Predicting the Critical Temperature of Superconductors

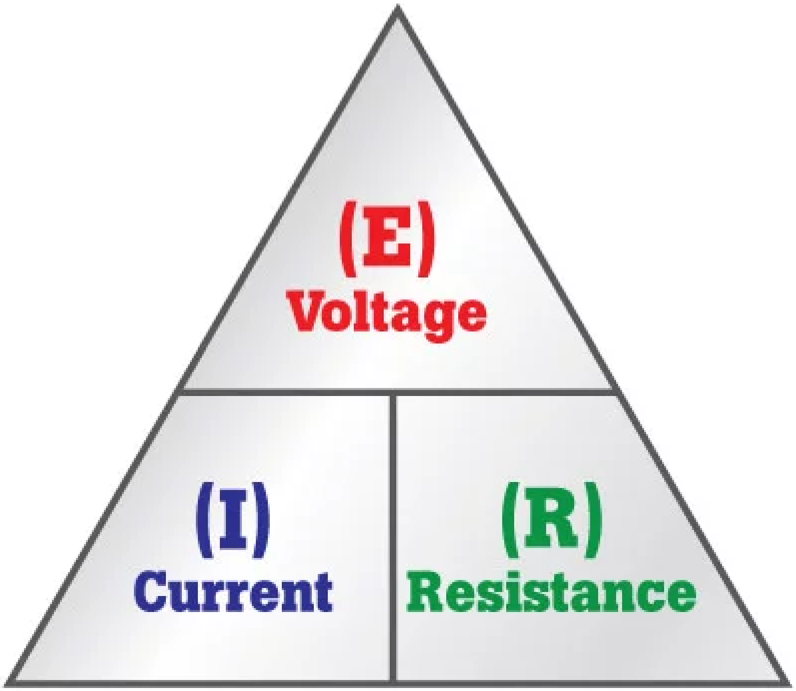
Yonaton Heit

Objective

- Predict the critical temperature for superconductors.
- Data Source:
  - UCI Machine Learning Repository database of superconductors and extracted properties
  - http://archive.ics.uci.edu/ml/datasets/Superconductivty+Data

What is a Superconductor

- Superconductor are materials with zero resistance.
- With zero resistance, electronic current can be maintained without external voltage
- With zero resistance, an electronic current can be maintained indefinitely.

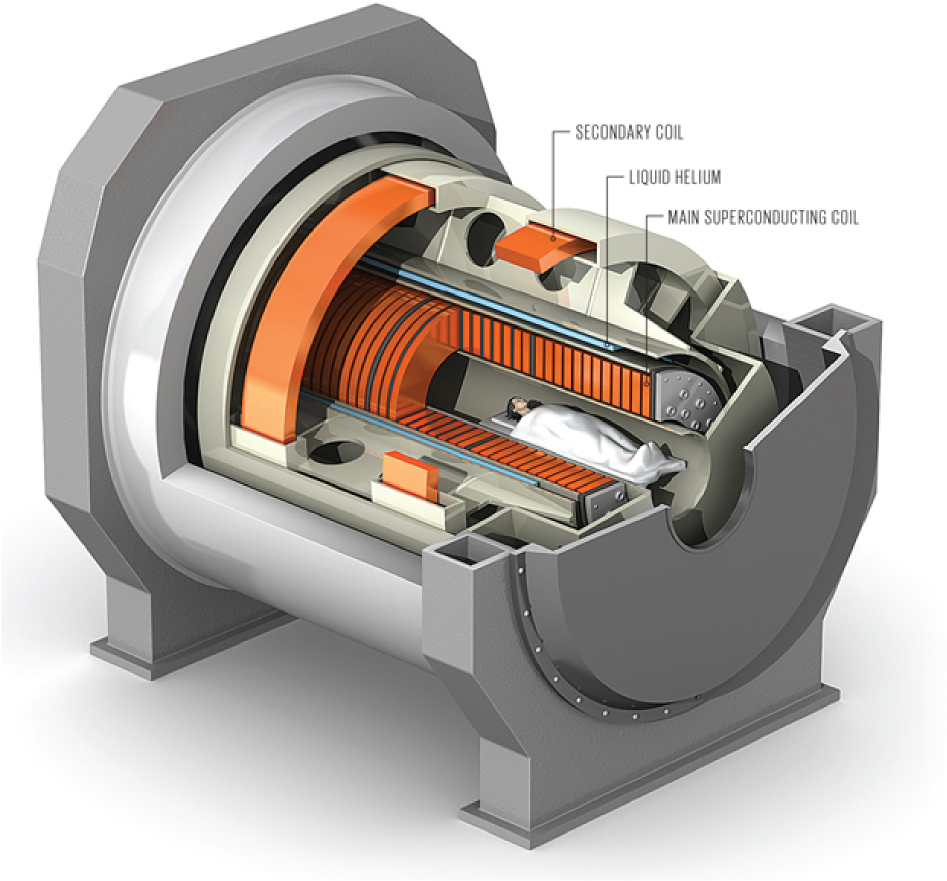


Source:https://www.fluke.com/en-us/learn/best-practices/measurement-basics/electricity/what-is-ohms-law

Applications for Superconductors

- Superconducting magnetic in Magnetic Resonance Imagining. (MRI)
- Superconducting coils in the Large Hadron Collider
- Superconductors could replace components in electronic powered systems [1]

MRI machine

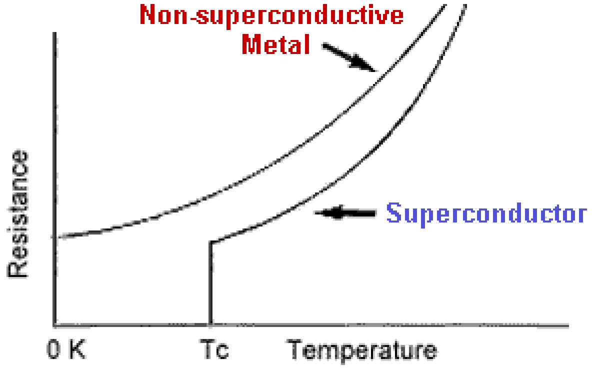


Source: https://phys.org/news/2013-10-world-powerful-mri-online.html

[1] W. V. Hassenzahl, et. al. *Proceedings of the IEEE*, **92** (10),1655-1674, 2004

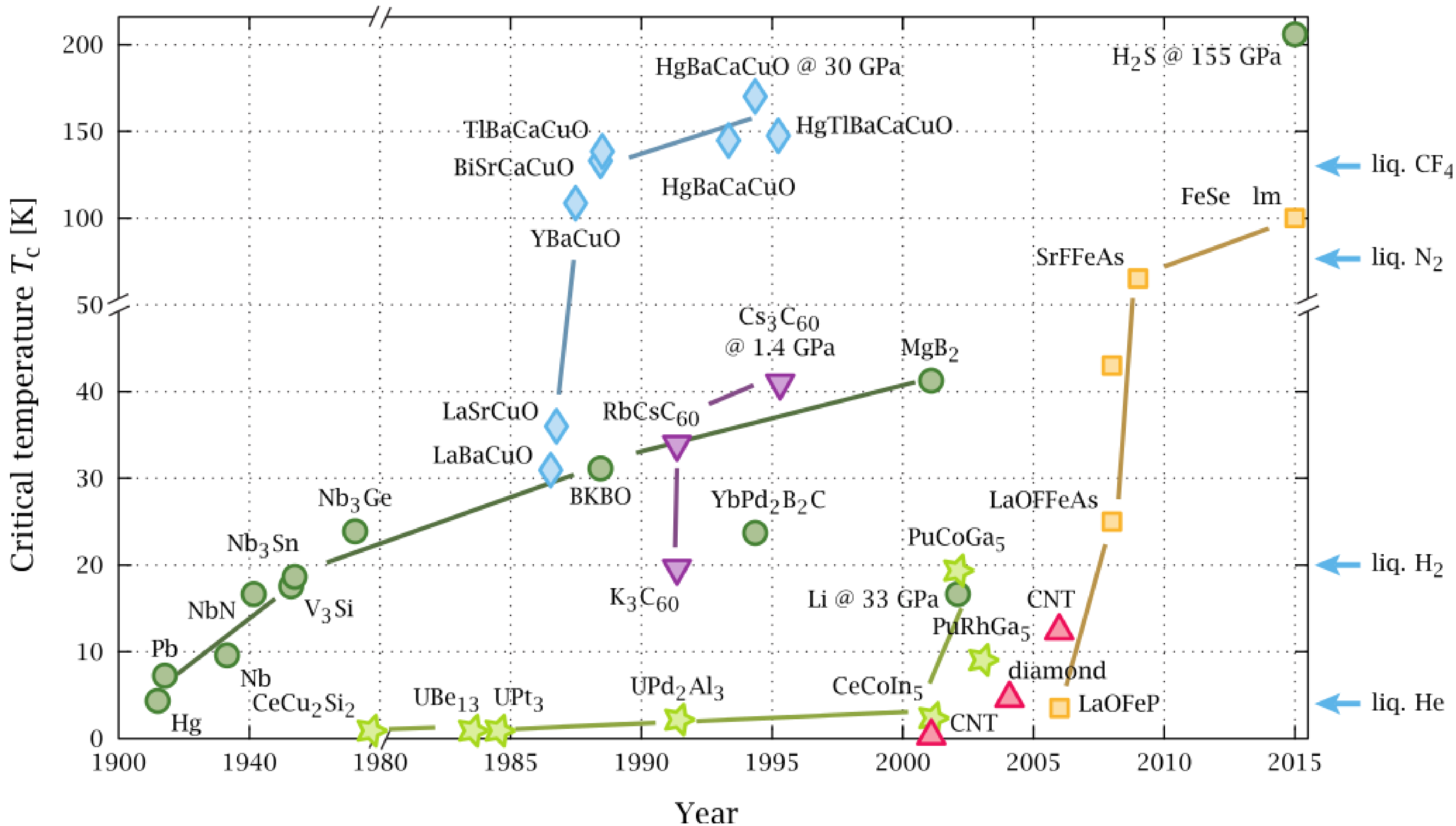
Critical Temperature

- Superconductivity can only be maintain below a certain temperature.
- This is called the critical temperature.
- Superconductors have to refrigerated in order to maintain superconductivity in, for example,
  - Liquid Helium (4 K)
  - Liquid Nitrogen (77 K)



Source: http://www.superconductors.org/tc_graph.gif

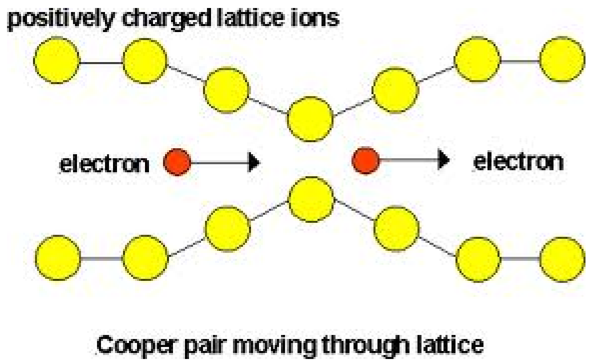
Critical Temperatures



Source: https://en.wikipedia.org/wiki/Superconductivity

Superconductor Theory

- There is not universal theory for superconductivity.[1]
  - In 1957, Bardeen, Cooper, Schrieffer (BCS) theory was proposed.[2]
    - Electrons are bound as Cooper Pairs.
    - Works well for low temperature superconductors (type I)
    - Not so well for higher temperatures superconductors which were later discovered (type II)
  - Other theories include
    - Resonating-valence-bond theory[3]
    - Spin fluctuation theory [4]



Source:https://physics.stackexchange.com/questions/126742/do-all-the-electrons-form-cooper-pairs-at-absolute-zero

[1] A. Mann, *Nature*, **475**, (21), 280-282, 2011
[2] J. Bardeen, et al. *Phys. Rev.* **106** (5), 162–164, 1957
[3] P.W. Anderson, *Science*, **235** (4793), 1196–1198, 1987
[4] P. Monthoux, et al. *Phys. Rev. Lett.*, **67** (24), 3448–3451, 1991

Data

- With no universal theory to predict critical temperature, we have to derive them from regression.
- Data set contains:
  - 21,263 superconductors
  - 81 feature
    - 8 properties derived from the elemental components
    - 10 statistical measurements determined from the 8 properties
    - 8 X 10 = 80
    - The final feature is the number of elements.
  - A complete list of the ratios elements for each superconductor (Not used by the original analysis by Hamidieh)

K. Hamidieh, *Computational Materials Science*, **154**, 346-354, 2018

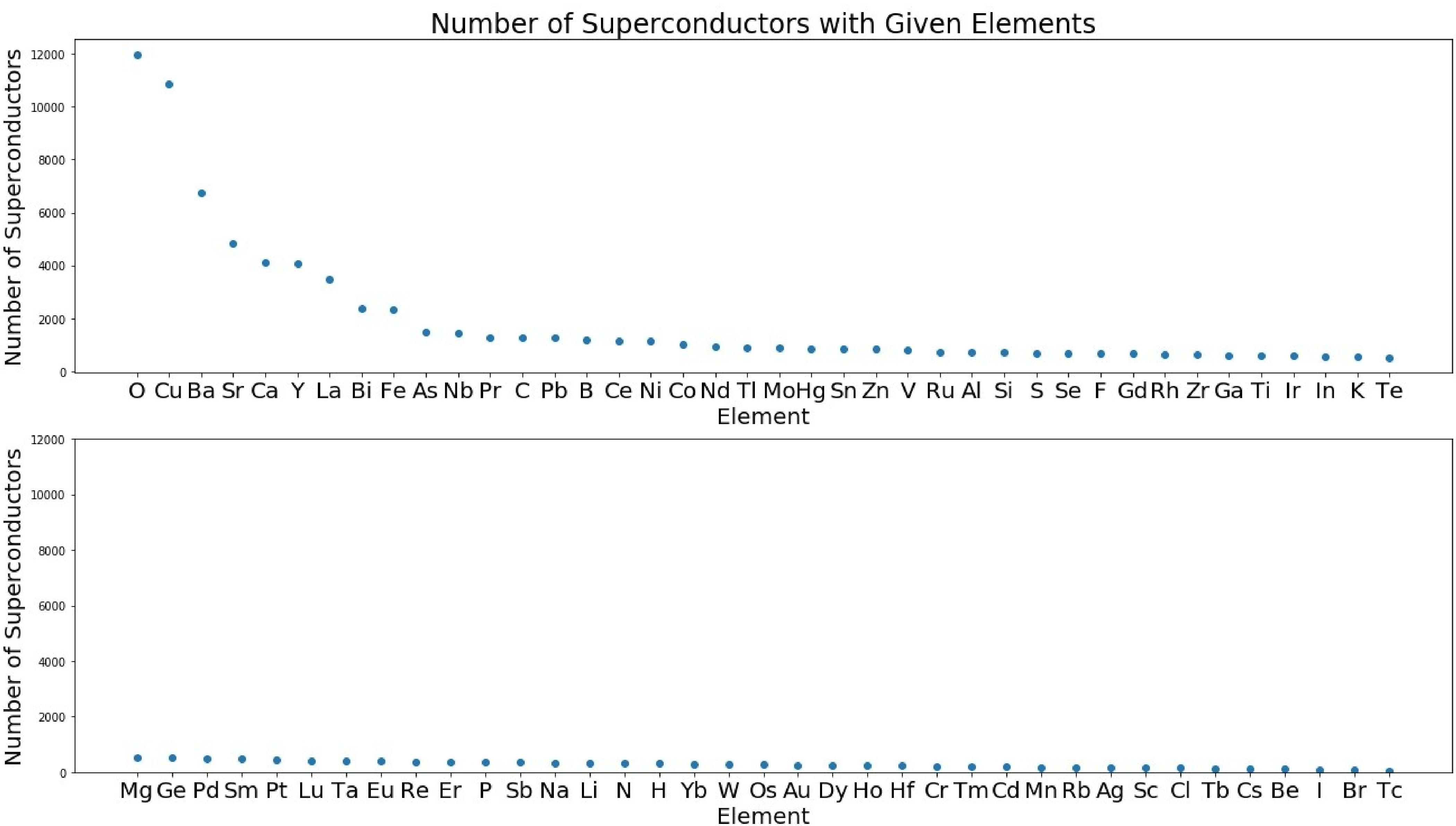
Attributes

Variable	Units	Description
Atomic Mass	atomic mass units (AMU)	total proton and neutron rest masses
First Ionization Energy	kilo-Joules per mole (kJ/mol)	energy required to remove a valence electron
Atomic Radius	picometer (pm)	calculated atomic radius
Density	kilograms per meters cubed (kg/m^3)	density at standard temperature and pressure
Electron Affinity	kilo-Joules per mole (kJ/mol)	energy required to add an electron to a neutral atom
Fusion Heat	kilo-Joules per mole (kJ/mol)	energy to change from solid to liquid without temperature change
Thermal Conductivity	watts per meter-Kelvin ($W/(m \times K)$)	thermal conductivity coefficient κ
Valence	no units	typical number of chemical bonds formed by the element

Feature & Description	Formula
Mean	$= \mu = (t_1 + t_2)/2$
Weighted mean	$= \nu = (p_1 t_1) + (p_2 t_2)$
Geometric mean	$= (t_1 t_2)^{1/2}$
Weighted geometric mean	$= (t_1)^{p_1} (t_2)^{p_2}$
Entropy	$= -w_1 \ln(w_1) - w_2 \ln(w_2)$
Weighted entropy	$= -A \ln(A) - B \ln(B)$
Range	$= t_1 - t_2 \ (t_1 > t_2)$
Weighted range	$= p_1 t_1 - p_2 t_2$
Standard deviation	$= [(1/2)((t_1 - \mu)^2 + (t_2 - \mu)^2)]^{1/2}$
Weighted standard deviation	$= [p_1 (t_1 - \nu)^2 + p_2 (t_2 - \nu)^2)]^{1/2}$

K. Hamidieh, *Computational Materials Science*, **154**, 346-354, 2018

Element analysis

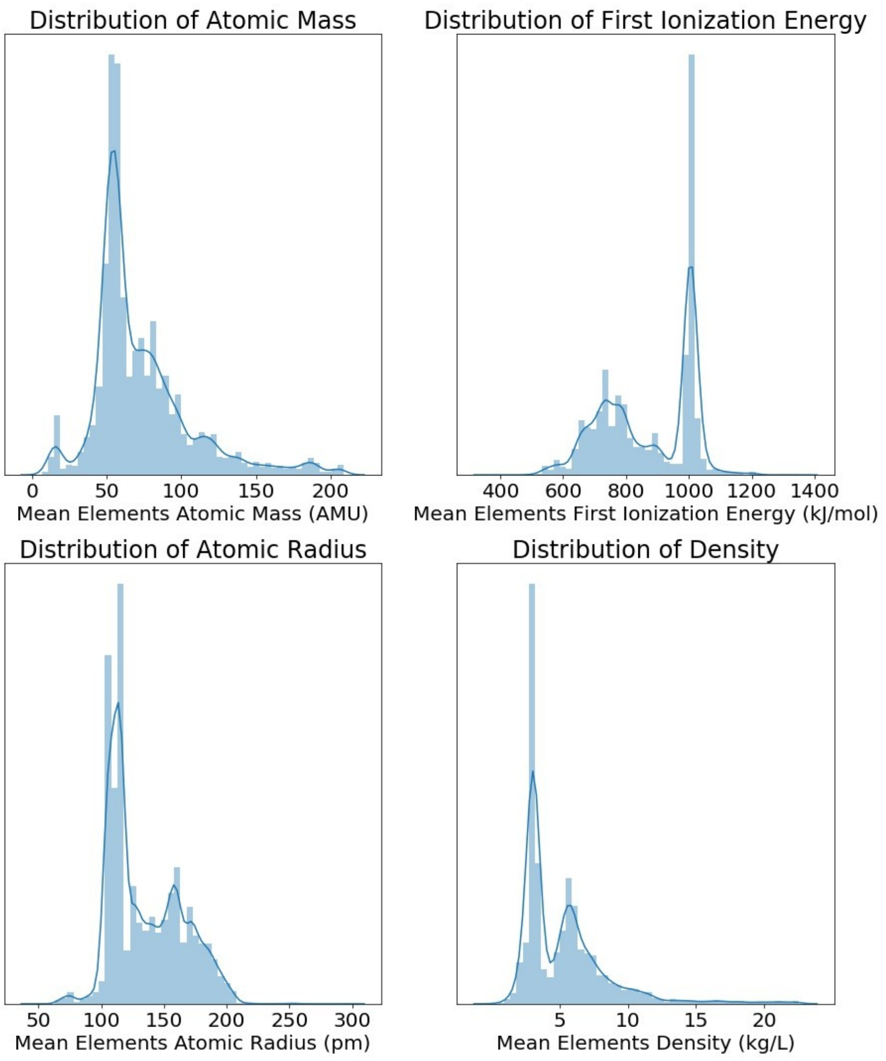


Element analysis

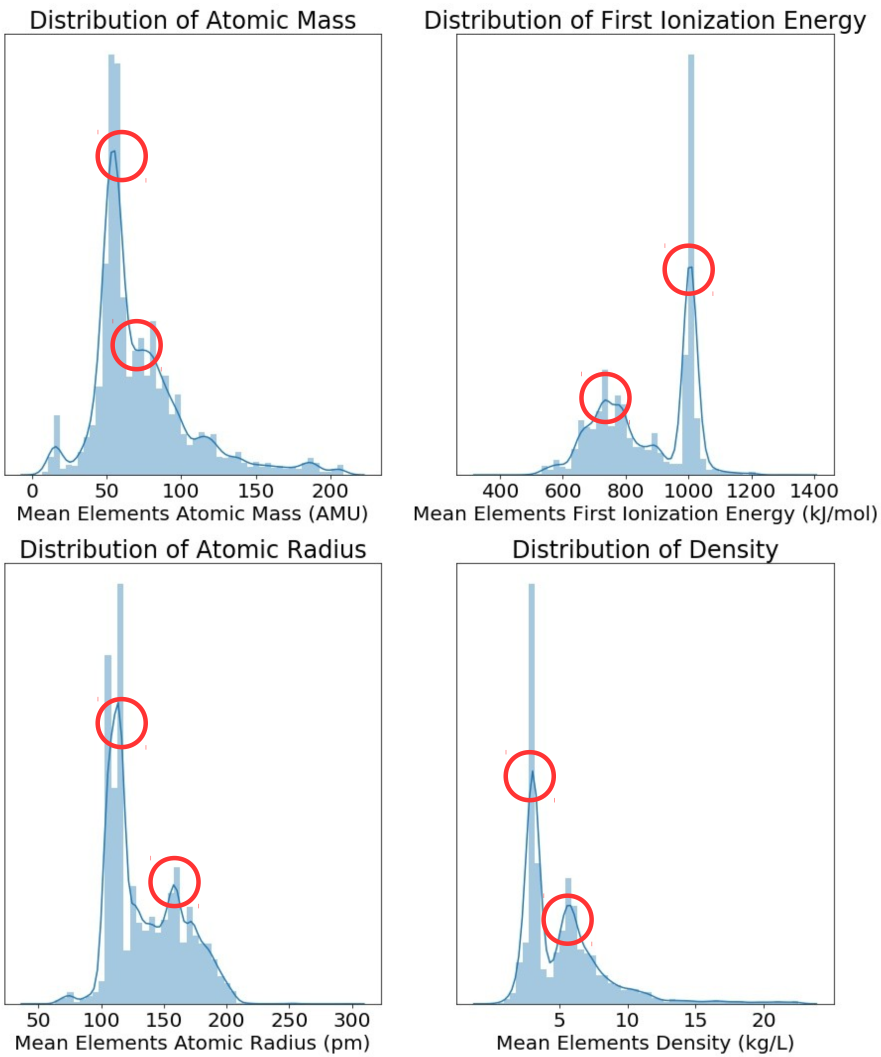
- There are 77 elements in the database.
- 60 elements appear in less than 5% of superconductors.
- Oxygen and Copper are the most common element
- 58.92% of superconductors have oxygen or copper.
- It would be interesting to see how well a model does using only elements.

Superconductors with Element		
	Percent	Number
Oxygen	56.27%	11964
Copper	50.97%	10838
Barium	31.75%	6751
Strontium	22.82%	4852
Calcium	19.34%	4112
Yttrium	19.16%	4075
Lanthanum	16.29%	3463
Bismuth	11.24%	2389
Iron	11.00%	2339
Arsenic	7.06%	1502

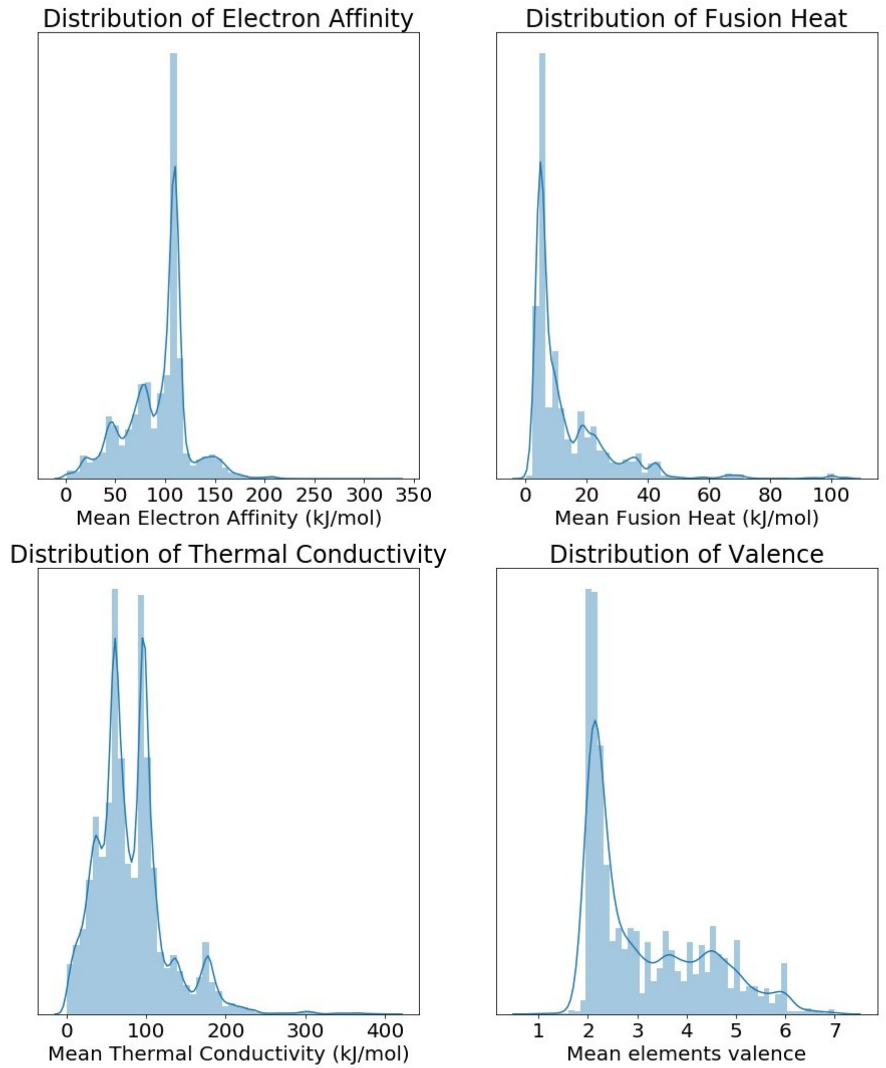
Property Distribution of the Weighted Mean



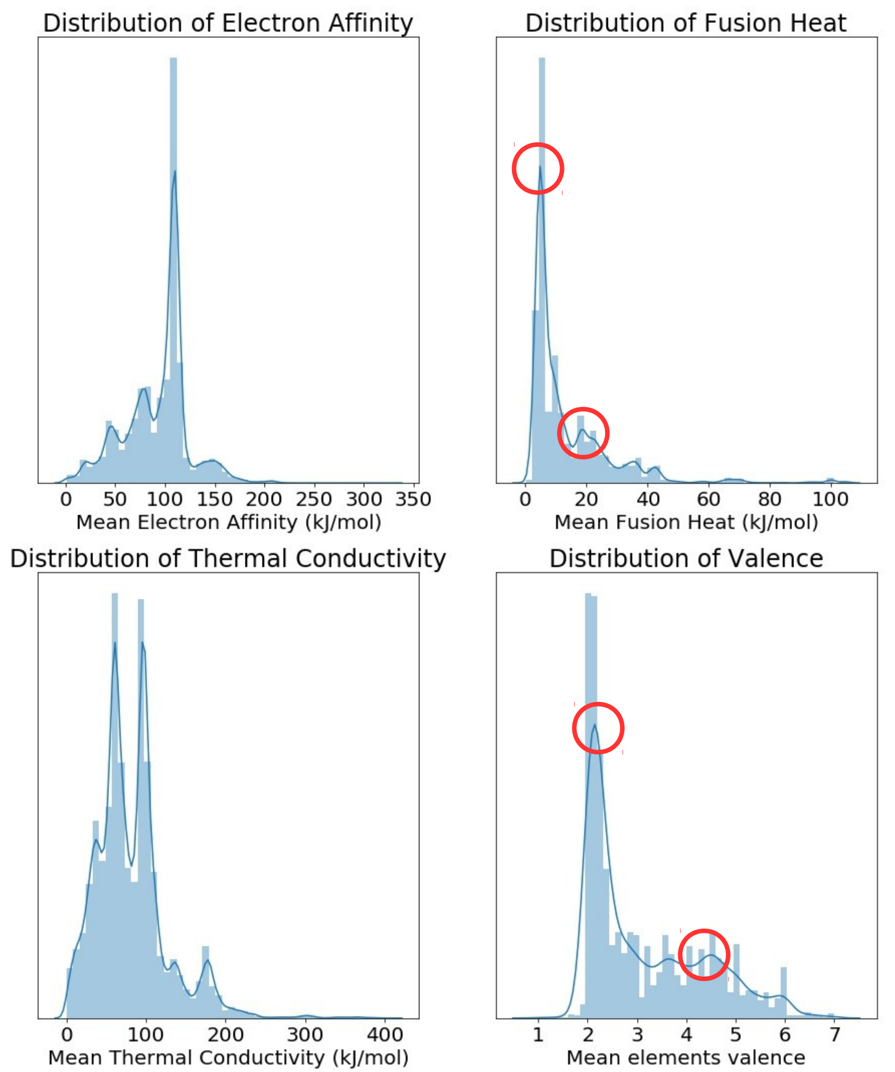
Property Distribution of the Weighted Mean



Property Distribution of the Weighted Mean

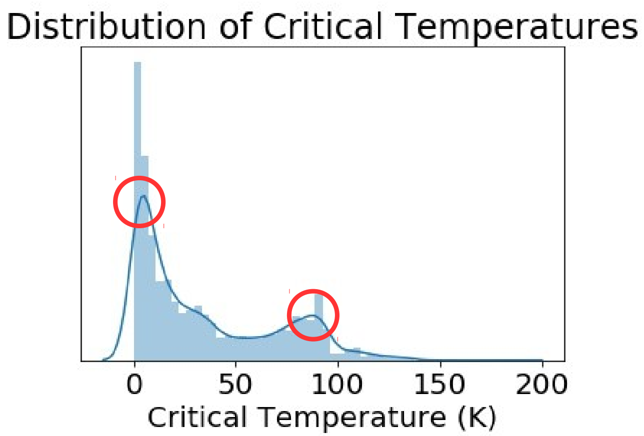


Property Distribution of the Weighted Mean



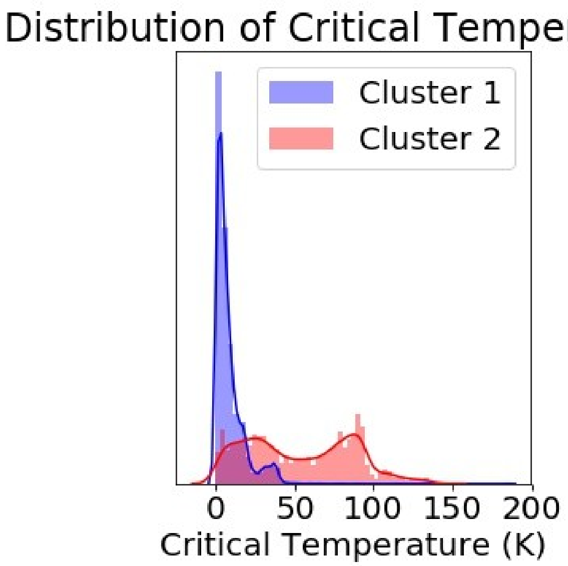
Critical temperature distribution

- The distributions suggest that there are two populations of superconductors in the data set.
- Possibly type I and type II superconductors.
- Type I are significantly lower temperature then type II.
- After clustering, there were too many type II superconductors in the lower temperature cluster for the clusters to be separated by type.

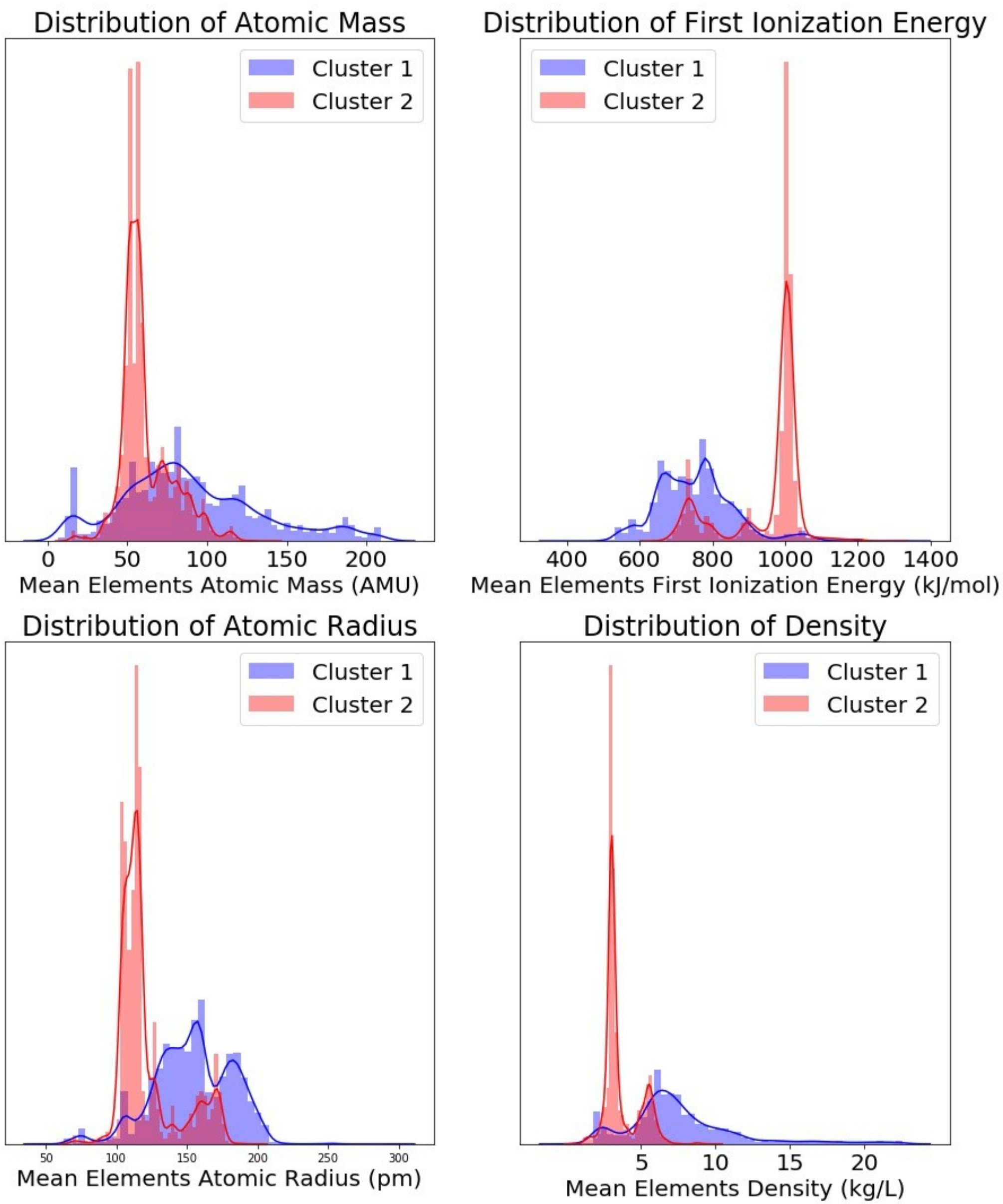


K-means Clustering

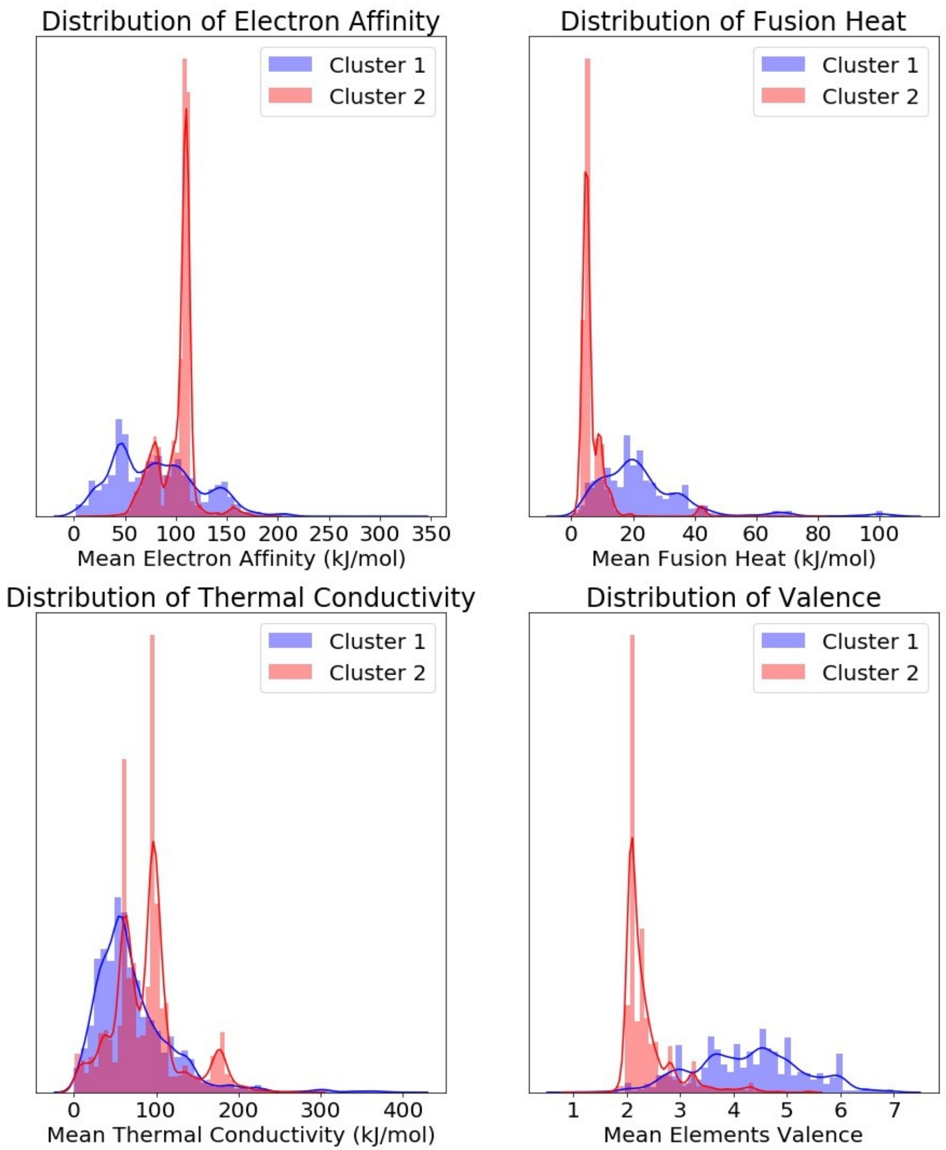
- Superconductors were clustered into two distinct clusters.
  - Cluster 1: 8792 superconductors
  - Cluster 2: 12471 superconductors
- Is there is any predictive value to clustering



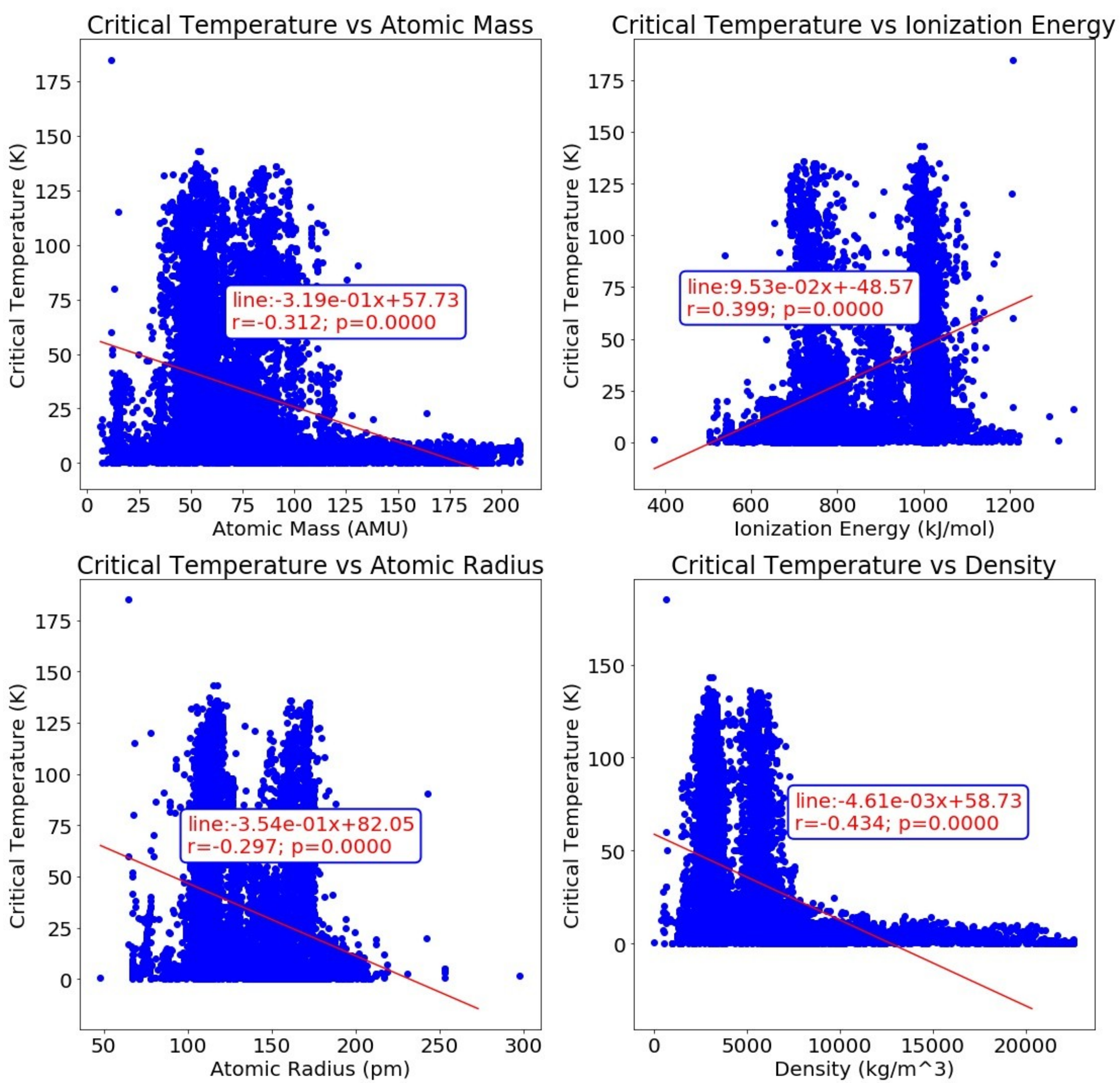
K-means Clustering



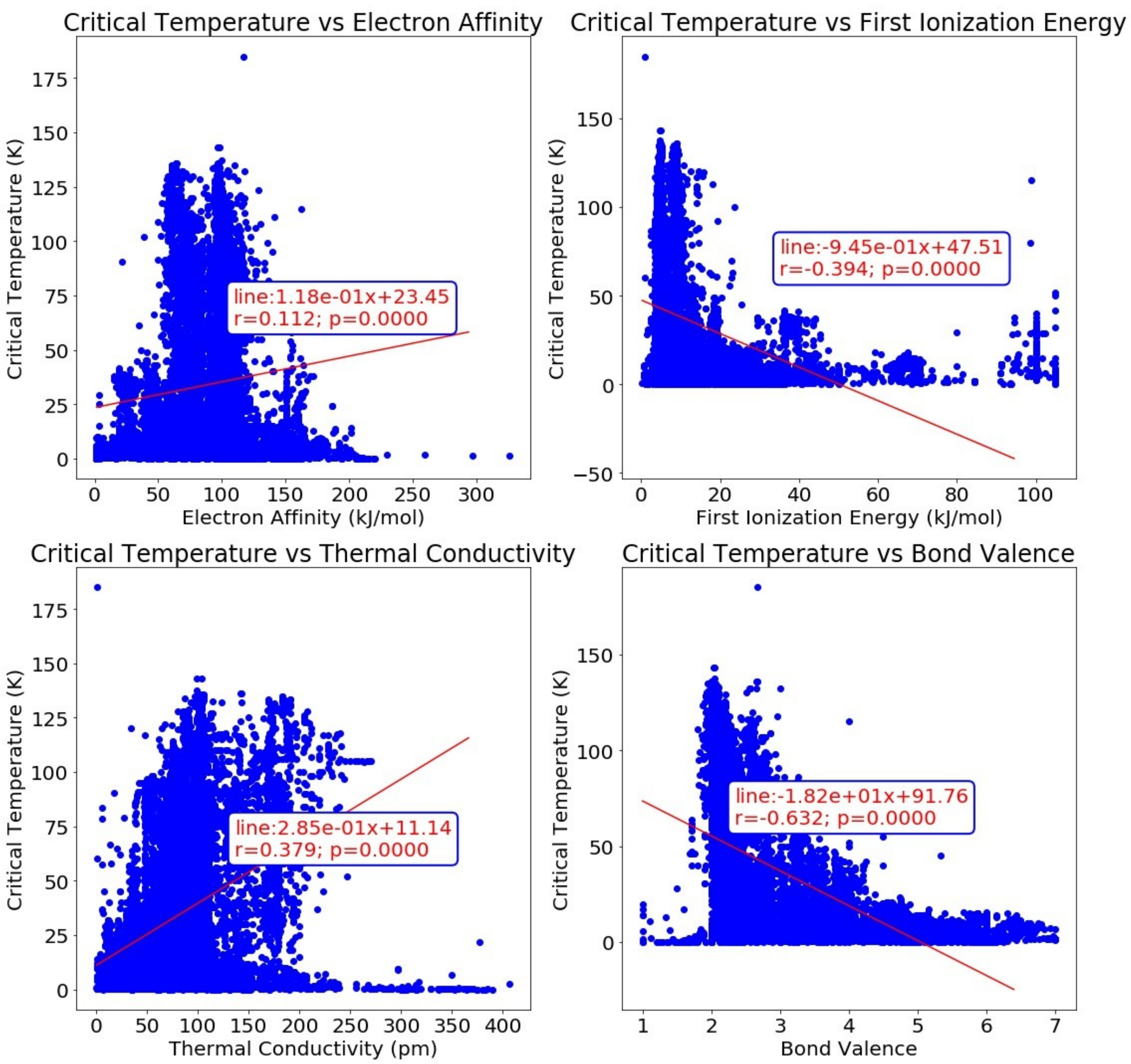
K-means Clustering



Scatter Plots of the Weighted Mean

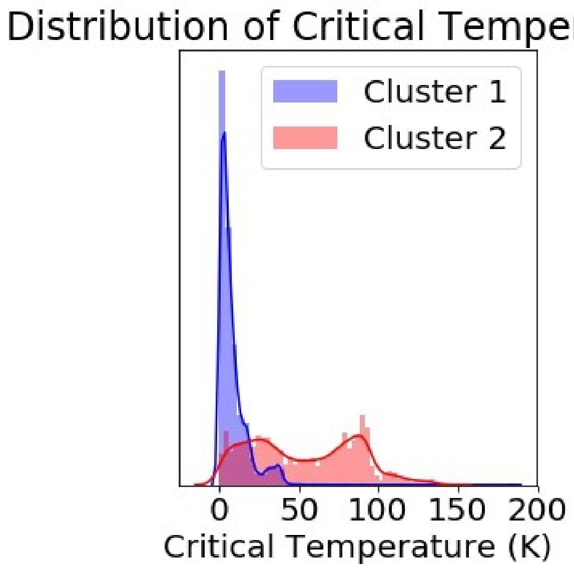


Scatter Plots of the Weighted Mean



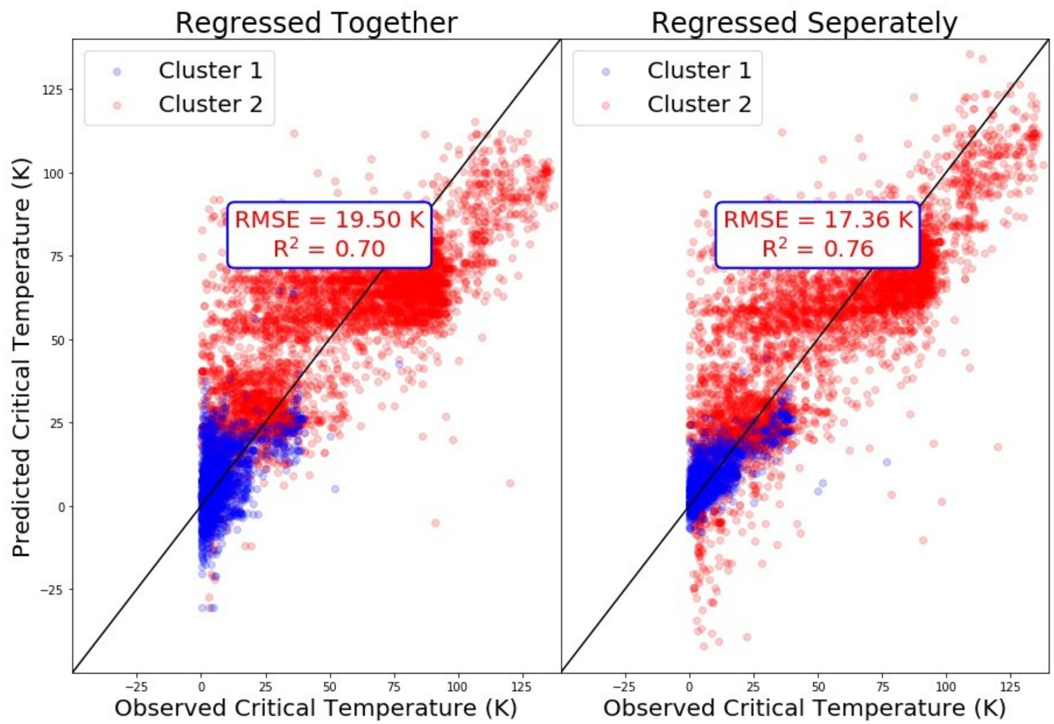
Data Separating

- 7000 data points from each cluster for training data set
  - Same number of data points per cluster to avoid unbalanced model.
- Two strategies used for each machine learning method.
  - Model clusters together.
  - Model each cluster separately
  - Allows us to see the effects of clustering.
- Two methods used for modeling:
  - Least Root Means Square Linear fit
  - Gradient Boosting



Linear Fit

- Performed using Scikit-Learn Python Libraries
- Features are scaled by
  - setting centering mean.
  - then dividing by standard deviation.



Modeling Cluster with Linear Regressions

- Modling Cluster separately reduced RMSE and R^2
  - Modeling Clusters Together: RMSE: 19.50 K

    R^2: 0.70
  - Modeling Clusters Separately: RMSE: 17.36 K

    R^2: 0.76
  - Cluster 1 Error is significantly decreased by modeling clusters separately
    - RMSE :9.58 K to 5.98 K
    - R^2: -0.21 to 0.53
  - Cluster 2 Error only sees smaller decrease:
    - RMSE :21.79 K to 19.71 K
    - R^2: 0.58 to 0.65

Gradient Boosting

- Performed using XGBoost Python Library
- 4-fold cross validation was performed using Scikit-learn Python Library
- XGBoost's Scikit-learn API was slow so custom API's were created.
  - This API was 5 times faster than the API build-in to XGBoost.
- Features were standardized
  - centered at the mean,
  - divided by standard deviation.
- Turned hyperparameters.

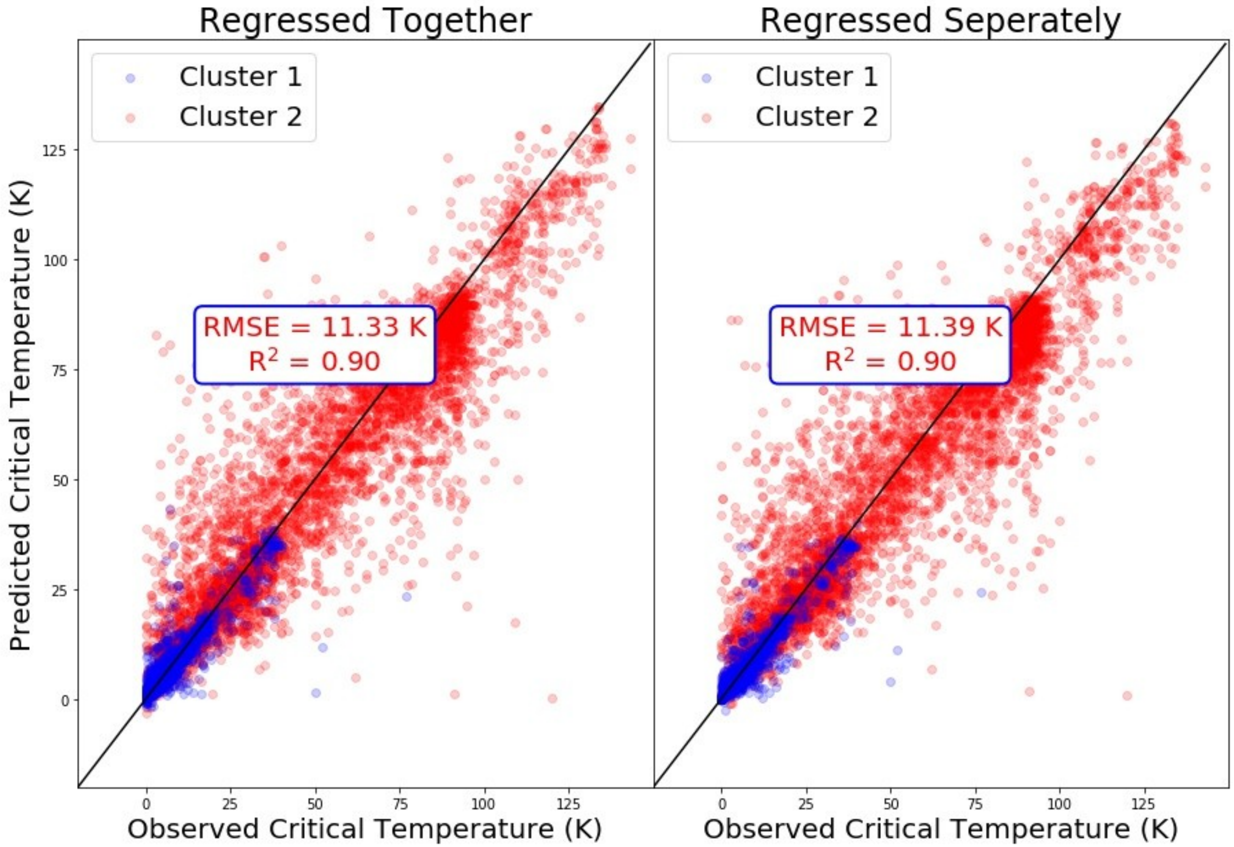
Tuning Hyperparameters

- Tuned hyperparameters with a random search of 200 possible combinations.

Parameters	Both Clusters	Cluster 1	Cluster 2	Values searched
n_estimators	230	130	139	1,2,3,…,699
min_child_weight	18	3	13	1,2,3,…,20
max_depth	17	18	12	1,2,3,…,20
learning_rate	0.40	0.28	0.34	0.01,0.02,0.03,…,1.00
subsample	1.00	0.75	1.00	0.50,0.75,1.00
solcample_bytrue	1.00	0.75	1.00	0.50,0.75,1.00

Gradient Boosting Results

- Gradient boosting performs significantly better than linear regressions.
- Modeling clustering separately had no affect on accuracy.



Important Features

Both Cluster		Cluster 1		Cluster 2	
Feature	Fraction Gain	Feature	Fraction Gain	Feature	Fraction Gain
Range thermal conductivity	0.594	Weighted mean atomic mass	0.253	Weighted mean thermal conductivity	0.460
Weighted geometric mean thermal conductivity	0.131	Range first ionization energy	0.114	Weighted mean valence	0.072
Standard deviation atomic mass	0.020	Weighted mean valence	0.054	Standard deviation atomic mass	0.062
Weighted mean valence	0.018	Weight geometric mean electron affinity	0.035	Weighted geometric mean valence	0.055
Weighted geometric mean valence	0.017	Mean first ionization energy	0.033	Weighted standard deviation electron affinity	0.036
Weighted standard deviation electron affinity	0.014	Weighted standard deviation thermal conductivity	0.026	Range atomic radius	0.014
Weighted range atomic mass	0.013	Mean density	0.024	Weighted entropy thermal conductivity	0.013

Conclusion

- We successfully created models to predict the critical temperature of superconductors from features derived from the properties of the elements.
- Best Model: RMSE: 11.33 K and R^2: 0.90
- Gradient boosting was more accurate than linear regression.
- Modeling the clustering separately improved accuracy for linear regression but not gradient boosting.
  - Separating by clustering may add flexibility to the linear regression model while gradient boosting (being non linear) was flexible enough.

Special Thanks

- Yogendra (Yogi) Pandey and Dipanjan (DJ) Sarkar, my mentors at Springboard, for all his help.
- Liam Doherty, for creating the template for these slides.

  https://github.com/dohliam/libreoffice-impress-templates