**Is a recommendation system too much for this issue?**

High-Level Info

* Personal Lines
* Excess and Supply Homeowners
  + Homes with coastal exposure
  + Homes with no prior insurance coverage
  + Difficult liability exposures
  + Homes with prior claims
  + Builders risk / homes under renovation
* Doesn’t fit the producer’s authority so referred to Markel underwriter
* **Find appetite guidance that Markel can provide to producers**
  + Risk pockets
    - Similar attributes in approvals

Useful Resources/Links

* <http://mymarkelglobal/Departments/underwriting/us-spec/Pages/Markel-Personal-Lines.aspx>
* <https://www.markelinsurance.com/specialty-property/excess-and-surplus>
* <https://www.markelinsurance.com/specialty-property>
* <http://mymarkelglobal/Departments/underwriting/us-spec/Markel%20Documents/MPL%20materials/Dwelling-fire-0416.pdf>

Plan

* Text mine the risk details and reason for binding notes fields to find risk buckets that correspond to an approval or decline
  + Treat each classification differently
  + Question –
* After finding similarities, build a recommendation system targeted for the producers to decide if they should send certain policies to the underwriter desks
  + apriori?
    - <https://stackabuse.com/association-rule-mining-via-apriori-algorithm-in-python/>

Questions

* Should I just use the binary classification in the decline column or come up with a new classification system that incorporates pending, converted, subj, etc.
  + Just use 2 classifications
* If producers get more commission if they bind the policy themselves why should we try to keep the policies away from the underwriters desks?
  + Underwriters can then focus on more complicated issues
* Who wrote the risk details column, the producer or underwriter?
* What year homes are considered too old for an approval? Is there a hard cutoff?
* Do all homes have online photos available?
* What are the factors for determining fire risk for homeowners insurance? Is being in a brushy area one of the biggest factors?
* How important is roof quality? Is roof risk included in updated home risk?
* How many losses will constitute for a decline? Is there a hard cutoff?
* How bad is a long response time in terms of risk?
* The approval/decline classification is very imbalanced, should I oversample the declines so that I can make better generalizations regarding the differing summary metrics?
* Although most of the risk details/reasons for decisions columns typically only include 1 risk type, do the underwriters check all risk types?
  + In other words do they go through a process
* How is an E&S home defined as E&S in the first place?
  + Is it from the risk itself or something before?

Specific Risks identified after analyzing the Risk Details Column

* Brush/Fire
* Alarm
* House systems updated
* Roof status
* Loss/claim history
* Online photo
* Response time
* Tried to use wordcloud on biggest reason for decision text but could not find any meaningful info and was led down a rabbit hole
  + Will ignore for now – maybe not necessary for risk buckets and/or recommendation system
* Bar chars of counts for different types of risks (approval and decline)
* Analyze each individual risk type
  + Home update
    - Look at years
    - How old are these homes?
    - Are they updated or not?

Recommender System

Either a hand drawn decision tree classifier or a recommendation training on binding data

* Rules
  + What is the response time?
  + Is the home in a brushy area?
    - If home is in brushy area and response time > 15-20 minutes, then most likely automatically decline
  + What are the loss runs?
  + Are the home systems updated?
    - Roof?
    - Alarm systems?
  + Are there online photos?