

Milestone 2: App Write Up

Ted Haley

Design Rationale

Data Wrangling

The wine magazine dataset includes 150,000 different wines from around the globe. Each entry includes the country of origin (country), a brief description of the wine that describes that taste (description), the winery designation of the series of wine (designation), a rating out of 100 (points), a price in USD per bottle (price), province in which the wine was made (province), sub-regional area of each province (region1 and region2), the twitter handle for the individual who rating the wine (tastertwitterhandle), the name of the individual who reviewed the wine (taster_name), the type of grape used to make the wine (variety), and the winery that produced the wine (winery). To make the dataset more usable and responsive, I removed rows that included any missing data. Additionally, the tool I am making is intended to be used by consumers, and as such, I wanted to only include wines that a realistic consumer would purchase, so I removed wines costing more than \$500 per bottle and I limited to the variety of wines to non-blended, commercially available varieties. I then discarded parameters that I did not think would be useful for the analysis, which included the removal of subregions, reviewer names, and twitter handles.

This left me with a complete dataset that includes countries (string, location), wine description (string), designation (string), rating points (integer), price in USD (integer, USD), province (string, location), variety (string), and winery (string).

Plot Choices

The challenge with this dataset is trying to determine what information is relevant to the consumer and when to show it. The majority of the data are string values, making quantitative analysis challenging, but has also allowed me to abstract new uses for the data.

The typical consumer is interested in getting the best deal on whatever they buy, so allowing them to see a comparison between price and goodness (rating) of many products is important. I believe this is a good first step, however not everyone has the same tastes in wine, therefore having a method to see what something tastes like is essential. Finally, if I were in a liquor store and wanted to buy one of these wines, I would need to

know exactly what the wine is called, what variety it is, where it's from, etc.

I decided on three outputs that I believe would best suite the needs of the consumer listed above. These plots include a word cloud that would describe the flavour profile of the selected wines, a value chart that would plot the top rated wines from a selected price range and would identify wines that are a good/bad value, and a rendered data table that would show the wines identified as good/bad value in the value chart.

Layout and Interaction

I decided to follow a familiar layout with the manipulation tools on the left hand side, and the mainplot area displaying an interactive plot that can be switched between tabs, along with a static data table below. The purpose of the tabs for the interactive plots is to not overload the user with too much information on one screen. Keeping the rendered data table static at the bottom of the page provides reference to the user between pages and allows them to know which wines are being analyzed.

The side panel includes 5 inputs. They include 2 range sliders (one for price, one for points), 2 multi-select dropdown menus (one for variety of wine, one for country), and 1 numeric input (for the number of wines to compare in the value chart).

These inputs allow the user to modify the price to their desired price range, as well as change the rating scale to fine tune the desired rating. These inputs modify all of the outputs, including the word cloud, the value chart, and the rendered data table. The wine variety selector is a dropdown selector with the ability to have multiple wine varieties selected at once. This is important as price, rating (points), and flavours all change fairly dramatically depending on the variety of wine. By selecting multiple wine varieties, one can see how the flavour profile changes and what variety of wines offer the best value. The country selector is again a multi-select dropdown menu that allows the user to select multiple countries. This selector is important because quality of wine varies greatly by region and country. Allowing the user to specify where their wine is from could affect the flavour profile of the wine, as well as its relative value. Finally, the numeric input allows the user to select how many wines they wish to compare on the value chart. While writing this report, I realized that this input should only be found on the value chart tab, as that is the only plot it relates to. This input is important because it gives control to the user to see if there are more wines within their desired criteria that may also be of good value.

The flavour profile word cloud is a plot that reads the description of the top 50 wines in the users desired criteria, and outputs the frequency of the most common descriptor words. I decided to have the flavour profile be the first plot on the main tab because price and score are irrelevant if the flavours of the wine do not match the users tastes. The flavour profile word cloud is sensitive to the price, points, variety, and country inputs. Each one of the input parameters has some effect on the flavour profile.

Once the user has found combination of flavours that match their taste, they may continue to the second tab where the value chart is found. The value chart is a scatter plot that plots the price and points rating of the top 'x' wines, where 'x' is a user defined parameter. A linear regression line is fit to the points, and the residuals are

measured to each point. The residuals, depending on if they are positive or negative, are colour coordinated to give a sense of value to each of the plotted points. This plot makes it easy for the user to identify a wine with good value (the best score with the minimum price), and look for it in the rendered data table below to seek more information. Every point on the scatter plot is has the name of the winery and the designation (title) of each wine.

The rendered data table remains static between tabs, however it is dynamic to all of the inputs noted above. The data table acts as a reference to the user between pages. By default the table organizes the wines by highest rated for the given price range (as it correlates to the value chart), however these are all modifiable by the user. The user can specify if they would like the data sorted by winery, designation, variety, country, points, or price. In addition, the entire dataset is searchable so the user can find any specific wine they may be looking for.

Changes from the Proposal

Reduction of Scope

The app layout from milestone 1 was admittedly a little ambitious. There had been a couple of features that I thought would look 'cool', but realistically didn't add very much substance. As a result, I decided to remove the world map from the app as it made it run more slowly, and didn't contribute much meaning to the consumer. The map was intended to show the quantity of the selected wines are from each country, however in reality this did not work very well as there are some countries that produce far more wine than others. This resulted in the US and various European countries consistently showing the same values. Removing the map from the app also meant that I would remove the inputs associated with it from the toolbar. I removed the continent, region, and flavour selection inputs from the toolbar as a result.

Screenshot 1

Wine Selection Tool

The main page with default settings.

Screenshot 2

Wine Selection Tool

Welcome to the Wine Selection Tool for the [Wine Ratings](#) dataset. Filter the dataset using the selection tools below:

Price Range (\$):

Rating (points):

Variety:

- Merlot
- Pinot Noir
- Shiraz

Country of Origin:

- France
- Germany
- Italy
- US

Value Chart Wines:

Flavour Profile **Value Chart**

Show **10** entries

	winery	designation	variety	country	province	points	price
1	Tenuta dell'Ornellaia	Masseto	Merlot	Italy	Tuscany	99	250
2	Tenuta dell'Ornellaia	Masseto	Merlot	Italy	Tuscany	99	250
3	Fattoria Petrolo	Galatrona	Merlot	Italy	Tuscany	98	135
4	Fattoria Petrolo	Galatrona	Merlot	Italy	Tuscany	98	135

Search:

The main page with modified price range setting.

Screenshot 3

Wine Selection Tool

Flavour Profile
Value Chart

Welcome to the Wine Selection Tool for the [Wine Ratings](#) dataset. Filter the dataset using the selection tools below:

Price Range (\$):

4 53 102 151 200 249 288 347 396 445 486

Rating (points):

80 82 84 86 88 90 92 94 96 98 100

Variety:

 Merlot Pinot Noir Shiraz

Country of Origin:

 France Germany Italy US

Value Chart Wines:

Show **10** entries

	winery	designation	variety	country	province	points	price
1	Le Macchione	Messorio	Merlot	Italy	Tuscany	93	225
2	Domaine René Bouvier	Les Fontenys Premier Cru	Pinot Noir	France	Burgundy	93	128
3	Louis Latour	Clos du Roi	Pinot Noir	France	Burgundy	93	140

Search:

The main page with modified point range setting.

Screenshot 4

Wine Selection Tool

Flavour Profile
Value Chart

Welcome to the Wine Selection Tool for the [Wine Ratings](#) dataset. Filter the dataset using the selection tools below:

Price Range (\$):

4 53 102 151 200 249 298 347 396 445 486

Rating (points):

80 82 84 86 88 90 92 94 96 98 100

Variety:

- Chardonnay
- Cabernet Sauvignon
- Malbec
- Merlot
- Moscato
- Pinot Grigio
- Pinot Noir
- Riesling
- Rosé

Show **10** entries

winery	designation	variety	country	province	points	price
1 Domaine Leflaive	Sous le Dos d'Ane Premier Cru	Chardonnay	France	Burgundy	93	150
2 Domaine Leflaive	Clavoillon Premier Cru	Chardonnay	France	Burgundy	93	145
3 Jarvis	Estate Grown Cave Fermented Finch Hollow Unfiltered	Chardonnay	US	California	93	130

Search:

The main page with modified wine variety settings.

Screenshot 5

Wine Selection Tool

Flavour Profile
Value Chart

Welcome to the Wine Selection Tool for the [Wine Ratings](#) dataset. Filter the dataset using the selection tools below:

Price Range (\$):

4 53 102 151 200 249 298 347 396 445 486

Rating (points):

80 82 84 86 88 90 92 94 96 98 100

Variety:

Country of Origin:

[France](#) [Germany](#) [Italy](#) [US](#)
[Australia](#) [Austria](#) [Argentina](#)
[Croatia](#)

Value Chart Wines:

Show **10** entries

Search:

	winery	designation	variety	country	province	points	price
1	Domaine Leflaive	Sous le Dos d'Ane Premier Cru	Chardonnay	France	Burgundy	93	150
2	Domaine Leflaive	Clavoillon Premier Cru	Chardonnay	France	Burgundy	93	145
3	Jarvis	Estate Grown Cave Fermented Finch Hollow Unfiltered	Chardonnay	US	California	93	130

The main page with modified country settings.

Screenshot 6

Wine Selection Tool

Welcome to the Wine Selection Tool for the [Wine Ratings](#) dataset. Filter the dataset using the selection tools below:

Flavour Profile
Value Chart

Price Range (\$):

4 53 102 151 200 249 298 347 396 445 486

Rating (points):

80 82 84 86 88 90 92 94 96 98 100

Variety:

 Chardonnay Riesling

Show 10 entries
Search:

winery	designation	variety	country	province	points	price
1 Domaine Leflaive	Sous le Dos d'Ane Premier Cru	Chardonnay	France	Burgundy	93	150
2 Domaine Leflaive	Clavoillon Premier Cru	Chardonnay	France	Burgundy	93	145
3 Dr. Pauly	Bernkasteler alte Badstube am Doctorberg Trockenbeerenauslese -	Riesling	Germany	Mosel	93	250

The main page with modified wine variety settings.

Screenshot 7

Show 10 entries

Search:

	winery	designation	variety	country	province	points	price
1	Domaine Leflaive	Sous le Dos d'Ane Premier Cru	Chardonnay	France	Burgundy	93	150
2	Domaine Leflaive	Clavoillon Premier Cru	Chardonnay	France	Burgundy	93	145
3	Dr. Pauly Bergweiler	Bernkasteler alte Badstube am Doctorberg Trockenbeerenauslese - 375 ml	Riesling	Germany	Mosel	93	250
4	S.A. Prüm	Graacher Himmelreich Vat 69 Eiswein	Riesling	Germany	Mosel	93	175
5	Jarvis	Estate Grown Cave Fermented Finch Hollow Unfiltered	Chardonnay	US	California	93	130
6	Olivier Leflaive	Les Pucelles Premier Cru	Chardonnay	France	Burgundy	93	130
7	Baron Knyphausen	Hattenheimer Wisselbrunnen Auslese	Riesling	Germany	Rheingau	93	167
8	Dr. Loosen	Erdener Prälat GG Reserve Alte Reben Trocken	Riesling	Germany	Mosel	93	236
9	Domaine Leflaive	Les Combettes Premier Cru	Chardonnay	France	Burgundy	93	225
10	Roux Père et Fils	Corton-Charlemagne	Chardonnay	France	Burgundy	93	164

Showing 1 to 10 of 56 entries

Previous

1

2

3

4

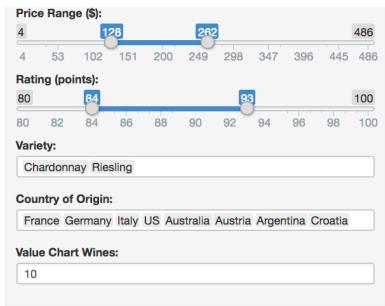
5

6

Next

The rendered datatable with varying size.

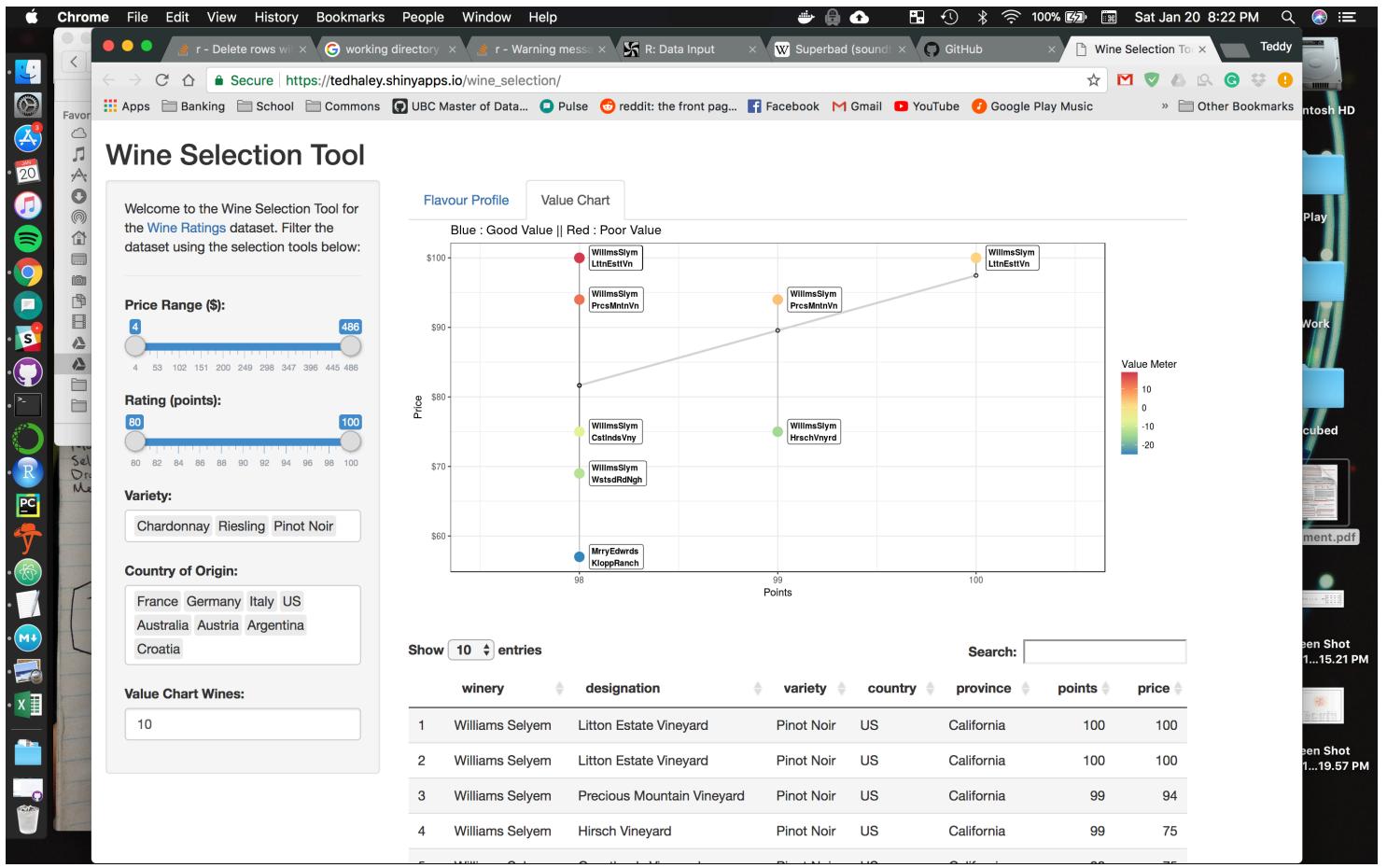
Screenshot 8



Show	25	entries	Search:				
winery	designation	variety	country	province	points	price	
1	Domaine Leflaive	Sous le Dos d'Ane Premier Cru	Chardonnay	France	Burgundy	93	150
2	Domaine Leflaive	Clavillon Premier Cru	Chardonnay	France	Burgundy	93	145
3	Dr. Pauyl Bergweiler	Bernkasteler alte Badstube am Doctorberg Trockenbeerenauslese - 375 ml	Riesling	Germany	Mosel	93	250
4	S.A. Prüm	Graacher Himmelreich Vat 69 Eiswein	Riesling	Germany	Mosel	93	175
5	Jarvis	Estate Grown Cave Fermented Finch Hollow Unfiltered	Chardonnay	US	California	93	130
6	Olivier Leflaive	Les Pucelles Premier Cru	Chardonnay	France	Burgundy	93	130
7	Baron Knyphausen	Hattenheimer Wisselbrunnen Auslese	Riesling	Germany	Rheingau	93	167
8	Dr. Loosen	Erdener Prälatt GG Reserve Alte Reben Trocken	Riesling	Germany	Mosel	93	236
9	Domaine Leflaive	Les Combettes Premier Cru	Chardonnay	France	Burgundy	93	225
10	Roux Père et Fils	Corton-Charlemagne	Chardonnay	France	Burgundy	93	164
11	Olivier Leflaive	Les Folatières	Chardonnay	France	Burgundy	93	137
12	Domaine Laroche	Réserve de l'Obéissance Les Blanchots Grand Cru	Chardonnay	France	Burgundy	93	145
13	Pol Roger	Extra Cuvée de Réserve Blanc de Blancs Brut	Chardonnay	France	Champagne	93	128
14	Pascal Bouchard	Les Clos Grand Cru	Chardonnay	France	Burgundy	93	137
15	Olivier Leflaive	Les Pucelles Premier Cru	Chardonnay	France	Burgundy	93	130
16	Domaine Leflaive	Les Folatières Premier Cru	Chardonnay	France	Burgundy	93	220
17	Domaine Leflaive	Sous le Dos d'Ane Premier Cru	Chardonnay	France	Burgundy	93	155
18	Pol Roger	Extra Cuvée de Réserve Blanc de Blancs Brut	Chardonnay	France	Champagne	93	128
19	S.A. Prüm	Graacher Himmelreich Vat 46 Eiswein	Riesling	Germany	Mosel-Saar-Ruwer	93	143
20	Maximin Grünhäuser	Abtsberg Beerenauslese	Riesling	Germany	Mosel-Saar-Ruwer	93	204
21	Stephan Ehlen	Erdener Treppchen Trockenbeerenauslese	Riesling	Germany	Mosel-Saar-Ruwer	93	144
22	Ruinart	Dorn Ruinart Blanc de Blancs Brut	Chardonnay	France	Champagne	93	180
23	Pol Roger	Extra Cuvée de Réserve Blanc de Blancs Brut	Chardonnay	France	Champagne	93	128
24	Maximin Grünhäuser	Herrenberg Eiswein	Riesling	Germany	Mosel-Saar-Ruwer	93	171
25	Salon	Le Mesnil Blanc de Blancs Brut	Chardonnay	France	Champagne	93	225

The complete main page with rendered data table.

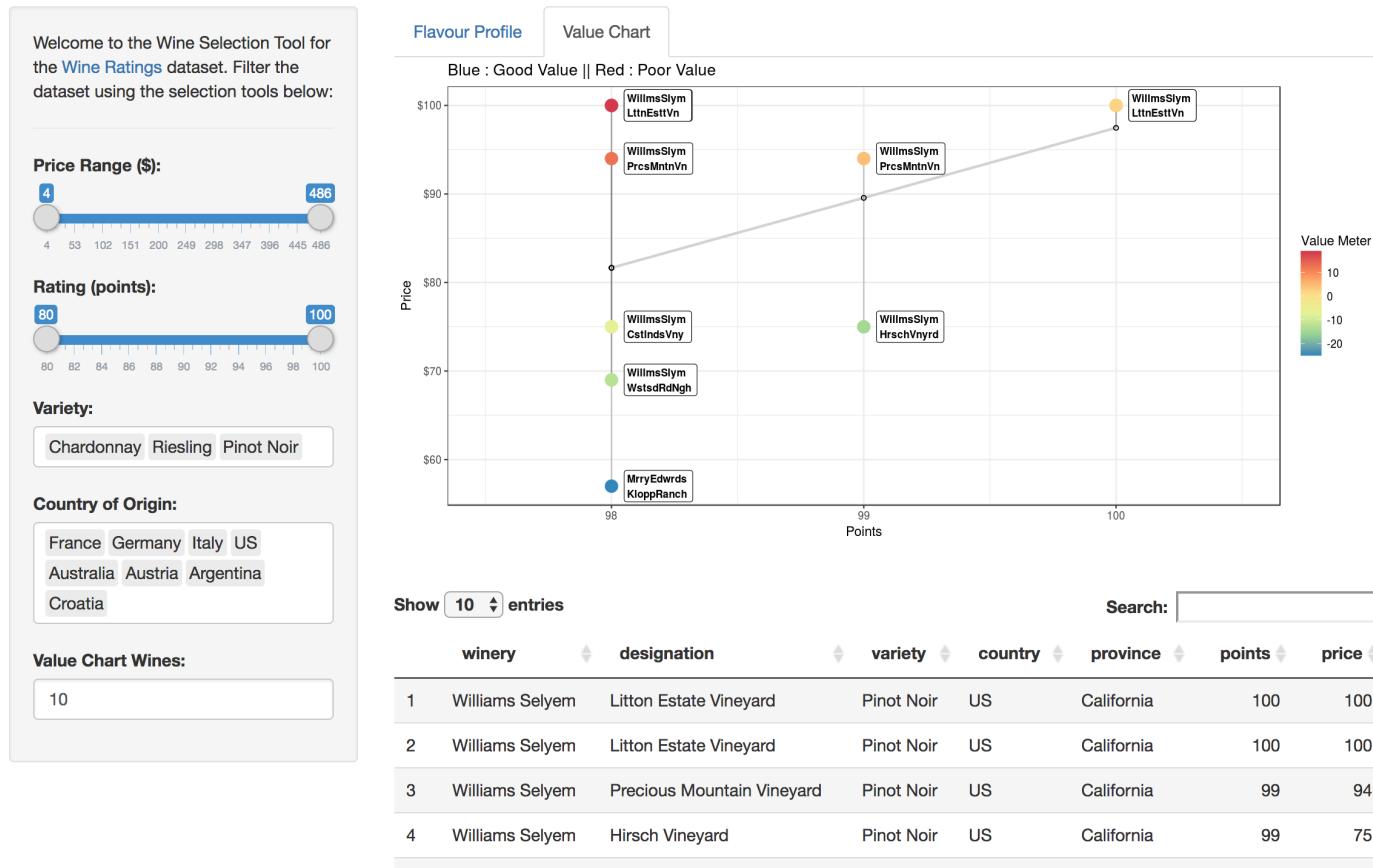
Screenshot 9



The second page with value chart and rendered data table (default settings).

Screenshot 10

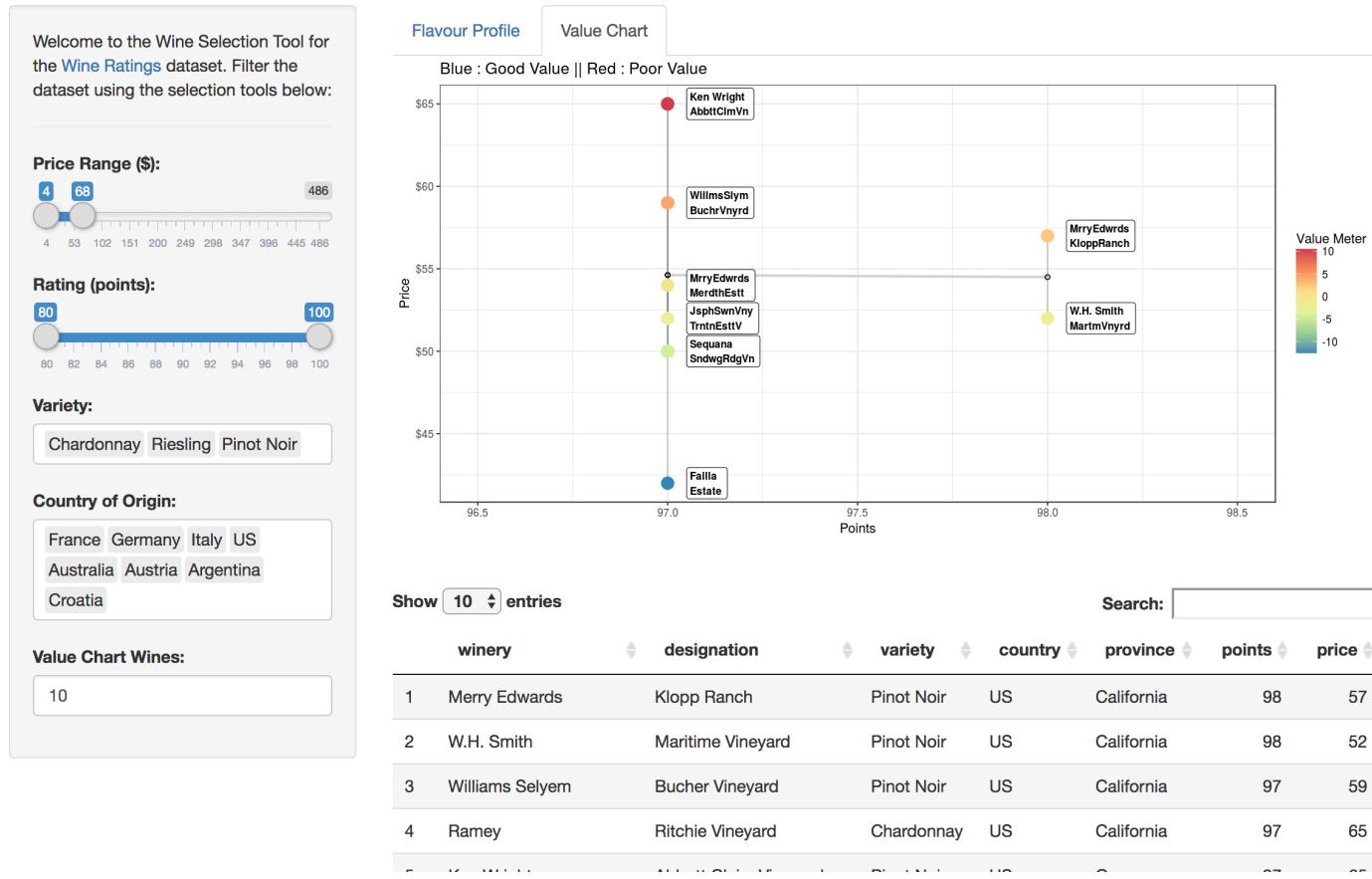
Wine Selection Tool



The value chart with modified price range.

Screenshot 11

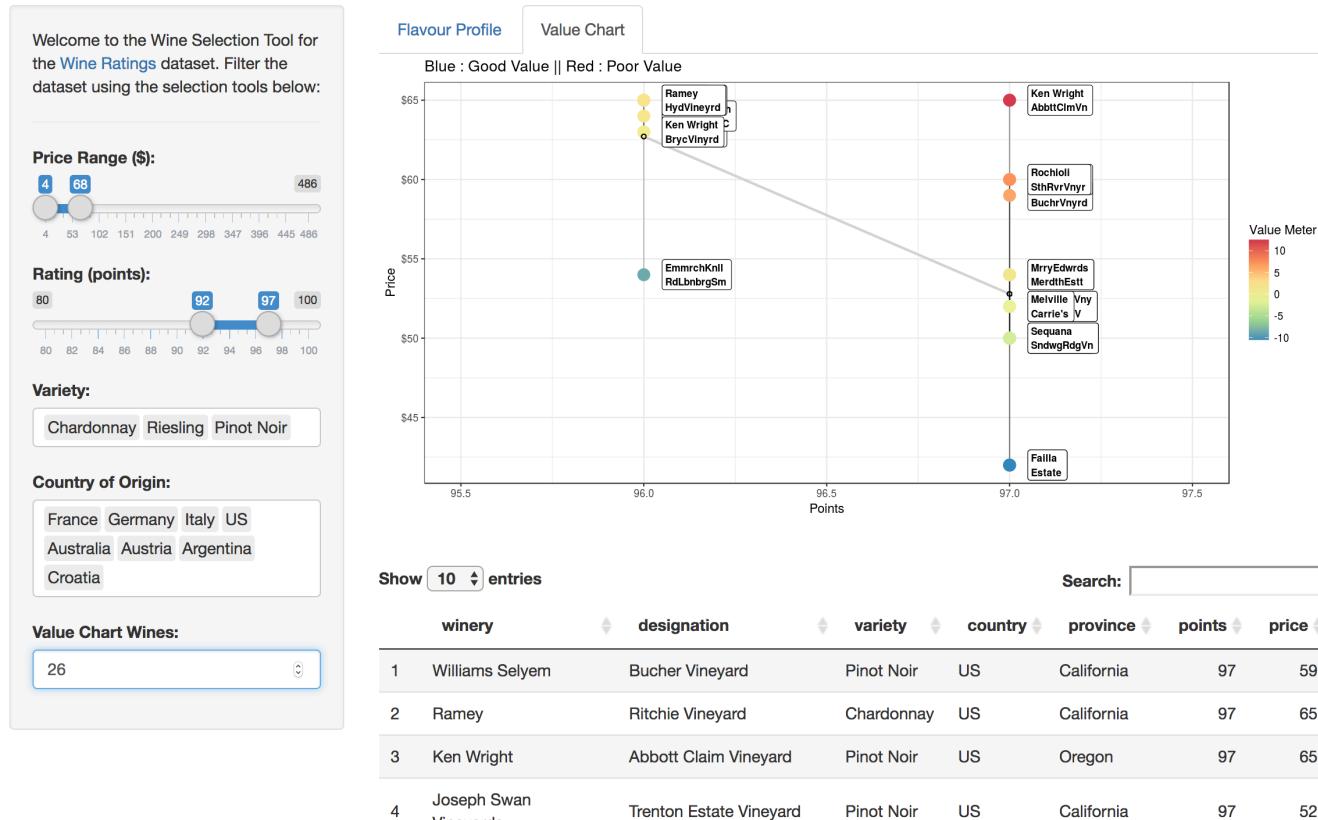
Wine Selection Tool



The value chart with modified point range.

Screenshot 12

Wine Selection Tool



The value chart with additional entries to compare.