



BUS 5100-93 (35749) Project Tutorial



Authors: [Madeleine Dela Cruz](#), [AndyMike Emata](#), [Claudia Perez De La Cruz](#), [Kelly Templeton](#), [Kelly Young](#)

Instructor: [Longwook Woo](#)

Date: 04/02/2024

Lab Tutorial

{aemata, mdelac65, cperezd, ktemple2, kyoung49} @calstatela.edu

04/02/2024

Anticipated Home Buying Trends by MBA Graduates Using SAC

Objectives

In this Lab Tutorial, you will perform the following:

1. Data Engineering exercise to acquire, transform dataset and model data from Zillow to SAP Analytics Cloud
2. Story Creation exercise to provide visualization on Geo Maps, Charts and Timeseries Analysis.
3. Create Regression Model to help predict the forecasted average prices of the houses.

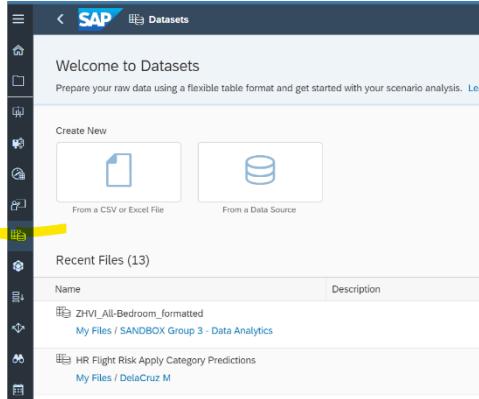
Platform Spec

- **Desktop Browser :** Google Chrome or Microsoft Edge
- **Operating System :** MS Windows 10 or higher or Apple OS Catalina (10.15 or higher)
- **Network Bandwidth :** 500-800 kbit/s per user
- **Screen Resolution :** XGA 1024x768 (high color) or higher. Widescreen: 1366x766 or higher
- **Minimum recommended browser cache size :** 250MB

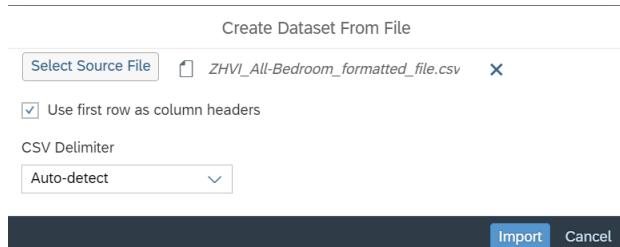
Data Modeling SAP Analytics Cloud

1. Open a web browser and go to SAP Analytics Cloud (SAC). <https://higher-education.us10.sapanalytics.cloud/>

2. Go to **Files** → **Shared with Me** → **Cohort 8 Group 3 FEMBA**
3. Download **ZHVI_All-Bedroom_FINAL.csv**
4. Go to **Data Set** on the left panel and click **From a CSV or Excel File**



5. Find the file you downloaded and click **Import**



6. Choose the folder where you will save your file and click **OK**
7. Once the file has been imported, go to **RegionName** column and change column name to **MetroName**

RegionID	SizeRank	MetroName	RegionType
394913	1	New York, NY	msa
753899	2	Los Angeles, CA	msa
394463	3	Chicago, IL	msa
394514	4	Dallas, TX	msa
394692	5	Houston, TX	msa

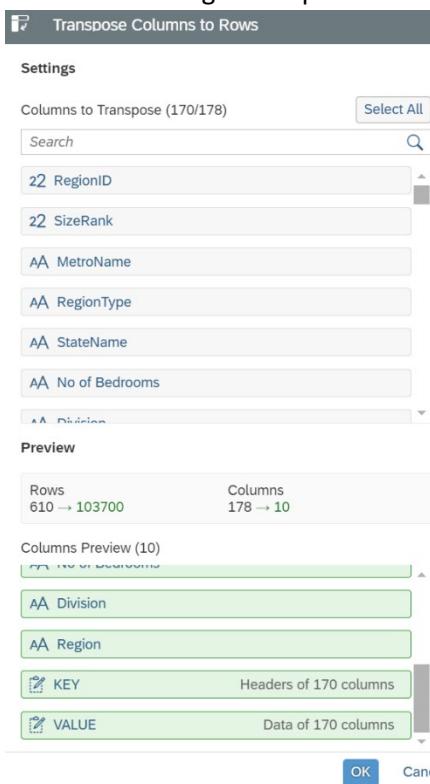
8. Highlight column 1/31/2010, click shift and highlight until 2/29/2024

1 st 1/31/2010	1 st 2/28/2010	1 st 3/31/2010	1 st 4/30/2010	1 st 5/31/2010	1 st 6/30/2010	1 st 7/31/2010
247597.0447	248228.9263	248761.7465	249118.1208	249179.9435	249912.1002	250279.0072
235676.2577	234557.0266	233224.6787	233240.6319	233834.7643	232758.9132	230819.8642
132347.7057	132243.3776	131441.6513	131561.4959	130904.0632	130811.7945	128928.5192
60630.58505	60285.33115	59623.7929	59350.39329	59102.20137	59096.81013	58920.01674
62923.65936	62770.0896	62517.84187	62381.92467	62162.68229	61887.94372	61337.80263
234159.6	235528.1229	236595.6882	237530.4137	238500.4762	239303.764	239342.5651
147918.8391	148053.7744	148170.9712	148256.6327	148028.3076	147668.5272	147044.6771
76412.50583	75587.71121	74937.6903	74354.34639	73669.15926	72793.10293	71907.6958
101980.2533	101513.1907	101235.3583	101278.1138	100755.9915	100097.6968	98962.21088
214012.8377	214730.8714	215267.3562	215899.9433	216275.7122	216731.8215	216668.5523
72524.98157	71335.21559	70016.73692	68833.9787	67663.68007	66540.524	64849.11737

9. Go to Actions and click Transpose Columns to Rows



10. Click OK at the right side panel.



11. The message below will pop-up, click Got it!

Sampled Data

Full data size: 103700 rows, 10 columns

Sampled size: 2040 rows

Any changes made to the sample data are applied to the full set of data during dataset creation.

Got it!

12. Go to Key Column and change to Date

13. Go to Value Column and change to HousePrice

AA KEY	AA VALUE
AA Date	AA HousePrice
1/31/2010	247597.0447
1/31/2010	235676.2577
1/31/2010	132347.7057
1/31/2010	60630.58505

14. Go to Data Set Overview and make sure that Measures and Dimension are set as below:

Dataset Overview

ZHVI_All-Bedroom_formatted_... 2040 row
13 column

Search

Output Columns

Measures (1/3)

- HousePrice NONE
- Latitude NONE
- Longitude NONE

Dimensions (10)

- RegionID
- MetroName
- RegionType
- StateName
- No of Bedrooms
- Division
- Region
- Date
- SizeRank
- City

- 15. Go to Date Column, click Details on the panel on the right, change Data Type to Date and Statistical Type to Continuous**

Viewing Details For: Date

Dimension Properties:

Description: No column specified

Column Details:

Data Type: Date

Statistical Type: Continuous

Conversion Format: MM/DD/YYYY

Data Distribution Validation

Sampled Rows: 2040 / 103700

Unique Values 167

Group Values Count

1/31/2010 and 165 other values	12
11/30/2023	8

- 16. Go to HousePrice Column, click Details on the panel on the right, change Aggregation Type to None and Statistical Type to Continuous**

Viewing Details For: HousePrice

Measure Properties:

(i) Measure properties are used in visualizations.

Measure Units:

Type a measure unit (limit 20 characters)

Aggregation Type:

NONE

Decimal Places:

2

Scale:

None Specified

Column Details:

Viewing Column Details For: HousePrice

Data Type: Decimal

Statistical Type: Continuous

17. Go to **MetroName** column and on the pop-up, click **Split on ","**

AA MetroName	AA MetroName_1	AA MetroNa...
New York, NY	NY	
Los Angeles, CA	CA	
Chicago, IL	IL	
Dallas, TX	TX	
Houston, TX	TX	
Washington, DC	DC	
Philadelphia, PA	PA	
Miami, FL	FL	
Atlanta, GA	GA	

18. Change MetroName_1 to City

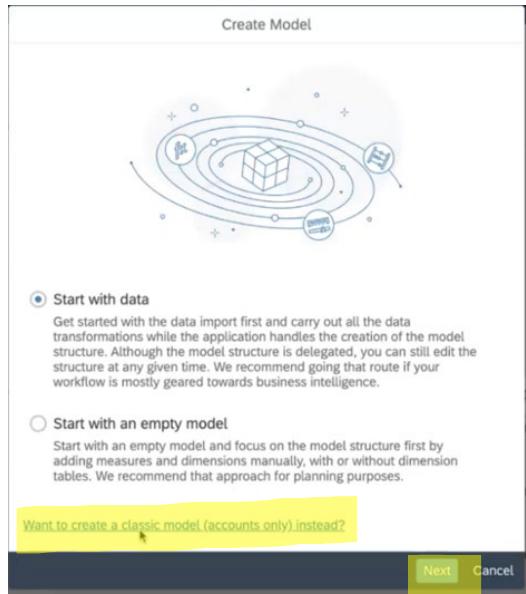
19. Delete MetroName_2

20. Click Save.

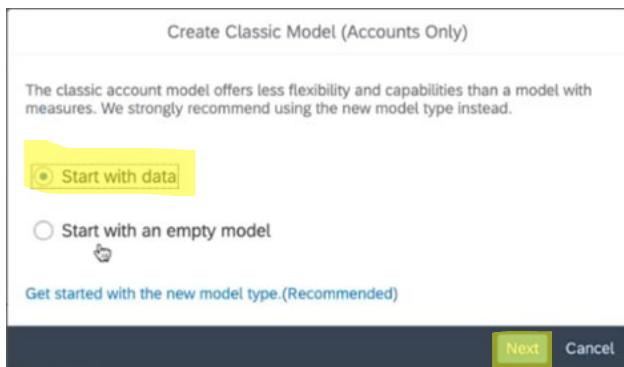
21. Got to **Menu → Modeler → Model**

The screenshot shows the SAP Modeler application interface. At the top, there's a navigation bar with icons for home, back, SAP logo, and Modeler. Below the bar, the title "Welcome to the Modeler" is displayed, followed by a subtitle "Prepare your raw data using a structured view and get started with your sc...". A horizontal menu bar includes "Models", "Public Dimension Tables", "Currency Conversions", and "Poin...". On the left side, there's a vertical toolbar with various icons. The main area has sections for "Create New" (with "Model" and "Live Data Model" options) and "Recent Files (5)".

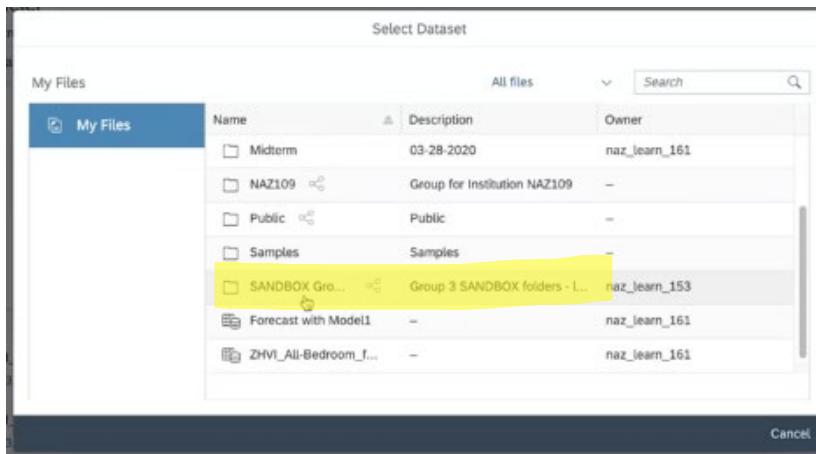
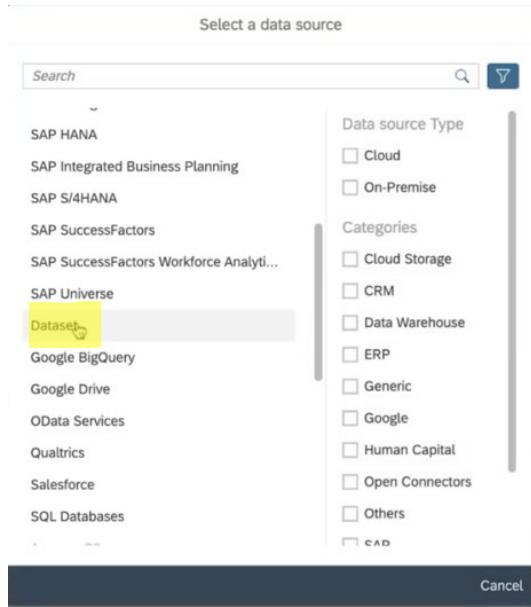
22. Click **Want to create a classic model (accounts only) instead?**



23. Click **Start with Data**



24. Click **Dataset** and find the dataset you have saved.



25. Click **MetroName** Column and under Actions, click **Geo Enrichment → Coordinates**

New Model*					
Output	Display	Actions			
Geo Enrich by: Coordinates MetroName RegionType Size					
Longitude	City		Coordinates	MetroName	RegionType
-73.9249	New York		Area Name	New York, NY	NY
-118.4068	Los Angeles			Los Angeles, CA	CA
-87.6866	Chicago			Chicago, IL	msa
-96.7667	Dallas			Dallas, TX	msa
-95.3885	Houston			Houston, TX	msa

26. **Geo by Coordinates** should look like below

Geo by Coordinates

Dimension Name*
MetroName_Location

Identifiers

Location ID
MetroName

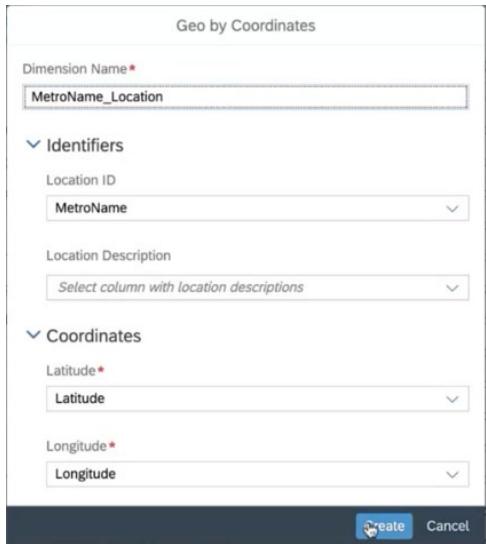
Location Description
Select column with location descriptions

Coordinates

Latitude*
Latitude

Longitude*
Longitude

Create Cancel



27. The panel on the right should look as below.

MetroName

Unique Values: 144 | Sampled Rows: 2000 | Data Type: Text

Modeling

Type: Generic

Property: Latitude

Property: Longitude

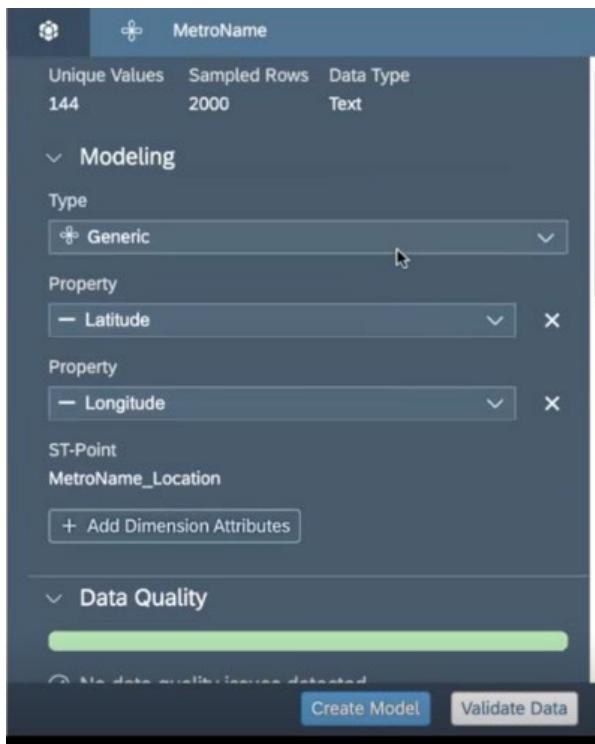
ST-Point
MetroName_Location

+ Add Dimension Attributes

Data Quality

0% of the data quality issues detected

Create Model Validate Data



28. Click **MetroName** Column and under **Actions**, click **Level-Based Hierarchy**



	RegionType	StateName	MetroName
C	msa	NY	New York, NY
les	msa	CA	Los Angeles, CA
	msa	IL	Chicago, IL
	msa	TX	Dallas, TX
	msa	TX	Houston, TX

29. Hierarchy should look like below, click OK.

Hierarchy Builder

Census_Region_Division

Region
Northeast, West, Midwest, South

Division
Middle Atlantic, Pacific, East North Central, West So...

StateName
NY, CA, IL, TX, VA, PA, FL, GA, MA, NC, WI, TN, OR, ...

MetroName
New York, NY, Los Angeles, CA, Chicago, IL, Dallas, ...

Generate a 100% unique bottom level for my hierarchy.

OK Cancel

30. The note will show below, click OK.

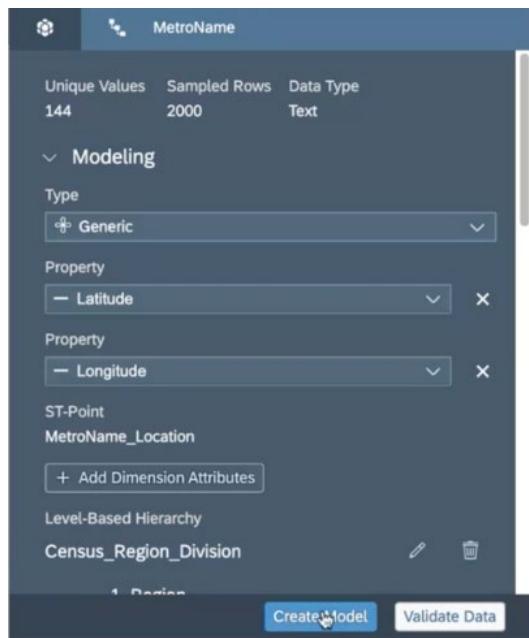
① Confirm Modeling Changes

You won't be able to change the name of your hierarchy after the hierarchy is created.

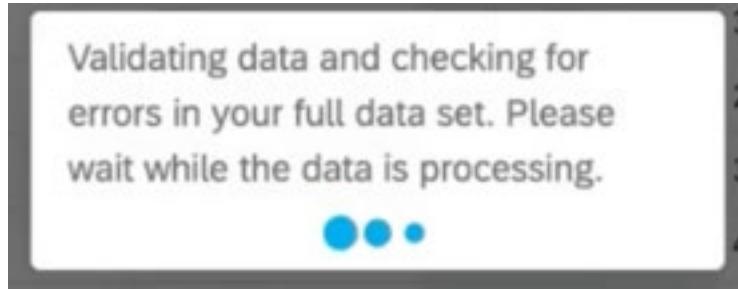
"Region" will become a property of "MetroName".
 "Division" will become a property of "MetroName".
 "StateName" will become a property of "MetroName".

OK Cancel

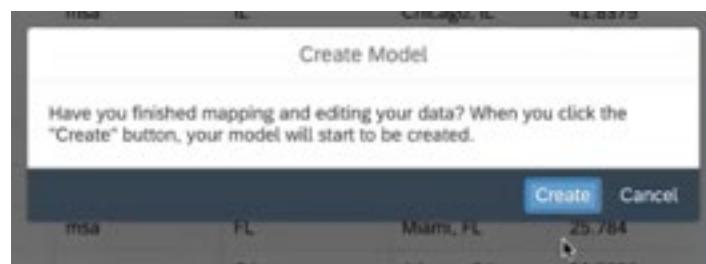
31. Right Panel should look like below. Click **Create Model**.



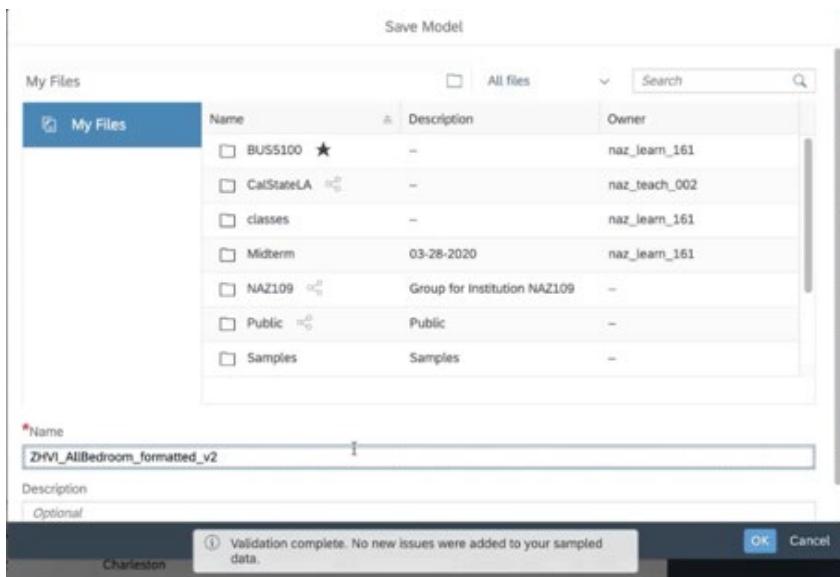
32. The note below will pop-up.



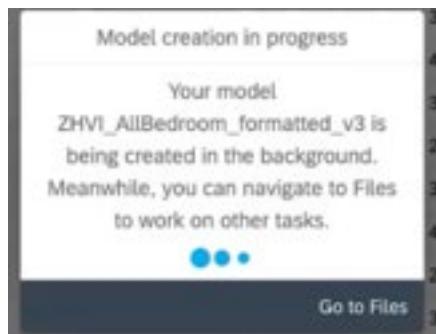
33. Click Create once the pop-up below appears.



34. Name your model and click OK.



35. This note will pop-up, wait for the model creation to finish.



Story Creation SAP Analytics Cloud

1. Select **Stories** from the menu on the left side of the screen. Then, select **Create New Responsive Page**.

Welcome to Stories

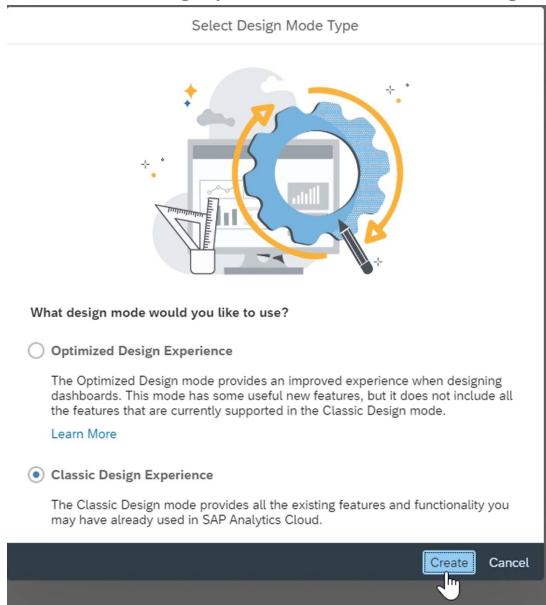
Create a story to visualize, summarize, explore, and monitor yo...

Create New

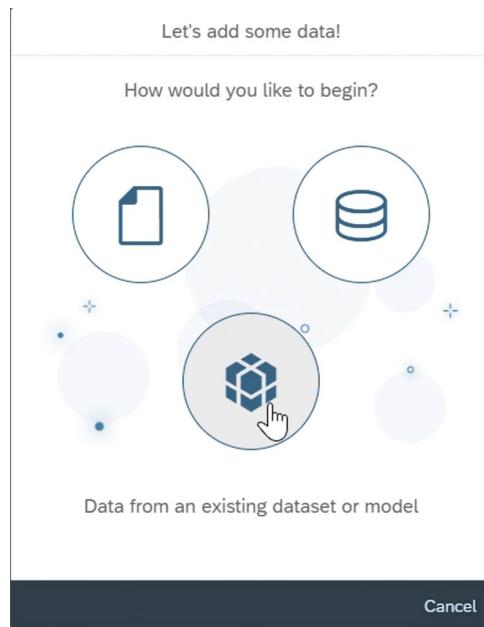
Responsive

Canvas

- a. In the following option, select **Classic Design Experience** and then select “Create”.



- b. On the upper left of the page, click “Data” and select the “Module” icon that looks like a cube in the options pop-up.



- c. Locate the model you created.

Select Dataset or Model

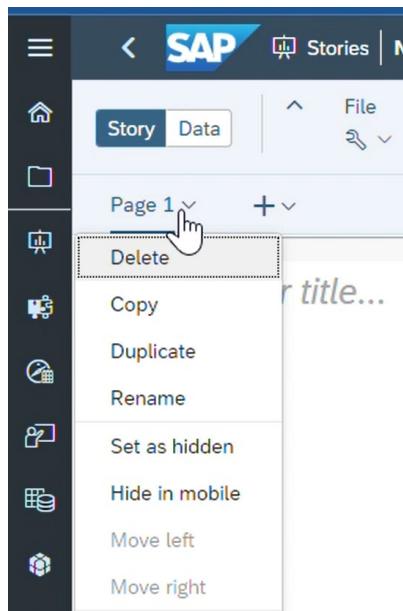
My Files / SANDBOX Group 3 - Data Analytics

All files	Search		
My Files	Name	Description	Owner
OLD Files	–	naz_learn_150	
Housing Value P...	–	naz_learn_168	
ZHVI_All-Bedroom...	–	naz_learn_150	
ZHVI_All-Bedroom...	–	naz_learn_161	
ZHVI_AllBedroom...	–	naz_learn_161	
ZHVI_AllBedroom... [ZHVI_AllBedroom_formatted_v3]	–	naz_learn_161	

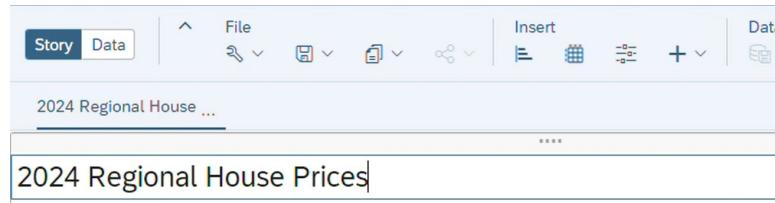
Cancel

Comparison Charts Home Price Analysis in SAP Analytics Cloud

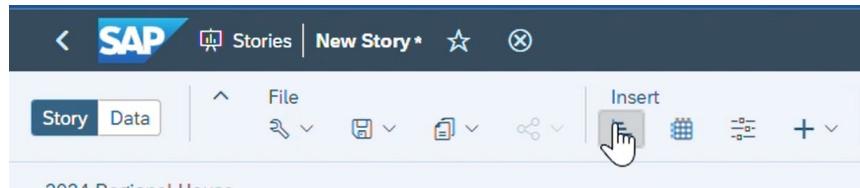
1. Click **Story**
2. Rename Page 1 of the Story. On the upper left corner, toggle to **Story** and then click on the downward chevron (looks like a “v”) of **Page 1** to select **Rename**. Rename the page “2024 Regional House Prices”.
3. Click Save and name your Story “Anticipated Home Buying Trends by MBA Graduates Using SAC”



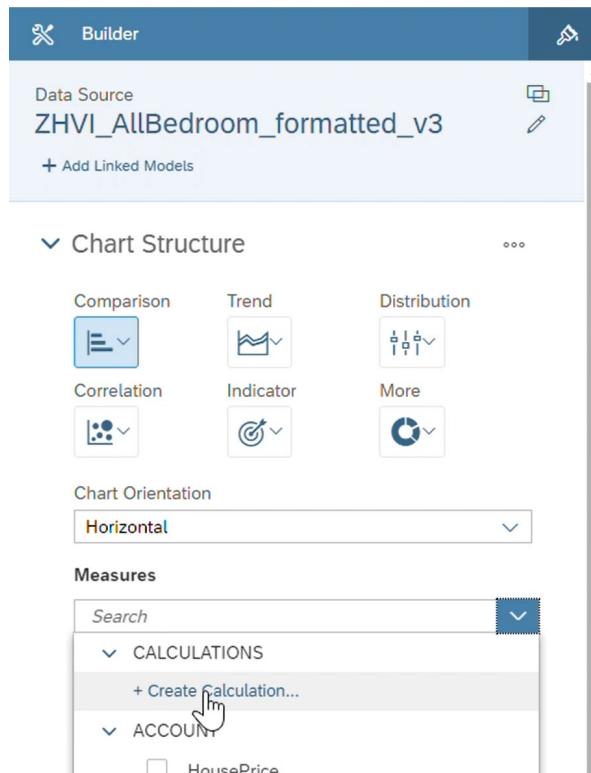
4. Title the **Lane** by clicking *Click to enter title...*
 - a. Name the Lane “2024 Regional House Price”



5. Create a Comparison Bar Chart for *2024 Regional House Price – 1 Bedroom* by selecting **Insert>Chart**.



- a. Create a Measure Calculation. In the **Builder** window, select **Measures>Calculations>+Create Calculation...**



- b. In the Calculation Editor pop-up, under **Type** select **Aggregation**.
- Under **Operation**, select *Average*
 - Under **Aggregation Dimensions**, select the following:
 - MetroName*
 - No of Bedrooms*
 - Name the calculation “Average_House_Price” and then select **OK**

Calculation Editor

Type: Aggregation Name: Average_House_Price

Properties

Operation: AVERAGE Measure: HousePrice

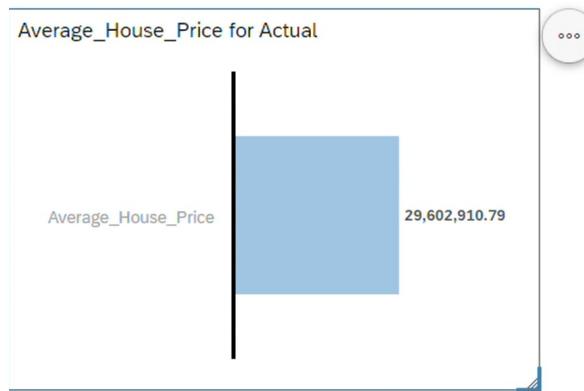
Aggregation Dimensions: MetroName, No of Bedrooms

+ Add a Dimension

Use conditional aggregation

OK Cancel

- c. The comparison bar chart should look like the below and have a price around 29,602,910.



- d. Under **Dimensions**, select Dimensions>MetroName>Region
e. Under **Filters**, select +Add Filters and add the following:

- i. **No of Bedrooms**, and in the pop-up check *1 Bedroom*

Set Filters for No of Bedrooms

Available Members	Selected Members
Show unbooked members <input type="checkbox"/>	1 Bedroom
Exclude selected members <input type="checkbox"/>	
<input checked="" type="checkbox"/> 1 Bedroom	
<input type="checkbox"/> 2 Bedroom	
<input type="checkbox"/> 3 Bedroom	

Clear Selection

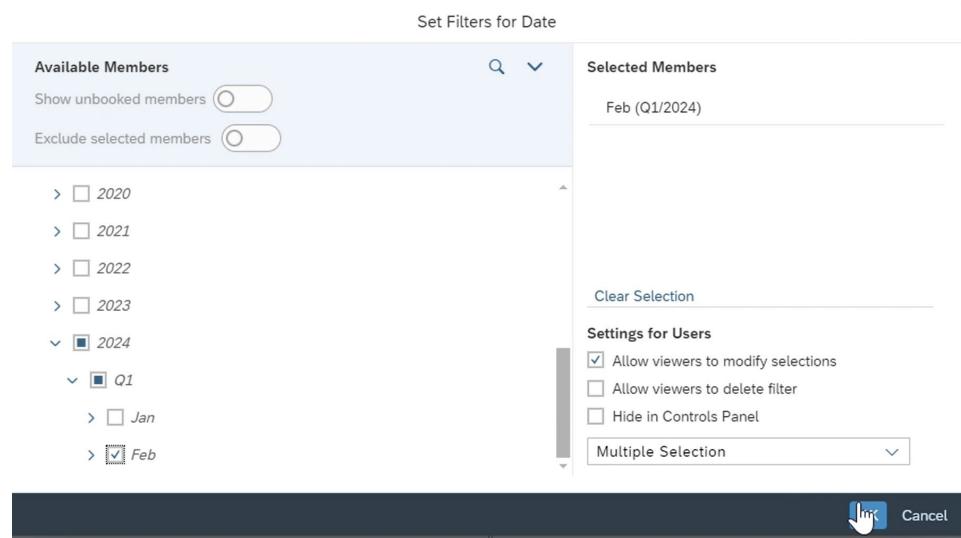
Settings for Users

Allow viewers to modify selections
 Allow viewers to delete filter
 Hide in Controls Panel

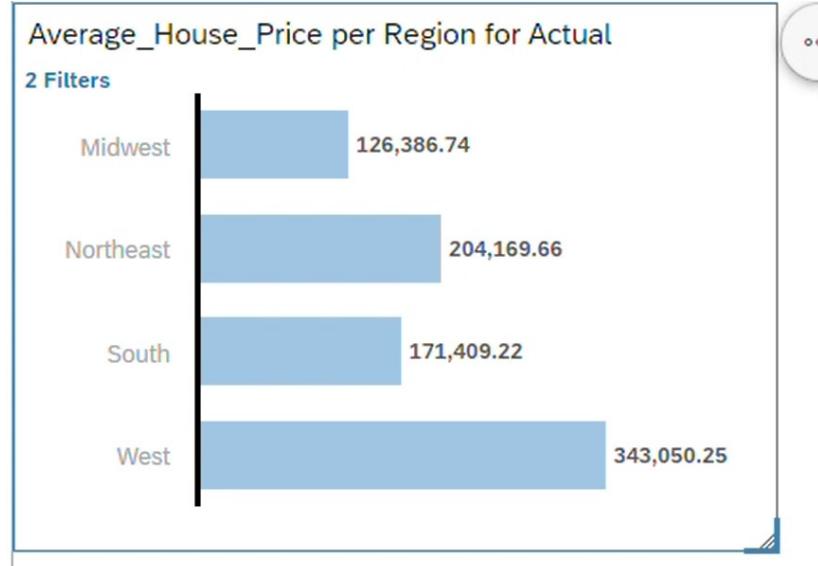
Multiple Selection

OK Cancel

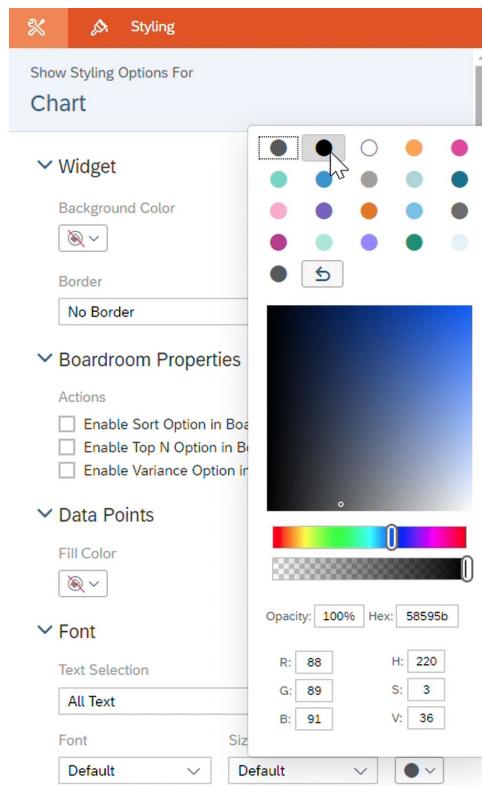
- ii. **Date (Member)**, and then check All>2024>Q1>Feb



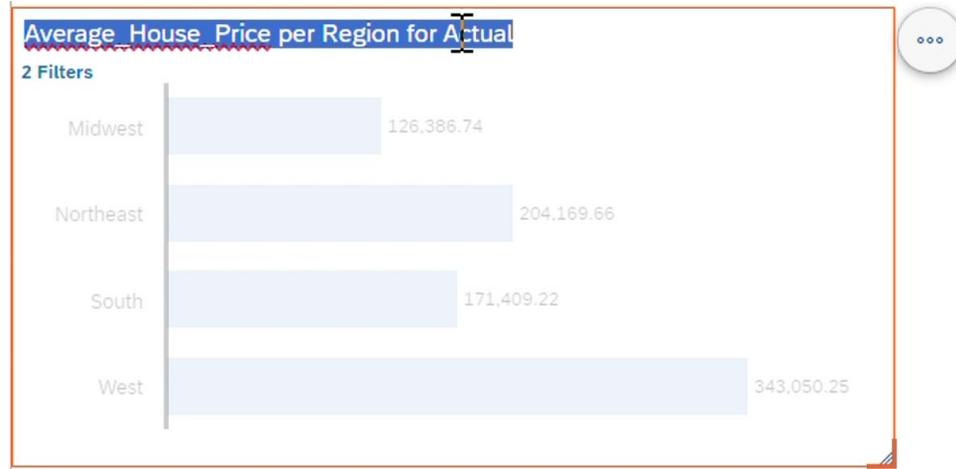
- iii. The bar chart should update to look like the below



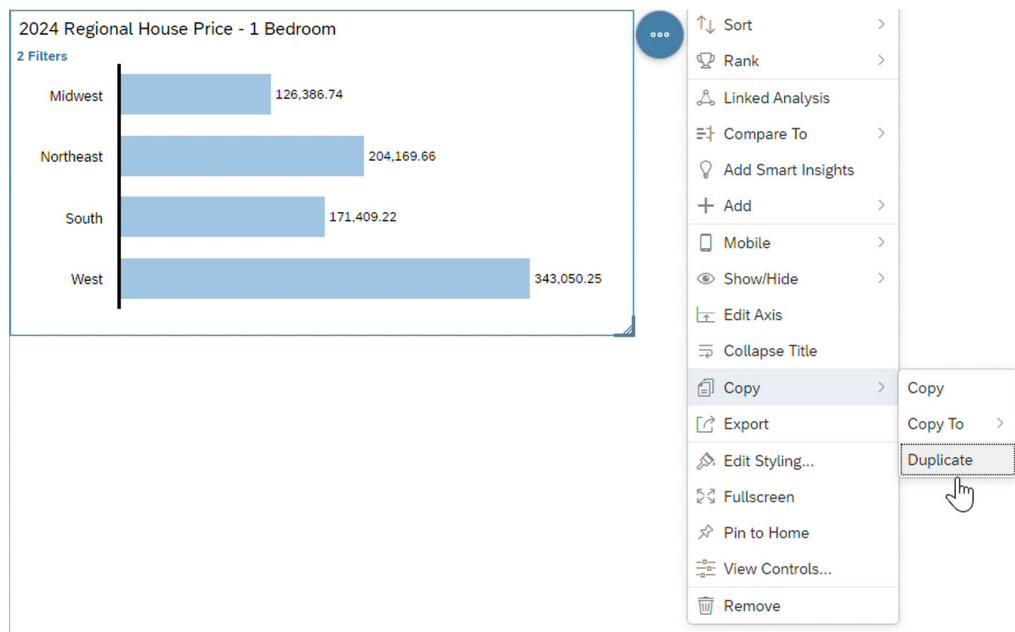
- f. Toggle to **Style**, change to **Font Color** by selecting the second option *Black*. NOTE: You will notice that the font color of the Regions (Midwest, Northeast, South, West) in the above bar chart will be darker and easier to read.



- g. Rename the bar chart by double-clicking the name and type “2024 Regional House Price – 1 Bedroom”



6. Create a second comparison bar chart of 2024 Regional House Price – 2 Bedrooms. Click the **Action** button (the icon with the three dots) next to the 2024 Regional House Price – 1 Bedroom.
- Select *Copy>Duplicate*



- b. In the **Builder** window, we will change the Filter for No of Bedrooms from **1 Bedroom** to **2 Bedrooms**.

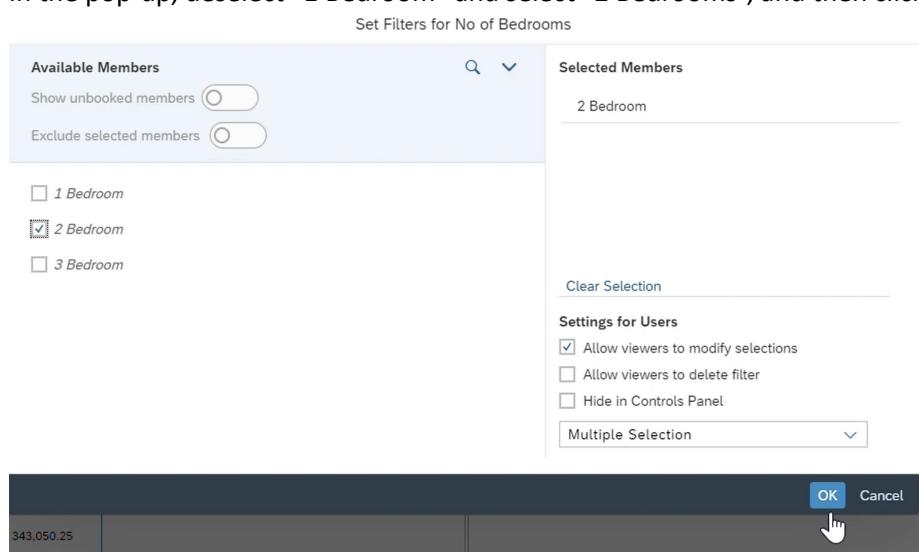
i. Double-click **Filters>No of Bedrooms**

Filters

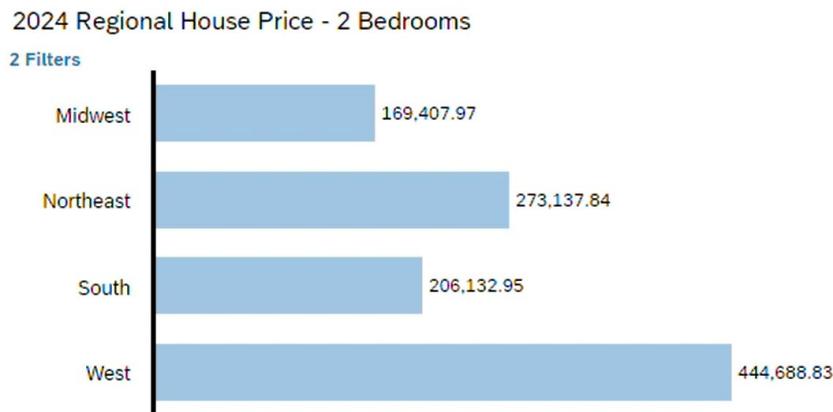
- Category (1)
- Actual
- Date (1)
- Feb (Q1/2024)
- No of Bedrooms (1)
- 1 Bedroom

+ Add Filters

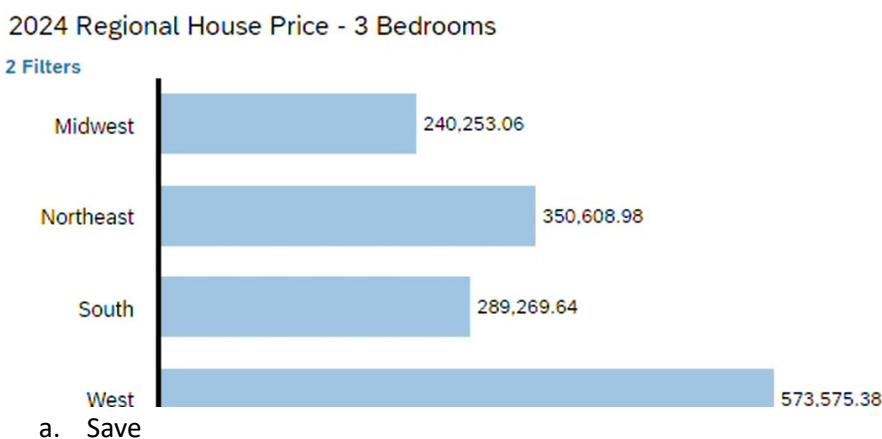
ii. In the pop-up, deselect “1 Bedroom” and select “2 Bedrooms”, and then click OK



- iii. Rename the chart to “2024 Regional House Price – 2 Bedrooms”
- iv. The resulting chart will look like the below image:



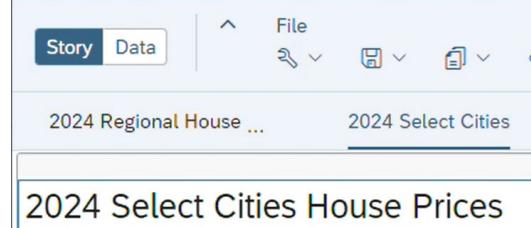
- 7. Create a third comparison bar chart for *2024 Regional House Price – 3 Bedrooms* by repeating **Step 5** above. The third chart will look like the below image:



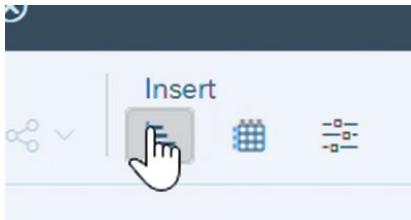
- a. Save

- 8. Add a second **Responsive Page**

- a. Rename **Page 1** as “2024 Select Cities”
- b. Rename **Lane** as “2024 Select Cities House Prices”



- 9. Create a Comparison Bar Chart for *2024 Select Cities House Prices – 1 Bedroom* by selecting **Insert>Chart**.

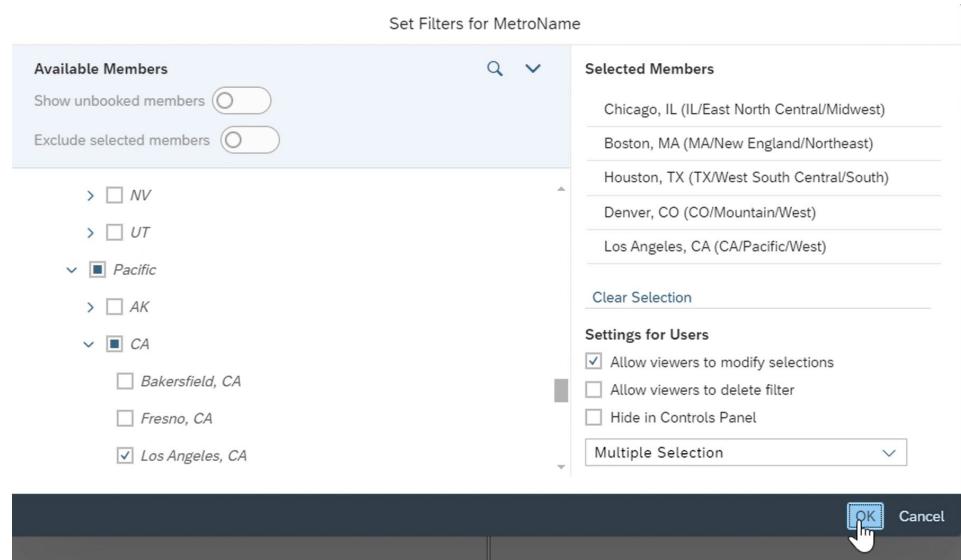


- a. In the **Builder** window, select the following:
- Chart Orientation>*Vertical*
 - Add Measures>*Average_House_Price*
 - Add **Filters**

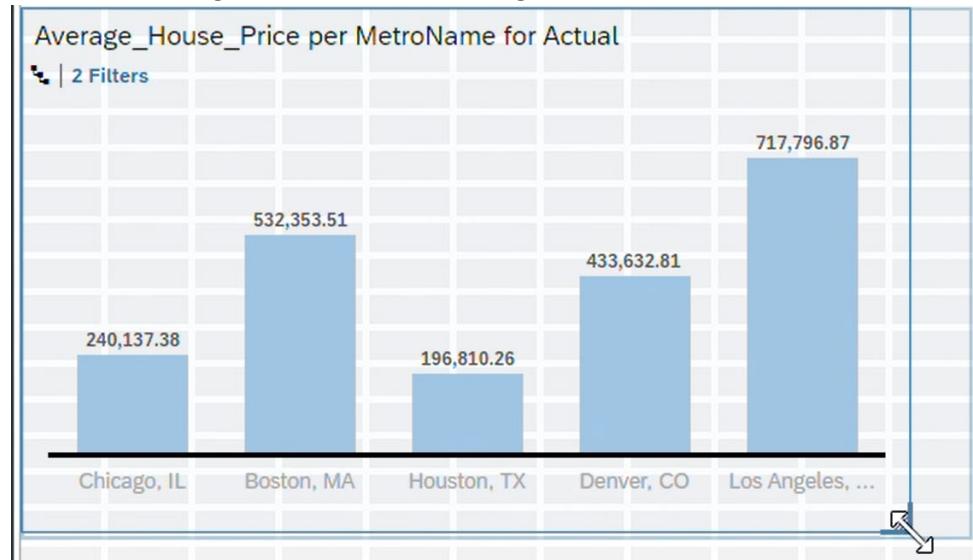
1. **MetroName**, in the Filter pop-up, select the following 5 MSAs:

A screenshot of the Power BI Builder window. The 'Measures' section shows 'Average_House_Price' selected. The 'Color' section shows a color palette with several colored dots. The 'Filters' section is expanded, showing a 'Category (1)' node with 'Actual' selected. A search bar is at the top of the filters list. Below it, there are options for 'Category', 'City', 'Date (Member)', 'Date (Range)', and a expanded 'MetroName' node which lists 'Midwest', 'Northeast', 'South', 'West', 'Mountain', and 'Pacific' categories.

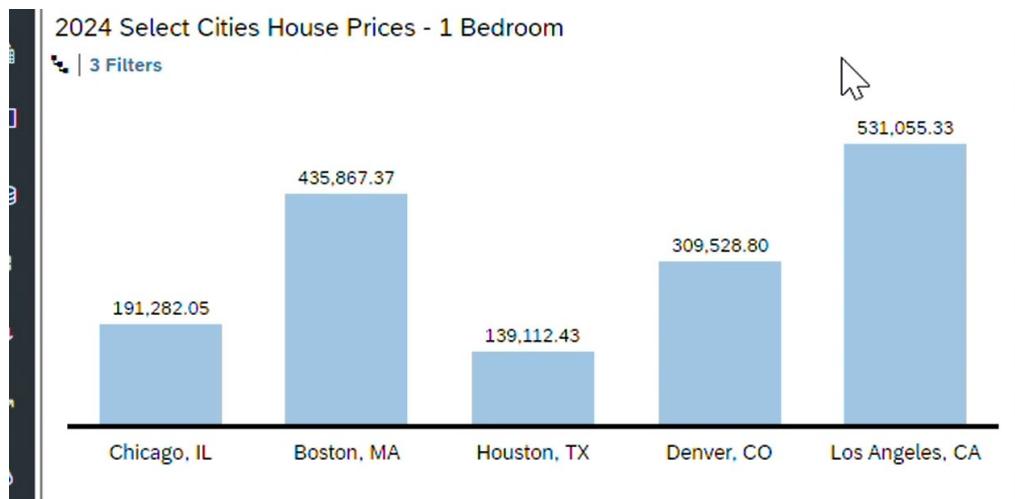
- Midwest>East North Central>IL>*Chicago, IL*
- Northeast>New England>MA>*Boston, MA*
- South>West South Central>TX>*Houston, TX*
- West>Mountain>CO>*Denver, CO*
- West>Pacific>CA>*Los Angeles, CA*



2. **Date (Member)**, and then check All>2024>Q1>Feb. Same step as we did in 4,e,ii.
- b. Resize the Chart by clicking the bottom right corner of the chart, and while the corner is still selected, drag the cursor over to the right.



- c. Toggle to **Style**, change to **Font Color** by selecting the second option *Black* – just like we did in Step 4,f.
- d. Rename the chart *2024 Select Cities House Prices – 1 Bedroom*

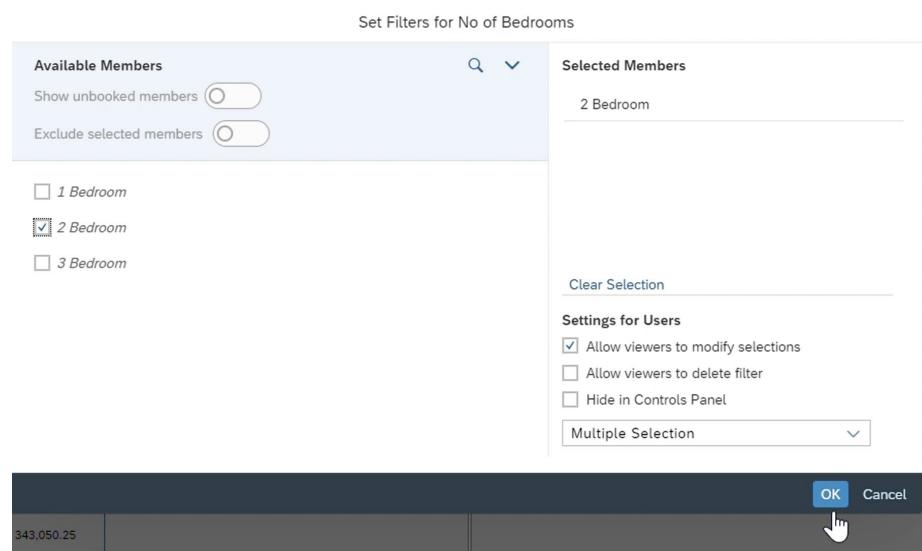


10. Create a second comparison bar chart of 2024 Select Cities House Price – 2 Bedrooms. This is similar to **Step 5**. Click the **Action** button (the icon with the three dots) next to the *2024 Select Cities House Price – 1 Bedroom*.

- Select *Copy>Duplicate*
- In the **Builder** window, we will change the Filter for No of Bedrooms from *1 Bedroom* to *2 Bedrooms*.
 - Double-click *Filters>No of Bedrooms*

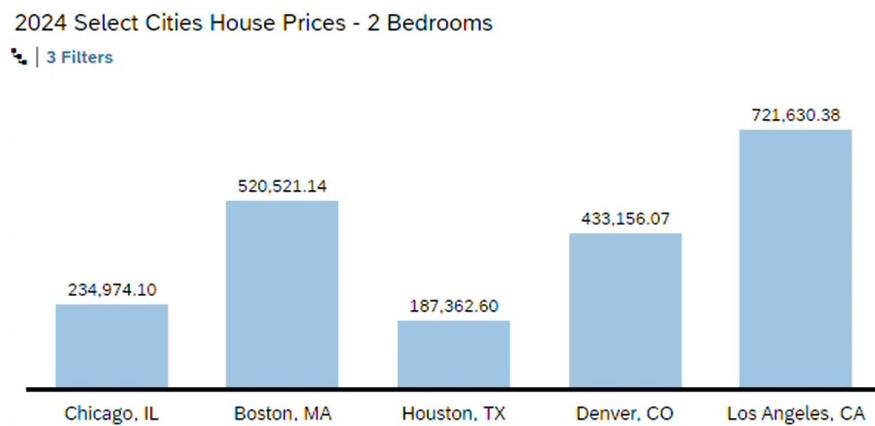


- In the pop-up window, deselect “1 Bedroom” and select “2 Bedrooms”, and then click OK



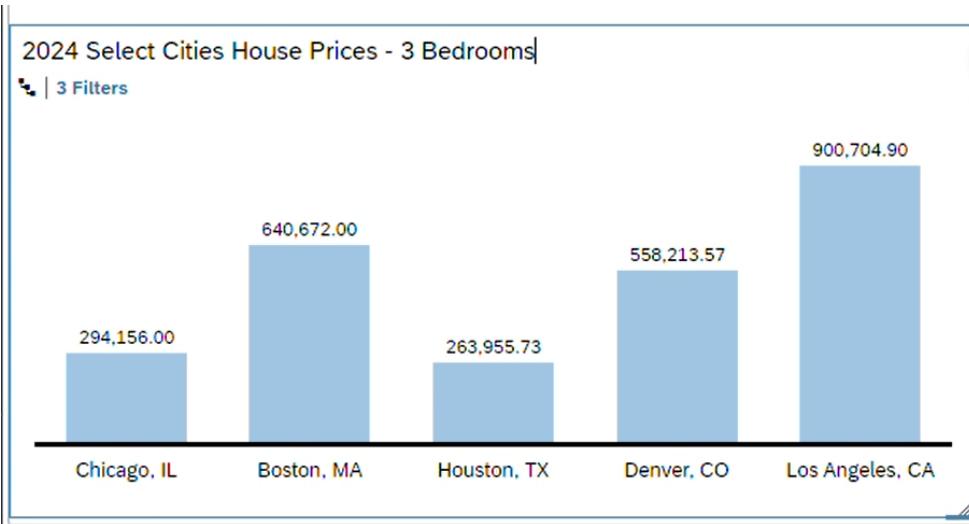
iii. Rename the chart to “2024 Select Cities House Price – 2 Bedrooms”

iv. The resulting chart will look like the below image:



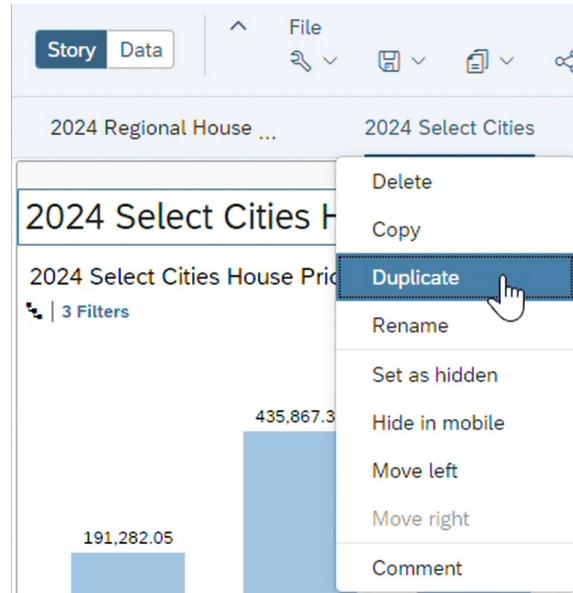
c. Save

11. Create a third comparison bar chart for *2024 Select Cities House Price – 3 Bedrooms* by repeating **Step 9** above. The third chart on the page will look like the below image:

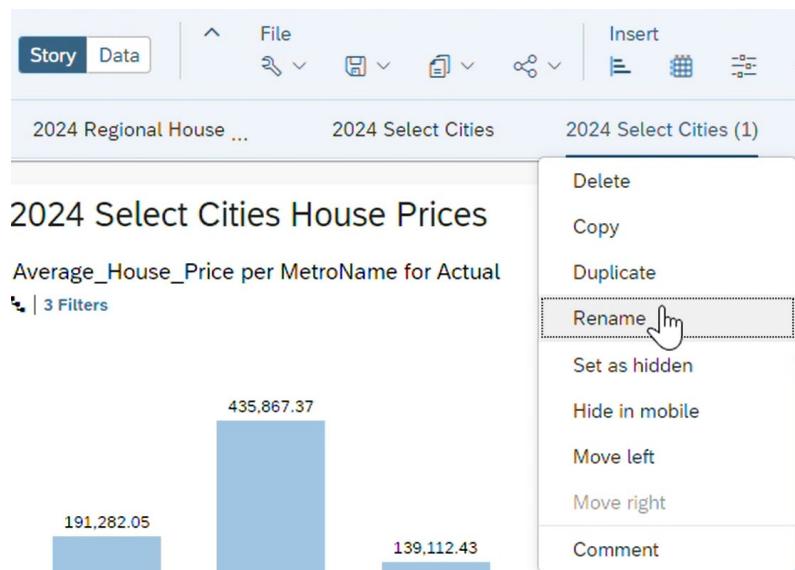


12. Create the **Page** for the *2014 Select Cities*.

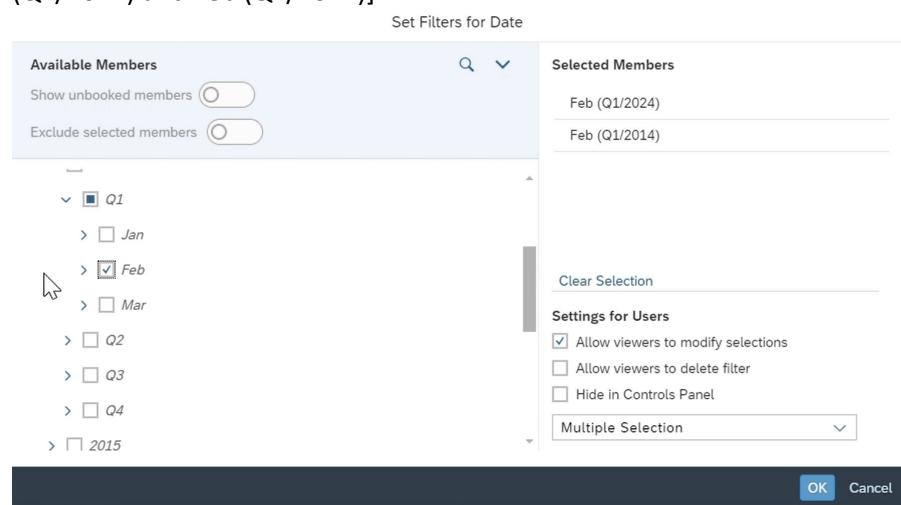
- First, you want to **Duplicate Page** *2024 Select Cities House Prices* by clicking the downward chevron next to the page's title (looks like a "v") and select **Duplicate**.



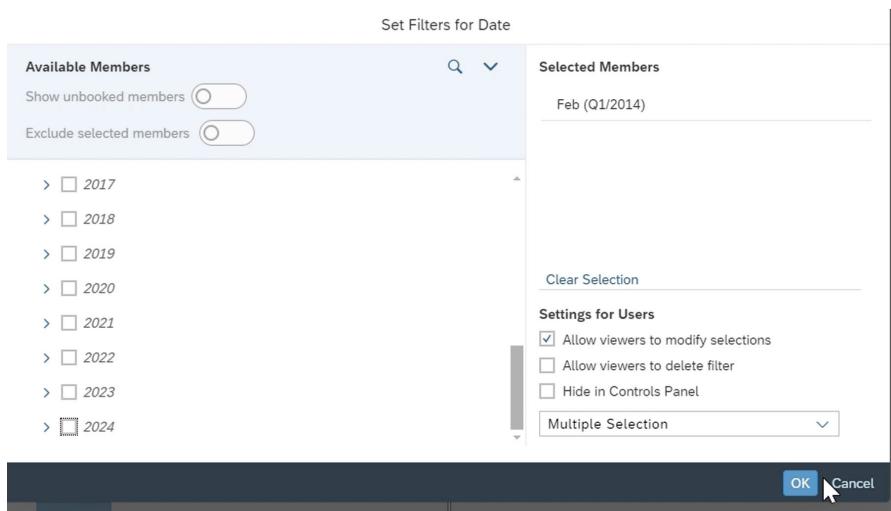
- Rename the duplicate page to “2014 Select Cities”



- c. Rename the Lane Title to “2014 Select Cities House Prices”
13. Select the top chart and in the Builder window, edit Filter>Date(1)
- Select All>2014>Q1>Feb. Note, you will see two Selected Members [Feb (Q1/2024) and Feb (Q1/2014)]

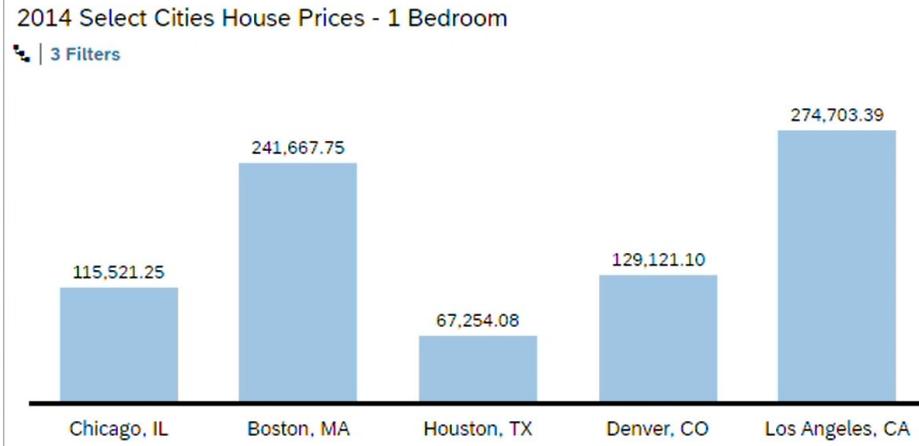


- Deselect All>2024>Q1>Feb. Note: you will now only see one value in the Selected Members section [Feb (Q1/2014)]



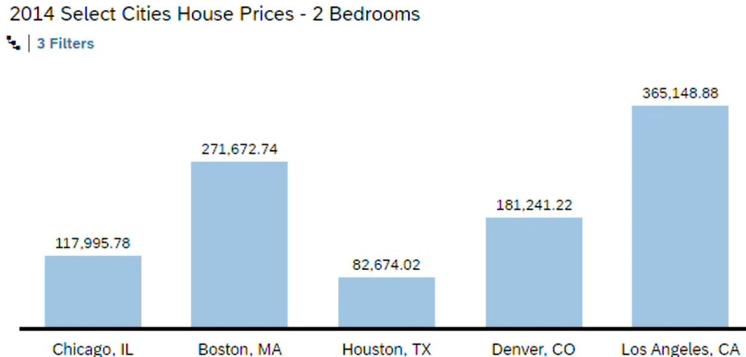
iii. Click OK

- Rename the chart “2014 Select Cities House Prices – 1 Bedroom”
- The top chart will look like the below image:



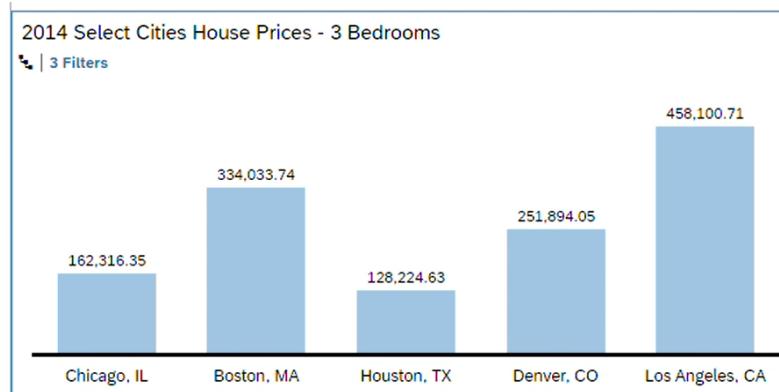
14. Repeat Step 12 for the middle chart

- Rename the chart “2014 Select Cities House Prices – 2 Bedrooms”
- The chart will look like the below image:



15. Repeat Step 12 for the bottom char

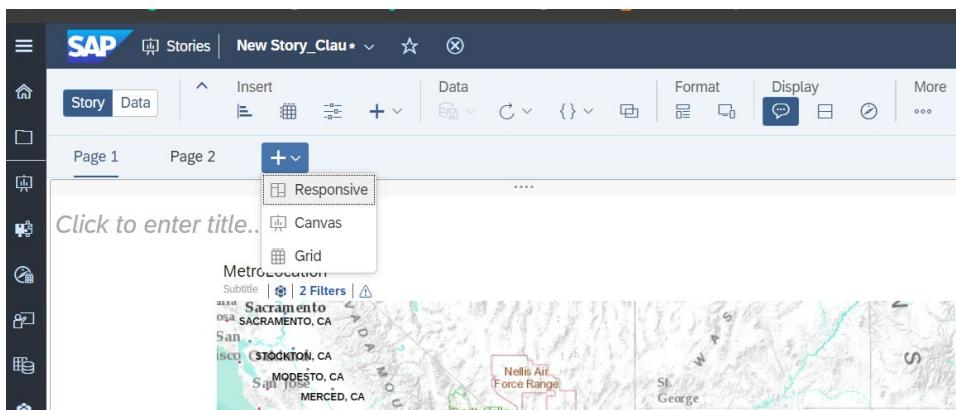
- Rename the chart “2014 Select Cities House Prices – 3 Bedrooms”
- The chart will look like the below image:



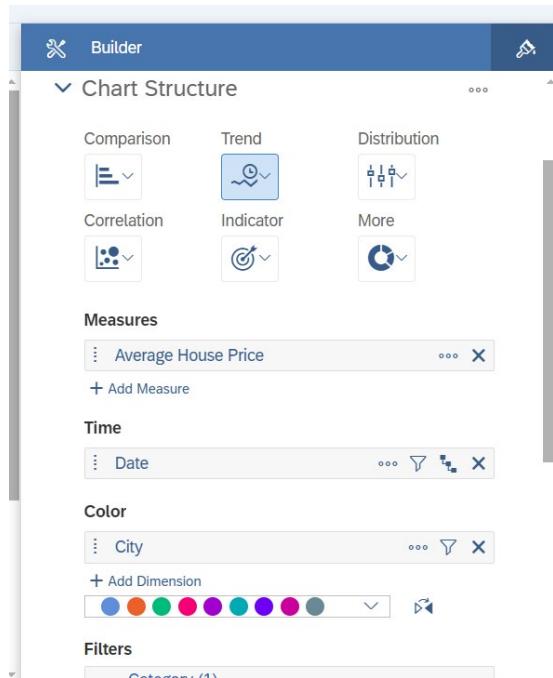
16. Save.

Timeseries Analysis

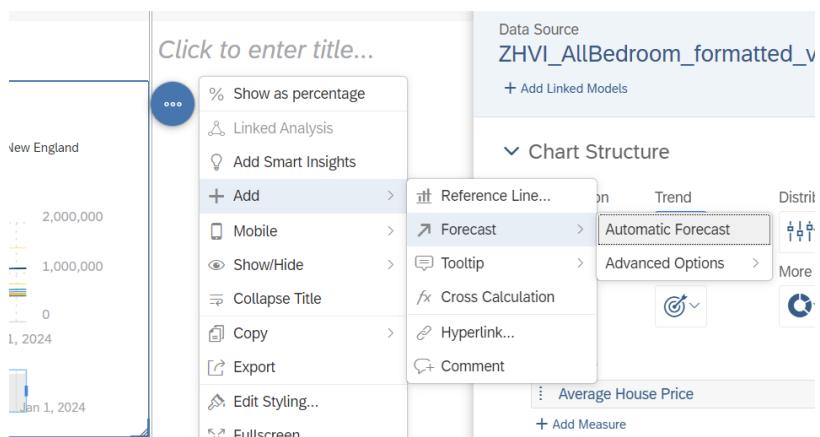
1. Add a new page and rename to “Timeseries Analysis”:



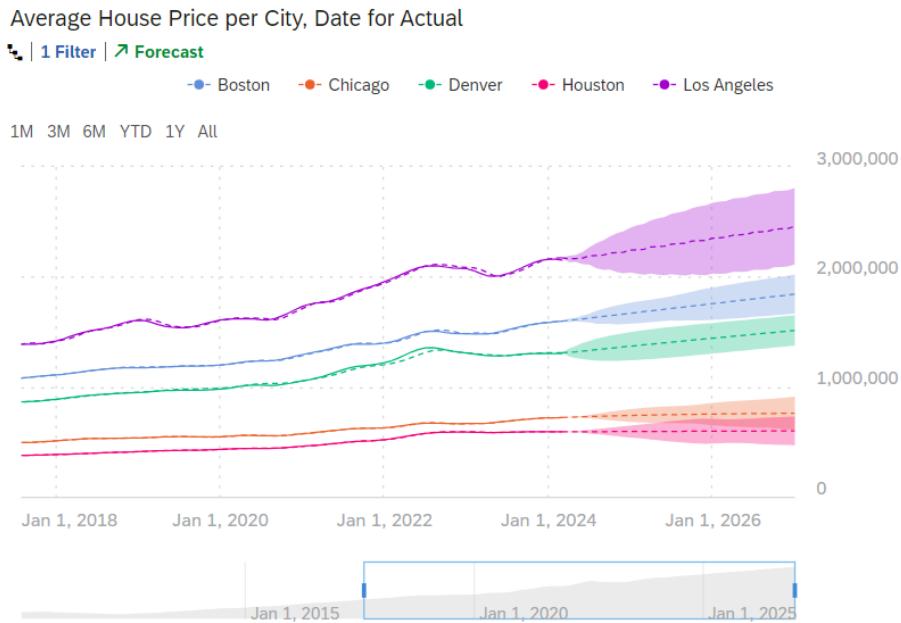
2. Insert a new chart. Extend the chart to the height and width of the lane.
3. In Builder:
 - a. Under Chart Structure, select **Trend**, then **Time Series**
 - b. In Measures, select **Average House Price**
 - c. In Time, select **Date**
 - d. Under Color, select **Add Dimension**, then select **City**
 - e. Filter the cities to: Boston, Chicago, Denver, Houston, and Los Angeles
 - f. Choose what colors you would like to use.



4. In the Chart options, hover over Add, then Forecast. Select **Automatic Forecast**.



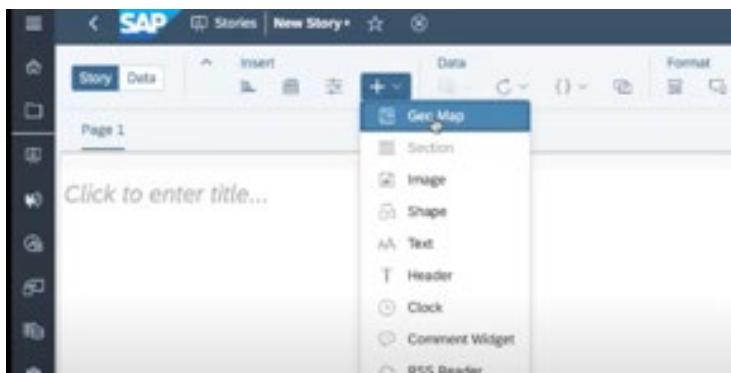
5. Your chart should look like the below:



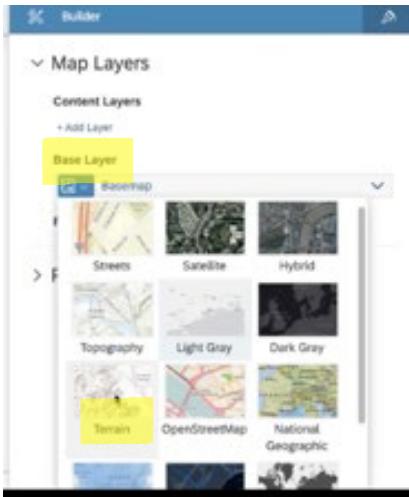
6. Click Save.

GeoMap

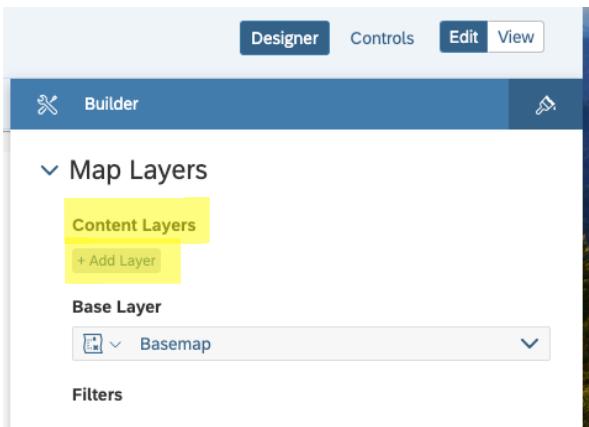
1. Add a new page and rename to “1 Bedroom Average Price”:



2. Select **Geo Map** from the + section.



3. In **Builder** select **Base Layer** and pick **Terrain**



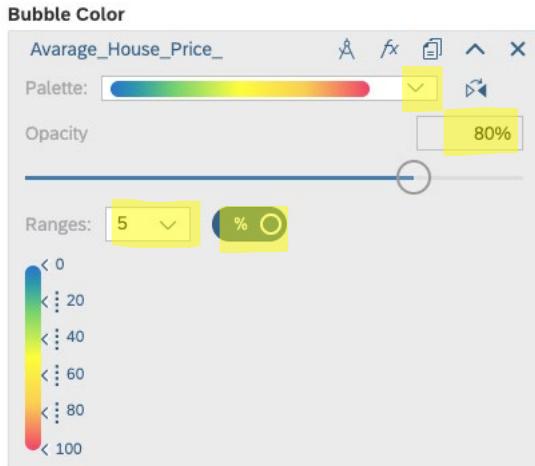
4. Select **Content Layers** and **add a layer**

The screenshot shows the Tableau Builder interface for creating a bubble layer. The top bar has a 'Builder' tab. Below it, 'Layer 2' is selected, indicated by a blue border. The 'Data Source' is set to 'ZHVI_AllBedroom_Final'. Under 'Layer Type', 'Bubble Layer' is chosen. In the 'Location Dimension' section, 'MetroName_Location' is highlighted with a yellow box. Other options like 'Search' and '---' are also visible. The 'Bubble Size' and 'Filters' sections are partially visible below. A link to 'Cluster Properties' is at the bottom.

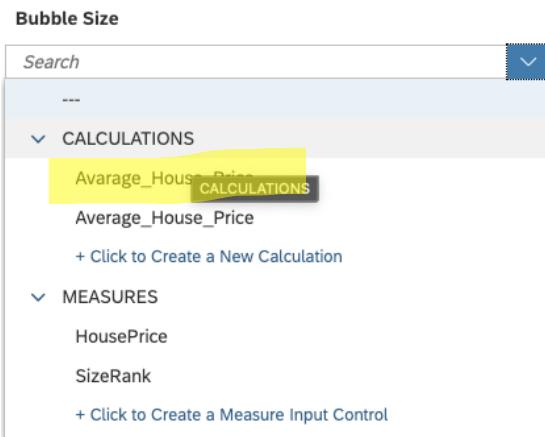
5. Select **Location Dimension** and choose **MetroName_Location**

This screenshot shows the continuation of the Tableau Builder interface. The 'Location Dimension' dropdown now shows 'MetroName_Location' selected. In the 'Bubble Color' section, 'Avarage_House_Price_' is highlighted with a yellow box. Other options like 'Search', 'CALCULATIONS' (which includes 'Average_House_Price'), and 'MEASURES' are visible. The 'Builder' tab is still at the top.

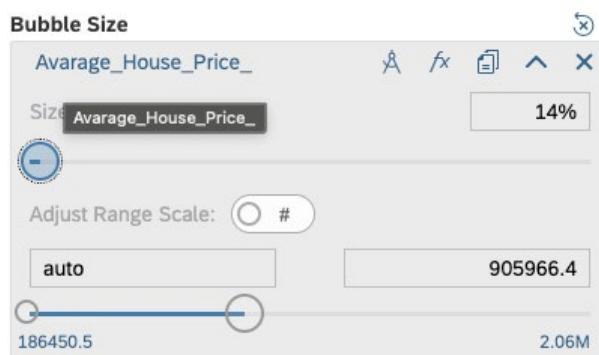
6. Under **Bubble color** select **calculation** and select **Average_House_Price_**



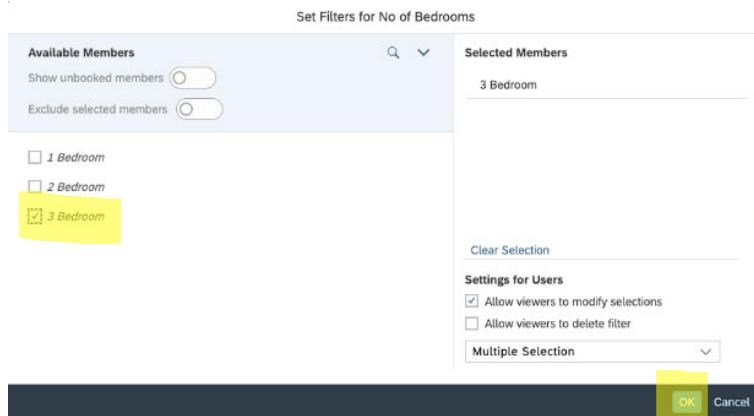
- a. Change the palette color to one of your liking and ensure that all option match the image above.



7. Under **Bubble size** select **calculation** and select **Average_House_Price**

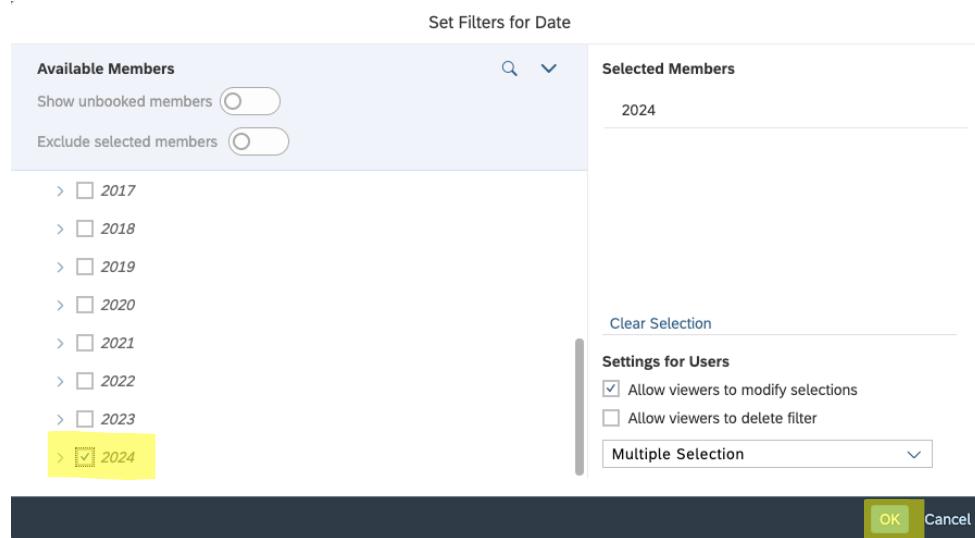


8. Select the **Bubble size** to **14%** and the price range to **900,0000**



9. Select **Filters** and select No Bedroom.

b. You can select **3 bedrooms** and press ok



10. Select **Filters** again and select **Date (member)**

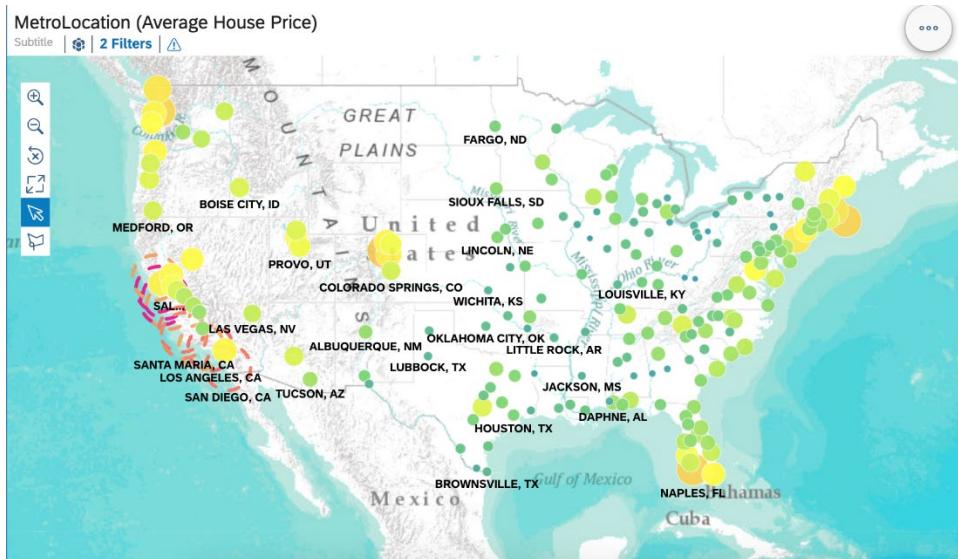
c. Select **2024** for the current year

The screenshot shows the Tableau Builder interface with the following configuration:

- Layer 2** is selected.
- Data Source**: ZHVI_AllBedroom_Final
- Layer Type**: Bubble Layer
- Location Dimension**: MetroName_Location
- Bubble Color**: Average_House_Price_
- Bubble Size**: Average_House_Price_
- Filters**:
 - No of Bedrooms (1) - 3 Bedroom
 - Date (1) - 2024
 - Category (1) - Actual
- Cluster Properties** button is visible.

11. Your **Builder** should look like the image above.

- Select OK
- Repeat steps 1-10 and recreate for 1-bedroom and 2-bedroom to get the average price in the geo map.



12. Click **Save**

Predictive Model: Regression Analysis

1. Create the Regression Analysis Model

The screenshot shows the SAP Predictive Scenarios application interface. At the top, there is a navigation bar with icons for search, ideas, notifications (with 23), and user profile. Below the header, a welcome message says "Welcome to Predictive Scenarios" and provides a brief description of creating predictive scenarios for business questions.

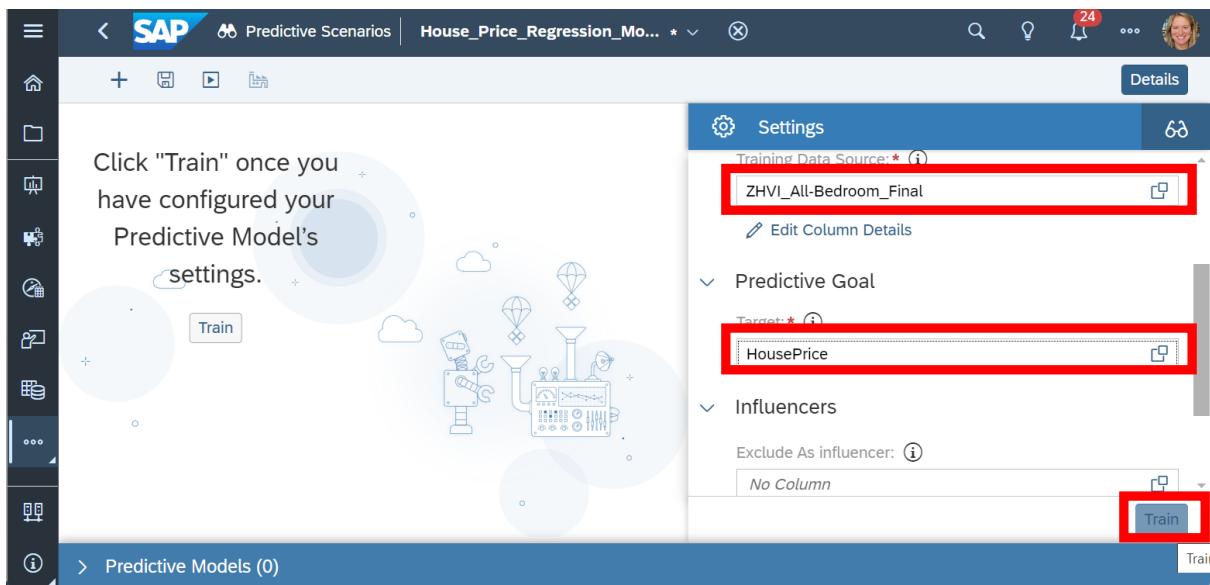
In the center, there is a "Create New" section with three options: "Classification" (represented by a tree icon), "Regression" (represented by a scatter plot icon, which is highlighted with a red box), and "Time Series Forecast" (represented by a line graph icon).

On the left side, there is a vertical navigation menu with several icons. Some of these icons are highlighted with red boxes: "Data Actions", "Allocations", and "Predictive Scenarios". The "Predictive Scenarios" icon is specifically highlighted with a red box and is also highlighted in the "Create New" section above.

A search bar at the top right says "Search In: Recent Files". Below the search bar, a message states "No recent files yet. Open a file and it will show up here."

At the bottom left, there is a label "A." followed by a red box.

2. Click on the left menu to select Predictive Scenarios, then choose Create New > Regression
3. Choose your folder to save your regression model, and name it, "House_Price_Regression_Model"
4. Select OK
5. Click under Training Data Source in Settings, and select the existing dataset "ZHVI_All-Bedroom_Final"
6. Click Train
7. This could take some time to complete, wait for the process to finish



You will see the Global Performance indicators as in the below:

The Influencer Contributions, or those variables that have the greatest influence on House Price, is shown below:

SAP Predictive Scenarios | House_Price_Regression_Model

Influencer Contributions

Influencer	Contribution
Date_Y	20.88%
No of Bedrooms	17.39%
Longitude	16.77%
Region	12.67%
RegionID	9.12%

Settings

Training Data Source: * ZHVI_All-Bedroom_Final

Predictive Goal

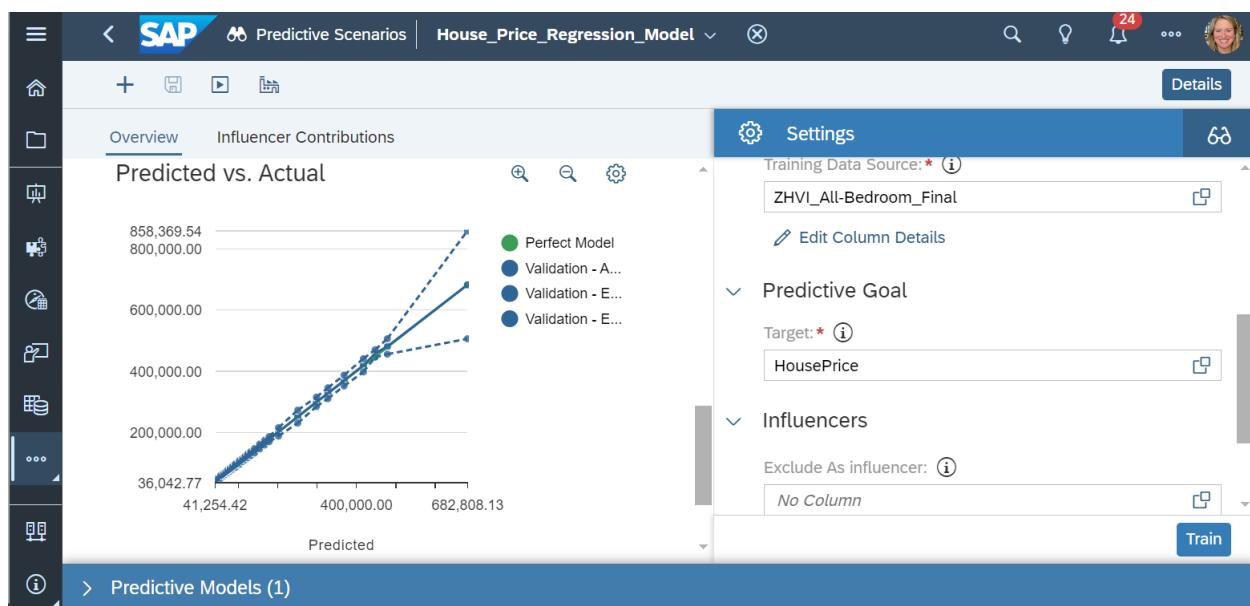
Target: * HousePrice

Influencers

Exclude As influencer: No Column

Train

The Predicted House Price vs. Actual House Price is shown below:



Click Save.

You have completed your Regression Model. Congratulations!

This is the end of Lab Instructions.