

The goal of this exercise:

Create a Power BI report that analyzes wine review data. Scores, pricing, points, and other contextual information about thousands of wines worldwide are included in the data collected from WineEnthusiast (Kaggle).

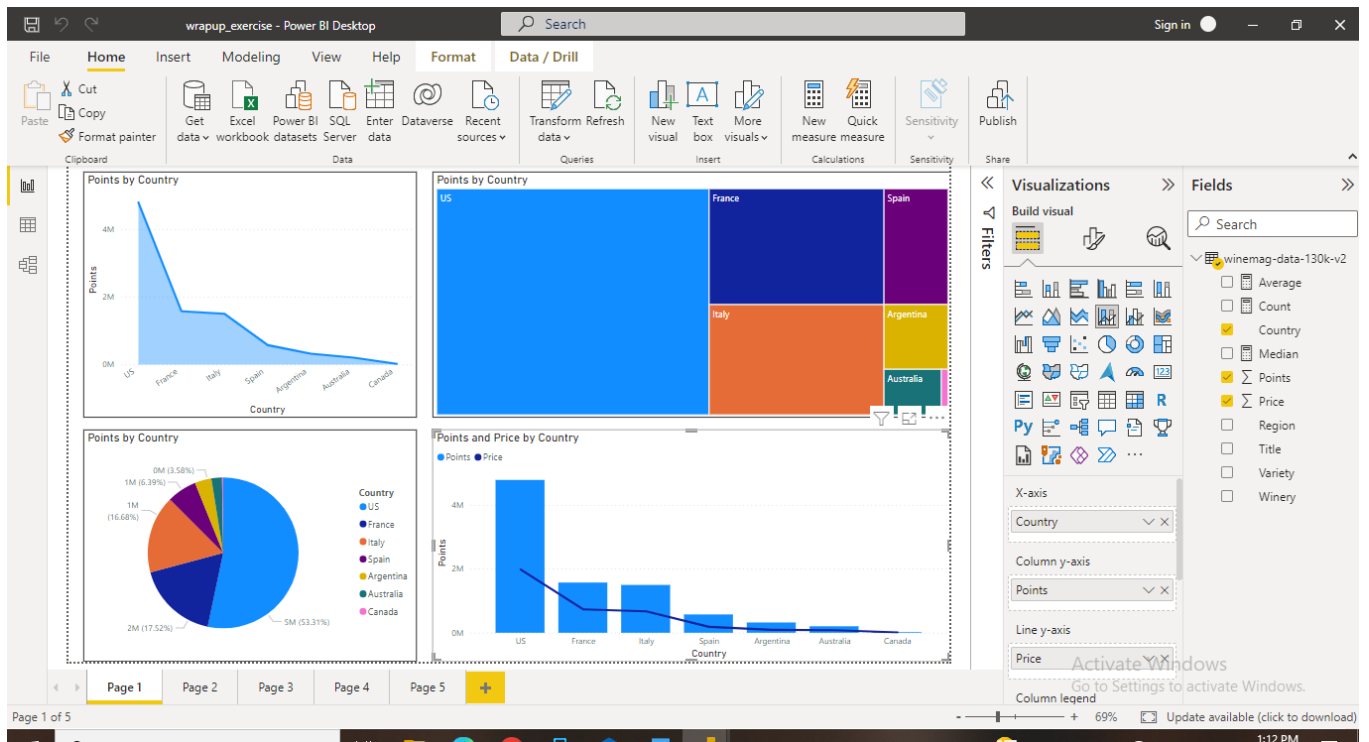
- Find out which wineries, countries, or regions have the highest points rating or which varieties have the highest cost.
- Try using a map visual to show the data, such as average points per country.

Since this was raw data, I first transformed it before loading. Then, I choose the columns (Country, Points, Price, Region, Variety, Winery) I want to keep to get my answers by selecting them in Power Query Editor. I also renamed wherever I felt necessary.

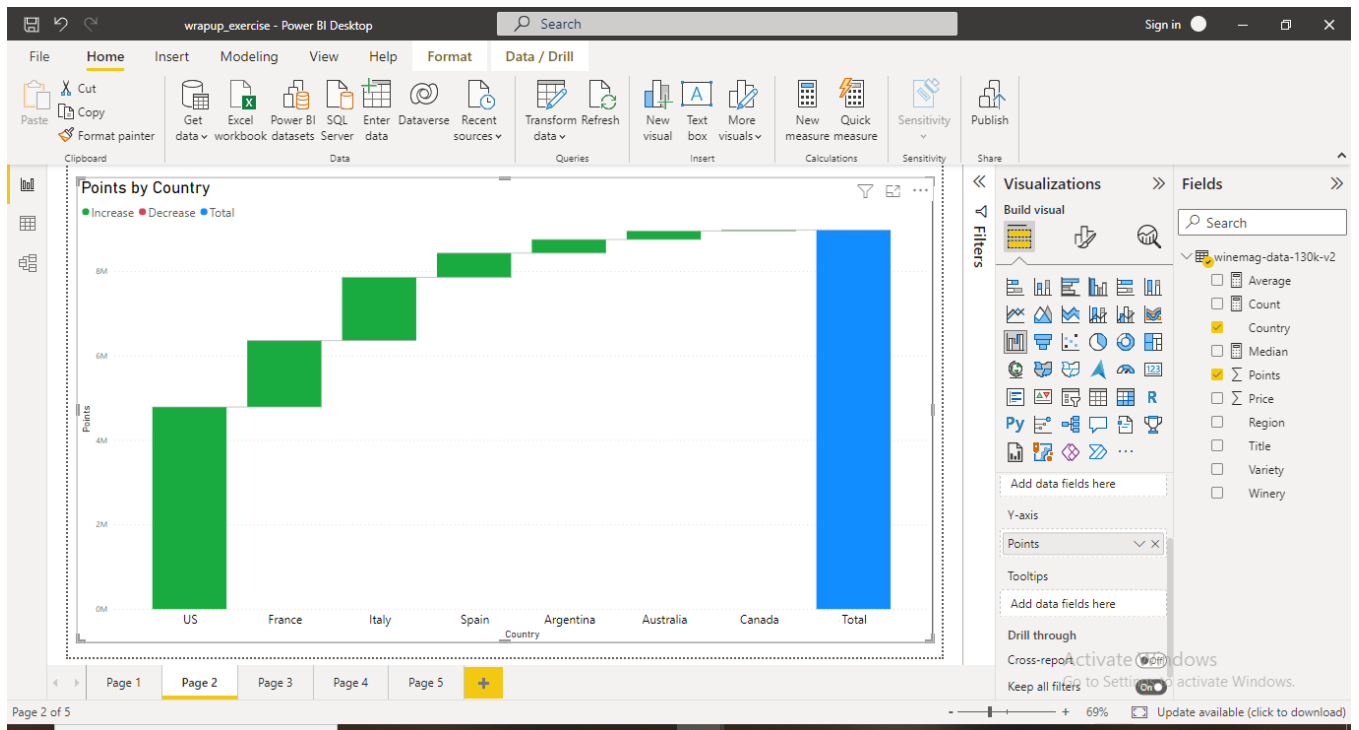
The screenshot displays the Power Query Editor interface. The main area shows a table with 24 rows and 5 columns: Country, Points, Region, Variety, and Winery. The data includes wine reviews from various countries like Italy, Portugal, US, Spain, France, Germany, Argentina, and the US, with specific regions and wineries listed. The right-hand pane shows the 'Query Settings' for 'winemag-data-130k-v2', including properties like Name and Applied Steps. The bottom status bar indicates '999+ ROWS' and 'Column profiling based on top 1000 rows'.

| | Country | Points | Region | Variety | Winery |
|----|-----------|--------|------------------------|--------------------|---------------------|
| 1 | Italy | | 87 Etna | White Blend | Nicosia |
| 2 | Portugal | | 87 | Portuguese Red | Quinta dos Avidagos |
| 3 | US | | 87 Willamette Valley | Pinot Gris | Rainstorm |
| 4 | US | | 87 Lake Michigan Shore | Riesling | St. Julian |
| 5 | US | | 87 Willamette Valley | Pinot Noir | Sweet Cheeks |
| 6 | Spain | | 87 Navarra | Tempranillo-Merlot | Tandem |
| 7 | Italy | | 87 Vittoria | Frappato | Terre di Giurfo |
| 8 | France | | 87 Alsace | Gewürztraminer | Trimbach |
| 9 | Germany | | 87 | Gewürztraminer | Heinz Eifel |
| 10 | France | | 87 Alsace | Pinot Gris | Jean-Baptiste Adam |
| 11 | US | | 87 Napa Valley | Cabernet Sauvignon | Kirkland Signature |
| 12 | France | | 87 Alsace | Gewürztraminer | Leon Beyer |
| 13 | US | | 87 Alexander Valley | Cabernet Sauvignon | Louis M. Martini |
| 14 | Italy | | 87 Etna | Nerello Mascalese | Masseria Setteporte |
| 15 | US | | 87 Central Coast | Chardonnay | Mirassou |
| 16 | Germany | | 87 | Riesling | Richard Bäck |
| 17 | Argentina | | 87 Cafayate | Malbec | Felix Lavaque |
| 18 | Argentina | | 87 Mendoza | Malbec | Gaucha Andino |
| 19 | Spain | | 87 Ribera del Duero | Tempranillo Blend | Pradorey |
| 20 | US | | 87 Virginia | Meritage | Quail@vremont |
| 21 | US | | 87 Virginia | Red Blend | Quail@vremont |
| 22 | US | | 87 Oregon | Pinot Noir | Acrobat |
| 23 | Italy | | 87 Sicilia | White Blend | Baglio di Pianetto |
| 24 | US | | 87 Paso Robles | Merlot | Bianchi |

Next, on Page 1, I created some basic visuals for my data. This includes an area chart, Pie Chart, Tree map, and Line and Stacked column chart. These charts show Points/Prices by Country.



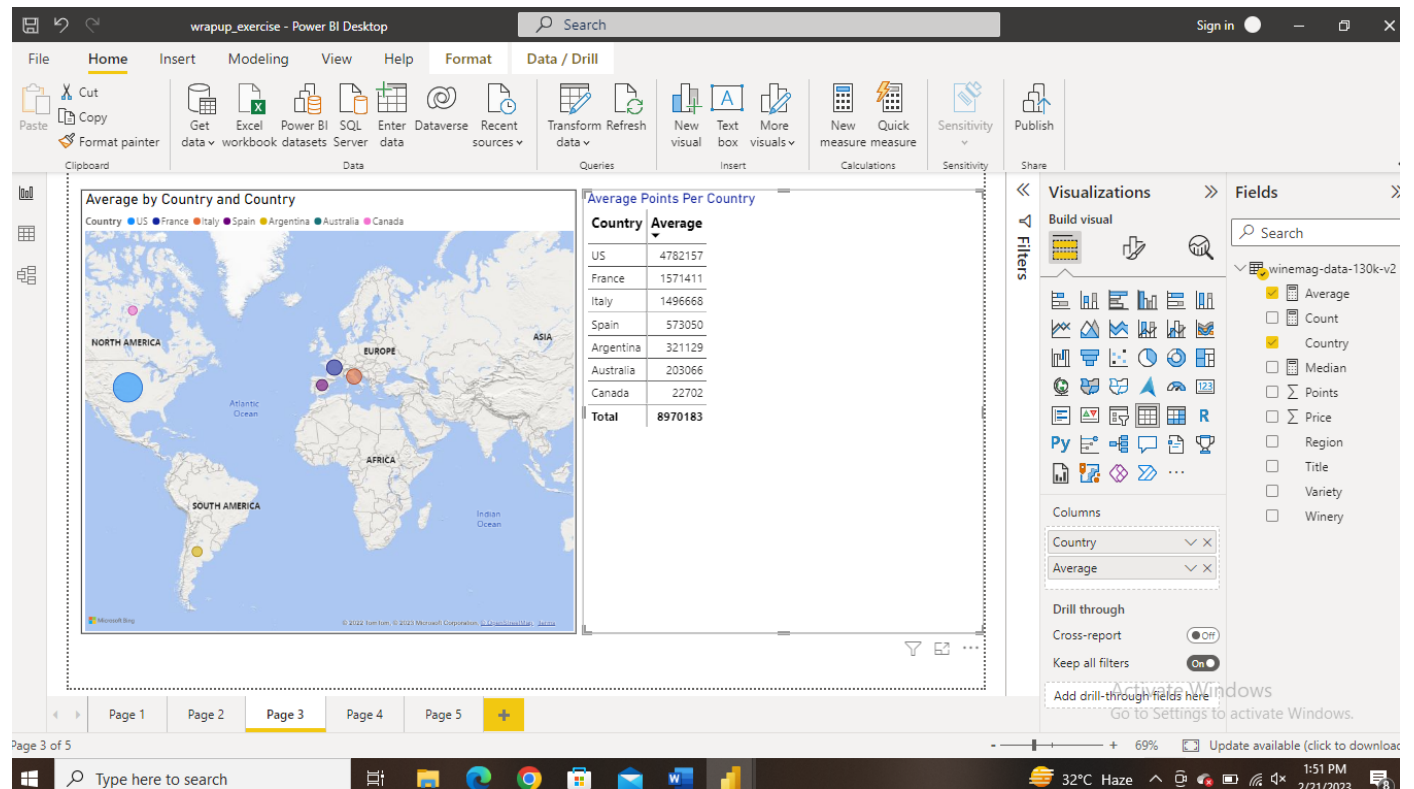
On page 2, I created another visual – waterfall chart to show Points by Country. The visual demonstrates that the US has the highest points (4782157) and Canada has the lowest (22702).



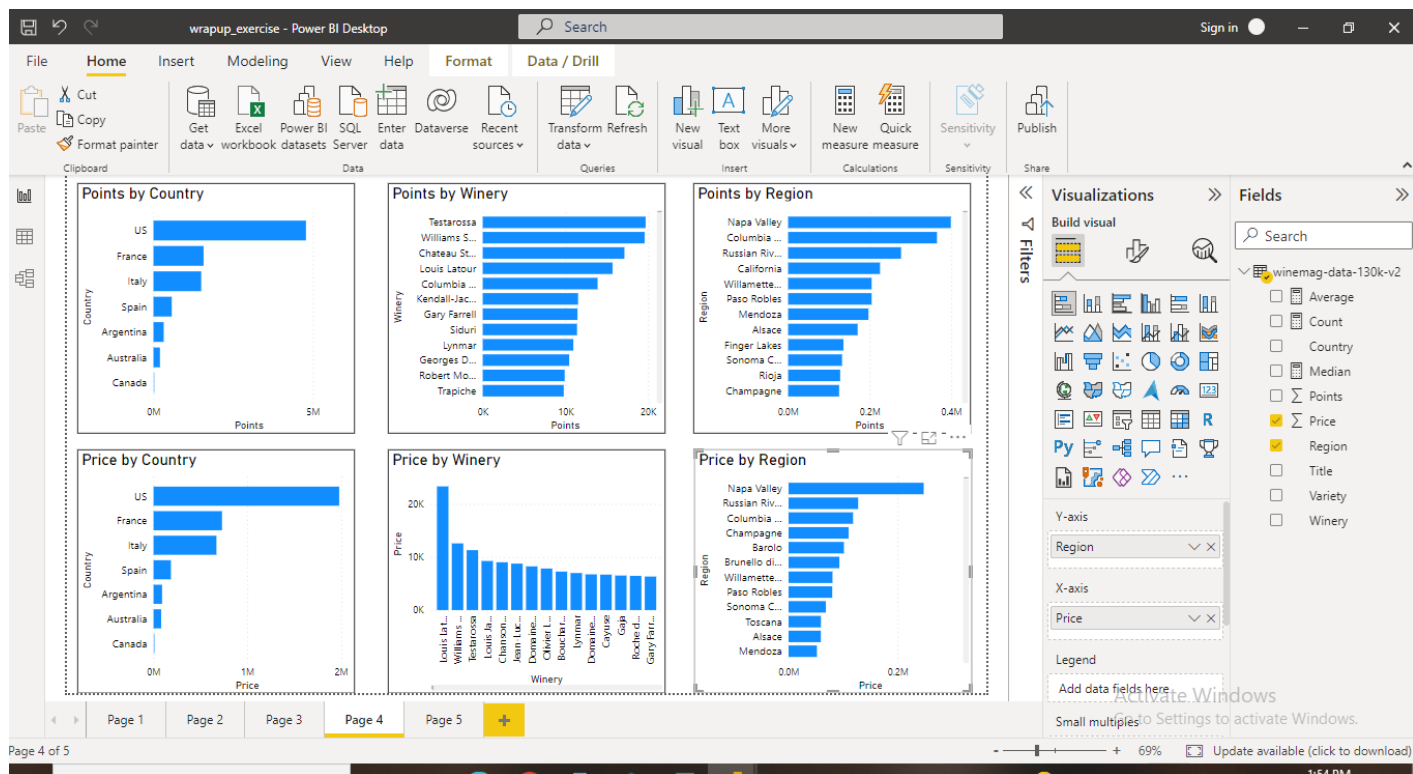
On Page 3, I created a new measure called Average using DAX. I used the formula mentioned below:

Average = `sum('winemag-data-130k-v2'[Points])`

Next, I created two visuals: a map and a table to show the average Points Per Country.



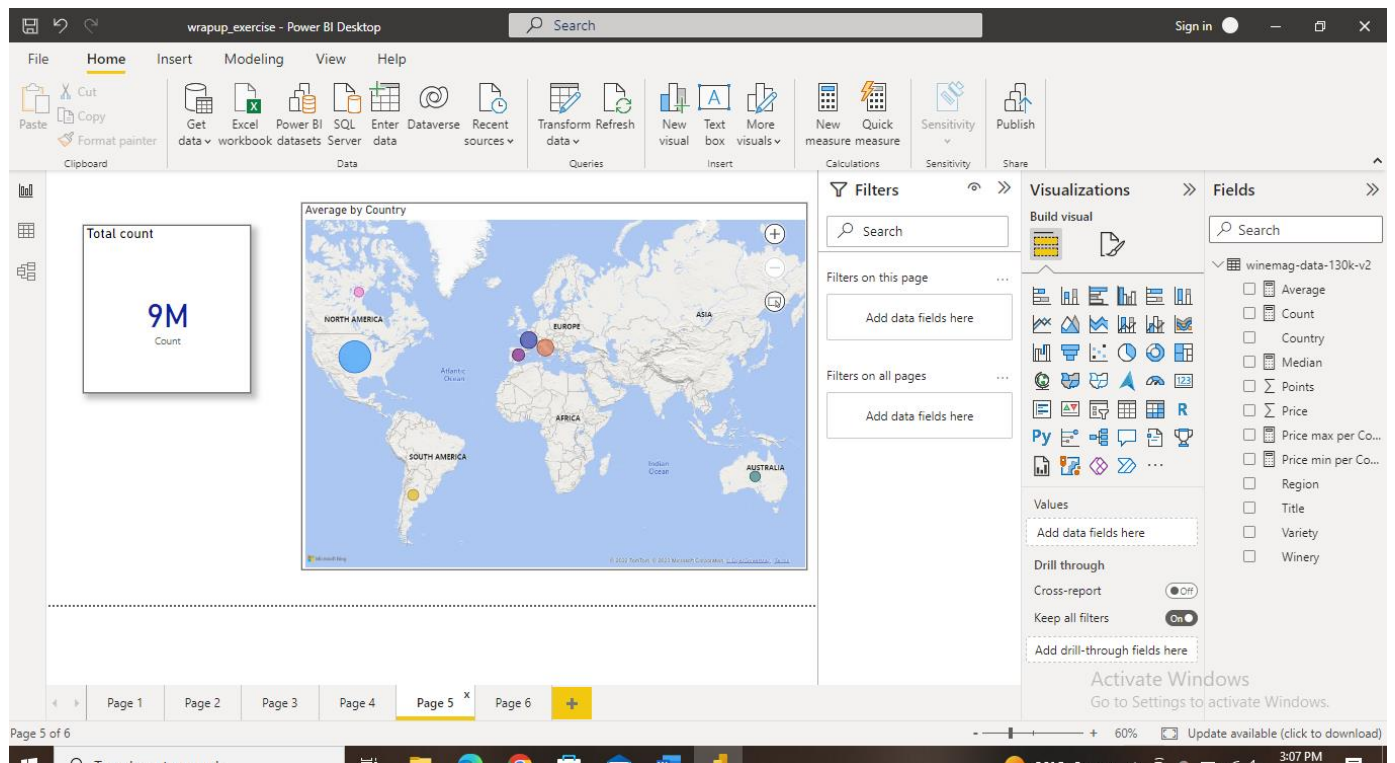
On the next page, since I wanted to know which countries, Wineries, and Regions have the highest points and Prices, I created a few more visuals with Stacked bar charts.



On page 5, I wanted to know quickly which country has the highest sales and also the total sales. For this, I created a number card and a map.

And to display the total Count, I created the explicit measure - **Count** that displays there are 9M rows in the Points column. For Count, I wrote the below DAX formula:

Count = ([Average])



Lastly, I wanted to find out the min., max., and median values of Price by Country. So I created Quick measures for min and max, whereas, for the median, I created a new measure. And that's the result: the US having the max price of 1978668 in the card was rounded off to 2M. Min price was of Canada; 9071.

