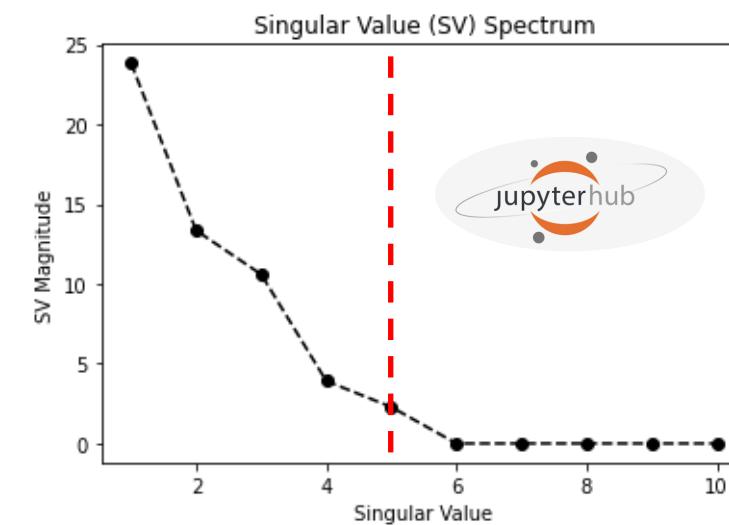


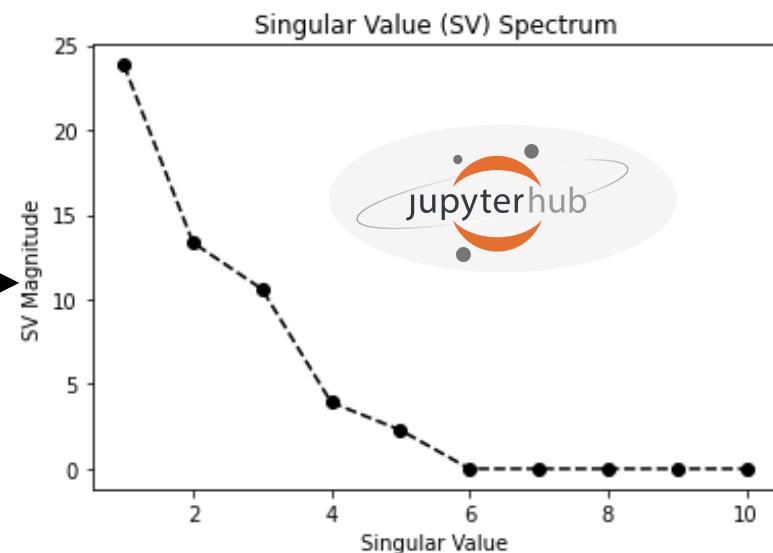
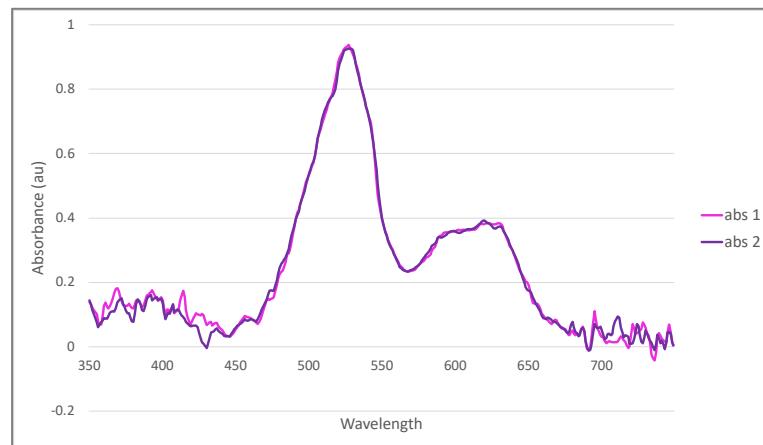
Online operation via JupyterHub or offline operation via Raspberry/Orange Pi

Collecting data for unknown dyes and performing singular value decomposition

Pts III and IV of Introduction to Spectroscopy and Data Science: How many dyes are in a collection of colorful samples?



Today you'll be taking on the role of a food chemist, tasked with identifying how many unique dyes are used to make 5 food samples that are all different colors.



In part 1, you'll collect spectroscopy data on your unknown samples.

You'll be using your hand-built spectrometers for this exercise. As you work through data collection, think about how your results might vary if you were using a commercial instrument.

In part 2, you'll use a data science technique to analyze your data and determine how many dyes are present.

All calculations will be done using Jupyter Notebooks, which are commonly used in industry and academic settings for programming.

Part III: Collecting spectroscopy data of your unknown mixtures

In this portion of the lab, you will collect absorbance spectroscopy data on samples made from combinations of unknown dyes.

Video instructions for spectrometer operation (optional):

<https://vimeo.com/661075622>

Password: Trimontana

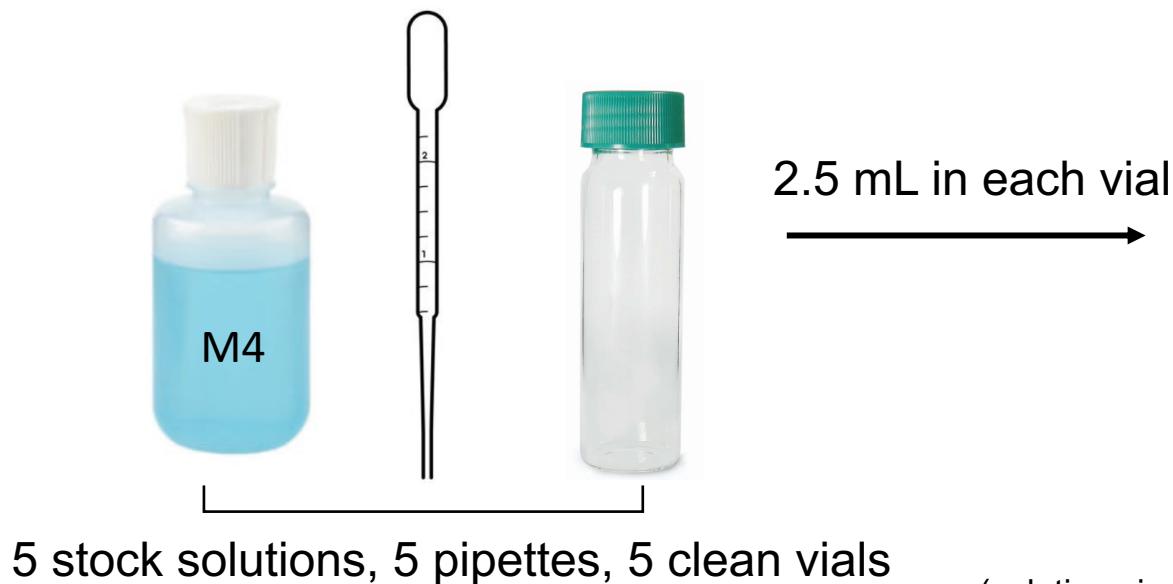
Collect your unknown samples

****Important note before starting:**

Do not dispose of any of your unknown solutions unless your instructor tells you otherwise.

Do not return your solutions to the stock solutions after you have taken your 2.5 mL fraction.

Do not use the same pipette for two different stock solutions.



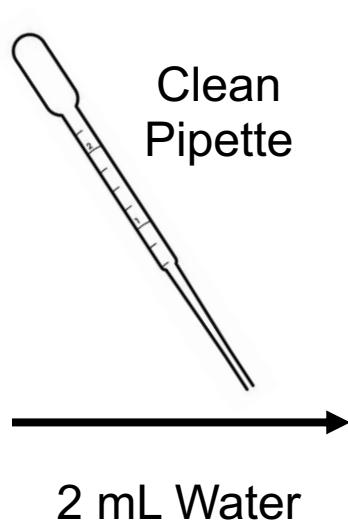
Example of 5 unknown samples
(solutions in your class may look different, as different dye combinations may be used for different courses)

**All solutions will contain between 2 and 6 dyes

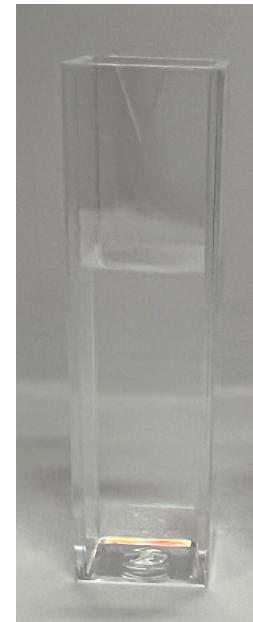
- 1.) Transfer ~2.5 mL of each unknown sample from the 5 stock solutions to 5 different vials (one vial for each unique stock, labeled M1 – M5). Each of the 5 samples should look different but will contain mixtures of the same dyes. Your job today will be to determine how many individual dyes are present across all mixtures in a set.

Prepare your blank cuvette (if needed)

2.) Make sure all your cuvettes are clean before starting and before measuring a new dye to ensure your unknown solutions are not contaminated with residual dye.



Clean Cuvette

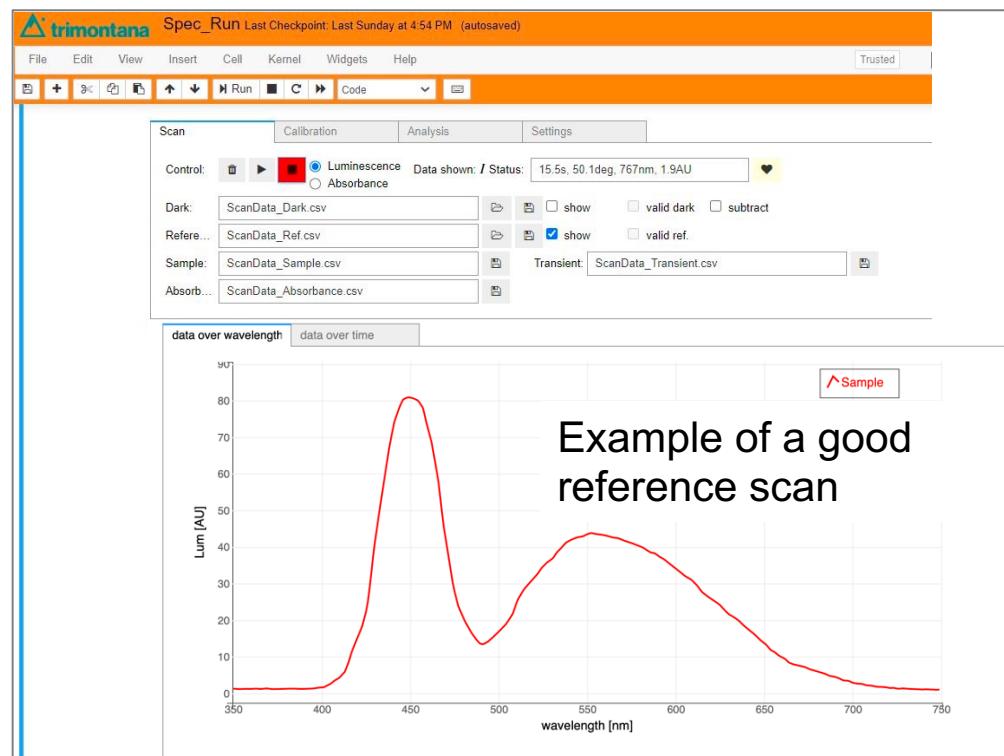


Background/Reference Sample

If you already have a 2 mL water background saved and have not bumped/moved your diffraction grating, you can skip steps 3 – 4 and move to slide 5 (step 5).

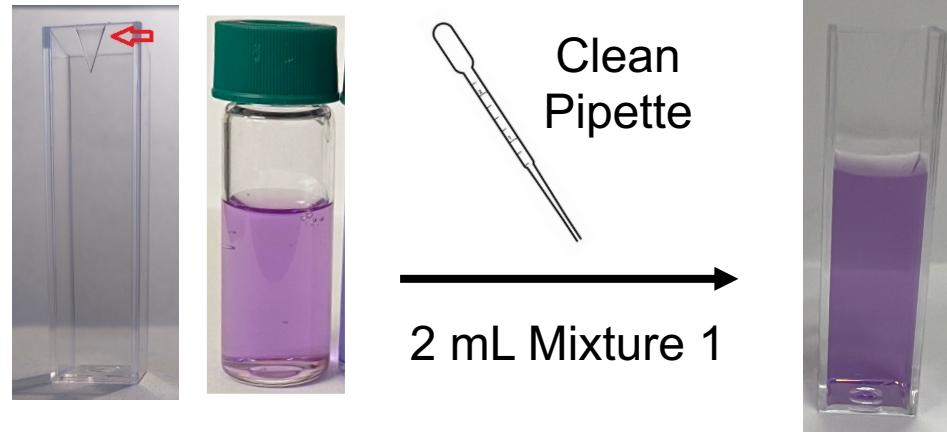
3.) Prepare a blank/reference cuvette with ~2 mL water (just like you did for the known food dye samples) and a sample cuvette with ~2 mL of your first unknown solution. Use a plastic pipette to measure out each sample for this part of the lab, making sure to use a separate pipette for each unknown mixture and your blank/reference sample.

Collect reference spectrum (if needed)



- 4.) Collect a background spectrum and save it as you did in the previous section of the lab (Part II, Slides 4 – 9). **If you already have a 2 mL water background saved and have not bumped/moved your diffraction grating, you can skip this step.** Remember to check your blank/background scan, which will appear in RED on the absorbance graph in the Trimontana Jupyter notebook. If there are any sharp spikes or peaks that deviate from your reference background spectra (included with your Trimontana spectrometers), you will need to re-collect this data.

Collect absorbance spectrum for your first sample



Clean Pipette
2 mL Mixture 1



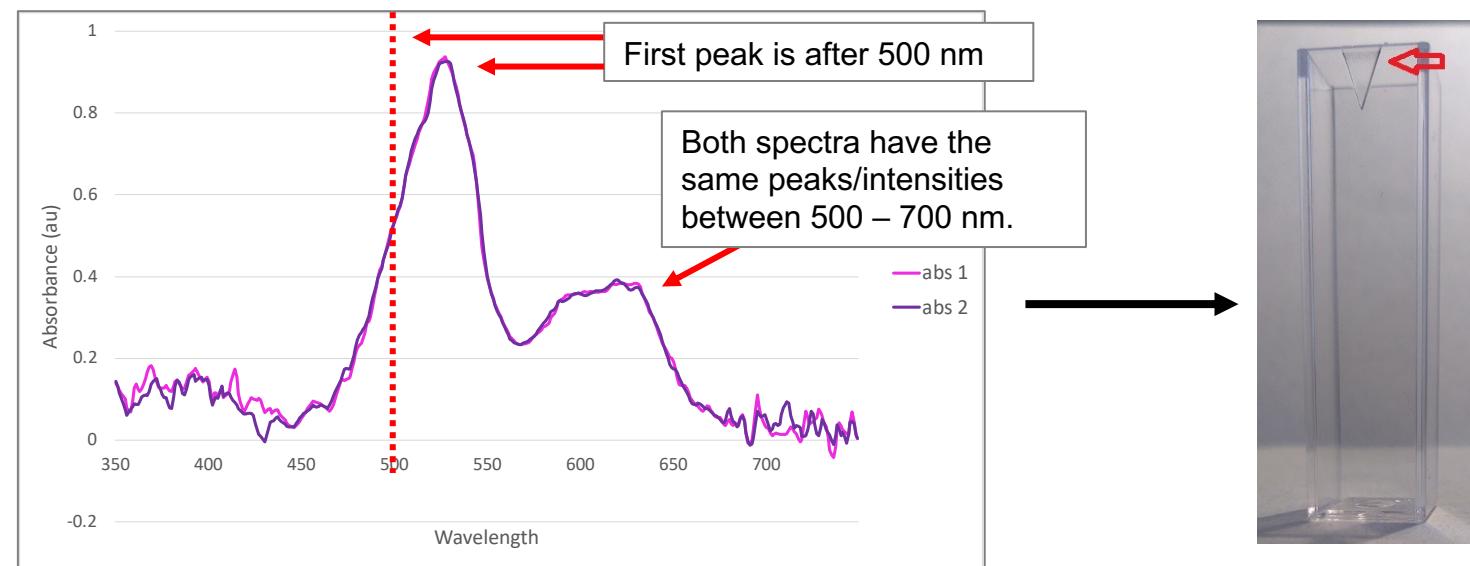
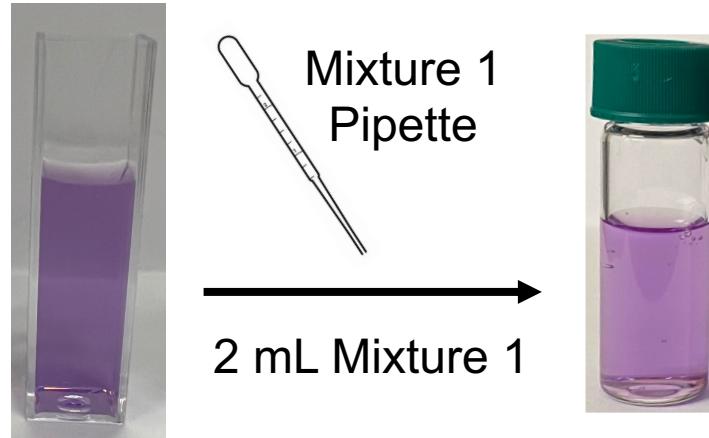
Remember to clear data before each new scan

Save data to Jupyter with a **file name for your group**. Then **download** from Jupyter and **save locally** on your computer as a **.xlsx** or **.csv** file). After data has been downloaded, you can repeat and collect a second spectrum.

5.) Using a **clean cuvette** and **clean pipette**, transfer ~2 mL of your first mixture to a cuvette.

6.) Collect **two absorbance spectra** for your first unknown solution. Make sure to **save data to Jupyter using a file name for your group** (e.g., Absorbance_group099.csv). You'll write all of your data to this filename. **Download and save each spectrum after it is collected** (refer to instructions from Part II, slides 10 – 13 for specific instructions on collecting and downloading data). Save your data as either a .xlsx (Excel) or .csv file.

Return sample to vial, plot data, and clean sample cuvette



7.) Once you have collected two absorbance spectra of your first unknown, return the 2 mL of solution to your original sample vial (NOT the stock solutions provided to the class) using your designated pipette for that solution.

If time allows, plot both spectra in Excel to make sure your two spectra for the same dye look similar (example shown above). If your data looks significantly different across the two scans (peaks are shifted or intensity is different by more than 0.1 au), recollect one of your data points.

Your data might look slightly different than the one shown above, as different courses will have different dyes. However, all dyes should have peak absorbances between 500 and 700 nm. **If the first peak in your data is before 500 nm, notify your instructor before proceeding, as this could indicate a problem with your diffraction grating.

8.) Clean the cuvette you used for your first unknown solution thoroughly with water, making sure the cuvette is completely dry before moving on to your next mixture.

Repeat data collection for remaining four samples



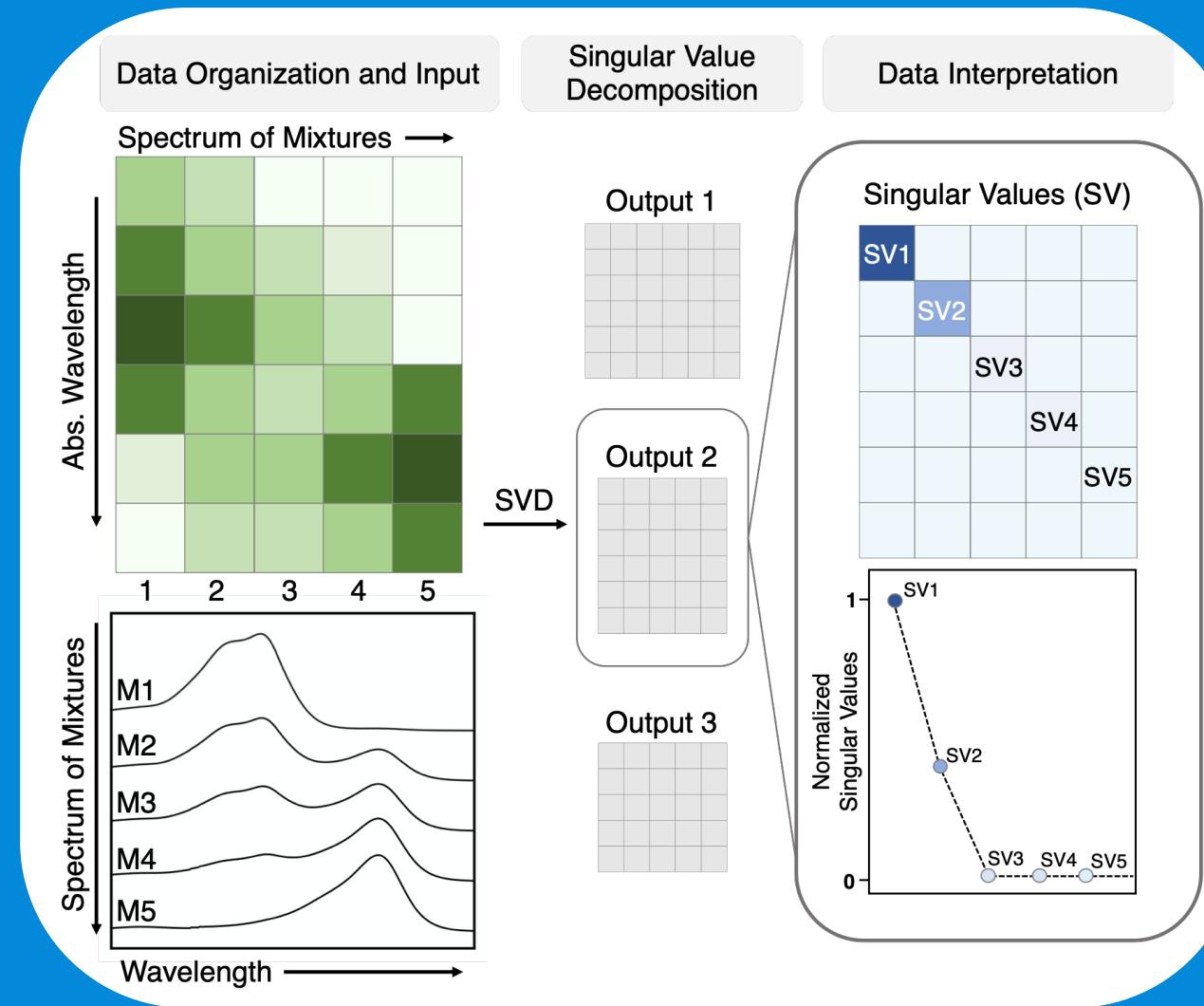
9.) Repeat steps 5 - 8 (slides 7 – 8) with the remaining four unknown dye mixtures. If you have not refreshed your browser with the Jupyter notebook for data collection, and you have not adjusted/bumped the diffraction grating for your spectrometer, you should not need to recollect and save your blank/reference data after you have completed steps 3 and 4 once.

Part IV: Data Analysis using Singular Value Decomposition

In this portion of the lab, you will apply a data science technique to analyze your spectroscopy data and answer the question:

How many dyes do you have in your mixtures?

Workflow for Part IV shown below



Organize your data in Excel

**See Excel tips slides if you need help copying, plotting, or formatting data in Excel

Column values should extend down to row 401

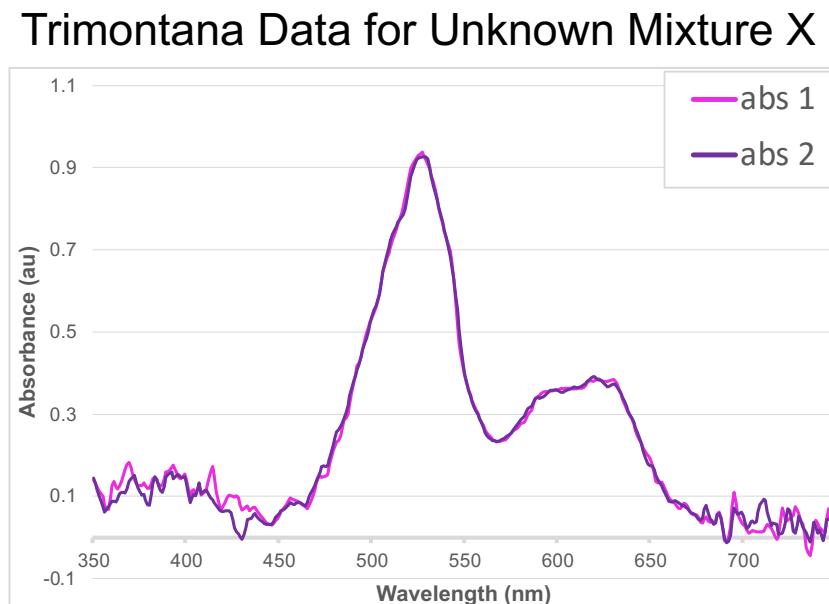
wavelength[nm]	Mix 1 abs 1	Mix 1 abs 2	Mix 2 abs 1	Mix 2 abs 2	Mix 3 abs 1	Mix 3 abs 2	Mix 4 abs 1	Mix 4 abs 2	Mix 5 abs 1	Mix 5 abs 2
350	0.06746	0.07395	0.0557	0.0526	0.05246	0.04425	0.06413	0.08846	0.06868	0.06519
351	0.06636	0.06871	0.05721	0.05017	0.04902	0.04444	0.06028	0.07891	0.063	0.06499
352	0.06675	0.06495	0.05674	0.06128	0.04313	0.03946	0.05909	0.0709	0.05881	0.06629
353	0.06547	0.05999	0.05045	0.06036	0.03605	0.03589	0.05823	0.06175	0.05432	0.06547
354	0.06251	0.05353	0.04913	0.04735	0.02876	0.03367	0.05458	0.05458	0.0495	0.06251
355	0.06579	0.05311	0.05876	0.04762	0.02844	0.03096	0.05598	0.05598	0.05091	0.06579
356	0.06845	0.0561	0.05412	0.05514	0.03154	0.02804	0.06476	0.06277	0.05767	0.07114
357	0.06385	0.06133	0.03727	0.0527	0.03849	0.02355	0.06948	0.06467	0.05954	0.07026
358	0.05776	0.06729	0.04404	0.04759	0.0467	0.01862	0.06785	0.06026	0.04813	0.06785

11 columns total, with 1 for wavelength, and the rest for 10 sets of absorbance values

- 1.) Organize your data in Excel by copying over all absorbance values. An example of what your organized data should look like is shown above.

Plot your data in Excel and estimate number of dyes

**See Excel tips slides if you need help copying, plotting, or formatting data in Excel



Plot data for all samples
(only one example is
shown here)

2.) Create an absorbance vs. wavelength graph for all five samples (example of one plot shown above).

3.) Before proceeding to the singular value decomposition (SVD) analysis steps, **hypothesize how many different dyes are in this sample based on absorbance features in the 500 - 700 nm range** (as this spectrometer is less accurate in the 400 - 500 nm range). What features of your data led you to this conclusion? **Hint: your answer should lie between 1 - 6 dyes.**

Estimated number of dyes = _____

Organize data for SVD analysis

4.) Create a separate .csv file with your absorbance data from 500 - 700 nm. This will serve as the data table that you'll input for SVD. Do this by copying and pasting the absorbance values from 500 - 700 nm for each unknown dye mixture spectrum. Do not include titles at the top of your columns. **The final .csv file should look something like the photo outlined in red (on the bottom right).**

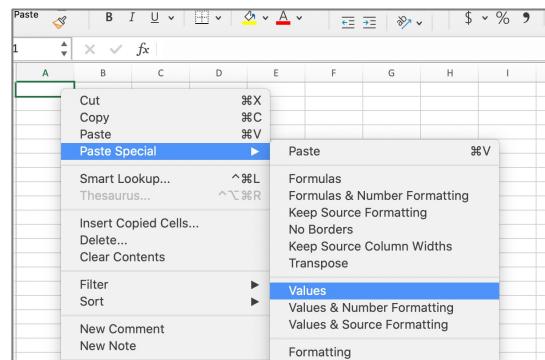
a.) Open Excel file with organized data

wavelength[nm]	Mix 1 abs 1	Mix 1 abs 2	Mix 2 abs 1	Mix 2 abs 2	Mix 3 abs 1	Mix 3 abs 2	Mix 4 abs 1	Mix 4 abs 2	Mix 5 abs 1	Mix 5 abs 2
350	0.06746	0.07395	0.0557	0.0526	0.05246	0.04425	0.06413	0.08846	0.06868	0.06519
351	0.06636	0.06871	0.05721	0.05017	0.04902	0.04444	0.06028	0.07891	0.063	0.06499
352	0.06675	0.06495	0.05674	0.06128	0.04313	0.03946	0.05909	0.0709	0.05881	0.06629
353	0.06547	0.05999	0.05045	0.06036	0.03605	0.03589	0.05823	0.06175	0.05432	0.06547
354	0.06521	0.05353	0.04913	0.04735	0.02876	0.03367	0.05458	0.05458	0.0495	0.06251
355	0.06579	0.05311	0.05876	0.04762	0.02844	0.03096	0.05598	0.05598	0.05091	0.06579
356	0.06845	0.05861	0.05412	0.05514	0.03154	0.02804	0.06476	0.06277	0.05767	0.07114

b.) Copy absorbance values from 500 – 700 nm

A	B	C	D	E	F	G	H	I	J	K	L	M
487	0.1246	0.1151	0.08267	0.0894	0.01589	0.01729	0.187	0.1109	0.05815	0.06802		
488	0.1227	0.1143	0.08503	0.09473	0.008688	0.0131	0.1147	0.1121	0.05708	0.06301		
489	0.1278	0.1091	0.0885	0.1016	0.004621	0.0194	0.1029	0.09925	0.05995	0.05692		
490	0.1204	0.1086	0.0883	0.1016	0.004621	0.0194	0.1029	0.09925	0.05995	0.05692		
491	0.1248	0.09104	0.0883	0.1064	0.008144	0.01879	0.05933	0.05698	0.05008	0.05835		
492	0.1334	0.1036	0.08863	0.1124	0.008813	0.02444	0.1118	0.111	0.05501	0.04944		
493	0.1403	0.1151	0.09707	0.1167	0.01473	0.02807	0.1236	0.1178	0.05702	0.05038		
494	0.143	0.1186	0.1016	0.1169	0.01815	0.0267	0.1193	0.1156	0.05196	0.04765		
495	0.1512	0.1197	0.1078	0.1204	0.02006	0.028	0.1193	0.1141	0.04924	0.04867		
496	0.164	0.123	0.1155	0.1263	0.02414	0.0286	0.1221	0.1132	0.04861	0.05176		
497	0.1665	0.1338	0.124	0.1241	0.02863	0.02827	0.1251	0.1211	0.04797	0.05935		
498	0.1704	0.1393	0.1272	0.1303	0.02241	0.0257	0.1257	0.1248	n.mn15	n.mn081		
499	0.1793	0.1452	0.1306	0.13302	0.02934	0.0345	0.13315	0.129	Cut	%X		
500	0.1918	0.1452	0.1395	0.1395	0.03302	0.02934	0.1345	0.13315	Copy	%C		
501	0.1975	0.1572	0.1354	0.1391	0.0184	0.03251	0.1561	0.1262	Paste	%V		
502	0.2017	0.1672	0.1486	0.1474	0.01931	0.03302	0.1615	0.1405	Paste Special	►		
503	0.2058	0.1768	0.1339	0.1465	0.01465	0.03372	0.1579	0.1673	Smart Lookup...	^%L		
504	0.2064	0.1806	0.1177	0.1433	0.05928	0.03523	0.1594	0.1732	Thesaurus...	^~%R		
505	0.2118	0.1869	0.1214	0.1231	0.05324	0.03917	0.1657	0.1768	Insert...			
506	0.2267	0.1956	0.1305	0.1335	0.06271	0.04907	0.1726	0.1721	Delete...			
507	0.2358	0.1938	0.138	0.1404	0.0636	0.05839	0.1845	0.173	Clear Contents			
508	0.2457	0.2194	0.1223	0.1387	0.06282	0.06591	0.1908	0.1768	Format Cells...	^%1		
509	0.2588	0.2437	0.1284	0.1473	0.07539	0.07951	0.2105	0.1947	Pick From Drop-down List...			
510	0.2753	0.2569	0.1268	0.1521	0.09045	0.09462	0.2348	0.2153	New Comment			
511	0.3023	0.2697	0.1269	0.1542	0.09723	0.09697	0.2457	0.2249	New Note			
512	0.3222	0.2756	0.1301	0.1485	0.1017	0.1027	0.2539	0.2396	Define Name...			
513	0.3283	0.2874	0.1393	0.1478	0.1026	0.1132	0.2574	0.2612	Hyperlink...	%K		
514	0.3396	0.307	0.1485	0.1519	0.114	0.1233	0.2828	0.2836				
515	0.3482	0.3234	0.1519	0.1704	0.1494	0.1392	0.2919	0.2939				
516	0.3842	0.3534	0.1619	0.1704	0.1494	0.1392	0.3191	0.3139				
517	0.4045	0.3754	0.1657	0.1699	0.1636	0.1515	0.323	0.3274				
518	0.4192	0.3919	0.1704	0.1756	0.1641	0.1678	0.3405	0.3371				
519	0.4356	0.4156	0.175	0.1756	0.1805	0.1871	0.3535	0.3528				
520	0.4526	0.4438	0.1828	0.1723	0.2068	0.2032	0.3748	0.3767				
521	0.4684	0.4707	0.1841	0.1811	0.2135	0.2163	0.3949	0.4004				
522	0.4779	0.4875	0.1851	0.1907	0.22	0.2309	0.4146	0.4133				
523	0.4901	0.5053	0.1873	0.1982	0.2317	0.2424	0.434	0.4281				
524	0.5068	0.5187	0.1927	0.2024	0.2456	0.2487	0.4507	0.446				

c.) Open a new blank spreadsheet and paste your absorbance values.



d.) Final spreadsheet should have values in columns A – J and extend down to row 201

A	B	C	D	E	F	G	H	I	J	K
1	0.1918	0.1452	0.1395	0.1356	0.03302	0.02934	0.1345	0.1315	0.06708	0.06449
2	0.1975	0.1572	0.1354	0.1391	0.0184	0.03251	0.1561	0.1262	0.06333	0.08356
3	0.2017	0.1672	0.1485	0.1474	0.01931	0.03035	0.1615	0.1405	0.05495	0.08648
4	0.2053	0.1768	0.1339	0.1465	0.03372	0.0322	0.1579	0.1673	0.05049	0.08063
5	0.2064	0.1806	0.1177	0.1433	0.05928	0.03523	0.1594	0.1732	0.04971	0.07201
6	0.2118	0.1869	0.1214	0.1231	0.05324	0.03917	0.1657	0.1657	0.0556	0.06623
7	0.2267	0.1997	0.1305	0.1235	0.06177	0.05497	0.1854	0.1721	0.06946	0.07377
8	0.2368	0.2103	0.128	0.1374	0.0636	0.0636	0.1936	0.1723	0.07561	0.08324
9	0.2457	0.2194	0.1223	0.1387	0.06282	0.06591	0.1908	0.1768	0.07742	0.08702
10	0.2588	0.2427	0.1284	0.1473	0.07539	0.07951	0.2105	0.1947	0.09103	0.1021
11	0.2753	0.2569	0.1268	0.1521	0.09045	0.09462	0.2348	0.2153	0.1062	0.1162
12	0.3025	0.2697	0.1265	0.1542	0.09723	0.09697	0.2457	0.2249	0.1069	0.1267
13	0.3223	0.2756	0.1301	0.1485	0.1017	0.1027	0.2539	0.2306	0.1117	0.1246

Login to course JupyterHub

Sign in

Username:

Password:

Sign in

This is a service provided by Trimontana
Teaching Solutions -- trimontana.tech
For a login request please contact us at
support@trimontana.tech
Instructions can be found [here](#)

5.) In this step you will log into the JupyterHub for your course.

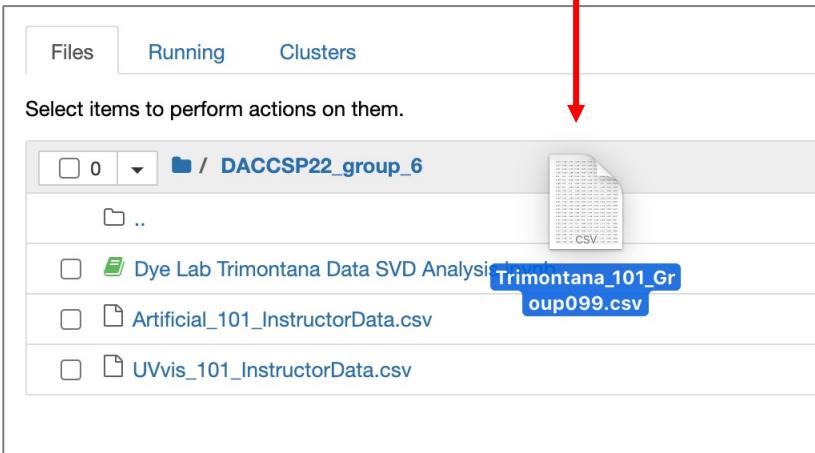
Open a web browser, and type the following address into the address bar: <http://hub.trimontana.tech>

This will take you to the log in screen for the Jupyter hub, the online repository in which the software is stored. Input your username and password. Then click “sign in”

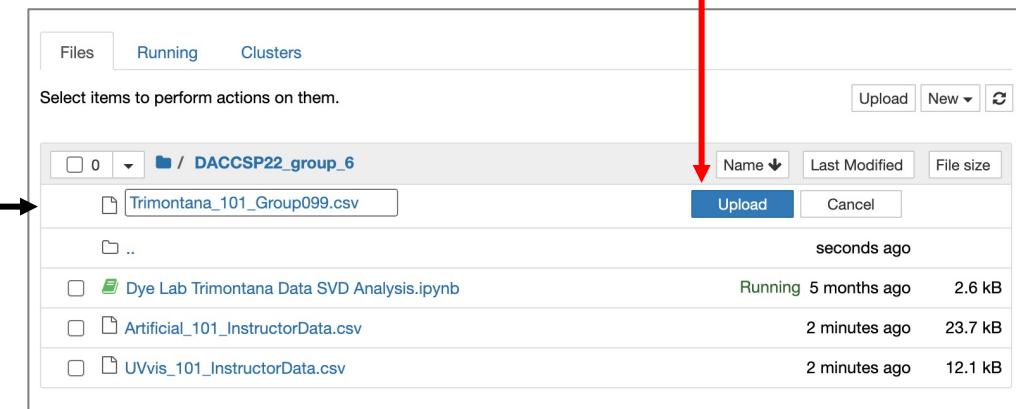
Username and password will be provided by your instructor

Upload your group's data to your folder

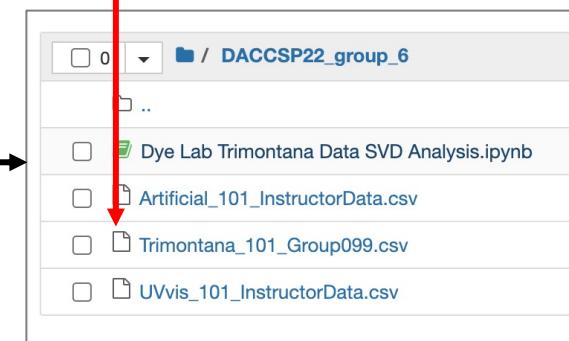
Drag and drop .csv



Select Upload



File should now appear in group folder

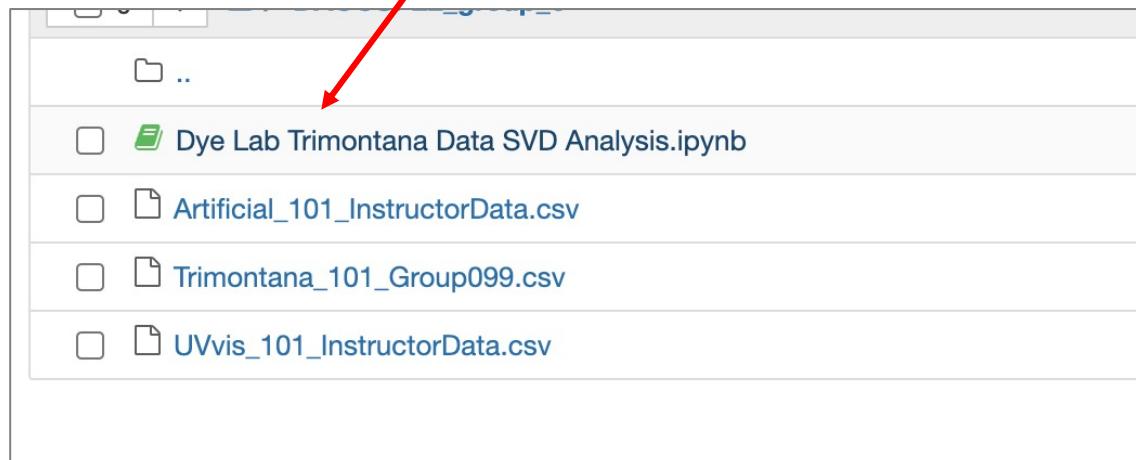


6.) Locate your group folder and upload your formatted data by dragging and dropping your data. If you have not already, name your data using the following format:

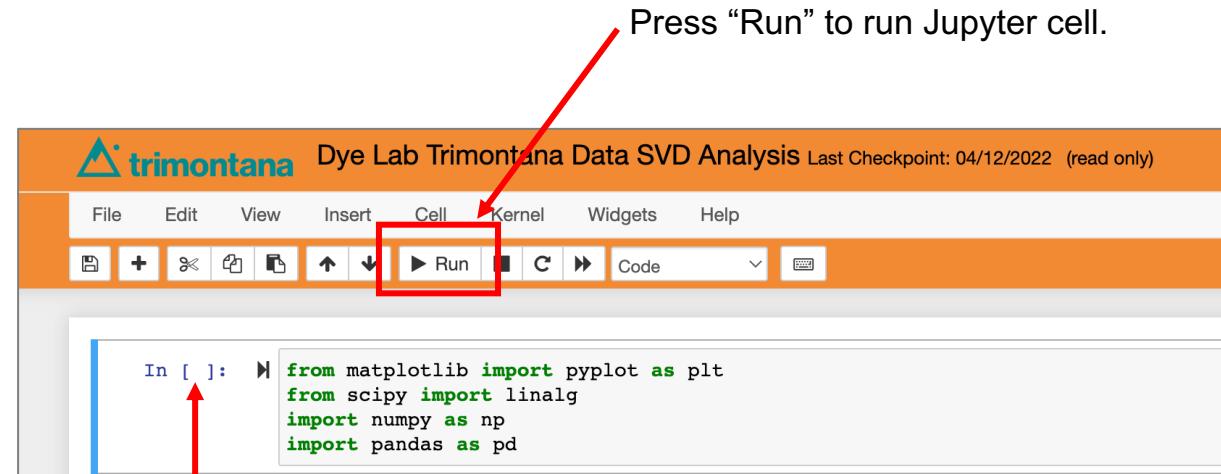
[DataType]_[CourseNumber]_[Group].csv

**In science and computational work, using organized naming conventions for your data is an important practice, as it allows you to more easily work with your data as you start to accumulate different data sets.

Open Jupyter Notebook for analysis and import Python libraries



7.) Open the “Dye Lab SVD Analysis” notebook in your group folder. You will follow along with the input and output for each section of the code in this notebook to perform your singular value decomposition.



8.) Run the first Jupyter cell (just as you ran your cells for data collection). This first cell will import the **Python libraries** you need to do your data analysis. A Python library contains collections of prewritten code that can be used to perform tasks without having to re-write code yourself. This allows you to simplify code you are writing.

Import data from your .csv file

```
In [1]: ┌─▶ from matplotlib import pyplot as plt  
      from scipy import linalg  
      import numpy as np  
      import pandas as pd
```

```
In [ ]: ┌─▶ # Import .csv of absorbance data  
      abs_data = pd.read_csv(r'INSERT FILE NAME.csv', header=None)  
      print(abs_data)
```



```
In [1]: ┌─▶ from matplotlib import pyplot as plt  
      from scipy import linalg  
      import numpy as np  
      import pandas as pd
```

```
In [ ]: ┌─▶ # Import .csv of absorbance data  
      abs_data = pd.read_csv(r'Trimontana_101_Group099.csv', header=None)  
      print(abs_data)
```



```
In [2]: ┌─▶ # Import .csv of absorbance data  
      abs_data = pd.read_csv(r'Trimontana_101_Group099.csv', header=None)  
      print(abs_data)
```

	0	1	2	3	4	5	6	7	8	9
0	0.142	0.029	1.151	0.235	0.624	0.142	0.029	1.151	0.235	0.624
1	0.147	0.029	1.162	0.236	0.627	0.147	0.029	1.162	0.236	0.627
2	0.151	0.029	1.174	0.238	0.632	0.151	0.029	1.174	0.238	0.632
3	0.156	0.029	1.186	0.239	0.636	0.156	0.029	1.186	0.239	0.636
4	0.160	0.030	1.197	0.241	0.640	0.160	0.030	1.197	0.241	0.640
..
296	0.012	0.008	0.002	0.002	0.002	0.012	0.008	0.002	0.002	0.002
297	0.011	0.008	0.002	0.002	0.002	0.011	0.008	0.002	0.002	0.002
298	0.010	0.007	0.002	0.002	0.002	0.010	0.007	0.002	0.002	0.002
299	0.010	0.007	0.002	0.002	0.002	0.010	0.007	0.002	0.002	0.002
300	0.009	0.006	0.002	0.002	0.002	0.009	0.006	0.002	0.002	0.002

[301 rows x 10 columns]

9.) Change the file path found in the second Jupyter cell. Then, run this second Jupyter cell to import your data from the .csv file you uploaded. You'll be doing this with the Python library Pandas, which we renamed "pd" when we imported it (for simplicity).

Notice that for this cell, we see an output. This is because we asked our code to print our imported data, using "print(abs_data)".

When programming, you do not have to print your data, but sometimes we do this to understand what our code is doing (here, to directly see what we're importing).

Calculate singular values using SciPy

```
In [3]: # Compute SVD
U, s, V = linalg.svd(abs_data, full_matrices=False)
print(s)

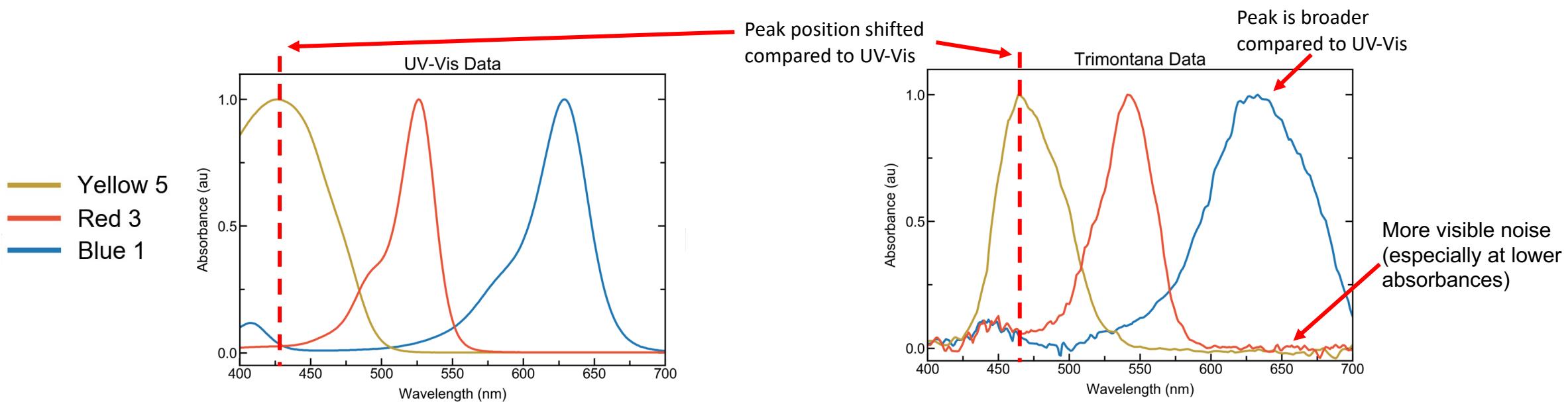
[2.38403839e+01 1.33475645e+01 1.05979117e+01 3.91544294e+00
 2.28138159e+00 8.38716397e-15 2.76771293e-15 1.00051700e-15
 4.75504827e-16 3.92644599e-16]
```

10.) Run the third cell to compute the singular values associated with your data. You'll be doing this with the Python library SciPy.

Considering noise when interpreting our data

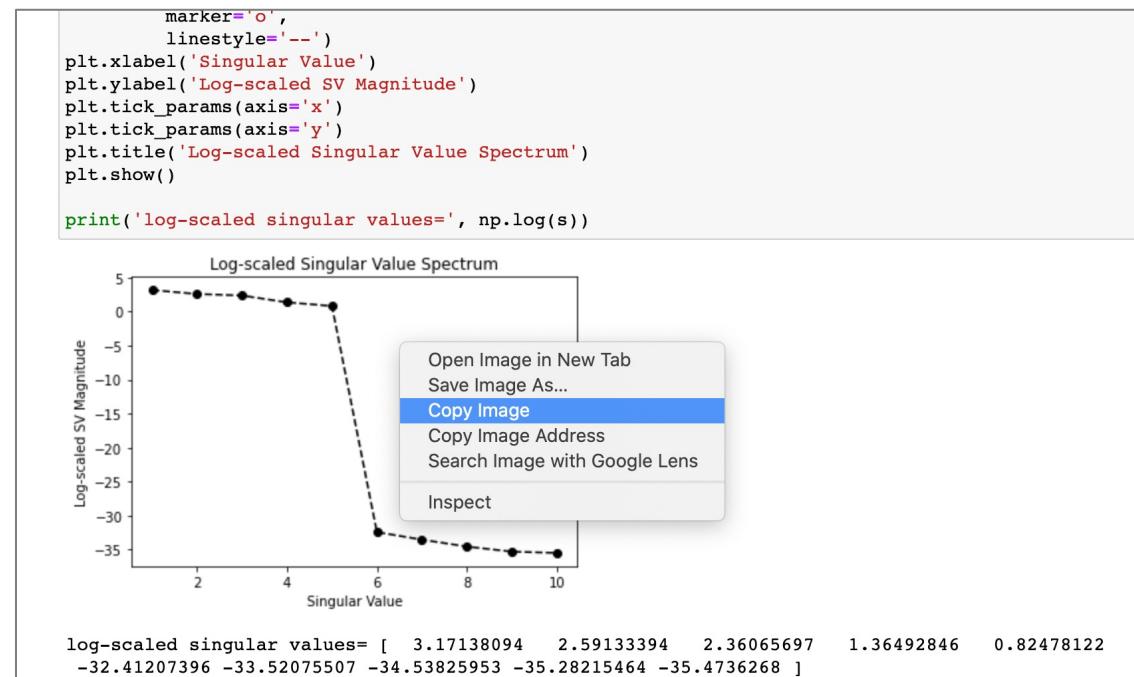
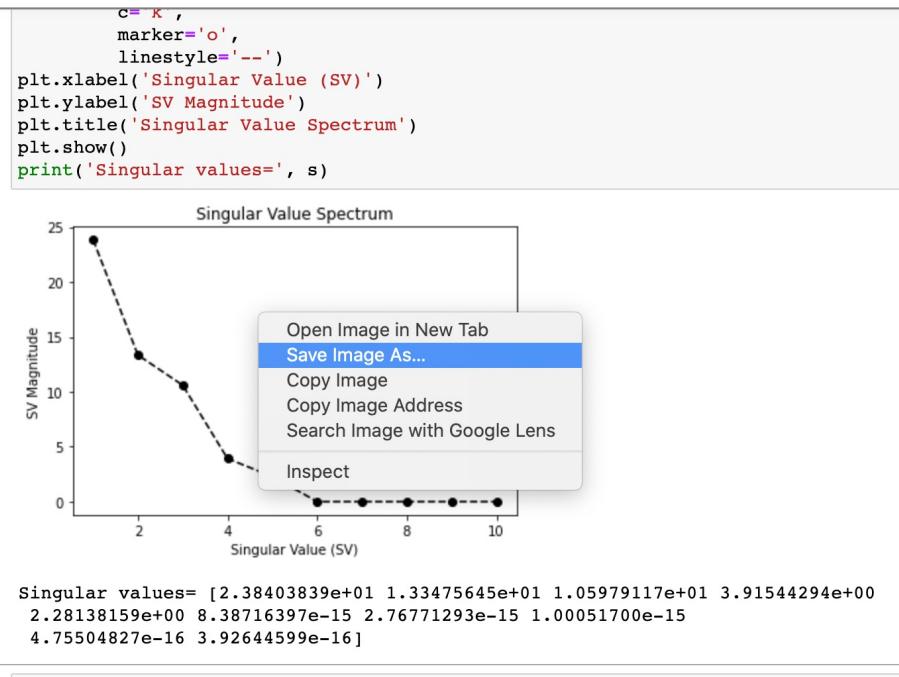
If our data was perfect, we would be able to easily distinguish between larger singular values corresponding more clearly to how many components we need to describe our solution, and smaller insignificant singular values on the order of zero.

However, our data from today is going to contain significant flaws. Just by looking at your plots, you can see fluctuations in the data from noise/the accuracy of our hand-built spectrometers. Comparing UV-vis and Trimontana Spectrometer data (shown below for the dyes from Part II) can help you better see some of these features.



Given this we will need to approximate which singular values should be considered and which should be excluded, as some non-zero singular values may be attributed to noise. There are several ways to do this, but your class will use the heuristics in the following slides (20 – 27) to do this.

Plot your singular values

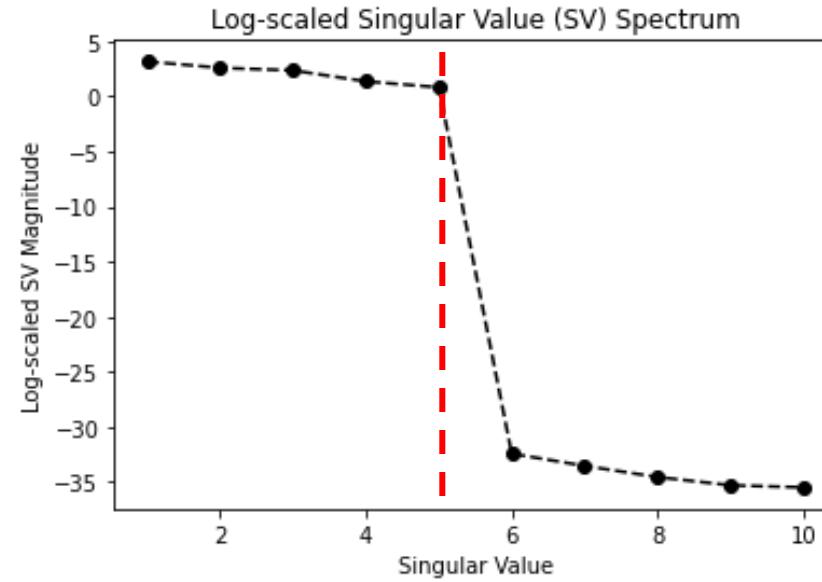
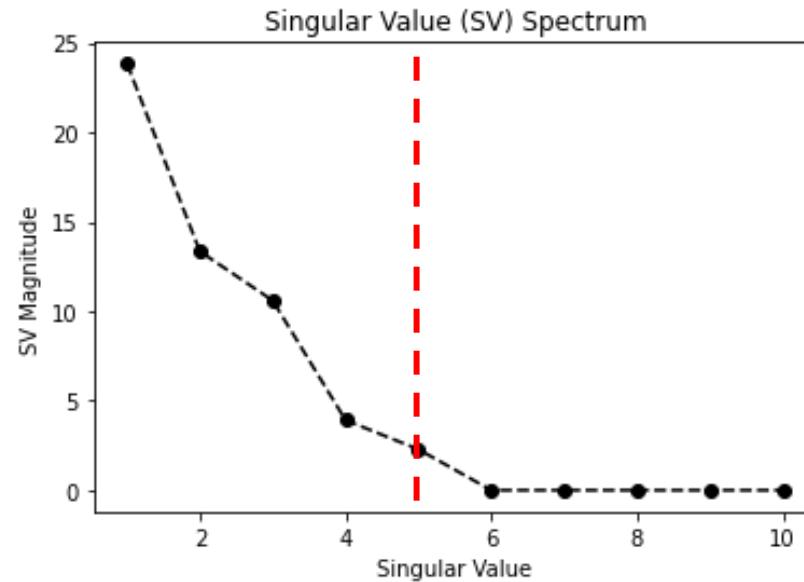


11.) Run the fourth and fifth cells to print and plot your data. Save your plots by either:

- (1) Right clicking on the plot and selecting “Save Image As...”
- (2) Copying your image into a Word document or your Excel file and saving it there
- (3) OR taking a screen shot of your plots.

Analyze your singular value plots to estimate significant values

You now have two plots for your singular values on a (1) standard and (2) natural log scale (this allows us to better distinguish smaller singular values, compared to the first singular value, which will be large).



Answer for this data set:
5 unique dyes

For your **SV spectrum**, find the point in your data where remaining singular values are close to 0.

For your **log scaled SV spectrum**, find the point where the remaining values are <0.

All singular values at or above this point will be considered significant and should give us a good approximation for how many dyes we have in our unknown mixtures. An example of this is shown above for a data set with 10 spectra analyzed via SVD, containing 5 significant singular values, and therefore 5 unique components across all unknown mixtures.

Confirm significant singular values (slides 23 – 26)

To better understand your analysis from the singular value plots (on the previous slide), you'll next determine **approximately what percentage of your data you're describing with these singular values**. You'll do this by finding the sum of all the singular values, then finding the sum of the significant/larger singular values you identified in part (a) above.

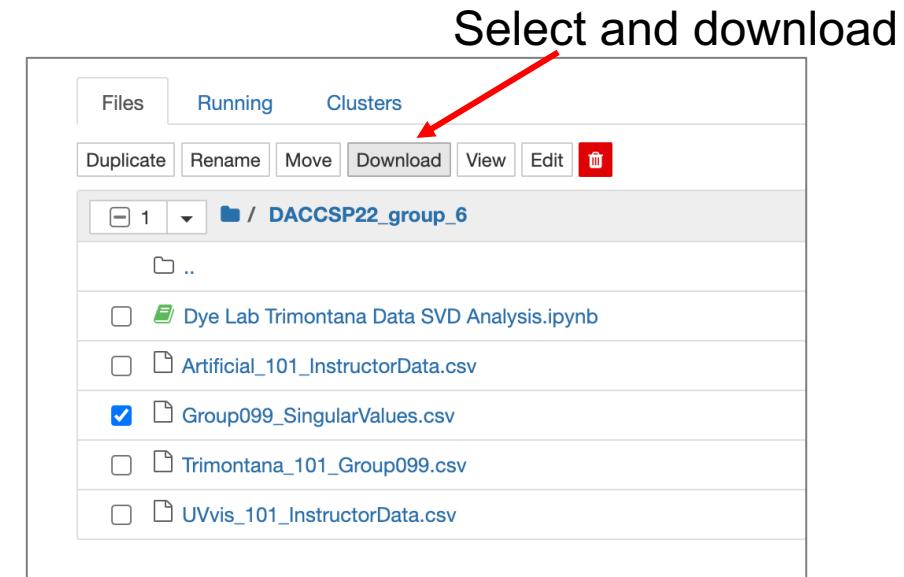
****The following slides will help you confirm how many components you have in your mixture. These calculations can be performed by hand. However, we'll walk you through how to download your data and perform the calculations in Excel. This is useful as it allows you to easily save/store your work electronically.**

Save singular values to a .csv file

```
-32.41207396 -33.52075507 -34.53825953 -35.28215464 -35.4736268 ]  
In [ ]: # Save singular values as a .txt file  
SV = np.array(s)  
np.savetxt("SingularValues.csv", SV)
```

↓

```
-32.41207396 -33.52075507 -34.53825953 -35.28215464 -35.4736268 ]  
In [ ]: # Save singular values as a .txt file  
SV = np.array(s)  
np.savetxt("Group099_SingularValues.csv", SV)
```

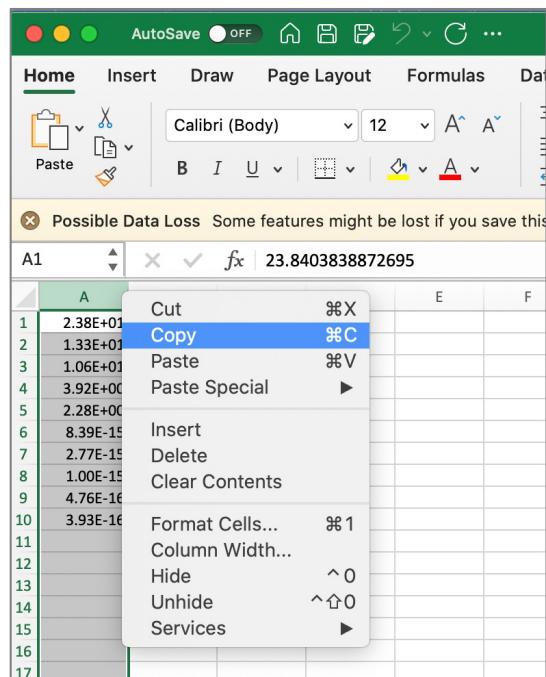


12.) Change the filename found in the final cell to indicate which set of singular values you are calculating (e.g., "Group099_SV.csv" or "UVVis_SV.csv"). Run the final cell to save your singular values as a .csv file.

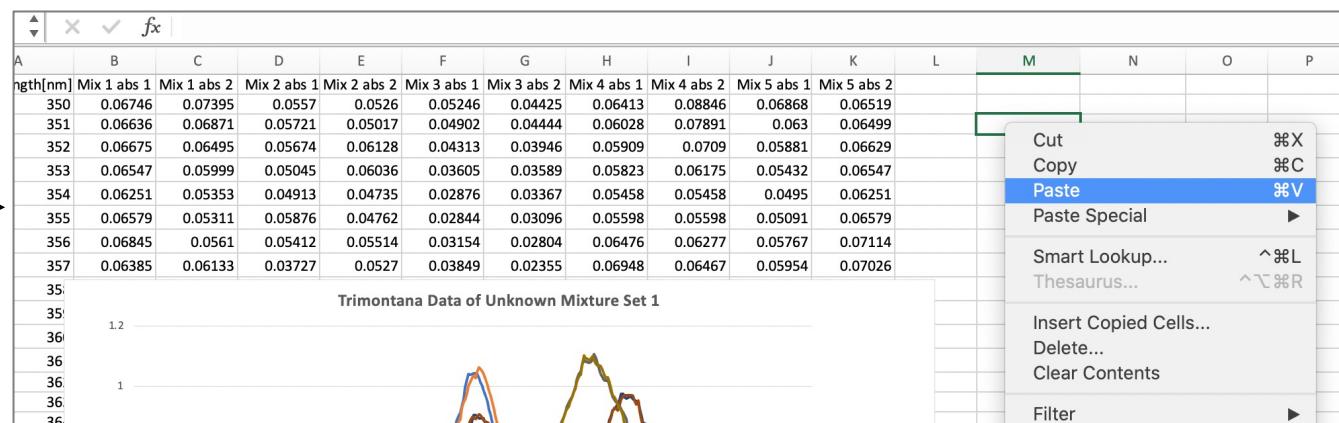
13.) Download the .csv file containing your singular values.

Copy your singular values to Excel for analysis

Open .csv and copy values



Open Excel file with your organized data and paste values



10 singular values now in Excel file

A screenshot of an Excel spreadsheet showing the raw singular values pasted into a new column. The data is organized into columns I, J, K, L, and M. A legend at the bottom indicates two series: 'Mix 1 abs 1' (blue line) and 'Mix 1 abs 2' (orange line). The data ranges from approximately 10^0 to 10^-16.

I	J	K	L	M	N
4 abs 2	Mix 5 abs 1	Mix 5 abs 2			
0.08846	0.06868	0.06519			
0.07891	0.063	0.06499			
0.0709	0.05881	0.06629			
0.06175	0.05432	0.06547			
0.05458	0.0495	0.06251			
0.05598	0.05091	0.06579			
0.06277	0.05767	0.07114			
0.06467	0.05954	0.07026			

14.) Open your .csv in Excel. We recommend copying this data into your original Excel file from **slide 11**, so that you have your plots/analysis in one place.

Note: The data contained in your .csv file will be your raw singular values (and NOT the log scaled singular values). These are the values you want to calculate the percentage of your data you are describing, as the log-scaled plot was just useful for visualizing trends in our data for smaller singular values.

Use Excel to calculate the sum of all singular values and significant singular values

a. Type “=SUM(” in a cell where you’ll calculate the total sum

M	N	O
	Total SUM	Sum Values 1-5
23.84038389	=SUM(
13.34756454	SUM(number1, [number2], ...)	
10.59791166		
3.915442941		
2.281381588		
8.38716E-15		
2.76771E-15		
1.00052E-15		
4.75505E-16		
3.92645E-16		

b. Click first cell, then drag cursor down to select cells you want to sum over

M	N	O
23.84038389	=SUM(M3:M12)	
13.34756454	SUM(number1, [number2], ...)	
10.59791166		
3.915442941		
2.281381588		
8.38716E-15		
2.76771E-15		
1.00052E-15		
4.75505E-16		
3.92645E-16		

10R x 1C

c. Close parentheses (here, the cell should read “=SUM(M3:M12)”)

Then press enter/return to populate the cell with your calculation. You can see both the answer in the cell as well as the calculation being done in the top left corner of your Excel window.

fx =SUM(M3:M12) ← Function for selected cell	
F	G
2 Mix 3 abs 1	Mix 3 abs 2
0.05246	0.04425
0.06413	0.08846
0.06868	0.06519
0.063	0.06499
0.05881	0.06629
0.06629	0.06517
0.06517	0.06517
23.84038389	53.98268462

Answer to calculation

d. Repeat in another cell for your significant singular values

K	L	M	N	O
abs 2		Total SUM		
06519		23.84038389	53.98268462	=SUM(M3:M7)
06629		13.34756454		
06547		10.59791166		
06251		3.915442941		
06579		2.281381588		
07114		8.38716E-15		
07026		2.76771E-15		
		1.00052E-15		
		4.75505E-16		
		3.92645E-16		

15.) Calculate the sum of (1) All singular values and (2) the larger/significant singular values determined from your plots on slide 21. You can do this using the “SUM” function in Excel (example shown above).

Use Excel to calculate the fraction of data described by your singular values

- a. Choose a cell where you'll calculate the fraction of your significant singular values (S_5 here) over the sum of all singular values (S_{total} here).

You can divide values in Excel cells using the “ / ” command. An example of this is shown below for cell Q3 divided by cell N3.

M	N	O	P	Q
	Total SUM (S_{total})	Sum Values 1-5 (S_5)		
23.84038389	53.98268462	53.98268462	=O3/N3	
13.34756454				
10.59791166				
3.915442941				
2.281381588				
8.38716E-15				
2.76771E-15				
1.00052E-15				
4.75505E-16				
3.92645E-16				

- b. Press enter/return to see the fraction of your values. The example data shown below was done using artificial data (which you'll calculate yourself later), so you should expect your Trimontana data fraction to be different.

	Total SUM (S_{total})	Sum Values 1-5 (S_5)	S_5/S_{total}
23.84038389	53.98268462	53.98268462	1
13.34756454			
10.59791166			
3.915442941			
2.281381588			
8.38716E-15			
2.76771E-15			
1.00052E-15			
4.75505E-16			
3.92645E-16			

- 16.) Divide the sum of your larger singular values by the sum of all your singular values. You can do this in Excel following the example shown above. For the Trimontana data, you should get a fraction above 0.9. This fraction indicates roughly what percentage of our data we can describe with the largest singular values (e.g. 0.95 = 95% of the data can be described using these singular values). Given the amount of noise associated with the Trimontana spectrometers, accepting the fewest number of singular values that will describe over 90% of your data should give you the approximate number of dyes in solution.

Repeat with UV-Vis and Artificial Data to see how noise affects your analysis

17.) Repeat steps 9 – 16 with the two data sets listed below. These data sets will be provided by your instructor, found in the course materials folder.



[UVvis_101_InstructorData.csv](#)



[Artificial_101_InstructorData.csv](#)

a. Absorbance data for the same 5 mixtures collected on a commercial UV-Vis Spectrometer (file name should start with “UVvis”)

b. Artificial/simulated mixtures of the same dye(s) (file name should start with “Artificial”)

Given the low amount of noise associated with the UV-Vis data and the Artificial Data when compared to the Trimontana Spectrometer, accepting the fewest number of singular values that will describe over 95% of your data should give you the approximate number of dyes in solution.