
PYTHON CODING TO CUSTOMIZE AND INTEGRATE WORKFLOWS WITH EQUIS

Southeast EQUIS User Conference

2024-03-08

R. Siebenmann, P.E., GISP

Key take-home message:

- EQUIS is a great platform to collect, manage, and store data and generate outputs.
- Python is versatile for data analysis and visualization.
- Four common use-cases:
 - Data Cleaning/Preprocessing
 - Data Processing
 - Charting
 - Chembox Labels



Python: A High-Level, General-Purpose Programming Language

- Easy to learn and versatile.
- User-friendly with simple syntax.
- Extensive libraries and modules.
- Emphasizes readability.
- Ubiquitous in data science.
- Essential tool in engineering.



Use Case 1: Python in Data Cleaning and Preprocessing

Task: Generate an EDD from a multi-column pivoted table

- Uncrosstab or unpivot complex data.
- Remove duplicates.
- Update values.
- Comparison with traditional methods (Visual Basic, Access databases).

	A	B	C	D	E	F	G	H
1	Device ID	Date/Time	CH4	CO2	O2	Balance	%LEL	Init. Static
2	CALCHKPM	#####	49.7	35.1	0	15.2	>>>>	-1.23
3	CALCHKPM	#####	49.4	35	0	15.6	>>>>	-1.29
4	CALCHKPM	#####	50.1	34.6	0	15.3	>>>>	-1.61
5	CALCHKPM	#####	49.1	34.3	0	16.6	>>>>	-1.49
6	CALCHKPM	#####	49.6	34.3	0.2	15.9	>>>>	-1.62
7	CALCHKPM	#####	49.8	34.7	0	15.5	>>>>	-1.78
8	CALKKHPM	#####	0	0.2	11	88.8	0	69.63
9	CALKKHPM	#####	0	0.2	11.1	88.7	0	66.95
10	CALKKHPM	#####	0	0.3	11	88.7	0	68.09
11	CALKKHPM	#####	0	0.6	10.8	88.6	0	68.31
12	CALKKHPM	#####	0	1.8	10.9	87.3	0	65.33
13	CALKKHPM	#####	0	0.3	11.2	88.5	0	65.48
14	KLFFLR01	#####	46.1	37.1	0.3	16.5	>>>>	-8.5
15	KLFFLR01	#####	45.9	36.6	0.2	17.3	>>>>	-8.56



	A	B	C	D
1	Device ID	Date/Time	parameter	results_str
2	CALCHKPM	3/27/2018 14:00	CH4	49.7
3	CALCHKPM	3/27/2018 14:00	CO2	35.1
4	CALCHKPM	3/27/2018 14:00	O2	0
5	CALCHKPM	3/27/2018 14:00	Balance	15.2
6	CALCHKPM	3/27/2018 14:00	%LEL	>>>>
7	CALCHKPM	3/27/2018 14:00	Init. Static Press.	-1.23
8	CALCHKPM	3/27/2018 14:00	Adj. Static Press.	-1.16
9	CALCHKPM	3/27/2018 14:00	Init. Diff. Press.	1.085
10	CALCHKPM	3/27/2018 14:00	Adj. Diff. Press.	1.003
11	CALCHKPM	3/27/2018 14:00	Init. Gas Temp.	0
12	CALCHKPM	3/27/2018 14:00	Adj. Gas Temp.	0

Use Case 2: Data Processing with EQuIS

Task: Update a custom field to report duplicate results

- Exporting an Analytical Results II table.
- Application of reporting logic using Python.
- Updating the database using SQL Server Management Studio.

Field and Laboratory Duplicates

For cases where a field duplicate sample result is present for the same parameter in a sample, a single representative concentration for the sample was selected as follows:

- If there was a detection in both samples, the higher concentration was selected;
- If there was a detection in one sample but not the other, the detected concentration was selected; and
- If there was not a detection in either sample, the lowest reported value was selected.



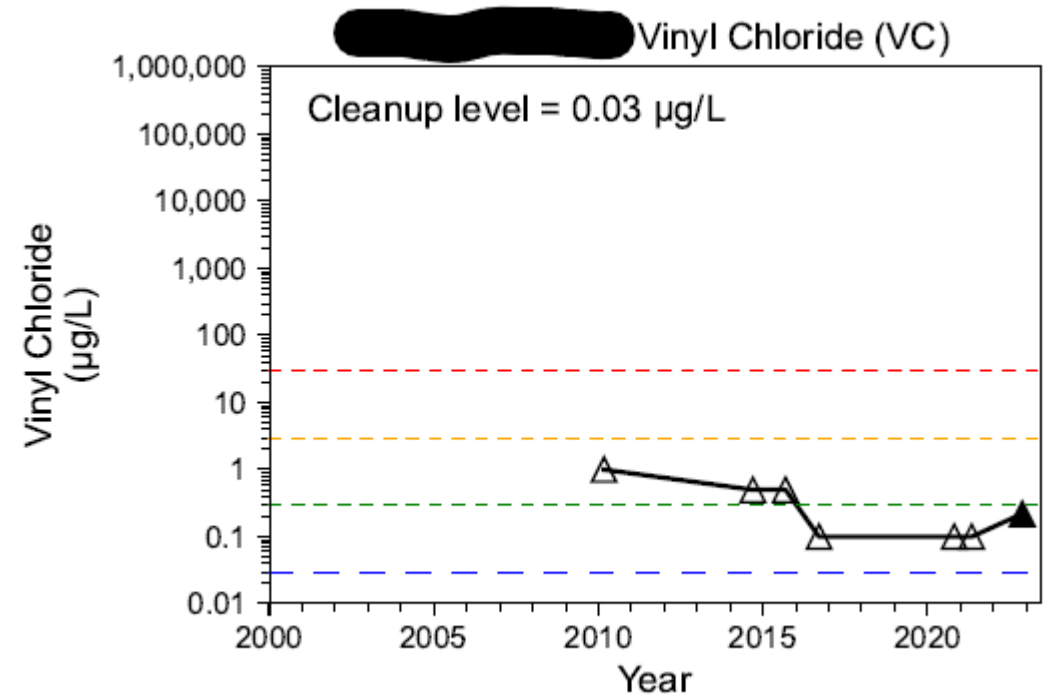
	A	B	C
1	TEST_ID	CAS_RN	exclude
2	653138	113507-82-7	
3	653137	113507-82-7	Exclude
4	653138	13252-13-6	
5	653137	13252-13-6	Exclude
6	653138	151772-58-6	
7	653137	151772-58-6	Exclude



Use Case 3: Charting with Matplotlib

Task: Generate Statistical Plots and Time-Series Plots

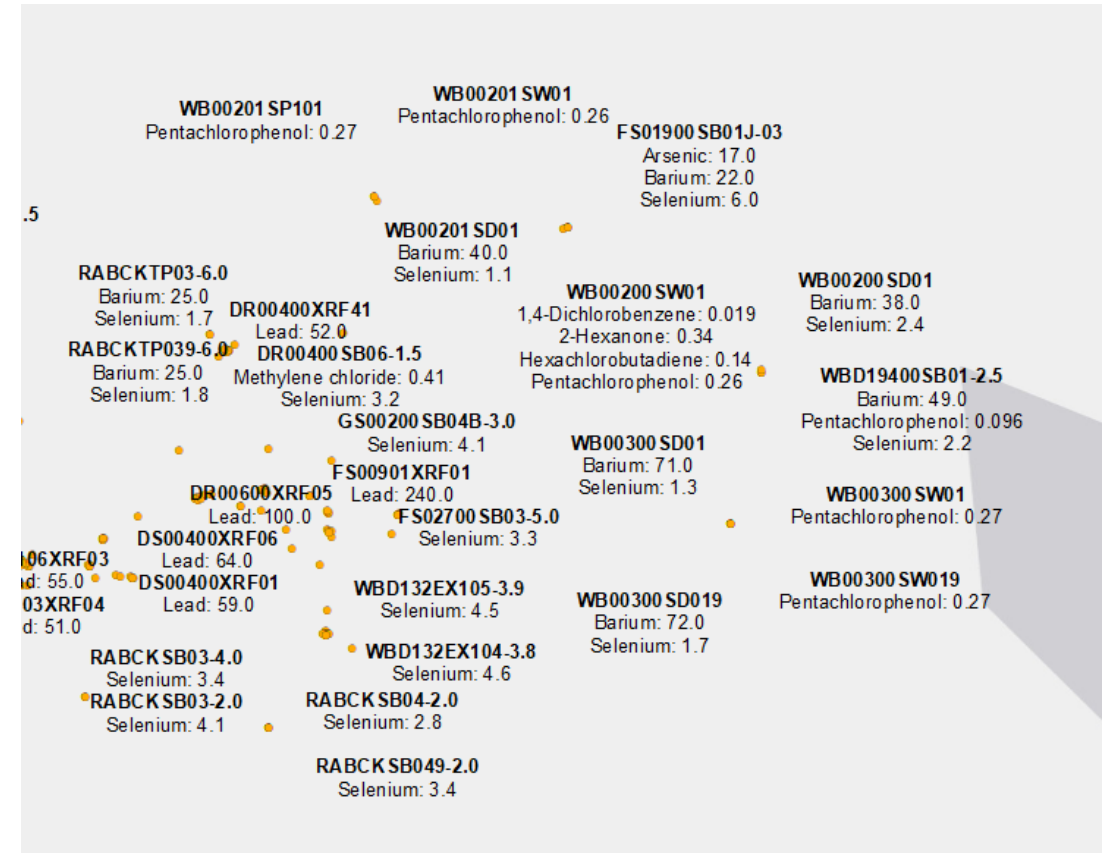
- Matplotlib library
- Creating charts for time-series data directly from EQulS exports.



Use Case 4 - Custom Chembox Labels for ArcGIS Pro

Task: Create “Chembox” style labels for ArcGIS Pro

- Creating custom labels with Python.
- Application and benefits of these labels in ArcGIS Pro.



Python Best Practices and Tips

- Uphold neat, consistent coding.
- Include comments and documentation.
- Manage errors and exceptions effectively.
- Match format to project complexity.
- Use YAML for storing keys and passwords.



Python and EQUIS: Empowering Your Data Workflow

- EQUIS has amazing out-of-the-box reporting functionality
- Python delivers another set of tools to perform analysis
- It is the "lingua franca" of data science.
- Gaining familiarity and investing in it is a great career move as part of life-long learning.

