# As one of the preparations for the forthcoming of recession, the Qualitative Research Department of Wells Fargo Risk Control was asked to conduct a study to understand its mortgage (home loans) portfolio risk landscape as a way for earlier detection to help customers and also help the bank to prepare for the potential credit loss.

# Study Objectives To identify customer segments which have different arrears behavior/propensity, and describe what the customer is like (not just what stage his/her arrears have reached The goal is to enable risk control to have greater understanding of why different customers go into arrears, and provide insights to drive development of differentiated approaches for different types of customers

# Methodology and Approach In order to separate customer base into meaningful clusters each with a different character (“homogenous within and heterogeneous without”). Segmentation model built on customers with active mortgage accounts with the bank currently. Methodology is basing on “k-means”: SAS FASTCLUS procedure used to obtain disjoint cluster analysis on the basis of distances computed from one or more quantitative variables, observations and divided into clusters such that every observation belongs to one and only one cluster, cluster centers are the means of the observations assigned to each cluster when the algorithm is run to convergence

# Findings When predicting future arrears, the single most predictive factor is whether the customer has been in arrears before. This leads to the conclusions that 1) Customers in arrears and/or who have been in arrears recently are segmented according to their arrears behavior (frequency and extent) 2) But since the big portion of population had no delinquency in last 12 months, study has focused on breaking this large population into manageable and meaningful segments 3) As a result, 14 distinguish segments were being identified as the below picture illustrated with their name characterizing each of their unique profile 4) Each segment would be followed up with different actions or passed to different departments for further research like Risk Strategy, Collections, and Operations

# My involvements It was at the beginning of data preparation and final stage of segment profiling which I involved at this study. My learning was tremendous given the research team and especially my mentor’s guidance. I had gotten a lot of support from the data team, whom conducted hours of interviews to explain to me the meanings of the attributes

# Data Preparation and Standardization: First I worked on identifying data with problems such as multicollinearity, outliners, missing or incomplete values. Next is the procedure of data sanitization and standardization. For example some of the missing values I needed to fill in the average values generated from the other non-missing observations. After that data was ready to be used in models and analysis.

# Custer Profiling: After distinguish clusters/segments been identified, I was asked to outline the descriptive profile draft for five of them. I started to take their qualitative characters/attributes, and I learnt to build the spider web chart with them as shown at below chart. This part of work was very new and exciting to me as it was related to both the science and arts. I got to use what I have learnt from my school AP Statistics class and also get to try the deductive methods I learnt from reading those Sherlock Holmes books. I felt like I was almost like a detective trying to deduce who they might be from what I saw at their surface (the data). Here below was an example work of what I outlined for one of the segments, of course was refined with my mentor’s help.

# Segment Name: Making Low Incomes Stretch

# Method: Analysis and reading with spider web chart below

# Hypothesis: Customers making low income stretch have low incomes, low property values and low Loan to property value. Although they have managed to keep their mortgage up-to-date, they are finding it difficult to make ends meet. Although they are not heavily in debt yet, they have missed payments on their other debts like credit card payments. And this is reflected in their credit score. They maybe on the edge to default on their mortgage with us if anything may happen at their circumstances like losing a job.

# 