1. The dataset I will be using for this assignment will be regarding the detail of every parking meter charge in the city of Seattle. It includes the transaction ID, meter code, transaction date, payment type, amount, duration, blockface name, side of the street and many more. <https://data.seattle.gov/Transportation/Paid-Parking-Transaction-Data/gg89-k5p6/data#>
2. My research question is regarding whether or not cash should be eliminated from parking meters. I want to see whether or not there are enough cash payments to still justify the existence of it as a payment method. First of all, I would be using clustering based on the time taken and payment amount as axes labels, and use the payment types as the label. We then find the SVM of it and test it on the remaining dataset. If we could find a good cluster, we will then notice that the amount of payment and time taken to park will have an effect on the type of payment. If there is none, then the payment type chosen is random.
3. A prediction of type of payment based off how long a person parks there and how much he paid.
4. I will split the dataset into training and testing. Afterwards, I will find an SVM of the clusters, and test it on the dataset and check how many predictions are correct.