**Python 5**

1. What is the difference between a tuple and a list?
2. How to iterate through a dictionary?
3. Explain the usage of decorators

**Data Science  25**

1. Explain what precision and recall are. What does high recall mean?
2. Explain what ROC and AUC are. What does it mean if a classifier’s AUC is high?
3. Explain what a false positive and a false negative are. Why is it important to differentiate these from each other?
4. What is selection bias, why is it important and how can you avoid it?
5. How would you screen for outliers and what should you do if you find one?
6. How will you deal with categorical feature in regression problem?

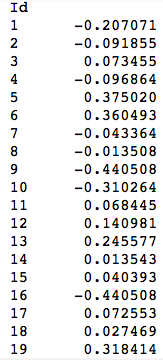
**Programing problem 30**

In this section, we will test your basic operation ability with Python Pandas, please open interview.ipython and finish the following task:

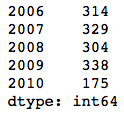
Q1. Get LotArea’s sum / mean / std



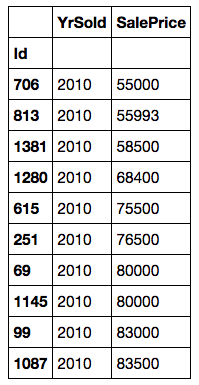
Q2. Perform z-normalization on LotArea



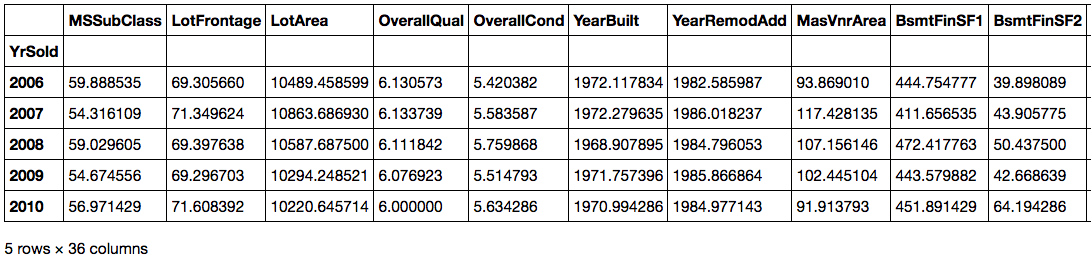
Q3. Get each class's amount in YrSold



Q4. Sort YrSold descending(2010-->2006), SalePrice ascending (cheap to expensive) , and fetch top 10 result



Q5. Groupby YrSold and calculate its mean



Q6. Create a new feature called YrTarget, if YrSold>2008, label it as 1, otherwise 0