

Workshop 8 – Pandas and Regular Expressions

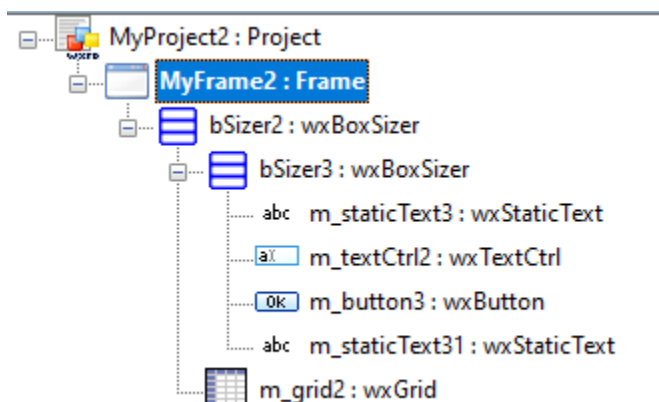
The lecture this week introduced Pandas and Regular Expressions. In this workshop, we will use both to analyse a csv file of Wine reviews

Pre: Install Pandas into your existing python environment

1. Check and install Pandas
 - > `conda env list`
 - > `conda activate Your-Environment`
 - > `conda list pandas`if no Pandas, then install it.
 - > `conda install pandas`
2. Use your existing PyCharm project or create a new one, and link it to your python environment.
 - Download and unzip the weekshop8.zip
 - Copy the extracted files into the new project.

Workshop Task – Using Pandas, load `part_wine_reviews.csv` into a DataFrame; and complete the following analysis task using Pandas and RE.

- Show **all wines (rows)** which are more likely to be "tropical" or "fruity", along with **the total number of such rows**.
 - Tips: Go through all the descriptions and find reviews that use the term 'tropical' or 'fruity'.

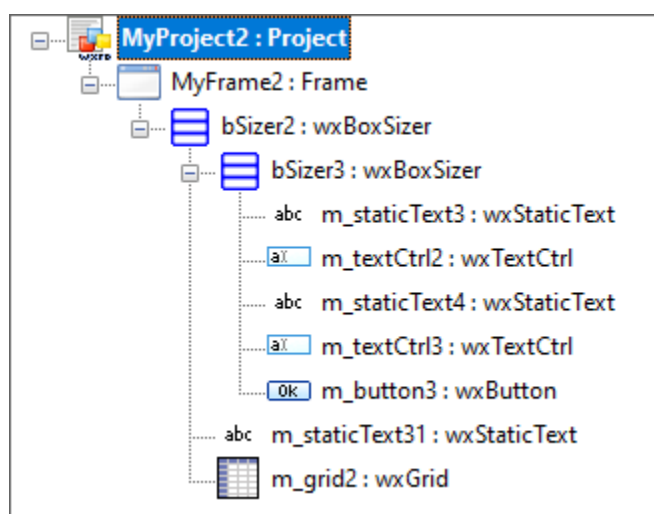


key word: The number of rows: 24

	points	price	province	region	variety	winery	
1	90	42.0	Loire Valley	Sancerre	Sauvignon Blanc	Cherrier Frères	This ripe, full-bodied wine has ripe gre
2	90	25.0	New York	Finger Lakes	Riesling	Dr. Konstantin Frank	While exuberantly fruity, almost tropic
3	91	19.0	Oregon	Willamette Valley	Pinot Gris	Trinity Vineyards	Massively fruity, this is an apple/pear t
4	91	45.0	Burgundy	Chablis	Chardonnay	Domaine Gérard Duplessis	This delicious, refreshing wine is textu
5	86	10.0	California	California	Cabernet Sauvignon	Belle Ambiance	Bright, light oak shadings dress up this
6	86	12.0	Beira Atlantico	nan	Baga	Adega de Cantanhede	The local Baga grape, better known for
7	86	13.0	California	California	White Blend	Kitchen Sink	This wine is dry and substantial, emph
8	86	20.0	Southwest France	Gaillac	Rosé	Château Clément Termes	This pale pink-grey colored wine is a b

Assignment:

- You are required to develop a simple GUI that **allows users to input a minimum and maximum price**. The GUI should then **display all wines (rows) with a price within this range**, along with **the total number of such rows**.



Min Price: Max Price:

The number of rows:

	country	points	price	province	region	variety	winery	
1	US	96	235.0	California	Napa Valley	Cabernet Sauvignon	Heitz	This tremendous
2	Spain	96	110.0	Northern Spain	Toro	Tinta de Toro	Bodega Carmen Rodriguez	Ripe aromas of f
3	US	96	90.0	California	Knights Valley	Sauvignon Blanc	Macauley	Mac Watson hor
4	US	96	65.0	Oregon	Willamette Valley	Pinot Noir	Ponzi	This spent 20 mc
5	France	95	66.0	Provence	Bandol	Provence red blend	Domaine de la Bégude	This is the top wi

- Files required for the submission:**
 - template_frame.py:** contains the code for GUI
 - template_frame.txt:** contains the same code as template_frame.py for Plagiarism Review.
 - run_app.py:** contains the code to run the template

4. **run_app.txt**: contains the same code as run_template.py for **Plagiarism Review**.
5. **Workshop8.fbp**: exported by wxFormBuilder
6. **execution_results.docx**: contains screenshots of the execution results.