Problem Set: SAS

1. **Files (each tab is a separate file):**
   1. Activity: Patient-level database that captures their presence in the health system.
      1. Patient\_ID: Anonymous patient identifier
      2. M\_ : Monthly activity of patients in the years 2010 – 2013 (months represented by numbers 01-12)
         1. 1 = actively enrolled
         2. 0 = not actively enrolled
   2. Diagnosis: Contains the dates and types of ICD (diagnosis) codes received by patients
      1. Patient\_ID: Anonymous patient identifier
      2. Diag\_Code: ICD code received
      3. Date: Date patient received ICD
   3. Rx: Contains the dates and days of supply of drugs prescribed to patients
      1. Patient\_ID: Anonymous patient identifier
      2. Rx Date: Date of prescription
      3. Drug: Name of drug prescribed
      4. Days\_Supply: Days of supply for prescription
2. **Questions:**
   1. Export to Excel a list of patients that are continuously enrolled for at least 12 months **at any point in time**.
   2. Export to Excel a list of patients that are continuously enrolled for **the most recent** 12 months.
   3. Of patients in “b”, flag the patients that have received at least 2 ICD codes of any listed below (i.e. it could be 2 of the same kind or 1 of multiple different ones) **in the most recent** 12 months.
      1. 299.00, 399.88, 495.01
   4. Of patients in “a”, flag the patients that have received at least 2 ICD codes of any listed below (i.e. it could be 2 of the same kind of 1 or multiple different ones) **with a 12 month washout on the first claim** (i.e. they are continuously enrolled for 12 months prior to their first claim) and 24 month look-forward period (i.e. they are continuously enrolled for 24 months after their first claim). These patients’ first ICD claim below will be referred to as “diagnosis claim” from here on out.
      1. 299.00, 399.88, 495.01
   5. Of patients in “d”, find the number of patients that were treated with the drugs (below) after their diagnosis (do not look further than 24 months after diagnosis).
      1. Abilify
      2. Seroquel
   6. Of patients in “e”, report the patients that have a 12+ month washout on their first treatment claim post diagnosis for
      1. Abilify
      2. Seroquel
   7. Of patients in “f”, report the mean/median days of supply they received in the first 12 months after their initial treatment claim for the drugs below (do not look further than 24 months after diagnosis).
      1. Abilify
      2. Seroquel

Problem Set 2: Client Question (in bulleted note form)

* The client has a product in development for a rare disease in which physicians recognize differences in severity, *but no guidelines exist* for our client to map individual patients directly to severity levels
* The client’s product will be targeted in patients with moderate to severe disease
* The client has a database that contains clinically-relevant information about the disease (i.e., flags indicating the presence of symptoms and a variable counting the total number of symptoms) and each patient has been rated as mild, moderate, or severe by an actual physician
* The client would like ClearView to extract the mental heuristic physicians are using when they label a patient with a severity
* **You have been tasked with extracting simple rules** (*1-3 per severity level*) to classify patients from the database. An example of a set of rules would be “patients that have fatigue OR exactly 2 symptoms are severe.”
* In your response email back please provide the following:
  + The rules you discovered that describe each severity level
  + The overall confusion matrix
  + A general description of your approach (only 1-2 sentences)