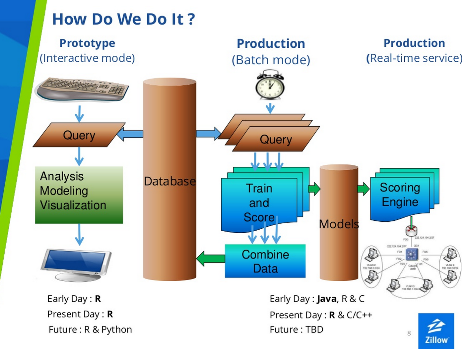
1. 
2. Work streams
   1. Theory: what are things that may cause a customer to churn (elements described below)
   2. One - Chris
      1. Creating analyzable file
         1. Create SQL script which transforms raw data in a file which can be analyzed by a statistical software program
            1. Pre processing
            2. Data re-shaping
   3. Two - jenn
      1. Communicate
         1. Goal
         2. Executive summary/brief methdology
         3. What does the data look like
            1. Lots graphs, descriptive
         4. The model chosen – briefly go over
         5. Model diagnostics
         6. Further research – how to improve model with more time/information
         7. Appendix – alternative models
         8. Details of methodology and models and technical stuff (sql as well)
   4. Three – Combine thoughts
      1. Revisit steps  1 & 2 based on findings which we have each discovered in phases one & two
      2. Model running – combine thoughts on various models
         1. Different specifications
         2. Possible different estimators
         3. Etc
   5. Four - Jenn
      1. Finish communications
   6. Put churn model in production – chris
      1. Translate churn model developed in stage three into SQL or write program in java/R to upload results into table which can be called each month

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Theory described above (draft)**

Goal: Given information of a an FOL ordering customer at month *t-1* we seek to predict whether or not that customer will purchase or not at month *t*… (data is lagged by one month)

1. Non Dynamic Variables (vars summed up of last 12 month period)
   1. Pricing
      1. Poor priced customers will not get manual overrides on website and will “see” the inaccurate price. Thus managed pricing is important.
      2. How to best measure? MDI, Straight Manual, Etc
      3. Is relationship Linear? (maybe sqrd term?)
   2. Size of Customer
      1. Larger customers receive more attention because of their very size, lots of unmeasured stuff here
      2. How to best measure? AVG sales per month of last 12 months of history? Total 12 month sales? Orders? Credit Limit
   3. OS
      1. A customer who has an outside salesman will receive more attention and less likely to churn
   4. Possible Interaction (OS\*Size)
   5. B2O
      1. Customers with high bid to order sales may not use the website as much? Is this true? Or does this simply correlate with some outside variable?
   6. High % of splits/backorders
      1. Customers who do a lot of splits and backorders will find the website more difficult because it doesn’t mirror the brick and mortar accessibility as easily.
   7. Number of functionality used (have they done an SSE on Mylist, or some of the other tools?, how many have they used? Are some more important than others?)
   8. Tenure on FOL & FEI
      1. Customers who are more used to doing business with Ferguson will excel on the website
      2. Customers who are more used to the website will continue to use the website (inertia)
   9. Number of users: customers who have many users are more likely to continue to order
   10. Pro plus customer: customers who have signed up for pro plus are more engaged. Although they may also just be larger
   11. Channel acquired customer (showroom, counter, traditional, etc)
   12. Number of Specials
   13. Breadth of products – number of distinct discount groups purchased in last year
   14. High Job Sales as PCT of total Sales
   15. AVG days between transactions (FOL & FEI)
   16. AVG days between SSEs
2. Dynamic Variables
   1. Month of year (seasonal effects): winter is generally slow and summer are generally busy
      1. Possible interaction with type of customer: some customers groups are more seasonal than others
   2. Change in offline sales, online sales over past *n* months
   3. Last month FOL sales of customer/average FOL sales per month of customer: was last month’s sales on FOL atypically low?
   4. Change in above for the past *n* months
   5. Days since last FOL transaction: shorter days more likely to order frequently
   6. Run of FOL purchasing months: longer run more likely to continue
   7. AOV at month *t* – Mean AOV: distance from overall AOV may be problematic: too high may mean all sales went through one order, too low may mean not many sales in general
   8. Change in total SSEs
   9. LM days between transactions/ AVG time between transactions