Approach

The basic idea was to get the latest trend for each patient. So i tried out different approaches as follows:

1. Taking the last year data and building cross tab from that. Leftover patients were imputed using all the data.
2. Taking last six month data and doing the similar process as previous one.
3. **Next i divided every year into first 4 months and last 8 months. So i started from last 8 months of 2013 and did the similar crosstab function and then leftover patients were added from first 4 months. Further leftover patients were added from last 8 months of 2012 year and subsequently until we get all the patients data. (FINAL MODEL)**
4. Giving weights to the every year events such that recent events have more weight.
5. Frequency wise segmentation of patient: patients which were frequent, their prediction was done from the recent data and for the rest of the patients we used the overall data.
6. I tried to create training and test such that i could solve it as a machine learning problem, so i took out the most frequent events and for each patient, features were generated as the count of that event for each patient. And the last entry for each patient as the target variable.

Quality checks performed / Errors found:

No errors were found.

Feature Extraction

1. From Date
    a. **Year**
    b. **Month**

Key Trends

Latest information about the patients was showing better results rather than taking data from 2011-12 year.

Model choice explanation:

I used a rule based method for this problem. I tried using markov but was not able to generate higher order markov chains. So rule based method, was performing best for me.