caso Sede Nombre Adm y dar clic en aceptar.
- Con esto se abre una nueva pestaña que contiene los valores de la columna seleccionada para filtrar, para el ejemplo se selección Medellín y finalmente Aceptar.

Ahora la vista previa de la base se modificó, solo contiene las observaciones en

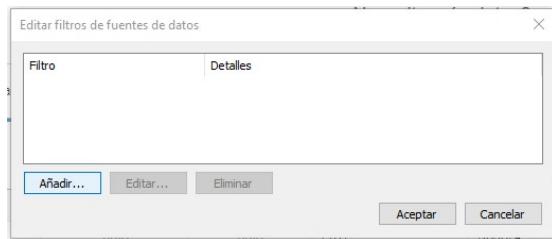


Figure 2.15: Crear un filtro

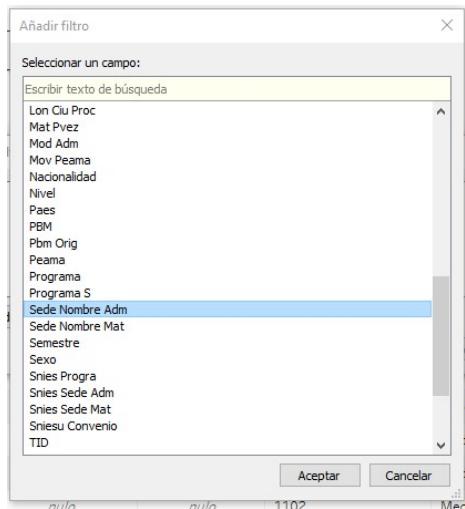


Figure 2.16: Nombre de los campos

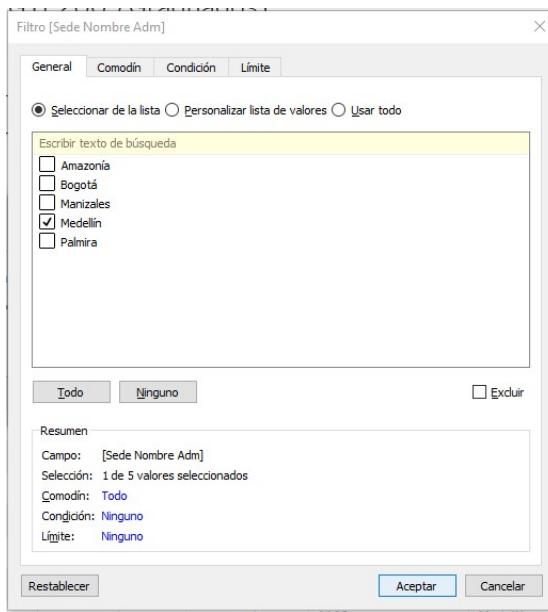


Figure 2.17: Valores de la columna seleccionada

las cuales se cumple el filtro aplicado, es decir donde Sede Nombre sea Medellín.

Este mismo tipo de filtros se puede aplicar a medidas, definiendo un intervalo para los valores o seleccionando un valor mínimo o máximo o un cálculo especial. Otra opción ofrecida por Tableau es hacer uniones entre tablas de datos, esto es útil cuando se tiene la información en distintas bases de datos con la misma estructura como en este caso que se tiene la información de estudiantes graduados desde el año 2009 hasta el 2020 semestre 1 a nivel de microdatos, es decir los archivos están separados.

Existen dos métodos básicos para combinar conjuntos de datos en Tableau la unión de columnas y la unión de filas, se pueden combinar las columnas de dos conjuntos de datos o bien filas de dos o mas conjuntos de datos, primero debe conectarse a las tablas que desea combinar cabe aclarar que estas tablas deben pertenecer al mismo archivo, en este caso como se está usando archivos Excel las tablas a unir deben ser dos hojas del archivo, a modo de ejemplo se creara un archivo en Excel que contenga en una hoja la información de los graduados en el año 2009 y en otra hoja los registros de los graduados en el año 2010 y se hará la conexión a los datos, cuando realice dicha conexión en el panel lateral izquierdo se ubica el nombre del archivo Excel, en este caso llamado “P2009-2010 Graduados” y más abajo el nombre de las dos hojas que contiene dicho archivo llamadas “P2009G” y “P2010G”, como se ilustra en la figura 2.18.



Figure 2.18: Nombre del archivo y hojas

La unión por columnas es útil cuando se quiere trabajar con dos columnas que se encuentran en diferentes conjuntos de datos, existen cuatro formas de realizar las uniones por columnas:

- Interior: devuelve únicamente los registros que están presentes en ambas tablas.
- Izquierda: devuelve todos los registros de la tabla de la izquierda y solo los registros que coinciden con la tabla de la derecha.
- Derecha: devuelve todos los registros de la tabla de la derecha y solo los registros que coinciden con la tabla de la izquierda.
- Exterior: devuelve todos los registros de ambas tablas.

En los conjuntos de datos que se están usando todos tienen las mismas columnas por lo que el interés se centra en realizar unión por filas y no por columnas para realizar este tipo de unión se tiene dos opciones:

1. Arrastrar y soltar; este método consiste en arrastrar la primera hoja al lienzo, arrastrar la otra hoja que se quiere unir y no soltar hasta que aparezca el cuadro “Unión de filas” en color naranja, solo se debe soltar la hoja cuando este cuadro aparezca.
2. Usando el panel “Nueva unión”; este panel se encuentra ubicado en el lateral izquierdo justo debajo del nombre de las hojas como se muestra en la figura 2.18. Para usar esta opción se debe hacer doble clic en este panel, se abre una ventana en la cual se debe asegurar que este seleccionado Específico (manual), se deben arrastrar las hojas a unir al espacio en blanco que tiene esta ventana, finalmente hacer clic en Aceptar.

Tableau crea dos columnas adicionales a las que contiene a la base que ayuda

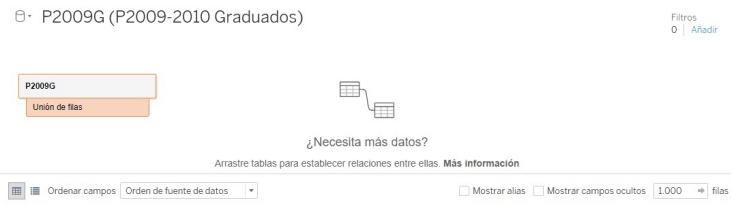


Figure 2.19: Unión: Método de arrastrar y soltar

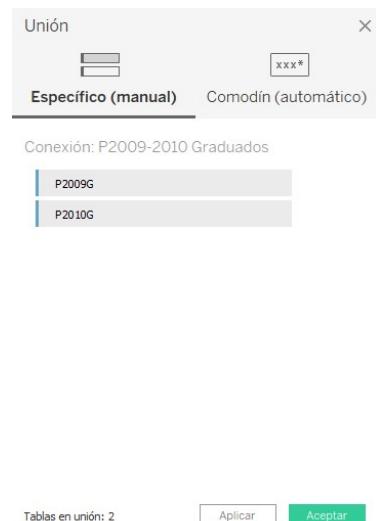


Figure 2.20: Unión: Método Nueva unión

a la identificación de la hoja y tabla a la que pertenecen las observaciones; si se edita la cantidad de filas que se muestra en la vista previa del conjunto de datos se puede observar los registros de ambas hojas.

Abc PROGRAMA	Abc PROGRAMA_S	Abc AREAC_SNIES	Abc CA_CINE	# CD_CINE	Abc AREA_CINE	Abc Sheet	Abc Table Name
Ciencias agrarias	Ciencias agrarias - Bo...	Agronomía, veterinar...	811		8 Agricultura, silvicultu...	P2009G	P2009G
Salud pública	Salud pública - Bogotá	Ciencias de la salud	912		9 Salud y bienestar	P2009G	P2009G
Ciencias agrarias	Ciencias agrarias - Pal...	Agronomía, veterinar...	nulo		nulo	P2009G	P2009G
Enfermería	Enfermería -Bogotá	Ciencias de la salud	913		9 Salud y bienestar	P2010G	P2010G
Ciencias humanas y e...	Ciencias humanas y e...	Ciencias sociales y hu...	312		3 Ciencias sociales, peri...	P2010G	P2010G
Ciencias humanas	Ciencias humanas - Bo...	Ciencias sociales y hu...	314		3 Ciencias sociales, peri...	P2010G	P2010G
Enfermería	Enfermería -Bogotá	Ciencias de la salud	913		9 Salud y bienestar	P2010G	P2010G
Artes	Artes -Bogotá	Bellas artes	nulo		2 Artes y humanidades	P2010G	P2010G

Figure 2.21: Columnas Nuevas

Siguiendo con un análisis detallado de lo mostrado por Tableau en la vista previa del conjunto de datos, se observan columnas problemáticas, la variable Ciu\_Nac presenta una combinación de números y texto como se muestra en la figura 2.4.1, existe una función llamada división personalizada que se puede ver al hacer clic en el menú desplegable de la columna, dicha función necesita un separador para hacer la división pero en este caso no es posible usarla ya que no existe separador alguno entre el numero y el texto, esto en un problema que no puede solucionarse desde Tableau.

\begin{figure}

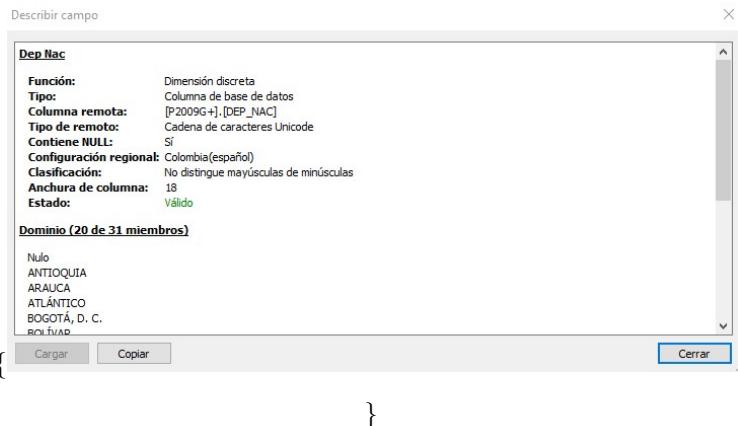
# P2009G+ ID	Abc P2009G+ TID	# P2009G+ Year	# P2009G+ Semestre	Abc P2009G+ Tipo Nivel	Abc P2009G+ Nivel	Abc P2009G+ Dep Nac	Abc P2009G+ Cod Dep Nac	Abc P2009G+ Ciu Nac	▼
1 CC		2009		1 Pregrado	Pregrado	CUNDINAMARCA	25	SOPÓ	
2 CC		2009		1 Pregrado	Pregrado	BOGOTÁ, D. C.	11	11BOGOTÁ D.C.	
3 CC		2009		1 Pregrado	Pregrado	BOGOTÁ, D. C.	11	11BOGOTÁ D.C.	
4 CC		2009		1 Pregrado	Pregrado	nulo	nulo	nulo	
5 CC		2009		1 Pregrado	Pregrado	ANTIOQUIA	5	05MEDELLIN	
6 CC		2009		1 Pregrado	Pregrado	BOLÍVAR	13	13CARTAGENA	
7 CC		2009		1 Pregrado	Pregrado	ANTIOQUIA	5	05MEDELLIN	
8 CC		2009		1 Pregrado	Pregrado	BOYACÁ	15	15TUNJA	

}

\caption{Problema columna Ciu\_Nac} \end{figure}

En este mismo menú desplegable se encuentra una opción llamada Describir, al dar clic en esta opción Tableau abre una ventana que muestra una descripción corta de la base de datos, algo similar a la función summary() de R. En la figura 2.4.1 se muestra la descripción de la variable Dep\_Nac, esta permite visualizar el tipo de campo en este caso es dícreto, contiene valores faltantes y en la parte inferior muestra una lista de los miembros más dominantes en este caso 20 de los 31 miembros totales.

\begin{figure}



\caption{Descripción columna Dep\_Nac} \end{figure} Hacia las ultimas columnas de la base de datos nos encontramos con un campo llamado Programa\_S, esta compuesto por el nombre del programa y la sede a la que pertenece estos dos atributos se encuentran separados por un guion, en este caso particular si es posible usar la función División, ya que el campo posee un separador.

Hacer clic en el menú desplegable de la columna de interés y seleccionar División.

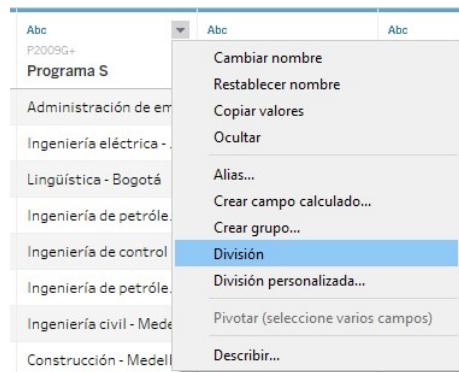


Figure 2.22: División de columnas

Con esto se obtiene una nueva columna llamada Programa\_S División 1 que contiene el nombre del programa, es decir que elimino todo lo que se encontraba a la derecha del guion (sede), como se puede observar en la descripción de esta nueva variable.

En el caso en que se quiera obtener ambas columnas, es decir una columna que contenga el programa y otra que contenga la sede es necesario usar División Personalizada, que también se encuentra en el menú desplegable de la columna.

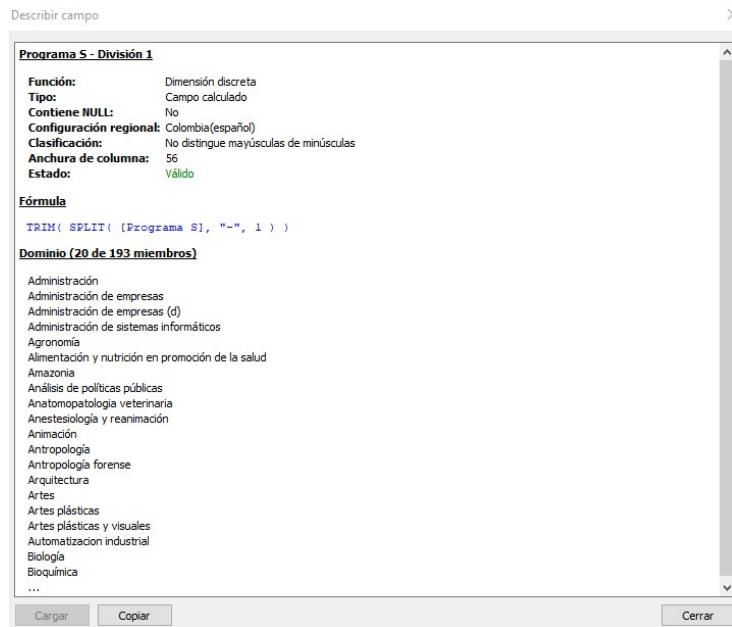


Figure 2.23: Descripción de división

1. Hacer clic en el menú desplegable de la columna y seleccionar división personalizada.
2. En el campo Usar separador escribir – que es el separador de la columna de interés, en el campo División Desactivada se debe seleccionar Todas, para poder obtener las columnas de programa y Sede.

Finalmente se obtiene tres columnas una que contiene el programa, otra la sede y una que aparentemente posee los espacios.

En este caso se considera correcto eliminar la columna División 3 ya que la información relevante de la columna original se encuentra almacenada en los campos llamados División 1 y 2.

Una funcionalidad importante que también que se ubica en el menú desplegable de las columnas de tipo numérico como Edad\_Mod es Crear grupos que permite agrupar las edades de los graduados en categorías, algo similar a los que se tiene en la columna Cat\_Edad.

1. Hacer clic en el menú desplegable de la columna Edad\_Mod y seleccionar Crear grupos.
2. El primer grupo estará conformado por las edades de 23 o menos años, para esto con ctrl sostenido y clic seleccionamos las edades que cumplen esta condición, luego clic en Grupo y se edita el nombre del grupo.

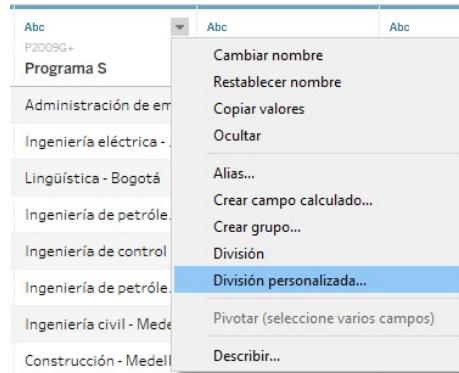


Figure 2.24: División personalizada

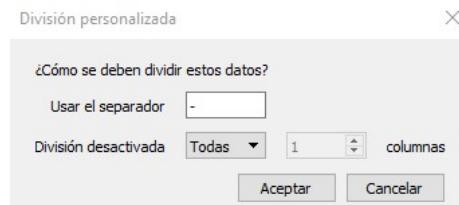


Figure 2.25: Ventana de división personalizada

Programa S - División 1	Programa S - División 2	Programa S - División 3
Administración de empresas	Bogotá	
Ingeniería eléctrica	Bogotá	
Lingüística	Bogotá	
Ingeniería de petróleos	Medellín	
Ingeniería de control	Medellín	
Ingeniería de petróleos	Medellín	
Ingeniería civil	Medellín	
Construcción	Medellín	

Figure 2.26: Campos divididos

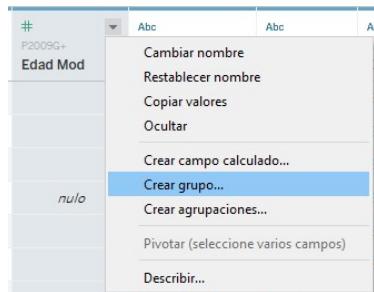


Figure 2.27: Crear grupos

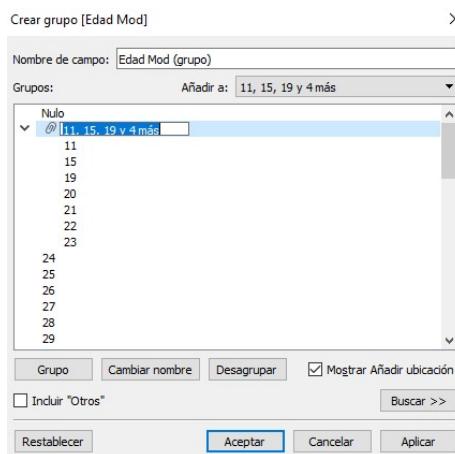


Figure 2.28: Creación de grupos de edad

3. Repita el paso anterior hasta crear las categorías de edad que considere pertinentes y luego de clic en Aceptar, con esto se obtiene una columna llamada Edad\_Mod (grupo), que asigna a cada observación el grupo que pertenece.

\begin{figure}

# P2009G+ Edad Mod	Grupo Edad Mod (grupo)
24	24 a 25 años
26	26 a 30 años
27	26 a 30 años
nu/o	nu/o
25	24 a 25 años
24	24 a 25 años
28	26 a 30 años
28	26 a 30 años

}

\caption{Agrupamiento de la columna Edad\_Mod} \end{figure} Para el campo Estrato\_Orig también es posible crear grupos para obtener la categorización mostrada en el campo Estrato.

Tableau no permite hacer una limpieza y preparación de datos, por tal razón la limpieza del conjunto de datos se realizó en R, se eliminaron columnas innecesarias, esto se podía hacer desde Tableau con la opción de ocultar columnas en la vista previa del archivo de datos, pero se decidió hacer en R ya que era necesario analizar la cantidad de valores faltantes por columnas y en base a esto se eliminaron; también se corrigieron espacios, mayúsculas, números y ortografía de algunas columnas de cadenas de caracteres ya que tenían varios valores que en realidad eran iguales pero por tildes, espacios o mayúsculas se contaban como diferentes. Se realizó una unión de todas las bases ya que se tenía a nivel de microdatos, el archivo a conectar con Tableau es llamado “Datos.xlsx”.

La primera observación al conectarse al conjunto de datos Datos.xlsx es que hay algunas columnas que no se tomaron como debería, por ejemplo, las variables relacionadas con la ubicación geográfica Latitud y longitud de la ciudad de nacimiento, para esto se debe,

1. Hacer clic en el icono “Abc” y en el menú que se despliega seleccionar Número(decimal).
2. Cuando el icono del campo sea “#”, hacer clic nuevamente en este icono y seleccionar Función geográfica, luego clic en latitud para el caso de la variable Lat\_Ciu\_Nac y longitud para la otra variable.

La variable Edad\_Mod debe ser numérica y no una dimensión discreta, por tanto se debe editar seleccionando Número (entero), la variable Snies\_Progra

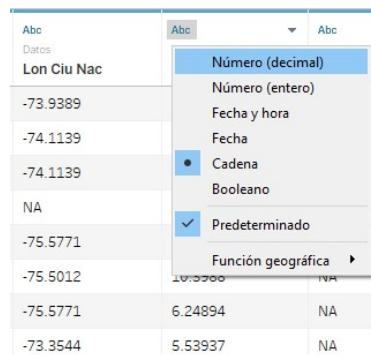


Figure 2.29: Cambiar tipo de columna

			Edad Mod	Cat E
Datos	Lon Ciu Nac	Número (decimal)	24	24 a
	-73.93889	Número (entero)	26	26 o
	-74.11392	Fecha y hora		
	-74.11392	Fecha		
	na/o	Cadena		
	-75.57706	Booleano		
	-75.50123	Predeterminado		
	-75.57706	Función geográfica		
	6,25 NA			
	-73,35436			
	5,54 NA			

Figure 2.30: Convertir a latitud un campo

es numérica y debe ser discreta ya que se refiere a una categorización de los programas, editela seleccionando Cadena en el menú desplegable del icono “#”. Luego de tener la base de datos lista hacer clic en Hoja1, recuadro que aparece en la parte inferior izquierda. Ya en el lienzo de trabajo en el panel lateral izquierdo en Tablas se ubican el nombre de todas las columnas de la fuente de datos como se mencionó en la sección 2.3, recuerde que los iconos azules se refieren a dimensiones y los verdes a medidas, en la parte inferior de este panel se encuentran ubicados cuatro variables numéricas dos de ellas son Semestre y Snies\_Progra que eran originales del conjunto de datos, pero hay otros dos campos llamados Datos (Recuento) y Valores de medias, estos dos campos fueron creados de manera automática por Tableau; Datos (Recuento) se refiere al total de filas de la base de datos en este caso 101.840, Valores de medida contiene un recuento de las dos variables numéricas leídas por Tableau, el Semestre y el total de datos, la suma de semestre es 151.469.



Figure 2.31: Nombre de columnas desde el panel Tablas

Realmente estas variables no son de interés en el análisis por lo que se ocultara Datos (Recuento) haciendo clic en el menú desplegable del campo y seleccionando ocultar, el campo Valores de medidas no es posible ocultarlo o eliminarlo, se dejara allí pero no se usara como campo para la realización de gráficos.

El campo Semestre debe ser arrastrado hacia la parte superior para convertirlo en dimensión, luego de verificar que los campos hayan sido leídos correctamente por Tableau es momento de iniciar con las visualizaciones.

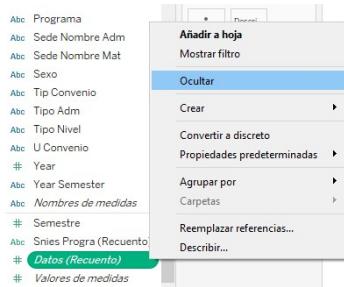


Figure 2.32: Ocultar columnas innecesarias

## 2.4.2 Análisis de datos

### 2.4.2.1 Gráfico de líneas

Se iniciara con gráficos similares a los presentados en la sección cifras generales y graduados la página de las estadísticas de la Universidad Nacional, en principio se hará un gráfico de líneas que muestre la evolución histórica de los estudiantes graduados en los periodos de 2009-1 a 2020-1.

1. Tome la columna Year-Semester, arrástrela hasta el estante columnas, tome nuevamente este campo y arrástrelo al estante filas.



Figure 2.33: Campos en los estantes columnas y filas

2. En el campo Year-Semester ubicado en el estante filas, haga clic en el menú desplegable y seleccione medida y recuento.

Con esto se obtiene un grafico de barras, donde cada barra representa el periodo y la altura de dicha barra en la cantidad de estudiantes graduados en ese periodo.

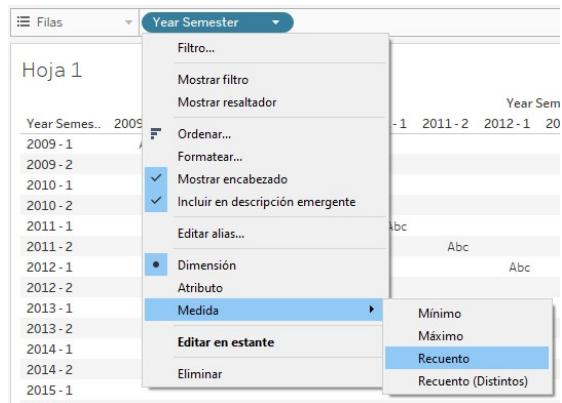


Figure 2.34: Editar agregación en el estante filas

3. Se quiere realizar un gráfico de líneas y no de barras, para cambiarlo en el estante marcas haga clic en el menú desplegable que en el momento se encuentra en “Automático” y seleccione línea.

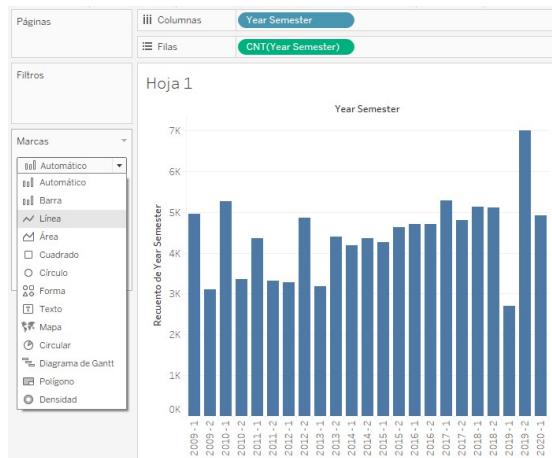


Figure 2.35: Cambiar la marca Automático por Línea

Hasta el momento la visualización se ve de esta manera.

4. Observe que hay un espacio vacío en el lienzo para ajustar la visualización a todo el lienzo, ubíquese en la barra de herramientas y haga clic en el menú desplegable de “Estándar” y seleccione “Ajustar anchura”, con esto el grafico de líneas ocupara todo el lienzo.
5. Observe que hay detalles como el título del gráfico, títulos de los ejes que no están claros, para agregar un título a la visualización hay dos opciones:

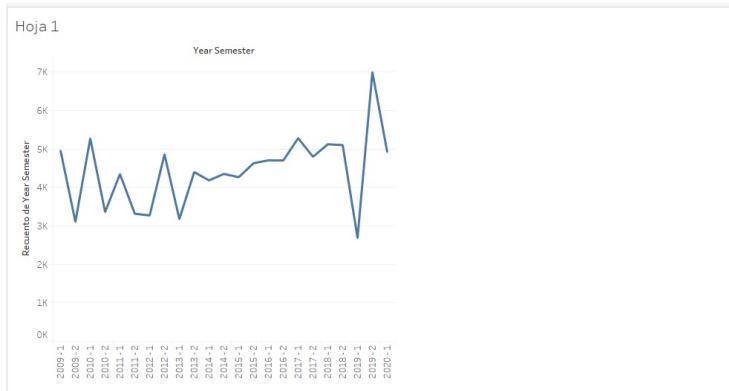


Figure 2.36: Vista previa de la visualización



Figure 2.37: Ajustar tamaño de la visualización

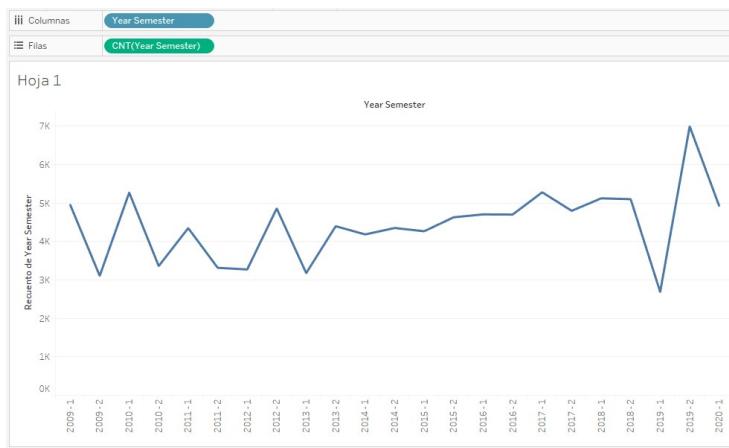


Figure 2.38: Vista en el lienzo completo

- Asignar un nombre a la hoja de trabajo en la parte inferior izquierda donde dice hoja 1, para esto haga doble clic sobre Hoja 1 y escriba el título que desea para su visualización, por ejemplo, evolución histórica del total de estudiantes graduados.
- Hacer clic derecho en el título de la visualización que en este momento es hoja 1 y seleccionar editar título, con esto se abre un cuadro de dialogo que permite escribir y editar el tipo de fuente, tamaño y color del texto.

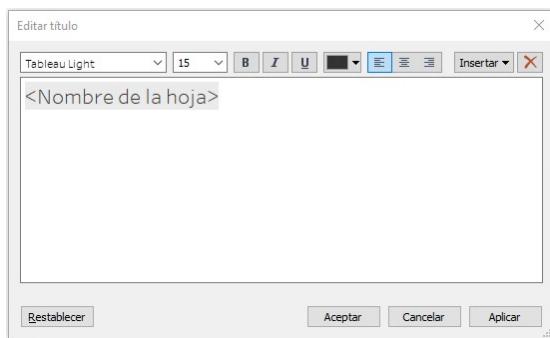


Figure 2.39: Editar título de la vista

Elimine el texto y escriba Evolución histórica del total de estudiantes graduados.

- Para editar el título del eje Y, haga clic derecho sobre el y seleccione editar eje, en el cuadro de dialogo cambie el título del eje por Número de estudiantes graduados y para el subtítulo haga clic en el cuadro Automático y escriba k: miles en el espacio para subtítulo.

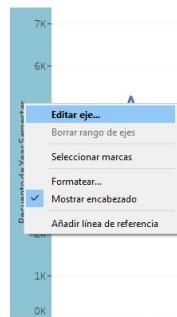


Figure 2.40: Editar eje Y

- Para el eje X que en este caso es llamado Year Semester no es posible editarlo como el caso del eje Y por tanto se debe ocultar, para esto haga clic derecho sobre esta etiqueta y seleccione Ocultar etiquetas de campo para columnas.



Figure 2.41: Editar título del eje Y

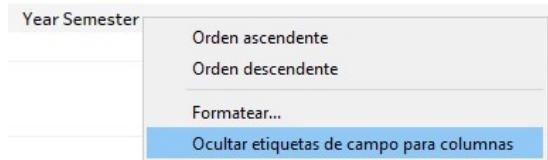


Figure 2.42: Ocultar título del eje X

Se obtiene la siguiente visualización.

8. Cuando se pasa el puntero por la línea, aparece un cuadro que contiene la información del periodo y el recuento de estudiantes graduados, pero la información en este cuadro no coincide con el nombre del eje Y.

Para que esta descripción sea más clara se debe cambiar Year Smester por Periodo y Recuento de Year Semester por Número de estudiantes graduados, en la tarjeta marcas haga clic en el recuadro que dice descripción emergente, con esto se abre una ventana que contiene la información del recuadro mostrado en la figura 2.44, en este cuadro cambie cambiar Year Smester por Periodo, Recuento de Year Semester por Número de estudiantes graduados y luego clic en Aceptar.

Ahora la descripción emergente de la visualización es mucho más clara.

Estos son los pasos básicos para hacer que las visualizaciones creadas en Tableau se vean claras y estéticas.

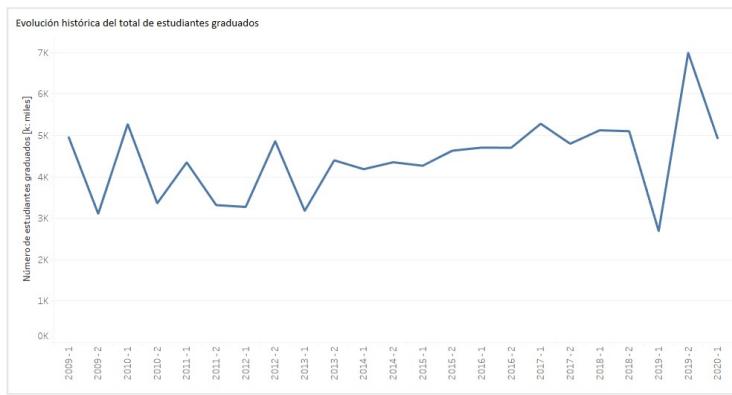


Figure 2.43: Vista previa de la visualización con algunas ediciones



Figure 2.44: Descripción emergente

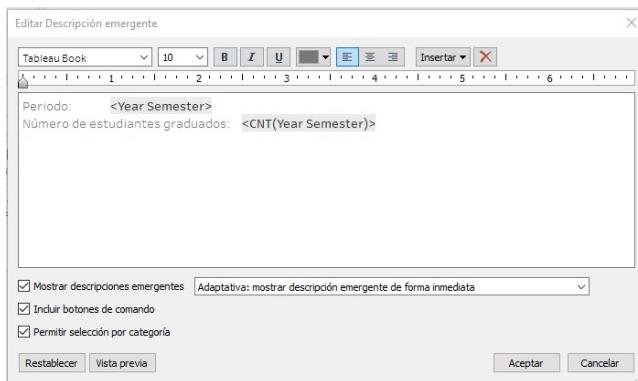


Figure 2.45: Editar información de la descripción emergente



Figure 2.46: Descripción emergente editada

#### 2.4.2.2 Gráfico de líneas segmentado por una dimensión

La siguiente visualización que se encuentra en la página de estadísticas de la Universidad Nacional de Colombia, en un gráfico de líneas por modalidad de formación, una tabla que contiene la misma información y un gráfico circular que contiene la información por modalidad de formación para el periodo actual es decir 2020-1.

1. Repita los pasos 1, 2, 3, 4 presentados en 2.4.2.1.
2. En el paso 5 mostrado en la sección 2.4.2.1, en el nombre de la hoja escriba Serie y en el título de la visualización escriba Evolución del número de estudiantes graduados por modalidad de formación.
3. Edite los ejes X y Y como se mostro en los pasos 6 y 7 de 2.4.2.1.
4. Arrastre el campo Tipo Nivel a color en el estante Marcas, con esto se crean dos líneas en el gráfico, una para los estudiantes graduados de Pregrado y otra para los graduados de Postgrado.



Figure 2.47: Gráfico de líneas por Tipo Nivel

5. Para cambiar los colores, clic en color, editar colores; en la ventana emergente que se abre seleccione Postgrado y clic en el color naranja, seleccione Pregrado y clic en el color verde, finalmente clic en aceptar.
6. En la leyenda de colores ubicada en el panel lateral derecho de la visualización, como se muestra en la figura 2.47, haga clic en el menú desplegable y seleccione Editar título, en el cuadro de texto borre Tipo Nivel y escriba Modalidad de formación y clic en aceptar.
7. Finalmente, la tarjeta de descripción emergente no es clara por lo que es necesario editarla, para esto siga el paso 8 de 2.4.2.1; cambiando Tipo Nivel por Modalidad de formación, Year Semester por Periodo y Recuento de Year Semester por Total Graduados.

La visualización obtenida es:

Otra opción para crear esta vista de líneas segmentada es duplicando la

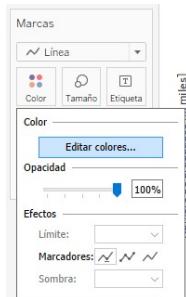


Figure 2.48: Editar colores

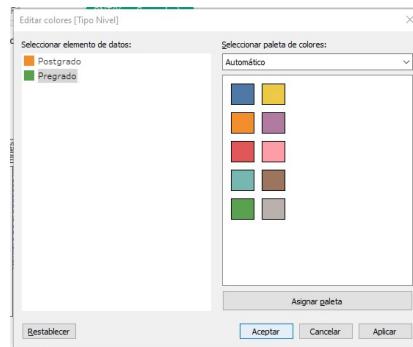


Figure 2.49: Asignación de colores a Tipo Nivel

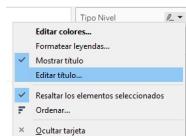


Figure 2.50: Editar título de leyenda de colores

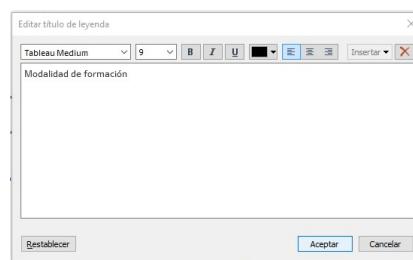


Figure 2.51: Asignación de título a leyenda de colores

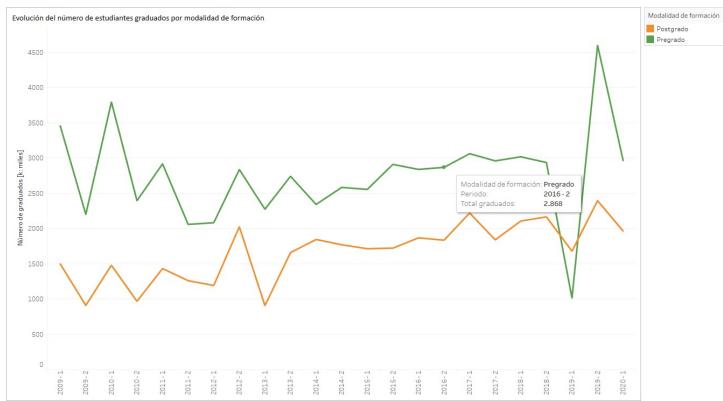


Figure 2.52: Visualización de la evolución de estudiantes graduados por modalidad de formación

visualización creada en la sección anterior, es decir el grafico de líneas, para duplicar esta vista ubíquese en la hoja llamada Evolución histórica del total de estudiantes graduados y en la barra de herramientas haga clic en el ícono que tiene dos hojas, con esto se crea una nueva hoja que contiene un duplicado del grafico de líneas.

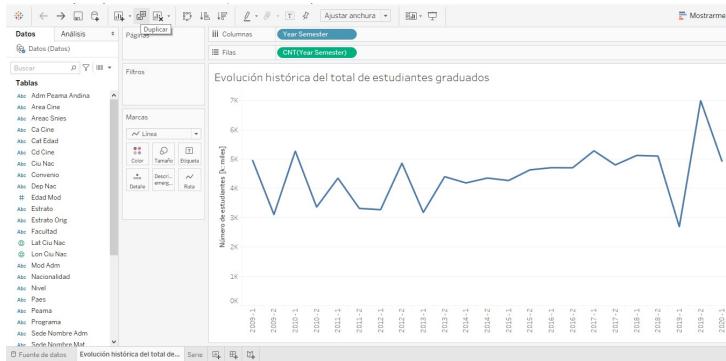


Figure 2.53: Duplicar hojas de trabajo

El título de la visualización y de la hoja duplicada es Evolución histórica del total de estudiantes graduados (2), por lo que debe cambiar el nombre de la hoja a Serie y el título de visualización a Evolución del número de estudiantes graduados por modalidad de formación. Finalmente repita los pasos 4, 5 y 6 mostrados en 2.4.2.2, la visualización se ve de esta manera.

Note que la descripción emergente no muestra la modalidad de formación a la que pertenece cada estudiante graduado por lo tanto se debe editar, agregando “Modalidad de formación: ” donde : y se encuentran separados con un Tab;

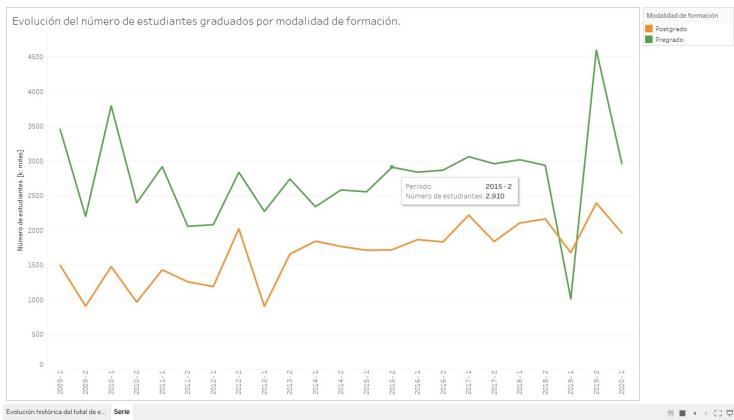


Figure 2.54: Gráfico de líneas segmentado

Modalidad de formación: debe ser escrito en el gris más oscuro disponible en la paleta y en negro.

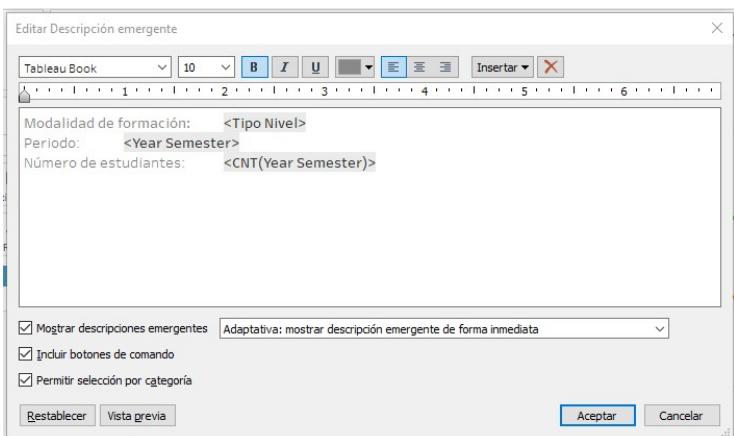


Figure 2.55: Editar descripción emergente

Finalmente se obtiene la visualización presentada en la figura 2.52.

Si se observa detalladamente la descripción emergente de la visualización presentada en la página de las estadísticas de la Universidad se identifica que contiene el porcentaje de graduados por modalidad para añadir esto a la descripción emergente se debe:

1. Crear un campo calculado, haciendo clic en el menú desplegable ubicado entre Datos y Tablas, seleccionar Crear campo calculado, asignar un nombre útil al cálculo por ejemplo Conteo de estudiantes graduados por periodo, en el panel en blanco escriba COUNT([Year Semester]) y clic en

Aceptar. Con esto se obtiene un nuevo campo que se ubica en la parte inferior del panel Tablas.

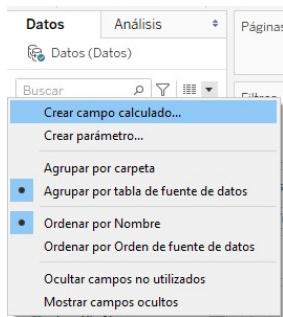


Figure 2.56: Crear campo calculado



Figure 2.57: Edición del campo calculado

2. Luego tener el campo calculado creado debe arrastrarlo a la tarjeta etiqueta ubicada en el estante Marcas.
3. En el menú desplegable del campo AGG(Conteo de estudiantes graduados por periodo), seleccione Añadir calculo de tabla.
4. En la ventana calculo de tablas en tipo de cálculo seleccione Porcentaje del total y en Calcular usando seleccione Tabla (abajo) y cierre la ventana.
5. Ahora la visualización muestra el porcentaje de estudiantes graduados por modalidad de formación, por ejemplo, para el periodo 2009-1 el 60.77% de los graduados son fueron de pregrado y el restante 30.33% de postgrado.
6. Para que estas etiquetas no se muestren sobre la línea si no en la tarjeta de descripción emergente Luego tener el campo calculado creado haga clic sobre descripción emergente y agregue esta línea de texto “Porcentaje del total: <% de total AGG(Conteo de estudiantes graduados por periodo)>”,

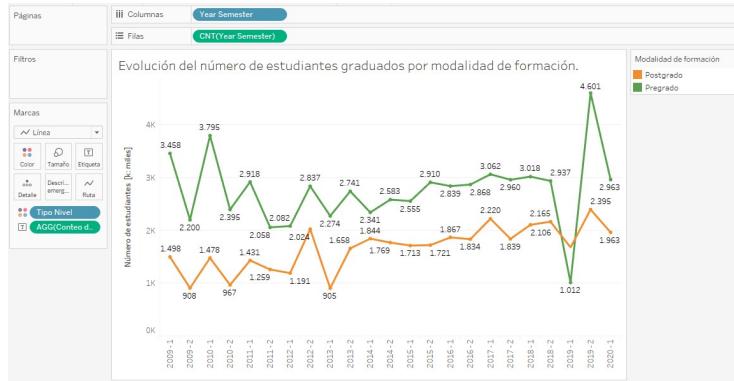


Figure 2.58: Etiquetas del total de estudiantes graduados

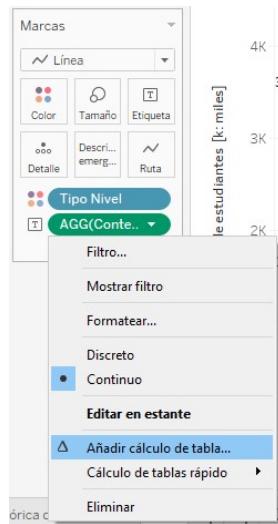


Figure 2.59: Añadir cálculo de tabla

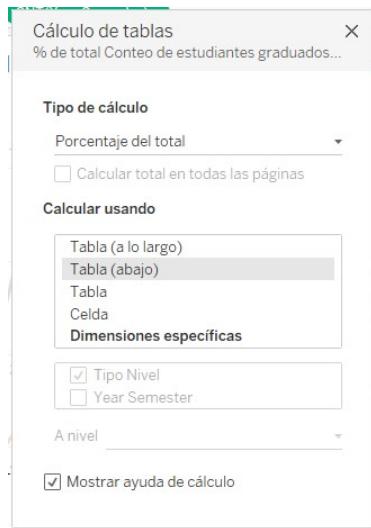


Figure 2.60: Seleccionar cálculo y alcance

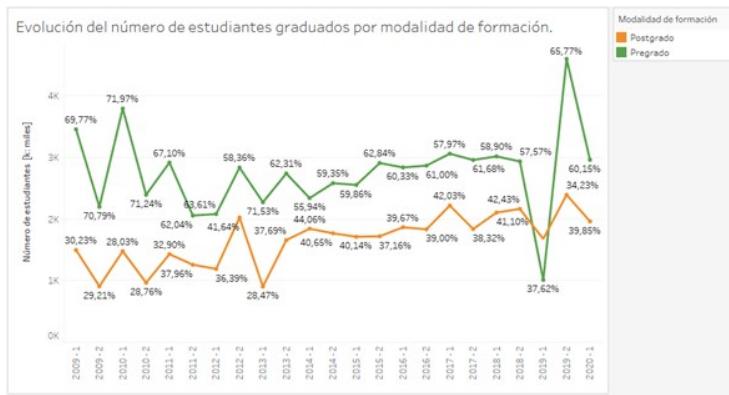


Figure 2.61: Visualización con etiquetas de porcentaje

edite el color y el tamaño para que coincida con lo demás que contiene esta tarjeta y clic en Aceptar.

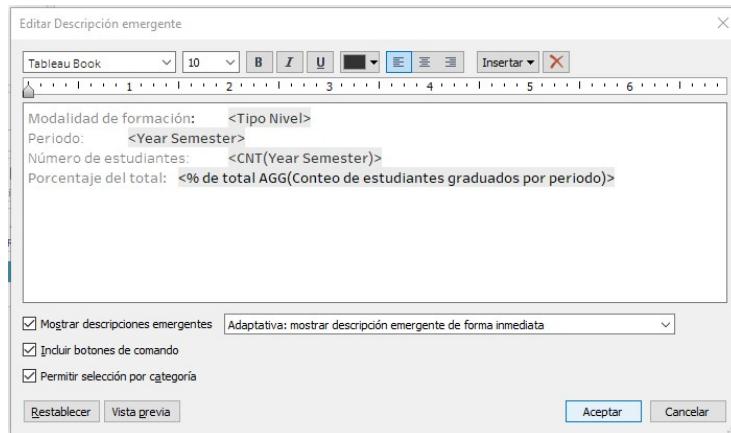


Figure 2.62: Añadir cálculo de tabla a la descripción emergente

- Haga clic en la tarjeta Etiqueta del estante marcas y desactive la opción Mostrar etiquetas de marca; finalmente la visualización se ve como se deseaba.

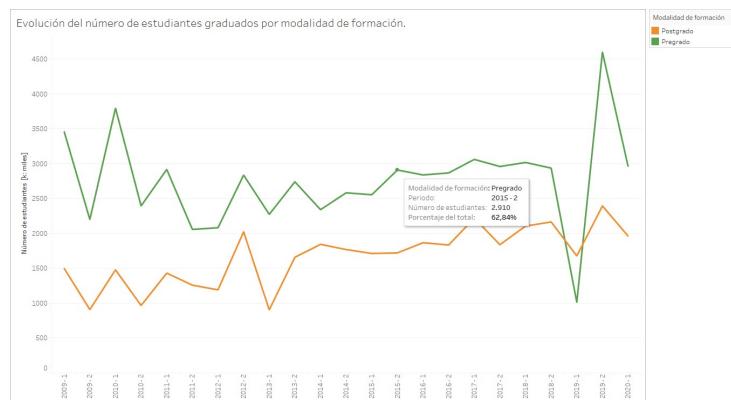


Figure 2.63: Visualizacion segmentada y con porcentajes

#### 2.4.2.3 Tablas de texto

Una forma útil y clara de mostrar los datos es usar tablas de texto, en este caso se presentara la forma de hacer una tabla de texto que contenga la información del número de estudiantes graduados por modalidad de formación.

- Agregue una nueva hoja de trabajo y cambie el nombre por Tabla.

2. Arrastre el campo Tipo Nivel al estante Columnas, los campos Year y Semestre cámbielos a cadena haciendo clic en el icono # y luego seleccione Cadena, arrastre estos campos modificados al estante filas, primero Year y luego Semestre.

Tabla		
		Tipo Nivel
Year	Semestre	Postgrado Pregrado
2009	1	Abc Abc
	2	Abc Abc
2010	1	Abc Abc
	2	Abc Abc
2011	1	Abc Abc
	2	Abc Abc
2012	1	Abc Abc
	2	Abc Abc
2013	1	Abc Abc
	2	Abc Abc
2014	1	Abc Abc
	2	Abc Abc
2015	1	Abc Abc
	2	Abc Abc
2016	1	Abc Abc
	2	Abc Abc
2017	1	Abc Abc
	2	Abc Abc
2018	1	Abc Abc
	2	Abc Abc
2019	1	Abc Abc

Figure 2.64: Arrastrar campos para crear tablas de texto

3. Ajuste el tamaño de la vista en la barra de herramientas seleccionando Vista completa.
4. Arrastre el campo Recuento por periodo a la tarjeta Texto en el estante marcas.

Filtros			Tabla		
Marcas			Tipo Nivel		
Year	Semestre		Postgrado	Pregrado	
2009	1		1.498	3.458	
	2		3.908	2.200	
2010	1		1.476	3.795	
	2		3.677	2.396	
2011	1		1.431	2.918	
	2		1.259	2.058	
2012	1		1.191	2.082	
	2		2.024	2.837	
2013	1		1.950	2.774	
	2		1.659	2.741	
2014	1		1.844	2.341	
	2		1.769	2.583	
2015	1		1.713	2.555	
	2		1.721	2.910	
2016	1		1.697	2.537	
	2		1.824	2.848	
2017	1		2.220	3.062	
	2		1.839	2.960	
2018	1		2.106	3.018	
	2		2.165	2.937	
2019	1		1.676	1.012	
	2		2.355	4.601	
2020	1		1.963	2.963	

Figure 2.65: Añadir el texto a la tabla

5. Los nombres de los campos utilizados en la vista no tienen nombres adecuados, por lo que se deben cambiar, desde el panel lateral Tablas seleccionando el menú desplegable del campo y haciendo clic en cambiar nombre, por ejemplo, Year debe ser cambiado por Año y Tipo Nivel por Modalidad

de Formación.

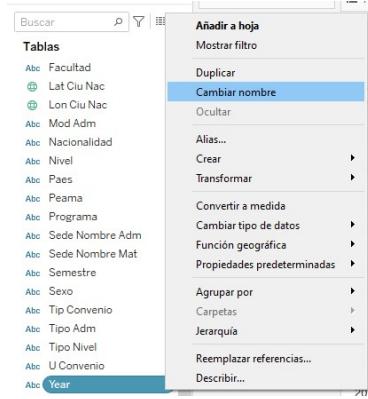


Figure 2.66: Editar nombres de los campos

6. El uso de filtros es útil para permitir que el usuario seleccione los años o semestres específicos que desea ver en la tabla, por lo cual se añadirá un filtro con el campo Año y otro con el campo Semestre. Arrastre el campo Año al estante filtro y en la ventana asegúrese de que todos los años estén seleccionados y luego haga clic en aceptar.

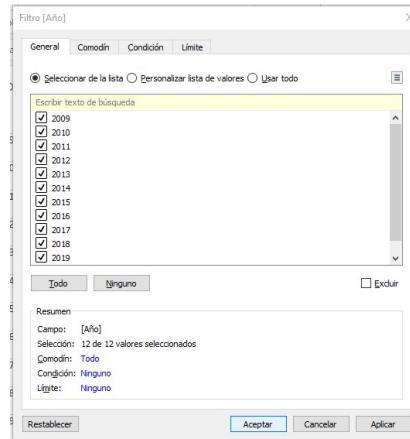


Figure 2.67: Añadir filtro

7. En el estante Filtros haga clic en el menú desplegable del campo Año y seleccione mostrar filtro, con esto aparece una tarjeta en el panel lateral derecho que contiene el filtro.

La tabla de texto se visualiza de esta manera.

8. En el panel lateral derecho esta el filtro que permite seleccionar los valores

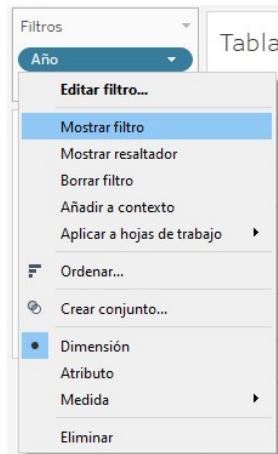


Figure 2.68: Mostrar el filtro

Tabla

Año	Semestre	Modalidad de formación	Año	
			Postgrado	Pregrado
2009	1		1.498	3.458
	2		908	2.300
2010	1		1.478	3.795
	2		967	2.395
2011	1		1.421	2.918
	2		1.259	2.058
2012	1		1.191	2.082
	2		2.024	2.837
2013	1		905	2.274
	2		1.658	2.741
2014	1		1.844	2.341
	2		1.769	2.583
2015	1		1.713	2.555
	2		1.721	2.910
2016	1		1.867	2.839
	2		1.834	2.868
2017	1		2.220	3.062
	2		1.839	2.960
2018	1		2.106	3.018
	2		2.165	2.937
2019	1		1.678	1.012
	2		2.395	4.601
2020	1		1.963	2.963

Figure 2.69: Vista previa de la tabla de texto

de años que el usuario desea ver, repita el paso 6 y 7 para crear un filtro con el campo semestre.

- Cambie el título de la visualización por Evolución del número de estudiantes graduados por modalidad de formación.

Finalmente se muestra la visualización obtenida que permite al usuario seleccionar diferentes valores para el campo año y semestre.

Evolución del número de estudiantes graduados por modalidad de formación			
Año	Semestre	Modalidad de formación	Semestre
		Postgrado	(Todos)
2009	1	1.498	3.458
	2	908	2.200
2010	1	1.478	3.795
	2	967	2.395
2011	1	1.431	2.918
	2	1.359	2.020
2012	1	1.191	2.058
	2	2.024	2.014
2013	1	905	2.061
	2	1.658	2.035
2014	1	1.844	2.274
	2	1.769	2.019
2015	1	1.713	2.341
	2	1.721	2.020
2016	1	1.867	2.883
	2	1.834	2.555
2017	1	2.220	2.910
	2	1.839	2.839
2018	1	2.106	3.062
	2	2.165	2.950
2019	1	1.678	3.918
	2	2.395	2.937
2020	1	1.963	1.012
			4.601
			4.963

Figure 2.70: Tabla de texto con filtros

#### 2.4.2.4 Gráfico circular

Los gráficos circulares son un recurso estadístico muy utilizado para representar porcentajes y proporciones, en este caso se hará un grafico circular que permita ver la distribución del total de estudiantes graduados por modalidad de formación para el periodo actual es decir 2020-1.

- Cree una nueva hoja de trabajo y llámela Distribución de graduados por modalidad de formación, periodo 2020-1.
- Arrastre el campo modalidad de formación a la tarjeta Color ubicada en el estante marcas.
- Cambie la forma de Automático a circular en el menú desplegable de este mismo estante.
- Arrastre el campo calculado llamado Recuento por modalidad de formación a la tarjeta ángulo.
- Ajuste el tamaño de la vista seleccionando Vista completa en la barra de herramientas, la visualización debe verse de esta manera.
- Se deben añadir etiquetas que indiquen la modalidad de formación correspondiente a cada color, para esto arrastre el campo Modalidad de formación a la tarjeta etiquetas.



Figure 2.71: Cambiar marca por circular



Figure 2.72: Vista previa grafico circular

7. Es importante visualizar los porcentajes que corresponden a cada modalidad, para esto se creara un nuevo campo calculado que contenga el conteo de los graduados por modalidad, como se mostro en 2.4.2.2 y llámelo Recuento por modalidad de formación.
8. Arrastre el campo calculado creado a la tarjeta etiqueta, se obtiene una etiqueta en el gráfico que muestra la cantidad de estudiantes graduados para todos los periodos por el nivel de formación.

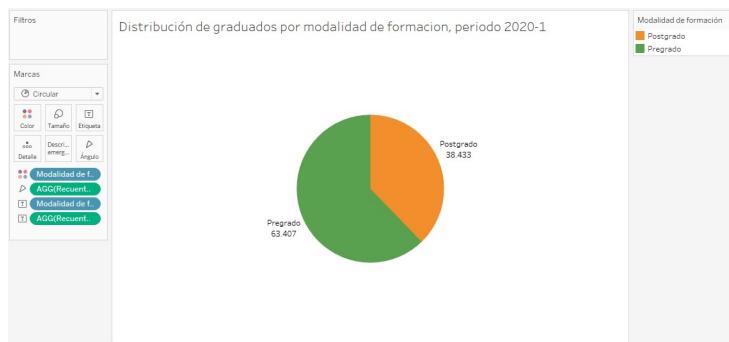


Figure 2.73: Añadir etiquetas

9. Se desea que esta visualización solo muestre los datos del periodo 2020-1, para esto cree un filtro con el campo Year semestre y seleccione únicamente el periodo 2020-1, no es necesario que muestre el filtro.
10. En realidad, no interesa que la visualización muestre el número de estudiantes graduados por nivel si no el porcentaje de esto, para esto se debe añadir un calculo de tabla en la etiqueta del campo calculado, haga clic en el menú desplegable del ultimo campo del estante Marcas y seleccione Añadir calculo de tabla.
11. En la venta de cálculos de tablas, en Tipo de cálculo seleccione Porcentaje del total y calcular usando seleccione Tabla (a lo largo).
12. Se debe editar la decepción emergente para que sea mas clara, cambie % de total Recuento por modalidad de formación junto con Tabla (a lo largo) por Porcentaje y Recuento por periodo por Número de graduados. Finalmente, se obtiene un gráfico circular que muestra el porcentaje de graduados por modalidad de formación y con una descripción emergente clara.

#### 2.4.2.5 Gráfico de barras

Estos gráficos son útiles para representar las frecuencias de las clases de alguna variable de interés, en este caso se realizará un gráfico de barras horizontales que muestre la frecuencia y el porcentaje de estudiantes graduados por nivel de formación para el periodo 2020-1.

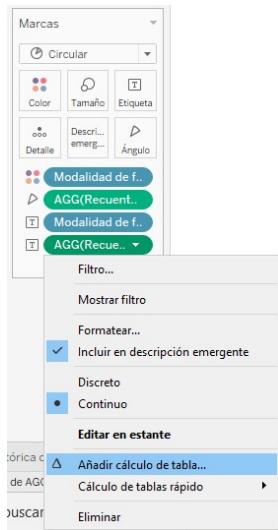


Figure 2.74: Añadir cálculo de tabla a las etiquetas

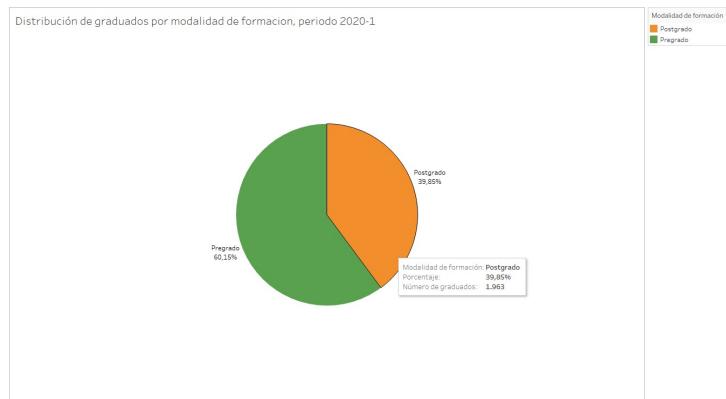


Figure 2.75: Gráfico circular para el periodo 2020-1

1. Cree una nueva hoja de trabajo y llámela Distribución de estudiantes graduados por nivel de formación, periodo 2020-1.
2. Arrastre el campo Nivel a filas y arrástrelo nuevamente desde el panel Tablas al estante columnas.
3. En el menú desplegable del campo Nivel ubicado en el estante columnas seleccione medida y luego recuento como se mostro en el paso 2 de 2.4.2.1. Debe obtener un grafico como este,



Figure 2.76: Gráfico de barras

4. Ajuste el tamaño de la vista seleccionando vista completa en la barra de herramientas.
5. Arrastre el campo Nivel a la tarjeta color ubicada en el estante marcas.
6. Edite los colores de cada barra como se mostro en el paso 5 de 2.4.2.2. Asigne el color verde a Pregrado, rojo a maestría, azul a especialización, amarillo a especialidades medicas y gris a doctorado.
7. Luego haga clic en el botón que señala orden descendente ubicado en la barra de herramientas. Hasta el momento su visualización debe verse así,
8. Observe que el recuento de nivel se esta haciendo para todos los periodos que contiene la base es decir desde 2009-1 hasta el 2020-1, para que solo muestre los del periodo actual debe añadir un filtro con el campo Year Semester y seleccionar únicamente el periodo de interés, no es necesario que muestre el filtro.
9. Edite el eje x cambiando el título, escriba como título Número de graduados; también edite la descripción emergente cambiando Recuento de nivel por Número de graduados.

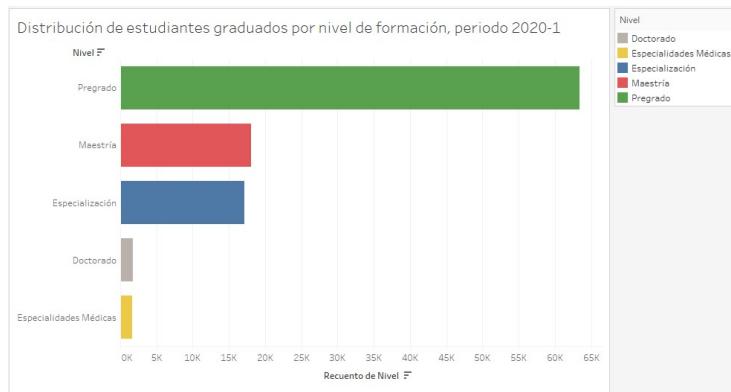


Figure 2.77: Gráfico de barras con colores asignados

- Como etiqueta de cada barra se debería mostrar el total de estudiantes graduados en ese nivel, para esto arrastre el campo Nivel a la tarjeta etiqueta, en el menú desplegable del campo seleccione Medida y luego recuento.

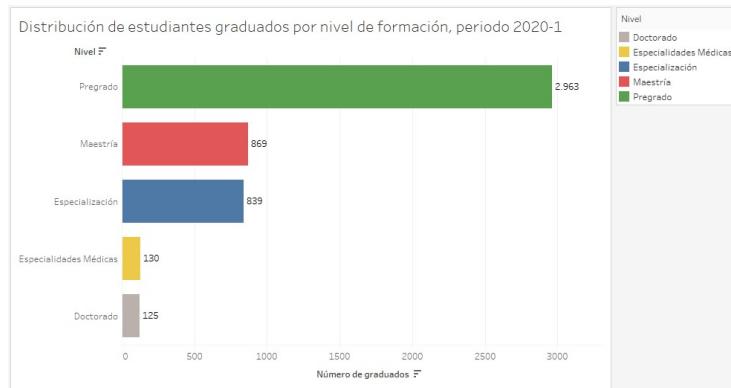


Figure 2.78: Gráfico de barras con etiquetas

- Es importante que en la descripción emergente se muestre el porcentaje de cada nivel de formación, para esto arrastre el campo Nivel a la tarjeta descripción emergente en el menú desplegable seleccione media y recuento.
- Haga clic en este mismo menú después de haber realizado en el paso anterior y seleccione añadir cálculo de tabla, en la ventana de cálculos de tabla en Tipo de cálculo seleccione Porcentaje del total y en calcular usando seleccione Tabla (abajo), en este momento la descripción emergente se ve así,
- Cambie Recuento de nivel por Número de graduados y % de total Recuento

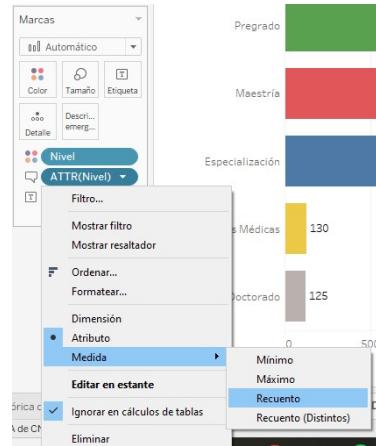


Figure 2.79: Añadir porcentaje a descripción emergente

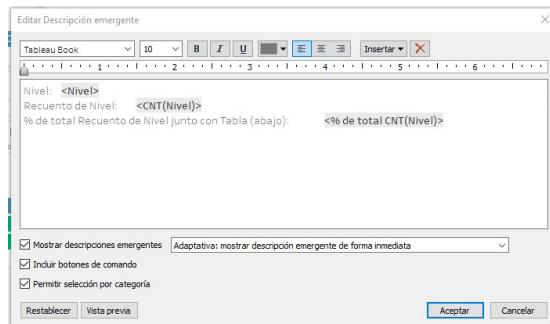


Figure 2.80: Edición de la descripción emergente

de nivel junto con Tabla (abajo) por Porcentaje y de clic en aceptar.

El gráfico de barras horizontales para el nivel de formación para el periodo 2020-1 se muestra a continuación.

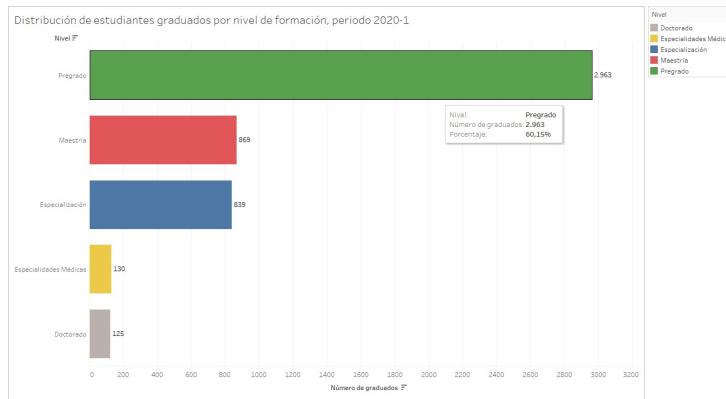


Figure 2.81: Gráfico de barras por nivel de formación, periodo 2020-1

#### 2.4.2.6 Mapas de árbol

Estos mapas usan rectángulos anidados para mostrar datos jerárquicos como parte de un todo. La forma cuadrada permite comparar más fácilmente tamaños relativos, para crear un mapa de árbol es necesario una medida y una dimensión. En este caso se hará un mapa de árbol que permita identificar cual sede de la Universidad Nacional tuvo más graduados a lo largo de los periodos que se tienen en la base de datos.

1. Cree una nueva hoja de trabajo y llámela Distribución de graduados por sede de admisión.
2. Arrastre el campo Sede Nombre Adm a la tarjeta color y nuevamente desde el panel tablas arrastre este mismo campo a la tarjeta tamaño.
3. En el menú desplegable de los campos ubicados en el estante marcas, seleccione medida y recuento, para ambas variables.
4. Arrastre nuevamente Sede Nombre Adm a la tarjeta Detalles.
5. El cuadrado mas grande y en el azul más fuerte identifica a la sede con mas graduados, que en este caso es la sede Bogotá, esta información se ubica en la descripción emergente, pero es necesario ubicarla como etiquetas de los cuadrados para que la visualización sea mas clara, para esto arrastre Sede Nombre Adm a la tarjeta etiqueta, presione la tecla ctrl y clic sobre alguno de los dos campos en verde ubicados en marcas y arrástrelo hasta etiqueta. Ahora su visualización muestra la sede de admisión y el total de estudiantes graduados.

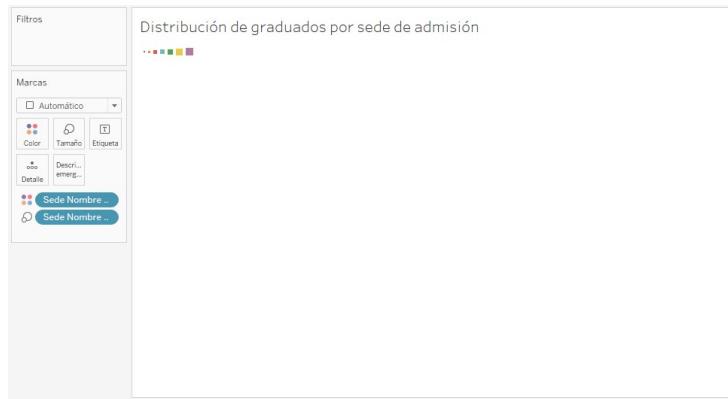


Figure 2.82: Añadir los campos a la tarjeta marcas

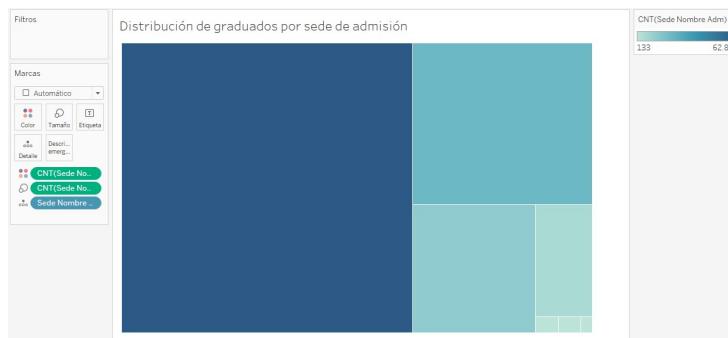


Figure 2.83: Mapa de árbol base



Figure 2.84: Añadir etiquetas al mapa de árbol

6. Los mapas de árbol son útiles para mostrar porcentajes del total, para esto repita los pasos 10 y 11 de 2.4.2.4.
7. Edite la descripción emergente para que el nombre de los campos sea más claro; cambie Sede Nombre Adm por Sede de admisión, Recuento de Sede Nombre Adm por Número de estudiantes graduados y % de total Recuento de Sede Nombre Adm junto con Tabla (a lo largo) por porcentaje.
8. Edite el nombre de la tarjeta ubicada en el lateral derecho haciendo clic en el menú desplegable y seleccione editar título, el nuevo título es Número de estudiantes graduados.

El mapa de árbol obtenido es claro y consistente con la información que se desea visualizar:



Figure 2.85: Mapa de árbol por sede de admisión

Una variación de los mapas de árbol que también permite mostrar la jerarquía que existe en los datos son el gráfico de burbujas y nube de palabras.

Para realizar un grafico de burbujas duplique la hoja distribución de graduados por sede de admisión y llámela Gráfico de burbujas para la distribución de graduados por sede de admisión, luego haga clic en el botón Mostrarme y seleccione burbujas agrupadas.

Note que la descripción emergente no está clara por lo que es necesario cambie Sede Nombre Adm por Sede de admisión, Recuento de Sede Nombre Adm por Número de estudiantes graduados y % de total Recuento de Sede Nombre Adm junto con Sede Nombre Adm por porcentaje; duplique el campo correspondiente a color y llévelo a etiqueta, ajuste el tamaño de la visualización a vista completa, edite el titulo de la leyenda de la clave de color, asigne como nuevo título porcentaje. Finalmente, su visualización al igual que el mapa de árbol permite identificar cual sede de la Universidad Nacional tuvo más graduados a lo largo de los periodos registrados en el conjunto de datos.

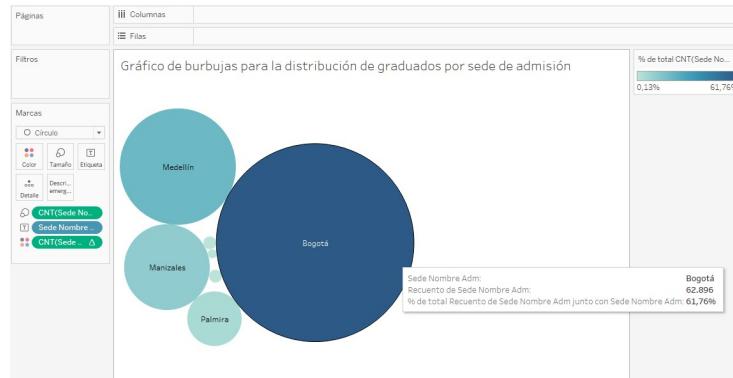


Figure 2.86: Gráfico de burbujas base por sede de admisión

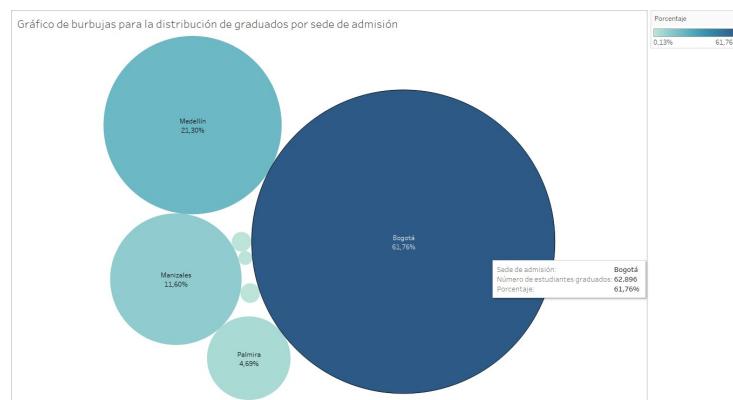


Figure 2.87: Gráfico de burbujas por sede de admisión

La otra forma de visualizar la distribución de los estudiantes graduados es usando una nube de palabras, esta muestra la frecuencia de uso. Nuevamente duplique la hoja en la que se creó el mapa de árbol y llámela Nube de palabras para la distribución de graduados por sede de admisión, en el estante marcas cambie automático por texto.



Figure 2.88: Gráfico base de nube de palabras por sede de admisión

En este tipo de gráficos no es conveniente mostrar el porcentaje, ya que el enfoque se debe dar al nombre de la sede, arrastre el campo CNT(Sede Nombre Adm) de la tarjeta etiqueta a la tarjeta detalles.



Figure 2.89: Nube de palabras por sede de admisión, sin porcentaje

Note que de la descripción emergente se eliminó el campo que mostraba el porcentaje, para recuperarlo en la descripción emergente al lado de Porcentaje escriba <% de total CNT(Sede Nombre Adm)>, que es la expresión que usa Tableau para calcular el Porcentaje.



Figure 2.90: Nube de palabras por sede de admisión

#### 2.4.2.7 Gráfico de control

Este tipo de gráficos se usan para evaluar cómo cambia un proceso en el tiempo. Se parecen mucho a los gráficos de líneas sencillos, pero se debe agregar una línea de promedio, un límite de control inferior y uno superior. Estos gráficos son una herramienta estadística de control de procesos, para determinar si un proceso de fabricación comercial está dentro o fuera de los límites. En este caso se considera importante visualizar si los graduados a través del tiempo siguen un proceso controlado, cabe resaltar que este proceso se puede ver afectado por muchos factores tales como la deserción, paros estudiantiles, la sede, considerando únicamente la sede de matrícula; la facultad y el programa a que pertenecen; se hará un gráfico general y luego un desglose de esta información mencionada para tener más claridad sobre el comportamiento de estas cifras.

1. Duplique la hoja que contiene el gráfico de líneas mostrado en 2.4.2.1.
2. Cambie el nombre de la hoja duplicada por Gráfico de control general.
3. Arrastre el campo calculado llamado Recuento por periodo al lado derecho del campo ubicado en el estante filas.
4. En el menú desplegable del campo AGG(Recuento por periodo) seleccione Eje doble.
5. En el estante marcas cambie línea por círculos para el campo AGG(Recuento por periodo).
6. Oculte la tarjeta ubicada en el lateral derecho.
7. Asigne el mismo color tanto a las líneas como a los puntos, en este caso ambas se dejarán en color azul.



Figure 2.91: Añadir campo calculado al estante filas

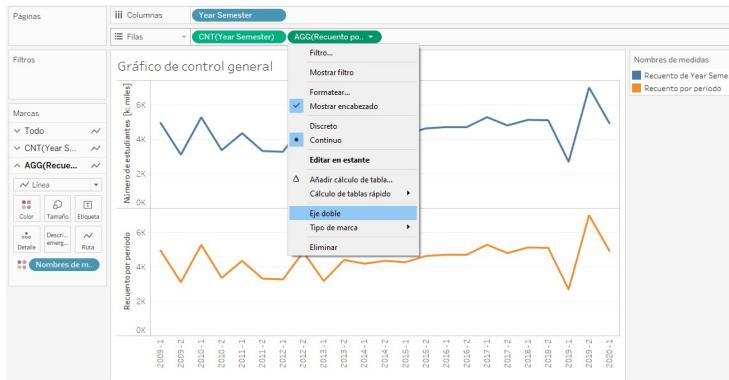


Figure 2.92: Asignar eje doble



Figure 2.93: Asignar marca circular

8. Haga clic en el panel Análisis ubicado en el lateral izquierdo, en personalizado arrastre línea de referencia a la visualización y suéltela sobre panel.

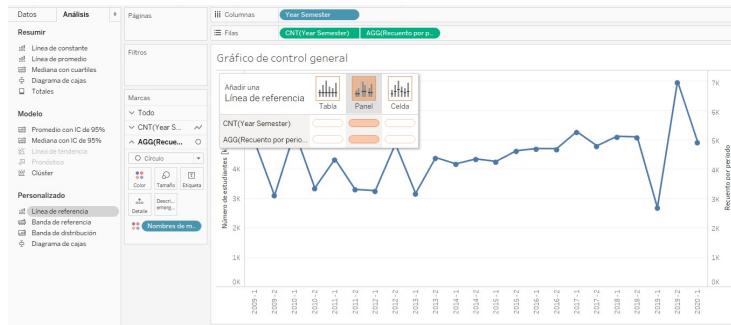


Figure 2.94: Añadir línea de referencia

9. En la ventana emergente asegúrese que en el campo valor este CNT(Year Semester) y promedio, puede agregar intervalos de confianza y dar formato a la línea, en este caso no es necesario ya que solo se debe agregar la línea de promedio.

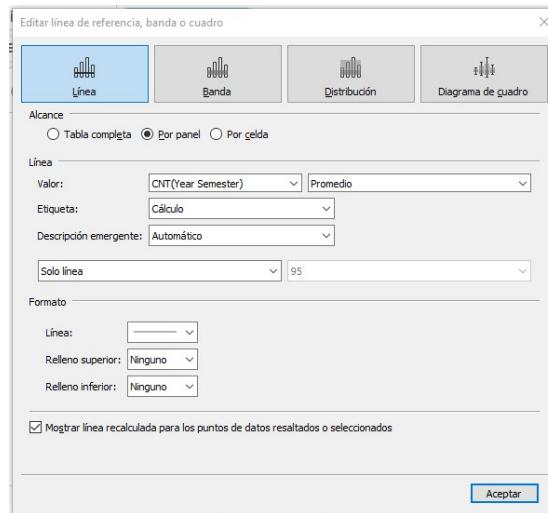


Figure 2.95: Editar línea de referencia

10. Es momento de implementar un parámetro que permita al usuario controlar cuantas desviaciones estándar por debajo y por encima del promedio se ubicaran los límites. En el menú del panel tablas seleccione crear parámetro, en la ventana emergente asigne el nombre de Desviaciones estándar al parámetro, tipo de dato cámbielo por entero, en valores permitidos seleccione lista y añada los valores de 1, 2 y 3, finalmente haga

clic en aceptar.

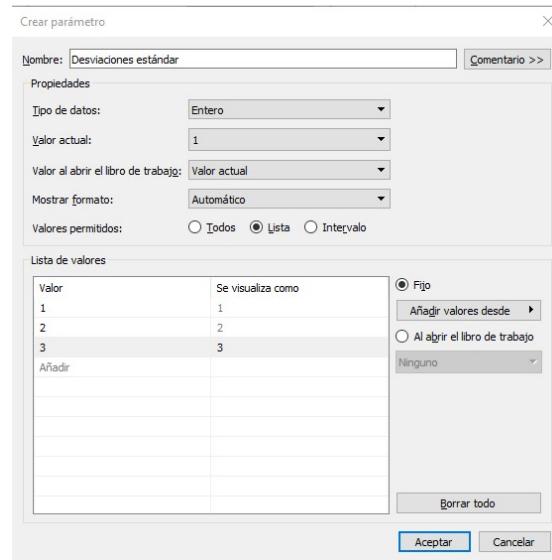


Figure 2.96: Crear parámetro para desviaciones estándar

11. Los límites inferior y superior serán campos calculados, nuevamente el menú del panel tablas seleccione crear campo calculado asigne el nombre de Límite superior y escriba  **$\text{AVG}(\text{COUNT}([\text{Year Semester}])) + \text{WINDOW\_STDEV}(\text{COUNT}([\text{Year Semester}])) * [\text{Desviaciones estándar}]$**



Figure 2.97: Límite superior

12. Para el límite inferior repita el paso anterior, pero en lugar de sumar la desviación estándar la debe restar.
13. Agregue los límites de control a la tarjeta detalle ubicada en el estante marcas.

14. Nuevamente desde el panel Análisis en personalizado arrastre banda de referencia hacia la visualización y suéltela en panel.
15. En la ventana emergente, en la sección banda desde para valor asigne límite inferior y mínimo, en la etiqueta seleccione ninguno; para la sección banda hasta para valor asigne límite superior y máximo, en la etiqueta seleccione ninguno, para el formato de la imagen seleccione línea punteada en color gris oscuro y un relleno gris más claro.

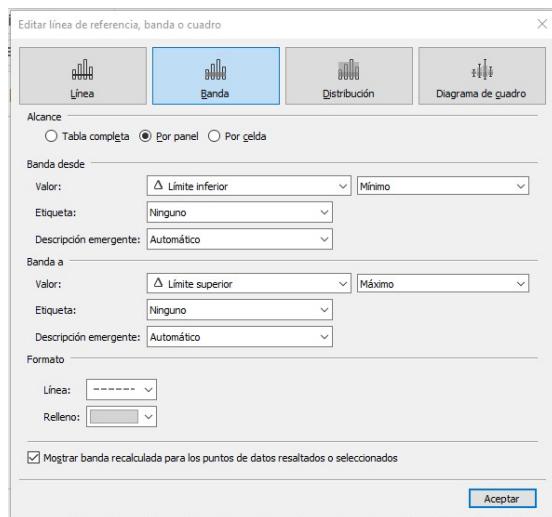


Figure 2.98: Editar banda de referencia

16. Para tener aun mas claridad en la vista e creara un campo calculado que identifique si el punto esta fuera o dentro de los límites de control. Cree un nuevo campo calculado y asigne el nombre de “¿Está dentro del límite de control?” y en el espacio en blanco escriba

**IF COUNT([Year Semester]) > [Límite superior]**

**or COUNT([Year Semester]) < [Límite inferior]**

**THEN “No” ELSE “Si”**

**END**

17. Asignar este campo calculado a la tarjeta colores del conteo que tiene como forma círculos, de el color rojo a la categoría No y verde a Si.
18. Muestre el control del parámetro creado en el paso 10, haciendo clic derecho sobre el parámetro que se encuentra ubicado en la parte inferior del panel tablas y seleccionando Mostrar parámetro.

Finalmente, la visualización obtenida permite identificar en que periodos se observa desviaciones fuertes en el número de graduados y el parámetro permite



Figure 2.99: Gráfico de control con una desviación estándar

al usuario controlar cuantas desviaciones estándar se aginan a los límites.

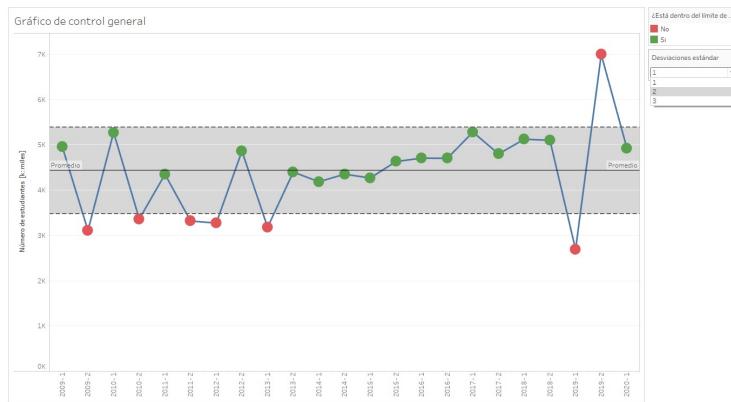


Figure 2.100: Gráfico de control con interacción de cantidad de desviaciones estándar

Cabe aclarar que lo observado en este gráfico de control tiene una información de trasfondo que puede hacer que las conclusiones obtenidas no sean completamente ciertas.

Ahora se realizará este mismo gráfico de control, pero discriminado por Sede de Matricula es decir la sede en la cual matricularon el último semestre del programa.

1. Duplique la hoja de trabajo que contiene el gráfico de control general.
2. Cambie el nombre de la hoja duplicada por Gráfico de control por sede de matrícula.
3. Arrastre el campo Sede Nombre Mat al estante filas, ubíquelo al lado

izquierdo de las dos medidas que ya se tienen en este estante.



Figure 2.101: Añadir campo Sede de matrícula a la visualización

4. Es necesario ajustar el tamaño a Vista completa y reducir la magnitud de los puntos, esto último lo puede hacer desde la tarjeta tamaño ubicada en el estante marcas.
5. Oculte el encabezado del eje Y, también oculte el encabezado ubicado en la parte superior del nombre de la sede de matrícula.

Se obtiene una visualización que permite ver el comportamiento de los graduados a través de un periodo de tiempo discriminado por la sede de matrícula.



Figure 2.102: Gráfico de control por sede de matrícula

Observe que para sede Caribe hay periodos en los cuales no se graduaron estudiantes, esto puede ser una causa de los puntos fuera del límite de control en la figura 2.100.

También es de vital importancia visualizar este comportamiento a nivel de facultades para cada sede, ya que no todas las facultades poseen el mismo número de graduados y por ende puede causar que el número de estudiantes graduados de la sede a al que pertenece esa facultad este por encima o por debajo de los límites de control.

1. Duplique esta última visualización creada y llame a esta hoja duplicada Gráfico de control por facultad.
2. Retire el campo Sede Nombre Mat del estante filas y llévelo al estante filtros, seleccione todas las sedes y haga clic en aceptar.
3. En el menú desplegable del filtro seleccione mostrar filtro, en este momento su visualización debe verse de esta manera.



Figure 2.103: Filtrar por sede de matrícula

4. Como solo se quiere mostrar una sede a la vez es necesario editar la lista desplegable del filtro que se ubica en el panel lateral derecho, haga clic en el menú desplegable y seleccione Valor Individual (lista).

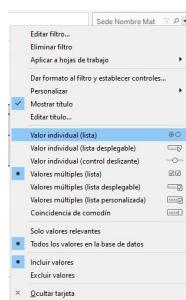


Figure 2.104: Edición del control de filtro

5. Edite el título del filtro, ponga como nuevo título Sede de matrícula.

6. Arrastre la variable Facultad al estante filas, ubíquelo al lado izquierdo de las dos medidas que ya se tienen en este estante.

Por ejemplo, si selecciona sede Palmira y 2 desviaciones estándar, su gráfico de control se ve de la siguiente manera.

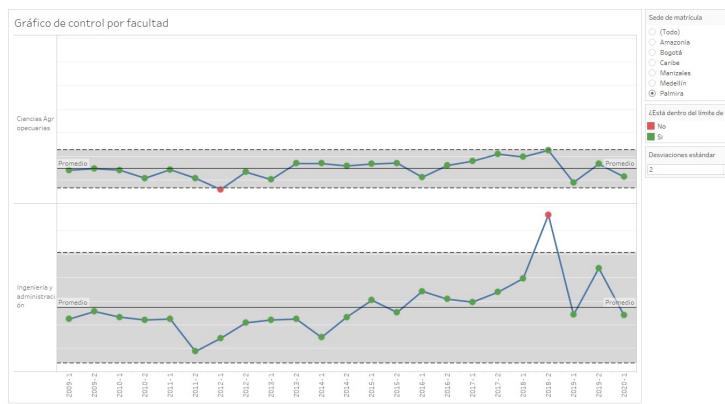


Figure 2.105: Gráfico de control por facultad

Es claro que la cantidad de graduados de la facultad de Ingeniería y administración tiene una tendencia de asenso y su promedio es superior a la facultad de Ciencias Agropecuarias. La facultad de ciencias agropecuarias no presenta una alta variabilidad en el número de graduados por periodo en la ventana de tiempo analizada.

#### 2.4.2.8 Gráfico de pendientes

Estos gráficos muestran los cambios en la clasificación o la posición de una dimensión desde un punto de inicio hasta un punto final, es útil para mostrar si una dimensión específica aumento o disminuyo entre dos puntos en el tiempo. Para los datos que se están analizando este tipo de gráficos pueden ser utilizados para mostrar la cantidad de estudiantes graduados por año de cada programa de las distintas sedes de matrícula, teniendo como año de inicio el 2009 y año final el 2019, cabe aclarar que el conjunto de datos posee registros para el año 2020 pero solo para el 1 semestre, por tanto, la comparación no será equitativa ya que no se tienen registros para los dos semestres del año como en el caso de 2009. Para crear un grafico de pendientes siga estos pasos,

1. Cree una nueva hoja de trabajo y llámela Gráfico de pendientes para la sede Amazonía.
2. Arrastre el campo Año al campo columnas y nuevamente desde el panel tablas arrástrelo al estante filas, allí seleccione medida y recuento.
3. Cambie la marca por línea, ajuste la visualización a vista completa.

4. Añada la variable año al estante filas y seleccione únicamente los años 2009 y 2019 ya que son los años que se quieren visualizar en el gráfico de pendiente. Hasta el momento su visualización se ve de esta manera,



Figure 2.106: Filtrar la visualización por los años deseados

5. Se quiere visualizar la sede Amazonía, por tanto, añada el campo Sede Nombre Mat a filtros y seleccione Amazonía.
6. Añada programa a la tarjeta color del estante marcas.
7. En la advertencia seleccione añadir todos los miembros.

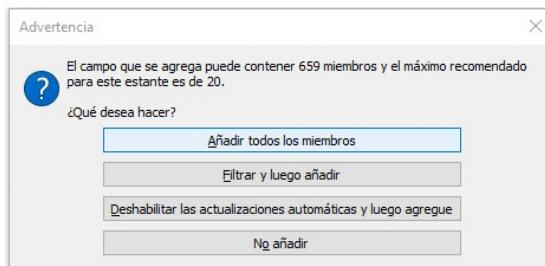


Figure 2.107: Añadir todos los miembros

8. Quite la selección de mostrar encabezado en eje y, haciendo clic derecho sobre este eje.
9. Nuevamente arrastre el campo Año a la tarjeta etiqueta ubicada en el estante marcas, seleccione medida y recuento en el menú desplegable del campo.
10. Edite la descripción emergente cambiando Recuento de Año por Número de graduados.

La visualización obtenida es muy simple, ya que para esta sede no se tienen programas específicos.

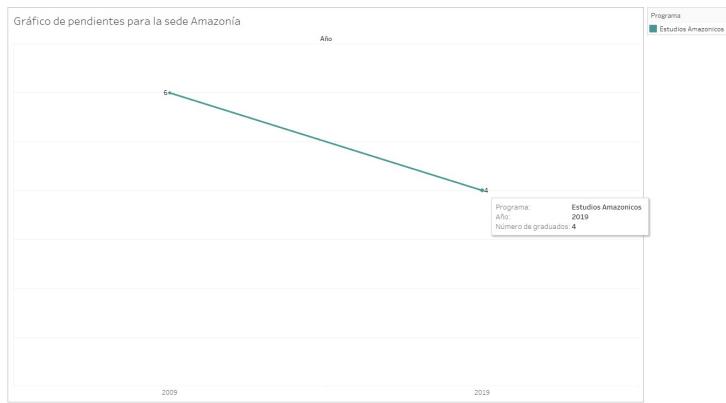


Figure 2.108: Gráfico de pendiente para la sede Amazonía

Para realizar la visualización para las demás sedes únicamente debe duplicar la vista anterior y cambiar el nombre de la hoja por la sede a la que pertenece, en el campo Sede Nombre Mat ubicado en la tarjeta filtros, haga clic derecho y seleccione editar filtro, en la ventana emergente seleccione Bogotá.

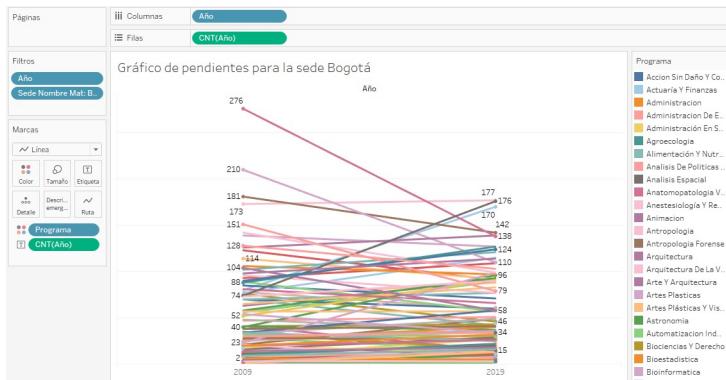


Figure 2.109: Filtrar por sede Bogotá

La gráfica en este momento es poco entendible ya que la para la sede Bogotá hay muchos programas y se están mezclando las dos modalidades de formación por lo que es útil agregar un filtro que permita seleccionar la modalidad que se quiere ver. Añada el campo modalidad de formación al estante filtros, seleccione ambas modalidades y haga clic en aceptar; muestre el filtro y seleccione valor individual (lista), si selecciona pregrado en modalidad de formación su visualización se ve así.

Note que aún hay saturación en la gráfica, el ultimo desglose que se puede hacer es filtrar la visualización por las facultades. Añada el campo Facultad al

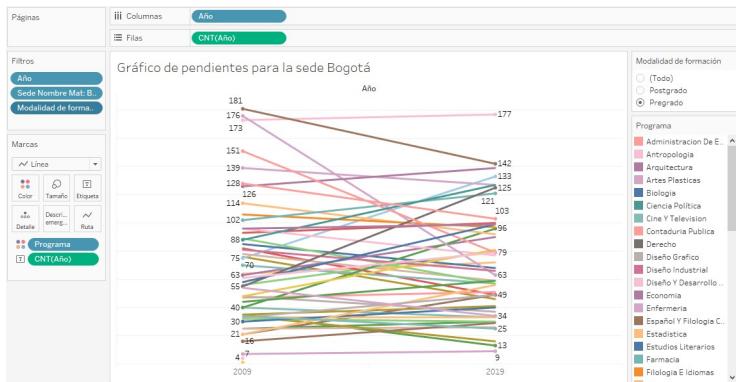


Figure 2.110: Añadir filtro de modalidad de formación

estante filtros, seleccione todas las facultades, muestre el filtro y finalmente seleccione valor individual (lista), a modo de ejemplo seleccione la facultad de ciencias, su visualización se verá así.

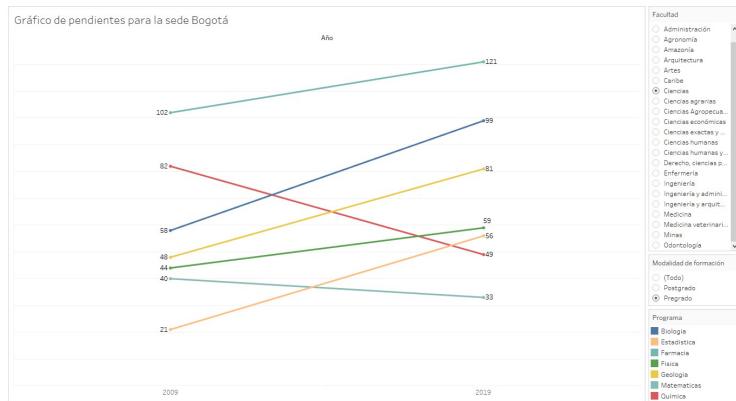


Figure 2.111: Gráfico de pendiente para la sede Bogotá

Esta visualización permite identificar que los programas de estadística, física, geología, biología y farmacia han aumentado en la cantidad de estudiantes graduados, mientras que los demás programas de los que se tienen registro que ofrece la facultad de ciencias han disminuido en el número de estudiantes graduados, cabe aclarar que en esta visualización no se está teniendo en cuenta lo que sucede con estos programas en los años intermedios.

Observe que en el lateral derecho donde se muestra el filtro de facultad se están mostrando todas las facultades que posee la universidad, cuando en realidad solo interesa mostrar las 12 facultades que pertenecen a la sede Bogotá, esto sucede ya que Tableau de manera predeterminada calcula los

filtros de manera independiente, es decir, cada filtro accede a todas las filas con independencia de la existencia de otros filtros, para solucionar esto existen los filtros de contexto, el filtro que defina como contexto tendrá acceso a todas las filas del conjunto de datos mientras que los otros filtros dependerán del filtro de contexto, es decir que solo tienen acceso a las filas que cumplen el filtro de contexto.

En este caso añadir el campo Sede Nombre Mat como filtro de contexto será de gran utilidad ya que el filtro facultad dependerá de los registros que coincidan con la sede seleccionada.

Para crear el filtro de contexto haga clic en el menú desplegable del campo Sede Nombre Mat y seleccione añadir a contexto.

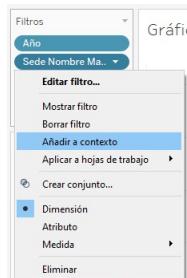


Figure 2.112: Añadir filtro de contexto

El campo sede se tornará de color gris y quedará al inicio de los demás filtros.

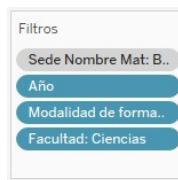


Figure 2.113: Filtro de contexto

Ahora debe ser editado el filtro facultad desde el menú desplegable de la leyenda del filtro, en el panel lateral derecho muestre el menú desplegable del filtro facultad y seleccione Todos los valores en contexto.

Ahora el gráfico de pendientes para la sede Bogotá solo contiene las facultades que pertenecen a dicha sede.

Para las demás sedes de la universidad puede duplicar la ultima hoja creada, cambiar su nombre por la sede a la que pertenece y en el filtro Sede Nombre Mat seleccionar la sede que quiere visualizar, al final obtendrá 6 gráficos de pendientes, uno para cada sede, estas hojas de trabajo podrán ser reunidas en una historia, este procedimiento se mostrara en la siguiente sección.

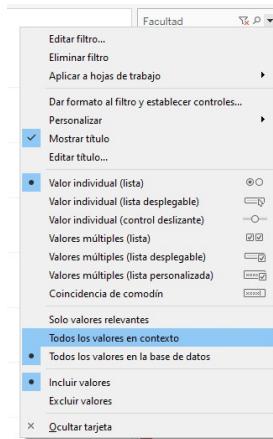


Figure 2.114: Edición filtro Facultad

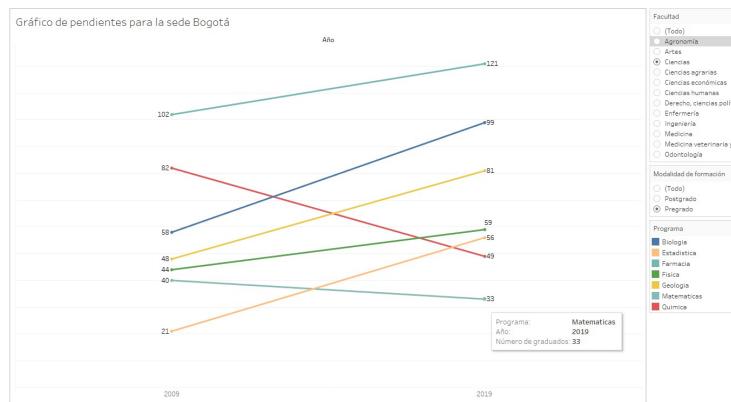


Figure 2.115: Gráfico de pendientes sede Bogotá con filtro de contexto

#### 2.4.2.9 Gráfico de bala

Este tipo de gráficos son una manera son una buena manera de mostrar la progresión en etapas hacia una meta, en este caso es de interés mostrar como avanza el número de graduados por Sede de matrícula hacia una meta establecida para cada una, estos valores objetivos se tomarán de manera arbitraria, para la sede Amazonía será 200, para Bogotá 65.000, Caribe tendrá un valor objetivo igual a 60, para Manizales 13.000, para la sede Medellín 23.000, finalmente el valor de meta para la sede Palmira será de 6.000.

La base de datos original no posee un campo que contenga estos valores de meta, por lo tanto es necesario la creación de un campo calculado que asigne el valor de meta a cada sede, para cree una nuevo campo calculado y como nombre asigne Meta de graduados, la expresión clave para asignar el valor de meta según la sede será IF y también se hará uso de una función que identifique si la Sede empieza por ciertas letras que se le indican, dentro del campo calculado escriba lo siguiente:

```
IF STARTSWITH([Sede Nombre Mat], "Ama") THEN "200"
ELSEIF STARTSWITH([Sede Nombre Mat], "Bo") THEN "65000"
ELSEIF STARTSWITH([Sede Nombre Mat], "Ca") THEN "60"
ELSEIF STARTSWITH([Sede Nombre Mat], "Ma") THEN "13000"
ELSEIF STARTSWITH([Sede Nombre Mat], "Me") THEN "23000"
ELSE "6000" END
```

Por ejemplo, en el primero renglón se le esta indicando que si el campo Sede Nombre Mat inicia con “Ama” asigne como meta 200, finalmente haga clic en Aceptar para guardar el campo.



Figure 2.116: Crear campo calculado con metas por sede

Este campo se guardo como cadena y en realidad es un número entero, cambie el tipo de dato por Número (entero), haciendo clic sobre el icono Abc de Meta de graduados y seleccione Número (entero), debe notar que el icono del campo

ahora es "#". Después de haber creado el campo calculado y cambiar el tipo de dato, es momento de iniciar con la creación del gráfico de bala.

1. Cree una nueva hoja de trabajo y cambie su nombre por Gráfico de bala.
2. Arrastre el campo Meta graduados a la tarjeta detalles del estante marcas, en el menú desplegable del campo seleccione Continuo.
3. Añada Year Semester a columnas y seleccione medida y recuento, el campo Sede Nombre Mat debe ser asignado a filas, ajuste el tamaño de la visualización seleccionando vista completa.

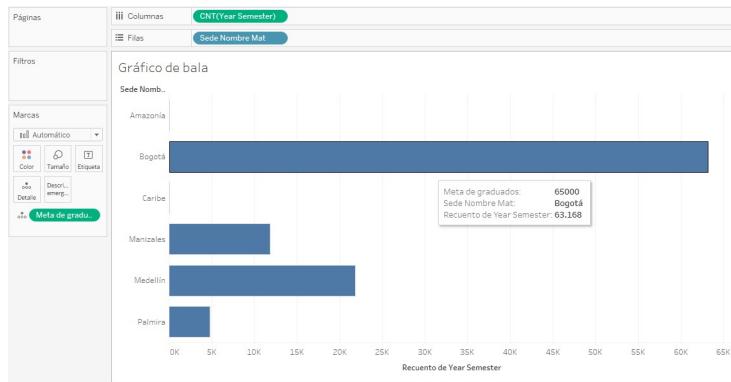


Figure 2.117: Paso 3 para la creación de un gráfico de barras

4. Edite la descripción emergente para que sea clara y concisa, también edite el eje x cambiando el título por Número de estudiantes graduados, oculte la etiqueta del eje y.
5. Es momento de añadir la línea de referencia que representa la meta por sede, para esto haga clic en el panel análisis en el lateral izquierdo, en personalizado arrastre línea de referencia hacia la vista y suéltela sobre celda, es necesario hacerlo sobre celda y no en las otras opciones que brinda ya que la meta de graduados es diferente para cada sede.
6. En la ventada de edición de la línea, para valor seleccione Meta de graduados y promedio, asegúrese que etiqueta este en ninguno, para la descripción emergente seleccione personalizado y en el espacio del lado derecho escriba Meta de graduados =, en botón del al lado con el símbolo > seleccione valor por último haga clic en aceptar para guardar los cambios.
7. Es momento de añadir una banda de referencia que muestre los porcentajes 50, 75 y 100, nuevamente desde el panel análisis y personalizado arrastre banda de referencia y suéltela en celda.
8. En la ventana seleccione Distribución, en el menú desplegable de valor seleccione porcentaje, en el espacio de porcentajes escriba 50; 75; 100, en porcentaje de seleccione Meda de graduados y promedio.

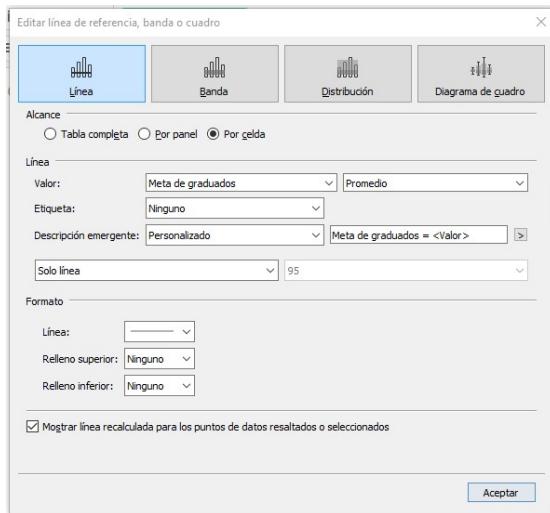


Figure 2.118: Edición de la línea de referencia para las metas

9. Para el campo etiqueta seleccione ninguno y para la descripción emergente seleccione automático.

La visualización obtenida presenta el avance de cada sede hacia la meta de estudiantes graduados, cada tono de gris indica el percentil, por ejemplo, el gris más oscuro representa el percentil 50. Estas metas fueron tomadas de manera arbitraria teniendo en cuenta la cantidad de estudiantes graduados por sede en todos los periodos para los cuales se tienen registro en el conjunto de datos

#### 2.4.2.10 Histogramas

Los histogramas son una representación gráfica de una variable en forma de barras, son útiles ya que proporcionan una vista general de la distribución de la población o de la muestra respecto a una característica, en este caso es útil mostrar el histograma para la variable edad ya que da una idea al usuario de la distribución de las edades que tiene los estudiantes que se graduaron en los períodos registrados en el conjunto de datos. Para crear un histograma siga estos pasos.

1. Cree una nueva hoja de trabajo y cambie su nombre por Histograma para edades.
2. Haga clic derecho sobre el campo Edad Mod, seleccione crear y Agrupaciones.
3. En la ventana para editar las agrupaciones como nombre del campo escriba Agrupaciones para edad, Tableau sugiere que el tamaño de las agrupaciones es 7.8, pero no es tamaño adecuado ya que los registros para la

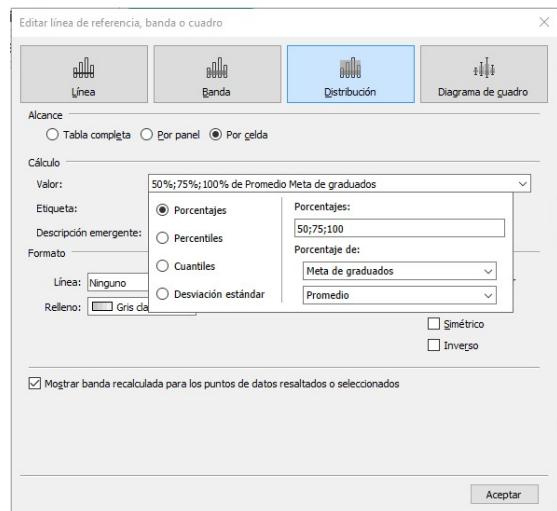


Figure 2.119: Edición de la banda de referencia para las metas

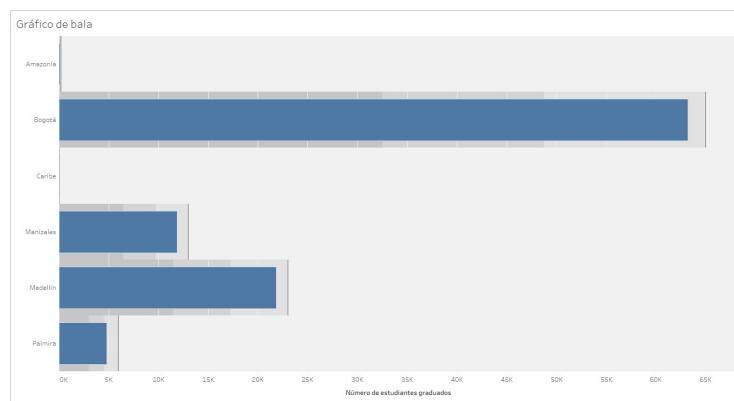


Figure 2.120: Gráfico de bala por sede de matrícula

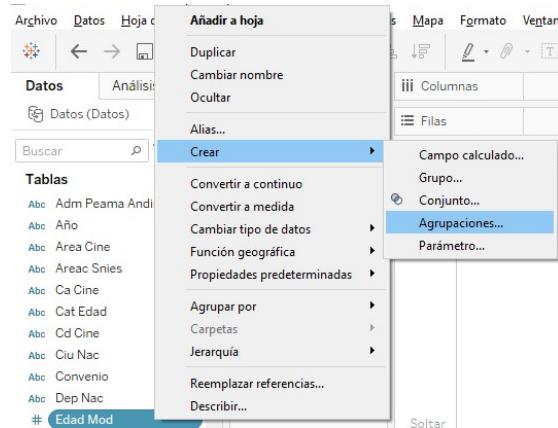


Figure 2.121: Crear agrupaciones

variable edad son números enteros, por tanto, el tamaño será 5, finalmente haga clic en aceptar.

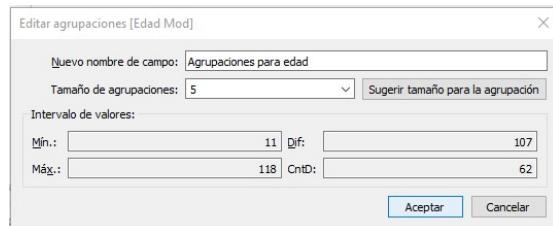


Figure 2.122: Editar agrupaciones

4. Añada el campo Edad mod al estante filtros, seleccione todos los valores y clic en aceptar.
5. En el menú desplegable del campo ubicado en filtros seleccione continuo se abrirá una ventana que permite editar el filtro seleccione intervalo de valores, cambie 11 por 15 y 118 por 70, esto con el fin de descartar esas edades atípicas ya que existe la posibilidad de que sean errores de digitación en los registros de la base de datos.
6. Arrastre el campo Edad Mod al estante filas, arrastre Agrupaciones para edad a columnas, en el menú desplegable del campo que ubico en filas seleccione medida y recuento; y para el campo en columnas seleccione continuo, su histograma debe verse así.
7. Haga clic derecho en el eje x y seleccione editar eje, en título escriba Edad; en la pestaña Marcas de graduación seleccione fijo, origen de graduación cero e intervalo de graduación cinco para marcas de graduación principal.

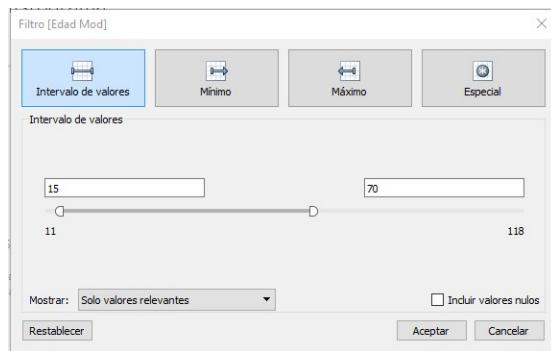


Figure 2.123: Filtrar valores atípicos de edad

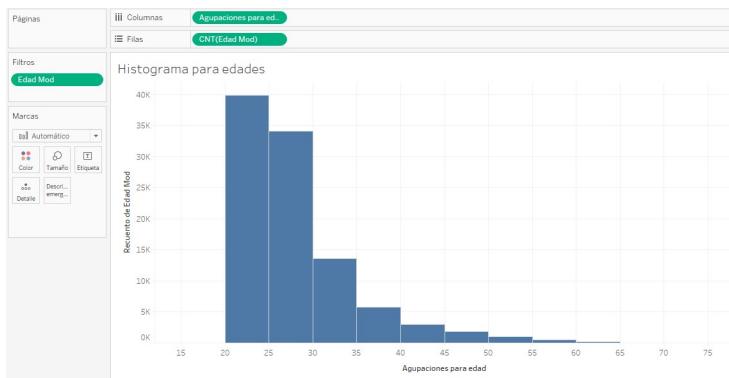


Figure 2.124: Histograma con un valor nulo



Figure 2.125: Editar eje x de agrupaciones

8. Cambie el título del eje y por Número de graduados y edite la descripción emergente.

El histograma obtenido permite concluir que las edades de los estudiantes graduados se concentran en el intervalo de 20 a 30 años.

Existe la posibilidad de añadir un filtro que permita visualizar este histograma para la sede de matrícula, y así evidenciar si la concentración de edades es la misma para todas las sedes de las cuales se tiene registro. Arrastre el campo Sede Nombre Mat al estante filtro, seleccione todas las sedes, muestre el filtro, cambie el nombre de la leyenda del filtro por Sede de matrícula, también seleccione valor individual (lista). Por ejemplo, si se selecciona como sede de matricula a Caribe se evidencia que la concentración de edades se traslada al intervalo de 30 a 35 años.

Otro factor que puede afectar esta distribución de edades es la modalidad de formación, se esperaría que para pregrado las personas sean más jóvenes que para postgrado, para añadir este nivel de detalle arrastre el campo modalidad de formación al estante columnas y suéltelo al lado izquierdo del campo agrupaciones para edad, también añada este campo a color.

#### 2.4.2.11 Box-plot

Estos diagramas son una herramienta estadística muy útil ya que permite visualizar la dispersión y simetría de una variable, permite la identificación de posibles valores atípicos. Para el conjunto de datos que se está trabajando

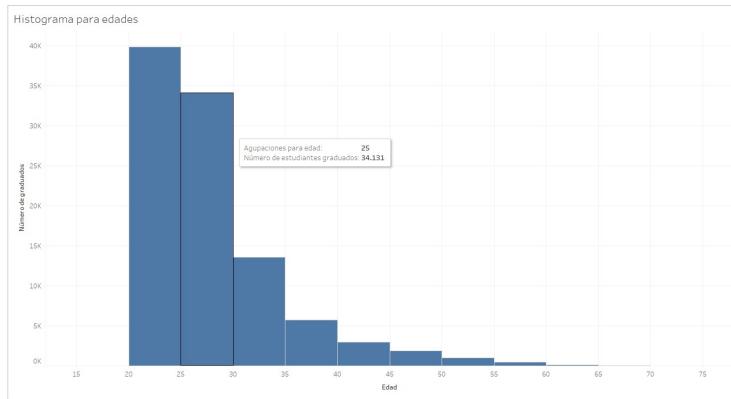


Figure 2.126: Histograma general para las edades

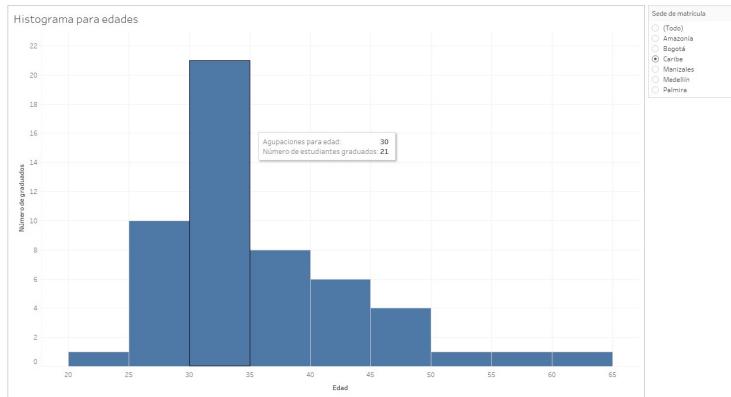


Figure 2.127: Histograma por sede para las edades

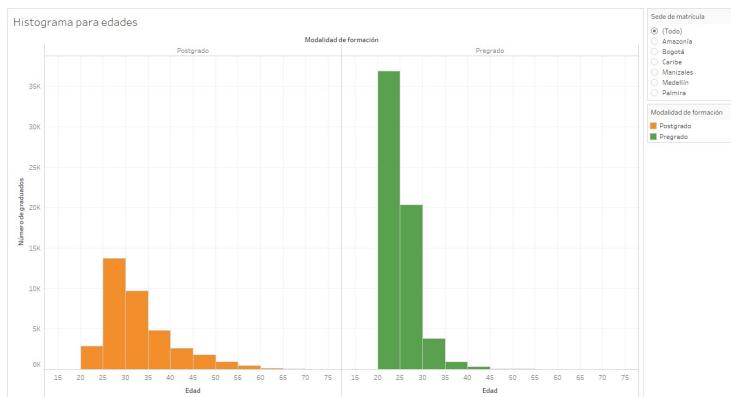


Figure 2.128: Histograma por sede y modalidad de formación para las edades

puede ser de interés realizar un box-plot para la variable edad, con el fin de identificar valores por fuera del límite inferior o superior, observar que tan simétrica es la variable y la dispersión de esta.

1. Cree una nueva hoja de trabajo y cambie su nombre por Boxplot para edad.
2. Repita los pasos 4 y 5 del histograma mostrados en 2.4.2.10.
3. Añada el campo Sede Nombre Mat al estante filtros, seleccione todo, muestre el filtro y cambie el modo de selección de la sede a Valor individual (lista), cambie el título de la leyenda por Sede de matrícula.
4. Añada el campo Nivel al estante columnas y Edad Mod a filas clic derecho sobre este último campo, seleccione medida y promedio.

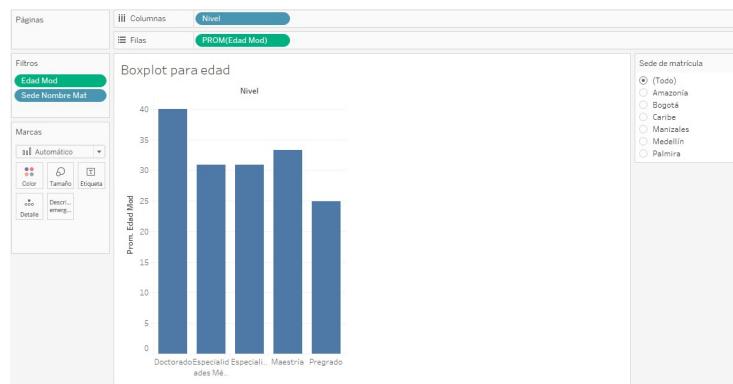


Figure 2.129: Paso 4 para la creación de un boxplot

5. Añada los campos Sede Nombre Mat y Year Semester a la tarjeta detalles.
6. Haga clic en el botón Mostrarme y seleccione diagramas de campos o valores.
7. Añada el campo Nivel a la tarjeta colores en el estante marcas y ajuste el tamaño de la vista a vista completa.
8. Cada punto de la visualización representa el promedio de las edades por periodo para cada nivel de formación según la sede que se elija en el filtro, es de interés mostrar las edades una a una y no el promedio para esto es necesario desagregar las medidas. Haga clic en la pestaña análisis y haga clic en Agregar medidas.
9. El eje Y de la visualización cambiara, ya que no representa el promedio de la edad si no la edad de cada estudiante. Finalmente edite la descripción emergente.

El gráfico realizado permite identificar la distribución de las edades por nivel de formación y realizar un filtro para la sede que se desea mostrar. Por

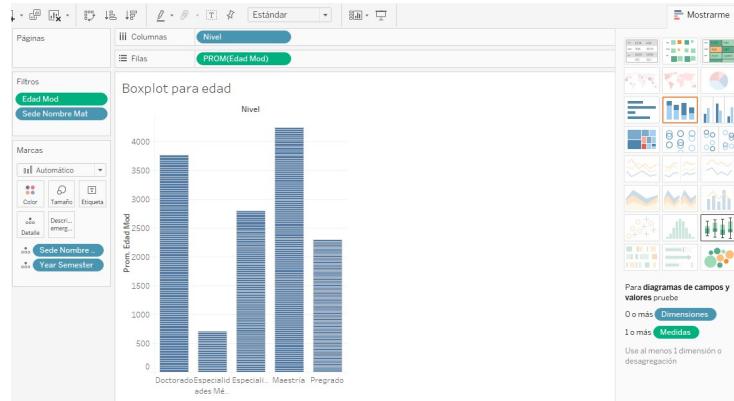


Figure 2.130: Uso del botón Mostrarme



Figure 2.131: Añadir color por nivel de formación

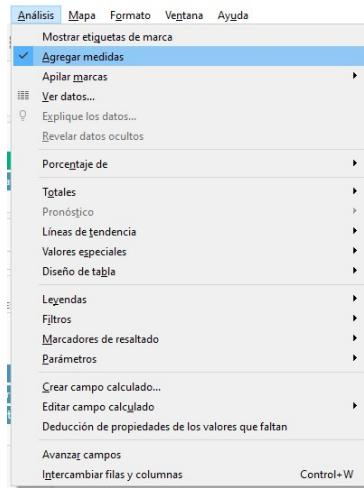


Figure 2.132: Desagregar medidas

ejemplo si se selecciona la sede Bogotá se observa un punto por debajo del límite inferior, al leer la descripción emergente de este punto se identifica que fue registrado para el periodo 2009-2 y la edad del estudiante de 15 años, esto puede deberse a errores al momento de digitar la información en la base de datos.

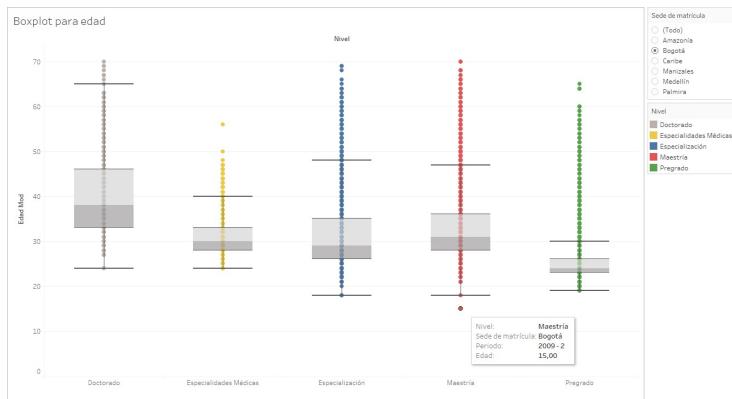


Figure 2.133: Boxplot para edad por sede y nivel de formación

#### 2.4.2.12 Mapeo de datos

La georreferenciación es una herramienta más apropiada cuando se desea mostrar la procedencia de los registros a nivel de ciudad, municipio o departamento; en el conjunto de datos que se está trabajando se tiene

información sobre la longitud y la latitud de la ciudad de nacimiento, el nombre de la ciudad y el departamento de los estudiantes.

Inicialmente se deben asignar funciones geográficas a los campos que contienen el departamento y la ciudad de nacimiento ya que estas fueron leídas por Tableau como cadenas, haga clic sobre el ícono “Abc” del campo Ciu Nac, seleccione función geográfica y Provincia/Municipio/Condado.

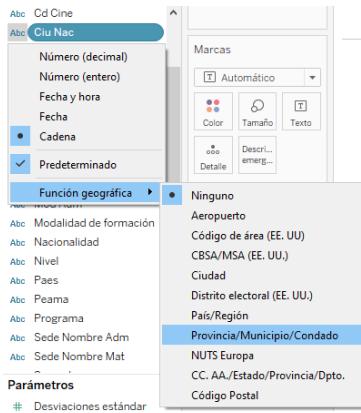


Figure 2.134: Asignación de función geográfica

Para el campo Dep Nac realice el mismo proceso, pero como función geográfica seleccione CC.AA./Estado/Provincia/Dpto. Debe notar que los iconos de estos campos cambiaron de “Abc” a un globo terráqueo que significa que ahora son variables con funciones geográficas.

Luego de tener listas las variables a usar es momento de hacer mapas, inicialmente se hará un mapa que contenga el total de graduados por departamento para el periodo 2020-1.

1. En una nueva hoja de trabajo llamada Total por departamento, periodo 2020-1 arrastre el campo Dep Nac a la tarjeta detalle, Tableau genera automáticamente los valores de latitud y longitud para los departamentos.
2. Haga clic en el botón 1 desconocido ubicado en la esquina inferior derecha, seleccione editar ubicaciones, se abrirá un cuadro de dialogo que permite identificar cual es el departamento que no coincide con los nombres que Tableau tiene internamente, se observa que la ubicación que no coincide es “NA”, es decir los datos vacíos.
3. Como son datos vacíos no es posible asignar una ubicación coincidente por lo tanto cierre el cuadro de dialogo mostrado en el paso anterior, nuevamente haga clic en el botón 1 desconocido y seleccione filtrar datos, esto con el fin de eliminar esos datos vacíos.
4. En Marcas cambie automático por Mapa, esto lo que hace es llenar el

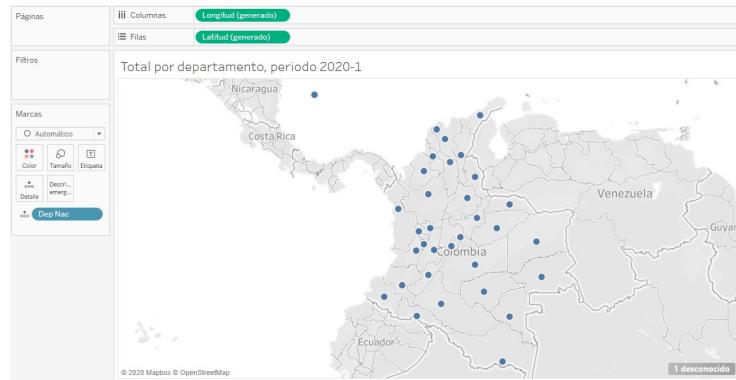


Figure 2.135: Mapa base por departamentos

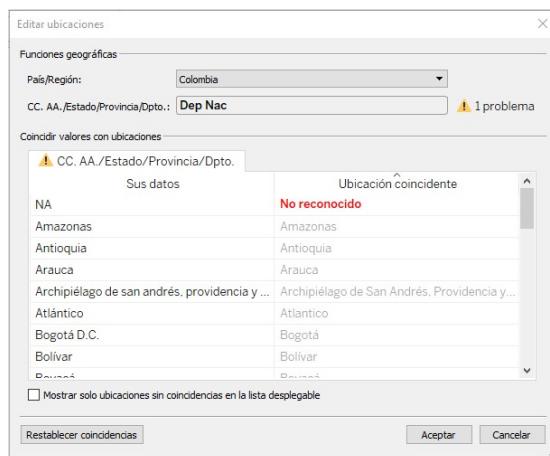


Figure 2.136: Editar ubicaciones

área que pertenece a cada departamento y no tener un solo punto por departamento.



Figure 2.137: Asignar áreas por departamento

5. Añada el campo Dep Nac a la tarjeta etiqueta, también debe añadirlo a color, en color seleccione medida y recuento.



Figure 2.138: Agregar color y etiquetas

6. Interesa mostrar los registros para el periodo 2020-1, para esto arrastre el campo Year Semester al estante filtros y seleccione únicamente el periodo 2020-1.
7. Edite la descripción emergente para que le quede de esta manera.
8. Edite el título de la leyenda de color por Total de graduados.

A continuación, se presenta la visualización obtenida, la cual permite identificar la cantidad de graduados por departamento para el periodo 2020-1, por ejemplo, para este periodo se graduaron 2 estudiantes provenientes del departamento de Vichada.

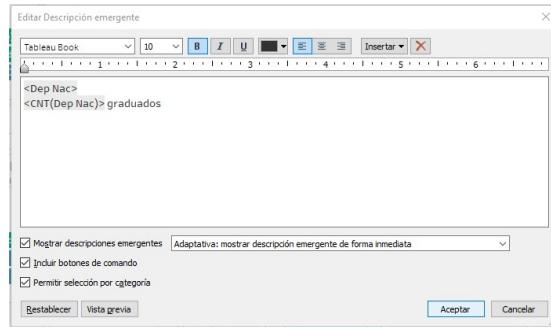


Figure 2.139: Editar adecuadamente la descripción emergente

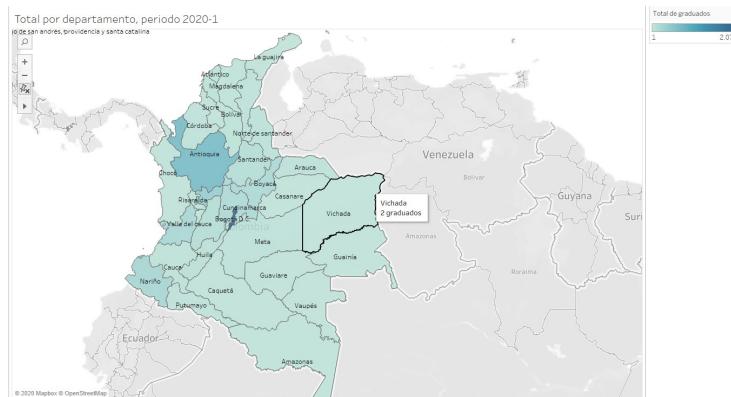


Figure 2.140: Total por departamento, periodo 2020-1

También es posible crear un mapa para mostrar el total de estudiantes graduados por municipio para el periodo 2020-1.

1. Cree una nueva hoja de trabajo y asigne como nombre Total por municipio, periodo 2020-1.
2. Arrastre el campo Ciu Nac y Dep Nac a detalle, obtendrá el mapa de Colombia con puntos que simbolizan cada ciudad para la que se tiene registros, en la esquina inferior encontrara el número de ciudades que Tableau no logró identificar que en este caso son 31, haga clic sobre ese botón y seleccione editar ubicaciones. Es posible que muchas ubicaciones no coincidan por tildes o mayúsculas por lo que es necesario ir asignando las ubicaciones. Por ejemplo, Cartagena de indias no se reconoció correctamente ya que la base interna de Tableau contiene solo Cartagena, por lo tanto, asigne Cartagena a Cartagena de Indias como ubicación coincidente, de esta manera asigne los demás que Tableau no logró reconocer.

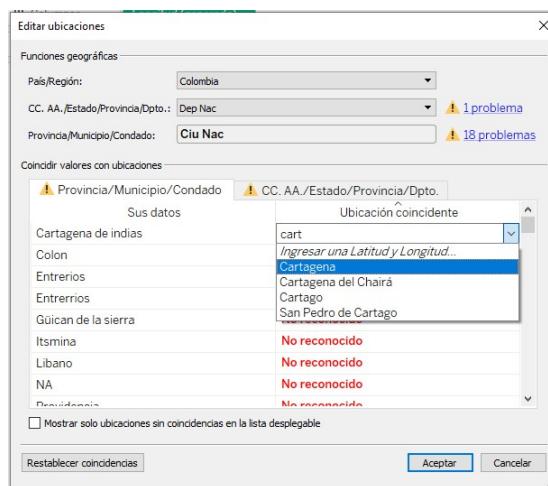


Figure 2.141: Editar ubicaciones para ciudades

3. Después de asignar todos los municipios conocidos haga clic en aceptar, es posible que aun queden valores desconocidos, los cuales pertenecen a valores faltantes en los datos por lo tanto haga clic sobre el botón de desconocido y seleccione Filtrar datos.
4. Nuevamente cambie la marca de automático a mapa para que las áreas se rellenen y añada el filtro de Year Semester seleccionando únicamente el periodo 2020-1.
5. El mapa se ve con muchos espacios ya que solo se llenan las áreas de las ciudades que pertenecen al periodo 2020-1. Añada el campo Ciu Nac a color y seleccione medida y recuento.

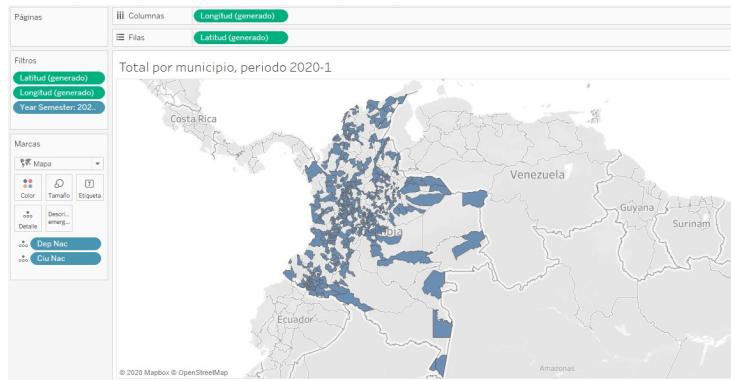


Figure 2.142: Rellenar el mapa por ciudades



Figure 2.143: Añadir color por ciudades

6. En la pestaña mapa seleccione capas de mapa.

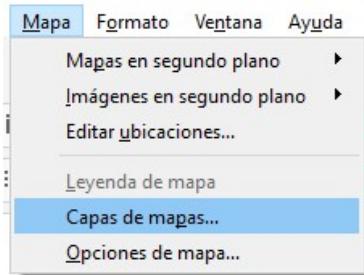


Figure 2.144: Editar las capas del mapa

7. Se mostrará un panel en el lateral izquierdo que permite seleccionar los atributos que se quieren mostrar en el mapa, en este caso se selecciona Nombres de estados/ provincias y Nombres de condados, esto hará que el mapa muestre el nombre de los municipios.

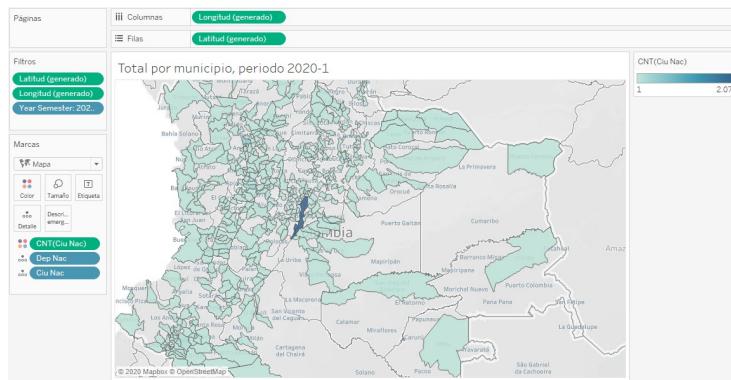


Figure 2.145: Agregar nombres de municipios

8. Edite la descripción emergente para que le quede de la siguiente manera.
9. Por último, edite el título de la leyenda de color.

Este mapa permite visualizar el municipio de nacimiento de los estudiantes graduados en el periodo 2020-1, por ejemplo 2 de los estudiantes nacieron en Inírida (Guainía).

#### 2.4.2.13 Gráfico de dispersión

Estos gráficos permiten identificar las relaciones existentes entre dos variables numéricas, en Tableau pueden ser creados agregando una medida al estante filas y otra al estante columnas, es posible añadir dimensiones para agregar detalle a su vista. El conjunto de datos con el cual se realizaron todas las

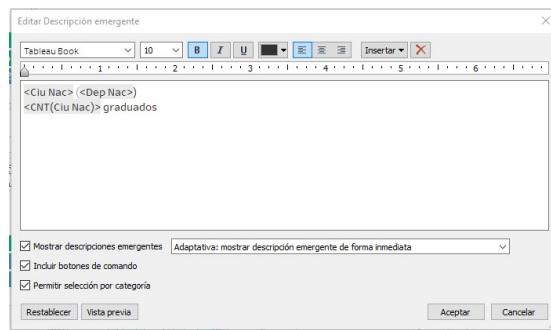


Figure 2.146: Editar la descripción emergente

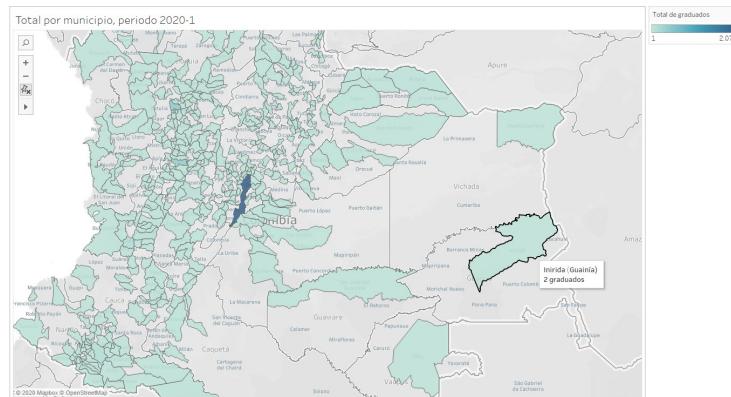


Figure 2.147: Total por municipio, periodo 2020-1

visualizaciones anteriores no posee dos variables numéricas que permitan hacer gráficos de dispersión por esto es necesario conectarse a una nueva fuente de datos, se hará en un libro nuevo ya que a pesar de ser gráficos importantes no están relacionados con el conjunto trabajado anteriormente, para crear un nuevo libro de trabajo haga clic en archivo y seleccione nuevo.

Conéctese a la fuente de datos Sample-Superstore.xls y seleccione la hoja Orders, este conjunto de datos contiene información sobre productos vendidos.

1. A la hoja uno de trabajo llámela Diagrama de dispersión para ventas y beneficios.
2. Arrastre el campo Sales a filas y el campo Profit a columnas y ajuste el tamaño de la vista.

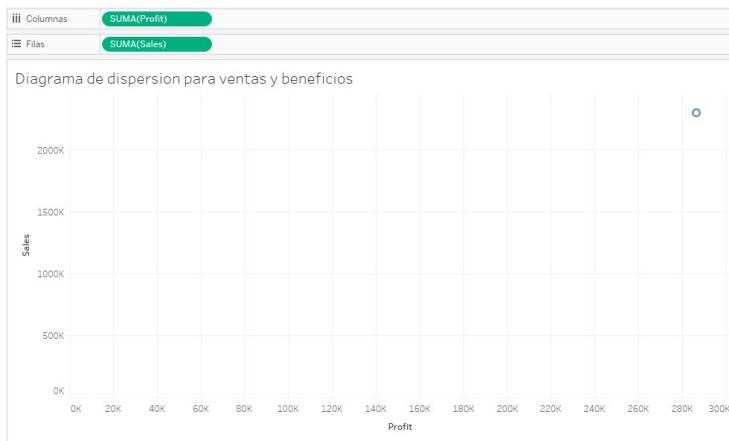


Figure 2.148: Diagrama de dispersión base

3. Por defecto Tableau usa la agregación de suma para las medidas numéricas para mostrar todos los puntos se deben desagregar las medidas, haga clic en la pestaña análisis y seleccione agregar medidas. Después de hacer eso su visualización contendrá todos los puntos para ventas y beneficios.
4. Este diagrama de dispersión tiene una mezcla de tres categorías de productos, por lo tanto, para agregar más detalle a la vista añada el campo Category a la tarjeta color.
5. Tableau permite añadir líneas de tendencia que representarían modelos de regresión lineal simple, desde el panel análisis en modelo arrastre línea de tendencia y suéltala en lineal, observe que Tableau proporciona más opciones tales como logarítmica, exponencial, polinómica y potencia.
6. Se agregarán tres modelos de regresión ya que la vista esta dividida en las categorías de los productos, para que la vista se divida y sea más clara añada el campo Category a columnas y ubíquelo a la derecha de Profit.

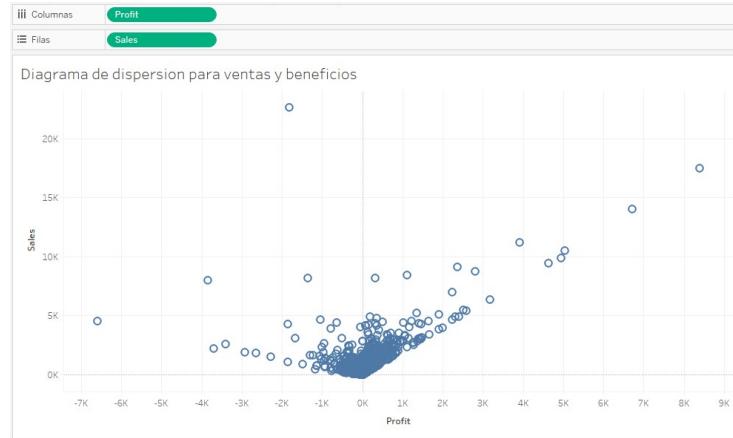


Figure 2.149: Diagrama de dispersión con medidas desagregadas

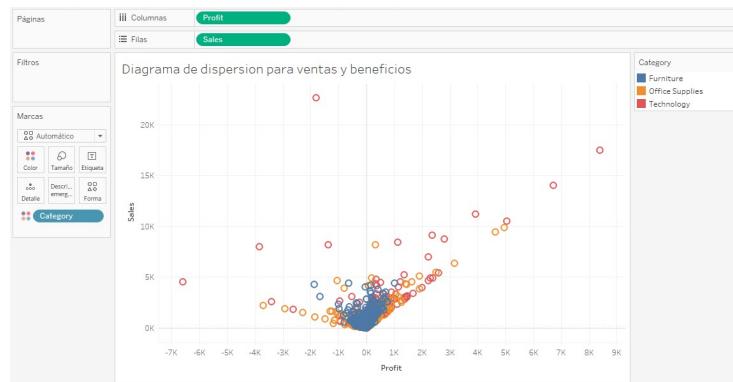


Figure 2.150: Añadir color según categoría

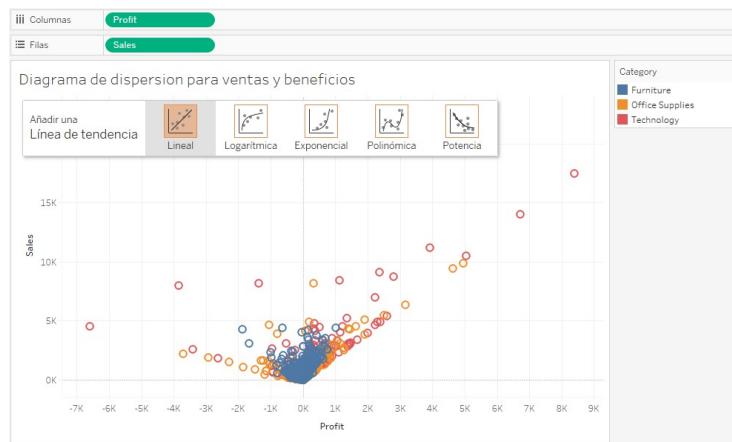


Figure 2.151: Opciones de línea de tendencia

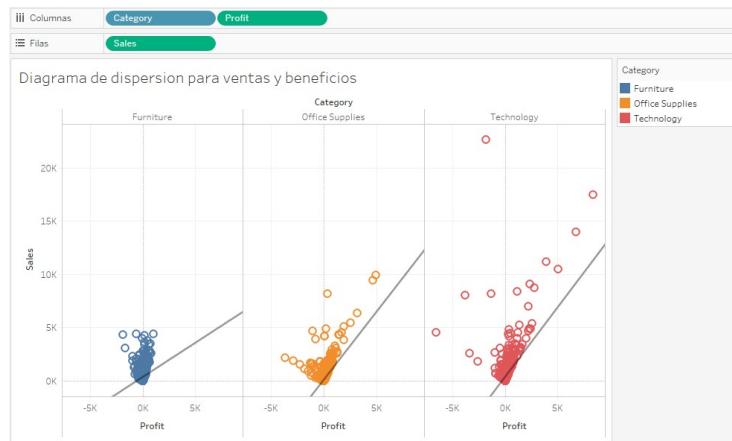


Figure 2.152: Dividir el eje según la categoría del producto

7. La descripción emergente de las líneas de tendencia ajustadas da información sobre la ecuación del modelo, la mediada de  $R^2$  y  $valor - p$ , valores útiles para realizar análisis estadísticos.

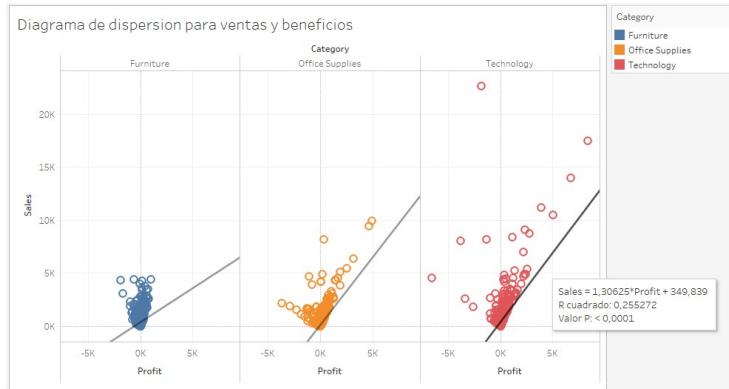


Figure 2.153: Descripción emergente de la línea de tendencia

8. Haciendo clic derecho sobre una de las líneas de tendencia y seleccionando describir línea de tendencia se abrirá un cuadro de dialogo que permite visualizar la ecuación del modelo, sus coeficientes, desviaciones estándar, valores del estadístico t y valores P.



Figure 2.154: Describir línea de tendencia

9. Haciendo clic derecho sobre una de las líneas de tendencia y seleccionando describir modelo de tendencia se abrirá un cuadro de dialogo que hace una descripción de como fue ajustado el modelo general.
10. Para editar el modelo de regresión ajustado, haga clic sobre una línea de tendencia y seleccione editar, en la ventaja emergente que se abre encontrarás diferentes opciones, como cambiar el tipo de modelo, hacer que los modelos se calculen de manera independiente según una dimensión y agregar objetos visuales como bandas de confianza.

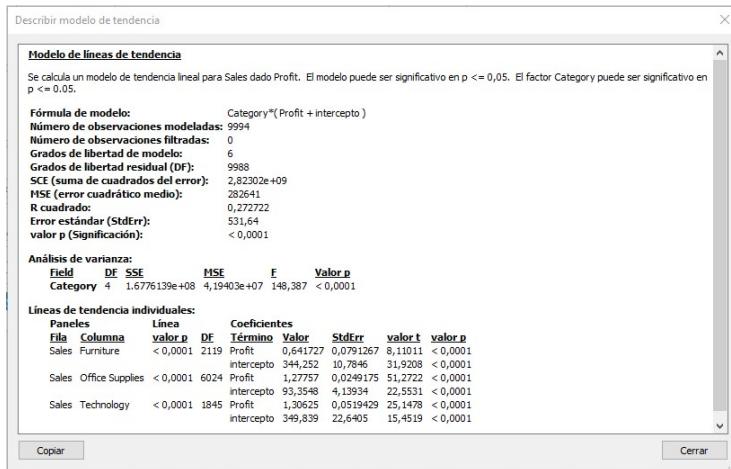


Figure 2.155: Describir modelo de tendencia

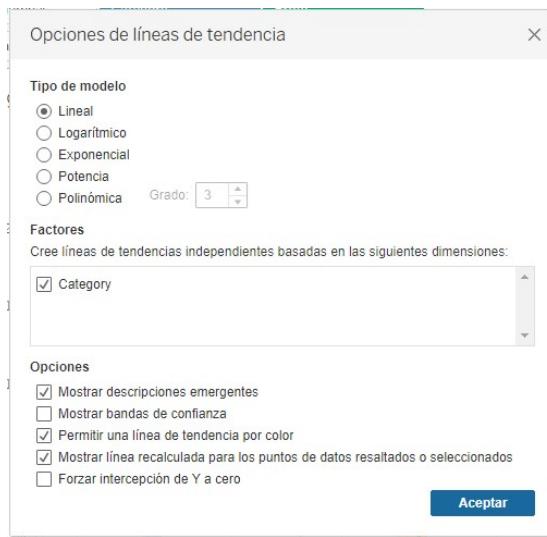


Figure 2.156: Opciones de líneas de tendencia

#### 2.4.2.14 Pronósticos para series de tiempo univariadas

Tableau permite generar pronósticos para una serie de tiempo, para esto es necesario tener un campo con el tipo fecha y como mínimo una medida, se inicia con un gráfico de líneas básico y se añade el pronóstico.

1. Cree una nueva hoja de trabajo y asigne como nombre evolución histórica de las ventas.
2. Arrastre el campo Order Date a columnas y Sales a filas, ajuste el tamaño a Vista completa.



Figure 2.157: Gráfico de líneas base con fecha discreta

3. En el menú desplegable del campo ubicado en columnas seleccione mes, observe que hay dos opciones para año, trimestre y mes, el primer bloque toma esto como discreto, por ejemplo, mes acumularía las ventas por mes de todos los años, pero si se selecciona mes del segundo bloque las fechas se trabajaran como continuas, es decir que se grafica cada mes de cada año.
4. Obtendrá una serie de tiempo mensual para las ventas.
5. Es momento de añadir el pronóstico de ventas, haga clic en el panel análisis y desde modelo arrastre pronostico y suéltelo en añadir pronostico.

Note que Tableau automáticamente añade una leyenda de color que indica cuales son las observaciones reales y las estimaciones, por defecto la línea de pronósticos se añada con un intervalo de confianza del 95%, esto puede ser editado al hacer clic sobre un punto de la línea de pronóstico y seleccionando editar, se abre un cuadro de dialogo que permite editar el pronóstico, es posible editar la duración del mismo, elegir la fuente de datos, decidir cuando meses se quieren ignorar para la realización del pronóstico, elegir el modelo y seleccionar el tamaño del intervalo de confianza. Para obtener información

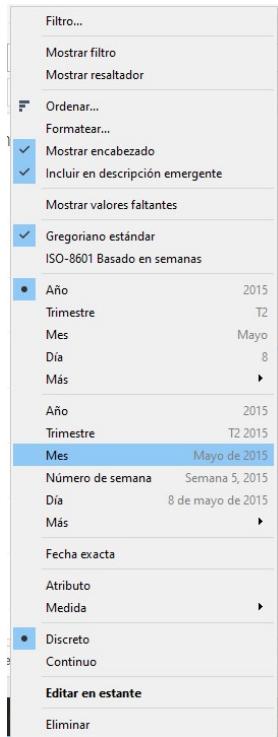


Figure 2.158: Opciones del campo fecha

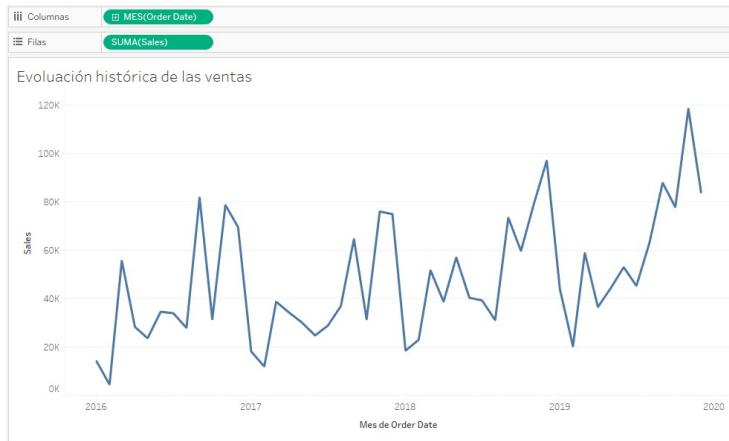


Figure 2.159: Serie de tiempo mensual para las ventas



Figure 2.160: Serie de tiempo mensual para las ventas con pronóstico

más detalla acerca de las opciones puede hacer clic en Obtener más información sobre las opciones de pronóstico, el cual lo llevara a la pagina de ayuda de Tableau y le proporcionara información detallada sobre las opciones.

La descripción del pronóstico puede ser obtenida haciendo clic derecho sobre la estimación seleccionar pronóstico y describir pronostico, se abrirá una ventana con dos pestañas una de resumen y otra de modelos.

La pestaña de resumen contiene una descripción general de los modelos de pronóstico que Tableau ha creado, también los patrones generales que encontró en los datos. Para obtener información mas detalla sobre lo que contiene esta pestaña puede hacer clic en Obtener más información sobre el resumen de pronósticos.

La pestaña modelos presenta información mas detalla sobre el modelo de pronóstico que Tableau ha creado; contiene el nivel, tendencia, temporada, métricas de calidad y coeficientes de suavizado, si se añada una dimensión a la vista se genera una fila en la tabla por cada nivel de la dimensión.

Nuevamente se encuentra disponible el botón para obtener información mas detallada sobre los modelos de pronósticos.

Después de crear todos los objetos visuales necesarios para implementar tableros e historias, es necesario dar formato al libro de trabajo en cuanto a fuentes, tamaño y color de títulos de hojas de trabajo, tableros e historias y descripciones emergentes, con el fin de que el libro de trabajo contemple la misma fuente en todas sus hojas.

Haga clic en el botón formato y seleccione libro de trabajo, en el panel derecho se ubican fuentes y líneas, aumente el tamaño de los títulos de las hojas de trabajo seleccionando tamaño 18, color negro y negrilla, haga lo mismo para los títulos de tableros e historias, si quiere puede cambiar el tipo de fuente, en

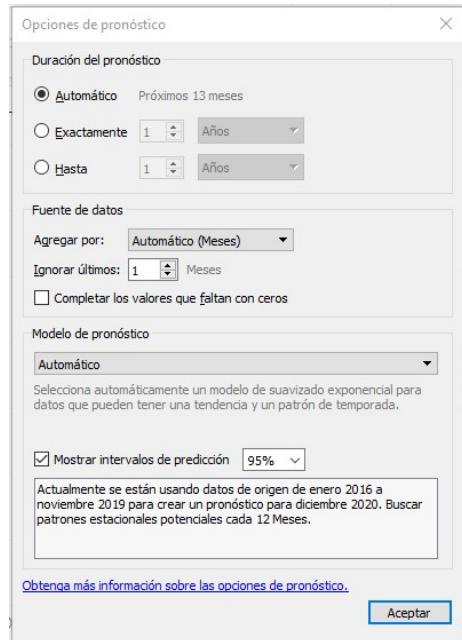


Figure 2.161: Opciones de pronóstico



Figure 2.162: Describir pronóstico - resumen



Figure 2.163: Describir pronóstico - modelos

ese caso se dejan los predeterminados por Tableau.

### 2.4.3 Compartir las visualizaciones

Las visualizaciones creadas usando la versión de escritorio de Tableau Public pueden ser compartidas como hojas individuales, dashboard o historias, en este caso se crearán dashboard que luego se integran a historias para replicar la forma de visualización que se publica en la página de estadísticas estadísticas de la Universidad Nacional de Colombia en la pestaña cifras generales y graduados. Es necesario crear ocho tableros con las visualizaciones que se crearon anteriormente y luego se incluirán en una historia con ocho puntos.

#### 2.4.3.1 Crear un dashboard

Al abrir la página web de estadísticas de la universidad y seleccionar cifras generales, graduados se identifica que el primer gráfico es una serie de tiempo que presenta la evolución histórica del total de estudiantes graduados, observe que solo posee una visualización por lo que no es necesario crear un dashboard y es posible agregar esta hoja directamente a la historia. Sin embargo, se quiere añadir una funcionalidad que permita ver los metadatos asociados a las cifras generales que fueron analizadas.

1. Cree una nueva hoja de trabajo y llámela Metadatos.
2. Cree un nuevo parámetro u signe como nombre Ver metadatos, seleccione cadena y lista, en valores escriba Ver metadatos y clic en aceptar para guardar los cambios.
3. Cree un nuevo campo calculado y llámelo Metadatos, en el espacio en blanco escriba [Ver metadatos] y haga clic en aceptar.

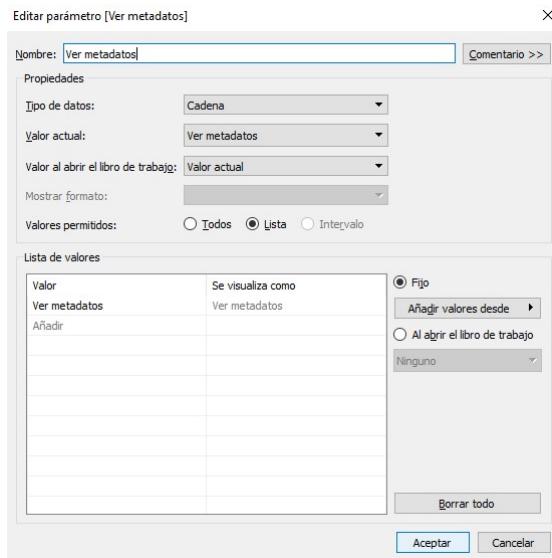


Figure 2.164: Crear parámetro para metadatos



Figure 2.165: Crear campo calculado para metadatos

4. Arrastre este campo calculado que creó a la tarjeta texto ubicada en el estante Marcas, haga clic en esta tarjeta y modifique le tamaño del texto a 18, finalmente, ajuste la anchura desde la barra de herramientas, finalmente asigne como nombre del dashboard Evolución graduados.
5. Cree un nuevo dashboard, ajuste el tamaño a automático, desde el panel hojas arrastre la hoja que contiene la evolución del total de estudiantes graduados.



Figure 2.166: Agregar hoja de evolución de graduados

6. Añada la hoja llamada Metadatos y suéltela en la esquina superior derecha, oculte el título de esta hoja.
7. Haga clic en el botón dashboard ubicado en la parte superior cerca a la barra de herramientas y seleccione Acciones.

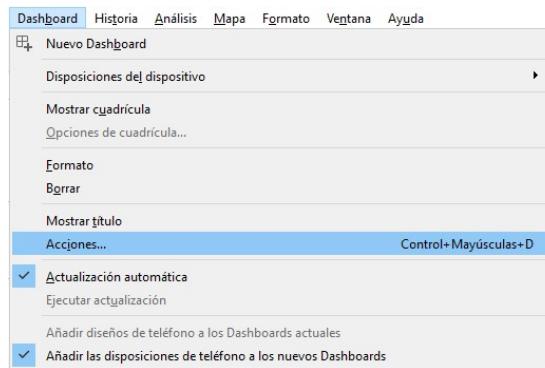


Figure 2.167: Agregar acciones a un dashboard

8. En el cuadro de dialogo Acciones haga clic en Añadir acción y seleccione URL.

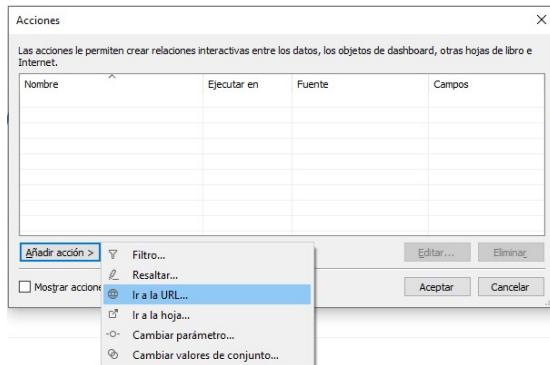


Figure 2.168: Elegir acción de URL

9. En el campo nombre escriba Metadatos, como hoja de origen establezca el dashboard llamado Evolución graduados y únicamente seleccione la hoja llamada metadatos, en la ejecución de la acción haga clic en seleccionar, en el campo URL pegue el enlace asociado a los metadatos ubicados en la página web de las estadísticas oficiales de la Universidad Nacional, por último, clic en aceptar.

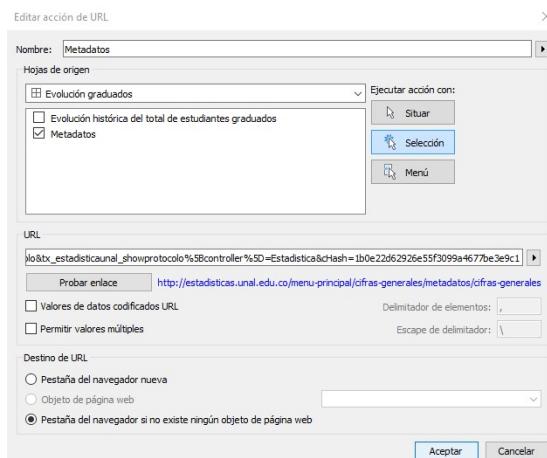


Figure 2.169: Editar acción de URL

Finalmente obtendrá un dashboard que permite visualizar la evolución de los estudiantes graduados en la Universidad Nacional de Colombia, con un botón ubicado en la esquina superior izquierda que permite ir a los metadatos alojados en la página web de las estadísticas de la Universidad.

La siguiente visualización presenta una combinación de una serie de tiempo, una tabla de texto y un diagrama circular con la distribución de los

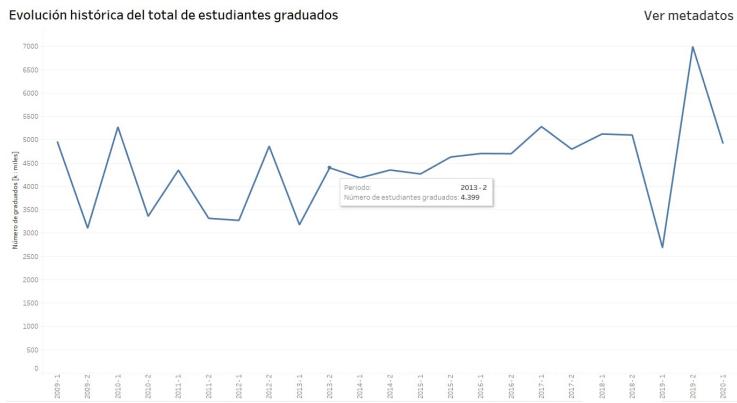


Figure 2.170: Dashboard para la evolución del total de graduados

estudiantes graduados en el periodo actual. Todas estas visualizaciones ya fueron creadas, pero es necesario agregar un botón que permita intercambiar entre la serie y la tabla.

1. Diríjase a la hoja de trabajo llamada serie, haga clic en el menú desplegable ubicado entre datos y tablas y seleccione crear parámetro.
2. En el cuadro de dialogo escriba como nombre del parámetro Seleccione visualización, tipo de datos seleccione cadena, en valores permitidos seleccione lista, finalmente en Lista de valores, escriba serie como el primer valor y, a continuación, tabla que son las dos visualizaciones que se desean integrar al dashboard.
3. Ahora se debe crear un campo calculado que incluya el parámetro, cree un nuevo campo calculado llámelo, Mostar hoja, en el cuadro de texto de la fórmula, escriba el nombre del parámetro que creó anteriormente y haga clic en aceptar.
4. Añada el campo calculado creado en el paso anterior al estante filtros, seleccione personalizar lista de valores escriba serie en el cuadro de texto y haga clic en el botón añadir elemento y clic en aceptar.
5. Repita el paso anterior para todas las hojas que deseé añadir al dashboard.
6. Muestre el parámetro haciendo clic en el menú desplegable del mismo y seleccione mostrar parámetro, para todas hojas en las que añadió el campo calculado al estante filtro.
7. Cree un nuevo dashboard, ajuste el tamaño a automático, desde el panel hojas arrastre la hoja que contiene el gráfico circular para la distribución de graduados por modalidad de formación para el periodo actual.
8. Desde el panel objetos arrastre un contendor de trazado vertical hasta el

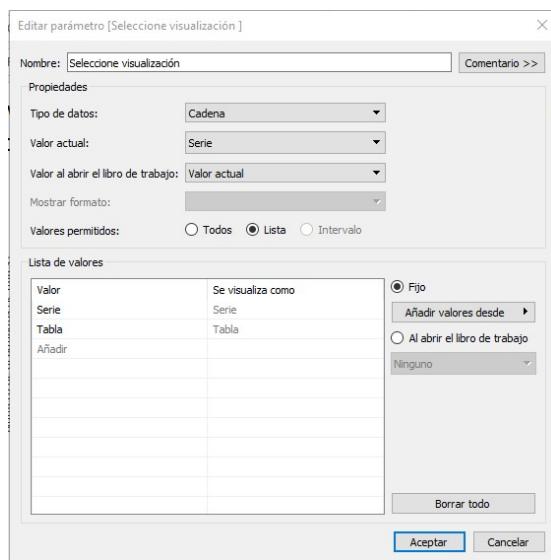


Figure 2.171: Editar parámetro de selección de visualizaciones



Figure 2.172: Crear campo calculado con el parámetro

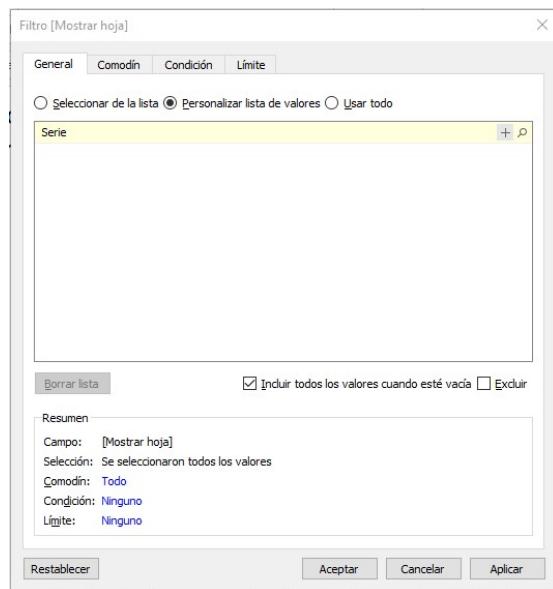


Figure 2.173: Editar filtro del campo calculado



Figure 2.174: Añadir una visualización al dashboard

dashboard y suéltelo sobre el rectángulo gris que ocupa la mitad izquierda del tablero.

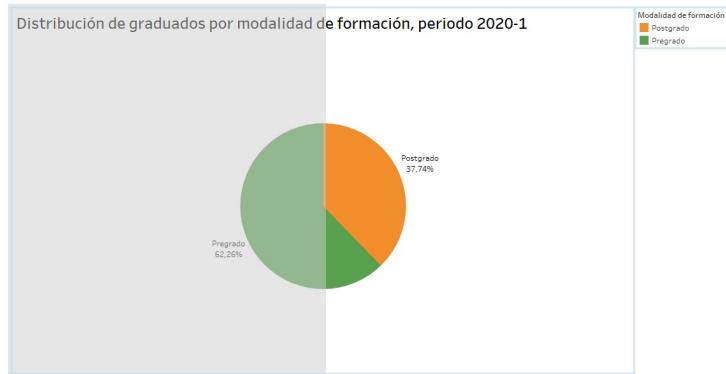


Figure 2.175: Agregar un contenedor de trazado vertical

9. Arrastre la hoja Serie al contenedor, arrastre también la hoja que contiene la tabla de texto y suéltela sobre el rectángulo gris claro que se crea debajo del título de la serie.



Figure 2.176: Agregar vistas al contenedor de trazado vertical

10. Debe ocultar los títulos de las hojas puestas sobre el contenedor, haga clic sobre el título de la serie, seleccione más opciones y luego título, esto debe hacerlo también para la tabla.
11. Desde el panel objetos arrastre texto y suéltelo en la parte superior del contenedor.
12. En el cuadro de dialogo editar texto escriba evolución histórica del total de estudiantes graduados por modalidad de formación, ajuste el tamaño de la letra a 18, seleccione negrilla y centrar el texto.



Figure 2.177: Ocultar títulos

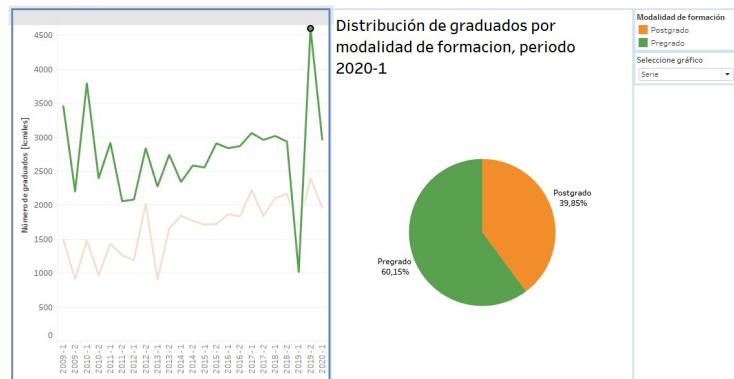


Figure 2.178: Agregar cuadro de texto

13. Cambie el nombre del dashboard a modalidad de formación, por comodidad y facilidad se eliminaron los filtros de la hoja tabla.
14. Agregue una acción de filtro como se mostro anteriormente.

A continuación, se presenta el dashboard obtenido que proporciona información clara sobre la modalidad de formación, permite intercambiar entre la serie y la tabla de texto y al hacer clic sobre Ver metadatos permite ir a los metadatos asociados a estas visualizaciones.

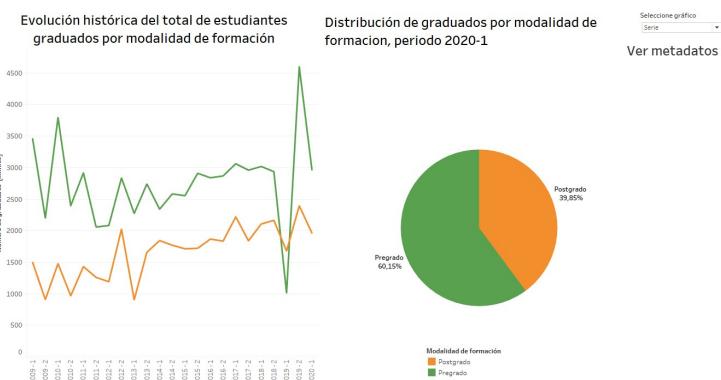


Figure 2.179: Dashboard para modalidad de formación

Oculte las hojas que ha usado en el dashboard anterior con el fin de que la barra de hojas de la parte inferior no se sature, haga clic derecho sobre el nombre de la hoja de trabajo y seleccione ocultar.

El siguiente dashboard a realizar corresponde a las estadísticas por nivel de formación, debe crear todos los objetos visuales que involucra este dashboard e incorporarlos al tablero como se mostro anteriormente, el dashboard para nivel de formación debe verse así,

Para el tablero que contiene las estadísticas por sede de admisión se decide no usar el gráfico de barras y en su lugar incluir el mapa de árbol, con el fin de mostrar visualizaciones diferentes pero que informan de manera clara.

Tablero para nacionalidad.

Estadísticas para lugar de nacimiento.

El siguiente tablero presenta las estadísticas por sexo de los estudiantes graduados.

Finalmente se realiza el dashboard con las estadísticas por áreas del conocimiento.

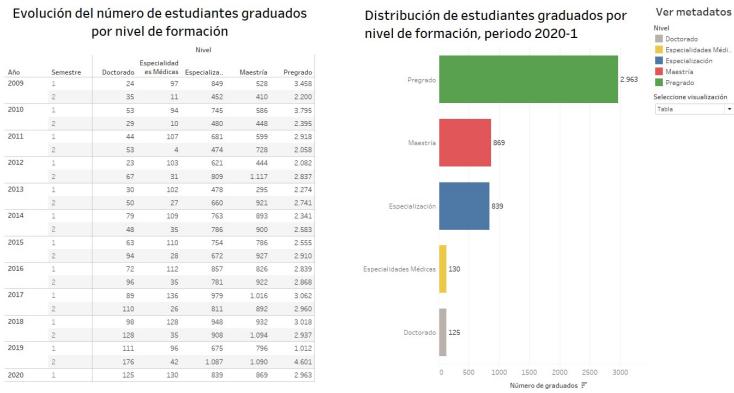


Figure 2.180: Dashboard para nivel de formación



Figure 2.181: Dashboard para sede de admisión

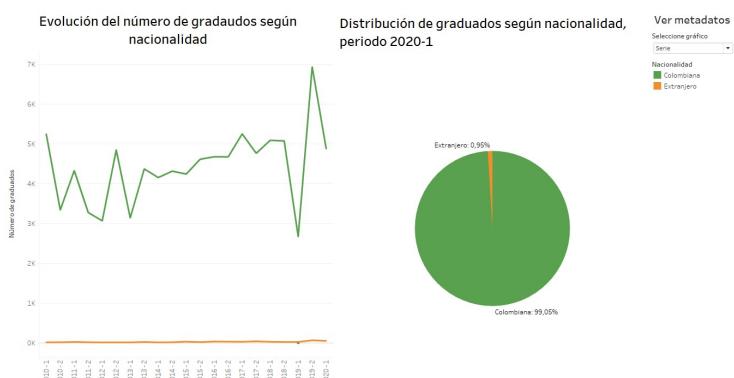


Figure 2.182: Dashboard para nacionalidad

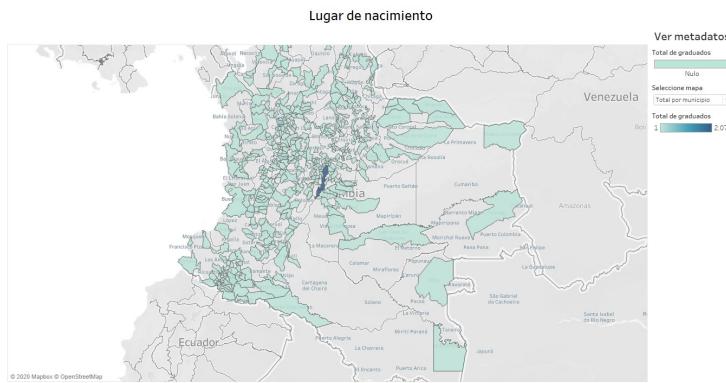


Figure 2.183: Dashboard para lugar de nacimiento

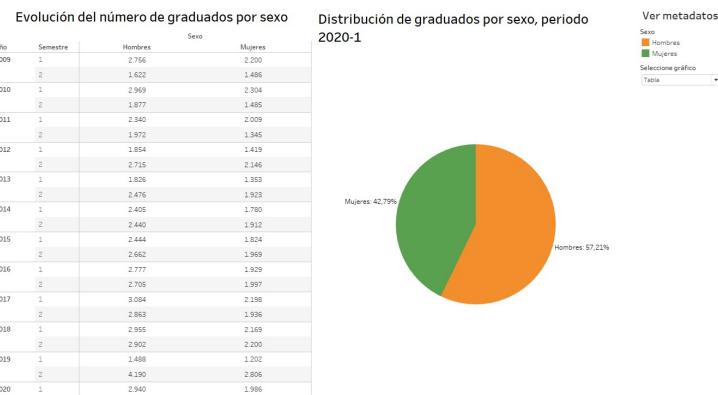


Figure 2.184: Dashboard por sexo

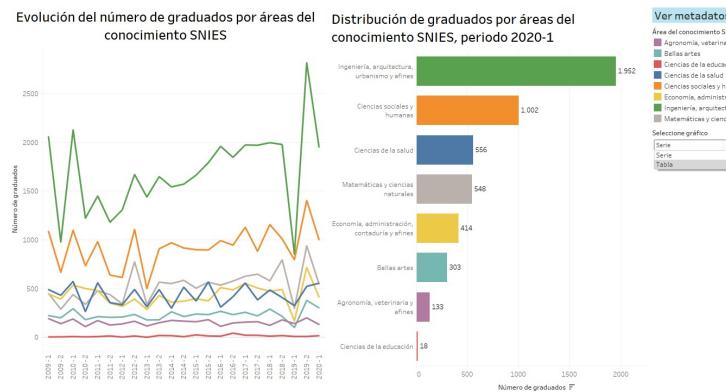


Figure 2.185: Dashboard por áreas del conocimiento SNIES

#### 2.4.3.2 Crear una historia

Como ya se menciono una historia es una secuencia de vistas o dashboards que se utilizan en forma conjunta para mostrar información, en este caso se hará uso de las historias para mostrar una secuencia de las visualizaciones y dashboards creados anteriormente y lograr una estructura similar a la mostrada en la página web de las estadísticas de la Universidad.

1. Haga clic en el icono de historia en la barra inferior donde se ubican los tableros creados, asigne como nombre a la nueva historia Graduados.
2. Ajuste el tamaño para que ocupe toda la pantalla del dispositivo.
3. En el cuadro gris añada como subtítulo Evolución graduados, en el panel lateral izquierdo en la pestaña historia encontrará un botón que tiene el texto En blanco, dicho botón es usado para crear más puntos en la historia, haga clic allí hasta tener 8 puntos en la historia. Añada estos subtítulos a los puntos de la historia en este orden Modalidad de formación, Nivel de formación, Sede de admisión, Nacionalidad, Lugar de nacimiento, Sexo, Áreas del conocimiento.
4. En el primer punto de la historia agregue el dashboard llamado Evolución graduados.



Figure 2.186: Añadir una visualización a la historia

5. A los demás puntos de las historias añada los tableros que fueron creados en la sección anterior.
6. Haga clic en la pestaña formato ubicada sobre la barra de herramientas y seleccione historia.
7. Se despliega un menú en el panel lateral izquierdo que permite formatear el aspecto de la historia, en la sección navegador haga clic en fuente, aumente el tamaño de la fuente y seleccione negrita; ahora haga clic en sombreado

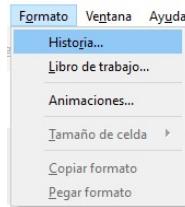


Figure 2.187: Formato de la historia

y más colores, agregue este código hexadecimal que coincide con el color usado en la página web “#636363”.

Se obtiene una historia que permite navegar entre los diferentes dashboards que contienen información sobre las estadísticas generales de la Universidad.



Figure 2.188: Historia: graduados

Oculte todos los dashboards involucrados en la historia, haga clic en la pestaña archivo y seleccione guardar en Tableau Public como...

Asigne un título al libro de trabajo y haga clic en guardar, con esto se abrirá una pestaña en su navegador que contiene la visualización publicada en su cuenta de Tableau Public, puede hacer clic en botón compartir y copiar el código para insertar en alguna página web o el enlace.

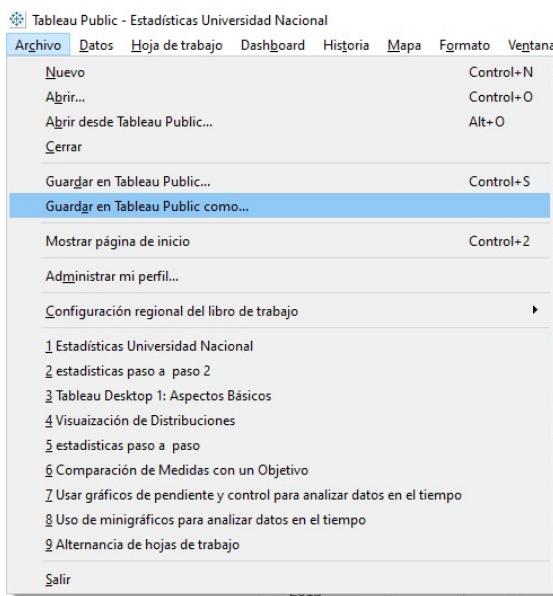


Figure 2.189: Guardar el libro de trabajo

# Chapter 3

## Power BI

### 3.1 Generalidades

#### 3.1.1 ¿Qué es Power BI?

Microsoft Power BI es una colección de servicios de software, aplicaciones y conectores que ayudan a las organizaciones a recopilar, administrar y analizar datos de una variedad de fuentes, a través de una interfaz fácil de usar. Funcionan en conjunto para convertir sus fuentes de datos no relacionadas en conocimientos coherentes, visualmente inmersivos e interactivos. Ya sea que sus datos sean un simple libro de trabajo de Microsoft Excel o una colección de almacenes de datos híbridos locales y basados en la nube, Power BI le permite conectarse fácilmente a sus fuentes de datos, limpiar y modelar sus datos sin afectar la fuente subyacente, visualizar o descubrir lo que es importante para compartirlo con su organización.

Esta aplicación fue concebida originalmente por Thierry D'Hers y Amir Netz del equipo de SQL Server Reporting Services en Microsoft. Fue diseñado originalmente por Ron George en el verano de 2010 y nombrado Proyecto Crescent, estaba disponible inicialmente para su descarga pública el 11 de julio de 2011 incluido con SQL Server Codename Denali. Más tarde renombrado a Power BI, Microsoft lo dio a conocer en septiembre de 2013 como Power BI para Office 365. La primera versión de Power BI se basó en complementos de Microsoft Excel: Power Query, Power Pivot y Power View. Con el tiempo, Microsoft también agregó muchas características adicionales como preguntas y respuestas, conectividad de datos de nivel empresarial y opciones de seguridad a través de las puertas de enlace de Power BI. Power BI fue lanzado por primera vez al público en general el 24 de julio de 2015.

En febrero de 2019, Gartner confirmó a Microsoft como líder en el “Cuadrante Mágico de Gartner 2019 para Análisis y Plataforma de Inteligencia de

Negocios” como resultado de las capacidades de la plataforma Power BI. Esto representó el duodécimo año consecutivo de reconocimiento de Microsoft como proveedor líder en esta categoría cuadrante mágico (a partir de 3 años antes de que se creara esta herramienta).

### 3.1.2 Principales ventajas de Power BI

- Intuitiva y fácil de usar

Permite crear informes y paneles básicos sin conocimientos técnicos, sobre todo en la versión Desktop. Aunque a medida que se profundice en la herramienta es necesario poseer una formación para extraer el máximo provecho de esta herramienta.

- Integración perfecta con Microsoft Excel

Es posible exportar y conectar fácilmente los datos de Excel con los paneles de Power Bi; esto es una gran ventaja, ya que muchas empresas usan hojas de cálculo de Excel como herramienta de análisis de datos.

- Permite hacer múltiples análisis complejos en un solo panel

Logrando una visualización única y muy atractiva. Además, los paneles se pueden publicar y compartir con toda la organización, lo que ayuda a fomentar la cultura de análisis de datos e inteligencia empresarial en la organización.

- Se actualiza constantemente

Esto genera un análisis de datos en tiempo real, permite a las organizaciones tomar decisiones basadas en lo que ocurre en el momento.

- Incorpora herramientas de Power View y Power Map

Lo cual permite visualizar imágenes y gráficos en tres dimensiones y visualización de datos en mapas geográficos. Existe una integración con ArcGIS Maps de ESRI, la compañía líder mundial en geomarketing, lo que permite un análisis espacial avanzado. Los mapas de ArcGIS incorporan capas de información sociodemográfica y otras variables, para proporcionar contexto a los datos y relacionar la ubicación con otras variables críticas.

- Herramienta de inteligencia artificial

No solo analiza lo que ha sucedido en el pasado y lo que está ocurriendo en la actualidad dentro de la organización, además permite detectar tendencias y hacer predicciones de lo que puede pasar a futuro.

- Almacena la información en la nube de Azure

Microsoft Azure es una plataforma de cloud computing o servicio en la nube alojado en la red global de centros de datos de Microsoft. Esto permite el acceso a los datos y paneles de Power BI desde cualquier lugar y dispositivo, con total seguridad y privacidad.

- Es compatible con múltiples fuentes de datos, Power BI soporta diferentes fuentes de datos, algunas de ellas son:
  - Ficheros Excel, CSV, PDF, etc.
  - Bases de datos relacionadas como SQL Server, MySQL, Oracle, entre otras.
  - Servicios de Azure.
  - Fuentes online como Google Analytics, etc.
  - Algunos conectores a servicios Web.
- Tiene una gran comunidad

Existe una gran comunidad de expertos y usuarios de Power BI a nivel mundial, que ayudan a resolver todas las dudas en miles de artículos foros y blogs, Power BI Community.

### 3.1.3 Principales desventajas de Power BI

- No tiene la capacidad de publicar informes con todos los datos asociados, lo cual significa que algunos datos pueden quedar fuera de las visualizaciones.
- Limitaciones de fuentes de datos para la versión gratuita.
- Límites de GB para las versiones no premium.
- Interfaz más compleja, pero de gran utilidad.

### 3.1.4 Productos de Power BI

- Power BI Desktop: es una aplicación de escritorio gratuita, se puede instalar directamente en su equipo y ayuda a explorar los datos de manera profunda y avanzada.
  - Las visualizaciones creadas son guardadas de manera local.
  - Múltiples conexiones a orígenes de datos tanto locales como basados en la nube tales como Dynamics 365, Salesforce, Azure SQL DB, Excel y SharePoint.
  - Preparación de datos a través de la herramienta Power Query.
  - Uso del lenguaje natural que permite hacer preguntas a Power BI sobre sus datos.
  - Uso del lenguaje DAX, para la creación de funciones como medidas rápidas, agrupación, entre otras.
  - Recursos de aprendizaje guiado
  - Ejemplos de código de las herramientas y tecnologías de Microsoft.
- Power BI Pro: es una licencia de Power BI que posee funciones más extensas que la versión gratuita.
  - Permite la colaboración entre miembros de un equipo mediante la creación de grupos de trabajo.
  - Crear, publicar y ver paquetes de contenido organizativo: de manera que periódicamente las personas que deseas recibas alertas con información y datos de preparados.

- Control de acceso a los datos con seguridad de nivel de fila para usuarios y grupos.
- Infinidad de conexión a fuentes de datos.
- El tamaño máximo de un conjunto de datos individual es de 1GB.
- Posee un almacenamiento máximo de 10GB por usuario.
- Cuenta con servicio en la nube.
- Inserción de contenido en otras interfaces, como las de Teams, SharePoint u otras aplicaciones SaaS.
- Power BI Premium: es una extensión de Power BI Pro que ofrece la posibilidad de obtener un mejor y más fiable rendimiento.
  - Inteligencia de negocio para empresas.
  - Análisis de macrodatos, informes en la nube y en el entorno local.
  - Recursos de almacenamiento y proceso en la nube dedicados.
  - El tamaño máximo de un conjunto de datos individual es de 10GB.
  - Posee un almacenamiento máximo de 100 TB.
  - Permite Almacenar datos de Power BI en Azure Data Lake Storage Gen2.
  - Seguridad y cifrado de datos.
  - Modelado de datos basado en IA usando AutoML, Cognitive Services y Azure Machine Learning.

### 3.1.5 Compartir el trabajo realizado en Power BI

Cuando se trabaja con la versión Pro o Premium se tiene diversas formas de compartir y colaborar con las personas de la organización.

- Guardar el área de trabajo de forma local en su computadora, haciendo clic en archivo, guardar como, añadir un nombre para el archivo y finalmente clic en guardar.



Figure 3.1: Guardar informe como archivo local

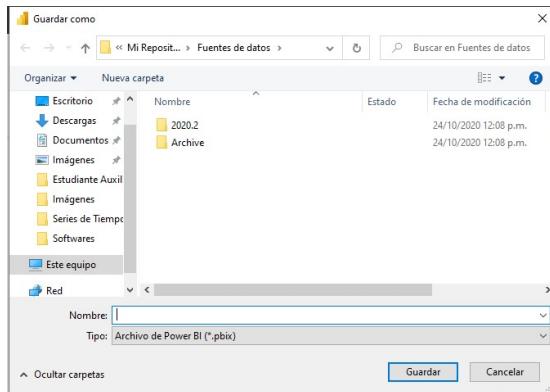


Figure 3.2: Asignar nombre al archivo para guardarlo

- Publicar informe en el servicio de Power BI, hacer clic en archivo y seleccionar publicar, se abre un cuadro de dialogo pidiendo el inicio de sesión en el servicio de Power BI.
- Insertar informes en páginas web, esto es posible hacerlo usando SharePoint Online o directamente en una página web al extraer la URL del informe.
- Imprimir o guardar en formato PDF, haciendo clic en archivo, seleccionar exportar y finalmente exportar a PDF.



Figure 3.3: Exportar informe a PDF

Al usar la versión gratuita de Power BI, es decir el producto Power BI Desktop solo es posible guardar el área de trabajo de manera local y exportar los informes como PDF.

## 3.2 Instalación de Power BI Desktop

La descarga de este software de visualización se realiza desde la aplicación de la tienda de Windows, hacer clic en instalar.



Figure 3.4: Instalar Power BI

Cuando se complete la descarga haga clic en iniciar, de esta manera ya tiene el software e su computador y se actualizara automáticamente.

## 3.3 Forma de navegación

Al momento de iniciar Power BI esta es la pantalla con la que se encuentra, aparece un cuadro de dialogo de introducción que en su panel lateral izquierdo contiene pestañas para conectarse a fuentes de datos, proyectos realizados con el software y la opción de abrir otros informes; en el panel central se muestra la opción para iniciar sesión y comprar licencias pagas de Power BI; finalmente, el panel derecho posee información sobre novedades, blogs, foros y tutoriales útiles sobre el uso de esta herramienta de visualización.



Figure 3.5: Pantalla inicial de Power BI

Cierre este cuadro de dialogo introductorio para explorar el entorno de creación, dicho entorno se verá así:

Contiene 5 espacios principales que le permitirán crear informes con diferentes

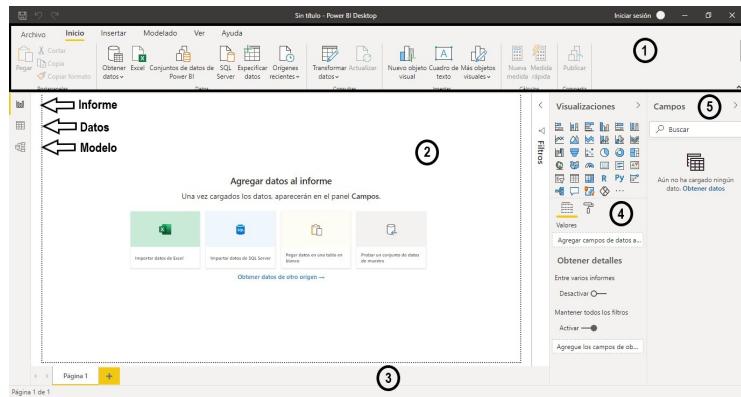


Figure 3.6: Entorno de creación

visualizaciones, el panel superior denominado 1 contiene algunos botones útiles tales como Archivo que permite guardar y exportar los informes creados, inicio que contiene herramientas asociadas a los informes como al conexión a fuentes de datos, transformación de datos, insertar elementos visuales, cuadros de texto entre otros, una sección llamada cálculos que permite crear medidas y finalmente el botón publicar. El botón insertar contiene opciones para insertar diferentes elementos, la pestaña modelado contiene herramientas para editar relaciones, cálculos, creación de parámetros, seguridad y preguntas y respuestas; la pestaña ver contiene elementos de diseño de página como temas, diseño para móvil, opciones de página y mostrar algunos paneles que se usen para crear visualizaciones; por ultimo se ubica la pestaña ayuda, dicha pestaña posee información sobre el software, aprendizaje guiado, videos tutoriales, soporte técnico, documentación, ejemplos y la comunidad de Power BI.

La sección número 2 representa la vista de informe o lienzo, este es el espacio donde se crean y organizan las visualizaciones, si hace clic en el ícono de datos se encontrará con la tabla de datos a la que se encuentra conectado, finalmente el ícono de modelo administra las relaciones existentes entre diversas fuentes de datos en el caso en que usted este conectado a varias fuentes y estas tengan alguna relación. La barra denominada área 3 es la pestaña de páginas, la cual permite navegar entre páginas y crear nuevas.

En el contenedor número 4 se ubican las visualizaciones que se pueden crear con Power BI, debajo de las opciones de visualización se ubican dos campos uno llamado valores y otro llamado formato con un ícono de rodillo, el primer campo contiene los estantes para ubicar las variables en los ejes X y Y, estos estantes cambian dependiendo de la visualización que se elija; el campo formato permite editar el color de los elementos en la visualización, título, leyendas, entre otras opciones.

Finalmente, en el panel campos se ubica el nombre de todas las variables que

contenga la base de datos, dichas variables se dividen en categóricas y numéricas, las categóricas no tienen icono asociado, mientras que las variables numéricas tienen asociadas un ícono de  $\Sigma$ , como se muestra a continuación.

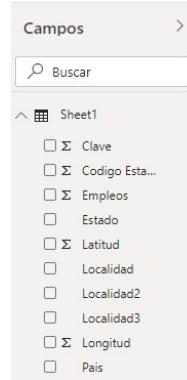


Figure 3.7: Campos

Observe entonces que los campos Estado, Localidad, Localidad2, Localidad3 y país con categóricos, mientras que los demás campos que contiene esa fuente son numéricos y se les asigna el ícono mencionado anteriormente.

## 3.4 Flujo de trabajo

### 3.4.1 Conexión a fuentes de datos

Antes de iniciar con la creación de visualizaciones y el análisis de datos debe conectar Power BI a estos, este software de visualización permite conectarse a muchos tipos de fuentes de datos, incluidas bases de datos locales, libros de trabajo de Microsoft Excel y servicios en la nube, puede explorar todas las opciones de conexión haciendo clic en obtener datos en la pantalla introductoria de Power BI.

En este caso la conexión a datos se hará a través de un archivo Excel, inicialmente se hará la conexión a las bases de datos de estudiantes graduados a nivel de microdatos para explorar algunas funcionalidades del panel Power Query.

1. Abrir Power BI y hacer clic en obtener datos.
2. En el cuadro de diálogo Obtener datos debe seleccionar Excel y hacer clic en conectar, al hacer esto se abre un cuadro de diálogo que permite navegar a través de las carpetas de su equipo para ubicar la localización de las bases de datos. Debe seleccionar una de las bases y dar clic en botón “Abrir”.

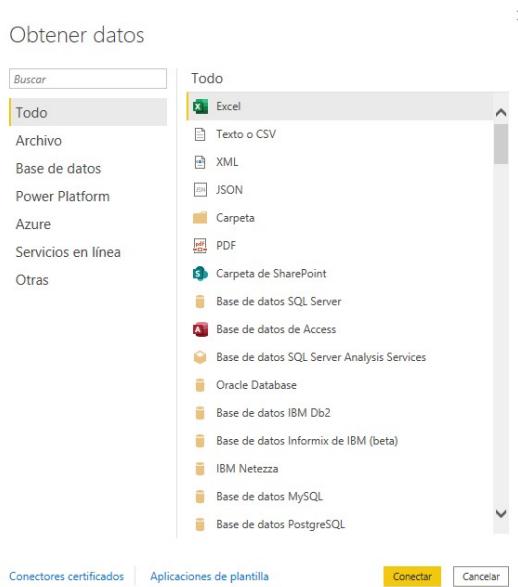


Figure 3.8: Obtener datos

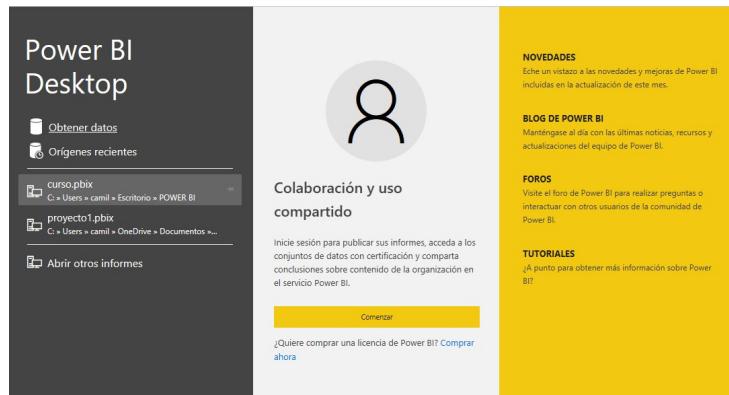


Figure 3.9: Pantalla inicial

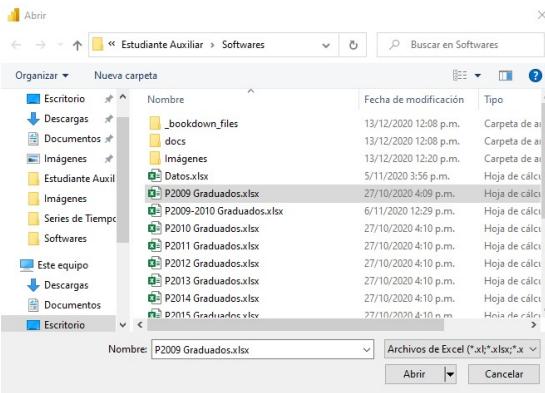


Figure 3.10: Navegación entre carpetas

3. Completando el paso anterior se abre un nuevo cuadro de dialogo llamado Navegar que en su panel izquierdo se ubica una lista de las hojas de trabajo que componen el archivo Excel, debe seleccionar una y luego hacer clic en Cargar.
  4. Después de cargados los datos haga clic en Transformar datos, ubicado en la barra de herramientas, obtendrá una nueva ventana llamada Editor de Power Query que permite realizar algunas transformaciones sobre la base de datos.

Este editor en su barra de herramientas contiene botones útiles que permiten interactuar con los datos, como transformaciones, creación de nuevas columnas, herramientas y ayuda. En el panel lateral izquierdo encontrara una lista de las tablas a las que se encuentra conectado, en este caso la hoja del archivo Excel a la cual se realizó la conexión es “P2009G”, en la parte central del editor se ubica una vista previa de los datos, por último, el panel Configuración de la consulta muestra las propiedades y los pasos aplicados para transformar la base de datos.

Observe que cada columna tiene asociado un icono que indica si es una variable numérica o de texto, puede editar este tipo de datos haciendo clic sobre el icono y seleccionando el tipo de dato, por ejemplo, la columna semestre debe cambiarse de numérica a texto.

Al hacer clic en el menú desplegable de cada variable encontrara información sobre los datos que contiene, opciones de orden como ascendente y descendente, eliminar vacíos y añadir filtros de texto, muy similar a lo encontrado en Excel cuando se quiere aplicar filtros a los datos. En la pestaña filtros de texto se tiene diversas opciones para filtrar los datos que permiten especificar si es igual o no, el inicio o final, entre otras opciones.

A modo de ejemplo puede seleccionar únicamente Antioquia y obtendrá una

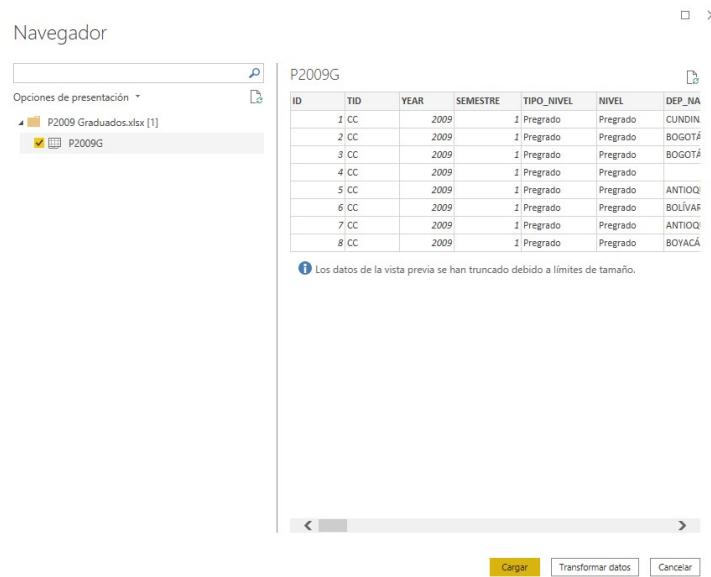


Figure 3.11: Navegador

The screenshot shows the Power Query Editor interface with the 'P2009G' query selected. The main area displays the same 8 rows of student data. To the right, there are several toolbars: 'Archivo', 'Nuevo', 'Transformar', 'Agregar columna', 'Vista', 'Herramientas', 'Ayuda', 'Consultas', 'Nuevo consulta', 'Configuración de origen de datos', 'Administrador de paquetes', 'Propiedades', 'Actualizar vista previa', 'Administrador', 'Dividir columnas', 'Agrupar columnas', 'Reemplazar los valores', 'Ordenar', 'Transformar', 'Tipos de datos flotante entero', 'Usar la primera fila como encabezado', and 'Combinar'. A 'PROPIEDADES' pane shows 'Nombre: P2009G' and a 'PASOS APLICADOS' pane showing 'Tipo cambiado'. The status bar at the bottom indicates '57 COLUMNAS, 999+ FILAS' and 'Generación de perfiles de columnas basada en las 1000 primeras filas'.

Figure 3.12: Panel Power Query

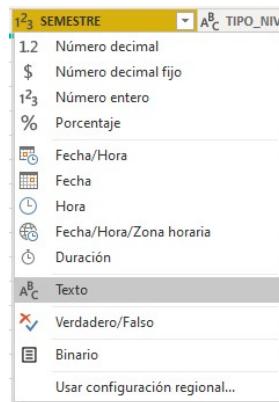


Figure 3.13: Cambiar tipo de dato

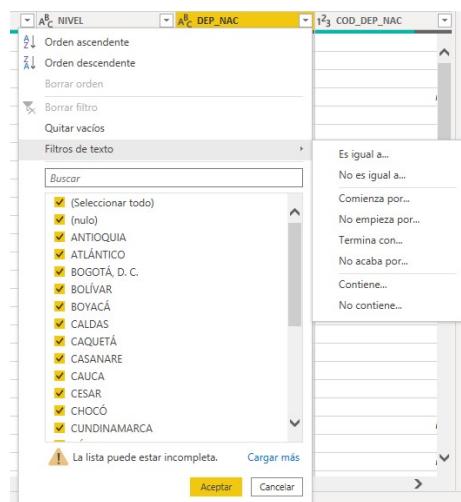


Figure 3.14: Menú de cada variable

base de datos que solo contiene las observaciones en las cuales se cumple el filtro aplicado, es decir donde DEP\_NAC sea Antioquia.

Haciendo clic derecho sobre el nombre de una variable se despliega un menú con diversas opciones, entre ellas copiar, quitar, duplicar, reemplazar, dividir, agrupar, entre otras; como se muestra en la figura 3.15.

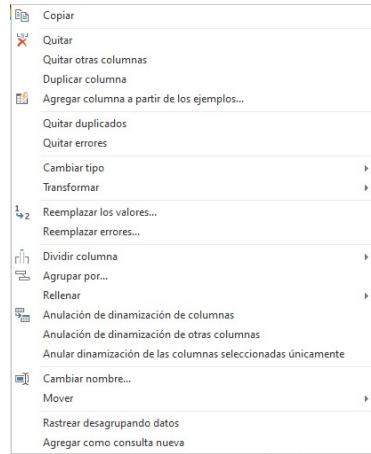


Figure 3.15: Opciones de cada variable

Al final de la base de datos se agregaron dos columnas llamadas columna56 y columna57 respectivamente, es necesario eliminarlas ya que no hacen parte del conjunto de variables a analizar, para esto haga clic derecho sobre el nombre de la columna y seleccione quitar.

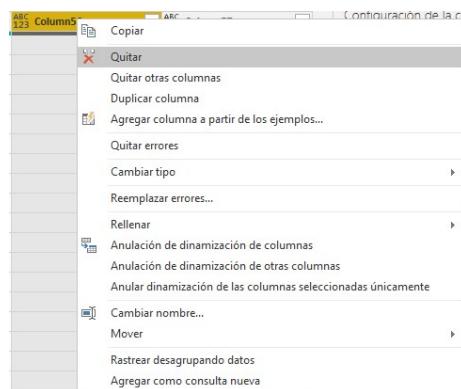


Figure 3.16: Eliminar columnas

La opción dividir columna presenta diferentes alternativas para realizar divisiones de columna tales como división por un delimitador, número de

caracteres, posiciones y algunas opciones de texto como mayúsculas, minúsculas y dígitos; tal y como se muestra en la figura 3.17.

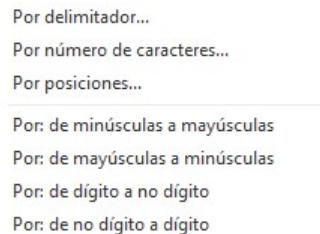


Figure 3.17: Alternativas para dividir columnas

La siguiente columna que presenta problemas es la llamada CIU\_NAC puesto que presenta una combinación de números y texto, para solucionar esto se debe usar la opción de división de dígito a no dígito.

- Haga clic derecho sobre la columna, en el menú desplegable seleccione dividir columna y de dígito a no dígito.

Figure 3.18: Dividir columnas usando de dígito a no dígito

Con esto se obtienen dos columnas, una de ellas contiene el número que en este caso corresponde al código del departamento y la otra variable contiene el nombre, tal y como se muestra en la figura 3.19.

Un detalle importante es que el primer registro que pertenece a la ciudad de Sopó no contenía números por lo tanto para esta fila no se hizo la división.

Cuando se tiene un separador en la columna es útil usar la división de columna por delimitador, en este caso la variable PROGRAMA\_S se compone del nombre del programa y la sede a la que pertenece, estos dos atributos se

	t <sub>3</sub> COD_DEP_NAC	A <sub>C</sub> CIU_NAC.1	A <sub>C</sub> CIU_NAC.2	t <sub>3</sub> COD_CIU_NAC	t <sub>2</sub> LON_CIU_NAC
1		25 SOPÓ		null	25758 -73,9388529
2		11 11	BOGOTÁ D.C.	11001	-74,11391872
3		11 11	BOGOTÁ D.C.	11001	-74,11391872
4	null	null	null	null	null
5		5 05	MEDELLIN	5001	-75,57705563
6		13 13	CARTAGENA	13001	-75,50122922
7		5 05	MEDELLIN	5001	-75,57705563
8		15 15	TUNJA	15001	-73,35436382
9		5 05	MEDELLIN	5001	-75,57705563
10		5 05	SALGAR	5642	-75,97543596
11		5 05	MEDELLIN	5001	-75,57705563
12		5 05	MEDELLIN	5001	-75,57705563
13		11 11	BOGOTÁ D.C.	11001	-74,11391872
14		15 15	MIRAFLORES	15455	-73,14486261
15		5 05	ARMENIA	5059	-75,78550446
16		73 73	CAJAMARCA	73124	-75,43379586
17		66 66	PEREIRA	66001	-75,71509992
18		66 66	PEREIRA	66001	-75,71509992
19		66 66	PEREIRA	66001	-75,71509992
20		11 11	BOGOTÁ D.C.	11001	-74,11391872
21		66 66	SANTA ROSA DE CABAL	66682	-75,62407439
22	null	null	null	null	null
23		11 11	BOGOTÁ D.C.	11001	-74,11391872
24	null	null	null	null	null

Figure 3.19: Vista previa de la columna Ciudad dividida

encuentran separados por un guion el cual se usara como delimitador para realizar la división.

- Haga clic derecho sobre la columna de interés, seleccione división y por delimitador.

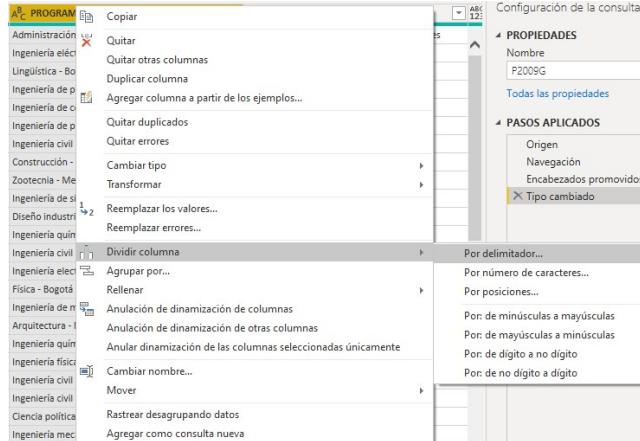


Figure 3.20: Dividir columnas usando un delimitador

- En la ventana emergente llamada dividir columna por delimitador seleccione personalizado y escriba un – en el siguiente espacio, por último, en la sección dividir en seleccione cada aparición del delimitador.

Al seleccionar cada aparición del delimitador se obtienen dos columnas, una de ellas contiene el programa y la otra la sede a la que pertenece.

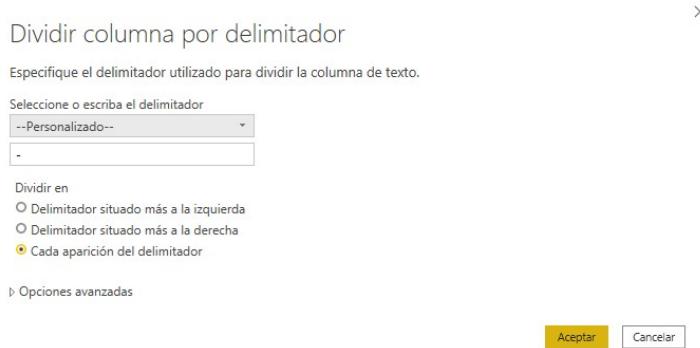


Figure 3.21: Configuración del delimitador

	A <sub>C</sub> PROGRAMA	A <sub>C</sub> PROGRAMA_S.1	A <sub>C</sub> PROGRAMA_S.2	A <sub>C</sub> AREAC_SNIES
1	19 Administración de empresas	Administración de empresas	Bogotá	Economía, administración, contaduría y afines
2	27 Ingeniería eléctrica	Ingeniería eléctrica	Bogotá	Ingeniería, arquitectura, urbanismo y afines
3	6930 Lingüística	Lingüística	Bogotá	Ciencias sociales y humanas
4	119 Ingeniería de petróleos	Ingeniería de petróleos	Medellín	Ingeniería, arquitectura, urbanismo y afines
5	6921 Ingeniería de control	Ingeniería de control	Medellín	Ingeniería, arquitectura, urbanismo y afines
6	119 Ingeniería de petróleos	Ingeniería de petróleos	Medellín	Ingeniería, arquitectura, urbanismo y afines
7	117 Ingeniería civil	Ingeniería civil	Medellín	Ingeniería, arquitectura, urbanismo y afines
8	120 Construcción	Construcción	Medellín	Ingeniería, arquitectura, urbanismo y afines
9	6926 Zootecnia	Zootecnia	Medellín	Agronomía, veterinaria y afines
10	6925 Ingeniería de sistemas e informática	Ingeniería de sistemas e informática	Medellín	Ingeniería, arquitectura, urbanismo y afines
11	5 Diseño industrial	Diseño industrial	Bogotá	Bellas artes
12	120 Ingeniería química	Ingeniería química	Medellín	Ingeniería, arquitectura, urbanismo y afines
13	25 Ingeniería civil	Ingeniería civil	Bogotá	Ingeniería, arquitectura, urbanismo y afines
14	6941 Ingeniería electrónica	Ingeniería electrónica	Bogotá	Ingeniería, arquitectura, urbanismo y afines
15	33 Física	Física	Bogotá	Matemáticas y ciencias naturales
16	118 Ingeniería de minas y metalurgia	Ingeniería de minas y metalurgia	Medellín	Ingeniería, arquitectura, urbanismo y afines
17	4126 Arquitectura	Arquitectura	Manizales	Ingeniería, arquitectura, urbanismo y afines
18	4125 Ingeniería química	Ingeniería química	Manizales	Ingeniería, arquitectura, urbanismo y afines
19	6915 Ingeniería física	Ingeniería física	Manizales	Ingeniería, arquitectura, urbanismo y afines
20	4121 Ingeniería civil	Ingeniería civil	Manizales	Ingeniería, arquitectura, urbanismo y afines
21	4121 Ingeniería civil	Ingeniería civil	Manizales	Ingeniería, arquitectura, urbanismo y afines
22	3140 Ciencia política	Ciencia política	Bogotá	Ciencias sociales y humanas
23	26 Ingeniería mecánica	Ingeniería mecánica	Bogotá	Ingeniería, arquitectura, urbanismo y afines
24	25 Ingeniería civil	Ingeniería civil	Bogotá	Ingeniería, arquitectura, urbanismo y afines
25				

Figure 3.22: Columna programa- sede dividida

La creación de grupos debe realizarse usando la pestaña agregar columnas y seleccionando columna condicional.

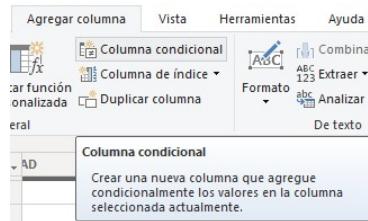


Figure 3.23: Crear columna condicional

Como muestra la descripción de esta opción lo que permite es crear una nueva columna que agrega los vaolres de manera condicional de una columna que se seleccione, en este caso la variable de interés es EDAD\_MOD, se crearan grupos para categorizar las edades. En la ventana emergente llamada agregar columna condicional se deben especificar las condiciones por las cuales se quiere agrupar, en ese caso debe llenar los campos como se muestra en la figura 3.24.

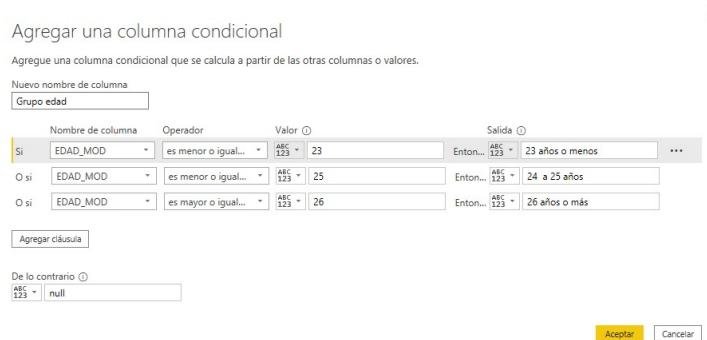


Figure 3.24: Agregar columna condicional

Al finalizar esta operación se crea una nueva columna llamada Grupo edad que contiene los grupos creados para la variable edad, dichos grupos fueron asignados de manera correcta y contiene algunos errores debido a que existen valores faltantes en la columna agrupada.

Otra funcionalidad importante que presenta el editor Power Query es la unión de bases de datos, estas alternativas se ubican en la barra de herramientas y el botón Combinar, las opciones disponibles son combinar consultas y anexar consultas, la primera permite hacer uniones por columna y la segunda permite hacer uniones por filas.

Las uniones por columna son útiles cuando se quiere trabajar con dos

ALIDAD	EDAD_MOD	CAT_EDAD	ABC Grupo edad	SEXO	ESTRATO_ORIG
1 null		24 24 a 25 años	24 a 25 años	Hombres	Estrato 2
2 null		26 26 o más años	26 años o más	Hombres	Estrato 2
3 null		27 26 o más años	26 años o más	Hombres	Estrato 2
4 null		null	null Error	Hombres	Estrato 4
5 null		25 24 a 25 años	24 a 25 años	Hombres	Estrato 3
6 null		24 24 a 25 años	24 a 25 años	Hombres	Estrato 2
7 null		28 26 o más años	26 años o más	Hombres	Estrato 3
8 null		28 26 o más años	26 años o más	Hombres	Estrato 3
9 null		27 26 o más años	26 años o más	Hombres	Estrato 6
10 null		27 26 o más años	26 años o más	Hombres	Estrato 3
11 null		29 26 o más años	26 años o más	Hombres	Estrato 4
12 null		28 26 o más años	26 años o más	Hombres	Estrato 3
13 null		29 26 o más años	26 años o más	Hombres	Estrato 2
14 null		41 26 o más años	26 años o más	Hombres	Estrato 2
15 null		27 26 o más años	26 años o más	Hombres	Estrato 2
16 null		28 26 o más años	26 años o más	Hombres	Estrato 2
17 null		25 24 a 25 años	24 a 25 años	Hombres	ND/NE
18 null		25 24 a 25 años	24 a 25 años	Hombres	ND/NE
19 null		25 24 a 25 años	24 a 25 años	Hombres	Estrato 2
20 null		24 24 a 25 años	24 a 25 años	Hombres	ND/NE
21 null		24 24 a 25 años	24 a 25 años	Hombres	ND/NE
22 null		null	null Error	Hombres	Estrato 1
23 null		29 26 o más años	26 años o más	Hombres	ND/NE
24 null		32 26 o más años	26 años o más	Hombres	ND/NE

Figure 3.25: Columna grupo edad

columnas que se encuentran en diferentes conjuntos de datos, existen seis opciones de realizar las uniones por columnas:

- Externa izquierda: devuelve todos los registros de la tabla izquierda y solo los registros que coinciden con la tabla de la derecha
- Externa derecha: devuelve todos los registros de la tabla derecha y solo los registros que coinciden con la tabla de la izquierda.
- Externa completa: devuelve todos los registros de ambas tablas.
- Interna: devuelve únicamente los registros que están en ambas tablas.
- Anti izquierda: devuelve únicamente los registros de la tabla izquierda.
- Anti derecha: devuelve únicamente los registros de la tabla derecha.

Los conjuntos de datos que se quieren analizar contienen las mismas columnas por lo que el interés se centra en realizar uniones por filas, se debe realizar la conexión a todas las fuentes de datos usando el botón nuevo origen, seleccionar Excel y cargar los datos, en el panel lateral izquierdo debe tener la lista de tablas a las que se conectó, como se muestra en la figura 3.26.

P20096	TÍD	ABC TIPO	YEAR	SEMESTRE	AÑO NIVEL
P20105	1	1 CC	2020	I Pregreso	Máster
P20116	2	2 CC	2020	I Pregreso	Máster
P20120	3	3 CC	2020	I Pregreso	Máster
P20130	4	4 CC	2020	I Pregreso	Diplomado
P20140	5	5 CC	2020	I Pregreso	Especialista
P20150	6	6 CC	2020	I Pregreso	Especialista
P20160	7	7 CC	2020	I Pregreso	Especialista
P20170	8	8 CC	2020	I Pregreso	Pregreso
P20180	9	9 CC	2020	I Pregreso	Pregreso
P20190	10	10 CC	2020	I Pregreso	Pregreso
P201910	11	11 CC	2020	I Pregreso	Pregreso
P202020	12	12 CC	2020	I Pregreso	Pregreso
P202025	13	13 CC	2020	I Pregreso	Pregreso
P202025	14	14 CC	2020	I Pregreso	Pregreso
P202025	15	15 CC	2020	I Pregreso	Pregreso
P202016	16	16 CC	2020	I Pregreso	Pregreso
	17	17 CC	2020	I Pregreso	Pregreso
	18	18 CC	2020	I Pregreso	Pregreso
	19	19 CC	2020	I Pregreso	Pregreso
	20	20 CC	2020	I Pregreso	Pregreso
	21	21 CC	2020	I Pregreso	Pregreso
	22	22 CC	2020	I Pregreso	Pregreso
	23	23 CC	2020	I Pregreso	Pregreso
	24	24 CC	2020	I Pregreso	Pregreso
	25				

Figure 3.26: Lista de tablas

Ahora para realizar la unión de todas las tablas debe:

- Hacer clic en el botón combinar ubicado en la barra de herramientas hacia el lado derecho, seleccionar anexar consultas y anexar consultas para crear una nueva.

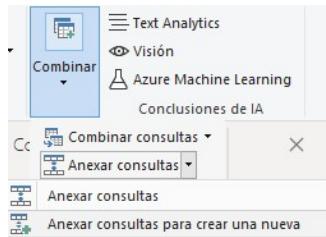


Figure 3.27: Combinar tablas de datos

- En la ventana Anexar seleccione tres o mas tablas, en el panel tablas disponibles haga clic sobre cada tabla y clic sobre el botón agregar, por ultimo haga clic en el botón aceptar para guardar los campos.



Figure 3.28: Anexar datos

En el panel Consultas obtendrá una nueva tabla llamada Anexar1 la cual posee la unión de todas las tablas a las que se tienen acceso, dicha tabla contiene 101.841 registros y 55 columnas.

Como se menciono en la de sección de Tableau los datos requieren una limpieza y preparación la cual se realizo desde R, dentro de los orígenes de datos disponibles en Power BI existe la opción de cargar scripts de R, por lo tanto, se usará esta forma de conexión a datos para cargar el conjunto de datos limpio y listo para realizar análisis.

Inicialmente es necesario verificar que Power BI se encuentre conectado a la versión en la que fue creado el script de R que se desea ejecutar, para esto debe:

1. Abrir Power BI Desktop, cerrar la ventana introductoria, seguido de esto hacer clic en archivo, seleccionar opciones y configuración y nuevamente clic en opciones.

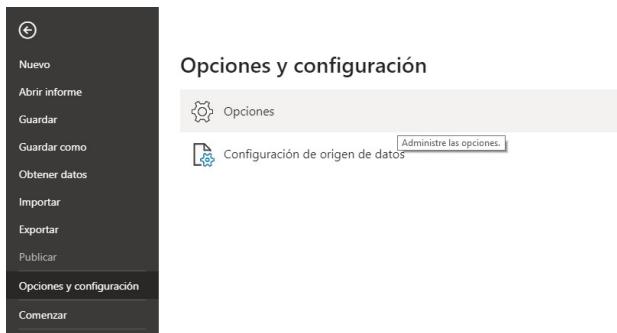


Figure 3.29: Opciones y configuración

2. Se abrirá una ventana emergente con las opciones de configuración disponibles, en el panel global seleccione scripts de R, luego de esto en el panel lateral derecho asegúrese de que el directorio se ubique en la versión en la cual fue creada el script y como IDE de R se ubique R Studio.

Cuando ya se han realizado las configuraciones necesarias haga clic en obtener datos y seleccione otras, script de R y clic en conectar.

Después de completado el paso anterior se abrirá una ventana llamada Script de R en la cual debe escribir su código R, dicho código puede ser copiado del archivo llamado Datacleaning.R ubicado en el repositorio de trabajo, debe tener cuidado de no copiar la última línea ya que en esa se extrae el archivo en CSV y no es necesario hacerlo ya que la tabla queda guardada en el entorno R.

Al hacer clic en aceptar para guardar los cambios Power BI iniciara la conexión con R para ejecutar las líneas de código escritas, nuevamente se abre una ventana llamada navegar que incluye todas las tablas almacenadas en el ambiente R, la que se debe usar es la llamada Graduados que contiene la unión de todas las tablas ordenadas y preparadas para el análisis, haga clic sobre esta tabla y finalmente clic en cargar.

Cuando se han cargado los datos en el panel campos se ubicarán todas las variables que contiene el conjunto de datos al que se encuentra conectado, haga clic en transformar datos con el fin de verificar el tipo de cada variable y hacer los cambios que sean necesarios, luego de verificar y realizar los cambios necesarios es momento de iniciar con las visualizaciones.

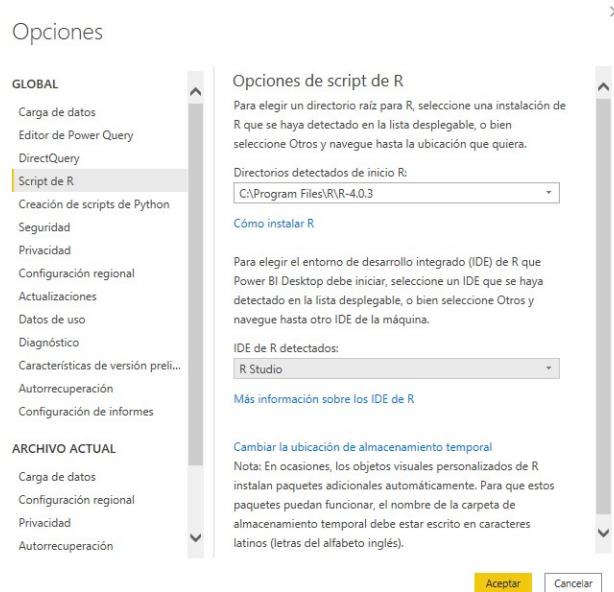


Figure 3.30: Opciones de script de R

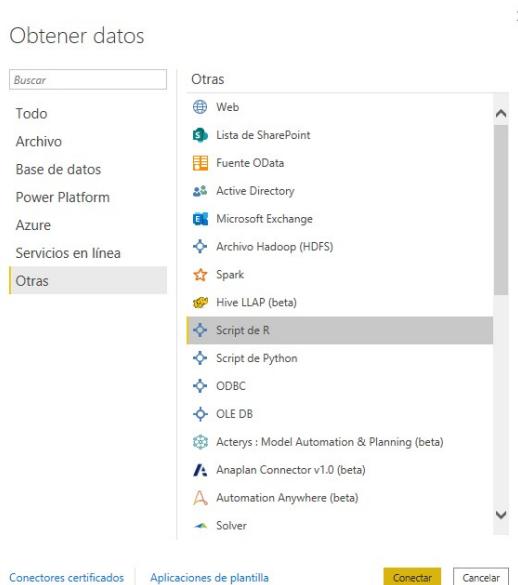


Figure 3.31: Obtener datos de script de R

Script de R

```

Script
"Salud Y Seguridad En El Trabajo")
Graduados$PROGRAMA <- str_replace_all(Graduados$PROGRAMA,
  "Seguridad Alimentaria(\sY\s)?(N|n)utricional",
  "Seguridad Alimentaria Y Nutricional")
Graduados$PROGRAMA <- str_replace_all(Graduados$PROGRAMA,
  "Salud(\sY\s)?(P|p)roducción Animal",
  "Salud Y Producción Animal")
Graduados$PROGRAMA <- str_replace_all(Graduados$PROGRAMA,
  "Vias(\sY\s)?(T|t)ransportes?",
  "Vías Y Transportes")
Graduados$YEAR_SEMESTER <- str_c(Graduados$YEAR, Graduados$SEMESTRE,
  sep = " - ")

```

El script se ejecutará con la instalación de R siguiente: C:\Program Files\R\R-4.0.3.  
Para establecer la configuración y cambiar la instalación de R que quiere ejecutar, vaya a Opciones y configuración.

Aceptar Cancelar

Figure 3.32: Script de R

Navegador

Graduados					
YEAR	SEMESTRE	TIPO_NIVEL	NIVEL	DEP_NAC	CIU_NAC
2009	1	Pregado	Pregrado	Cundinamarca	Sopo
2009	1	Pregado	Pregrado	Bogotá D.C.	Bogotá d.c.
2009	1	Pregado	Pregrado	Bogotá D.C.	Bogotá d.c.
2009	1	Pregado	Pregrado		null
2009	1	Pregado	Pregrado	Antioquia	Medellín
2009	1	Pregado	Pregrado	Bolívar	Cartagena
2009	1	Pregado	Pregrado	Antioquia	Medellín
2009	1	Pregado	Pregrado	Boyacá	Tunja
2009	1	Pregado	Pregrado	Antioquia	Medellín
2009	1	Pregado	Pregrado	Antioquia	Saigar
2009	1	Pregado	Pregrado	Antioquia	Medellín
2009	1	Pregado	Pregrado	Antioquia	Medellín
2009	1	Pregado	Pregrado	Bogotá D.C.	Bogotá d.c.
2009	1	Pregado	Pregrado	Boyacá	Miraflores
2009	1	Pregado	Pregrado	Antioquia	Armenia

Los datos de la vista previa se han truncado debido a límites de tamaño.

Cargar Transformar datos Cancelar

Figure 3.33: Navegador

# **Chapter 4**

## **Flourish**

Some *significant* applications are demonstrated in this chapter.

### **4.1 Example one**

### **4.2 Example two**



# **Chapter 5**

# **Conclusiones**

We have finished a nice book.



# Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2020). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.18.