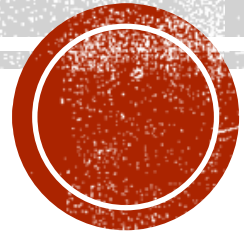
TABLEAU - MANISHA

YOUR DATA!



CareerNow
LEARNING

IMPORTANT TASKS

Microsoft Excel
ma Separated Valu

Dataset: **Spotify.csv**

ABOUT DATA

Description:

This dataset contains a comprehensive list of the most famous songs of 2023 as listed on Spotify. The dataset offers a wealth of features beyond what is typically available in similar datasets. It provides insights into each song's attributes, popularity, and presence on various music platforms. The dataset includes information such as track name, artist(s) name, release date, Spotify playlists and charts, streaming statistics, Apple Music presence, Deezer presence, Shazam charts, and various audio features.

ABOUT DATA

Key Features:

- ✓ track_name: Name of the song
- ✓ artist(s)_name: Name of the artist(s) of the song
- ✓ artist_count: Number of artists contributing to the song
- ✓ released_year: Year when the song was released
- ✓ released_month: Month when the song was released
- ✓ released_day: Day of the month when the song was released
- ✓ in_spotify_playlists: Number of Spotify playlists the song is included in
- ✓ in_spotify_charts: Presence and rank of the song on Spotify charts
- ✓ streams: Total number of streams on Spotify
- ✓ in_apple_playlists: Number of Apple Music playlists the song is included in
- ✓ in_apple_charts: Presence and rank of the song on Apple Music charts
- ✓ in_deezer_playlists: Number of Deezer playlists the song is included in
- ✓ in_deezer_charts: Presence and rank of the song on Deezer charts

ABOUT DATA

Key Features:

- ✓ in_shazam_charts: Presence and rank of the song on Shazam charts
- ✓ bpm: Beats per minute, a measure of song tempo
- ✓ key: Key of the song
- ✓ mode: Mode of the song (major or minor)
- ✓ danceability_%: Percentage indicating how suitable the song is for dancing
- ✓ valence_%: Positivity of the song's musical content
- ✓ energy_%: Perceived energy level of the song
- ✓ acousticness_%: Amount of acoustic sound in the song
- ✓ instrumentalness_%: Amount of instrumental content in the song
- ✓ liveness_%: Presence of live performance elements
- ✓ speechiness_%: Amount of spoken words in the song
- ✓ Potential Use Cases:
- ✓ Music analysis: Explore patterns in audio features to understand trends and preferences in popular songs.
- ✓ Platform comparison: Compare the song's popularity across different music platforms.
- ✓ Artist impact: Analyze how artist involvement and attributes relate to a song's success.
- ✓ Temporal trends: Identify any shifts in music attributes and preferences over time.
- ✓ Cross-platform presence: Investigate how songs perform across different streaming services.

FAMILIARIZE WITH THE TABLEAU INTERFACE AND TERMINOLOGY

What are the key components of the Tableau interface, and how do they assist in data visualization?

Answer: The key components of the Tableau interface include the Data Source tab for data connection, the Data Pane for field management, the Worksheet where you create visualizations, and the Dashboard for combining multiple visualizations. Tableau's interface assists in data visualization by providing a user-friendly platform for data exploration, transformation, and the creation of various charts and dashboards.

CONNECT TABLEAU TO DIFFERENT DATA SOURCES

How can you connect Tableau to the Spotify dataset, which is available in Excel format?

Answer: To connect Tableau to the Spotify dataset in Excel format, follow these steps:

a. Open Tableau and select "Connect to Data."

b. Choose "Microsoft Excel" as the data source.

c. Locate and select the Spotify dataset Excel file.

d. In the Data Source tab, you can preview, clean, and transform your data before using it in visualizations.

Explain the steps to import data from a text file (e.g., CSV) into Tableau for analysis.

Answer: To import data from a CSV file into Tableau:

a. Open Tableau and select "Connect to Data."

b. Choose "Text File" as the data source.

c. Locate and select the CSV file containing your data.

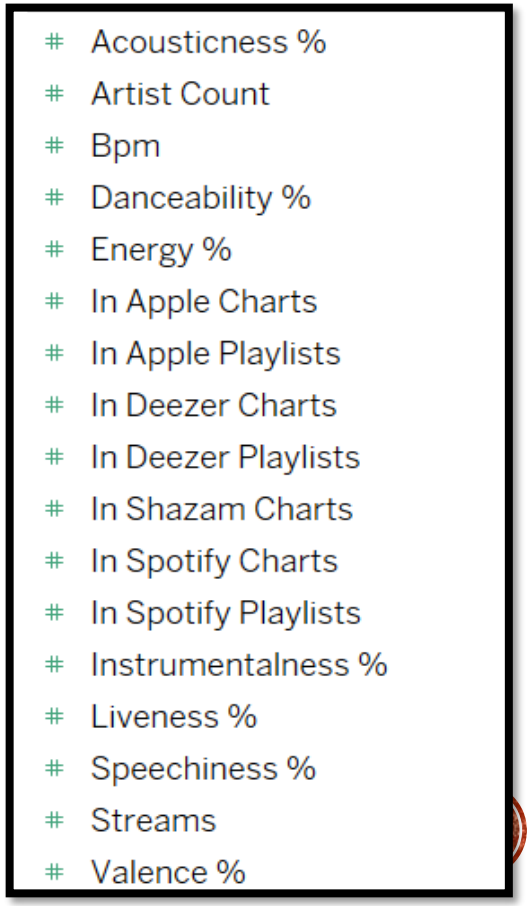
d. Follow the instructions to specify the delimiter and format settings.

e. Tableau will load the data, allowing you to proceed with analysis.

LEARN ABOUT DIMENSIONS AND MEASURES, AGGREGATION OPTIONS, CUSTOM SHAPES, AND FILTERING OPTIONS IN TABLEAU

Provide definitions and examples of dimensions and measures using columns from the Spotify dataset.

Answer: In the Spotify dataset, dimensions are categorical or qualitative data, while measures are numerical or quantitative data. For example, "artist_name" is a dimension, and "streams" is a measure.



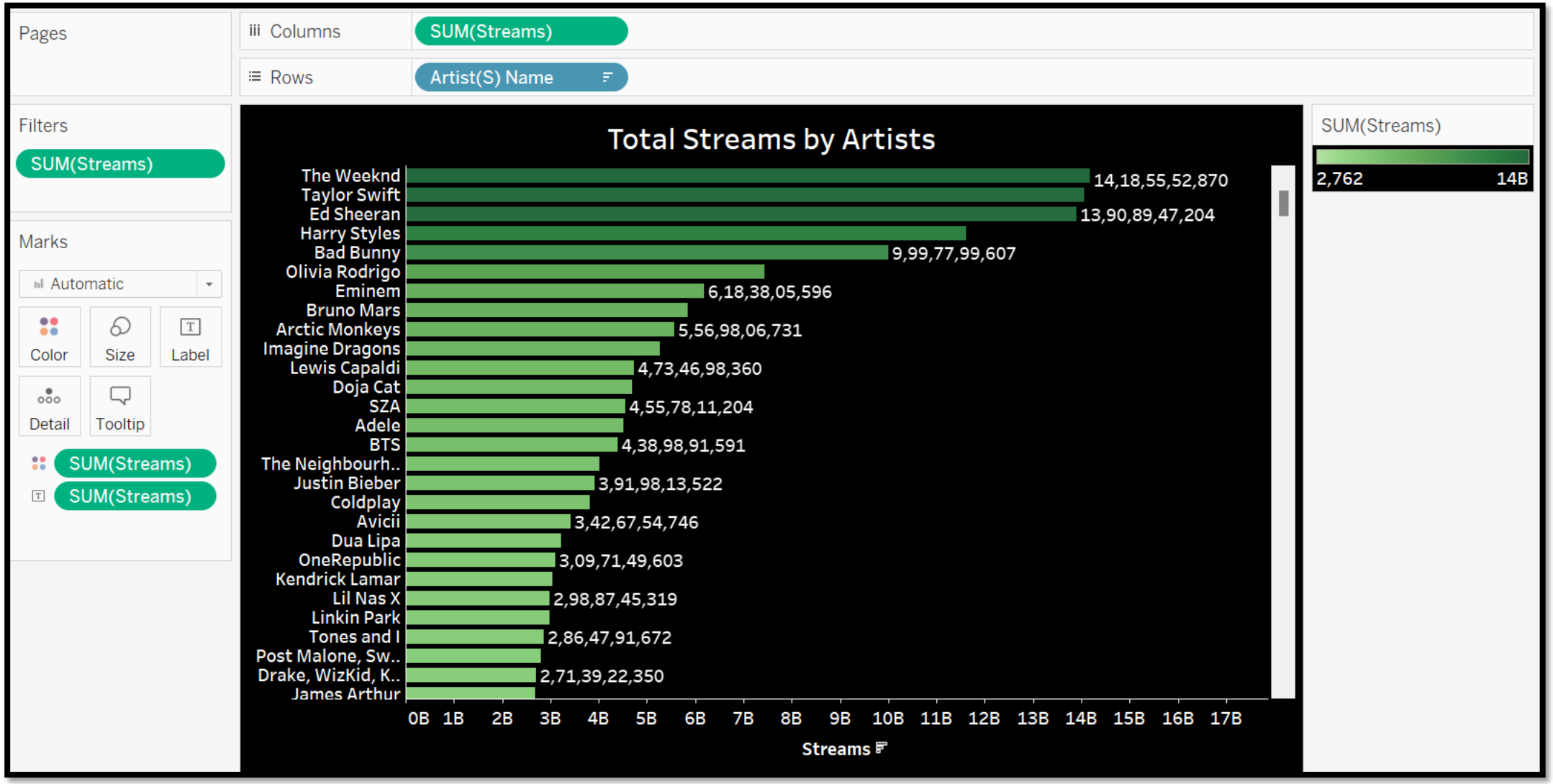
How can you aggregate data in Tableau, and why is it important when creating visualizations?

Answer: You can aggregate data in Tableau by using functions like SUM, AVG, COUNT, etc. Aggregation is crucial in visualizations because it summarizes data, providing insights and patterns. For instance, you can use SUM(streams) to see the total streams per artist.

Steps:

1. On the Worksheet, drag the following fields to the appropriate shelves:
    1. Drag "artist(s)_name" to the Rows shelf.
    2. Drag "streams" to the Columns shelf.
2. Your worksheet should now have rows for each artist's name and a column for the total streams.
3. By default, Tableau may have already applied the SUM aggregation to the "streams" field. If not, you can explicitly apply the SUM aggregation by doing the following:
    1. Click the dropdown arrow next to "streams" in the Columns shelf.
    2. Select "Measure" and then choose "Sum."
4. Now, you should see a bar chart with artists on the x-axis and the total streams on the y-axis.
5. To sort the artists by total streams in descending order, you can click on the "Total Streams" axis and choose "Sort" > "Descending."

How can you aggregate data in Tableau, and why is it important when creating visualizations?



Describe how to use custom shapes in Tableau to represent different data points uniquely.

Answer: Custom shapes in Tableau allow you to use custom images or icons instead of default marks. To use custom shapes, you can assign a field (e.g., "artist_name") to the Shape shelf, then import custom shape images for each unique value.

Here are the steps:

- **Prepare your custom shape images:**
  - Before you start, make sure you have custom shape images that you want to use for each unique data point. These images could be in PNG, JPEG, or other common image formats.
  - It's a good practice to name your image files to correspond to the unique values in your dataset. For this example, let's assume you have custom images named after artists in your dataset.
- **Open Tableau and connect to your dataset:**
  - Launch Tableau and connect to your Spotify dataset as you normally would.
- **Create a new worksheet:**
  - Click on the "Worksheet" tab to start creating a new worksheet.
- **Assign a dimension to the Shape shelf:**
  - Drag the dimension you want to use for custom shapes (e.g., "artist_name") to the "Shape" shelf, which is usually located just below the "Rows" and "Columns" shelves.
- **Edit Shape Palette:**

Describe how to use custom shapes in Tableau to represent different data points uniquely.

- In the "Shape" shelf, you will see a list of shapes that correspond to the unique values in your dimension (e.g., artists). Initially, these shapes will be generic.
- Click the dropdown arrow next to the dimension name in the "Shape" shelf, and then select "Edit Shape."

- **Assign custom shapes to unique values:**
  - In the "Edit Shape" dialog, you can manually assign custom images to each unique value in your dimension. For each artist (unique value), click on the shape icon and select the custom image file corresponding to that artist.

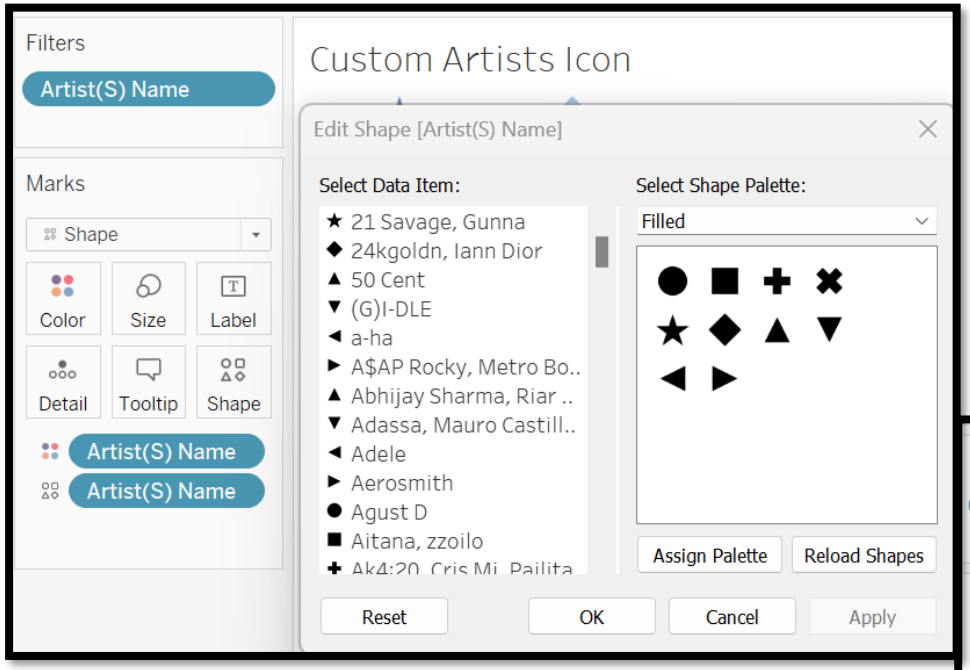
- **Apply the custom shapes:**
  - Once you've assigned custom shapes to the unique values, click "OK" to close the "Edit Shape" dialog.

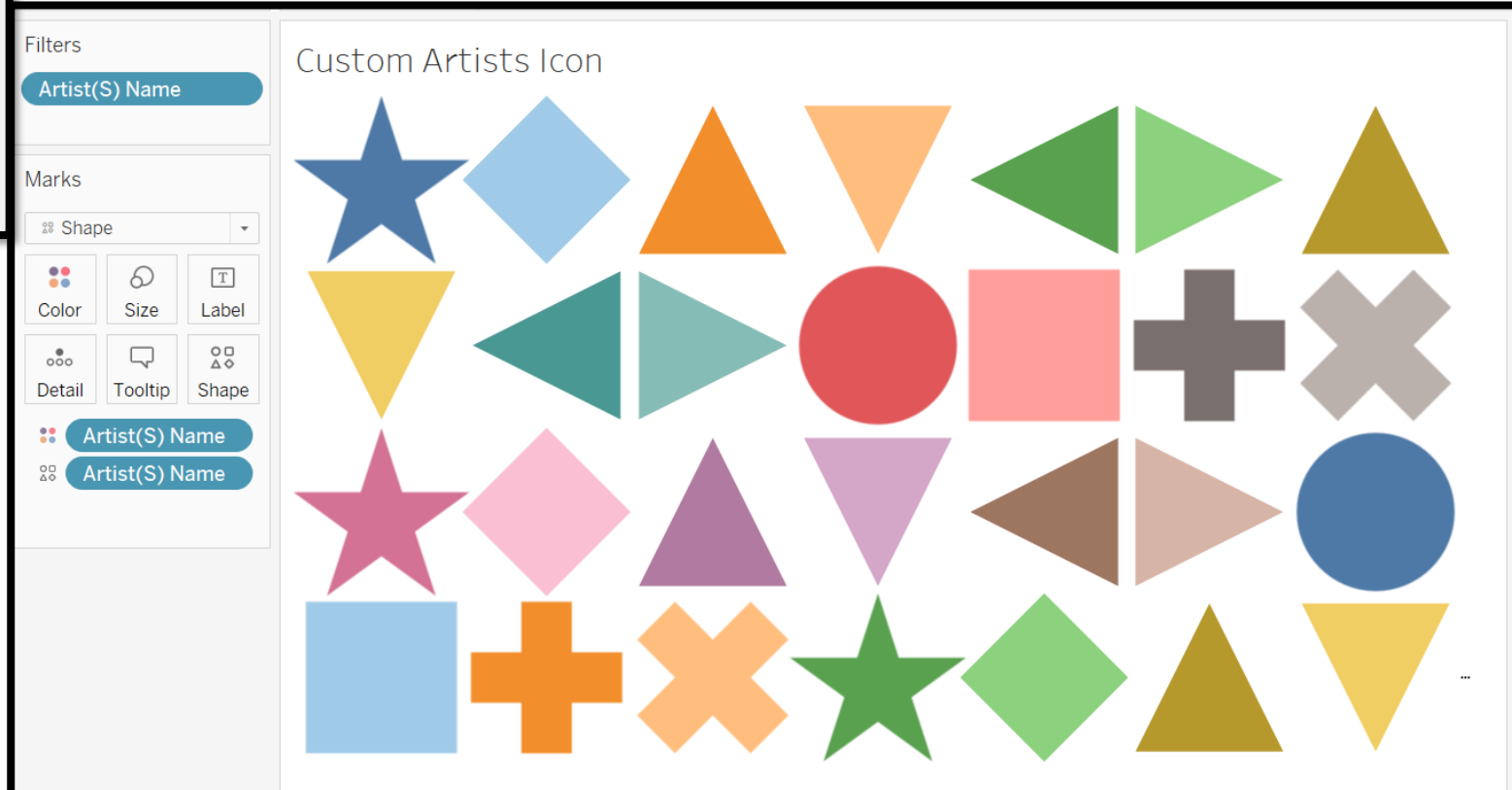
- **Create your visualization:**
  - Now, you can create your visualization as usual. For example, you can drag measures and dimensions to the "Rows" and "Columns" shelves to create a scatter plot or any other chart type.
  - As you do this, you'll notice that each data point (e.g., each artist) is represented by its custom shape instead of the default marks.

- **Format and customize the visualization:**
  - You can further format and customize your visualization, including labels, tooltips, and axis titles, as needed.

- **Save and share your customized visualization:**
  - Once you're satisfied with your visualization, save your Tableau workbook. You can then share it or publish it to Tableau Server or Tableau Public as desired.

Describe how to use custom shapes in Tableau to represent different data points uniquely.



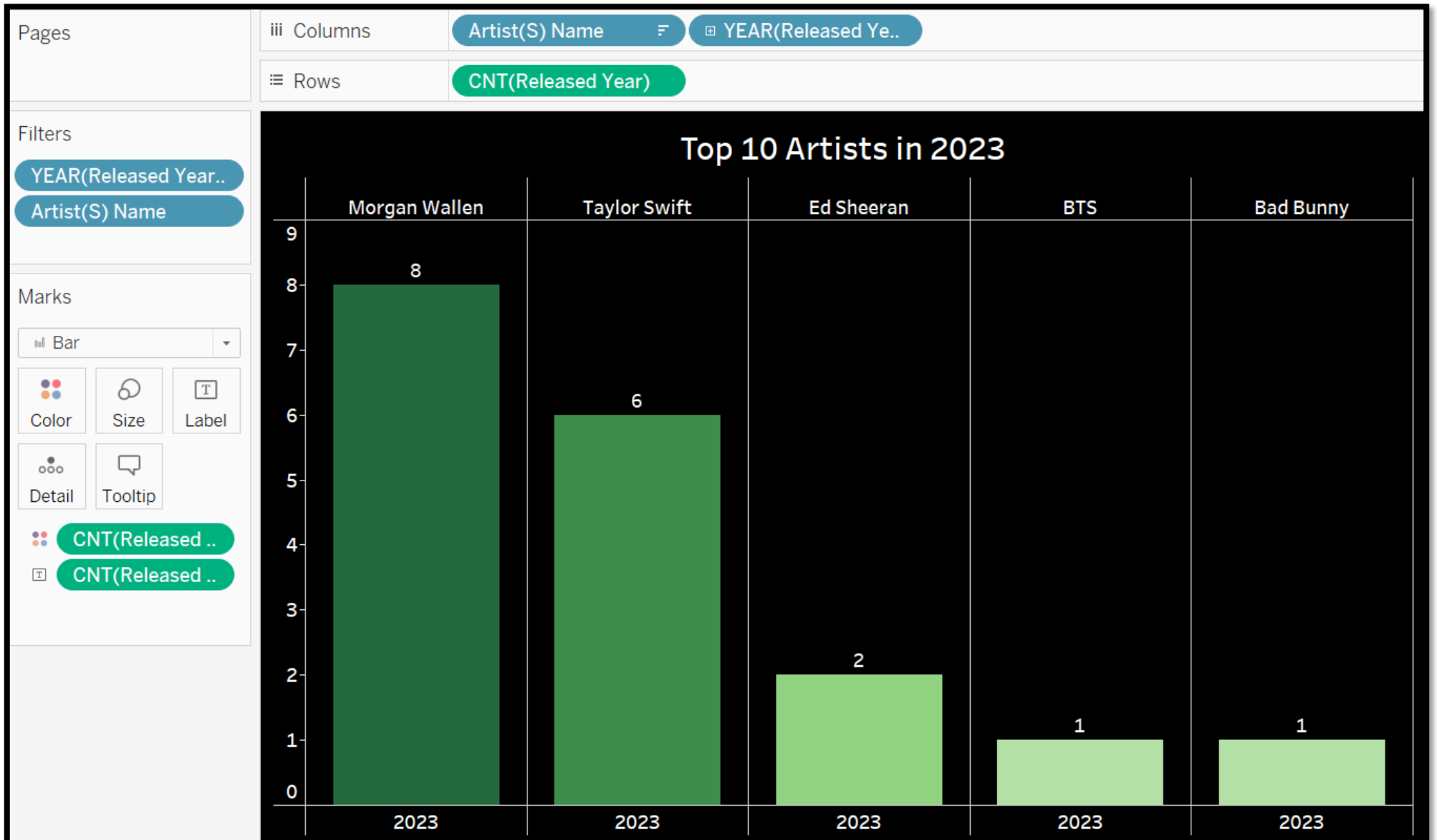


Explain the process of creating a filter in Tableau to interactively control the data displayed in a visualization.

Answer: To create a filter in Tableau, drag a dimension (e.g., "released_year") to the Filters shelf. Users can then interactively select values to filter the data displayed in the visualization.

1. To create an interactive filter for the "Released Year," drag the "Released Year" dimension from your dataset to the "Filters" shelf.
2. A filter control will appear, allowing you to select the years you want to filter by. You can customize this filter by right-clicking it and choosing "Edit."
3. In the "Edit Filter" dialog box, you can choose how you want the filter to behave. You can set it to be a single value dropdown, multiple value dropdown, a slider, or a range of options, depending on your preference. Adjust the settings to fit your needs.
4. You can also choose whether the filter affects one worksheet or multiple worksheets if you have multiple visualizations in your Tableau workbook.
5. Click "OK" to apply your filter settings.
6. Now, when you interact with the filter control, it will dynamically adjust the data displayed in your visualization based on the selected years. For example, if you select the year 2023, the bar chart will update to show data only for that year.
7. You can also add more filters for other dimensions or measures in a similar way to create a more interactive dashboard with multiple filter controls.
8. To view your visualization in Tableau, click on the "Show Me" tab and choose the type of chart or visualization you prefer. You can further customize the appearance of your chart and dashboard as needed.

Explain the process of creating a filter in Tableau to interactively control the data displayed in a visualization.



CREATE VISUALIZATIONS USING VARIOUS CHART TYPES IN TABLEAU

Generate a bar chart in Tableau to visualize the number of tracks released each year from the Spotify dataset.

Answer: To create a bar chart, follow these steps:

a. Drag "released_year" to Columns and "track_name" to Rows.

b. Change the Marks card to Bar.

c. Drag "track_name" to Label on the Marks card.

You'll have a bar chart showing the number of tracks released each year.

Generate a bar chart in Tableau to visualize the number of tracks released each year from the Spotify dataset.

Answer: To create a bar chart, follow these steps:

a. Drag "released_year" to Columns and "track_name" to Rows.

b. Change the Marks card to Bar.

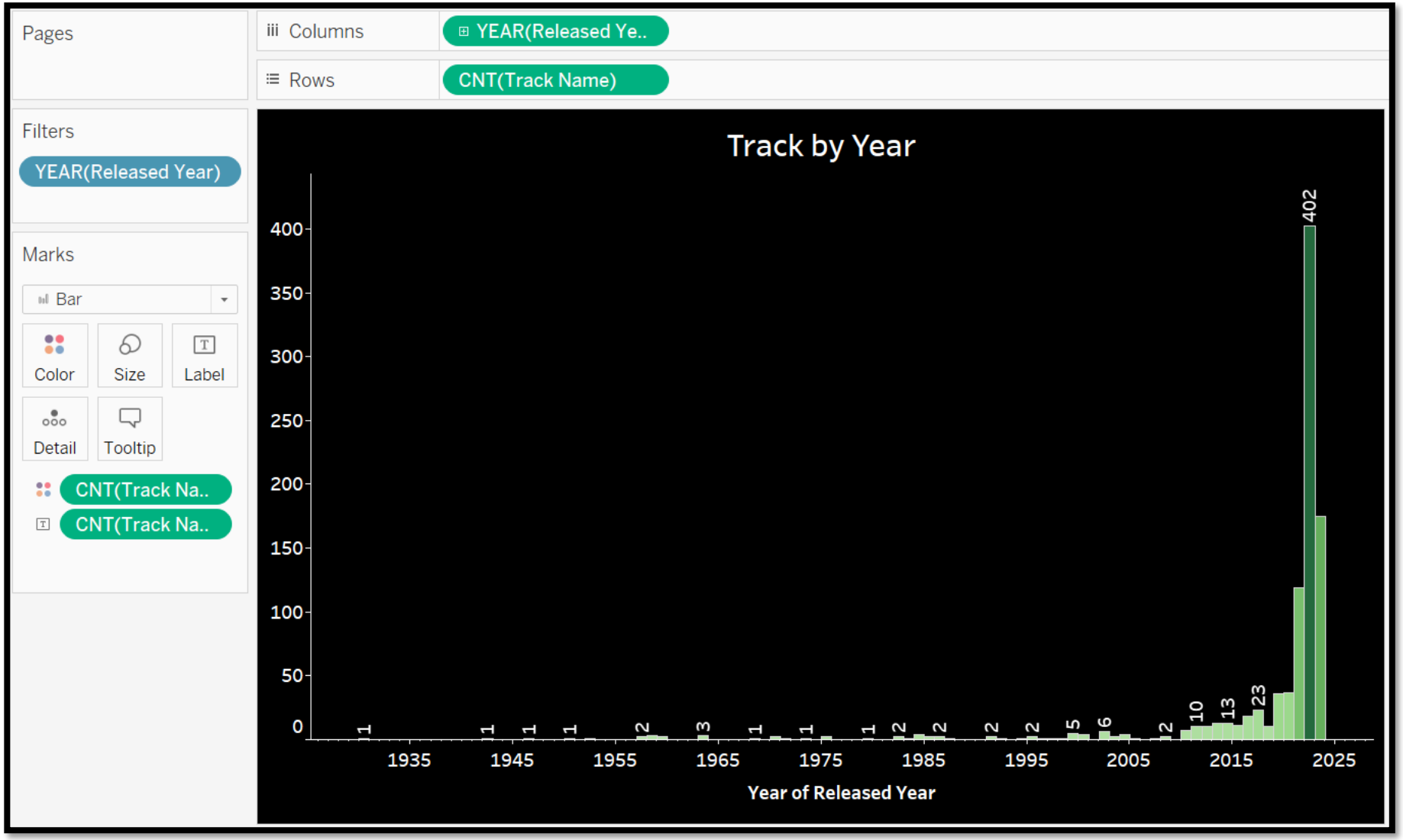
c. Drag "track_name" to Label on the Marks card.

You'll have a bar chart showing the number of tracks released each year.

Steps:

1. Drag the "released_year" dimension to the Columns shelf.
2. Drag the "Number of Records" measure to the Rows shelf.
3. Tableau will automatically create a bar chart with the number of tracks released each year.
4. Customize your chart as needed. You can change the chart title, adjust the colors, and format the axes.
5. If you want to show the specific number of tracks on each bar, drag the "Number of Records" measure to the Label shelf on the Marks card.

Generate a bar chart in Tableau to visualize the number of tracks released each year from the Spotify dataset.



Create a pie chart to display the distribution of danceability percentages for all tracks.

Answer: To create a pie chart:

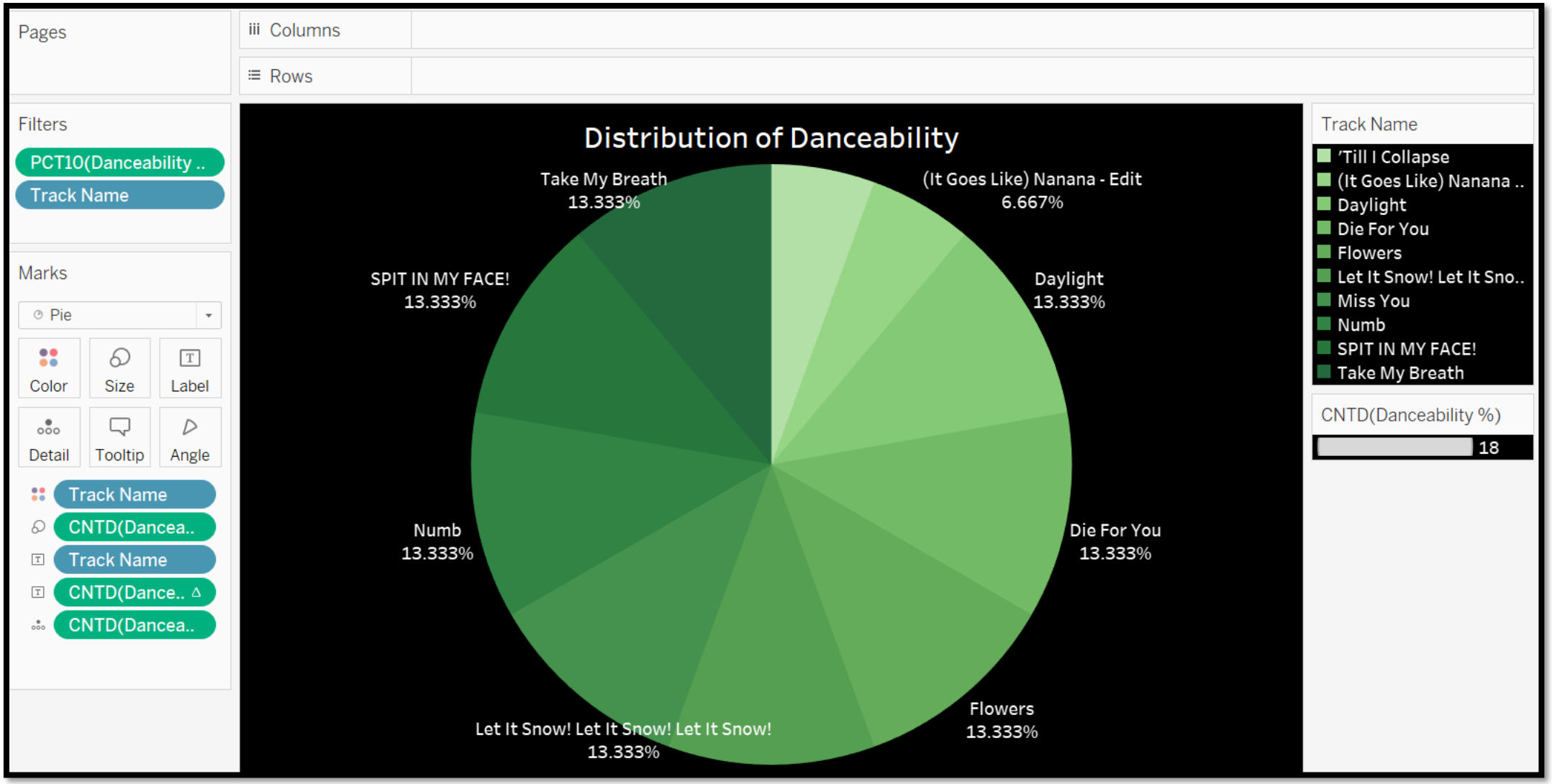
a. Drag "danceability_%" to Rows and "track_name" to Color on the Marks card.

b. Change the chart type to Pie.

Steps:

1. Drag the "danceability_%" measure to the Rows shelf.
2. Drag the "track_name" dimension to the Color shelf on the Marks card. This will help differentiate each segment of the pie chart by track name.
3. To change the chart type to Pie, locate the "Show Me" panel on the right side of the Tableau interface.
4. In the "Show Me" panel, you should see a pie chart icon. Click on it, and Tableau will automatically convert your visualization into a pie chart.
5. Tableau will generate a pie chart displaying the distribution of danceability percentages for all tracks, with each segment representing a different track.

Create a pie chart to display the distribution of danceability percentages for all tracks.



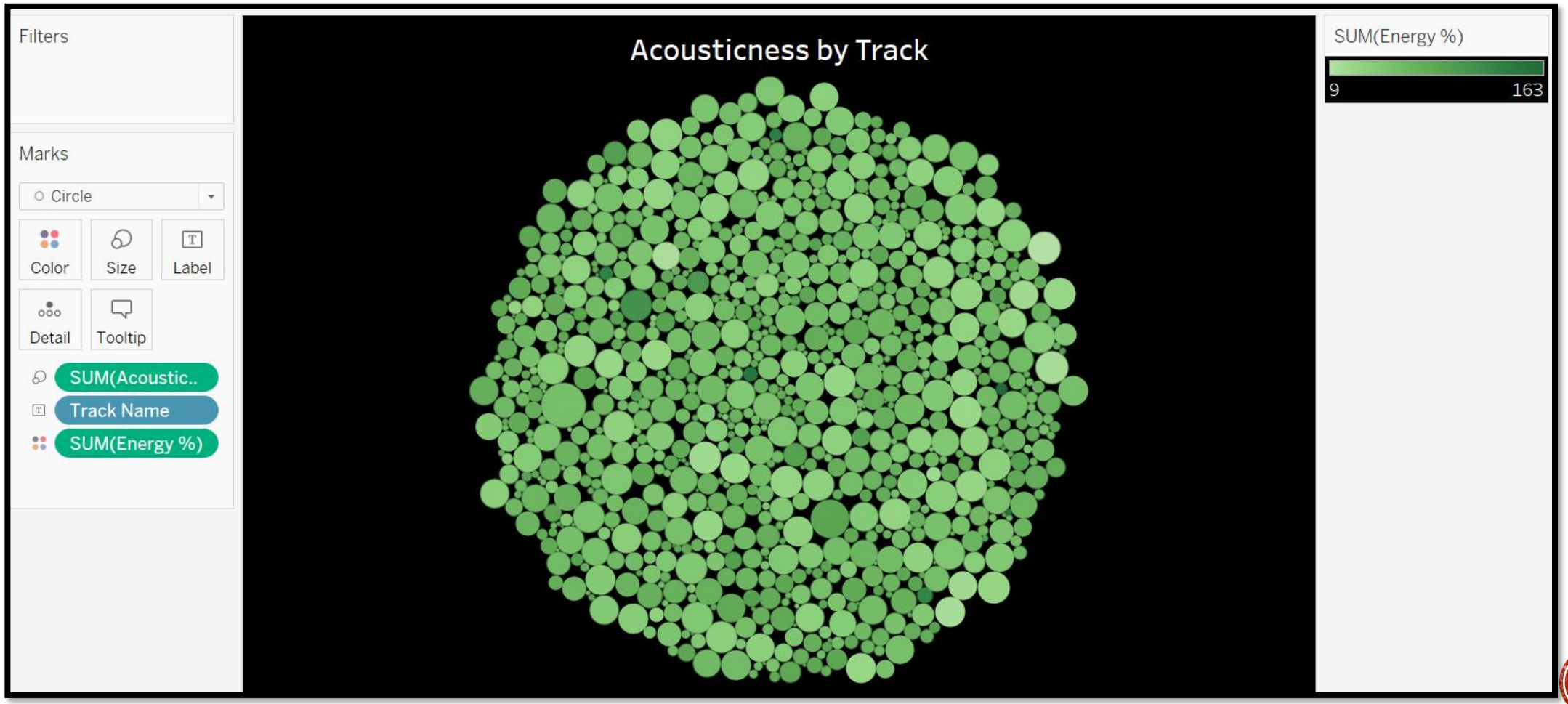
Using the Spotify dataset, build a bubble chart to show the relationship between energy percentage and acousticness percentage of tracks.

Answer: To create a bubble chart:

a. Drag "energy_%" to Columns, "acousticness_%" to Rows, and "track_name" to Label on the Marks card.

b. Adjust the size of the bubbles by dragging "streams" to Size on the Marks card.

Using the Spotify dataset, build a bubble chart to show the relationship between energy percentage and acousticness percentage of tracks.



FORMAT VISUALIZATIONS IN TABLEAU

How can you adjust the axis lines and number formats to improve the readability of a chart in Tableau?

Answer: To adjust axis lines and number formats, right-click the axis and select "Edit Axis." You can customize axis labels, number formats, and tick marks.

Describe the steps to add titles and tooltips to a visualization in Tableau and explain their importance.

Answer: To add titles and tooltips, right-click on the worksheet and select "Edit Title" or "Edit Tooltip." Titles provide context, and tooltips offer additional information when users hover over data points.

Provide guidelines for formatting the entire Tableau workbook, including color schemes and fonts, for a professional look.

Answer: To format the workbook professionally:

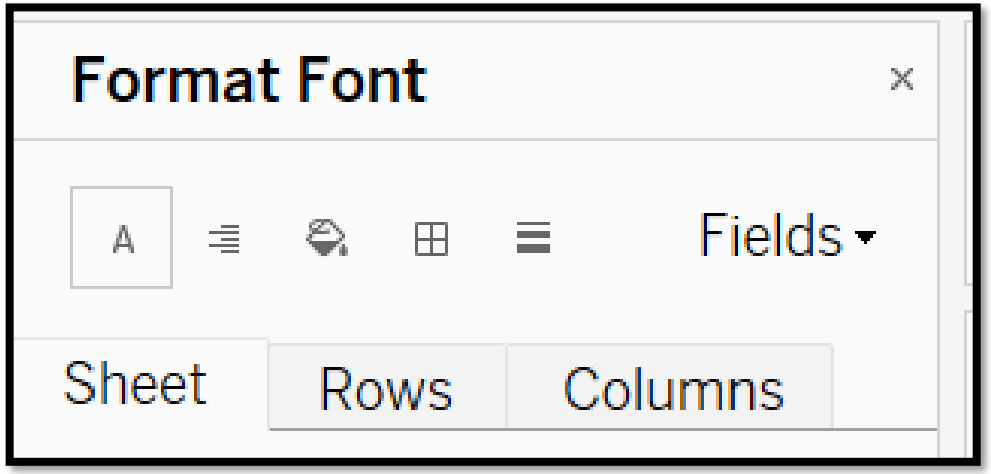
a. Use a consistent color scheme.

b. Choose readable fonts and font sizes.

c. Ensure alignment and spacing are consistent.

d. Consider the use of borders and shading for emphasis.

e. Test the workbook's appearance on different devices for readability.



SOME SITUATION BASED QUESTION FOR DS_SALARIES

1. Salary Comparison by Experience Level:

Create a bar chart to compare the average salary for different experience levels (e.g., Junior, Mid-Level, Senior) in the data science field.

2. Remote Work Analysis:

Build a pie chart to visualize the distribution of remote work ratios (e.g., Fully Remote, Partially Remote, On-Site) among data science-related job positions.

3. Salary Trends Over the Years:

Create a line chart to show how the average salary for data science jobs has changed over the years (using the work_year column).

4. Top Paying Job Titles:

Create a bar chart to identify the top 10 job titles that offer the highest average salaries for data scientists.

SOME SITUATION BASED QUESTION FOR DS_SALARIES

5. Salary Comparison by Employment Type:

Build a stacked bar chart to compare the average salaries among different employment types (e.g., Full-Time, Part-Time, Contract) for data science roles.

6. Geographical Salary Analysis:

Create a map visualization to explore the average salaries for data science jobs in different company locations. Use the company_location column and display salaries as color-coded data points on the map.

7. Company Size vs. Salary:

Generate a scatter plot to analyze the relationship between company size (small, medium, large) and average salaries for data science positions.

SOME SITUATION BASED QUESTION FOR DS_SALARIES

8. Salary Currency Conversion:

Create a calculated field to convert salaries to a common currency (e.g., USD) based on the salary_currency column. Then, visualize the converted salaries.

9. Residence vs. Company Location:

Build a bar chart to compare the locations of employees' residences (employee_residence) with the locations of their companies (company_location) for data science jobs.

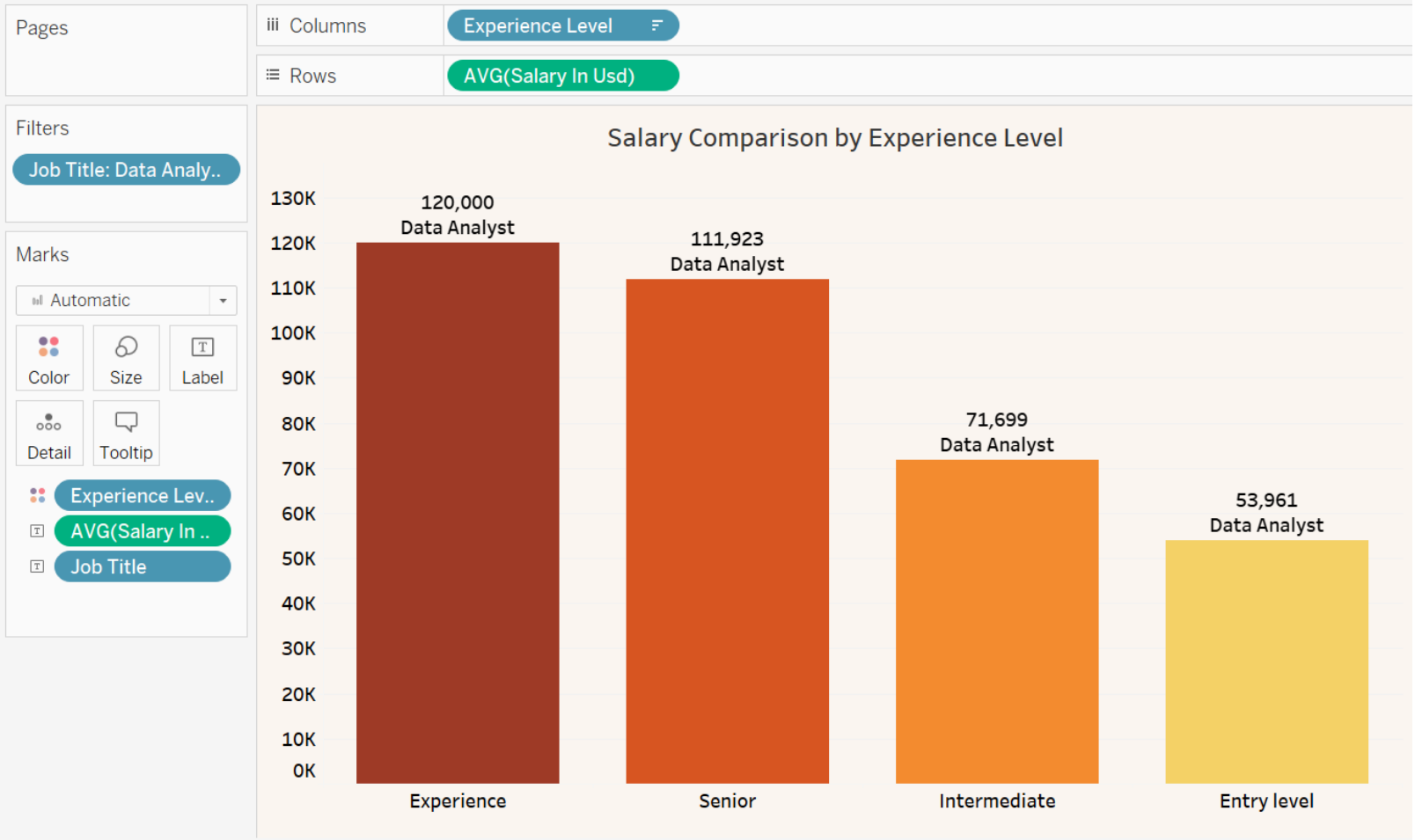
10. Salary Distribution:

Create a histogram to visualize the distribution of salaries for data science-related positions. This can help identify salary ranges and outliers.

Solutions for ds_salaries

1. Salary Comparison by Experience Level:

Create a bar chart to compare the average salary for different experience levels (e.g., Junior, Mid-Level, Senior) in the data science field.



2. Remote Work Analysis:

Build a pie chart to visualize the distribution of remote work ratios (e.g., Fully Remote, Partially Remote, On-Site) among data science-related job positions.

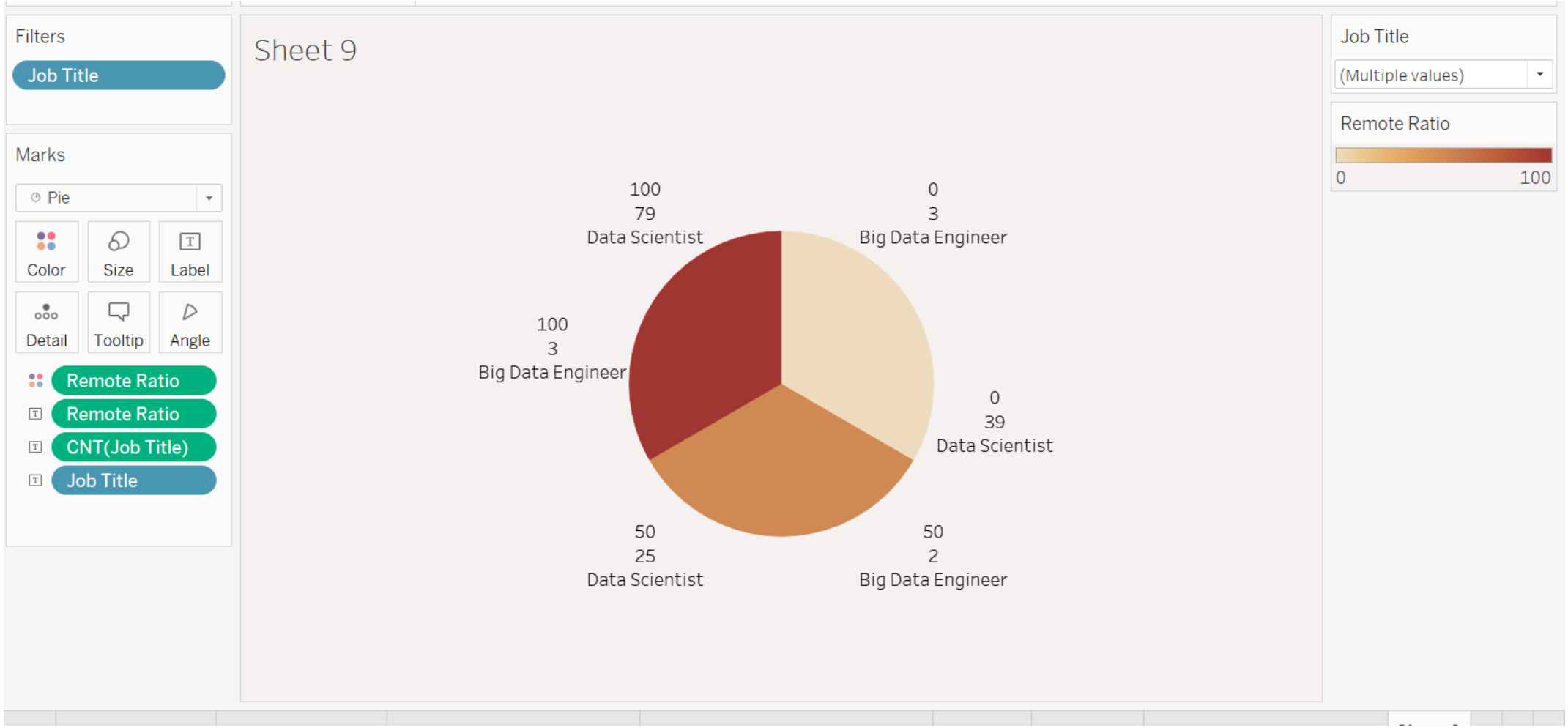
Steps:

1. Drag the "remote_ratio" field to the "Columns" shelf instead of the "Rows" shelf.

2. Tableau might still aggregate it as a measure (e.g., SUM(remote_ratio)). In the "Columns" shelf, right-click on the pill representing "remote_ratio" and choose "Dimension."

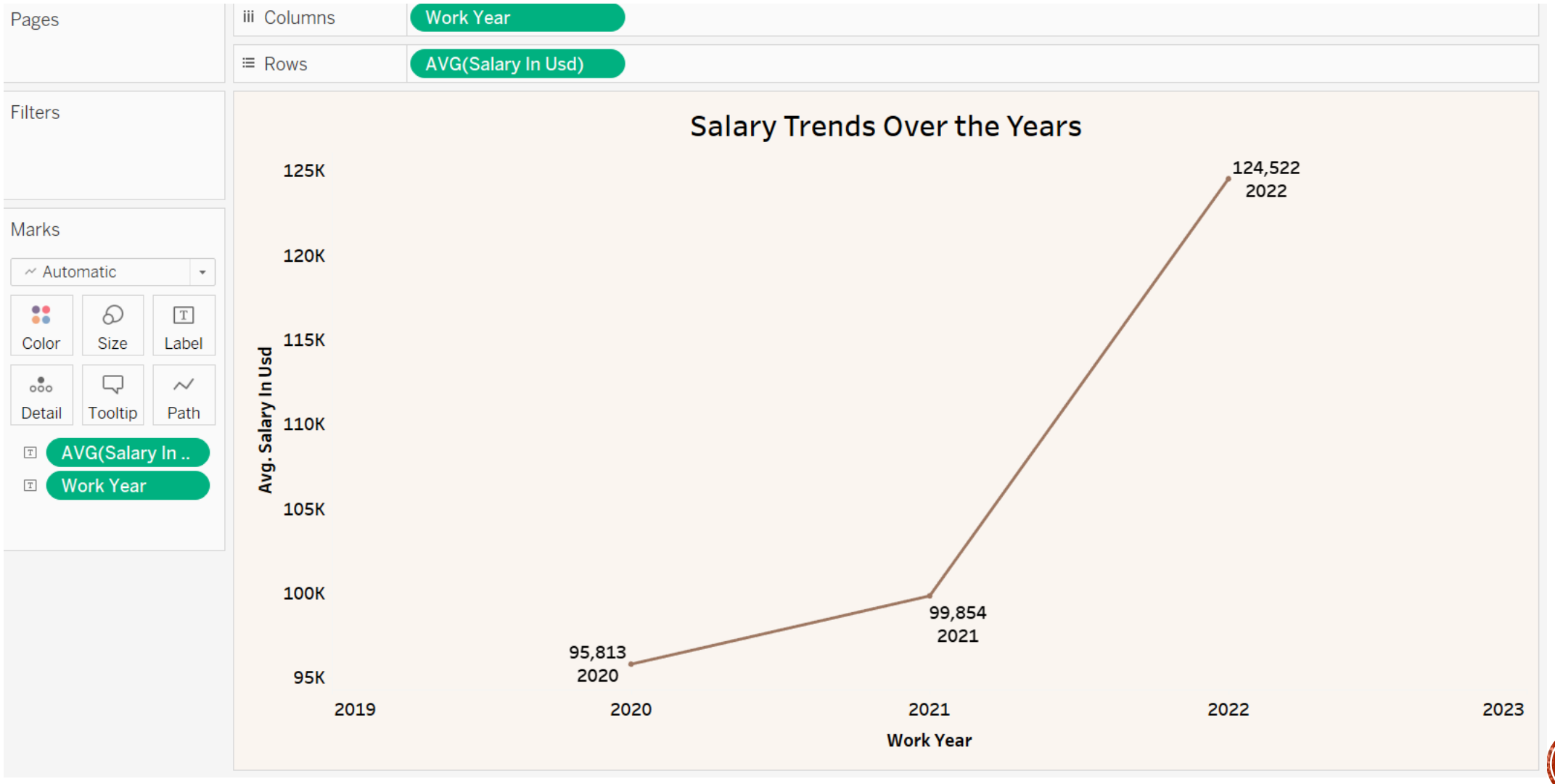
3. Now, drag the same "remote_ratio" field to the "Label" shelf on the Marks card to label the pie chart slices.

4. You should now have a single pie chart showing the distribution of remote work ratios.


Filters
Job Title
Marks
Pie
Color
Size
Label
Detail
Tooltip
Angle
Remote Ratio
Remote Ratio
CNT(Job Title)
Job Title
Sheet 9
100
79
Data Scientist
0
3
Big Data Engineer
100
3
Big Data Engineer
0
39
Data Scientist
50
25
Data Scientist
50
2
Big Data Engineer
Job Title
(Multiple values)
Remote Ratio
0
100

3. Salary Trends Over the Years:

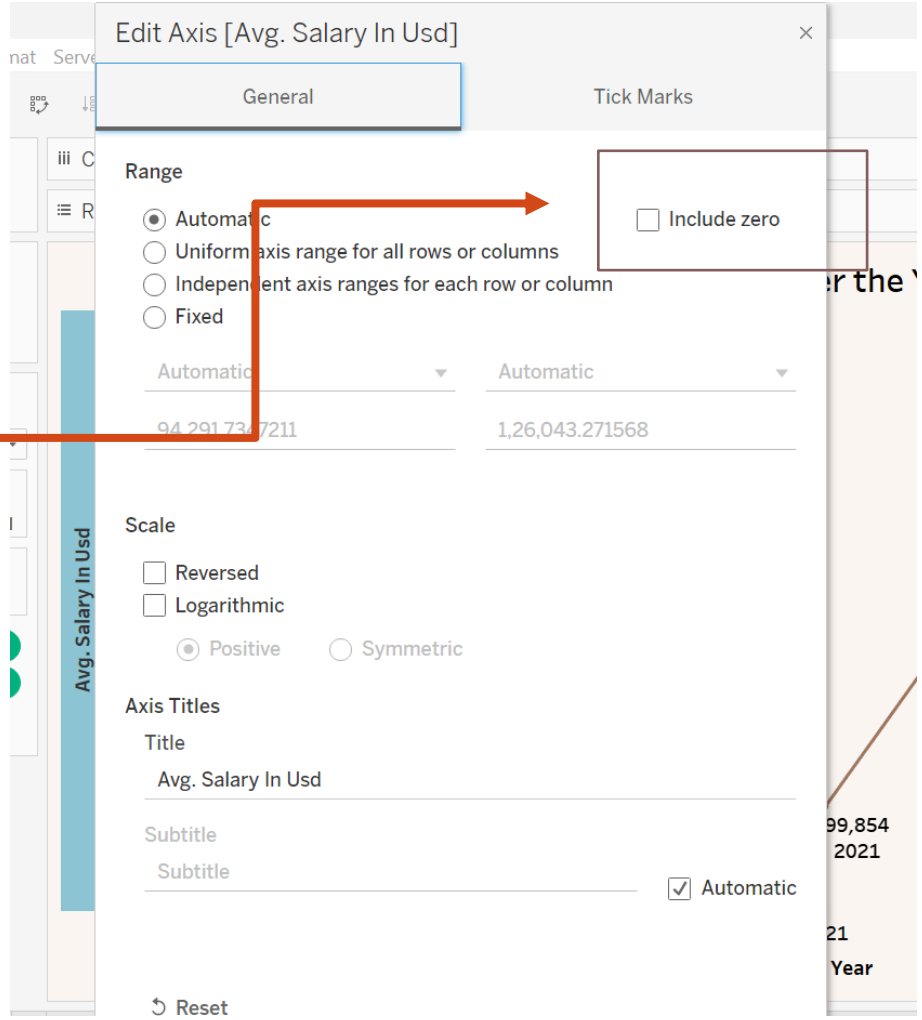
Create a line chart to show how the average salary for data science jobs has changed over the years (using the work_year column).



3. Salary Trends Over the Years:

Create a line chart to show how the average salary for data science jobs has changed over the years (using the work_year column).

Right Click on Left Axis and Uncheck this box



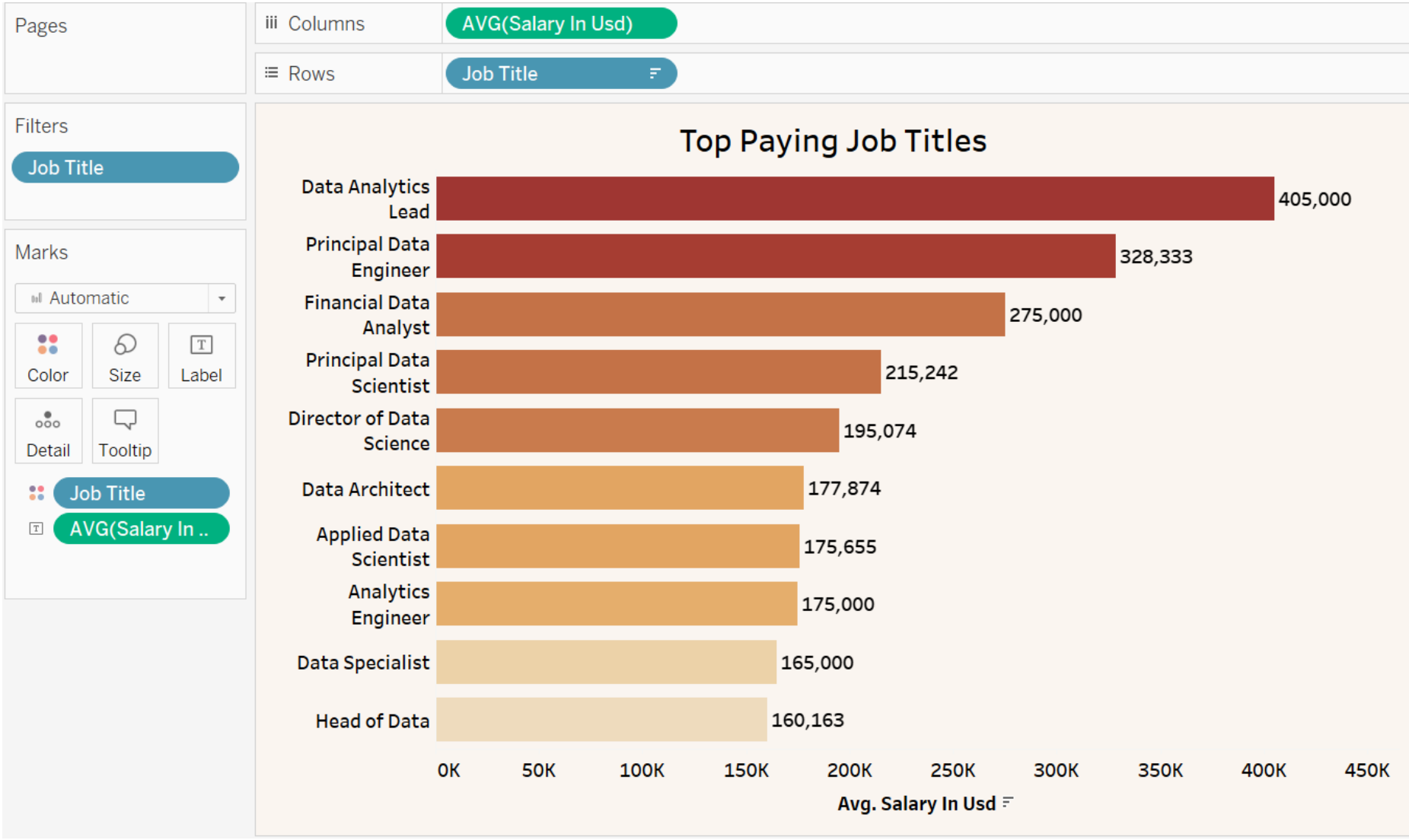
1.Drag the "work_year" dimension from the data source panel to the Columns shelf at the top of the workspace. This will set the years as the x-axis of your line chart.

2.Drag the "salary_in_usd" measure from the data source panel to the Rows shelf below the Columns shelf. This will set the salary as the y-axis of your line chart.

3.Tableau will automatically aggregate the salary measure as a SUM. To calculate the average salary, right-click on "SUM(salary_in_usd)" in the Rows shelf and select "Average" under the "Measure" option.

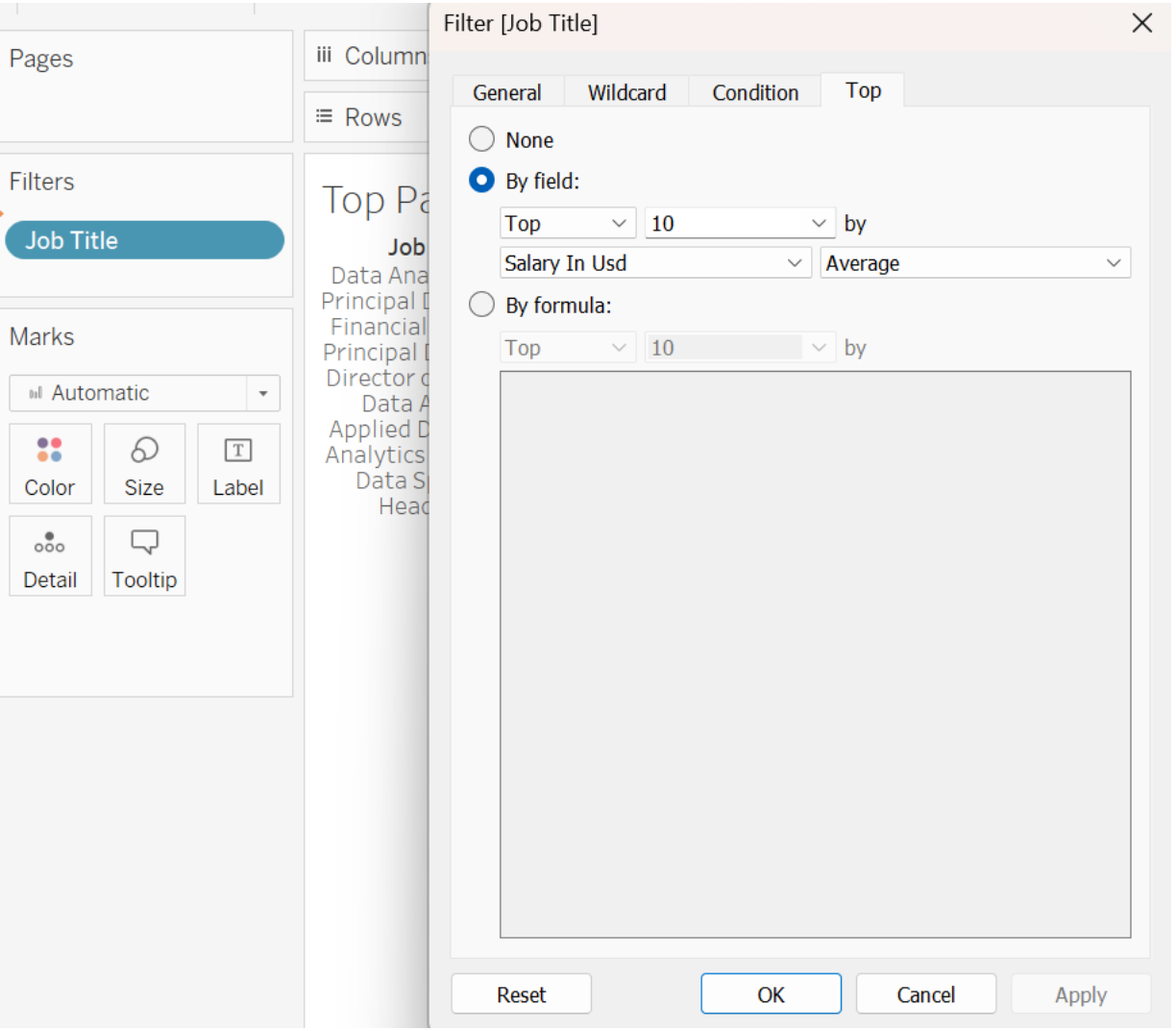
4. Top Paying Job Titles:

Create a bar chart to identify the top 10 job titles that offer the highest average salaries for data scientists.



4. Top Paying Job Titles:

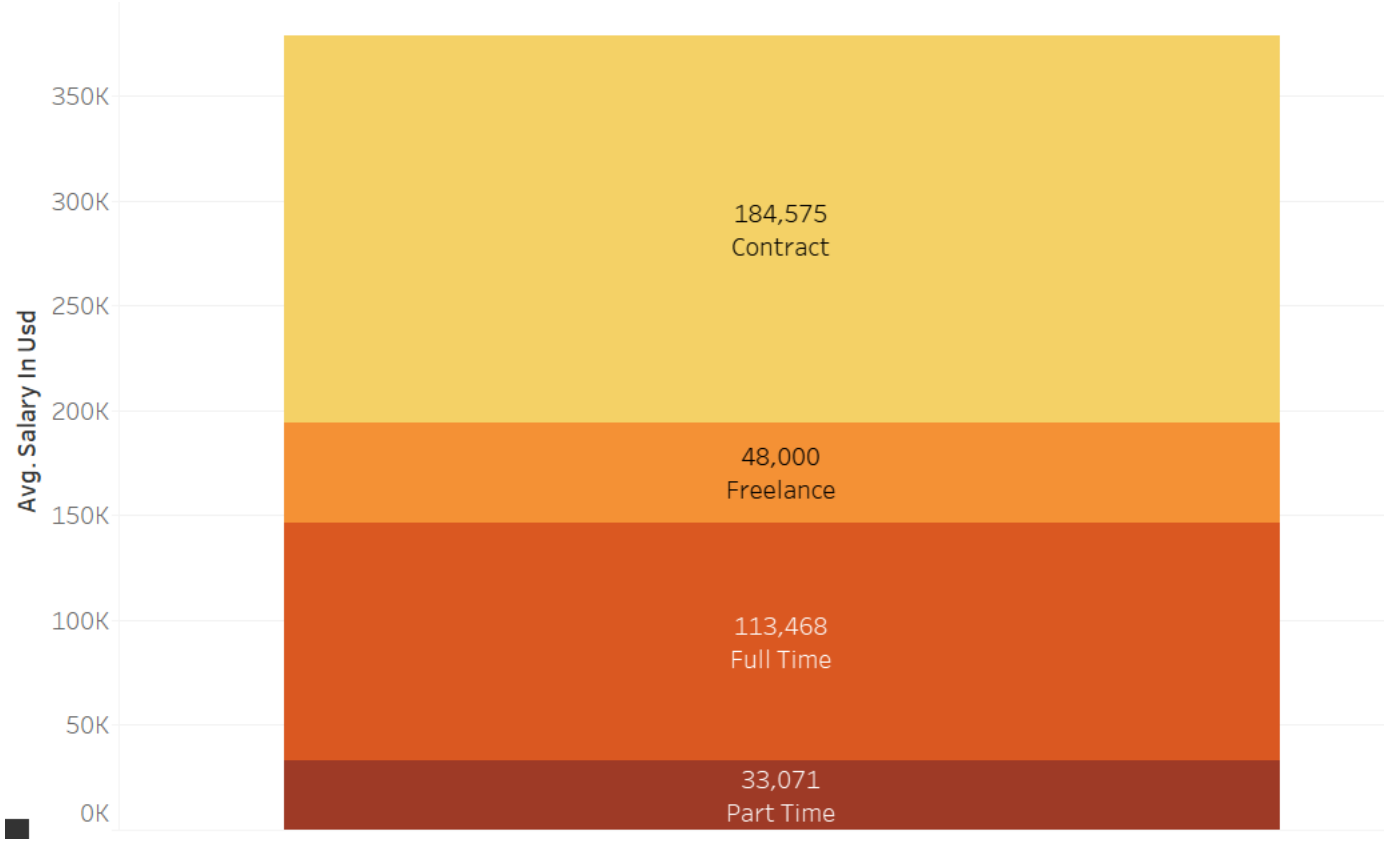
Create a bar chart to identify the top 10 job titles that offer the highest average salaries for data scientists.

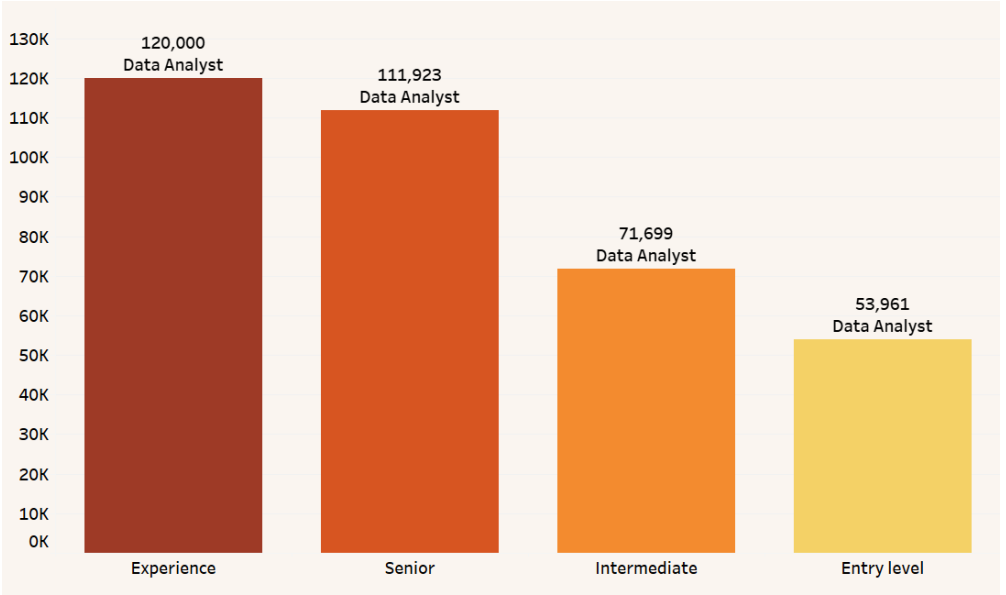


1. Drag the "Job Title" dimension to the "Rows" shelf.
2. Drag the "Salary" measure to the "Columns" shelf.
3. Make sure the "Salary" measure is set to "Average" by right-clicking it in the "Columns" shelf and selecting "Measure" > "Average."
4. Drag the field that indicates whether the job is related to data science to the "Filters" shelf.
5. Select the option for "Data Science" or a similar value to filter for data science-related jobs only.

5. Salary Comparison by Employment Type:

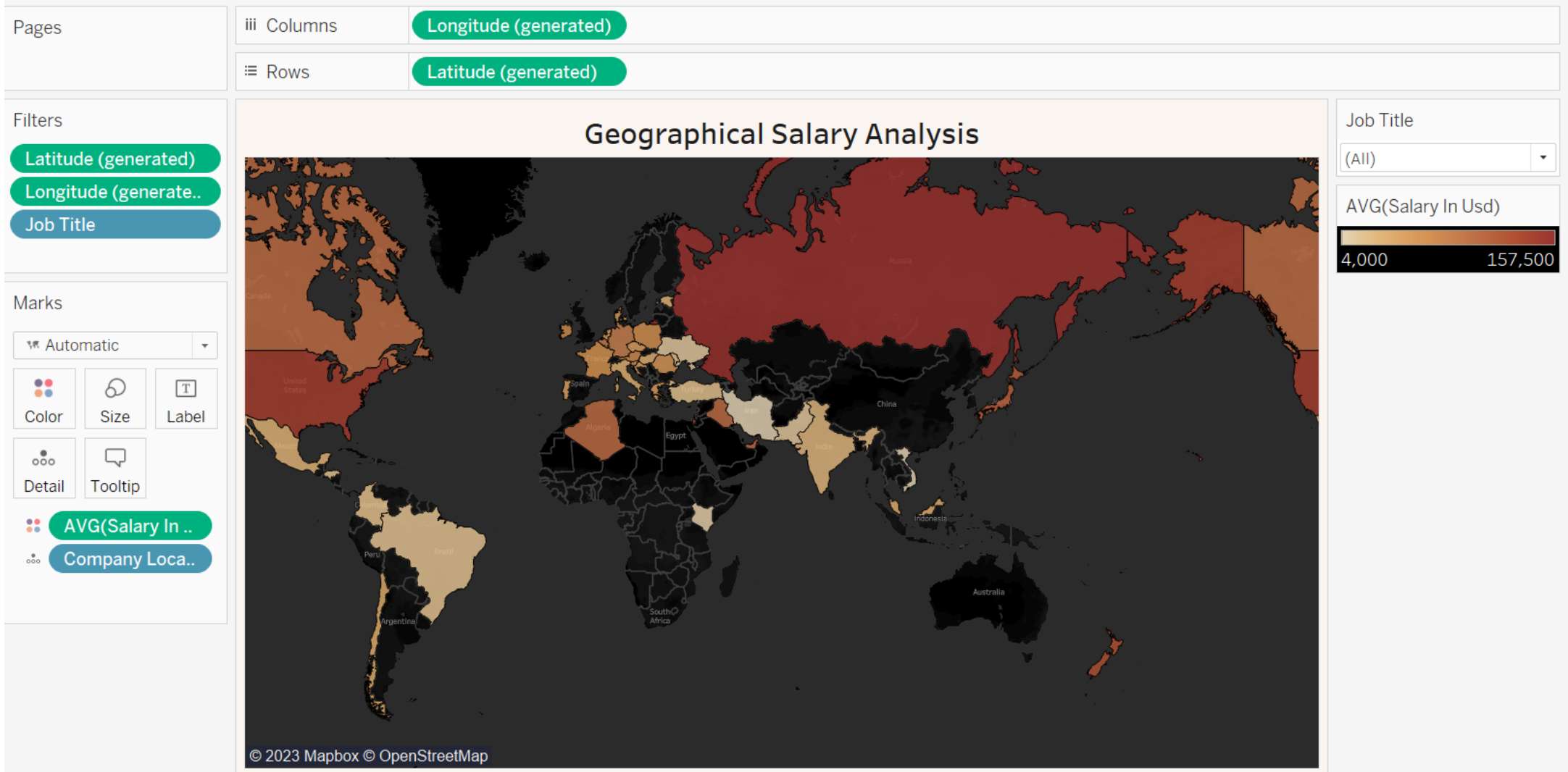
Build a stacked bar chart to compare the average salaries among different employment types (e.g., Full-Time, Part-Time, Contract) for data science roles.





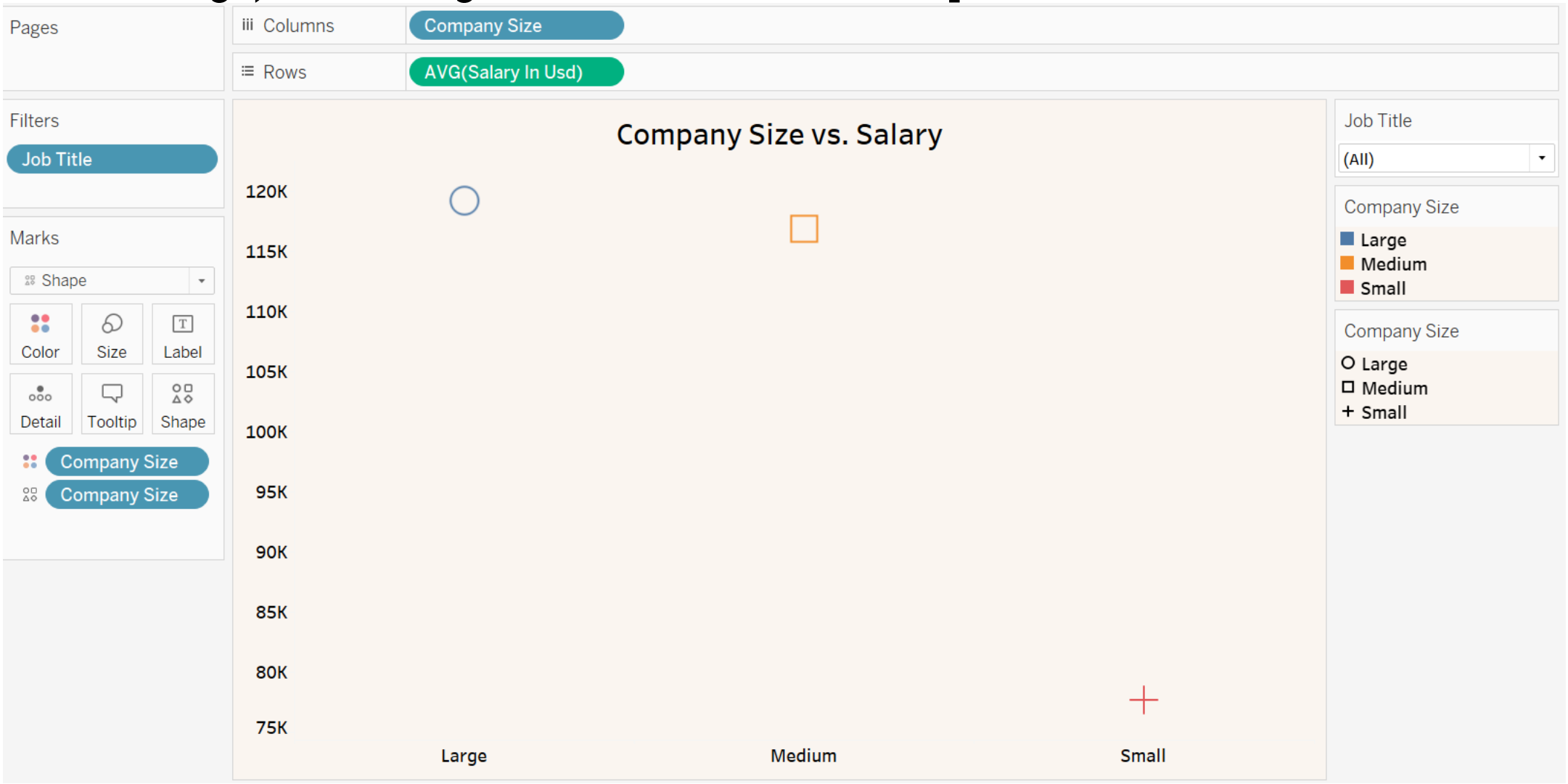
6. Geographical Salary Analysis:

Create a map visualization to explore the average salaries for data science jobs in different company locations. Use the company_location column and display salaries as color-coded data points on the map.



7. Company Size vs. Salary:

Generate a scatter plot to analyze the relationship between company size (small, medium, large) and average salaries for data science positions.



SOLUTIONS FOR DS_SALARIES

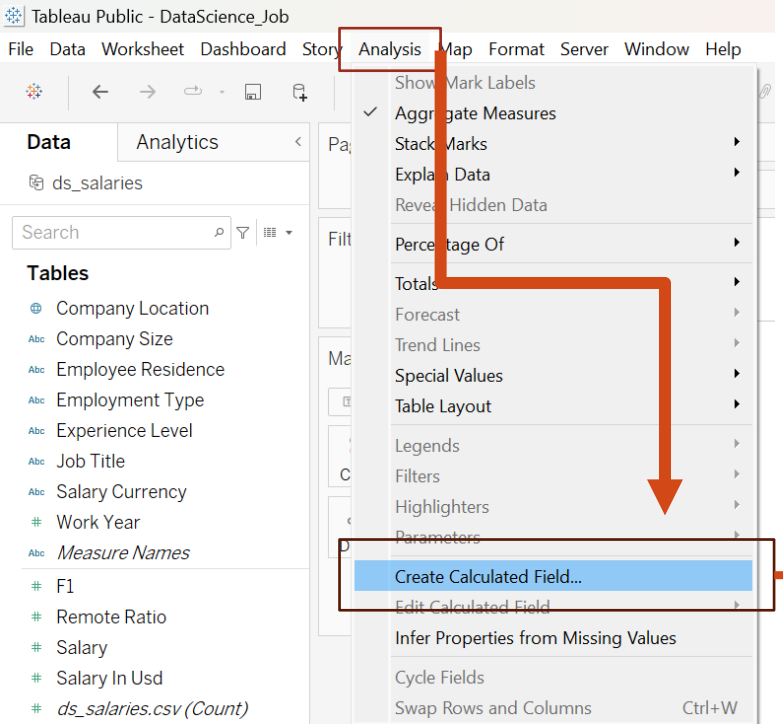
8. Salary Currency Conversion:

Create a calculated field to convert salaries to a common currency (e.g., USD) based on the salary_currency column. Then, visualize the converted salaries.

Assuming the following exchange rates:

1 EUR = 1.17 USD
1 GBP = 1.32 USD
1 HUF = 0.0036 USD
1 INR = 0.014 USD
1 JPY = 0.0091 USD
1 MXN = 0.049 USD
1 CAD = 0.78 USD
1 CNY = 0.16 USD

1 PLN = 0.23 USD
1 SGD = 0.73 USD
1 DKK = 0.16 USD
1 BRL = 0.18 USD
1 TRY = 0.12 USD
1 AUD = 0.72 USD
1 CHF = 1.09 USD
1 CLP = 0.0013 USD



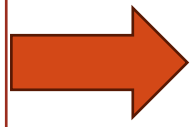
Salary Update

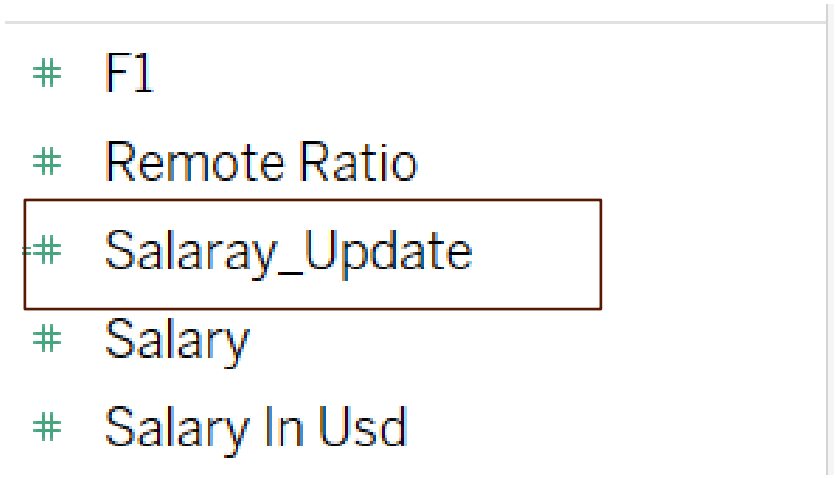
```
CASE [Salary Currency]
    WHEN 'EUR' THEN [salary] * 1.17
    WHEN 'GBP' THEN [salary] * 1.32
    WHEN 'HUF' THEN [salary] * 0.0036
    WHEN 'INR' THEN [salary] * 0.014
    WHEN 'JPY' THEN [salary] * 0.0091
    WHEN 'MXN' THEN [salary] * 0.049
    WHEN 'CAD' THEN [salary] * 0.78
    WHEN 'PLN' THEN [salary] * 0.23
    WHEN 'SGD' THEN [salary] * 0.73
    WHEN 'DKK' THEN [salary] * 0.16
    WHEN 'BRL' THEN [salary] * 0.18
    WHEN 'TRY' THEN [salary] * 0.12
    WHEN 'AUD' THEN [salary] * 0.72
    WHEN 'CHF' THEN [salary] * 1.09
    WHEN 'CLP' THEN [salary] * 0.0013
    WHEN 'CNY' THEN [salary] * 0.16
    ELSE [salary]
END
```

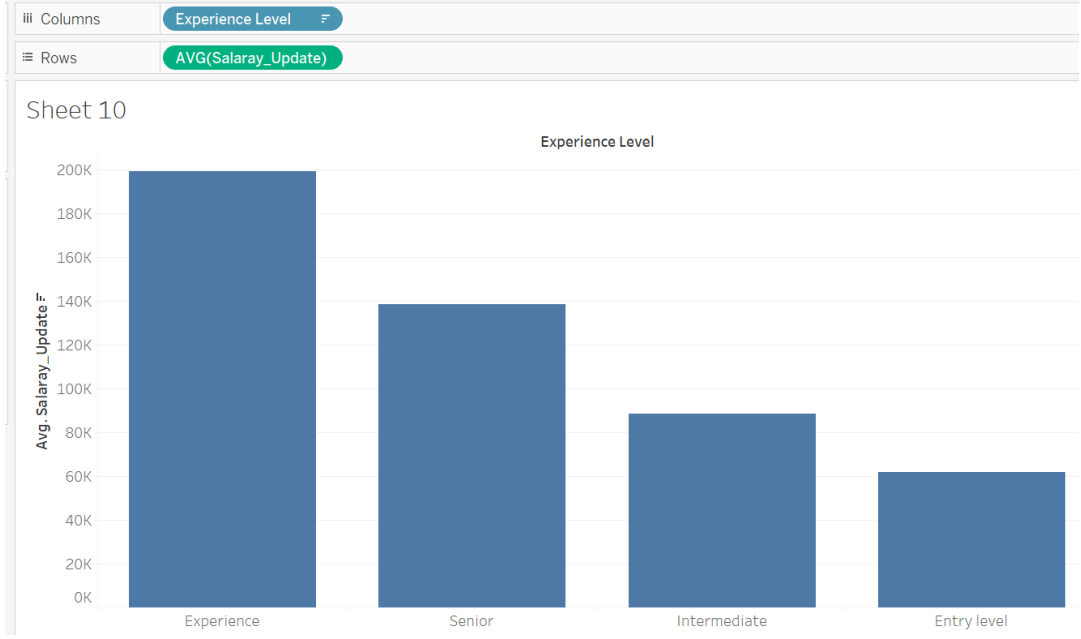
The calculation is valid.

Apply OK

```
CASE [Salary Currency]
   WHEN 'EUR' THEN [salary] * 1.17
   WHEN 'GBP' THEN [salary] * 1.32
   WHEN 'HUF' THEN [salary] * 0.0036
   WHEN 'INR' THEN [salary] * 0.014
   WHEN 'JPY' THEN [salary] * 0.0091
   WHEN 'MXN' THEN [salary] * 0.049
   WHEN 'CAD' THEN [salary] * 0.78
   WHEN 'PLN' THEN [salary] * 0.23
   WHEN 'SGD' THEN [salary] * 0.73
   WHEN 'DKK' THEN [salary] * 0.16
   WHEN 'BRL' THEN [salary] * 0.18
   WHEN 'TRY' THEN [salary] * 0.12
   WHEN 'AUD' THEN [salary] * 0.72
   WHEN 'CHF' THEN [salary] * 1.09
   WHEN 'CLP' THEN [salary] * 0.0013
   WHEN 'CNY' THEN [salary] * 0.16
    ELSE [salary]
END
```

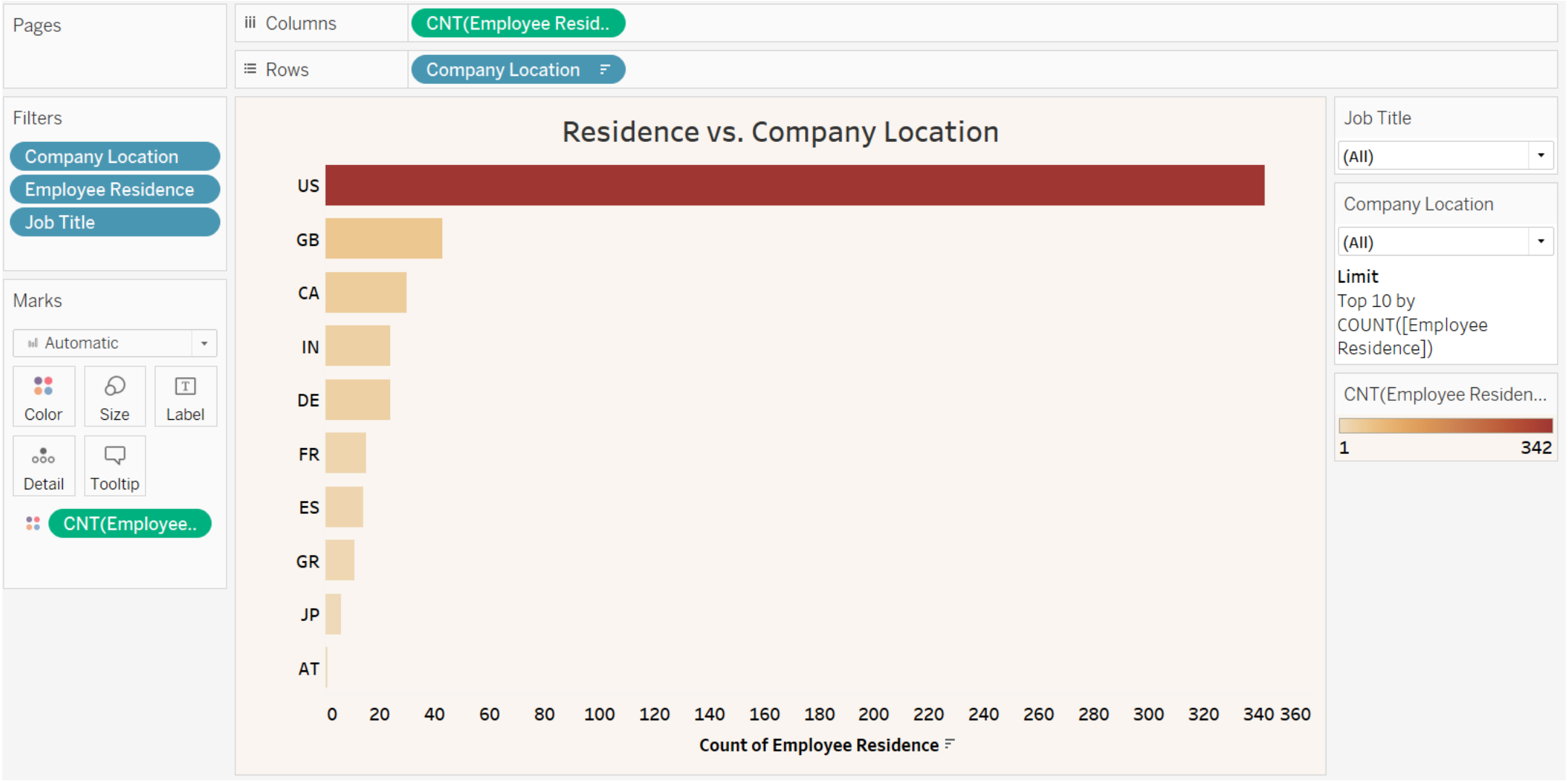






9. Residence vs. Company Location:

Build a bar chart to compare the locations of employees' residences (employee_residence) with the locations of their companies (company_location) for data science jobs.



10. Salary Distribution:

Create a histogram to visualize the distribution of salaries for data science-related positions. This can help identify salary ranges and outliers.