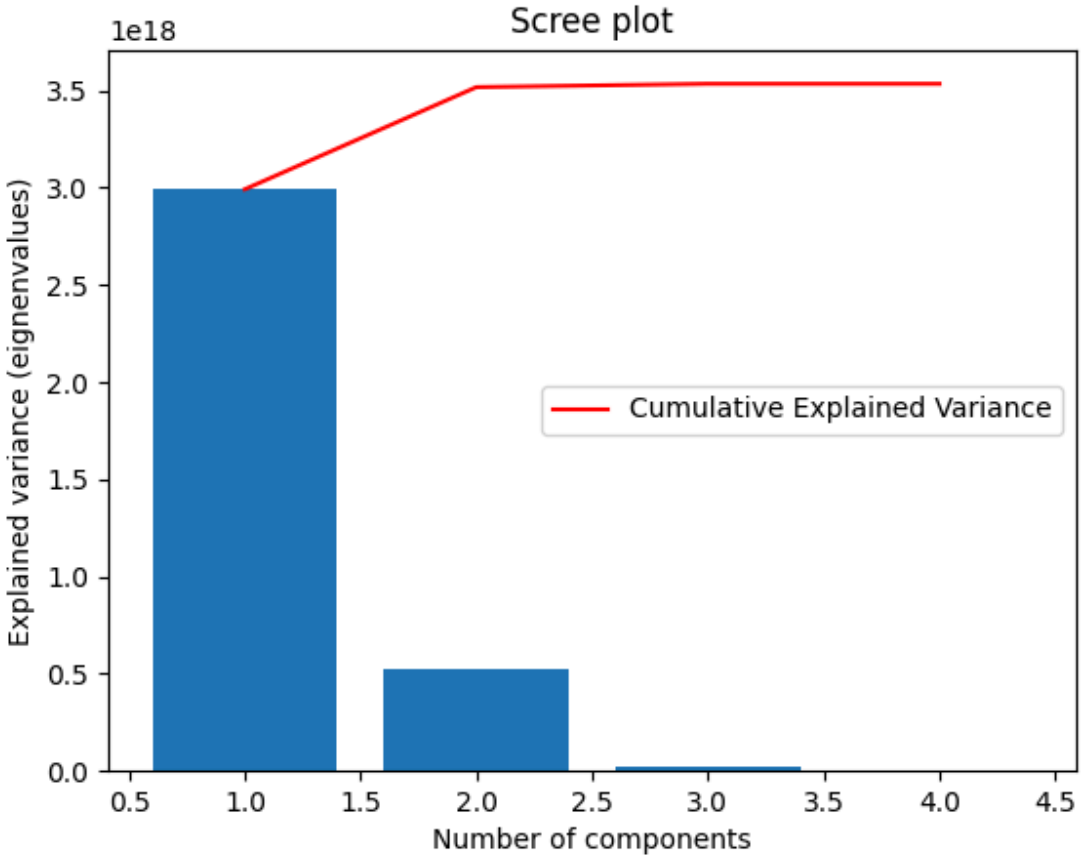
Principal Component Analysis or PCA is a method to compress or find the most relevant data in a data set without losing as much information as possible. It takes an existing data set and creates a new one that is more concise and represents the relevance of each of the features. For machine learning, it is useful for finding the most important features of a data set.

PCA is split into two parts.
Part 1: A PCA algorithm forms another table where each column is a “Principle Component”. Additionally, the columns are organized in descending order by importance (so the most important column would be the first one). Each Principal Component also has a “variance” value that represents how much of the data set it represents.

When checking for variance from the given data set, we got a list of `[0.84639744, 0.14871473, 0.00488448, …]`. This means that the first two columns already contain over 98% of the original data, so the rest of the columns have been discarded. This data is also shown in a bar graph below. The third component is barely visible and the fourth one cannot be seen at all. Note that there were a lot more components but they were removed to prevent the bar graph from being too wide.



Part 2: Each principal component contains a list of values that represents the relevance of each feature. It is ordered so the first value represents the first feature, the second value represents the second, and so on.

Using the two Principal Components from before, we created a dictionary for each of them where the keys are the feature names and the values are the feature's absolute values. The dictionary is then sorted from highest to lowest.

Principal Component 1:
`{'stcpb': 0.7079658396507997, 'dtcpb': 0.7057643493460535, 'Sload': 0.026084691665062904, …}`

Principal Component 2:
`{'dtcpb': 0.7082058173476752, 'stcpb': 0.7060060310060156, 'Sload': 6.641792324154551e-05, …}`

Both Principal Components show that "stcpb" and "dtcpb" are the most relevant while the remaining values begin to near zero. Note that this does not necessarily mean "stcpb" and "dtcpb" are the most relevant for other Principal Components. For example, if we did the same thing with the third Principal Component, then we would have gotten:
`{'Sload': 0.999657347116697, 'dtcpb': 0.018464025389709544, …}`
Where "Sload" is the most relevant feature for it.

Sources:

https://machinelearningmastery.com/feature-selection-machine-learning-python/
https://machinelearningmastery.com/calculate-principal-component-analysis-scratch-python/
https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html#sklearn.decomposition.PCA
https://towardsdatascience.com/pca-clearly-explained-how-when-why-to-use-it-and-feature-importance-a-guide-in-python-7c274582c37e
https://www.jcchouinard.com/pca-scree-plot/

https://www.youtube.com/watch?v=5vgP05YpKdE
https://www.youtube.com/watch?v=FD4DeN81ODY
https://www.youtube.com/watch?v=TJdH6rPA-TI